

# Developing English Conversation Chatbot Using Dialogflow

Aliv Faizal Muhammad  
Politeknik Elektronika Negeri Surabaya  
Surabaya, Indonesia  
aliv@pens.ac.id

Dwi Susanto  
Politeknik Elektronika Negeri Surabaya  
Surabaya, Indonesia  
dwi@pens.ac.id

Akhmad Alimudin  
Politeknik Elektronika Negeri Surabaya  
Surabaya, Indonesia  
alioke@pens.ac.id

Farah Adila  
Politeknik Elektronika Negeri Surabaya  
Surabaya, Indonesia  
adila1198@mb.student.pens.ac.id

Moh. Hasbi Assidiqi  
Politeknik Elektronika Negeri Surabaya  
Surabaya, Indonesia  
hasbi@pens.ac.id

Salim Nabhan  
Universitas PGRI Adi Buana Surabaya  
Surabaya, Indonesia  
salimnabhan@unipasby.ac.id

**Abstract**—Despite the fact that the use of English language is inevitable nowadays, some people find it difficult to learn. Moreover, the advance of technology has brought impacts on the way how people learn the language such as speech recognition technology that is used to develop learning tools. However, some existing learning media have several shortcomings. Therefore, to bridge this gap, such media to learn and practice the language are needed. This paper focuses on developing English conversation chatbot using speech recognition and artificial intelligence technology with Dialogflow platform as the artificial intelligence engine. Evaluations toward the chatbot were done by the expert review to identify the achievement of indicators and the users to know the level of responses accuracy. The result indicated that almost all agents have reached all indicators and most of the responses had accuracy rate of 100%. It is expected that the availability of the chatbot helps students to improve their conversation skills.

**Keywords**—Learning Media, Chatbot, Artificial Intelligence

## I. INTRODUCTION

English is a language widely used to connect and engage with citizens worldwide. English is used in almost all fields in the modern age as it is today, including technology, schooling, tourism, industry, politics, trade, among others. And the ability to talk English is very essential in the growth of this age of globalization while confronted with intense competition. However, some people get difficulty in learning English. Among the four English skills, students claim that the ability to speak English is more challenging than the other three English skills, writing, listening, and reading [1]

The ability to speak English has become the issue in the language education. Nevertheless, Heriansyah [2] noticed that the problems that often existed during the teaching and learning process when speaking English were 1) students often had no ideas of what to say, so they appeared to be silent, 2) they were often nervous and awkward, not relaxed when making mistakes, 3) students are afraid of making mistakes, as their classmates will laugh at them.

Besides, teachers often apply direct conversation between two or more people and are text-books oriented. The practices included doing the text conversations and reciting them directly. As a result, students got bored since teachers

constantly enforced this method [3]. Besides, students also prefer to do "silent practice" in reading conversation texts while learning speak English skills and some just memorized them. In addition to this, as today's learners are typically digital natives, the nature of students' learning languages have shifted from traditional way toward digital one through technology integration in their literacy practices [4]. Involving multimodalities such as audio and visual by involving technology in teaching and learning English can be beneficial. These new perspectives support the use of technology in teaching and learning English language.

Concerning the using of web-based speech recognition technology, Dio [5] developed learning media to practice English conversation skills. Web-based application using speech recognition was also developed as medium to practice speaking [6]. Students can then learn their English skills in particular digital environment which can be done wherever and anytime. However, the learning media have some shortcomings such as limited vocabulary and pronouncing the words that appear on the website. Therefore learning language using media is needed such as using speech recognition technology namely chatbot technology.

Chatbot is a conversational agent which interacts with users using natural language [7]. The chatbot is made with the Natural Language Processing (NLP) system which is a branch of artificial intelligence. Therefore, the authors develop English Conversation Chatbot by using the Dialogflow platform as the artificial intelligence engine. It is hoped that this chatbot can enhance students' English speaking skills and facilitate them to have digital experiences in their conversation.

## II. THEORITICAL BASIS

### A. Dialogflow

Dialogflow is a platform for natural language understanding that facilitates the design and integration of conversational user interfaces into mobile applications, web applications, devices, bots, interactive voice response systems, etc. By using Dialogflow, it can provide users with a new and interesting way to interact with the products they

produce. Dialogflow can evaluate specific forms of consumer data, including text or audio samples (e.g. from mobile or voice recording). It can also respond to customers in a variety of ways, either by text or by synthetic speech [8]

#### B. Branching Story

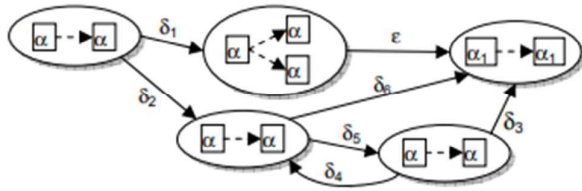


Fig.1. Example of story graph

A branching story structure is a story graph - a directed graph of arc- connected nodes reflecting user choices. Every potential route through the graph represents a story to tell the user. The user's sense of control over the story's progress is constrained in the branching story graph by the number of arcs along a specific direction. Figure 1 gives an example of a story graph. The system commences non-interactively with system-controlled characters performing  $\alpha_1$  and  $\alpha_2$  actions. The user then decides to perform either the  $\delta_1$  or the  $\delta_2$  operation. If  $\delta_1$  is selected, system-controlled characters perform  $\alpha_3$ ,  $\alpha_4$ , and  $\alpha_5$  actions. In the absence of any user action, a  $\epsilon$ -transition is taken [9].

#### C. Chatbot

Chatbot is a computer program designed to communicate with humans through text messages or speech. Usually, chatbot is also equipped with artificial intelligence and natural language processing which makes it an intelligent computer program that can answer human-specific questions. Chatbot is built on the basis of subjects that have been modeled on the knowledge base. Many current chatbots are based on topics and issues that somebody would like to solve for personal or business purposes [7].

#### D. Conversation Design

Conversation Design is a language-based design in human conversation (similar to how object design on pen and paper is a language-based design). The more interfaces an ordinary human interaction has to teach fewer users how to use it. This is a fusion of many disciplines of design including voice user interface design, interaction design, visual design, motion design, audio design, and UX writing. Conversation Design's role as an architect is to map what users can do in space while taking into account user needs and technological constraints. In detailed design specifications, they organize conversations, decide the flow and underlying logic which reflect complete user experience. They are collaborators with stakeholders and developers who can turn to develop and bringing to life experience [10].

#### E. Designing Conversation Storyline

In doing this research, it was found reference related to the design of the conversation storyline in this chatbot. This conversation storyline uses situational conversation, which is a type of conversation that is based on certain situations given. The following is a reference related to the work of this research.

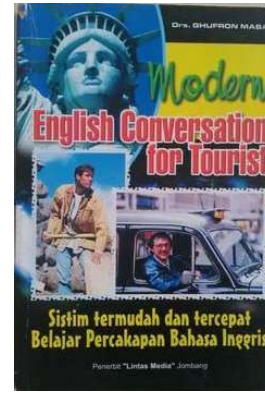


Fig.2. Modern English Conversation for Tourist Cover

This book contains English conversations containing situational conversations. The book can help improve English skills especially for tourists who want to travel between countries. It presents problems in the field of tourism, how to speak words, as well as interesting information and stories that use vocabulary that has been adapted to the current conditions of society [11].

#### F. Precision, Recall, and Accuracy

	Actual Positive	Actual Negative
Predicted Positive	TP	FP
Predicted Negative	FN	TN

(a) Confusion Matrix

$$\text{Recall} = \frac{TP}{TP+FN}$$

$$\text{Precision} = \frac{TP}{TP+FP}$$

$$\text{True Positive Rate} = \frac{TP}{TP+FN}$$

$$\text{False Positive Rate} = \frac{FP}{FP+TN}$$

$$\text{Accuracy} = \frac{TP+TN}{\text{Total}}$$

(b) Definitions of metrics

Fig.3. Machine Learning Evaluation Metrics

Classification specifies cases as either positive or negative. Classification decisions may be reflected in a system known as an uncertainty matrix or a contingency table. The confusion matrix has four categories: True Positive (TP) is a correctly labeled positive example. False Positive (FP) refers to an example that is mistakenly classified as positive. Real Negative (TN) is the negative which is correctly defined as negative. Finally, False Negative (FN) refers to positive, incorrectly labeled negative examples [12].

The Confusion Matrix is shown in Figure 3(a). Confusion matrix can be used to create points in ROC space or in the PR space. The journal determines the metrics used in each space given the uncertainty of the matrix, as shown in Figure 3(b). One plot is False Positive Rate (FPR) on the x-axis in the ROC space, and True Positive Rate (TPR) on the y-axis. FPR measures a small portion of negative examples that are mistakenly considered positive. TPR analyses a small portion of the positive samples which are labeled correctly. One recall plot in the PR space is on the x-axis, and the y-axis is precision [12].

A definition for each metric is given in Figure 3(b). The journal deals with metrics as functions that act on the underlying confusion matrix, which defines points in the ROC or PR space [12].

### III. METHODOLOGY

#### A. Study Existing

Existing studies carried out by looking for similar application references. Then obtained the advantages and disadvantages of the application. As a result, we found English conversation training applications, Speechace, Duolingo, and Eviebot.

#### B. Making a Conversation Storyline

In making this conversation storyline, first is determining the overall theme of the conversation, then determining each of its scenes and the setting of the place as well as character to illustrate each scene. After that it is made a sample conversation with the expected response taken based on the reference sources described earlier.

TABLE I. CONVERSATION STORYLINE

Scene	Character & Background Scene
Making a passport at the Immigration Office	Background scene : Immigration office Character : Officer
Exchange currency to Singapore Dollar	Background scene : Foreign Exchange Bureau Character : Officer
On the way to the airport by taxi	Background scene : Taxi Passenger seat Character : Driver
Boarding Passes at The Airport	Background scene : Boarding pass gate Character : Boarding Pass Officer
Buy an EZ link card to use the MRT in Singapore	Background scene : Passenger service Character : Passenger Service Officer
Asking about the direction of the MRT line in Singapore	Background scene : Airport Character : Other passenger
Check in Hotel	Background scene : Receptionist Character : Receptionist Officer
Getting Food Recommendations in Singapore	Background scene : Inside MRT Character : Other passenger
Exploring Lor Mee at Xin Mei Xiang stall	Background scene : Old Airport Road Food Centre (in front of Xin Mei Xiang stall) Character: Waitress
Exploring Castella Cake at the Chong Pang Old-time Flavor stall	Background scene : Old Airport Road Food Centre (in front of Chong Pang Old-time Flavor stall)

Scene	Character & Background Scene
	Character: Waitress
Getting to know someone at Merlion Park	Background scene: In front of Merlion Park Character : Other passenger
Buy Souvenirs at Bugis Street	Background scene : Bugis Street Character: Seller

#### C. Branching Storyline

After creating a conversation storyline, the next step is sketching the storyline branching route. The goal is that to anticipate when the user enters input answers that are not in accordance with the expected answers that have been made. This branching storyline was created with an open-source tool, Twinery. The following is one sample branching storyline shown in the following image.

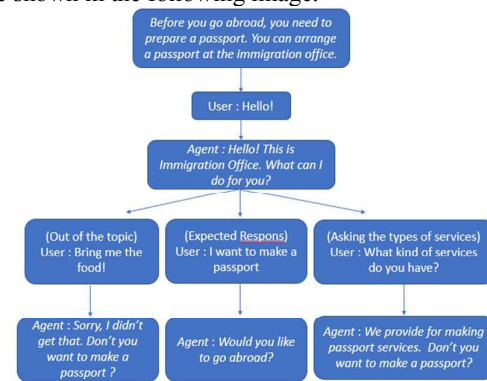


Fig.4. Branching Storyline Sample

#### D. Making an Agent

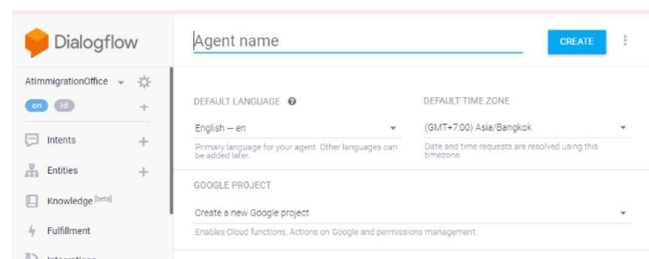


Fig.5. Dialogflow Console Agent

The first step that must be done before creating a chatbot in Dialogflow is to create an agent first. Agent is a natural language understanding module that will later train and understand human language, so the conversation becomes more natural. This agent will be created as many as 12 according to the scene in the storyline that has been made.

### E. Making an Intent

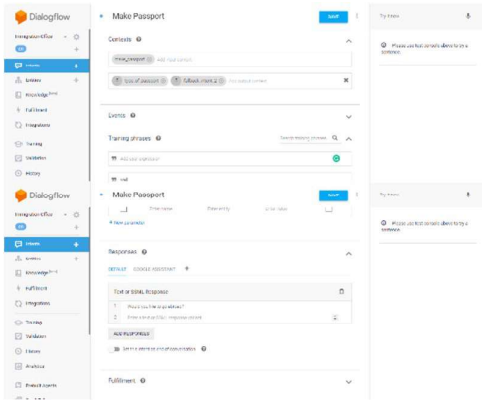


Fig.6. Intents on Dialogflow Console

The figure shows the making intent of a console dialogflow agent. In intent, there is a collection of several questions from users that contain the same topic. If the user sends a question, the agent will match the question that entered the data with the intents. Then, the agent will send the correct answer. If a question is sent that does not match or does not match the intent that was made, then the agent will send an answer from the fallback intent.

### F. Intent Parameter

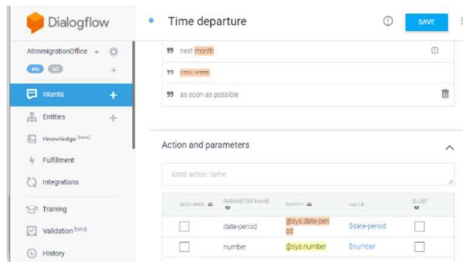


Fig.7. Intent Parameter

The figure shows the intent parameters in a console dialogflow agent. Inside there are parameters, entities, and values. These parameters must be checked in the required section. For example, if the user enters input data that does not contain dates or numbers as shown in the image (according to entities that have been checked), then this system will ask again the time of departure according to the conversation topic. Meanwhile, Entity is a tool to identify and extract data spoken by users, such as time, date, place, weather, etc. In addition, Entity can also be customized according to the conversation topics needed in the entities menu.

### G. Making an Entity

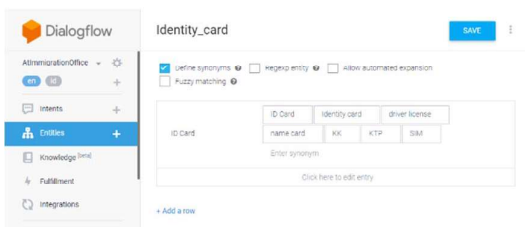


Fig.8. Entities on Dialogflow Console

The figure shows the creation of an entity in an agent Dialogflow. The creation of this entity aims to identify and extract data spoken by the user. The entities menu functions to create your own entities that do not yet exist in system entities. In system entities, there is already automatic data extraction in the form of time, date, place, weather, country, etc.

### H. Making Context

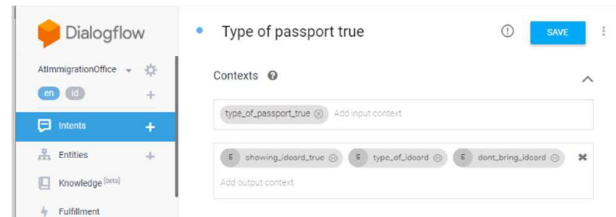


Fig.9. Context on Dialogflow Console

The figure shows the context in a console Dialogflow agent. Context is a tool used to regulate conversation flow so that conversations become structured and ordered. In this section, the developer determines his own name in the contexts section. There are 2 types of context, namely: input context and output context.

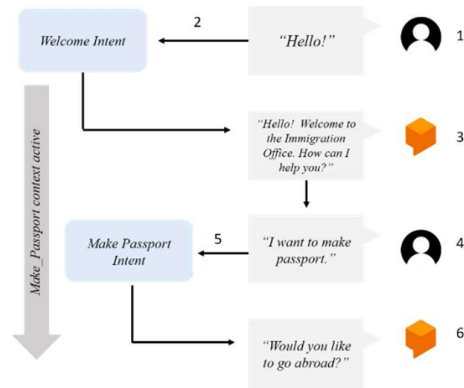


Fig.10. Context Flow

In the figure shows an example of the flow of conversation that will occur. The user first says "Hello", then, this agent will match the response with the intent welcome. The intent has a make passport output context, so the context is active. Then the agent asks "How can I help you?". Users respond with "I want to make a passport." Then the agent will match the response with the intent make passport. The Intent has a make\_passport input context. Then, the agent responds as shown. The role of context here, connecting various intents. Thus, branched conversations will become structured and ordered.

## I. Test Console

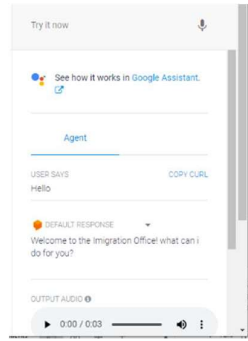


Fig. 11. Test Console on Dialogflow Console

After compiling the conversation flow through context, it is then simulated through the test console located on the side of the dialog flow console. The aim is to ensure that the conversations in the chatbot are structured.

## J. Web Demo

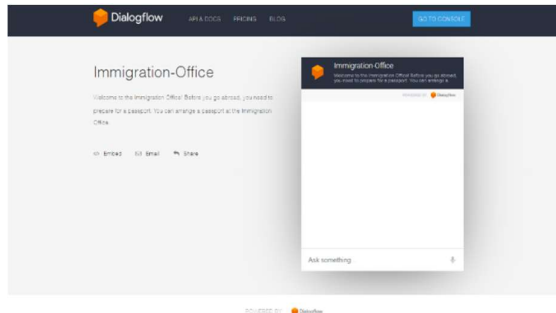


Fig. 12. Web Demo

In the figure is a web demo located on the integration menu in the Dialogflow console. This web demo will be used as a simulator or platform to be tested on users so they can interact with chatbots that have been made. The testing mechanism is the user will get as many as 12 web demo links. As previously explained, there will be 12 agents according to the number of scenes. So, each agent has a different demo web menu.

## K. Conversation History

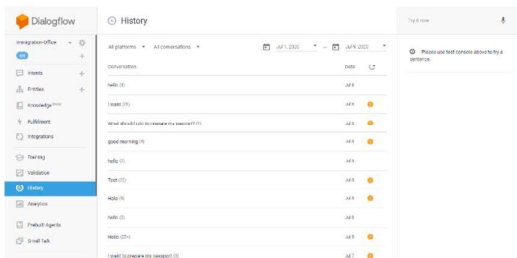


Fig. 13. History Menu

The figure shows that after the chatbot is tested on the user, the conversation will be recorded while the user interacts. On the console dialog flow history menu, there is a history of data conversations between users and agents such as date, time, platform type, number of interactions, to display conversations that are not in accordance with the

intent that has been made. Then, the analysis will be carried out with a precision, recall, and accuracy approach. So we get the accuracy or suitability of the agent in answering the response from the user.

## IV. TESTING

The test is carried out to the expert review and users. The component is tested by expert review to know whether there are indicators of the achievement of someone who can speak English in the chatbot conversation content that has been made. Meanwhile, the component tested by the user is the level of accuracy or suitability of the agent in responding to sentences submitted by the user.

### A. Testing of Expert Review

In this part of the test, an expert review is carried out through Google Form, which contains indicators of the achievement of someone in speaking English as follows:

1. Greetings
2. Welcoming
3. Introduction
4. Asking and Giving Information
5. Congratulations
6. Apologizing and Making Excuses
7. Giving Advice or Suggestion
8. Giving Order or Instruction
9. Expressing Requests & Prohibitions
10. Expressing Agreements & Disagreements
11. Thanking and Accepting Thanks
12. Partings

These indicators are classified for each agent, whether the agent contains the indicators mentioned. Review testing was conducted by the two experts in the field of English education. Furthermore, the test results can be seen in the Appendix. It can be concluded:

- Agent Immigration-Office

The agent has reached almost all indicators except for indicators congratulations.

- Money-Changer Agent

For these agents, indicators that have not yet been reached are introduction, congratulations, and partings.

- Go-to-the-airport agent

For these agents, indicators that have not yet been reached are welcoming, introduction, congratulations, apologizing and making excuses, and partings.

- Agent Boarding-Passes-at-The-Airport

For these agents, indicators that have not yet been reached are welcoming, introduction, congratulations, apologizing and making excuses, and partings.

- Agent Getting-EZ-Link-Card

For these agents, indicators that have not yet been reached are welcoming, congratulations, giving advice or suggestions, and giving orders or instructions.

- Agent Asking-MRT-Line

For these agents, indicators that have not yet been reached are welcoming, introduction, and congratulations.

- Agent Check-In-Hotel



For these agents, almost all indicators are achieved except indicators congratulations.

- Agent Getting-Food-Recommendations

For these agents, indicators that have not yet been achieved are introduction and congratulations.

- Agent Exploring-Food-at-Singapore

For these agents, indicators that have not yet been achieved are introduction and congratulations.

- Agent Exploring-Food-at-Singapore-Part-II

For these agents, almost all indicators are achieved except indicators and congratulations.

- Going-Vacation-to-Merlion-Park Agent

For these agents, indicators that have not yet been reached are welcoming, congratulations, and giving advice or suggestions.

- Agent Buy Souvenirs at Bugis Street

For these agents, indicators that have not yet been achieved are indicators of congratulations, thanking and accepting thanks, and partings.

The results of testing by an expert review can be concluded that almost all agents have reached all the indicators someone can speak English. However, it is not with congratulations indicators. In the future, it can be used as a suggestion to create and develop new scenes with indicators of congratulations.

## B. Testing of Users

As explained earlier, in this part of the test, approaches precision, recall, and accuracy are used. So that the accuracy or suitability of the agent is obtained in response to sentences submitted by users. The mechanism of this test is that the user try to interact with the agent through a demo or demo web platform located in the Dialogflow integration console section. Then, during interaction, the conversation is be recorded in the history menu section. Next, the data in the history menu is analyzed.

- Agent Immigration-Office Test Results

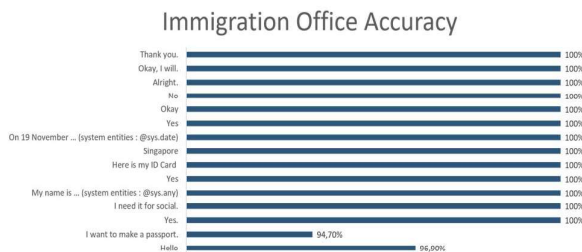


Fig.14. Agent Immigration-Office Accuracy

The figure explains that almost all responses to Agent Immigration-Office have an accuracy rate of 100%. Except for the "Greetings" and "I want to make a passport" response. Each gets an accuracy rate of 96.9% and 94.7%.

- Agent Foreign-Exchange-Bureau Test Results

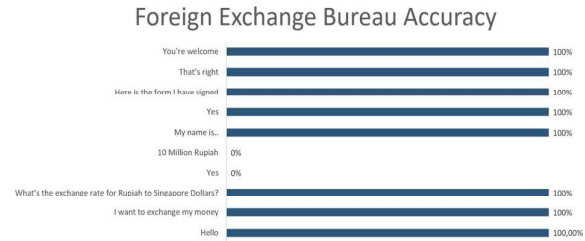


Fig.15. Agent Foreign-Exchange-Bureau Accuracy

The figure explains that almost all responses to the Foreign-Exchange-Bureau Agent have 100% accuracy. Except for the "Yes" and "10 Million Rupiah" responses. In this response get an accuracy rate of 0%.

- Agent Go-to-The-Airport Test Results

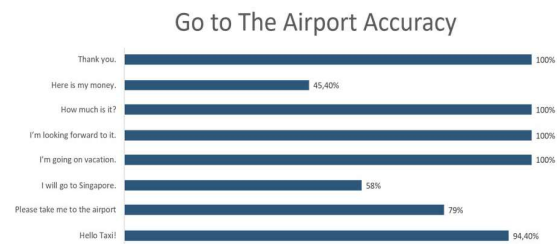


Fig.16. Agent Go-to-The-Airport Accuracy

The figure explains that there are 4 responses on Go-to-The-Airport Agents that have an accuracy rate of 100%. In addition, the responses "Hello Taxi", "Please take me to the airport", "I will go to Singapore", and "Here is my money" get 94.4%, 79%, 58%, and 45.4%.

- Agent Boarding-Passes-at-The-Airport Test Results

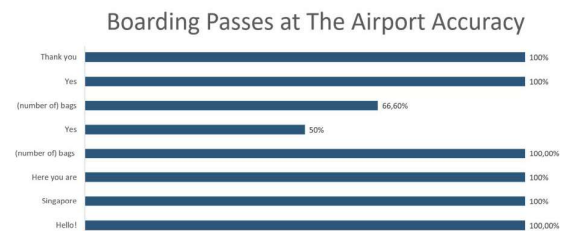


Fig.17. Agent Boarding-Passes-at-The-Airport Accuracy

The figure explains that almost all responses on Agent Boarding-Passes-at-The-Airport have a 100% accuracy rate. Except for the "Yes" and "Number of bags" responses. In this response, the accuracy rate was 50% and 66.6%, respectively.

- Agent Getting-EZ-Link-Card Test Results  
Getting EZ Link Card Accuracy

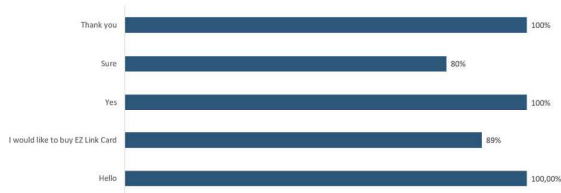


Fig.18. Agent Getting-EZ-Link-Card Accuracy

The figure explains that almost all responses on Agent Getting-EZ-Link-Card have a 100% accuracy rate. Except for the responses "I would like to buy EZ Link Cards" and "Sure". In this response, the accuracy rate was 89% and 80%, respectively.

- Agent Asking-about-MRT-Line Test Results  
Asking MRT Line Accuracy

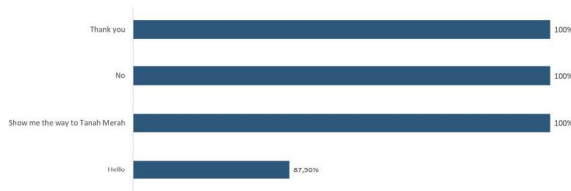


Fig.19. Agent Asking-about-MRT-Line Accuracy

The figure explains that almost all responses to the Asking-about-MRT-Line Agent have an accuracy rate of 100%. Except for the "Greetings" response, which gets an accuracy rate of 87.50%.

- Agent Check-In Hotel Test Results  
Check-In Hotel Accuracy



Fig. 20. Agent Check-In Hotel Accuracy

The figure explains that all responses to Agent Check-In-Hotel have a 100% accuracy rate.

- Agent Getting-Food-Recommendations Test Results  
Getting Food Recommendations Accuracy

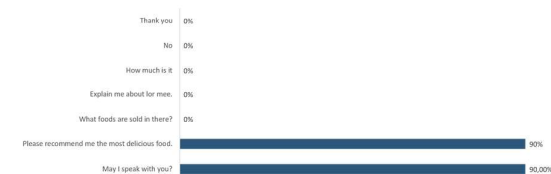


Fig. 21. Agent Getting-Food-Recommendations Accuracy

The figure explains that almost all responses to the Getting-Food-Recommendations Agent have an accuracy rate of 0%. Except for the responses "Greetings" and "Please Recommend me delicious food" which gets an accuracy rate of 90%.

- Agent Exploring-Food-at-Singapore Test Results  
Exploring Food at Singapore Accuracy

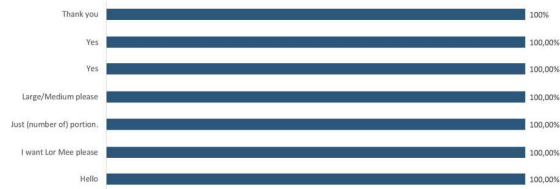


Fig. 22. Agent Exploring-Food-at-Singapore Accuracy

The figure explains that all responses to Agent Exploring-Food-at-Singapore have a 100% accuracy rate.

- Agent Exploring-Food-at-Singapore-Part-II Test Results  
Exploring Food at Singapore Part II Accuracy

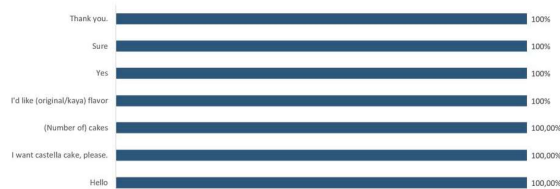


Fig. 23. Agent Exploring-Food-at-Singapore-Part-II Accuracy

The figure explains that all responses to Agent Exploring-Food-at-Singapore have a 100% accuracy rate.

- Agent Going-Vacation-to-Merlion-Park Test Results  
Going Vacation to Merlion Park Part II Accuracy

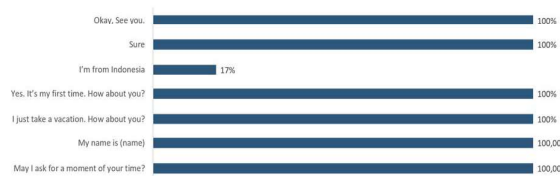


Fig. 24. Agent Going-Vacation-to-Merlion-Park Accuracy

The figure explains that almost all responses in Agent Going-Vacation-to-Merlion-Park have an accuracy rate of 100%. Except for the response "I'm from Indonesia" which gets an accuracy rate of 17%.

- Agent Buy-Souvenirs-at-Bugis-Street Test Results  
Buy Souvenirs at Bugis Street Accuracy

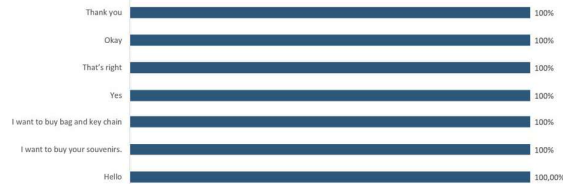


Fig.25. Agent Buy-Souvenirs-at-Bugis-Street Accuracy

The figure explains that all responses to Agent Buy-Souvenirs-at-Bugis-Street have a 100% accuracy rate.

## V. CONCLUSION

The English conversation chatbot using speech recognition and artificial intelligence technology with Dialogflow platform as the artificial intelligence engine has been developed and evaluated by the expert review and the users. From the results of this research, it can be identified that through storyline design and branching conversations, making chatbots, and testing, the level of accuracy of the response of an agent increase if the variation of the vocabulary data set is increased, and the conversation branching path is expanded more. This media can be utilized by language learners to practice English conversation. Besides, there is still a space to develop further this application.

## REFERENCES

- [1] Y. Bani Ahmad, "Student Perceptions on English pronunciation after course phonetics and phonology," *Judika (Jurnal Pendidik. Unsika)*, vol. 6, no. 1, pp. 12–16, 2018.
- [2] G. Amaral *et al.*, "Speaking problems faced by the English department students of Syiah Kuala University," *Ling. Didakt.*, vol. 6, no. 1, pp. 37–44, 2012, doi: 10.1017/CBO9781107415324.004.
- [3] E. A. Amalo, I. D. Agusalm, and C. D. Murdaningtyas, "Developing visual novel game with speech-recognition interactivity to enhance students' mastery on English expressions," *J. Sos. Hum.*, vol. 10, no. 2, pp. 129–130, 2017, doi: 10.12962/j24433527.v10i2.2865.
- [4] S. Nabhan and R. Hidayat, "Investigating literacy practices in a university EFL context from multiliteracies and multimodal perspective: A case study," *Adv. Lang. Lit. Stud.*, vol. 9, no. 6, pp. 192–199, 2018, doi: <http://dx.doi.org/10.7575/aiaa.all.v.9n.6p.192>.
- [5] D. E. Pratama, "Pengembangan aplikasi berbasis web dengan speech recognition sebagai media pembelajaran latihan conversation bahasa Inggris," *Politeknik Elektronika Negeri Surabaya*, 2019. [Online]. Available: <https://redaksi.pens.ac.id/2019/07/25/development-of-web-based-application-with-speech-recognition-as-english-learning-conversation-training-media/>. [Accessed: 01-Jul-2020].
- [6] A. F. Muhammad, D. E. Pratama, and A. Alimudin, "Development of Web Based Application with Speech Recognition As English Learning Conversation Training Media," in *IES 2019 - International Electronics Symposium: The Role of Techno-Intelligence in Creating an Open Energy System Towards Energy Democracy, Proceedings*, 2019, pp. 571–576, doi: 10.1109/ELECSYM.2019.8901594.
- [7] Accenture Interactive, "Chatbots in customer service," 2016.
- [8] "Dialogflow basics," *Google Cloud*. [Online]. Available: <https://cloud.google.com/dialogflow/docs/basics>. [Accessed: 02-Jul-2020].
- [9] M. O. Riedl and R. M. Young, "From linear story generation to branching story graphs," in *Proceedings of the 1st Artificial Intelligence and Interactive Digital Entertainment Conference, AIIDE 2005*, 2005, pp. 111–116.
- [10] "Conversation design—What is conversation design?," *Designguidelines.withgoogle.com*. [Online]. Available: <https://designguidelines.withgoogle.com/conversation-design/what-is-conversation-design.html#>. [Accessed: 02-Jul-2020].
- [11] G. Maba, *Modern English Conversation for Tourist*. Jombang: Lintas Media, 2000.
- [12] L. A. Staeheli and D. Mitchell, "The relationship between precision-recall and ROC Curves," in *Proceedings of the 23rd International Conference on Machine Learning*, 2006, pp. 546–559, doi: 10.4135/9780857021113.n29.