English Learning Assistant Application with a Translation Approach Using Rule Based System

Daniel M. D. U. Kasse, Patrisius Seran and Christa E. B. Bire
Politeknik Negeri Kupang, Penfui, Kupang, NTT, Indonesia

Keywords: English Learning, Dictionary, Rule Based, Translation.

Abstract: English has moved to improve its linguistic status, which is starting to act as a second language and language of instruction. This is often found in the social sphere, especially in youth, and education. The use of English also has an impact on education and the economy. In the field of education, the availability of sophisticated translation tools online opens the way for students to translate a text more easily. However, in real conditions, Indonesian students still find it difficult to practice their English skills. The purpose of this research is to make an application as a tool to learn how an English sentence can be translated into Indonesian. This application not only displays the translation results but also displays the types of words, sentence patterns to produce English translations and materials.

1 INTRODUCTION

The use of English in Indonesia in general has a great influence, although it is still a foreign language. English has moved to improve its linguistic status, namely starting to act as a second language and language of instruction. This is often found in the social sphere, especially youth, and education. The use of English also has an impact on education and the economy.

In the field of education, the availability of sophisticated translation tools on-line opens the way for students to translate a text more easily, media that expose them to English such as programs on television, music, commands on social media applications and their gadgets. can increase their knowledge. on how to understand English text and convert information into Indonesian. However, problems can also arise from tools and media that help them understand English, such as if students receive translations that do not match the true meaning after they enter sentences in the source text, without considering polysemy, transliteration, and culture. the use of language, the idioms used and so on, the result of the translation may not be accepted because it does not convey the idea that the author of the source text is trying to convey.

English teaching in Indonesia must be improved if you want to get better results. Student needs should be the focus of attention in teaching. The success of teaching English depends on student achievement in terms of the goals that have been determined before the program starts. English teachers must be open- minded and ready to keep learning and strive for better teaching outcomes. All the necessary information and knowledge must be used to make teaching English successful. English teachers should be able to choose teaching materials such as books, journals, audio-video tape recorders and cassettes, independent access and computerized language teaching to facilitate language learning so that students can achieve effective language learning (Richards, 2001:230). We need to be aware that language teaching methods can change over time as can fashion. But we can always judge whether a particular method is suitable for our purposes. It is important for us to be open-minded and ready to try new methods to improve the quality of teaching English. In real conditions, Indonesian students still find it difficult to practice their English language skills in communicating. The purpose of this research is to create an application that not only displays the results of the translation but also how an English sentence is translated into Indonesian.

Kasse, D., Seran, P. and Bire, C.
English Learning Assistant Application with a Translation Approach Using Rule Based System.
DOI: 10.5220/0012064600003575
In *Proceedings of the 5th International Conference on Applied Science and Technology on Engineering Science (iCAST-ES 2022)*, pages 1079-1083
ISBN: 978-989-758-619-4; ISSN: 2975-8246

2 RULE BASED SYSTEM

Rule Based System (RBS) is an expert system that uses rules to present knowledge. This RBS theory uses a simple technique, starting with the basic rules that contain all the knowledge of the problems encountered which are then coded into if-then rules containing data, statements and initial information. The system will check all the if condition rules that define the subset, conflict sets that exist. If found, then the system will perform a then condition. This loop will continue until one or two conditions are met, if the rules are not found then the system must exit the loop (terminate).

In the application of the Rule Based System in the process of translating sentences from English to Indonesian, it is based on the rules of English grammar. English has several sentence patterns, namely:

- Pattern 1: Simple Sentence

  One independent clause (SV.). Example: Mr. Potato Head eats monkeys.
  I refuse.

- Pattern 2: Compound Sentence

  Two or more independent clauses. Example: Mr. Potato Head eats them for breakfast every day, but I don't see the attraction.

- Pattern 3: Complex Sentence

  One independent clause PLUS one or more dependent clauses. Examples: He recommends them highly because they taste like chicken when they are hot.

- Pattern 4: Compound-Complex Sentence

  Two or more independent clauses PLUS one or more dependent clauses. Example: Mr. Potato Head said that he would share the secret recipe; however, if he does, Mrs. Potato Head will feed him to the piranhas, so we are both safer and happier if I don't eat monkeys or steal recipes.

To be able to read input from the user, a parsing process with context free grammar is used. Most systems for modeling constituent structures in English or other natural languages are using Context Free Grammar or CFG.

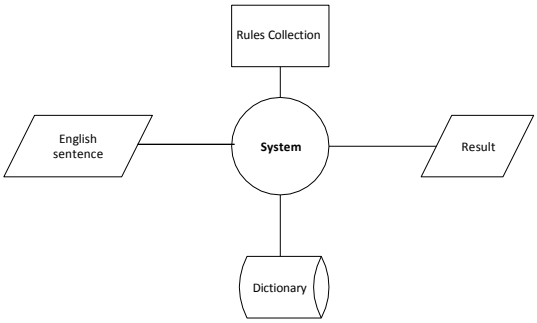

Figure 1: Rules-based translator system.

For a Context Free Grammar has four parameters (technically called 4-tuple)

- **N**: A collection of non-terminal symbols (or variables){NP, VP, PP}
- ∑: Terminal symbol set {det, noun, verb,..}
- P: The production set, which is expressed in the form A → α where A is a non-terminal symbol {<NP> → det noun}
- 4. S: Start symbol

The parsing method used is top down. The step is to find all the appropriate grammar rules. In the top- down parser there is a strategy called depth-first which looks for the appropriate grammar for each of the first and subsequent entries.

3 RESULTS AND DISCUSSION

3.1 System Implementation

The translator system is implemented using four modules that have different functions. Here are the four modules along with a brief description.

3.1.1 Modul Indentifikasi

The initial stage is the preprocessing stage which is carried out in this module. At the initial stage, a checking process will be carried out on user input, whether it is a word or not (in this case it is a phrase or sentence). Input in the form of words will be translated on the translator module while input in the form of phrases or sentences will undergo a preprocessing process. The preprocessing stage includes changing abbreviations (contractions) such as the word I′ ll to be changed to I will. In the simple future tense, there is the use of tobe going to which is changed to will. The tokenization stage is carried out to get tokens in the form of words

which are then converted into a single form such as the word kicks which is converted into the word kick. The next stage is checking the words and tenses. Input that meets the criteria (in this case, all words are in the word dictionary database and the tenses used include simple present tense and simple future tense) will be given to the parser module to perform the scanning process (parsing). The initial stage is the preprocessing stage which is carried out in this module. At the initial stage, a checking process will be carried out on user input, whether it is a word or not (in this case it is a phrase or sentence). Input in the form of words will be translated on the translator module while input in the form of phrases or sentences will undergo a preprocessing process. The preprocessing stage includes changing abbreviations (contractions) such as the word I′ ll to be changed to I will. In the simple future tense, there is the use of tobe going to which is changed to will. The tokenization stage is carried out to get tokens in the form of words which are then converted into a single form such as the word kicks which is converted into the word kick. The next stage is checking the words and tenses. Input that meets the criteria (in this case, all words are in the word dictionary database and the tenses used include simple present tense and simple future tense) will be given to the parser module to perform the scanning process (parsing).

3.1.2 Modul Parser

This module implements syntax rules in context free grammar. The process for analyzing syntax rules is called parsing. The parsing method used is top down with a depth first strategy (top down depth first parser). The parser will analyze the syntax structure of the input provided by the user so that the structure of the user input is obtained.

Table 1: Syntax rules.

Non terminal		Terminal			
<S>	→	<NP>	<VP>		
<S>	→	aux	<NP>	<VP>	
<S>	→	wh	<NP>	<VP>	
<S>	→	wh	aux	<NP>	
<S>	→	wh	aux	<NP>	<VP>
<NP>	→	det	noun		
<NP>	→	pron			
<VP>	→	verb	<NP>		
<VP>	→	verb	inf	<NP>	
<VP>	→	verb			

The symbol to the left of the arrow is a non-terminal symbol and to the right of the arrow is a terminal symbol. Table 2 is a description of some of the symbols used.

Table 2: Symbol Description.

Simbol	**Arti**
<S>	Sentence
<NP>	Noun Phrase
<VP>	Verb Phrase
adj	adjective
pron	pronoun
aux	auxiliary
det	determinant
not	negative word
noun	noun
verb	verb
wh	wh-question
propnoun	proper noun
inf	to infinitive

Pada proses *parsing*, sistem akan melakukan pencocokan seluruh aturan sintaks terhadap input user. Berikut merupakan contoh dari proses *parsing* (dengan konsep *parsing tree*) dalam *context free grammar*.

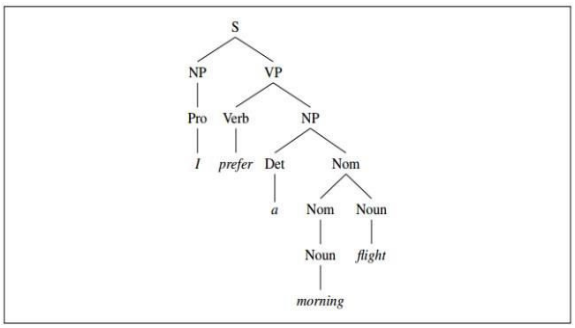

Figure 2: Parsing tree.

From the results of the parsing process, the input structure from the user is pron verb det noun. In the parsing process there are several processes

including 'expand', 'match', and 'backtrack'. Expand is the process of replacing non-terminal symbols into terminal symbols according to the existing production rules in the syntax rules such as <NP> to pron. Match is the process of matching user input word types with syntax rules. Backtrack is the process of going back to a non-terminal symbol from a previous production.

3.1.3 Modul Translator

In this module, the MD-DM pattern rules are implemented, namely the Explaining Explained word pattern which is commonly found in English texts is changed to the Explained - Explained word pattern which is commonly found in Indonesian texts. An example of the MD pattern is the big house. Table 3 shows the MD-DM pattern rules used.

Table 3: Aturan Pola MD-DM.

No.	**JK1**	**JK2**	**JK_1**	**JK_2**
1	Aux	Not	Not	Aux
2	Adj	Noun	Noun	Adj
3	Det	Noun	Noun	Det
4	Propnoun	Noun	Noun	Propnoun
5	Noun	Noun	Noun	Noun

For user input in the form of word units, the translation will be carried out directly without going through these stages. The results of this module are in the form of word tokens in Indonesian.

3.1.4 Modul Materi

At this stage, the system will display the material according to the type of words obtained from the identification.

3.2 System Testing

In the process of testing the system with a black box can be determined by studying the input and output. In this test, the focus will be on testing from the original language (English) to the target language (Balinese) whether it is in accordance with the expected results based on the system design and the suitability of the system interface. Table 4 is some of the data used in testing the system using the black box method.

Table 4: Black Box Method Test Data.

No	Data Sample	Keterangan	Task	Status
1	I Love You	Correct input (simple sentence pattern)	Process and display the translation results in Indonesian	OK
2	I was happy	Correct input (simple sentence pattern)	Process and display the translation results in Indonesian	OK
3	My elder brother became an engineer in 1988	Input that has a complex sentence pattern	Processes and displays a warning that Input has a complex sentence pattern	OK
4	was happy I	Input that has an unrecognized pattern	Processes and displays a warning that the Input has pattern not recognized	OK

4 CONCLUSION

From the implementation of the concept of translating English text to Indonesian text using the Rule Based method as well as from the results of translation testing as above, several conclusions can be drawn, namely:

- This English to Indonesian text translator can translate sentences in "simple sentence pattern" quite well.
- This English to Indonesian text translator is able to recognize the types of words in the entered sentence.
- This research is still in progress so it is not yet fully completed

REFERENCES

E. Gultom, "English Language Teaching Problems in Indonesia," *7th Int. Semin. Reg. Educ.*, vol. 3, (2015).

J. Sudrajat and T. H. Pudjiantoro, "Aplikasi Penterjemah Bahasa Indonesia Ke Bahasa Inggris Dengan Menggunakan Metode Terselia (Supervised Learning),"
J. Comput. BISNIS, vol. 2, no. 1, (2008).

M. Nadhianti, "An Analysis of Accuracy Level of Google Translate In English- Bahasa Indonesia and Bahasa Indonesia-English Translations," *Sastra Inggris-Quill*, vol. 5, no. 4, (2016).

R. A. Sandra, "From English to Indonesia: Translation Problems and Strategies of EFL Student Teachers - A Literature Review," *Int. J. Lang. Teach. Educ.*, vol. 2, no. 1, (2018), doi: 10.22437/ijolte.v2i1.4520.

S. Simaremare, D. E. Silalahi, P. Saut Raja Sihombing, and Y. Kristina Sinaga, "The Accuracy of the Translation of English Verb Phrase into Indonesian Using Google Translate," *JETAFL Publ.*, vol. 7, no. 1, (2021).

S. Wahyuningsih and M. Afandi, "Investigating English speaking problems: Implications for speaking curriculum development in Indonesia," Eur. J. Educ. Res., vol. 9, no. 3, (2020), doi: 10.12973/EU-JER.9.3.967.

Y. Farahsani, I. P. Rini, and P. H. Jaya, "Google Translate Accuracy in Translating Specialized Language From English to Bahasa Indonesia", Jan. (2021), doi: 10.2991/assehr.k.210120.156.