Abstract

This study analyzes the learning experience of studying foreign languages (FLs) and designs a personal voice assistant to support users. Learning FL is a big challenge. The challenge of learning a new language including how language is used is 'influenced by' or is a direct result of the context in which communication takes place. The author's self-experience as an international student informs this study and provides access to a broad range of peers at the institution who have similar experiences. On the other hand, by utilizing voice assistant (VA), users build a relationship with VA and learners can understand multiple language backgrounds and cultures. Through in-depth interviews, participants—recruited from campus—were asked questions about why one wants to learn FL, how one learns FL, which tools one used to support learning, the challenges one meets with this experience, etc. This research asks about one's learning experience and shares the researcher idea with VA utilizing it in supporting learning FL. This study does two sets of interviews with same group participants in order to collect interviewees' experiences and feedback through the design process. In a sample size of 18 participants, this research understands how culture, background and environment influence FL learners' experiences during studying a new language. Some interviewees live in foreign places and feel anxious using FL to communicate with native speakers. Some interviewees live in native and want to know more about FL's different cultures locally. With this study, this author designs a personal voice assistant and supports learners with learning foreign languages.

Keywords: Foreign Language, Voice Assistant, User Experience, Enjoyment, Human Computer Interaction

SUPPORT FOR LANGUAGE LEARNERS BY DESIGNING A PERSONALIZED VOICE ASSISTANT

by

Ai Ni Cheng

BDes., National Kaohsiung Normal University, 2018
M.B.A., National Kaohsiung Normal University, 2021

Thesis
Submitted in partial fulfillment of the requirements for the degree of
Master of Fine Art in Design.

Syracuse University
August 2023

Table of Content

List of Appendix

List of Illustrative Materials

Introduction

Foreign language (FL) is an effective skill, it can support people to communicate with multi-group populations from different backgrounds, cultures, and experiences. FL learning is a task, learners sometimes find it hard to find a person to practice with, sometimes don't understand the target language background culture, and some students feel anxious to begin speaking. Those facts lead FL learners to challenge themselves to learn and experience something new.

FL learning anxiety for international students

Foreign language anxiety (FLA) is widely researched in language teaching research and it has negative correlation for students (Li 2022; Tsiplakides and Keramida 2009; Razak, Yassin, and Moqbel 2019). This anxiety causes students to have lower active and lower performance than non-anxious students in class (Tsiplakides and Keramida 2009).

Communication apprehension, fear of negative evaluation, and test anxiety are three components of FLA that have been identified (Tsiplakides and Keramida 2009). Students feel unconformable and fear making language errors when using FL in communications due to limits in listening and speaking skills (Tsiplakides and Keramida 2009). Students feel confused when they face unfamiliar script, term, or text; students can't understand the meaning and can't continue talking, writing, or reading, and therefore their anxiety increases (Razak, Yassin, and Moqbel 2019). Unfamiliar cultural aspects also cause anxiety for international students (Razak, Yassin, and Moqbel 2019). Because they are not able to decipher the target language's background culture,

this fact leads them to limit understanding of the meaning of several patterns in communication (Razak, Yassin, and Moqbel 2019).

FL Learning Enjoyment

Positive emotion has a positive effect on FL learning (Zhang and Tsung 2021; Li 2022). Positive emotion and enjoyment go beyond a pleasant feeling and are foundational concepts of positive psychology (Li 2022). This pleasant feeling arises from learners engaging in FL activities and they are more engaged or willing to keep going in FL (Li 2022); this positive feeling decreases negative emotions, people feel more focused, and they report an increase in their stress resistance (Zhang and Tsung 2021). For a learner, fun, interest, creativity, joy, and happiness are included in FL enjoyment and those emotions increase learners' motivation to learn FL (Li 2022; Zhang and Tsung 2021).

Voice Assistants (VAs) Enjoyment

Involving voice assistants in FL learning can increase learners' enjoyment. Voice Assistants (VAs) are often used as a form of personal assistant for many people. Chat with social robots have positive social contributions to FL learners (Alm and Nkomo 2022; Huang, Hew, and Fryer 2022; Goda et al. 2014). Timeliness, ease of use, and personalization are three categories of chatbot technological affordance (Huang, Hew, and Fryer 2022). The degree of agreement about willingness to engage in a conversation with chatbots is increased by analysis results from second language learners (Alm and Nkomo 2022). Students are more willing to contribute to conversations with chatbots and they increase the number of conversations that students participated in (Goda et al. 2014).

Moreover, this supported tool, VA, enhances FL learners' positive emotion, enjoyment, FLA is decreased, and learners can personalize their learning experience. Personalizing voice assistants enhance users' enjoyment (Poushneh 2021; McLean, Osei-Frimpong, and Barhorst 2021). Because VA provides social existence and knowledge, users feel VA is close and will build a friendship with VA (McLean, Osei-Frimpong, and Barhorst 2021). VA provides personality and this fact drives users to feel satisfied during voice interaction (Poushneh 2021). As for VA's personality traits close to the customer need, one will feel enjoyment because people are "goal-oriented and select media that fits their needs" (McLean and Osei-Frimpong 2019). On the other hand, in order to meet the need of the user, VA remembers one's needs and learns to become personalized to enhance the interaction relationship (McLean, Osei-Frimpong, and Barhorst 2021). That is to say, the more the VA is personalized, users have more satisfied relationships with VA, and users can engage and become more willing to use this tool in the future.

In summary, the contributions of this study are shown as follows. First, there are more and more people researching voice assistants, however, this research area is still limited right now (Al Shamsi, Al-Emran, and Shaalan 2022). This thesis explores a wide range of this research area. Second, this research utilizes designer self-experience, an FL learner, to set the design challenge. By using personal experience, this design can more effectively reach out to the target audience and it can decrease domination assumption (Costanza-Chock 2020). Third, by researching and designing in progress, this design is based on the target users' experience to create the project. By understanding the target users' experience, this study brings a new scope of knowledge exchange design which means the databases are collected from local users. This

fact can support learners learning FL locally and learners can gain updated knowledge from other users.

The thesis is structured as descriptions. The next section develops the progress of the study and describes the methodology of this research. In the following section, the background to the study and reviews of prior research on FL and VA are proposed. In the final section, I discuss the implications of these findings and conclude by explaining the limitations and suggest directions for future research.

Methodology

The study implements voice parameters research (Appendix 1) and literature reviews (Chapter 3) of the voice assistant to comprehend prior existing research on FL learning. Along with one-on-one in-depth interviews with open-end questions, this author understands the target audience's learning and related tools based on her experience. Following prior research and two-set interviews, a personalized voice assistant design is conscientiously crafted (Figure 1).
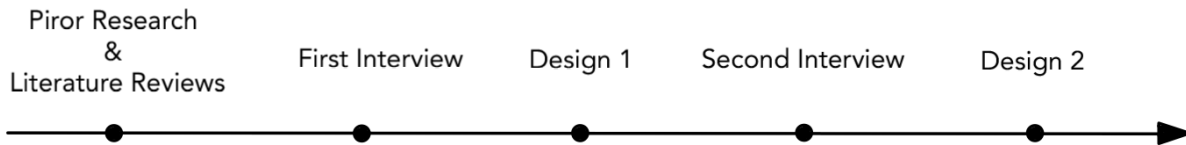
Figure 1 Study Description

Prior Research: Enhance Users' Enjoyment by Changed Voice Parameters

In the prior research, this author wants to understand if various voice assistants' (VAs') voice parameters (i.e., pitch range and speech rate.) may influence users' enjoyment and preference. VAs feature various voices as Siri has five kinds of recordings in the English (American) channel. Since Siri has diversified voices. A non-binary voice (i.e., iOS 15.4, recording five) with modified pitch ranges and speech rates by AUTO-TUNE software is used in this research. Through a web-based survey conducted on Qualtrics, participants—recruited from Amazon Mechanical Turk (MTurk)—were asked their feelings after listening to four patterns, which included higher pitch range, lower pitch range, higher speech rate, and lower speech rate. I ask about their enjoyment of listening to those patterns across five distinct categories:

motivation, goal, proof, evaluation, and potential. In a sample size of 114, this study discusses how the responses differ between the distinct categories and how they link to various VA voice parameters.

In this prior study, this author demonstrates that various VA voice parameters have no significant effect on users' motivation, goals, proof, and evaluation. Users' potential is slightly affected by the changeable voice parameters of this human-computer interaction; however, it is still not significant. With this investigation, I indicate that the pitch range and speech rate of the voice assistant, Siri iOS 15.4 recording five, has no effect on the enjoyment of their user's experience (see Appendix 1).

In short, prior research shows that personalizing VA can support FL learning, and utilizing modified voice parameters might have a positive effect on user experience. The following research recruits target audiences from Syracuse University (SU) to have to gain deeper understanding of their experience during learning an FL.

Participants

The participants were recruited from the SU campus and invited to join an in-depth interview. The number of interviewees is eighteen with a background in FL-learning. The majority of participants were studying at English Language Institute (ELI) (62%, 11/18). Specifically, 72% (13/18) of participants had little knowledge of English or were intermediate English learners.

Interviews

The 40 minutes interview will be taken place at the location of the interviewee's choice on the campus of SU (e.g.: Bird Library, Barnes, The Warehouse, etc.). Participants need to be at lease 18 years old and should have experience learning FLs to help the improvement of modified voice assistant. Two sets of interviews are included in this study. Further research and design are influenced by volunteers' feedback and comments.

During the first set of interviews, participants were asked questions about their motivation and experience of learning FL: why they wanted to learn FLs, how they learned FLs, and which tools were most effective in supporting their learning (Appendix 2). This research inquiries about individuals' learning experiences and communicates the researcher's concepts to virtual assistants to apply in facilitating the FLs' learning.

After understanding the target audiences' feedback and comments, the personalized VA design is considerately crafted. Following the second interview (Appendix 3), participants are invited to explore the introductory prototype. During the interview, eight interviewees are recruited from the same group of prior participants. Eight interviewees are asked to interact with the designed VA prototype. The designed VA is improved based on user testing.

Carefully Crafted Personal Voice Assistant Design

To begin with, the author utilizes a storyboard (Figure 2) and a video (Appendix 4) to describe a personal experience during learning FL. This story describes FL learners' social anxiety when facing unfamiliar FL slang in an FL conversation. In following the story, a

personalized VA is introduced in the video and the character talks with this personal VA to

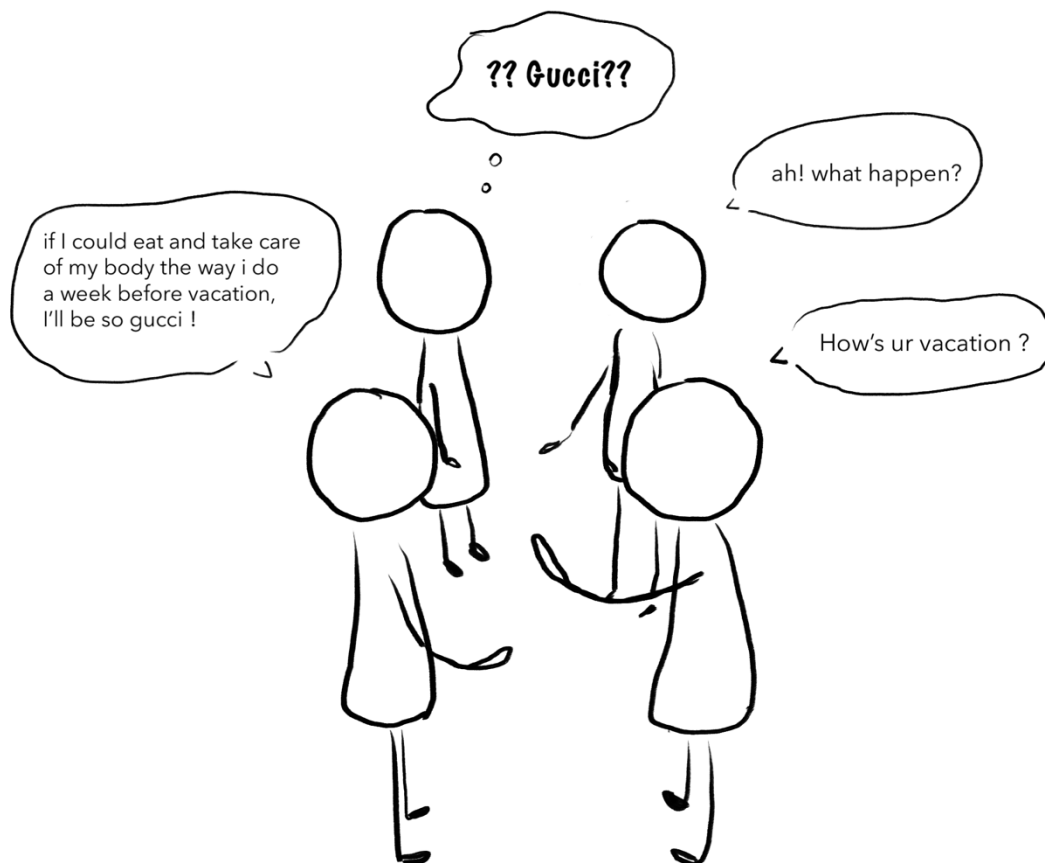understand the conversation pattern and the slang meaning.

Figure 2 First Prototype Storyboard

Next, the participants are asked to interact with a physical voice assistant and have oral

communication with the VA as the character in the story. The physical VA is built and the

participants are asked to interact with this VA to experience the design progress.

Design One Approach

This section introduces a custom-designed VA that references a local cultural event to support language learning in detail. The prototype makes use of a slang as an example and shares a story about the slang with a video. Additionally, the overall design includes a physical prototype, a tactile 'stress ball-like input device that allows users to 'squeeze and talk', to record snippets of conversations in social settings. For this physical prototype, I was able to create the VA via Python code so that interviewees could experience interacting with the product.

Storytelling Video

In the video, Annie, an FL learner, is learning English and chatting with colleagues in English after class. A colleague narrated one's feelings and said: "If I can eat and take care of my body the way I did a week before vacation, it will be so Gucci!" Annie feels so confused at that point, cannot join the conversation, and keeps thinking about what the meaning of Gucci is (see Figure 2).

At this point, Annie squeezed the Kinakaian device in her bag several times in order to document this conversation. Kinakaian records the conversation and documents 20 seconds as a period, i.e. Kinakaian documents 10 seconds before and after Annie squeezed. After finishing a whole day class, Annie goes back home to interact with the Kinakaian device (Figure 3). For more video information, please see appendix 4.

Kinakaian	Hi Annie, what can I help you with today?
Annie	Hi K, what's the meaning of Gucci?
Kinakaian	Gucci is a slang term that is used as an exclamation of approval or as an expression of goodwill.
Kinakaian	It can also be used as an adjective to describe something that is going well is in good condition.
Kinakaian	Gucci is derived from the luxury Italian fashion brand Gucci and is used t indicate that something is good, going well, or is in good taste. It is often as a slang alternative to the words good or great.
Kinakaian	Would you like to do some practice using this term?
Annie	Yes.

Figure 3 First Prototype Storyboarding Text

Physical Prototype of Voice Assistant

Stress Ball

This design utilized a stress ball to form a language-support voice assistant. Because a stress ball reduces the stress levels, it is a self-contained nursing way to support users when they are anxious (Kasar, Erzincanli, and Akbas 2020). By searching for the smallest speaker on the market, this author uses 3D software, Rhino, to sketch the stress ball layout and prototype (figure 4).
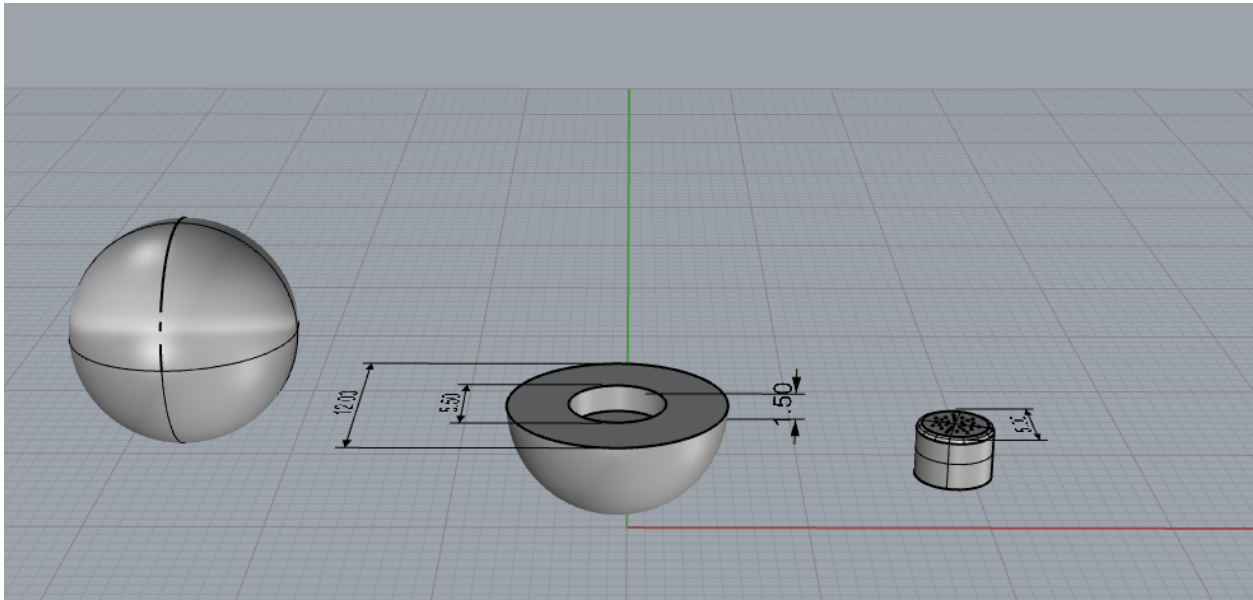
Figure 4 First Prototype 3D Model

This prototype is created by Makergear 3D printer. By inputting material, Nylon, in this printer, it prints out a flexible texture and users can squeeze and play with it (Figure 5).

Figure 5 Physical Prototype of VA

The reason that one's grip and pinch strength are affected by various factors, and those factors are also different from person to person (Habibi et al. 2013). There were several prototype sizes created. In order to figure out different people groups of different size balls, this research asked interviewees their volunteers to grip those balls (Figure 6).

Figure 6 Volunteers grip different size ball

<u>Speaker</u>

In order to create the experience of interaction with voice assistants, this researcher learns Python-code skills and produces a speak-and-respond pattern experience (figure 7).
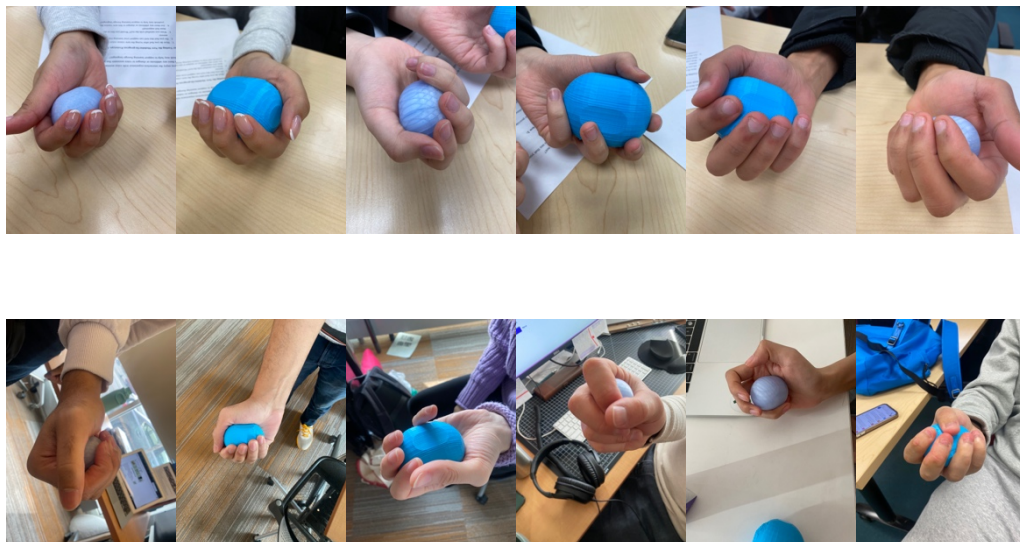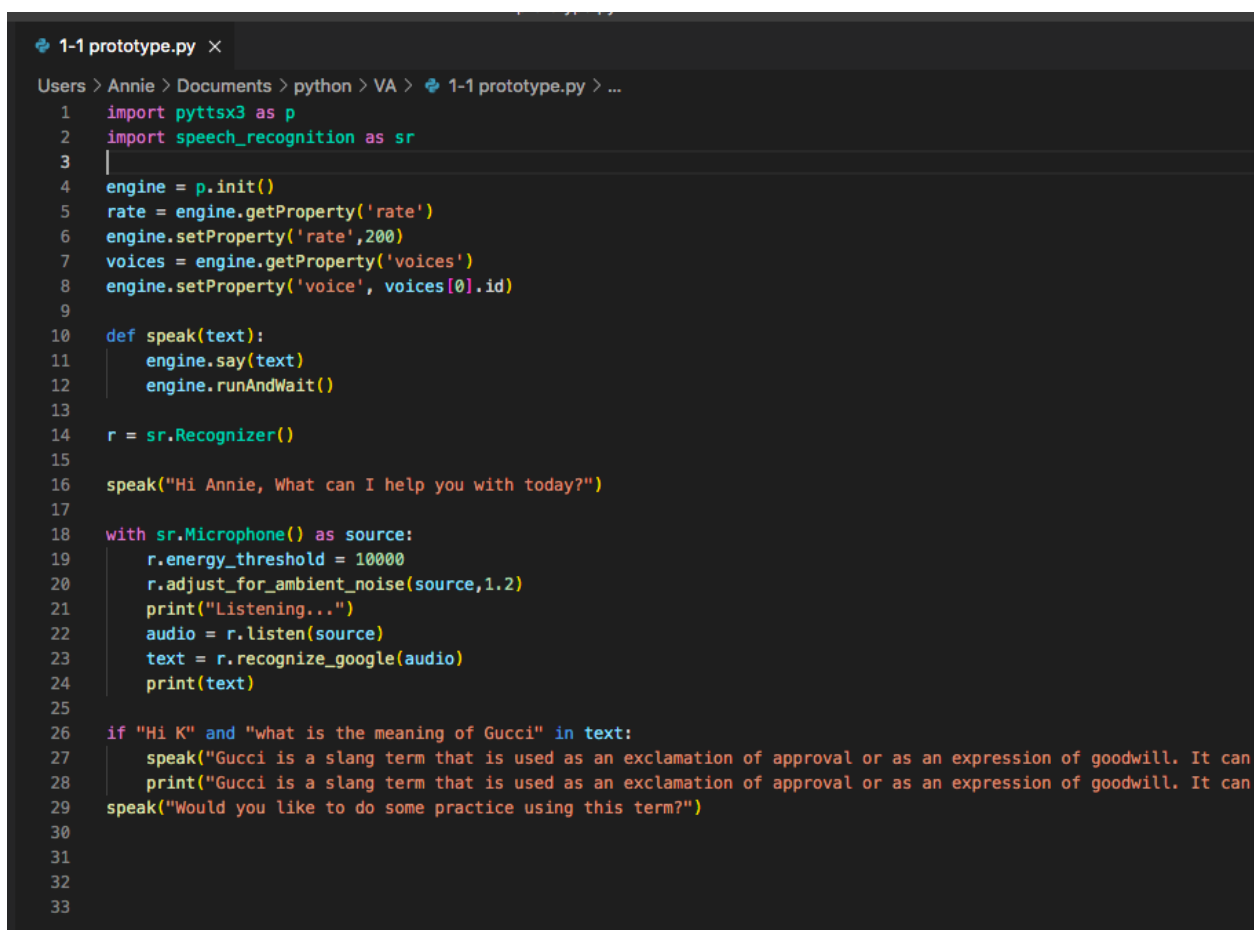
```python
import pyttsx3 as p
import speech_recognition as sr

engine = p.init()
rate = engine.getProperty('rate')
engine.setProperty('rate',200)
voices = engine.getProperty('voices')
engine.setProperty('voice', voices[0].id)

def speak(text):
    engine.say(text)
    engine.runAndWait()

r = sr.Recognizer()

speak("Hi Annie, What can I help you with today?")

with sr.Microphone() as source:
    r.energy_threshold = 10000
    r.adjust_for_ambient_noise(source,1.2)
    print("Listening...")
    audio = r.listen(source)
    text = r.recognize_google(audio)
    print(text)

if "Hi K" and "what is the meaning of Gucci" in text:
    speak("Gucci is a slang term that is used as an exclamation of approval or as an expression of goodwill. It can
    print("Gucci is a slang term that is used as an exclamation of approval or as an expression of goodwill. It can
speak("Would you like to do some practice using this term?")
```

Figure 7 First Prototype VA Python Code

After creating this function, this author connected the code to the speaker so that interviewees can talk to the VA; by utilizing squeeze, the user arouses or wakes up the VA and begins this conversation experience.

In the second interview, participants' feedback and comment knowledge are received after experiencing the first design prototype to improve the further design. The finalized first and second interview is documented in the result section with the further design solution. Along with participants' knowledge and working with MFA Design specialists, a successfully supported voice assistant that enhances FL learners' learning experience is created.

Literature Review

Learning FL language (The reason why one want to learn FL)

Intercultural contact is one of the main motivations for learning a foreign language (Dörnyei, Csizér, and Németh 2006; Eom and Braithwaite 2023). Foreign language is a tool that can succeed in one's understanding of cultural interests, the experience of multiple communities and linguistic self-confidence (Dörnyei, Csizér, and Németh 2006). By understanding the local language, one can learn more about contemporary pop culture, the linguistic and cultural diversity of the local community (Eom and Braithwaite 2023).

On the other hand, learning a foreign language is used as a form of building a bridge of communication(Olwi 2022; Álvarez Valencia and Michelson 2022). In television advertisements, advertisers use a foreign language as a marketing tool to communicate to broad target users (Olwi 2022). In the required language knowledge and strategic competence in communication, language knowledge referred to the linguistic information available in one's brain (grammatical, textual, and pragmatic knowledge), and strategic competence referred to the use of the capacity of that linguistic information (one enable to achieve cognitive management goal, assessing communicative resources, and planning language use) (Álvarez Valencia and Michelson 2022).

Álvarez Valencia and Michelson also point out, communicative competence is not only needed to understand linguistic, social, cultural, and psychological knowledge (Álvarez Valencia and Michelson 2022). But also one needs to know the rules of appropriate and acceptable

deployment of language (Álvarez Valencia and Michelson 2022). Therefore, in order to have successful communication, language learners need to study both linguistic knowledge and need to understand the language's broad range of backgrounds.

The Challenge for International Students to Learn FL

The challenge of FLs is related to intercultural communicative competence. Communicative competence is needed to understand the linguistic, social, cultural, and psychological knowledge (Álvarez Valencia and Michelson 2022). One also needs to know the rules of appropriate and acceptable deployment of language (Álvarez Valencia and Michelson 2022). By understanding the linguistic and cultural diversity of the local community, one can learn more about the local language, contemporary pop culture, and so on (Eom and Braithwaite 2023).

For a MFA student, such as myself, that grew up in Taiwan and then moved to upstate New York, I was able to learn English prior to entering this country. However, often these skills are deficient given that a large percentage of conversational English is only understood at a deeper level. Having an understanding of slang, local cultural events, and developing a trained ear to grasp the regional dialect all take time.

Because communication is linear and happens in real-time, there is often little or no time to "pause the conversation so that a particular word or phrase can be deciphered" to understand the meaning at a deeper level. There is also the challenge of gaining an understanding of local traditions (such as the annual St. Patrick's Day parade) which has regional significance given the large percentage of Irish Americans that live within the greater Syracuse community.

Each of these factors suggests there is a gap between current language learning software and the creation of a tool that truly deals with language usage within a local or regional context.

Related App to support FL exited now

In order to better understand the platforms that are currently on the market that are specifically aimed at supporting FL, in Figure 8-12, this research documents nine software that interviewees mentioned they have used to support their learning experience.

Logo		
Software Name	Linguee	Duolingo
Intro and Function	This app defines the meanings of words on web pages; shows translations and examples of the vocabulary. By highlighting terms on sites and clicking sharing, this app can catch the word and show the dictionary nearby. There are several translations it supports: German to English, French	Duolingo provides more than 40 languages and it support learners by lessons. Users can select languages, level of experience, the reason for learning this language, duration of learning, and the location they would like to learn. This App provides questions and answers to support users

	to Spanish, Russian to Japanese, to English, Portuguese to Italian, Dutch to Chinese and so on. Users can use this app during learning or searching information online easily.	remembering words. This software also provides voice and image to describe the meaning of the term. Users can replay the pronunciation of specific terms and can do practices with new terms.
Medium	Mobile App	Mobile App
App store ranking	iOS: 4.9 (416 reviews) Android: 4.8 (81K reviews)	iOS: 4.7 (16k reviews) Android: 4.5 (14.5K reviews)

Figure 8 Platforms that support FL learning 1

Logo		
Software Name	Rosetta Stone	Grammarly

Introduction and Function	Rosetta Stone utilizes a mimic way of learning one's native language to support foreign learners. This software brings out the point that language learners should learn a new language through daily life. By repeatedly reciting information by showing the same image, same vocabulary, and sentences, target learners will have the memory and then bring this language into one's minds. Moreover, by utilizing AI, this app records the users' voices to identify and correct users' pronunciation. By providing different levels of courses, learners can meet one's personal needs during learning.	Grammarly is a typing assistant to check grammar mistakes in English texts. The user can upload a Word document or type it in on the Grammarly webpage. This software will suggest grammar use, spellcheck, and punctuation check on users' writing. After finalizing the writing, users can directly download the article in different formats of files.
Medium	Mobile App and App for Smartphone Keyboard	Browser Extension, Webpage, Mobile App and App for Smartphone Keyboard
App store ranking	iOS: 4.7 (297 reviews) Android: 4.7 (355K reviews)	iOS: 4.4 (807 reviews) Android: 4.4 (179k reviews)

Figure 9 Platforms that support FL learning 2

Logo		
Software Name	Tandem	Google Translate
Introduction and Function	Tandem aims to create an international community where language learners can exchange knowledge of their native language and learn with each other. This software provides more than 160 languages. Users can use instant texts, voice messages, audio calls, or video calls to chat with partners. Users can reach partners by using location, age, and gender filters.	Google Translate can translate up to 108 languages. By selecting a term or highlining a paragraph, a translated language will pop up nearby. Customers can get the text in multiple languages and software provides immediate translations. Users can use a camera to enter the contain, and Google Translate will decrypt into different languages users choose. Moreover, it can translate information by documenting audio so that the audience can turn on this software during an FL conversation.

Medium	Mobile App	Mobile App and Web Page
App store ranking	iOS: 4.5 (3988 reviews) Android: 4.4 (30.3K reviews)	iOS: 3.7 (6592 reviews) Android: 4.2 (865K reviews)

Figure 10 Platforms that support FL learning 3

Logo		
Software Name	Netflix	YouTube
Introduction and Function	This program offers a wide variety of TV shows, films, anime, and talk shows in more than 25 languages. Those broadcasts will show differently as users' locations, and people can view films locally. Users can watch thousands of videos with different languages in the living room; FL learners can understand	YouTube is a sharable video online platform. Users can upload one's own created videos and share their views, comments, and notes. FL learners can subscribe to various channels from FL native speakers, to learn about their channel information and knowledge personally. There

	how FLs are used in conversations and simulate the FL environment in scenarios. Moreover, learners are able to follow subtitles to get a whole understanding. By using this software, users become familiar with how native speakers talk and get used to the speed rate of daily speaking. People will feel stressless learning languages by watching entertainment broadcasts; learners can replay the same video several times to understand the draft.	are multiple online YouTube channels to teach language lessons; there are TV shows and movies in the learner's target language and one can learn with video by transcripts and subtitles. One can leave their question or information in common and exchange knowledge together to have a better understanding of the broadcast. The changeable speed rate function on YouTube is helpful for viewers to slow down when feeling unfamiliar with the speaker's speed rate; Users are also able to speed up when they would like to get the whole idea in a short time.
Medium	Mobile App and Web Page	Mobile App and Web Page
App store ranking	iOS: 3.7 (341.9K reviews) Android: 4.3 (13.9M reviews)	iOS: 4.7 (30.2M reviews) Android: 4.2 (148M reviews)

Figure 11 Platforms that support FL learning 4

Logo	
Software Name	Wlingua learning Spanish
Intro and Function	Wlingua Spanish learning software provides a more personalized setup based on users' native languages for learning Spanish. This app asks for setting up the user's native language on the first page and then providing a suitable lesson via some questions. Wlingua Spanish brings up to 500 lessons which include vocabulary pronunciation, definition, and practicing using the term in sentences. By providing real conversation records, learners can listen to the audio and select the answer to respond. Wlingua Spanish brings multi-situation to lessons, such as how to answer phone calls naturally, and how to recognize different pronunciations in a variety of cases. This software provides language not only based on nationality but location. Take Spanish as an example, this software includes European Spanish and Mexican Spanish at the same time so that users can more ideally learn deeper with this new language.
Medium	Mobile App and Web Page

App store ranking	iOS: 4.8 (5 reviews) Android: 4.8 (60 K reviews)

Figure 12 Platforms that support FL learning 5

These software products are predominantly app-based and learners can gain knowledge easily using smart devices. However, interviewees point out that app-based instruction can be stressful and decrease the motivation to work with the app on a continual basis. Utilizing YouTube and Netflix to learn an FL is effective, however, subtitles displayed in the native language are easier to understand while watching videos but decrease the learning possibility. To a large the extent, users focus less on learning a FL and are more engaged in the story or plot.

In order to better understand other products that successfully support FL learners, related software and technology provided by prior research are documented in this study. Those tools show a positive effect on FL learning in existing research (Please see Figure 13-15).

Logo		
Software Name	Memrise: speak a new language	TED Talks
Intro and Function	Memrise focuses on learning a new language via conversation and daily life situations. By providing a variety of conversational options, such as travel, education, and food, users are able to learn how to chat about specific topics. This software brings images to support understanding as well. Moreover, learners can watch some local videos to understand how the local population speaks this language. An Al chatbot tutor answering a question without a time limit is also helpful for users during the learning experience.	TED Talks contain speeches from speakers with professions in education, business, science, tech, creativity, and nature. Speakers came from different backgrounds with different languages used. They share their experience and knowledge on the podium and the audience can listen to the speech in person or online. Speakers sometimes use slides to vivid their speech. Learners can learn a new language followed by new knowledge in an interesting subject. In the Ted video, subtitles are also available in 100+ languages. Users can change the speed or pause & start whenever they want. On some platforms, such as YouTube and Facebook, viewers are welcome to share

		their comments and exchange ideas with others.
Medium	Phone, Tablet, Chromebook and smart TV	In-person speech, Phone, Tablet, smart TV and Video Player
App store ranking	iOS: 4.8 (16.7K reviews) Android: 4.6 (1.42M reviews)	iOS: 4.6 (10.1K reviews) Android: 4.0 (214K reviews)

Figure 13 Platforms that support FL learning 6

Logo		
Software Name	WhatsApp	Oculus VR

Introduction and Function	WhatsApp is a social networking mobile application that can send messages, have video calling, and build social networks with others. Sharing videos, photos, files, and social contexts can be achieved through this software. Moreover, according to Alamer and Khateeb's research, the social networking mobile application, WhatsApp, has a positive significance on leaner's FL learning motivation. The learner uses WhatsApp to chat via FL and feels satisfied after utilizing this software. One can increase vocabulary, knowledge, reading, writing, speaking, and comprehension of the FL (Alamer and Al Khateeb 2023).	Oculus VR is a Virtual Reality(VR) experience technology that includes a headset, power bundle, and application. Users are able to experience certain apps, and games in 3D visual, tactile, and audio experience. The headset has an OLED display with visual and audio functions; the power bundle has a tactile function and control bar. Players can use Oculus VR to interact with game experiences that imitate a physical world. FL learners can use this tool to support their learning via experience FL environment. A chatbot in Oculus VR can chat with users like an in-depth FL conversation. Moreover, research shows that these are positive effects on FL learners' engagement, engrossment, and immersion during interactions with Oculus VR (Nicolaidou, Pissas, and Boglou 2023). FL learners' vocabulary performance has a higher score through practice with VR applications (Nicolaidou, Pissas, and Boglou 2023).

Medium	Phone and Tablet	VR headset, 2 Touch Controllers, and PC
App store ranking	iOS: 4.7 (12.2M reviews) Android: 4.3 star (180M reviews)	Amazon: 4.7 (83,644 ratings)

Figure 14 Platforms that support FL learning 7

Logo	![Babbel logo]	![HelloTalk logo]
Software Name	Babbel	HelloTalk
Intro and Function	Babbel is a popular language training application (Mohamed, 2020) that includes 14 languages for learners to experience. With Babbel, customers can strengthen	HelloTalk is a popular language training application (Mohamed 2022) and is the original language exchange app. FL learners can practice languages with native speakers and a variety of subject

	their habits by customizing one's learning experience. With a clean and refreshed software interface, a beginner learns vocabulary with some memory stages: repeating after listening to the vocabulary, understanding image knowledge with related words, imitating real-word chat, and building one's learning habit. Voice recognition in this software supports learners' accurate pronunciation. Revisiting and re-entering unfamiliar lessons enhance learners' long-term retention of memorizing the FL. Personalize learning paths based on the user's schedule so that learners can have a fixable learning experience.	rooms can be created by users. Customers are able to interact with FL and celebrate cultural events via this app; one can also share posts like social media functions with the public. On the personal information page, learners can share their own language, target FL they would like to learn and reach out to the target partner who has the same experience. Livestreams and Voiceroom support users to learn from natives with no pressure.
Medium	Phone, Tablet, Chromebook and Smart Watch	Phone and Tablet
App store ranking	iOS: 4.7 (509.6K reviews) Android: 4.5 (904K reviews)	iOS: 4.6 (37K reviews) Android:3.7 star (195K reviews)

Figure 15 Platforms that support FL learning 8

Conversation, daily life situations and social-constructivist support FL learners (Al-Hoorie and Hiver 2020; Gonz 2015; Alamer and Al Khateeb 2023; Nicolaidou, Pissas, and Boglou 2023). In Computer-assisted language learning, text, audio, and video SCMC, email, forums and bulletin boards, social networks, and games effectively promote learners in linguistic education (Gonz 2015). Social networking increases students' enjoyment and motivation while using it during the learning period; students' participants become more convinced of learning the FL via social-constructivist support (Alamer and Al Khateeb 2023). Moreover, emotional attachment and a focus of attention during FL have a positive influence on computer-assisted language learning (VR and mobile applications) because illustration knowledge is shown and users can interact with short-term programs without time-limited (Nicolaidou, Pissas, and Boglou 2023).

Therefore, this research includes conversation, social-constructivist support and daily life situations in following design section. Through this research with various interviewees, this author was able to assess which current FL software products are successful and understand the real need from target users. In addition, with the introduction of a VA, this study was able to assess how they can support user needs during everyday social interaction.

Voice Assistant

VA's are voice-controlled mobile or personal devices which allow users to communicate in FL. They allow users to make queries and have a deeper emotional connection with native speaking individuals. Examples of current VA's include products such as Google Assistant,

Microsoft Cortana, Amazon Alexa, Apple Siri and Yandex Alice (Arlinghaus and Ollermann 2022; Al-Kaisi, Arkhangelskaya, and Rudenko-Morgun 2021; Le Pailleur et al. 2020).

During VA's interaction, users prefer more personalized voice assistants (Machado et al. 2012; van Doorn et al. 2017; Khan and Iqbal 2020). VA personalized relations during interaction are more effective than impersonalized such as changed pitch value (Machado et al. 2012). Social robots, including VAs, during the interaction between users are seeking to develop in a socially enriching way (van Doorn et al. 2017). By personalized users' engagement during the interaction, this fact will raise the relationship between users and social robots (Khan and Iqbal 2020).

VAs' Positive Effect of FLs

A personal voice assistant as a supported tool for FLs learners because there are positive effects on learning FLs by using VAs (Al-Kaisi, Arkhangelskaya, and Rudenko-Morgun 2021; Arlinghaus and Ollermann 2022). In FLs learning, there is support for cheating with technology because one can learn about language production and complexity (Golonka et al. 2014). Aided by voice assistants with practicing speech, learners have a positive effect on an independent study of foreign languages (Al-Kaisi, Arkhangelskaya, and Rudenko-Morgun 2021). Moreover, the social behavior of spoken language interaction between users and humanizing voice assistants has a positive effect on users' experience (Hu et al. 2022; Golonka et al. 2014; Malodia et al. 2022). That is to say, by utilizing voice assistants (VAs) as a supported tool, users build a relationship with VAs and learners understand multiple language backgrounds and cultures. This

fact supports one's learning experience and the users will understand the local or regional context.

Moreover, chat robots encourage learners' social behavior by providing effective, open and coherent communication (Huang, Hew, and Fryer 2022). This tool can support students with speaking, listening, reading and writing skills (Huang, Hew, and Fryer 2022). Learners can talk with robots on multiple subjects and learn about negotiation strategies (Huang, Hew, and Fryer 2022). In Huang, Hew & Frye's research, chatbots increase FL learners' grammar, listening and writing skill; there is no research report that chatbots have a negative effect for language learners' learning result (Huang, Hew, and Fryer 2022).

Chatbot can support students' learning, but researchers point out that unnatural voices of this machine make users feel challenged (Huang, Hew, and Fryer 2022; Goda et al. 2014). This fact decreases students' positive emotion when interacting with social robots (Goda et al. 2014). AI has limited knowledge and users feel disappointed when robots can't reply or recognize the question after one type in (Yang and Lee 2019). Due to these facts, users feel it is unnatural to have a long conversation with a chatbot, and decrease the interest of learning FL (Huang, Hew, and Fryer 2022).

Users' Experience (UX) and VA's Voice

Enjoyment, trust and perceived ease of use (PEOU) are three categories that perceive users' use VA (Al Shamsi, Al-Emran, and Shaalan 2022). Enjoyment has a significant positive influence on perceived usefulness (PU) ; that is to say, enhancing the degree of VA's enjoyment,

the users believe that using this tool will enhance one's job performance (Al Shamsi, Al-Emran, and Shaalan 2022; Davis 1989). In Al Shamsi, AI-Emran & Shaalan research, they also mentioned that the more the PU is, students will be more conscious about using VA as a tool that can support one's learning and will be more willing to use this tool; students have more motivation to use VA, the situation of utilizing VA in education will increase.

Enjoyment is defined as a good experience (Brandtzæg, Følstad, and Heim 2018). Enjoyment is the effect produced by functional and nonfunctional interactions and reflects the users' feelings during the experiment. It includes enjoying the interaction while using the voice assistant such as feeling nervous, satisfied, anxious, etc (Urakami, Sutthithatip, and Moore 2020).

The frequency use of VA can influence emotional aspects of this voice-controlled device's user experience (Arlinghaus and Ollermann 2022). Players who have positive enjoyment, such as fun, pleasure, and enjoyment, with the target device often tends to affect their motivation to interact with the device continually (Kono and Araake 2022).

In summary, the contemporary approach of learning FL is focusing on positive emotion during learning. VA can provide positive feedback and emotion by offering personalized relations and social companionship. Understanding how to enhance user's enjoyment of VA and recognizing how to personalize VA to meet FL learners' needs are the main subjects in this research.

Result

This chapter describes the results from the first and second interviews and feedback from specialists. To understand the research target audience, this researcher conducted a first interview with participants recruited from campus and mainly focused on ELI. Through first interview, this research explores how culture, background, and environment influence learners' experiences while studying a new language.

After that, this author reached out to the same interviewees to understand their opinion with the new version voice assistant design approach. This research is also supported by MFA in Design specialists. Based on their knowledge and experience, this research pitches a better VA design for FL learners.

First Interview: The Motivation of Learning FLs

From the first interview, this research looks to understand each participant's motivation to learn FLs. Interviewees want to learn how to communicate with the native population in multiple situations or scenarios. The presence of multiple languages in the interviewee's hometown poses a challenge for people, hindering their ability to communicate effectively. In order to succeed in schools, workplaces, and daily life, interviewees want to learn how to communicate with the native population and they need the skill of speaking. In short, the learning motivation for participants will be communicating with others.

First Interview: Skills to Learn as a Beginner & The Challenges One Faces Learning FLs

When learning a FL, individuals begin with speech because speech is more important and it is typically easier to grasp than writing and reading. Also, interviewees want to speak with others to practice these skills and in general would like to have conversations on a frequent basis. However, finding the right person and the right time to talk and practice with a FL is a challenge many interviewees face. Because interviewees are not typically actively speaking with the native population, it can be difficult to frequently practice their speaking skills. On the other hand, some interviewees live in foreign places and feel anxious using FL to communicate with native speakers because they are afraid of saying something wrong, pronouncing it unclearly, or speaking too slowly. Some interviewees find it hard to know about local cultures due to linguistic gaps. Since they do not understand the local culture, environment, and background, they fear there could be something wrong with their speech and fear that native speakers will not understand them. Sometimes, the learners do not understand the meaning, accent, or vocabulary, during a conversation, and these facts cause them stress. Therefore, they will be unwilling to use this FL, and this decreases their self-confidence.

To support various gaps while learning an FL, this design aims to support target users' daily oral knowledge and basic cultural understanding. This research is not for FL learning beginners but for learners who want to have a more successful FL conversation experience.

Prototype Feedback

In the second set of interviews, this researcher reached out to the same group of participants (n=8) in ELI to understand their experience of the design process (see Appendix 3). By sharing this story and prototype, this author wants to understand their experience about whether this designed VA can support learning FL.

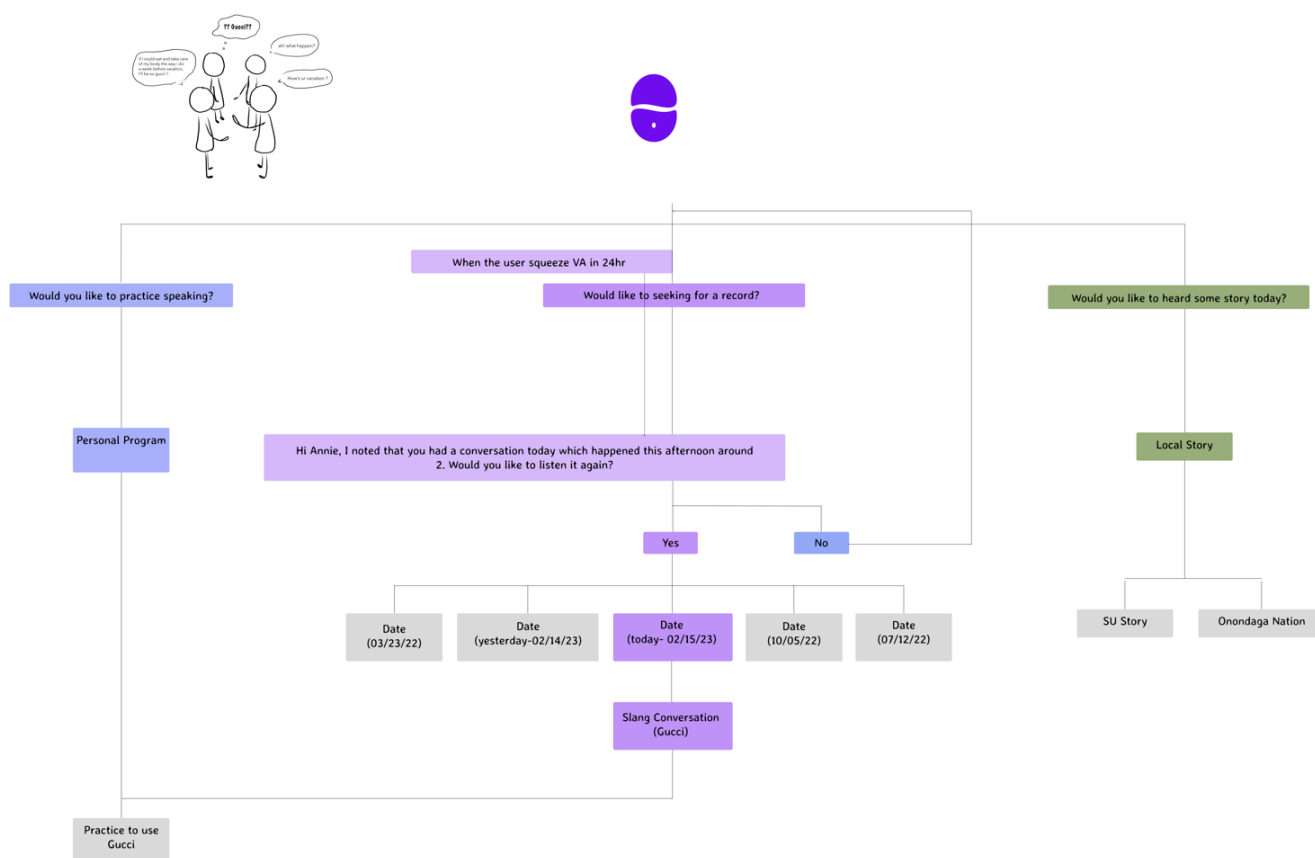<u>Structure of VA Design</u>

Figure 16 Structure of VA Software

After understanding the user's experience of the prior research and interview, this research builds a structure of the VA software (Figure 16).

The structure is shown and introduced to how the user can utilize this program. By creating a personalized VA, the user can create a personal account through a setting in the Kinakaian. This application connects with the voice assistant and users can see the visual structure of this interface. In the application, one can type in the user's name, native language, and the language the user wants to use; You can also type in VA's name, select a language, and a voice the voice assistant will use.

Besides, it provides two plans to the users: learning about different local cultures or supporting local stories. At the "learning different local cultures" section, users can select a location on a world map and learn a deeper level of this area (this research takes Syracuse City, NY, as an example); at the "supporting local stories" section, user can share one's story with one's the native language and experience by selecting a specific location on the world map (this section also take Syracuse City, NY, as an example).

Moreover, in the "learning different local cultures" section, users can see three categories: event, program, and local story. At the event section, the user, Annie, revisited the Gucci story to understand the meanings of this slang; at the program section, Annie practices with VA by using the term Gucci in conversation. The local story section will be supported by the local population and Annie can read or listen to the local story in this section. Annie can also add local stories at the beginning by selecting "support local stories". That is to say, users can exchange their stories with other users and learners can learn those stories locally.

<u>Features take Advantages from Successful Applications</u>

This design uses Figma to create the software prototype and involves advantages from

successful applications to have better support for users (Figure 17-22).
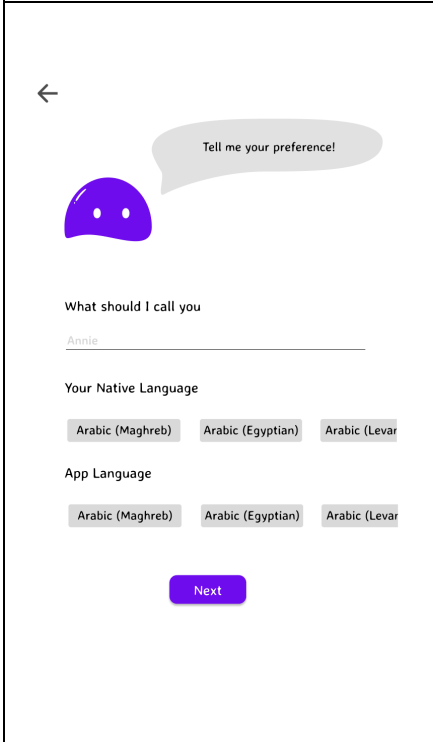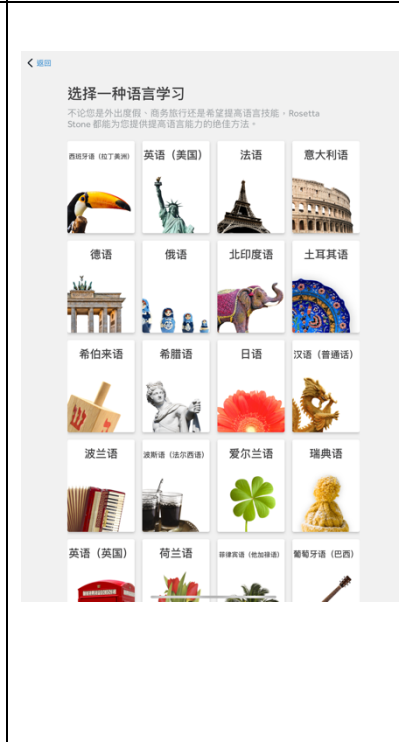
Kinakaian	Successful Application: Duolingo	Successful Application: Rosetta

Figure 17 Customize Language Section

Kinakaian	Successful Application: Duolingo	Successful Application: Rosetta

Figure 18 Utilize Illustration Support Understanding

Kinakaian	Successful Application: Duolingo

Figure 19 Sentence Completion Practice Section
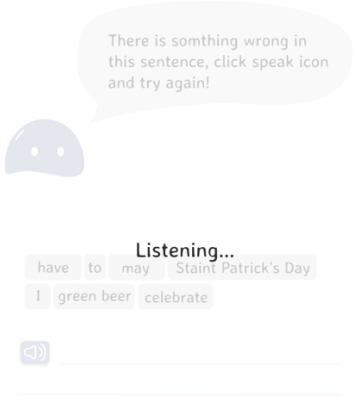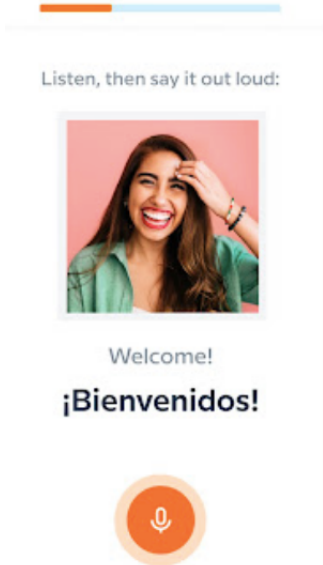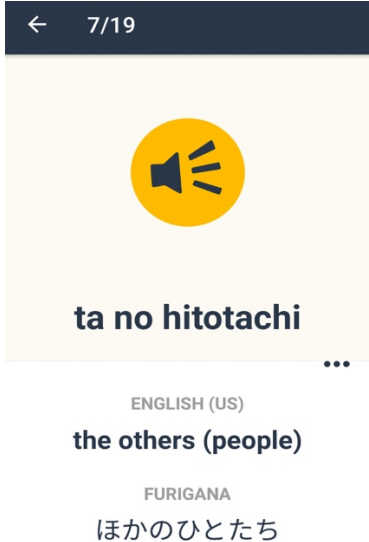
Kinakaian	Successful Application: Babble	Successful Application: Memrise
There is somthing wrong in this sentence, click speak icon and try again! Listening... have to may Staint Patrick's Day I green beer celebrate	Listen, then say it out loud: Welcome! ¡Bienvenidos!	7/19 ta no hitotachi ENGLISH (US) the others (people) FURIGANA ほかのひとたち

Figure 20 Oral Practice Section

Kinakaian	Successful Application: Memrise	Successful Application: WhatsApp
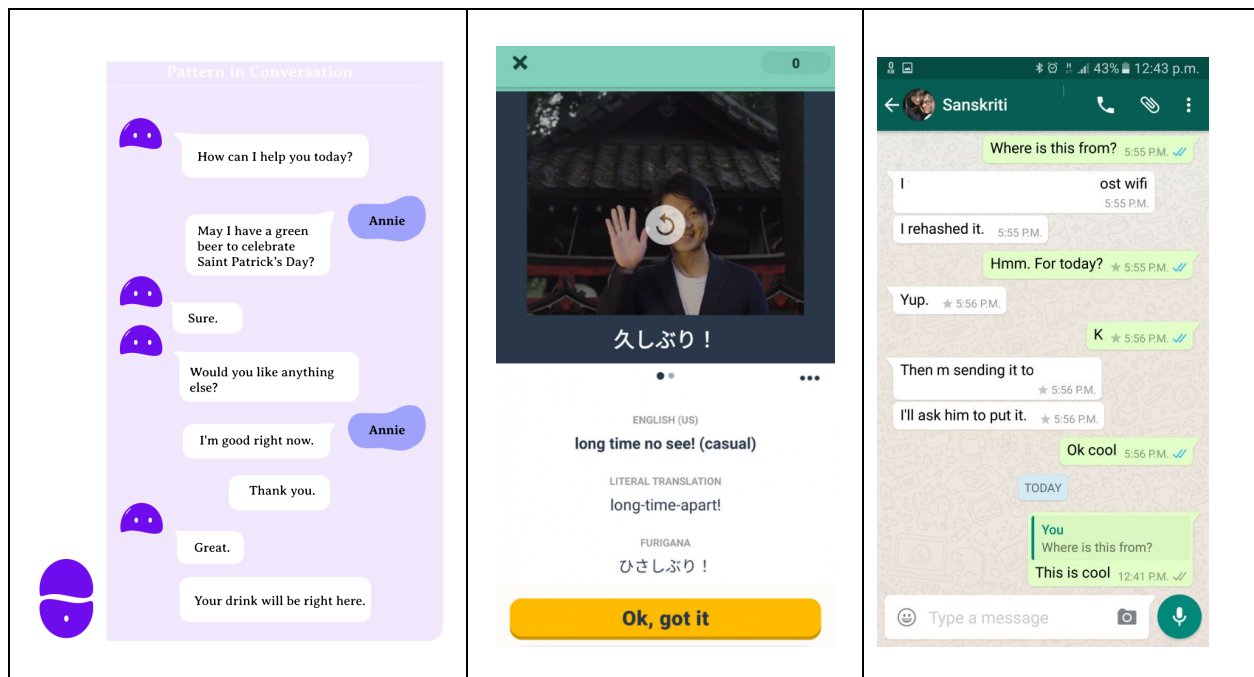
Figure 21 Social Practice Section

Kinakaian	Successful Application: Memrise
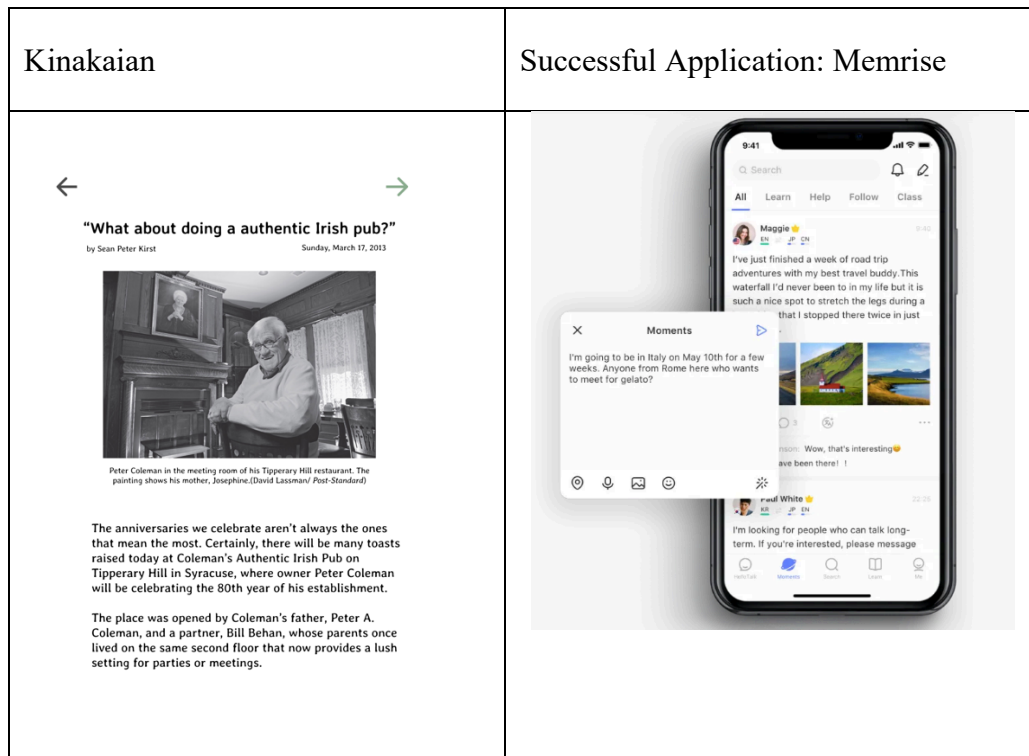
Figure 22 Personal Story Sharing Section

This design creates a knowledge-exchangeable database system. Users can share their own stories by selecting the location and sharing their knowledge (Figure 23). Learners from other locations can learn from this person's story by selecting this person's location on a map (Figure 24).
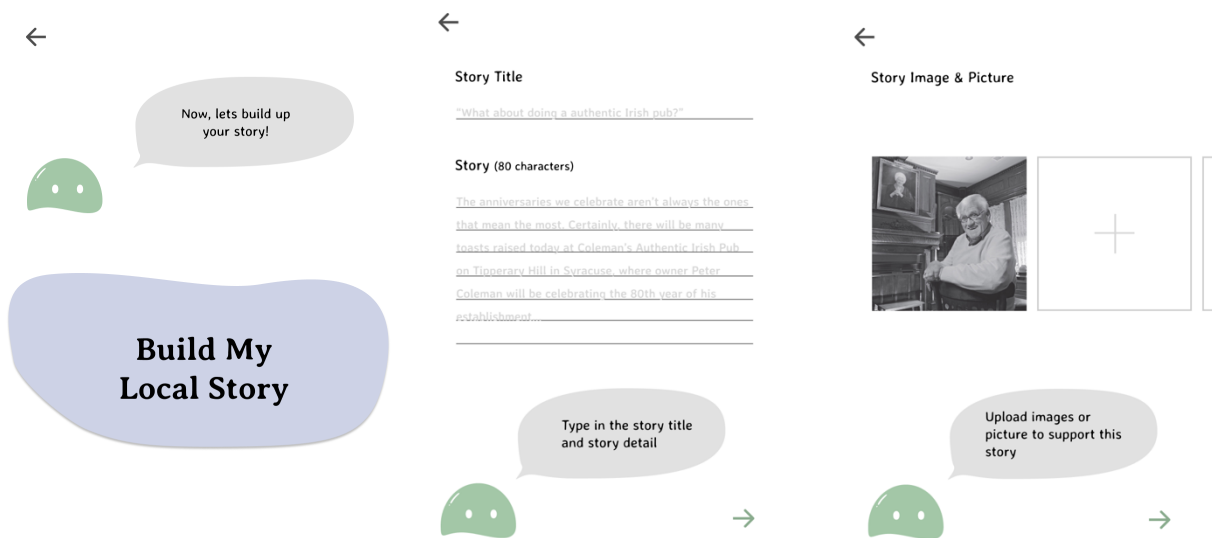
Figure 23 Users Build Own Story

Figure 24 Users Learn from Local by Selecting Location on Map

Five out of eight interviewees mentioned that they felt the local story is helpful, and

interesting, and it's important for one to know about different cultures and backgrounds. Six

out of eight participants felt the slang understanding and practice section is supported. Because

one can know about the meaning, it is good for beginners. Learning in a non-formal way is

also interesting. They feel that understanding background knowledge of FL is interesting

during learning.

Story Aspect

Storytelling in this design is a beginner's guide. This researcher play the video to

interviewees in order to understand if this design can be a way to support their FL learning. After

watching the video, they will understand when, why, and how the learner utilizes this tool to

support their learning. In the video, this author shot a group of people chatting and using the

term, Gucci, as slang to express their feelings (appendix 4). There is a FL learner, Annie, in the

group. She is confused during a chat with friends and has little understanding during and after the conversation. Therefore, she uses Kinakaian to document this pattern. After she heading home, she asks Kinakaian what is the meaning when people talk about this term.

Here are some notes from interviewees and professors after watching this video. One mentioned the term, Gucci, utilized in the story is not appropriate and this slang is not so familiar in daily life. However, slang terms or local culture knowledge used in this design is successful, because that knowledge changes from time to time, and this knowledge is seldom documented in the teaching of a FL.

Viewers feel confused to understand when and how the VA works in the video as well. Understanding the term the user learns is a base part of this design; learners will want to know the background or meaning. Practice speaking in learning FL is also an important aspect that should be more focused on in this design. How this practice will succeed in a conversation should be shown as well.

According to comments, this author switched Gucci to "Saint Patrick's Day" as an example to describe the design experience (Figure 25-30).
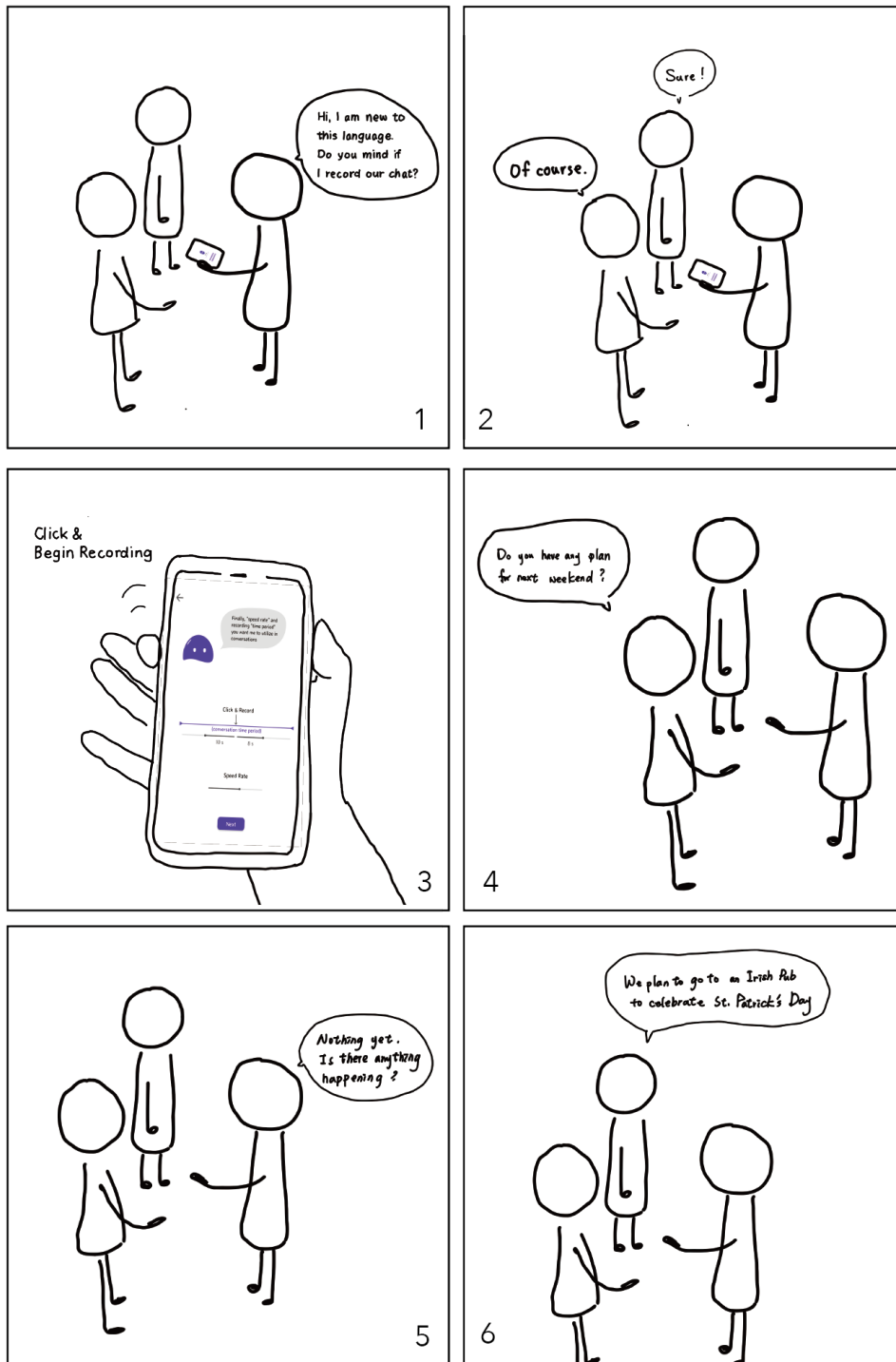
Figure 25 Saint Patrick's Day Storyboarding 1

Figure 26 Saint Patrick's Day Storyboarding 2

Figure 27 Saint Patrick's Day Storyboarding 3

Figure 28 Saint Patrick's Day Storyboarding 4
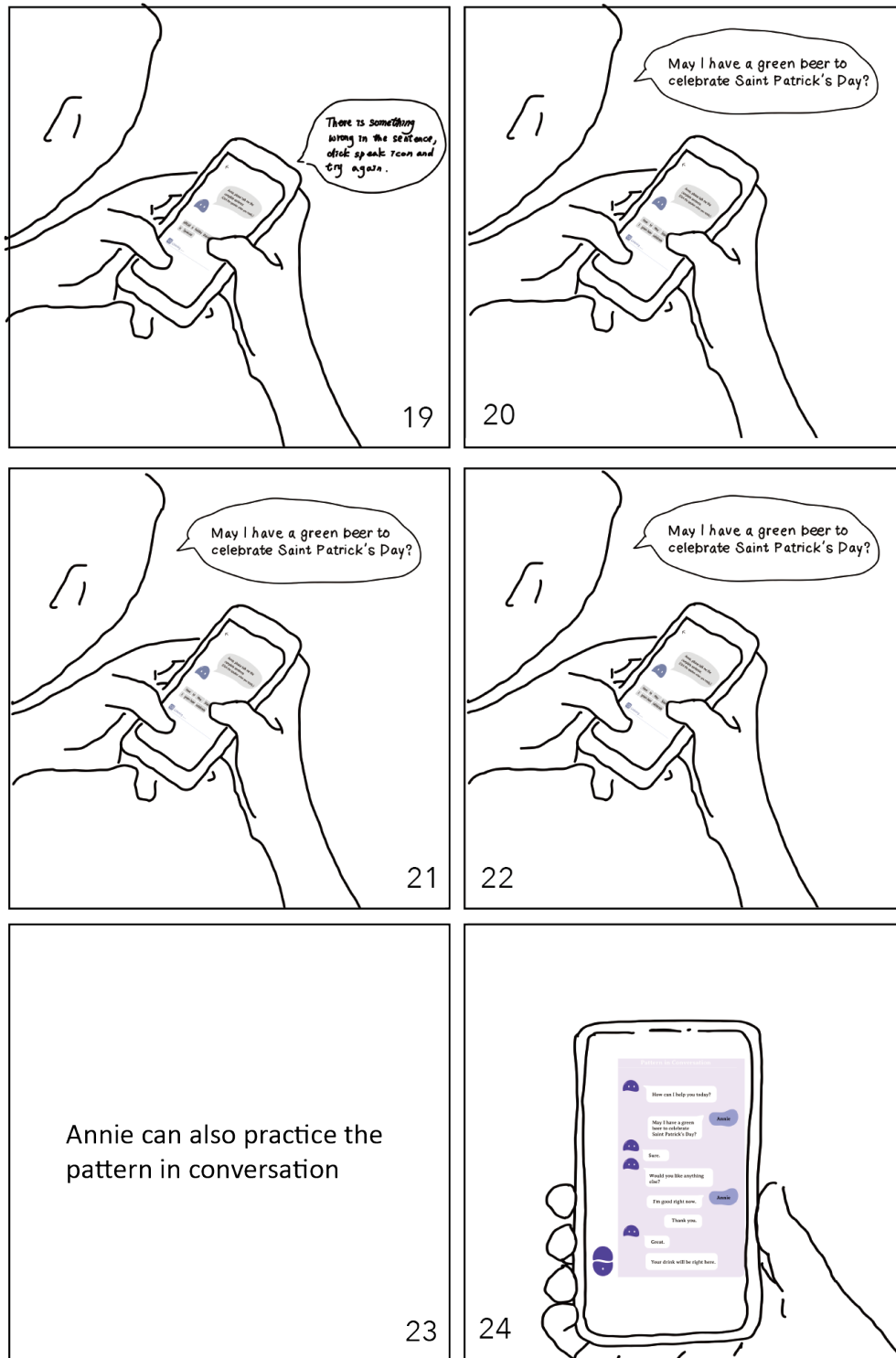
Figure 29 Saint Patrick's Day Storyboarding 5
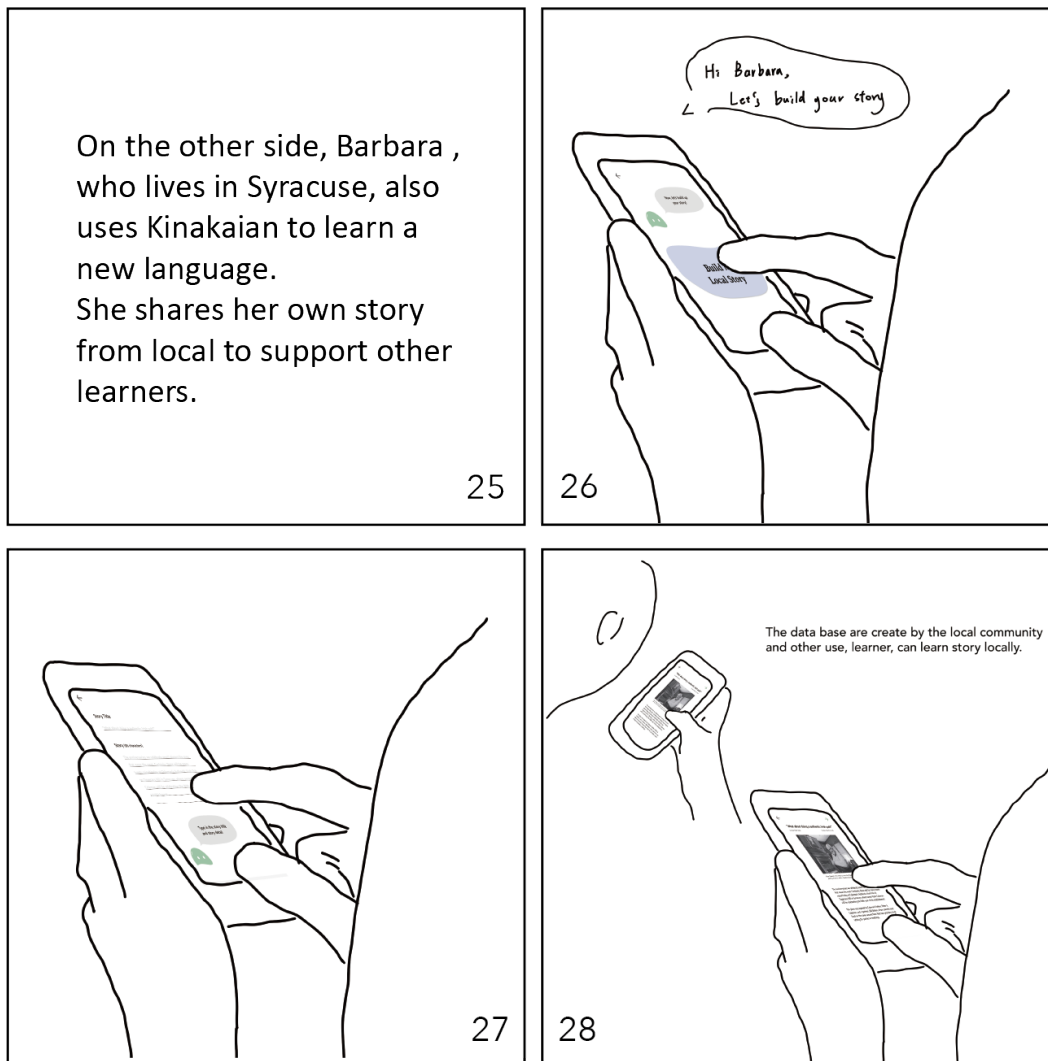
Figure 30 Saint Patrick's Day Storyboarding 6

"Saint Patrick's Day" is a holiday celebrated across the US and in many other English-speaking (or formerly English-colonized) places (particularly in places where large numbers of folks are descended from Irish immigrants.) The holiday is celebrated in Syracuse and this author experienced this culture personally. According to the reason above, this research recreated a video and sketch and described how learners learn about the "Saint Patrick's Day" experience by interacting with VA (please see Appendix 5 for color version, and Appendix 6 for video link).

Additionally, the new version still keeps the slang but changes it into "Lit" (figure 31).
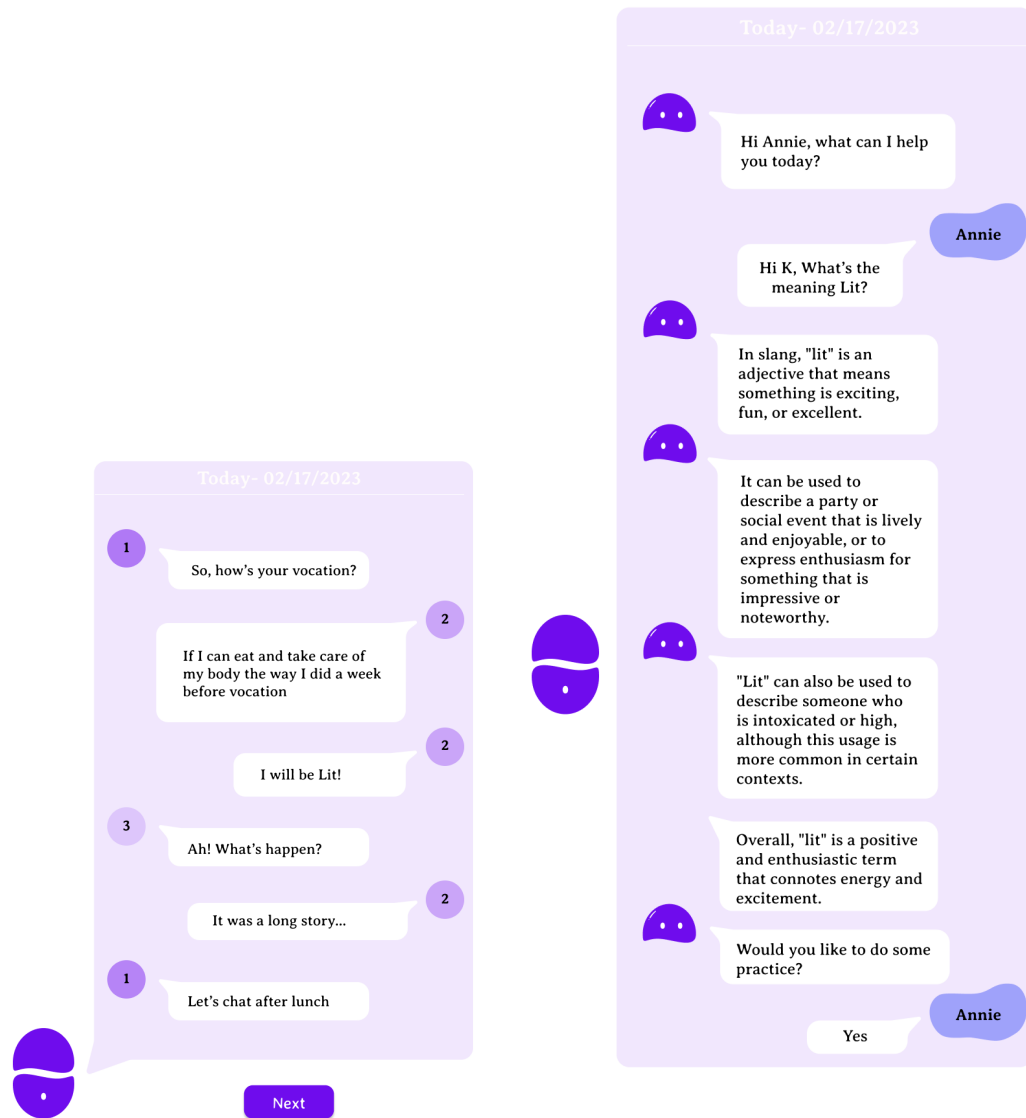
Figure 31 Slang Used in VA Conversation

There are several steps users need to set up at the beginning. In order to enhance personal needs, this VA asks users about their preference of voice, native language, target learning FL, and speed rate when this VA is talking (figure 32).

Figure 32 Personal Setting for Application and VA

Additionally, interviewees mentioned that if the user can change the recording length, it will be helpful because some users will want to document a shorter chat or longer speech. So, this program adds a function of setting the recording time slot section (figure 33).
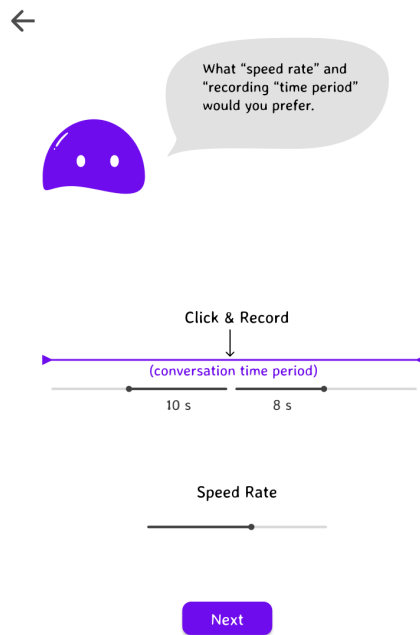
Figure 33 Personal Setting for Recording Time Period and VA Speech Speed Rate

This design uses a Map to search for locations where one would like to learn the language (figure 14). According to suggestions from environmental and interior design professor, this author should note that some people can't read maps and there should be some way to support this gap. Therefore, this program provides more options for location searching in order to meet the needs of users (figure 34).
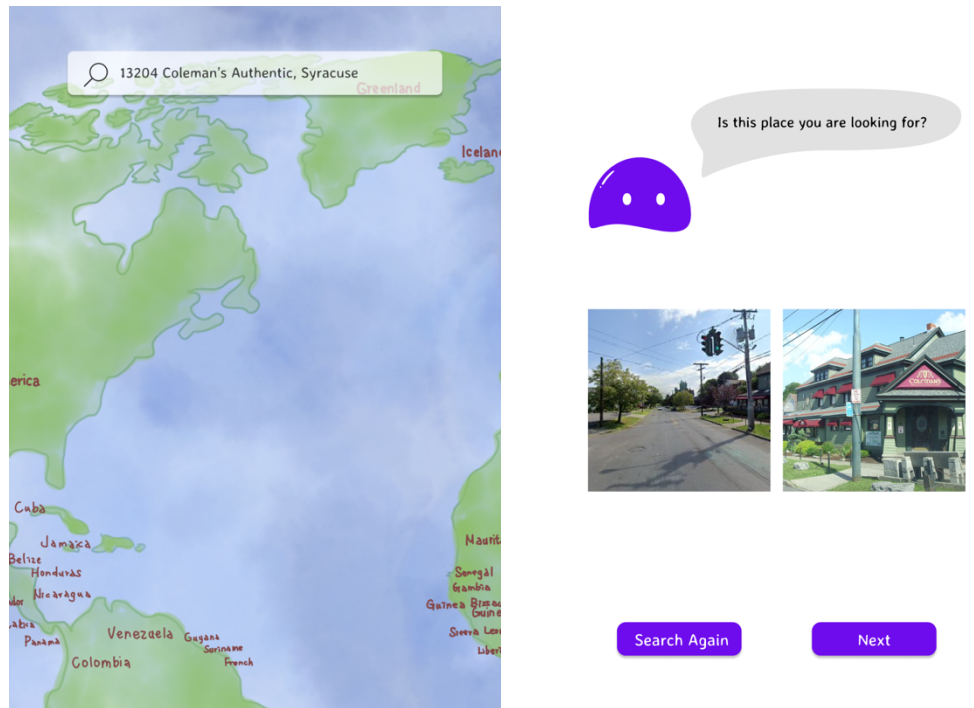
Figure 34 Location Searching Option

VA Support Learners

Base on target audiences' voices, visual support (such as pictures) to define a term would benefit their understanding. Because images and graphic symbols support understanding of the knowledge. On the other hand, this tool will be good to use during a lecture. Especially when the professor is a non-native speaker who speaks English and mostly use English to teach their class. Because pronunciation or accent can be different, this fact makes FL learners find it challenging to comprehend those courses. Through this design, the changeable accent of VA can support this experience. Furthermore, this tool decreases anxiety and makes one feel less stress. Talking to

the VA is less stressful than talking to a real person. They can use this tool without a time limit and bring it everywhere in one's pockets.

In accordance with that feedback, this author has provided a section that VA can help users search related images to define a term (figure 35). A practice section has been added to this program. Learners can use the term one just learned to complete a sentence (figure 36).
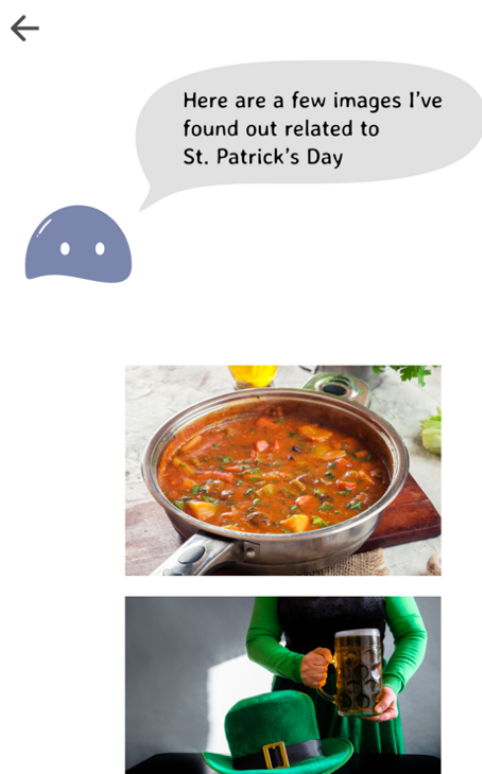
Figure 35 Visual Aids to Support Understanding

Figure 36 Practice Session 1
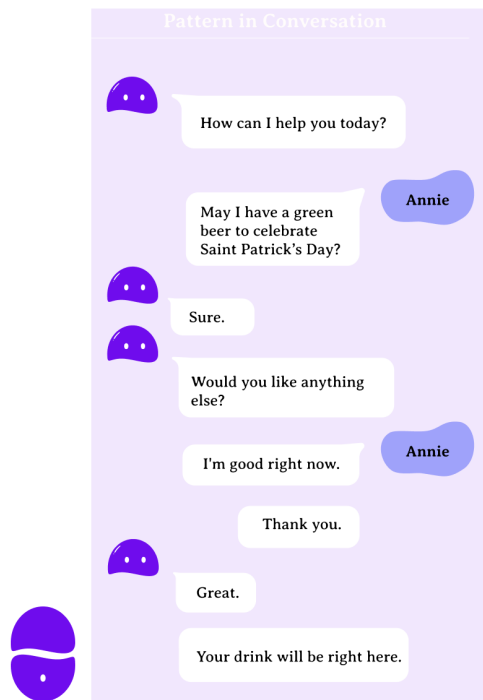
Figure 37 Practice Session 2

In Figure 37, Kinakaian shows learners can use this sentence, which had learned, to practice using it in specific situations. Moreover, users are able to practice other terms with this personal VA such as slang, and use native language to chat with Kinakaian (Figure 38).
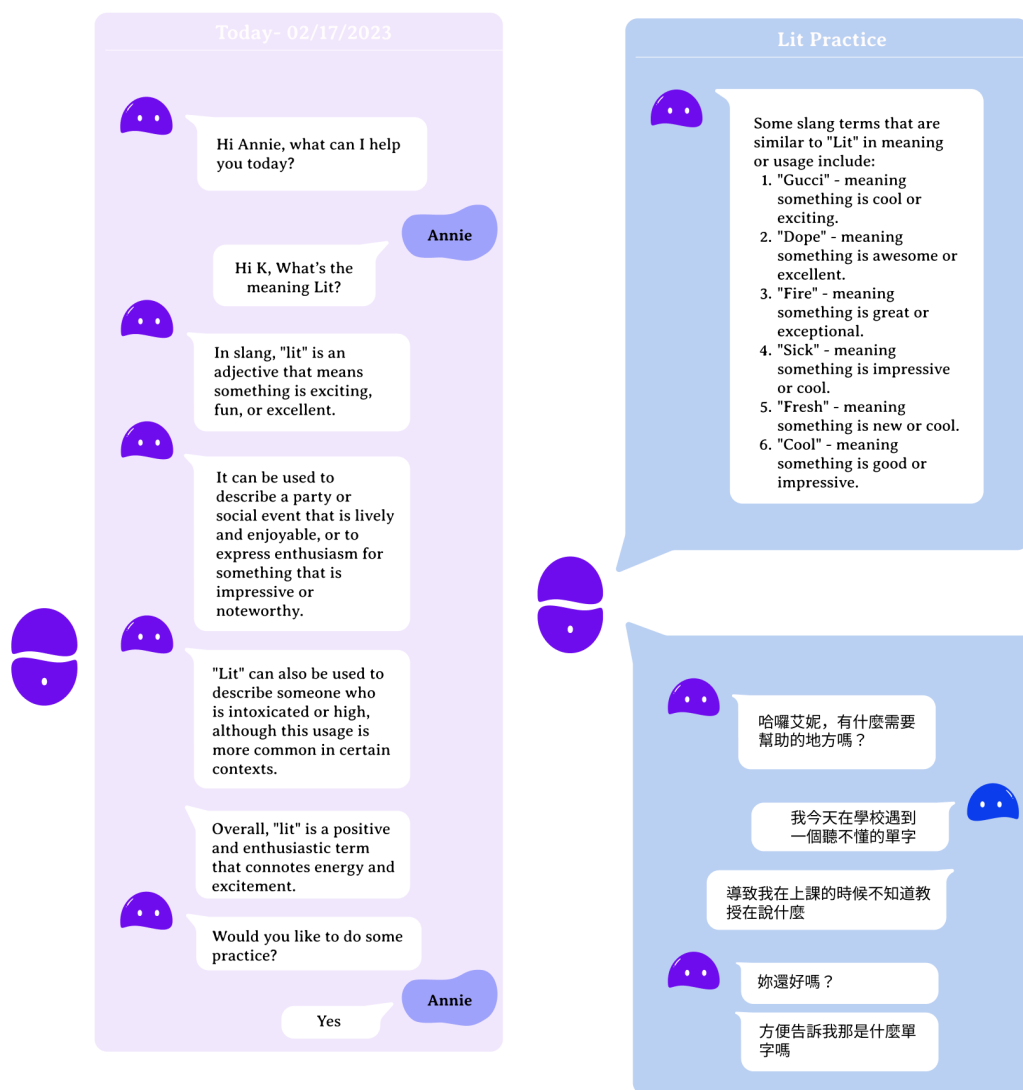
Figure 38 Practice in Variety Way

Appearance of the VA

At the physical prototype aspect, some noted that the squeeze ball is too big and not suitable for people to carry on. Squeeze ball might not be a good way to be a VA because one doesn't want to squeeze one's language learning teacher. Moreover, this tool can be a cell phone charm that users can easily carry.

Based on those responses, this author has reduced the physical porotype and moves to a more functional way to record conversation. By clicking one cell phone, users are able to document the chat easily (figure 39).
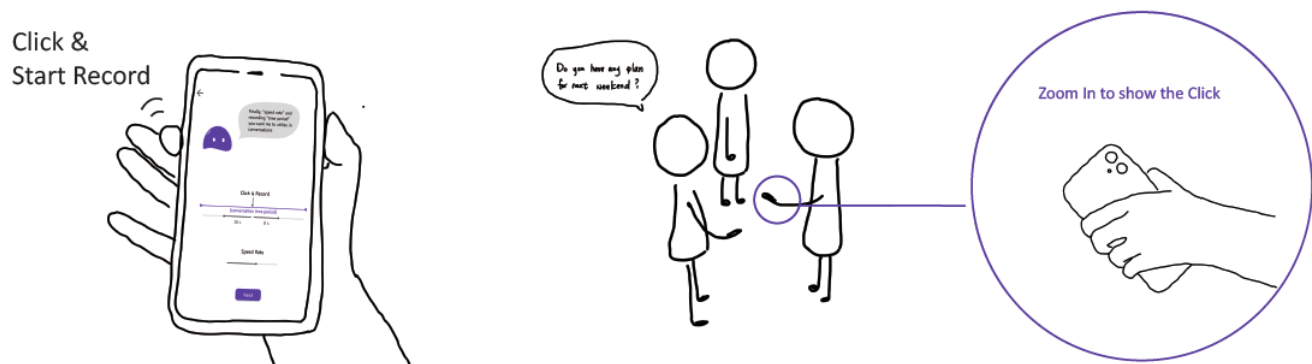
Figure 39 Way to Document and Record Conversations

Privacy Concerns

Privacy concerns are significantly prioritized in HCI research. People should understand how this VA will record, and who will be recording when it is recording. This VA introduces two functions for careful craft privacy concerns: limited time period function and declaration of audio notification tags in the product.

Limited time period function: this VA will record the conversation before and after the click, which means that if one clicks the phone, this VA will record 20 seconds before and after the

click. In order to achieve this function, the VA will ask users' awareness if they agree to have their speech recorded. After gaining permission, VA will record 20 seconds, save it to the Random-Access Memory (RAM) and then delete it after 20 seconds. When the user clicks and begins recording, this VA will move the 20-second recording that was recorded before clicking from the RAM to Intel Optane Persistent Memory (PMEM) in order to document it for later use. On the other hand, users can also set how long this RAM time slot should be documented (Figure 33). Also, this author shows at the beginning of the video that Annie asks permission to record before a group chat (Figure 25).

Declaration of audio notification tags in the product: in order to get noticed that the user is going to record this chat by other colleagues in a public conversation, users should turn on the audio function in this product. According to Ayoade, some countries, like Japan for instance, force manufacturers of portable cameras to incorporate mandatory built-in audio features; this way ensures that individuals cannot discreetly capture images of others without their awareness (Ayoade, 2007).

In summary, Kinakaian contains growing privacy and security concerns in part in order to ensure better usefulness for the public. The limited time period function and the declaration of audio notification tags in the product can provide more flexibility way for people to cover their needs.

Limitations and Future Research Directions

This study is limited by the coding skill of the author which impeded continuous interaction with the VA. Lacking of knowledge about computer science limited the VA's function of big data collection. Future studies may consider employing a computer engineering expert to create a more functional or natural VA interaction. The study was also limited by a general lack of access to any form of a comprehensive data-base for local knowledge. If the target audience could choose a specific location and learn the local knowledge, it might be more successful and bring more enjoyment for the user while learning a FL. Moreover, this research recruited participants from the student population at Syracuse University which in turn limited other target audiences from various backgrounds and cultures. Therefore, future research can and should focus on individuals other than college students so as to better understand other people's experiences as FL learners.

Conclusion

This research understands the target audience's FL learning experience captures their needs of utilizing VA support learning and designs a personal voice assistant. The findings of this study lend mixed support to the FL learner that has successful effects on extending one's FL knowledge aspect. This author shares one's experience as a challenge and shares ideas with target users to understand their needs. This design was worked with the target audience in progress in order to capture a variety of aspects of learners' experience. By utilizing this design, target audiences feel less anxious if this VA is nearby, they can practice FL by interacting with VA, and customers can learn updated local stories and knowledge without a time limit. Given the growing technological research, study, and product rollout of social robots, this author hopes that this research can support HCI studies and retailers to build more diversity and resilience in the target audience experience study and market environment.

Appendix 1

This study understand about target users' enjoyment about VA's variety pitch value and speed rate. This research is a poster and accepted by HCI INTERNATIONAL 2023, 25TH INTERNATIONAL CONFERENCE ON HUMAN-COMPUTER INTERACTION.

Utilizing Different Voice Value to Understand Voice Assistant Users' Enjoyment

Ai-Ni Cheng

Syracuse University, Syracuse NY 13244, USA
900 South Crouse Ave. Syracuse, NY 13244, USA
`aicheng@syr.edu`

Abstract. This study aims to understand if various voice assistants' (VAs') voice parameters (i.e. pitch range and speech rate.) that may have effect on users' enjoyment and preference. VAs follows various voices as Siri has five kinds of recordings in the English (American) channel. For the reason that Siri has diversified voices, this research aims to use a non-binary voice (i.e. ios 15.4, recording 5) as a sample and use AUTO-TUNE software to modify their pitch ranges and speech rates. Through a web-based survey conducted on Qualtrics, participants—recruited from Amazon Mechanical Turk (MTurk)—were asked their feelings after listening to 4 patterns, which included higher pitch range, lower pitch range, higher speech rate, and lower speech rate. I ask about their enjoyment of listening to those patterns across five different categories: motivation, goal, proof, evaluation, and potential. In a sample size of 114, this study discusses how the responses differ between the different categories and how they link to various VA voice parameters. In this study, I demonstrate that various VA voice parameters have no significant effect on users' motivation, goals, proof, and evaluation. Users' potential is slightly affected by the changeable voice parameters of this human-computer interaction however it is still not significant. With this study, I indicate that the pitch range and speech rate of the voice assistant, Siri iOS 15.4 recording 5, has no effect on the enjoyment of their user's experience.

Keywords: Human-Computer Interactions, Voice Assistant, User Experience, Nonverbal Perception, Enjoyment, Voice Parameter, Siri.

1 Introduction

1.1 Voice Assistants (VAs) and VA's Voice

Voice Assistants (VAs) are often used as a form of personal assistant for many people. VAs are voice-controlled mobile or personal devices which allow users to communicate with, to make some queries and have emotional connection VAs, such as Google Assistant, Microsoft Cortana, Amazon Alexa, Apple Siri and Yandex Alice [1–3].

During VA's interaction, users prefer more personalize voice assistants [4–6]. VA has personalized relation during interaction are more effective than impersonalized such as changed pitch value [4]. Social robots, included VAs, during the interaction

between users are seeking to develop in a socially enriching way[5]. By personalized users' engagement during the interaction, this fact will raise the relationship between users and social robots [6].

On the other hand, VAs' voice might be a way to increase VAs' personalization because voice is a way to transport information [7, 8]. Pitch range and speech rate are the fundamental characteristics of the voice that indicate personality [7]. People who have loud, rapid, and broad-pitched voices are most likely extroverts [8]. Moreover, voice is important for people's enjoyment. Niculescu et al (2011) state that using different pitches can increase overall enjoyment and interaction quality [9]. That is to say different voices' value in voice assistants might influence user experiences.

1.2   Users' Experience (UX) and VA's Voice

Enjoyment is defined as a good experience [10]. Enjoyment is the effect produced by functional and nonfunctional interactions and reflects the users' feelings during the experiment. It includes enjoying the interaction while using the voice assistant such as feeling nervous, satisfied, anxious, etc. [11].

The frequency use of VA can influence emotional aspects of this voice-controlled device's user experience [1]. Players who have positive enjoyment, such as fun, pleasure, and enjoyment, with the target device often tends to affect their motivation to interact with the device continually [12].

In short, this study aims to use enjoyment as a standard to value the users' satisfaction with various voice parameters of VA. By understanding the relationship between user's enjoyment and VA's multiple pitch range and speech rate, this research hypothesizes that VA's various voice parameters can influence users' experience and enhance their enjoyment of the interaction.

2   Methodology

This author has completed the Collaborative Institutional Training Initiative (CITI) Training and taken the Social/Behavioral Research Course (Appendix 1). This research is approved by Syracuse University Human Research Protections Program and the Institutional Review Board (IRB) authorization has been given (Appendix 2). After receiving IRB approval for the study protocol, changed voice parameters were created using AUTO software, and a web-based survey was conducted using Qualtrics.

2.1   Voice Parameters Created

Study voice parameters were created through AUTO software, Voice Recorder Pro, an audio editor platform widely used for signer, musicians, and creators. This software supports sharable files in iTunes and iCloud Drive, and has an affordable interface for users to record audio, audio mixing, modify voice value, etc. [13]. To select a more neutral voice for this study, I used IOS 15.4 Siri recording 5 in English (American) as a sample. Recording 5, according to Apple, was recorded by a non-binary person and

as the reason to provide more diversified-choice voices for users [14]. As a second step, I studied the frequency pattern consumers used and chose a pattern with questions because the question pattern is classified in high-arousal target audiences emotion [15]. I recorded the pattern by asking how's the weather and recorded Siri's response "The forecast is calling for a mix of rain, partly cloudy skies, and clear skies this week. Daytime temperatures will fluctuate between 43°F and 66°F, with overnight lows between 28°F and 52°F." Moreover, the speakers' voice affects personal attributions due to being raised or lowered by 20% of pitch or expanded or compressed by 30% of speech rate [7]. Therefore, I utilized the AUTO software to modify this record into four different pieces of audio: raised 20% of pitch, lowered 20% of pitch, expanded 30% of speech rate and, expanded 30% of speech rate.

2.2   Participants Recruitment

Study participants were recruited through Amazon's Mechanical Turk (Mturk), a data collection platform widely used for survey or experimental studies. Despite some constraints, many studies have found that the MTurk sample and the benchmark survey have no statistically significant difference [16]. To select the most appropriate survey participants for our study, only MTurk panelists with a previous Human Intelligence Tasks (HIT) approval rating 95% or higher were allowed to participate in the survey. As a second requirement, since our survey is conducted in English and audio pattern recorded in the United States, participation was restricted to MTurkers located in the United States. Each participant who successfully finished the survey to the end received a $0.5 reward.

2.3   Survey and Data Cleaning Process

Participants will be asked to answer survey questions about their enjoyment of voice assistant voice recordings. The survey will be based on Urakami et al research [11], the effect of naturalness of voice and empathic responses on enjoyment, attitudes and motivation for interacting with a voice user interface ,and Kono & Araake's research [12], Is it Fun?: Understanding Enjoyment in Non-Game HCI Research, to build questionnaires to measure participants' enjoyment. The questionnaire will measure how much participants enjoyed the interaction answering on 5-point Likert scales ranging from 1 (Strongly disagree) to 5 (Strongly agree) regarding enjoying the interaction; the questionnaire will measure subjective enjoyment, the questionnaire statements are in forms such as: "I enjoyed ~," "~was fun," "How enjoyable ~," and "~was enjoyable."

   The questionnaire used five classifications to understand the enjoyment: background and motivation, objective and goal, support and proof, results and evaluation, and application and potential [12]. Background and motivation refer to the reason why fun is required or important in VA' voice pitch or speed rate; objective and goal refer to VA's voice pitch or speed rate aims to make users feel fun; support and proof refer to VA's voice pitch or speed rate is a supportive element to strength the pleasure of HCI; results and evaluation refer to VA's voice pitch or speed rate in HCI is enjoyable;

7. Apple W, Streeter LA: Effects of Pitch and Speech Rate on Personal Attributions, Journal of Personality and Social Psychology (1979).
8. Eden BL: Clifford Nass and Scott Brave's Wired for Speech: How Voice Activates and Advances the Human-Computer Relationship. Journal of the Association for History and Computing (2006).
9. Niculescu A, Dijk B, Nijholt A, See S: The influence of voice pitch on the evaluation of a social robot receptionist (2011).
10. Brandtzæg PB, Følstad A, Heim J: Enjoyment: Lessons from Karasek. In: Blythe M, Monk A (eds) Funology 2. Springer International Publishing, Cham, pp 331–341 (2018).
11. Urakami J, Sutthithatip S, Moore BA: The Effect of Naturalness of Voice and Empathic Responses on Enjoyment, Attitudes and Motivation for Interacting with a Voice User Interface. In: Kurosu M (ed) Human-Computer Interaction. Multimodal and Natural Interaction. Springer International Publishing, Cham, pp 244–259 (2020).
12. Kono M, Araake K: Is it Fun?: Understanding Enjoyment in Non-Game HCI Research (2022).
13. Linfei Ltd: Voice Recorder Pro, App Store (2022).
14. Axon S: HEY SIRI - Apple will add fifth US English Siri voice in iOS 15.4. In: ars TECHNICA. https://arstechnica.com/gadgets/2022/02/apple-will-add-fifth-us-english-siri-voice-in-ios-15-4/ (2022).
15. El Ayadi M, Kamel MS, Karray F: Survey on speech emotion recognition: Features, classification schemes, and databases. Pattern Recognition 44:572–587. https://doi.org/10.1016/j.patcog.2010.09.020 (2011).
16. Buhrmester MD, Talaifar S, Gosling SD: An Evaluation of Amazon's Mechanical Turk, Its Rapid Rise, and Its Effective Use. Perspect Psychol Sci 13:149–154. https://doi.org/10.1177/1745691617706516 (2018).

Appendices

Appendix 1. Collaborative Institutional Training Initiative (CITI) Training

Completion Date 17-Sep-2021
Expiration Date 16-Sep-2024
Record ID 45139862

This is to certify that:

Ai-Ni Cheng

Has completed the following CITI Program course:

Not valid for renewal of certification through CME.

Basic/Refresher Course - Human Subjects Research
(Curriculum Group)
Social/Behavioral Research Course
(Course Learner Group)
2 - Refresher Course
(Stage)

Under requirements set by:

Syracuse University

CITI
Collaborative Institutional Training Initiative

Verify at www.citiprogram.org/verify/?w5cb96a31-6fe2-451c-9136-8eb813f64abf-45139862

Appendix 2. Institutional Review Board (IRB) Authorization

Syracuse University

TO: James Fathers
DATE: November 4, 2022
SUBJECT: **Determination of Exemption from Regulations**
IRB #: 22-334
TITLE: *Design a Personalize Voice Assistant by Utilizing Different Voice Value to Understand the Enjoyment of Voice Assistant Users*

The above referenced application, submitted for consideration as exempt from federal regulations as defined in 45 C.F.R. 46, has been evaluated by the Institutional Review Board (IRB) for the following:

1. determination that it falls within one or more of the eight exempt categories allowed by the organization;
2. determination that the research meets the organization's ethical standards.

It has been determined by the IRB this protocol qualifies for exemption and has been assigned to category **2**. This authorization will remain active for a period of five years from **November 4, 2022** until **November 3, 2027**.

CHANGES TO PROTOCOL: Proposed changes to this protocol during the period for which IRB authorization has already been given, cannot be initiated without additional IRB review. If there is a change in your research, you should notify the IRB immediately to determine whether your research protocol continues to qualify for exemption or if submission of an expedited or full board IRB protocol is required. Information about the University's human participants protection program can be found at: http://researchintegrity.syr.edu/human-research/. Protocol changes are requested on an amendment application available on the IRB web site; please reference your IRB number and attach any documents that are being amended.

STUDY COMPLETION: Study completion is when all research activities are complete or when a study is closed to enrollment and only data analysis remains on data that have been de-identified. A Study Closure Form should be completed and submitted to the IRB for review (Study Closure Form).

Thank you for your cooperation in our shared efforts to assure that the rights and welfare of people participating in research are protected.

Tracy J. Cromp
Tracy Cromp, M.S.W.
Director

DEPT: VPA – School of Design, 100 Crouse Dr, Syracuse, NY 13244   **STUDENT:** Ai-Ni Cheng

Office of Research Integrity and Protections
214 Lyman Hall, 100 College Place
Syracuse, NY 13244   T: 315.443.3013
                                                                   orip@syr.edu

Appendix 2

This section included two sets of in-depth interview survey.

In- depth Interview 1

Do you have experience learning foreign languages?

1. What languages have you learned?

2. Please tell me about your experience: why do you want to learn a foreign language?

3. Do you prefer learning alone or with other people? Why?

4. Could you please describe your experience during learning a new language? (Learning

alone- How do you practice? Why?; Learning with others- Who do you talk to? Why?)

5. Have you hired someone or attended classes to support your learning? What was that

experience like? Describe your ideal practice environment.

6. How do you practice speaking your new language in a supported way? What were the

difficulties you experienced?

7. Have you found any helpful tools to support your learning? (note, software, app, etc.) (If

yes, please answer the question a. and b)

      a. How often do you use this tool?

      b. Why do you feel this tool is useful?

8. What are the biggest challenges in learning a foreign language? What have you tried that

didn't work?

9. Are there learning approaches you'd like to try, but haven't been able to? What are they

and what's the reason that they are difficult to try/access?

10. What are the most difficult things about learning a foreign language?

11. What kind of support or learning language tools do you wish were available?

After Testing the Modular/In-progress Prototype:

1. How's your experience after trying the modified voice assistant?

2. Do you feel this tool can support your learning experience? Please provide reasons for the

answer.

3. Do you feel the modified speed rate supports your learning experience? Please provide

reasons for the answer.

4. Describe an ideal environment in which you would like to interact with this experience.

5. Were you satisfied with the tool? Would you like to change anything about them to make

them feel supported?

6. Do you enjoy the interaction experience with voice assistants? Why or why not?

7. Are there any additions or changes to voice assistants that you would like to see that you

think may help to support learning foreign languages?

After Testing the New Modular/In-progress Prototype, In-depth Interview 2:

1. How do you feel after trying the new vision voice assistant?

2. Do you feel this tool can support your learning experience? Please provide reasons for the

answer.

3. Were you satisfied with the tool? Would you like to change anything about them to make

them feel supported?

4. Are there any additions or changes to this new version that you would like to see that

you think may help to support learning foreign languages?

Appendix 3

These interview notes was documented base on the interviewees described.

Survey 1

Interviewee learned Spanish as a foreign language in school and did not feel super

familiar. The reason why this person would like to learn is because one wants to learn to

converse with more people in different places. Interviewee feel learning along and with other

people all support learning. One can write alone; listen, speak and with other people. The

experience was like doing assignments and taking quizzes in class. One feels attending a class

may be helpful cus it can practice and have fun. One's ideal environment is that there is an

activities-based place for a long term and one can reach a goal during learning.

Interviewee prefers speaking first to learn a language cus this person believes speaking is

easier than reading and writing. Apps, Rosetta Stone and Linguee, support learning. Interviewee

think the fill in blank of voc is not work for everybody. The challenge was hard to do it all the

time and stay focused on it. One would like to try a class with one on one teacher on-line, which

is not like a robot and can tell what is wrong. One doesn't have time to try a tutor and also don't know where to assess.

After experiencing the idea

The idea makes sense but needs to concern the danger of universal translation. This tool is practical. Ex: the user can ask how to use we vs. me. The modified speed rate plug-in can support learners when one wants to learn native speaker fast speech and can use the speed during learning by increasing speed.

Would like to use this tool during personal travel and would like to have text with this tool. Also like to use this device with headphones without text and can have a chat.

Survey 2

Learning Hindi, Marathi and basic Gujarat as foreign language. Want to communicate better in one's hometown cus there is so much diversity. Prefer learning with others and it will help to remember. Talk with friends, family with people who know the language and those people will know how to pronounce specific words. Try to learn language by oneself but have not yet got time to do it. Learning by practicing speaking with others and the person can correct oneself.

People are the best way to learn language and would like to talk weekly. The biggest challenge is not getting a chance to speak every day, and trying to watch foreign language shows but not work. Would like to try watching foreign language shows with subtitles every day again.

Understanding the rules of foreign language is the most difficult thing. Wish a app with translation during conversation is available and this tool can support during chat.

<u>After experiencing the idea</u>

Feel this idea is helpful and convenient. Users can ask this VA during learning. The tool is always with the user and the speed should be always slow. The speed plug-in should be put in the background. The interactive experience is good and one likes to try when it is available. This tool should connect other functions such as smart home or google map. Also, there should be a translated picture showing foreign and native language.

Survey 3

Learning language as a second language. This interviewee would like to learn in America for a future career and also would like to pass the test which requires applying to American University. One prefers learning alone cus this person finds it hard to talk with unfamiliar people and also cus one feels one's foreign language is not fluent. The experience during learning is remembering voc and using APP Tandem to chat with people online. This app can reach out to people who would like to learn the target user whose native language is one's learning foreign language. Users can add friends and will not feel stress cus they are all unfamiliar with each other's language. Practicing talking with people is helpful.

One participates in classes for learning and for tests. Those classes focus on listening, speaking and writing skills for TOFEL. Also, there is native English speaker to teach conversation class, and students can talk with native speakers. The biggest challenge this

interviewee meets is grammar during talking. One was trying to listen to music but it isn't so useful cus it is hard to focus and also it is not a complete conversation. One feels listening is the hardest part and wish there is a way to practice listening which is interesting, not stressful and can let the user become familiar with the English environment.

After experiencing the idea

The experience with reply is interesting. One can understand the different cultural backgrounds and learn how to respond more locally like native speakers, like how to call people in a polite way. This tool can support learning and it feels closer to reality and more emotional. It is like a friend, not so formal and can provide different choices for users. The speed rate setting can get more understanding of the speaker's meaning and can increase the speed rate to familiarity with native speakers' speech rate.

The ideal environment of this experience would be like using it along for practice and also can use it in school when meeting some situation don't understand. The thing one would like to add is the accent, if the accent is changeable, it will be helpful and one can understand this language more. The recent voice assistant is not so close, so would like the VA like a friend and feel closer and would like to see this tool can provide more culture knowledge and diversity.

Survey 4

Interviewee has experience learning foreign languages which are English and Spanish. The reason one learns languages is because they live outside of their hometown and need the skill for daily life. One prefers learning alone at the beginning and prefers learning with others

for conversation, non-formal speaking. Interviewee feel learning to write it differently. One practices a new language by talking with people, one-self or reading and watching movies. Those persons this participant would like to talk to are family, friends cus they are more familiar. Also, one would like to watch YouTobe to learn from YouTuber. Learning English since 4 y-o and had a personal coach in primary school. Use software Citrstone to learn to speak and use Dolingual to remember vocabulary. They have voc tests and questions to respond to. However, one feel learn in the native environment is more helpful cus one needs to use this language in daily life.

About the personal coach, it was a useful experience. The coach teaches how to speak and pronounce. She gave immediate feedback but sometimes this experience was stressful cus she always fixed her eyes on the interviewee. The most difficult part during the experience is to remember the spelling and grammar. The learning approach one would like to try will be to join the SU language partner. However, one always feels it is stressful and uncomfortable with social interaction. One wishes that the human brain of the chip is available and doesn't need to remember or learn a different language.

After experiencing the idea

It is a good idea. Users can use Siri to help spelling and to improve chatting skill with the device. One feels this tool can support learning because it is continued and can have quick feedback. The modified speed rate is supported at the beginning. After that, one familiar with native speaker speed, the speed rate might not be a problem. One would like to use this too when writing papers, reply to an email to a professor, and hope there is a copy pass function in software. Also, this tool can learn how to pronounce and spell words. Interviewee enjoy the

interaction and like it look like a friend and can help personal projects. The addition would be to teach how to spell voc.

Survey 5

Interviewee has experience learning foreign languages: English, Spanish and French. The reason one wants to learn is because one likes to know more about knowledge and would like to work in another place. One prefers learning with others because they can learn from mistakes. Learning a new language is difficult and learning in class make it harder cus people in class have little knowledge with this language and they can't distinguish the mistake during talking. Living in hometown is hard to find a chance to practice and talk with foreign language speakers. Practice conversation with people but don't know how to pronunciate in the right way. Watching English videos with subtitles is supportive. The biggest challenge in learning foreign language is getting ideas for writing.Gammer is the most difficult thing in this experience. Would hope that there is a robot that can write an essay for the interviewee.

<u>After experiencing the idea</u>

It is a good idea and good to practice with this voice assistant. The VA can teach and bring new knowledge to the user. It can correct sentence pronunciation and specific words. Want to use it at home. The speech part is important. Also, it should correct the user with the correct word. Would like to see different look such as different color, there are different options that can suit different people.

Survey 6

Interviewee has experience learning foreign languages: English. The reason one wants to learn is related to one's future career. One prefers learning with others because speaking is important. Also, speaking is difficult and living with someone who has an English background will be helpful. The ideal environment will be to experience the culture. Pronunciation is hard for japan cus there is no L sound in one's mother language. Using subtitles with 2 languages when watching a film is a good way to learn. Personal accent make it hard to let friends understand what one is saying. Remember voc is the most difficult thing in this experience. Hope there is a conversation partner or AI can talk with.

After experiencing the idea

Indeed need it! Siri is hard to distinguish one's speech. Would like to use it while driving a car. Would like to add quiz part like one can give a sentence quiz with pronunciation and speak, there is grading and cheer up part.

Survey 7

Interviewee has experience learning foreign languages: English. The reason one wants to learn foreign languages is because it can connect with different cultures and know how people think. One prefers learning with others and likes to talk and practice with friends and in class can talk and practice the language every day. As a beginner, one feels learning alone with music and movies is supportive. Had hired a professor and it increased the interviewee's focus, made it easier to get attention in this subject and got chances to speak with the native speaker. Essay

writing is hard and writing every day in class is helpful. Grammy is also a hard part in English. Writing is hard because one doesn't understand the culture. Gammer and remembering new words are the most difficult things in this experience. Hope to gain a new brain.

After experiencing the idea

Great! Voice assistant interaction with users is helpful. If it can gain the memory one's conversion detail will be helpful. Would like to use it at home. This tool is not a tool but a friend and hope it can remember what we had talked about.

Survey 8

Interviewee has experience learning foreign languages: English. The reason one wants to learn foreign languages is because one wants to improve language skills to study in SU. One prefers learning with others because native speakers are more fluent with this language and native speakers are friendly when they know one is not a native speaker. However, people might not correct interviewee's grammar and this fact does not improve one's learning experience. Also, living in China, the interviewee's hometown, makes learning English difficult because it is hard to know about American culture. One feels visit foreign culture and chat with different accent background foreigners is helpful. Long sentences, speed fast and accent makes the interviewee hard to understand. Translation apps such as google translate is useful and also one read newspaper on-line or on paper is useful. Accent is the biggest challenge in learning cus one don't understand other people with assent. Voc is the most difficult thing in this experience.

After experiencing the idea

This interviewee don't get a chance to experience this idea.

Survey 9

Participants have experience of learning foreign languages and learning English and Chinese. Cus one has Chinese and Taiwanese friends and would like to talk with them. Prefer learning with other cus other can correct one's mistake. In one's hometown, one does not find a chance to use this foreign language to speak and listen. The ideal environment will be to speak with English speakers so that one can know the culture differently and think in a different way. App, hello talk, can chat with native speakers. The biggest challenge for this interviewee is the pronunciation. One feels hard to speak like a native speaker and doesn't understand what native speakers are talking about due to the pronunciation and speed. Also, one feels uncomfortable in a foreign space, this fact makes one feel unwilling to speak. Joining a relationship with male/boy will be helpful and working in the language native country will also be helpful. The most difficult thing about learning is the motivation to continue. The supportive tool one wishes is to forget the hometown language and change one's brain.

<u>After experiencing the idea</u>

VA is a good way for speakers but not for writing. It can be used to practice listening and daily use. The modified speed part is useful and also would like to put the plug-in section into the background, and it will be helpful cus users can get used to the situation and not rely on it. Hope it can help to remember the vocabulary and the voice assistant can remember/record the user's words.

Survey 10

The participant has experience of learning English as a foreign language cus need to interact in school. Prefer learning alone when writing and doing homework; prefer learning with others when talking. One feels stress when talking with others. Music as a tool to support learning and can chat with people who know the language from one's hometown. Spelling is the most challenging part and one is afraid of saying the wrong thing. And, not understanding jokes and slang brings another challenge to continue the conversation. Watching a TV show but not needing to spend a long time will be helpful for learning. The most difficult thing about learning is remembering vocabulary. The supportive tool one wishes is available is having self motivation and there is something to push oneself.

After experiencing the idea

This idea will increase my conversation skill and it is convenient for people. If it can also provide examples like how this situation should be, one conversation with a different culture such like that will be helpful. The speed part is good and it can also use different moods like easy mood with slow speed and increase the speed with different levels of mood. Also, would like to add something that can help people to remember vocabulary cus vocabulary use in conversation is difficult. Hope there is not only a learning mood but there is others mood that can push the user. The participant enjoy this idea cus never get this idea before. It is so interesting. Want other look and want to bring it in dairy life, such as shopping or in class. Would like to use it along, at home and carry it everywhere.

Survey 11

Interviewee have experience of learning English in order to pass the TOEFL test, and would like to get a job. One's language experience is learning in school and the classes are all taught in English such as meth. One feels taking class is not so helpful cus the interviewee feels the goal for the class is to pass the exam and don't get a chance to have a conversation. The ideal environment for learning is speaking in daily life, however, the conversation part is difficult cus one is afraid of speaking in wrong words and doesn't understand the meaning that others say. The website, grammarly, is supportive and also likes to use social media and read one's commend. The culture difference is the biggest challenge because sometimes using a translation machine translates the wrong meaning of the language. Going out to work in an English place and making friends with native people will be the learning approach this interviewee wants to try. The most difficult thing about learning is doing daily practice and continuing speaking.

After experiencing the idea

Siri doesn't understand one's pronunciation and spelling, if this VA can give suggestions it will be useful. The speed rate part can improve one to familiar with how native speaker talk. Hope there is a way to remember don't know words. Want to make chat with this because talk with people is important.

Survey 12

Having experience of learning English cus language is a tool for conversation with others. Prefer learning with others because others can help the interviewee to find the mistake

during conversation. Also, learning alone is slow. Reading paper and memory is not helpful for memory and sometimes one finds it stressful. Also, one fears of wasting others time because the interviewee feels herself is not a fluent speaker, and this fact makes it hard to share one's idea with others. Confidence is hard to build so that leads to fear of speaking with others in this foreign language. Taking a part time job in America will be the learning approach but one doesn't have time. The most difficult thing about learning is confidence.

After experiencing the idea

If this friend can start the conversation first and push the user to conversation with this friend it will be great cus sometimes this interviewee feels disappointed if the conversation just end very soon. The modified speed rate is important for this interviewee cus it can build confidence, increase the skill of talking and also familiarity with the link word. Hope this VA can change another way to talk about the same thing if the user doesn't understand it. For the participant, this is not a tool, it should be a friend, a partner and one likes to build a relationship with this friend. This idea can be a decoration and it can also change outside color and look like.

Survey 13

One has experience with learning French, Spanish, German, Russian and Korean. One attends class in school for the required, and is also interested in learning different knowledge. Enjoy learning along with App for reading articles and also enjoy talking with classmates. However, in class was hard to find way to use France because the class used an exchange partner way to support learning. The partner only would like to practice English which is this

interviewee's native language. In the app, there is complete game and can play it around 15-20 min to reach the goal (voc and writing). The interviewee feels that speaking in a language is important and practice with a group is supportive. Also, people can exchange language in this learning group. One read newspaper to learn French. Dairy newspapers support learning and social media such as twitter is useful. The story way of articles helps one learn.

The biggest challenge in learning is practice and speaking with someone. Interacting is useful and can be practiced. Would like to try software and also go local but they are expensive. Local places can learn the national way to use the language, also every place has different slang. Wish to have an online speaking partner such as facetime and can exchange and learn other languages.

<u>After experiencing the idea</u>

If this idea has a culture or local language it will be helpful. Like VPN, the user can choose a location to learn the local language. The context part is helpful and learning about "How's going" meaning in local language not only definition (translate) will highly support. Also if there is exchange data base like no data and build by local people will be great.

About the shape, don't want this shape like decorate but like something can wear or can bring on shirt.

Survey 14

One learns Spanish in middle school because school is required. Also, learning a foreign language is fun and one can learn about different cultures. One prefers learning with others because it can practice speaking. One had hired a tutor due to business needs and to learn italian, and also learn mandarin with her wife. One practice is speaking verbally, reading articles, and practicing writing in school. In school, they were not allowed to speak native language and needed to speak Spanish, which was taught in class. Also, one feels that facial expression is important to understand in language.

App, Wlingua is interested in learning and ot can play a record which is recorded by other people and play a video mimic as a sample. This app help to focus to the conversation and stay in a language. It is useful. The biggest challenge is understand the culture, cus culture lead to different ways to talk. It is also helpful to chat with people who know about the language. In one's experience, one learn Spanish in class and there was a girl who came from Spain and one learn a lot background with her. One would like to try is to do some practice that can immediately reply and can correct the grammar and do some reminder for this person. App might be convenience when walking around or doing work. The support tool will be visual on screen and one can type in both language.

After experiencing the idea

Love the idea, one time one talked with a student and this student used "first of all" instead of first. This student don't know "first of all" in word is not polity. One feel if it (VA)

can have tone or slang part, it will be very helpful. "How's going" part is useful because the learner might not know about the culture and don't what it means in this language. This tool can support learning in different culture, refresh the memory and bring the culture conversation right a way. This tool should be focus on no-formal way; culture, dairy way to support. One will like to use it at a nice quick room by oneself with a visual refisher, book, nearby and can ask question at the same time. EX: can you talk me how people use soccer or what situation…

The future if voice assistant will supportive and one's son like to talk with VA. This VA should be updated quickly and should have some relax or dairy mood, also it should be friendly with the outlook.

Survey 15

Have experience learning English which was used for a test. Prefer learning alone cus the school taught this language only for those exams so they taught this language in books. Teachers taught how to remember vocabulary, read articles and answer questions in exams. They also taught grammar and repeated from memory of several articles. One feels this is not so useful and not helpful for conversation. One practice speaking skill in order to pass TOEFL test and attended in cram school, there was a tutor taught how to "win" the toefl speaking test. He will correct the pronunciation and mimic how the local people talk in American. One finds that google translate is useful cus it can translate article and one can read faster. The biggest challenge is that in school the way they teach such as score and exam let this interviewee feel lost self-confident. One pointed out it is more important to talk with people to learn a language. Want there is a AI and can remember the conversation and can record those conversations.

<u>After experiencing the idea</u>

Not sure how this idea will work, as a Data Science background, would like to know how to collect data. If this VA can open a conversation, it might be helpful. This can app something like repeat with the robot so that the user can mimic the conversation and use in real life. Also, if this VA can push the user to practice, it will helpful cus one said one is a lazy person and need someone to push. One feel this idea should have more of a different scene or practice environment. Also, it should able to correct the user if the user said something wrong. Not like the SIRI right now only says "sorry, i don't understand."

Survey 16

Have experience learning English. Would like to pass the test. One feel one's English skill is not so well cus the learning way in one's experience was for exams and when one came to Syracuse feel one can't use this language to fluently converse. However, on the other hand, one feel getting score can let one feel satisfied during learning experience. One use App to remember vocabulary and there is translation nearby so that one can see the meaning at the same time. Also, in this app, there was a long sentence practice. This part is helpful cus one can become familiar with the environment. If there is a Artificial intelligence can be like a tutor and can remember users' speech, one can feel helpful and the AI can correct the user or change the habit the user speaks.

<u>After experiencing the idea</u>

This idea might be similar with Alexa and seem too similar as voice assistant and don't feel supported. If this VA can upload an article and can read the article, it will be good to support learning because sometimes one can't read some specific word and the connected speech sometimes is hard to understand and hard to use in the real world. One feel the part of accent is not so useful and doesn't need this part.

Survey 17

Interviewee had learned English, Japanese and Korean before. One learned English in school and that school hired some native english speaker to teach all courses such as math, science, technology and so on. One feel this experience is useful. One compare with learning other languages experience, such as Korean, one learned it by oneself and remember voc in book, there was no people that can conversation and practice with. One feel learning also is important cus one need to remember some textbook information, such as vocabulary, reading and so on, to pass the test; on the other side, if one can have chance to talk with other, one feel it will be helpful to practice pronunciation. Interviewee had experience taking native speaker class and feel it worth to learn English. One tried APP to remember vocabulary or watch netflix.

BTW, netflix has subtitle and it is useful. One will try to watch it English only and watch again with English subtitles, if still unsure with the meaning, one will use two subtitles, which is English and chinese to support understand the show. The biggest difference one faces is in foreign country, one can't understand the speed and connecting words during conversation. Also, some slang is hard to understand and it change from time to time, and watching YouTube might be helpful?

Hope there is an international Alexa and can be nearby always.

After experiencing the idea

Like the idea, and if there is a way to remember one's conversation and can make some suggestions afterwards, it might be helpful. One feel the speed changeable part is interesting, one can speed up and used to the language. Would like to use this tool everywhere and would like to carry it all the time. If it can has something that can analyze the record and there is a weekly contact should which part is more wrong, it will be helpful.

Survey 18

One learned English and Japanese before. Don't like the way cram schools teach these two languages. After coming to America, one feel the language learned before is not so useful cus one couldn't find chance to practice speaking. One prefer learning with other so that one can practice speaking; also, one enjoy study along so that it can more focus on remembering vocabulary. One attended cram school in order to pass TOEFL , however, this is not so helpful and one can learn how the local people use term or how one speak. One also had experience attending GMAP, GRE courses, it was not a good experience and not supported. One read news daily in order to use to the language. Hope Siri can teach English and it has more function with dairy life conversation and vocabulary practice.

After experiencing the idea

This idea is nice but it is similar with Siri and want it can be more humanized and more close to dairy life. One feel speed part might be helpful and can support pronunciation. One might prefer use this tool at home in own space and if there is a correct section that can correct one's speech it will be helpful.

Second Interview

Survey 2-1

The local story is interesting, and participants like it a lot. If there was a translation of the local story part, it would be helpful. One feel satisfied with the local story and learn about how it came from, voc learning and so on. If there is a part can change speed, it will be better. Want to see more option and story outside Syracuse.

Survey 2-2

One feels local stories are helpful because one can learn the culture in this way. The culture story, history and slang is supported. One feel this can support pronunciation. One is thinking about if the local story part the author or writer wrote some wrong information, if other can correct it. Everyone can change the local story.

Survey 2-3

Participants feel the culture of local history is important and one also likes the slang part. Learning in a non-formal way is interesting and this idea is a good way to learn language from a cultural aspect. Pronunciation with different words is helpful.

Survey 2-4

Interviewee feels that the local story is interesting and defining the meaning of the slang part is also useful. Like the way one uses in real life and it can be practiced. Like the way the voice assistant said it and if there is a pic can show up nearby, it will be helpful. If the user ask VA about some word, and it can show some images, the user can know more what it mean.

Survey 2-5

One feels this is a good idea and feels it is useful for beginners and the slang part might more for middle learners. As a middle English learner, feeling this can be used during class when listen non-native speaker speak English. One feel hard to understand different accent or pronunciation said by non-native speakers of speaking English. This interviewee also point out the squeeze ball is too big. One note that doesn't use translation cus it is not good for learners and one also wants this app information to connect to the laptop so that one can recheck when working on the laptop. Also, providing an advanced level is needed.

Survey 2-6

Interviewee mentioned the record part is good when one feels it is difficult to understand the meaning. The record time slot can be changeable so that one can record longer or shorter. If one has a class with a different background, one can check around with this APP. The participant want this app to record 24 hrs and this tool can look like a cell phone charm so that one can use it easily. This app support users' memory and should concern about if the memory are full in this tool.

Survey 2-7

This tool needs to think about someone uncomfortable for recording. The size of the prototype is too big. If it can support lecture reading, it will be helpful. Should show something that others can know about what I am recording. This tool support voc practice and if there is a practice for speaking voc, it will be great.

Survey 2-8

One feels this design is needed. With this APP, people feel stressless when they don't understand communication. People revisit when going home and feel relaxed to learn about this new language. Also, speakers with this new language can get more ideas and background knowledge, and one will not be impolite to the other during chatting. Talking to a voice assistant is also not stressful and you can make mistakes without thinking about others' judgment.

Appendix 4

First Prototype Video Image.

So, how's your vacation?
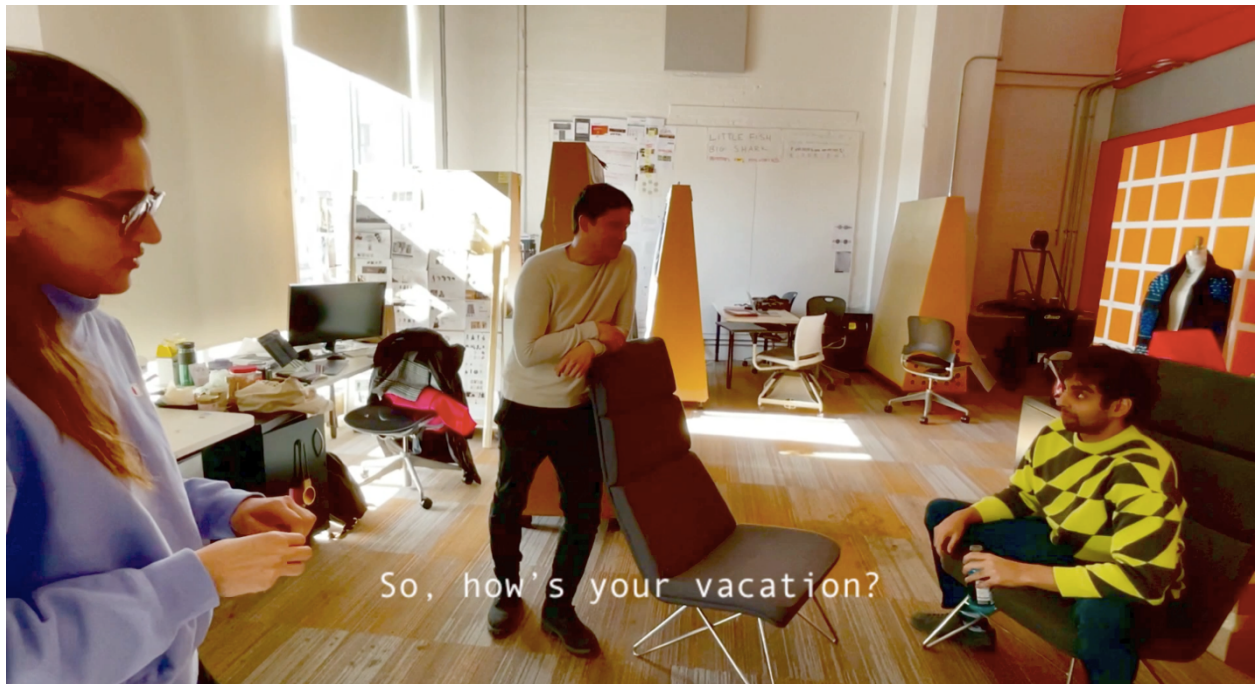
If I can eat and take care of my body

the way I did a week before vocation

WHAT HAPPEN
WHAT HAPPEN
WHAT HAPPEN
WHAT HE TALKING ANOUT
WHAT IS THAT MEANING
??????GUCCI??????

Hi Annie, What can I help you with today?

Hi K, what's the meaning of Gucci?

Gucci is a slang term that is used as an

exclamation of approval or

as an expression of good will

It can also be used as an adjective

to describe something that

is going well or is in good condition

Gucci is derived from the luxury Italian fashion

brand Gucci and is used to indicate that something

is good, going well, or is in good taste

It is often used as a slang alternative

to the words good or great

Would you like to do some practice?

Yes

Appendix 5

FAQ Session

Practice Session

Annie can also practice this pattern in a conversation.
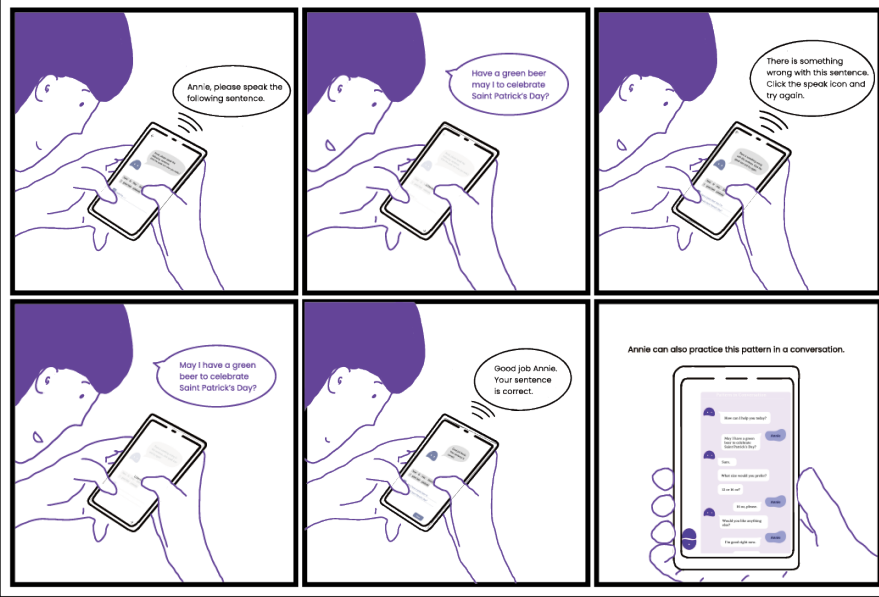
Sharing of Personal Stories

On the other side, Barbara, who lives in Syracuse, also uses Kinakaian to learn a new language. She share her own story from local to support other learners.

The database are created by local community and learners can learn diverse stories locally.

Appendix 6

Video Link -

https://youtu.be/yEOVLEcDq5Q

Figma Link-

https://www.figma.com/file/R7oaGHrYhMY2XkS12PCWBK/Kinakaian?node-

id=0%3A1&t=zDzvig1koUwYUtBs-1
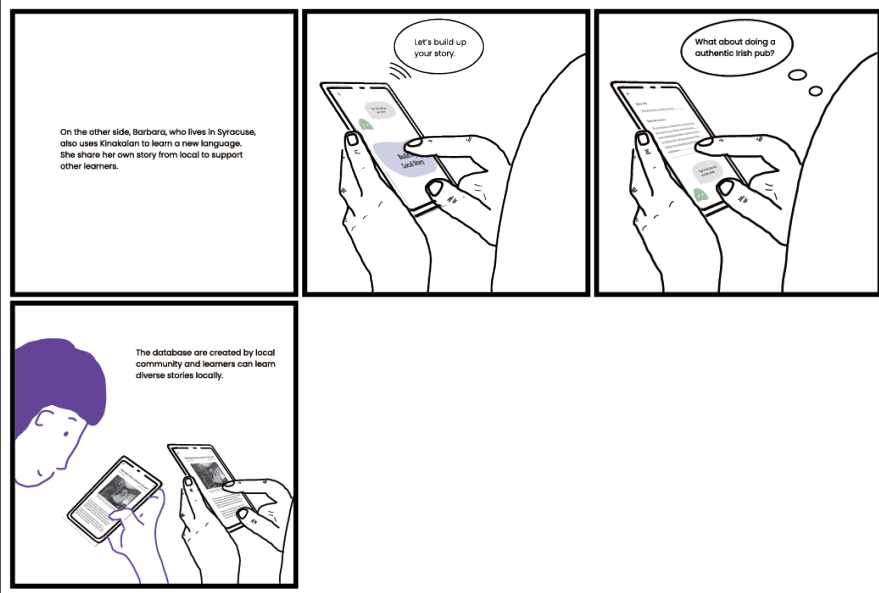
Reference

Al Shamsi, Jawaher Hamad, Mostafa Al-Emran, and Khaled Shaalan. 2022. "Understanding Key Drivers Affecting Students' Use of Artificial Intelligence-Based Voice Assistants." *Education and Information Technologies* 27 (6): 8071–91. https://doi.org/10.1007/s10639-022-10947-3.

Alamer, Abdullah, and Ahmed Al Khateeb. 2023. "Effects of Using the WhatsApp Application on Language Learners Motivation: A Controlled Investigation Using Structural Equation Modelling." *Computer Assisted Language Learning* 36 (1–2): 149–75. https://doi.org/10.1080/09588221.2021.1903042.

Al-Hoorie, Ali H., and Phil Hiver. 2020. "The Fundamental Difference Hypothesis: Expanding the Conversation in Language Learning Motivation." *SAGE Open* 10 (3): 2158244020945702. https://doi.org/10.1177/2158244020945702.

Al-Kaisi, Alisa N., Alla L. Arkhangelskaya, and Olga I. Rudenko-Morgun. 2021. "The Didactic Potential of the Voice Assistant 'Alice' for Students of a Foreign Language at a University." *Education and Information Technologies* 26 (1): 715–32. https://doi.org/10.1007/s10639-020-10277-2.

Alm, Antonie, and Larian M. Nkomo. 2022. "Chatbot Experiences of Informal Language Learners: A Sentiment Analysis." Chapter. Research Anthology on Implementing Sentiment Analysis Across Multiple Disciplines. IGI Global. 2022. https://doi.org/10.4018/978-1-6684-6303-1.ch050.

Álvarez Valencia, José Aldemar, and Kristen Michelson. 2022. "A Design Perspective on

Intercultural Communication in Second/Foreign Language Education." *Journal of International

and Intercultural Communication*, May, 1–20. https://doi.org/10.1080/17513057.2022.2066152.

Arlinghaus, Clarissa Sabrina, and Frank Ollermann. 2022. "Hey, Siri®! Ok, Google®! Does

Talking to Voice Assistants Enhance Emotional Aspects of Mobile Phone User Experience?" In

Mensch Und Computer 2022, 382–88. Darmstadt Germany: ACM.

https://doi.org/10.1145/3543758.3547544.

Brandtzæg, Petter Bae, Asbjørn Følstad, and Jan Heim. 2018. "Enjoyment: Lessons from

Karasek." In *Funology 2*, edited by Mark Blythe and Andrew Monk, 331–41. Human–Computer

Interaction Series. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-

68213-6_21.

Costanza-Chock, Sasha. 2020. "Introduction: #TravelingWhileTrans, Design Justice, and Escape

from the Matrix of Domination." *Design Justice*, March.

https://designjustice.mitpress.mit.edu/pub/ap8rgw5e/release/1.

Davis, Fred D. 1989. "Perceived Usefulness, Perceived Ease of Use, and User Acceptance of

Information Technology." *MIS Quarterly* 13 (3): 319–40. https://doi.org/10.2307/249008.

Doorn, Jenny van, Martin Mende, Stephanie M. Noble, John Hulland, Amy L. Ostrom, Dhruv

Grewal, and J. Andrew Petersen. 2017. "Domo Arigato Mr. Roboto: Emergence of Automated

Social Presence in Organizational Frontlines and Customers' Service Experiences." *Journal of Service Research* 20 (1): 43–58. https://doi.org/10.1177/1094670516679272.

Dörnyei, Zoltán, Kata Csizér, and Nóra Németh. 2006. *Motivation, Language Attitudes and Globalisation: A Hungarian Perspective*. Multilingual Matters.

Eom, Minhee, and Jean Braithwaite. 2023. "Motivation to Learn Korean as L2 or L3 in a Bilingual, Bicultural Community." *International Journal of Multilingualism*, February, 1–18. https://doi.org/10.1080/14790718.2023.2170381.

Goda, Yoshiko, Masanori Yamada, Hideya Matsukawa, Kojiro Hata, and Seisuke Yasunami. 2014. "Conversation with a Chatbot before an Online EFL Group Discussion and the Effects on Critical Thinking." *The Journal of Information and Systems in Education* 13 (1): 1–7. https://doi.org/10.12937/ejsise.13.1.

Golonka, Ewa M., Anita R. Bowles, Victor M. Frank, Dorna L. Richardson, and Suzanne Freynik. 2014. "Technologies for Foreign Language Learning: A Review of Technology Types and Their Effectiveness." *Computer Assisted Language Learning* 27 (1): 70–105. https://doi.org/10.1080/09588221.2012.700315.

Gonz, Marta. 2015. "Conversation Analysis in Computer-Assisted Language Learning." *CALICO Journal* 32 (3): 569–94. https://doi.org/10.1558/cj.v32i3.27568.

Habibi, Ehsanollah, Meghdad Kazemi, Habibollah Dehghan, Behzad Mahaki, and Akbar Hassanzadeh. 2013. "Hand Grip and Pinch Strength: Effects of Workload, Hand Dominance,

Age, and Body Mass Index." *Pakistan Journal of Medical Sciences* 29 (1(Suppl)). https://doi.org/10.12669/pjms.291(Suppl).3535.

Hu, Peng, Yeming Gong, Yaobin Lu, and Amy Wenxuan Ding. 2022. "Speaking vs. Listening? Balance Conversation Attributes of Voice Assistants for Better Voice Marketing." *International Journal of Research in Marketing*, April, S016781162200026X. https://doi.org/10.1016/j.ijresmar.2022.04.006.

Huang, Weijiao, Khe Foon Hew, and Luke K. Fryer. 2022. "Chatbots for Language Learning— Are They Really Useful? A Systematic Review of Chatbot-Supported Language Learning." *Journal of Computer Assisted Learning* 38 (1): 237–57. https://doi.org/10.1111/jcal.12610.

Kasar, Kadriye Sayin, Saadet Erzincanli, and Nesat Tolga Akbas. 2020. "The Effect of a Stress Ball on Stress, Vital Signs and Patient Comfort in Hemodialysis Patients: A Randomized Controlled Trial." *Complementary Therapies in Clinical Practice* 41 (November): 101243. https://doi.org/10.1016/j.ctcp.2020.101243.

Khan, Shahzia, and Mubeena Iqbal. 2020. "AI-Powered Customer Service: Does It Optimize Customer Experience?" In *2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)*, 590–94. Noida, India: IEEE. https://doi.org/10.1109/ICRITO48877.2020.9198004.

Kono, Michinari, and Koichi Araake. 2022. "Is It Fun?: Understanding Enjoyment in Non-Game HCI Research." arXiv. https://doi.org/10.48550/arXiv.2209.02308.

Le Pailleur, Félix, Bo Huang, Pierre-Majorique Léger, and Sylvain Sénécal. 2020. "A New Approach to Measure User Experience with Voice-Controlled Intelligent Assistants: A Pilot Study." In *Human-Computer Interaction. Multimodal and Natural Interaction*, edited by Masaaki Kurosu, 197–208. Lecture Notes in Computer Science. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-49062-1_13.

Li, Chengchen. 2022. "Foreign Language Learning Boredom and Enjoyment: The Effects of Learner Variables and Teacher Variables." *Language Teaching Research*, May, 136216882210903. https://doi.org/10.1177/13621688221090324.

Machado, Sheron, Emília Duarte, Júlia Teles, Lara Reis, and Francisco Rebelo. 2012. "Selection of a Voice for a Speech Signal for Personalized Warnings: The Effect of Speaker's Gender and Voice Pitch." *Work* 41: 3592–98. https://doi.org/10.3233/WOR-2012-0670-3592.

Malodia, Suresh, Nazrul Islam, Puneet Kaur, and Amandeep Dhir. 2022. "Why Do People Use Artificial Intelligence (AI)-Enabled Voice Assistants?" *IEEE Transactions on Engineering Management*, 1–15. https://doi.org/10.1109/TEM.2021.3117884.

McLean, Graeme, and Kofi Osei-Frimpong. 2019. "Hey Alexa … Examine the Variables Influencing the Use of Artificial Intelligent In-Home Voice Assistants." *Computers in Human Behavior* 99 (October): 28–37. https://doi.org/10.1016/j.chb.2019.05.009.

McLean, Graeme, Kofi Osei-Frimpong, and Jennifer Barhorst. 2021. "Alexa, Do Voice Assistants Influence Consumer Brand Engagement? – Examining the Role of AI Powered Voice

Assistants in Influencing Consumer Brand Engagement." *Journal of Business Research* 124
(January): 312–28. https://doi.org/10.1016/j.jbusres.2020.11.045.

Mohamed, Hassan A. 2022. "Using Mobile Applications to Enhance ESL Learning." M.A., Ann
Arbor, United States.
https://www.proquest.com/docview/2742629452/abstract/63E2833FBF1A4E24PQ/1.

Nicolaidou, Iolie, Petros Pissas, and Dimitrios Boglou. 2023. "Comparing Immersive Virtual
Reality to Mobile Applications in Foreign Language Learning in Higher Education: A Quasi-
Experiment." *Interactive Learning Environments* 31 (4): 2001–15.
https://doi.org/10.1080/10494820.2020.1870504.

Olwi, Alyaa. 2022. "The Use of Foreign Languages in Saudi Television Advertisements." *British
Journal of Middle Eastern Studies*, October, 1–13.
https://doi.org/10.1080/13530194.2022.2132215.

Poushneh, Atieh. 2021. "Humanizing Voice Assistant: The Impact of Voice Assistant
Personality on Consumers' Attitudes and Behaviors." *Journal of Retailing and Consumer
Services* 58 (January): 102283. https://doi.org/10.1016/j.jretconser.2020.102283.

Razak, Norizan Abdul, Amr Abdullatif Yassin, and Marwan Saeed Saif Moqbel. 2019.
"Investigating Foreign Language Reading Anxiety Among Yemeni International Students in
Malaysian Universities." *International Journal of English Linguistics* 9 (4): 83.
https://doi.org/10.5539/ijel.v9n4p83.

Tsiplakides, Iakovos, and Areti Keramida. 2009. "Helping Students Overcome Foreign Language Speaking Anxiety in the English Classroom: Theoretical Issues and Practical Recommendations." *International Education Studies* 2 (4): p39. https://doi.org/10.5539/ies.v2n4p39.

Urakami, Jacqueline, Sujitra Sutthithatip, and Billie Akwa Moore. 2020. "The Effect of Naturalness of Voice and Empathic Responses on Enjoyment, Attitudes and Motivation for Interacting with a Voice User Interface." In *Human-Computer Interaction. Multimodal and Natural Interaction*, edited by Masaaki Kurosu, 244–59. Lecture Notes in Computer Science. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-49062-1_17.

Yang, Heetae, and Hwansoo Lee. 2019. "Understanding User Behavior of Virtual Personal Assistant Devices." *Information Systems and E-Business Management* 17 (1): 65–87. https://doi.org/10.1007/s10257-018-0375-1.

Zhang, Lubei, and Linda Tsung. 2021. "Learning Chinese as a Second Language in China: Positive Emotions and Enjoyment." *System* 96 (February): 102410. https://doi.org/10.1016/j.system.2020.102410.

VITA

Area Experience

Engaging in design areas for over seven years; cooperating with multi-experts in problem-solving and design research. Expert in User Experience (UX) and User Interface (UI) especially for artificial intelligence products and industrial design. Proficient in innovation ideas based on user research and user-centered design. Good at communicating with team members both in person and online. Providing leadership to attract and implement creative rewards and recognition in group projects. Having a passion for develop prototype ,and focusing on design data-driven, human-centric experiences.

Technical Skills

Adobe Suite (Illustrator, Photoshop, Dreamweaver, InDesign, Fresco), 3D Software (SolidWorks, Rhino, AutoCAD, PRO-E(Creo), SketchUp), Figma, WordPress, Wix, Miro, Industrial Sketch, HCI, UX, UI, Python, Microsoft Word, PowerPoint, Excel, Windows, Mac OS

Education

August 2021 – August 2023

Master of Design (MFA), Syracuse University, NY, United States.

March 2020- July 2021

English Language Institute (ELI), Syracuse University, NY, United States

March 2020- June 2021

Master of Human Resources and Knowledge Management (MBA), National Kaohsiung Normal University, Kaohsiung, Taiwan

August 2015 - June 2018

BDes in Industrial Design, National Kaohsiung Normal University, Kaohsiung, Taiwan

August 2014- July 2015

Wood Based Material and Design, National Chiayi University, Chiayi, Taiwan

Work Experience

August 2021- June 2022

Teaching Assistant, Syracuse University, New York, United States

• Familiar with knowledge of web design, media prints & hand in machine shop

• Support 6 design projects and achieved 40% teams' design awards in a semester

• Perform work involving new concepts and ideas for achieving missions

• Collaborate with engineer expert and develop mock-up solutions

March 2020 - June 2021

Junior Graphic Designer and Editor, National Kaohsiung Normal University, Taiwan

• Publish the 100 NKNU Magazine

• Sustain local market regional planning via pitching virtual-friendly maps

<u>August 2016- September 2017</u>

Assistant Engineer , FULL WAY TECHNIQUE CO., LTD, Taiwan

• Responsible for AutoCAD engineering graphics

• Worked with group and achieve Chi Mei Corporation Building

• Build and develop prototypes by using 3D printing

Publication & Awards

• HCI Research Article Publish at HCI INTERNATIONAL 2023, Denmark

• 3rd Prize at Intelligence ++ Competition 2022, Syracuse University

• Best Creative Award at McDonald's Design Competition 2019, Taiwan

• Exhibited in Young Designers' Exhibition 2018, Taiwan