

Interactive AI for Linguistic Education Built on VR Environment Using User Generated Contents

WooHyun Park*, DoJin Park¹, ByungJune Ahn¹, SeokHun Kang³,

HaengYeong Kim⁴, RaeHyeon Kim⁵, JaeHeum Na⁶

*Department of Electrical and Computer Engineering,

¹ College of Information & Communication Engineering,

³ College of Biotechnology and Bioengineering, ⁴ College of Engineering,

⁵ College of Natural Science, Sungkyunkwan University, Suwon, Korea

⁶ College of Computer Science, Kangwon National University, Gangwon, Korea

cgvtt@skku.edu

Abstract— Currently, the market for Korean language education is over 4 trillion won, and 20 trillion yen in Japan and 50 trillion in China.

Traditional markets used to learn languages through native speaker, such as offline academies, burns, English, etc. In addition, in the case of English, it has become just a subject for studying for tests, not only to communicate.

This causes pressure on young people, resulting in poor learning efficiency. To learn language as a means of communication, the environment needs to be created in order to acquire language.

There are the language learning programs in the virtual reality environment. such as 'Learn Languages VR by Mondly' and 'House of Language VR'. In the case of Learn Languages program, it supports many environment and teaches many different languages. However, depending on the stated scenario, the degree of freedom of speech is low.

In the case of 'House of Language VR' program, the degree of freedom of speech is significantly reduced by matching words to words. Also, it is less realistic because it is in progress through animation programs produced for infants.

In learning languages, learning to speaking in a real-life situation with native speakers is required for advanced skill.

In this study, we intend to produce a learning program using VR devices and interacting with native speakers.

Keywords— Virtual Reality, Game, AI, Voice Recognition, Education

I. INTRODUCTION

Existing Korean language education has been focused on exams. After 2009, conversation-oriented language learning systems such as 'Seong won School' and 'Yanaudu' were launched. However, the lesson was also conducted through Internet lecture or video lecture. In learning the language, you need to communicate with native speakers in real situations to grow your skills. In this study, we want to make a learning program using realistic feeling of VR device and mutual feedback with native speakers.

We plan to make a profit in two ways.

The first is the in-game coin system. Coins are required for users to use new content in addition to the built-in content, which can only be obtained through cash settlement or native speaker mode recording. We will make profits by selling coins at an affordable price.

The other is voice data sales. Current voice data collection is done in part-time format, which causes the company to pay a lot of labor costs, and its speed is also low. We will use the program's Recorder mode to solve this problem, and we will make a profit by producing cheaper and massive voice databases.

II. RELATED WORK

A. Unity

Unity is a game production engine that makes cinematic content and gameplay sequences extremely easy and straightforward.

There are many reasons to use Unity among several game production engines, the first of which is that Unity is free if you do not use it for commercial purposes. In fact, if you look at the Unity web page, you can see that the Personal version is available for free if the revenue does not exceed \$ 100,000.

The second reason is the existence of 'Asset Store'.

The Asset Store is Unity's online platform where other Unity users create commercial models and sell already modeled assets at a very low cost. While the team can not get the model they want, it is very useful to create a demo version of the Asset Store's inexpensive models when there are no team members with modeling skills and no financial resources to outsource modeling.

The third reason is that the project made by Unity is very convenient to change the platform.

You can easily build a VR platform based on your configuration as well as change your PC project to Android or iOS platform. Through the advantages of Unity, we were able to carry out the project more efficiently.

III. APPROACH

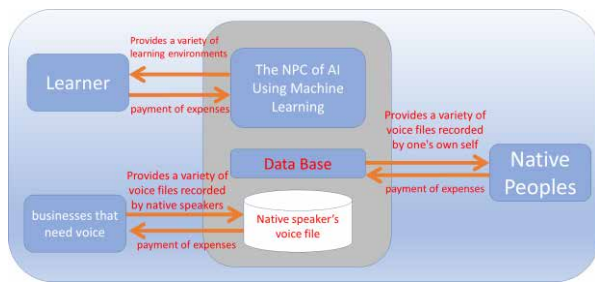


Figure 1. The structure of System

Existing VR learning programs are stored in the program database. It was possible to learn with only information.

But the program we proposed will be censored by the user and used as learning data. So it enables users to be more diverse and in a more realistic environment.

It can be used significantly in education areas of research.

A. Voice RSS

The key of this study is 'dialogue'. It is said that the basic conversation is composed through api.ai, but there is a restriction that the text is finally exchanged. So we decided to change the dialogue from api.ai to text form to audio form to realize more realistic dialogue.

Text to Speech (TTS) is a technology that converts text into audio form. It allows the site called 'Voice RSS' to implement this technology easily. This site supports API using HTTP.

If you request by including text that you want TTS in URL, you can simply receive file in audio format. You can also adjust the language, speed, codec, length, etc through the parameters of the NVP format. Above all, 500 sentences of the day are supported free of charge.

The figure below shows a diagram of the TTS system[6].

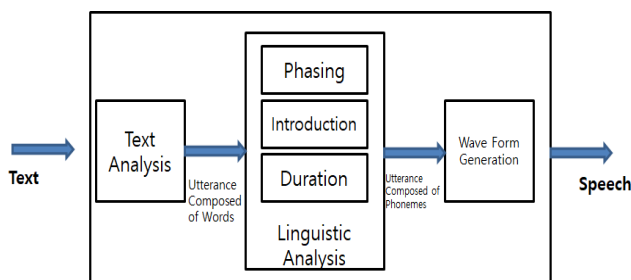


Figure 2. TTS System

B. Api.ai

The default dialog script was created using api.ai.

Api.ai is a free chat bot program provided by Google. It is easy to integrate with various services of Google and has an

intuitive UI. (It has been renamed Dialogue Flow since November 2017, but it is described as api.ai in the text).

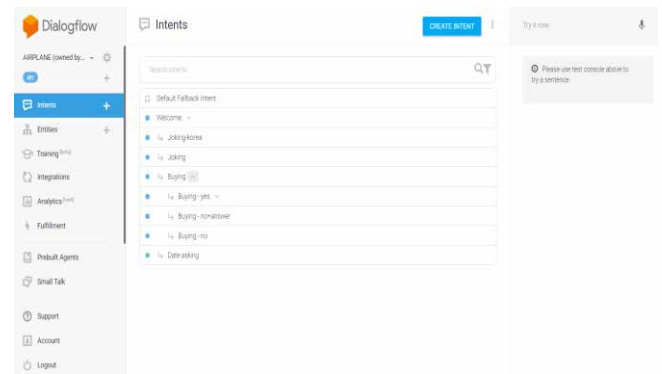


Figure 3. Api.ai – Main Screen

The left side shows the menus, the center is the function of the selected menu, and the right side is the space to test the dialogue that has been created. The most used menus in the study are Intents and Entities.

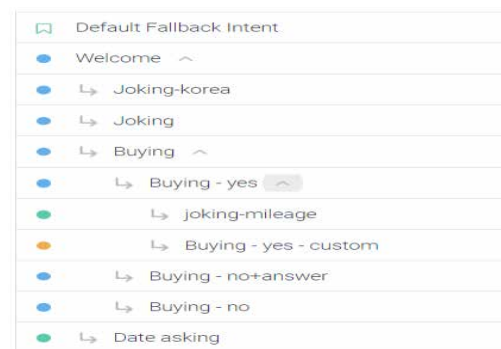


Figure 4. Api.ai – Intents In Hierarchy Structure

Intents are menus that make up the structure of the conversation itself

Figure 4 is part of the dialogue script created in the study.

The "Welcome" and "Buying" tabs in the figure 4 consist of the user's text and the computer's answer. Therefore, when the user's sentences are divided into several options, they are arranged horizontally in the Intents, and the sentences that can be used only after a certain sentence is arranged vertically as sub-tabs of the preceding sentence. However, if you select any of the tabs horizontally arranged to continue the conversation, it is possible to move back to the option in the next sentence. (Ex: Joking - korea -> Joking available).

Additionally, the colored circle on the left indicates the frequency of the conversation. Thus, if two horizontally arranged tabs can receive the same user's sentence, setting a higher frequency on one side increases the likelihood that the dialogue will progress to that tab. The frequency can be set in five ways.

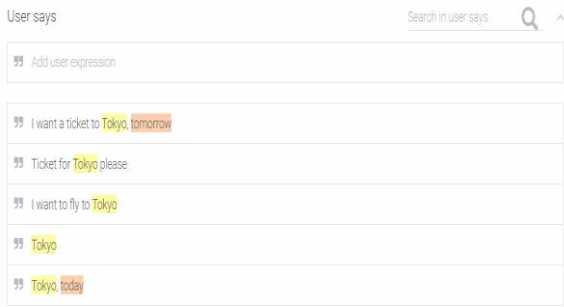


Figure 5. Api.ai – user says

First, you can determine the user's sentence necessary to proceed with the conversation on that tab. api.ai uses the built-in automated expansion to take the same type of sentence as the same.



Figure 6. Api.ai – text response

Next, you can set the computer's answer according to the user's sentence. If you set multiple answers, a single sentence is printed.

REQUIRED	PARAMETER NAME	ENTITY	VALUE	IS LIST	PROMPTS
☒	destination	@destination	\$destination	☐	Where is your d...
☒	date	@sys.date	\$date	☐	When do you wan...
☒	time	@time	\$time	☐	Which flight is ...
☒	class	@class	\$class	☐	Which class do ...
☐	Enter name	Enter entity	Enter value	☐	—

Figure 7. Api.ai – parameter

There is also a function called parameter for user's convenience. A parameter is a kind of variable used in a conversation, and can be set to include in conversation in progress. If there is no corresponding element in the user's sentence, the dialogue is again conducted through the set question, and when it is obtained, the corresponding data is stored and the computer's answer is outputted. A parameter is used in this way to extract information from a user's sentence and continue conversation. Below is an example used in this study.

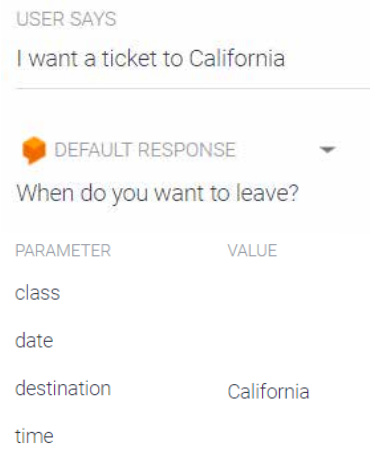


Figure 8. Api.ai – default response, parameter value

In Figure 8, we have specified four essential parameters: Class (date of flight), date, destination, time. If the user does not include the word "destination" in the sentence, the word "Where is your destination?" And the word "date" are not included. In the section that sets up the user's sentence (see Figure 8), orange on Tokyo and red on today and tomorrow indicate that each is recognized as a parameter of destination and date.

Entities are menus that manage the parameters used previously. The configuration is as follows.

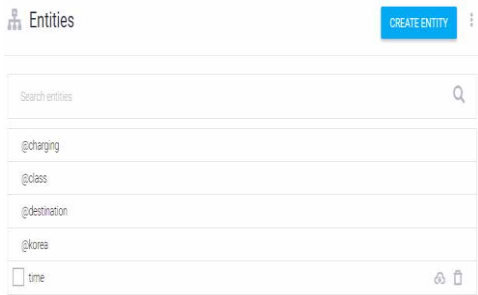


Figure 9. Entities1

As shown in Figure 9, specify the variable name of the overall parameter and click on the created tab to input additional information.

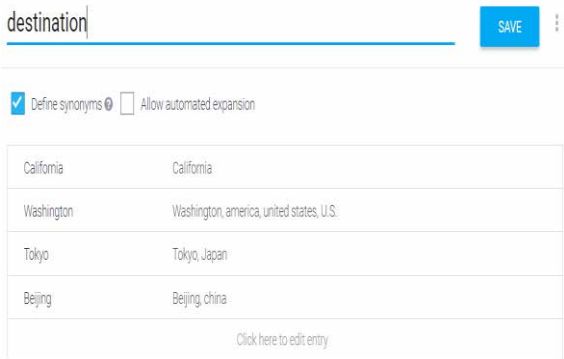


Figure 10. Entities2

Figure 10 shows the screen after selecting one tab. In order to purchase a ticket as a parameter of destination required during the dialogue progress of this study, the destination is limited to California, Washington, Tokyo, and Beijing.

Allowing Define Synonyms allows api.ai to recognize multiple words as the same and proceed with the conversation.

In Figure 10, Washington, America, united states, U.S are all recognized as Washington.

Automated expansion allows users to recognize words in the api.ai itself as having the same properties as the same parameter without having to set them up separately. In Figure 10, it is limited to four destinations, but if selected, it can be used as an element of the override attempt parameter of Tokyo, New York etc.

The resulting conversation script is stored in JSON format and can be extracted and used separately. In this study, we created a dialogue script using api.ai in the same way and extracted the JSON file and used it to produce the program.

We applied api.ai Unity SDK to integrate pi.ai into Unity development environment.

The api.ai Unity SDK contains some basic elements for using the chatbot module.

1. Google cloud voice recognition class applicable to Android environment.
2. Built-in speech recognition class applicable to all platforms including PC.
3. A transmission class that sends the results in api.ai location in the form of webhook.
4. Classes that receive result values processed by api.ai in json format.

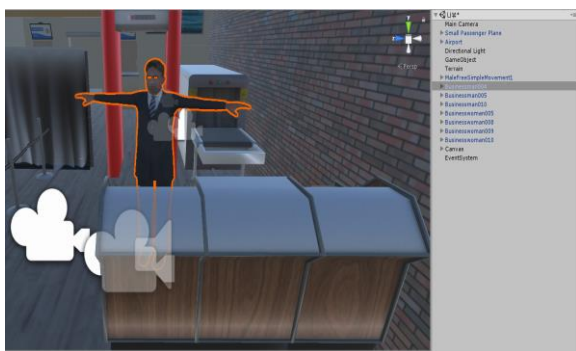


Figure 11. Character Model – Businessman004

First, a module that processes the received json format data as a character string and then output from the unity graphic character object. In order to use Voice RSS and api.ai described above in Unity, I wrote a simple script and put it in a businessman004 who is responsible for the airport checkpoint. This script converts the text answer received via api.ai into an audio file using the Voice RSS API and then uses the audio source component of the scripted object to speak. On the other hand, we configured a simple UI to communicate the words using api.ai. I created a text box and a button in the upper right and connected them with businessman004. By doing this, pressing the button will send

the contents of the text box to api.ai and the answer will be completed by businessman004.

IV. RESULT

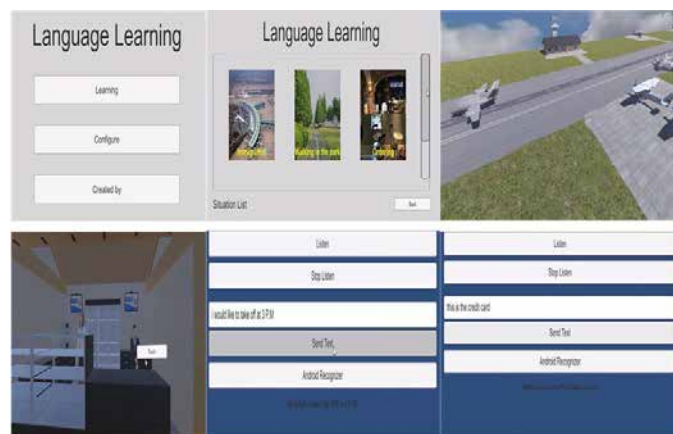


Figure 12. The Sequence of Screen Results

As a result of the study, we have developed an environment that can freely communicate with chat-bot in a foreign language in VR graphic environment implemented by Unity.

Because The Unity Engine cost no more than \$ 300 to purchase graphical objects in the asset store, which is a free market form, and did not cost any initial costs to use the various resources of the api.ai SDK and Google Cloud platform.

In our research, that we was disappointed in completing the research was that the user did not have the ability to enter new dialogue. Most chatbot api including api.ai, do not release the database to view internal data. Instead, we want to input the dialogue in the development environment implemented in the web page. Therefore, there is a method of transferring the dialogue entered by the user with the executable code in the web page and transmitting it as a webhook type feed.

Figure 12 shows the user's final learning process by running the app. Based on text and voice recording, one can see a scene where one is booking tickets with an artificial robot at the airport. It is expected to help greatly in learning English education using our system.

V. CONCLUSIONS

As a result, this study has created a value that is much more free when compared with the existing VR foreign language learning products, so that it can experience practical foreign language conversation.

Therefore, the results of this study have achieved the characteristic of foreign language learning environment that allows free conversation. This result has a new market value even when there is no new input function by the user. For further development, first, we developed a variety of chatbot agents with the possibility of expanding the scope of dialogue by the ambassador designer, and then applying more gentle graphics to satisfy the user's cognitive requirements.

It will be improved to be sellable.

REFERENCES

- [1] <https://www.oculus.com/experiences/gear-vr/1272636489423125/>
- [2] <http://fox3d.com/vr>
- [3] <https://www.superdataresearch.com/market-data/virtual-reality-industry-report>
- [4] <https://www.technavio.com/report/global-education-technology-online-language-learning-market>
- [5] <http://lugenzhe.blog.me/220906205447>
- [6] https://en.wikipedia.org/wiki/Speech_synthesis



RaeHyeon Kim is now a undergraduate Student at Sungkyunkwan University in South Korea.



WooHyun Park is now a Doctor Student at Sungkyunkwan University in South Korea.
His research interests lie in artificial intelligence.



DoJin Park is now a undergraduate Student at Sungkyunkwan University in South Korea.



HaengYeong Kim is now a undergraduate Student at Sungkyunkwan University in South Korea.



JaeHeum Na is now a undergraduate Student at Kangwon National University in South Korea.



SeokHun Kang is now a undergraduate Student at Sungkyunkwan University in South Korea.



ByungJune Ahn is now a undergraduate Student at Sungkyunkwan University in South Korea.