TALKING KRISHNA – IMPLEMENTATION OF DATABASE RESPONSIVE SPEECH BUILDER

[1] Janin Islam Khusnul Ainun Nur, [2] Hestiasari Rante, dan [3] Budi Berlinton Sitorus
[1] Department of Creative Multimedia Technology
Politeknik Elektronika Negeri Surabaya
[2,] Department of Creative Multimedia Technology
Politeknik Elektronika Negeri Surabaya
[2,]Informatics Department, IBI Kwik Kian Gie, Indonesia

e-mail: [1] janinislamu@gmail.com , [2] hestiasari@pens.ac.id , [3] budi.sitorus@kwikkiangie.ac.id,

ABSTRACT

The Z and Alpha generations are the children who are born in the millennial age, where they cannot avoid live without gadgets connected to the Internet. Being raised in this era of technological advancement leads the children in this generation to have a broad connection without a limit distance. However, language may become a barrier. Therefore, the need to be capable in the universal language becomes higher. This paper presents the development of an interactive English-language educational game using the voice recognition technology in Google Assistant. The game is taking Krishna as the main character to attract children. The game will give the children the opportunity to dialogue with Krishna. It is hoped that the application can train and improve the ability of children to speak English well and correctly. In addition, the introduction of Krishna character can help to deliver the moral values that are presented in Krishna's life story.

Keywords—Google Assistant, Krishna, learning English, story, voice recognition

1. INTRODUCTION

In today's sophisticated era, education plays the most important role in the development of a nation. Education is considered as a community capital to build a foundation concerning the development and growth of this nation. Nowadays, the competition to open "international" classes with various names has been started from kindergarten to higher education program. From the level of kindergarten to high school or the equivalent, there are those who use the name immersion class, bilingual class, and International Standard School Stubs. Apart from these diverse names, the program was created with the aim that later school graduates in Indonesia can compete globally. The government also supports the existence of this international standard educational movement by issuing the regulation of National Education System Law No. 20 of 2003, Article 50, Paragraph 3 (Suharso,2014).

The knowledge, understanding and experience of education in elementary schools always need to be refreshed and enriched, given the knowledge and concepts of education in elementary school are the result of dynamic human thinking, changing because of the influence of the situation and living conditions of humanity in general. The concept of education is always changing along with the demands of the times and civilization of humanity in the world in various aspects of life (Taufiq, 2014).

One of program that always being used to achieve international competition is the use of English as the language of instruction in schools. Even the kindergarten and elementary school students now have been taught English intensively to improve the level of education in Indonesia. Teaching early childhood education has many benefits, among which children can

learn more efficiently and gain more knowledge through activities that can be more interesting to learn, through arts, musics, sports, and any other interactive activities, then the content given early is easier to be embedded than when they are adults.

Therefore, in order to motivate Indonesian society about competing globally without having to be limited by language barriers, an interactive English-based application is created. The application is a game that utilizes voice recognition technology as an implementation of the practice of speaking English. It can train the ability of the user to practice speaking English well and correctly. In addition, the introduction of Krishna's character of this application is also a means to help convey some moral values that have been lost due to current technological developments. Krishna character is a god who is respected by Hindus and becomes a model for many people. That way, it is expected that children in this generation can understand moral values and can apply them in their daily lives, also can keep abreast of technological developments that exist with English language training that has been trained early.

II. RELATED WORKS

A. Talking Tom and Talking Angela

Talking Tom is a game for children, which presents the figure of a male cat that become the user's virtual pet. Talking Tom offers many exciting features; one of them is ability to imitate the voice of the user. When the user says a word or a sentence or even if user only screaming, Tom repeats it. This feature uses voice recognition, where the user says something then the application will capture the sound.

Talking Angela is similar to Talking Tom, except that Talking Angela is presenting a female cat character.

Idea of building Talking Krishna is based on those apps , but it should not only repeating or imitating the user’s voice, but indeed it must be able to dialogue with the user. When the user asks a question, the character will answer it.

B. SimSimi

SimSimi is a friendly yellow cute pullet, that you can chat with. This application is for online chatting with a computer or commonly called Chatbot/Messenger bot. For every word typed, it will search database of responses and spits out a random one. The way the app replies to the messages you send is dependent on the words the bot has associated with the text of what you just sent. For example, greeting SimSimi with a “hello” is likely to yield a “hello” back, because it has recognized this as a greeting that would usually yield the same response.

C. Krishna Tales

Referring to the title, this application will share about Krishna's Story and the story of this application is based on a book written by Subhojit Sanyal, named Krishna Tales. The book shows a few stories about Krishna in his childhood. However, in this application, the story that used for the storyline of the Talking Krishna is Krishna and Kaliya, Krishna and the Cow of Gokul and The Rising of Govardhan Mountain. Conversation in this application discusses the three stories that have been taken.

D. Google Assistant

Google Assistant is a software based on voice recognition from Google. It is made as a virtual assistant to make it easier for people to meet their daily needs.

Voice recognition itself is divided into two types, namely speech recognition and speaker recognition. Speaker recognition is an identity recognition process that is recognized through sound media, by recognizing the depth of sound, the speaker intonation and others. While speech recognition is the introduction of every word spoken by the speaker to a PC without the need to recognize the valid identity of the speaker. Speech recognizers aim to extract the lexical information from the speech signal independently

of the speaker by reducing the inter-speaker variability.

Unlike the speaker, recognition is the recognition of identity claimed by someone from his voice, a particular series can be voice intonation, sound depth, et cetera. Speech recognition is a process that is carried out the computer to recognize words spoken by someone regardless of the identity of the person concerned. Speech recognition implementations such as voice commands to run a computer application (Meillisa, 2008).

Speaker recognition is concerned with extracting the identity of the person (Tiwari, 2010). Speech Recognition is a method or technology for recognizing sound by recognizing the identity of the speaker where the process is translated from oral language into writing (Anusuya; & Katti, 2010). With another name, the technology is called Automatic Speech Recognition (ASR).

ASR can change the sound captured by the device into a written form. People need a tool to turn writing into a voice and vice versa. This ASR system can capture the incoming voice signal and conclude by paying attention to the frequency of the incoming signal. A word that is entered, for example, the word x is pronounced then the frequency that has entered is calculated the number of waves that can translate the signal (Gruhn,2011).

III. RELATED WORKS

The project "Talking Krishna" is an interactive application to invite elementary school students to practice speaking English directly. In this project, there are three main points of work that have been shown in Figure 1. So, in this book, the explanation that will be displayed is about the making of voice from the "Talking Krishna" application. Moreover, how to arranges the responsive speech builder by Google Assistant. Without any discussion about making the full application and the graphics for making Krishna characters.

Fig. 1. The Flowchart of Talking Krishna Job Description

This application voice would be built by Google Assistant that is a virtual assistant from Google. The advantages of Google Assistant are it can take two-ways conversation so that it can be used efficiently for various people. How to use Google Assistant by connecting two supporting applications from Google that is Google Action and Dialogflow.

A. Google Action

Google Action serves as a platform for the data we will design in Dialogflow. So Dialogflow itself is an agent that functions to organize the data that will be received and captured in the program. Dialogflow is a machine learning that can understand what users are going to say about the program we create.

In the process of creating Talking Krishna, using the toolkit that Google has set up. So the first thing to do is enter as a user into Google Actions, which is a Google-provided platform for voice recognition. After logging into the Google account as a user, open the Google Action's website to start working on the project. In the initial view, there may be an option to create a project then, press the Add / Import Project button to create a new project. Google Actions here serve as a platform provided by Google for voice recognition creation.

B. Dialogflow

Using the stories that have been selected in the book as determining sentence predictions or questions. At once with the response to the answers that will be the response to the application. Prepare for predicting questions that will be submitted by the user. Possible questions are made based on three Krishna stories in the book that has been determined. The example seems like story one:

Story: “One day, Krishna’s friends were forcing him to go to the nearby river for a dip. Krishna was unwilling to go as he knew that a very huge snake, Kaliya, had made the river its home."

TABLE I. The Possible Question in Talking Krishna

Question	*Similar Question*
Why Krishna does not swim	*Why Krishna unwilling to swim*
	Why Krishna does not want to swim
Who is Kaliya	*What is Kaliya*

After the questions are arranged as in Table I, the next process is to enter the data that has been prepared at Dialogflow. In the process of inputting the question and response data, the most important panel is intents. This panel of intents can arrange the distribution of questions and statements that have been compiled beforehand to be designed into a single unit with the response, shown in Figure 2.

Next is the step to build the possible answer to the questions that we already had. As in table 2, which displays the response form or answer for the user from the question that has been expressed. One question with some similar questions has the same response. If the same question has a different response, the Dialogflow invites the answer when the user asks the question that has been made.

Fig. 2. The Intents from Dialogflow

After designing the response to all the questions in the story. The next process is to enter data that has been created in the Dialogflow. The first step is, in the Dialogflow view remains on the full window then scroll down to find the Responses section and then write down the response that has been designed. Again, each response or answer only applies to one part of the question. Also, if not, then when the user submits one of the questions, the answer that will appear randomly.

To make the application easy in recognizing spoken sentences we make a keyword. Then the keyword is made to facilitate the application, with the facility in Dialogflow namely Entities. In this section, a grouping of each word is likely to be frequently spoken by the user by referring to the previously determined story. Can be seen in figure 3, where grouping words based on the main word in the sentence in the story.

Try the word chosen is a simple word, where students sitting on the elementary school can understand what is being said. Because the goal of forming the "Talking Krishna" application is elementary school students.

After all the previous steps have been carried out, it is necessary to do the integration process between Dialogflow and Google action. Set the integration in the Integration panel in Dialoglow. Integration is a panel for connecting Dialogflow to other applications that have worked with Google. As available in the Integration panel between other Viber, Facebook Messenger, and others. After connecting with Google Action, then do a simulation on the project then users can find out the results of the work that has been made. And on Google actions, the settings can be

made regarding various things by utilizing panels that have been provided by Google.

Fig. 3. The Entities of Dialogflow

IV. RESULTS

In this section is a trial process, and the testing that has been done is twice as much testing. The following will explain the analysis of each experimental result. In the first trial was carried out on the researcher, because there were still errors in the final project. However, the fourth trial was carried out to the intended target, because this project had achieved a successful status.

A. First Trial

In this first experiment the dialogflow creation process after passing the project creation phase and entering the intents section. Then after making intents, the data entry process is carried out. Data entry is a collection of phrases about the selected Krishna story that will be displayed in the attachment. Also, in one story all phrases and answers are included in one intent as in Figure 4. When the application has reached the simulation stage using a simulator on a PC and Google Assistant on a mobile phone, there is an error. At the time of trying the user gives a question, but the application cannot respond and does not recognize the intent of the user.

Fig. 4. Intents in Dialogflow for First Trial

In Figure 5 during the experiment, the application cannot recognize the response that has been given by the user and does not capture the intent of the user. In the first experiment, the application cannot detect the user's words due to the inclusion of the possibilities expressed by the user in the form of only phrases. So, the application cannot recognize words rather than users. The application cannot detect the insertion of phrases performed on intents because the phrase itself does not provide specification questions for the application. So that the application cannot recognize the speech given by the user.

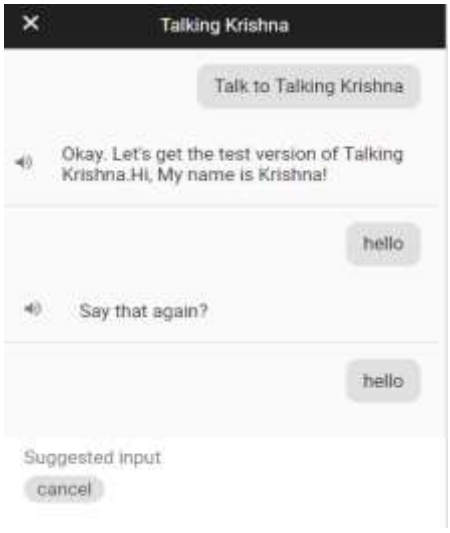

Fig. 7. The Result of First Trial

B. Second Trial

Fig. 6. Dialogflow Intents from Second Trial

In the second experiment, the process was still the same as the first experiment. Start the steps to make Dialogflow and the creation of google action is the same as the previous step. The second difference in experiment location is on the data entered on intents, wherein the first experiment the data entered is a phrase. So in this second experiment is a question and a statement. Queries and questions are made from each sentence in the predefined story. As in Figure 6, the second experiment is done by trying to set the phrase to be changed to a statement or question.

Fig. 7. The Result of Second Trial

Therefore, the predictions submitted by users become more specific, making it easier to detect answers. By setting all answers in one intents section. And at the time of the experiment produced images 7.

Figure 7 is the result of conducting a second experiment, where there is an error in answering when the user gives a question or statement. Questions or statements given by users do not match the answers that have been determined.

In the second experiment, the change from the first experiment is done, which is to change the incoming voice input from a phrase into a question and statement. This experiment produces answers that do not match the question or statement given by the user. This is because in setting the answer question put it into one. Questions or statements in this experiment are arranged in one part of the intents without being distinguished by their specifications. The answers are arranged into one in the intents without any link between the questions that exist with the answers. Thus, this causes the experiment to be carried out and the user gives a question or response statement rather than the application that is not as expected.

C. Final

After the experiment, a trial will be conducted on five elementary school students in table 3. This trial aims to find out how far the design of this application is running. And knowing the lack of words or questions specified in Dialogflow. The things that will be tested are the pronunciation and vocabulary that was chosen for the story and application.

1) The Story Voice-Over

a) Pronunciation

The process of analyzing the pronunciation of the user does not have a problem. The application can still capture the questions and statements expressed by the user. Even though 10 users have

different accents and accents, in English they can be identified. So there is no obstacle in the pronunciation of the user.

However, when the experiment was conducted outside the room the arrest of the voice of the user's conversation was slightly disturbed. This is caused by air factors that are around the user, such as wind factors and the circumstances of the people around. These two things can be factors that can interfere with the use of the application. So, the situation around the user can influence the success factor of the conversation on the application.

b) Vocabulary

In the analysis of the sentence or word used, unfortunately, seven students could not understand the sentence in the story, that shown in figure 9. The choice of words that are too difficult makes users unable to understand the meaning of the story given in the video. The factor is that seven out of ten people think that English is difficult and they do not like English.

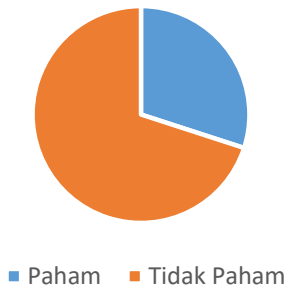

Figure 8. Percentage of Understanding of Vocabulary

Due to the difficulty of the user in understanding the contents of the story, the user cannot ask questions about the story in the application. The question raised by the user is a question that is the identity of a Krishna. Examples of expressed questions such as: "Who are you, where do you live, where you from, and other fundamental questions.

However, three out of ten other students who tried "Talking Krishna" asked several questions which were part of the video being played. And students can understand the story given in the video. They are asking for things they have not understood in the video to Krishna in English. That is because these three people have more abilities because they take regular English language tutoring. So that they can understand the meaning of the story displayed.

V. CONCLUSION

In this chapter explains the conclusions of the final project work that has been done. In this chapter also contains suggestions for the final project that has been done to improve this final project research.

The conclusions that have been obtained from the results of this final project are:

1. The words used in this application are too complicated for elementary school students.
2. The user of this application must be accompanied by parents, teacher, or anyone who can speak English.
3. Users of this application are recommended for grade 4 and above elementary students.

REFERENCES

[1]Suharso. 2014. Pembelajaran Bahasa Inggris Melalui CONTENT-BASED INSTRUCTION. FBS Universitas Negeri Yogyakarta

[2]Agus Taufiq, 2014. Pendidikan di Anak SD. Universitas Pendidikan Indonesia. Sukasari.

[3]Gressia Mellisa. 2008. PENCOCOKAN POLA SUARA (SPEECH RECOGNITION) DENGAN ALGORITMA FFT DAN DIVIDE AND CONQUER. Institut Negeri Bandung. Bandung.

[4]Ganesh Tiwari, 2010,, "Text Prompted Remote Speaker Authentication: Joint Speech and Speaker Recognition/Verification System".

[5]M.A.Anusuya; S.K.Katti,2010. Speech Recognition by Machine: A Review. India.

[6]R.E. Gruhn; Minker, W; Nakamura, S. 2011. Statistical Pronunciation Modeling for Non-Native Speech Processing, Signals and Communication Technology.