A Conversational Agent for English Writing Skill Learners

Akihito Nakano
Osaka Institute of Technology
Osaka, Japan
e1c16069@st.oit.ac.jp

Thanaruk Theeramunkong
SIIT, Thammasat University
Pathumthani, Thailand
thanaruk@siit.tu.ac.th

Hirotoshi Taira
Osaka Institute of Technology
Osaka, Japan
hirotoshi.taira@oit.ac.jp

***Abstract*— This paper presents a design and a preliminary implementation of a conversational agent to support learning English writing for non-native speakers. As one of the most important components in improving English writing skills is to understand how to write English passages that are complied with the standard English grammar. We construct a chatbot that utilizes a number of well-established platforms and services, such as Ginger Grammar API, LINE platform, and Heroku database platform, and its hosting service platform. Towards response generation, we introduce a set of rules for generating explanation. A chatbot is created and its performance is analyzed. The chatbot extracted words related to tense and location from the text input. The chatbot responds to questions that match with the focused tense. As another feature, if the location is detected, the system will show some tips related to that location. If the input sentence is grammatically incorrect, the system will rewrite the sentence to comply with the grammar and presents the cause of incorrectness as feedback to the user. In addition, if there is a paraphrased expression in the entered text, it will be indicated.**

Keywords—chatbot, conversational agent, English writing skill

I. INTRODUCTION

In recent years, with the development of information and communication technology and the spread of smartphones, communication by chat applications such as LINE and Facebook Messenger becomes popular rapidly, where the demand for chatbots is increasing. A chatbot is a program that performs conversation like a human. After the ELIZA [1] was developed in 1966, a lot of chatbots have been developed for some special tasks such as an agent for a restaurant reservation, and a chatbot that handles inquiries instead of a person.

This time we focused on several grammar checkers and grammar correction tools available in the market, including Grammarly, ProWritingAid, WhiteSmoke, Ginger, JetPack, PaperRater, Slick Write, 1Checker, etc. These applications are very useful for revising sentences with a few simple explanations. Most of these tools aim to correct grammar and check wording, not to learn English to acquire a second language. However, towards efficient English writing skill acquisition, the following features are preferable.

1. The system should help correct spelling.
2. The system should help suggest suitable words.
3. The system should help suggest suitable phrases.
4. The system should help support grammar correction.
5. The system should help provide explanation on errors.
6. The system should interactively provide guidelines for English learners.
7. The system should provide English grammar knowledge.
8. The system should analyze where are the grammatical errors and how to solve them.

In [2], Swain advocates that just teaching only knowledge to the language is insufficient for second language acquisition, and practice by using language-related knowledge is also necessary. We need to focus on the output of knowledge to become speak English well. Even though non-native speakers learn English for many years, most of them cannot use English fluently, due to unskilled vocabulary usage, unfamiliar grammar structure, cultural difference, etc. To solve this problem, we propose a new learning environment using LINE [3], a popular communication tool where interactive and conversational learning as well as personalized coaching is possible using Messaging API [4]. Our writing-supported chatbot can help correct spelling, suggest suitable words/phases, support grammar correction, explain errors, interactively provide guidelines for English learners, such as English grammar knowledge with error analysis.

II. IMPLEMENTATION

A. Chatbot architecture

The chatbot is implemented using Python 3.6 as a development language, Heroku as a platform, MySQL as a database. For interface and other functions, it uses the LINE Messaging API, Ginger Grammar API, Google translation API, and Power Thesaurus. The LINE Messaging API enables two-way communication among LINE users. The Ginger Grammar API is used to check the input sentence whether it complies with the pre-defined grammar and whether it includes any wrong spelling or not. The chatbot library enables simple conversations. Heroku is a hosting service called PaaS (Platform as a Service) that makes it possible to publish the developed application and make full use of the service. It's like a rental server. Network infrastructure, servers, operating systems, databases, and other platforms are already available, we can focus on building application development. Heroku has 150 add-ons, and databases using MySQL have been added as add-ons.

B. Development function

This chatbot has the following three functions.

1) *Translation*

When using the translation function, for example, if you want to translate the sentence “I am Japanese” from English to Japanese, you need to type “I am Japanese trans”. You need to type “(space) trans” at the end of the sentence you want to translate. Fig. 1 shows the result of some translations.

Fig. 1. Translation sample

2) *Grammar correction function*

Grammar correction function is always active, and there is no need to type any words at the end like the translation function. If the text input by the user has something to be corrected, the chatbot will feedback together with the correct text, the wrong text and what was wrong.

3) *Rephrase function*

Rephrase function is also always active. If there is a paraphrasable phrase in the text input by the user, the chatbot will feedback with it. Fig. 2 shows the message returned from the chatbot when "I will go to library" is entered.

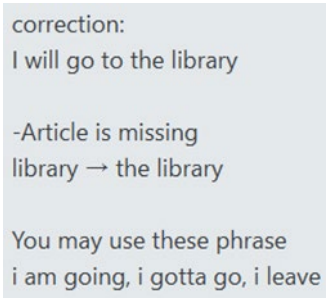

Fig. 2. Response sample

C. System structure

Fig. 3 shows the system structure of this chatbot. There is also a database in Heroku. Fig. 4 shows the detailed structure between LINE, Heroku, and APIs.

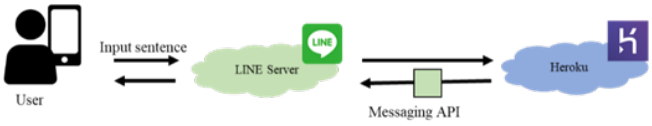

Fig. 3. System structure

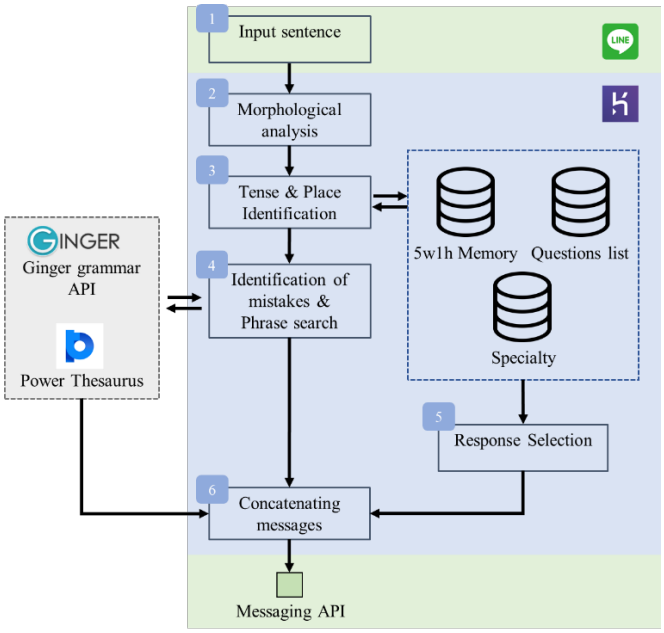

Fig. 4. Internal structure in Heroku and LINE Database

1. The user inputs a sentence. The input text is formatted into a format that is easy to process.

2. Formatted text is divided into word levels and parts of speech are given to the divided words. The parts of speech given are as shown in Fig 5.

3. Determine the tense from words that are divided and given parts of speech (auxiliary verbs, verbs), and select all questions that fit the tense. There are three types of tense to distinguish: past, present, and future.

4. In addition, the word corresponding to the location (library, university, Kanchanaburi, Osaka, etc.) is searched from the divided words. If there are words in the database of Specialty, the tips related to the location is output from the database of Specialty (Only 47 prefectures in Japan are supported).

5. The same question as the word tense determined in step 3 is randomly selected from the database of 5w1h. Interrogative word of the selected question is flagged so that the same question is no longer asked. If there is a place word in the divided words, rewrite the database of 5w1h value of “whr” in the database from 0 to 1.

6. Compare text input by the user with text modified using the Ginger grammar API and point out what was grammatically incorrect. Then it uses the modified text to extract a paraphrase from the power thesaurus.

7. Finally, the items selected, listed, and extracted in 3, 4, 5 and 6 are reformatted into a format that can be used by the messaging API and linked as a message returned to the user.

EX: existential there
IN: preposition or conjunction, subordinating
JJ: adjective or numeral, ordinal
JJR: adjective, comparative
JJS: adjective, superlative
MD: modal auxiliary
NN: noun, common, singular or mass
NNP: noun, proper, singular
NNPS: noun, proper, plural
NNS: noun, common, plural
PRP: pronoun, personal
PRP$: pronoun, possessive
RB: adverb
RBR: adverb, comparative
VB: verb, base form
VBD: verb, past tense
VBG: verb, present participle or gerund
VBN: verb, past participle
VBP: verb, present tense, not 3rd person singular
VBZ: verb, present tense, 3rd person singular
WDT: WH-determiner
WP: WH-pronoun
WP$: WH-pronoun, possessive
WRB: Wh-adverb

Fig. 5. Type of part of speech

Fig. 5 shows the database configuration. This database has two tables. First one is a table of questions list for a user. Second one is a table of recording whether the questions have been used before. The first question list table is shown in the Fig. 5. Each question divided into three types of questions: past, present, and future. There are 5W1H questions. The second table, for example, if place word (library, University, Kanchanaburi, etc.) is included in a sentence input by the user, rewrite status of "where (whr)" from 0 to 1 (Fig. 6).

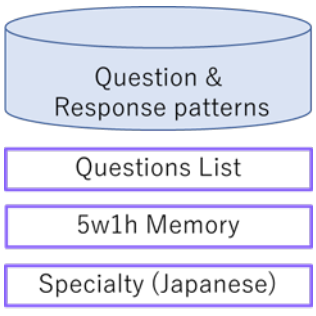

Fig. 6. Database configuration

id	sentence	tense	attribute
1	What did you do there?	1	0
2	what are you going to do there?	2	0
3	What will you do there?	2	0
4	What are you doing now?	3	0
5	Where did you go?	1	3
6	Where are you going?	2	3
7	Where will you go?	2	3
8	Who did you go with?	1	4
9	Who is going with you?	2	4
10	Who is with you?	3	4
11	How was it?	1	5
12	How is it going?	3	5
13	How's it going?	3	5
14	when did you go there?	1	1
15	when are you going to there?	2	1

Fig. 7. Question list

user_id	what	whn	why	whr	who	how
[redacted]	0	0	0	0	0	0
[redacted]	0	0	0	0	0	0
[redacted]	0	1	0	1	0	1

Fig. 8. 5W1H memory status

III. Testing Experiment

In fact, this chatbot was tested by 10 non-native speaker students. When the user greets, the chatbot says, "Where have you been recently? Or are you planning to go?". If the user answers this, the conversation continues. After the experiment was completed, we took a five-grade evaluation survey on the dialogue with the chatbot. In addition, we took a free description type questionnaire on the dialogue with the chatbot.

IV. Result and Discussion

Table 1 shows the contents of the questionnaire and the results of the questionnaire conducted after the experiment. In addition, Table 2 shows the results of the free description type questionnaire that wrote what the user noticed using this chatbot.

There was an opinion that it was fun to have a conversation because the chatbot points out what was wrong with the sentence sent by the user and what other ways to say it. On the other hand, there were opinions that the user didn't get a response fitting the sentence sent by the user and variations of conversation are not enough. This chatbot system is the side where the user answers, and the chatbot is the questioning side. Therefore, it is considered that the user's question was not answered because there was no answer prepared in advance.

Identify applicable funding agency here. If none, delete this text box.

From the questionnaire results, it found that there are people who are bored with the conversation with the chatbot. we should think of reactions that more please users.

Table 1. Results of questionnaire

Question	Evaluation average
Is this chatbot friendly?	3.7
Is this chatbot's reply easy to understand?	4.4
Did you feel boring with chatbot's reply?	3.6
Do you think want to use this chatbot again?	3.6
Do you think this chatbot is suitable for learning English writing?	4.0

Table 2 Comment about conversation with a chatbot

- I thought it is better that there are more bigness responses.
- Sometimes, we don't communicate very well.
- It's fun to have a conversation while studying what is wrong.
- I think the chatbot has pointed out a grammatical error, but sometimes, the explanation is incorrect.
- It is interesting to show the knowledge of prefectures.
- I think it's more fun when there are many variations of conversation.
- I want to see English questions and Japanese questions together.

V. CONCLUSION

In this paper, we designed and preliminarily implemented a conversation agent to supports learning English writing for non-native speakers. After experimenting with a chatbot dialogue system, it was found that grammar correction and paraphrase functions were highly evaluated. However, we found that it is not the system that the user wants to talk again.

In the future, we will work to improve the accuracy of error indication in grammar correction. we would also like to increase the number of conversations and develop a system that enables more natural conversation between the chatbot and the user.

REFERENCES

[1] Weizenbaum, Joseph. "ELIZA---a computer program for the study of natural language communication between man and machine." Communications of the ACM 9.1 (1966): 36-45.

[2] Swain, Merrill. "The output hypothesis: Just speaking and writing aren't enough." Canadian modern language review 50.1 (1993): 158-164.

[3] Line, https://line.me/ja/

[4] Line Messaging API, https://developers.line.biz/ja/services/messaging-api/