

## **THE ERROR ANALYSIS OF GOOGLE TRANSLATE AND BING TRANSLATOR IN TRANSLATING INDONESIAN FOLKLORE**

Jufriadi<sup>1</sup>, Amalia Asokawati<sup>2</sup>, Magfirah Thayyib<sup>3</sup>  
<sup>123</sup>State Islamic Institute of Palopo, Palopo, Indonesia  
<sup>3</sup>magfirah\_thayyib@iainpalopo.ac.id

### **ABSTRACT**

The objective of this research is to find out the lexical errors made by Google Translate and Bing Translator in translating Indonesian folklore "Princess Tandampalik" and "Sigarlaki and Limbat." The research applied the qualitative method. The data were analyzed using hybrid taxonomy of error analysis from Vilar, et al. The results of this research show that Google Translate made 103 errors in total which consist of 12 missing words, 19 errors in word order, and 64 incorrect words, and 8 unknown words. Meanwhile, Bing Translator made 95 errors which consist of 5 missing words, 1 error in word order, 88 incorrect words, and 1 unknown word. Incorrect word is the most frequent error found in the translation resulting from Google Translate and Bing Translator with a total of 152 errors. The incorrect words mainly occurred in the translation of adjectives and adverbs in which Google Translate and Bing Translator mostly translated them into noun form. Thus, it can be concluded that both translator machines' performances are not different because they have their advantages and disadvantages.

Keywords: Error Analysis, Translation Error, Translator Machine

### **INTRODUCTION**

Translation is a field of study that deals with translation theory, description, and application. Newmark (1988) stated that translation is a process of rendering the meaning of a text into another language in the way that the author intended the text. Because it examines translation both as an inter-lingual transfer and as intercultural communication, translation study can also be described as an inter-discipline that touches other diverse fields of knowledge, including comparative literature, cultural studies, gender studies, computer science, history, linguistics, philosophy, rhetoric, and semiotics. Additionally, translation enables people from all over the world to communicate effectively. It serves as a channel for knowledge transmission, an agent of cultural heritage, and is critical to developing a global economy.

In the translation process, some principles have to be considered by the translator to be able to bridge and maintain the linguistic and extra-linguistic aspects from one language to another language. It is because every language has its particular system. The principles of

good translation are it does not deviate from the content of the source language; it is easy to be understood by the readers; the use of sentences follows the rules of the target language; the translation concerns the disclosure of the content more than the appearance of speech; and it does not seem to be a translation but rather an original work (Sayogie, 2014). For sure, these kinds of principles should be considered in translating even when using a translator machine.

The existence of translator machines cannot be ignored when discussing the advancement of translation technology. Even though it cannot completely replace the role of human translators, it cannot be denied that the existence of this translator machine can aid the work of those who require translation results in a timely and efficient manner. The translator machines can help people to translate short text to quite long text. The existence of a translator machine enhances the significant role of translation in world communication, knowledge transmission, and economic development. The translator machine is also able to translate any kind of text and speech model.

Even though translator machine has many speech recognition models, it still has its limitation. The accuracy of the output of the translator machine is questionable. The translator machine cannot identify the context and cannot understand cultural terms, jokes, or homonyms, which are not part of everyday vocabulary. Translator machine only translates and produces the language by following the mechanism of the system. Therefore, translator machines cannot fulfill the whole principles of a good translation and it will influence the result of the translation. That is why the translation resulting from translator machines usually contains some errors. Vilar, et al. (2006) have shown that translation errors can be in the form of missing word, word order error, incorrect word, unknown word, and punctuation error. This hybrid taxonomy of error analysis by Vilar, et al. can be used to assess the quality of translator machine performance. Meanwhile, some people often accept the translator machine result without re-checking the validity of the translation output.

The outputs of translator machines have been previously studied by some researchers. Fadilah (2017) focused on finding and recognizing the translation errors in semantic aspects of an Instagram translator. Susanti (2018) also studied the types of lexical errors produced by Instagram translator and how it represents genuine language. While Veronika (2018) compared the performance between Instagram translator and human translator. Adiputra (2019) analyzed the translation errors caused by Google Translate and Bing Translator in terms of semantic elements according to Koponen's theory. Sukma (2019) focused on the Youtube Auto-Generated Subtitled performance in translating content for Vogue Magazine Channel. Those previous studies provide useful knowledge of the performance and limitation of the translator machine. They focused on assessing the performance of translator machines by looking at the errors found in the translation results.

The translation results from the translator machines still need to be studied further to add more knowledge on the performance and limitation of the translator machines. This research focused on analyzing the lexical errors in the translation results of Indonesian folklores from Google Translate and Bing Translator. Google Translate and Bing Translator are two translation engines that are widely used by people in general. The translation tool of Google recently can translate more than fifty languages, one of which is from Indonesian to English and vice versa while Bing Translator which is previously known as windows live translator provides free text and website translations (Almahasees, 2018).

The Indonesian folklore is chosen in this research because it is essential to introduce Indonesian culture using good translation with simple diction/lexicon. Folklores commonly use more advanced words or forms such as idioms that should be translated fine. It suits one kind of translation that is idiomatic translation instead of form-based translation. Idiomatic translation or meaning-based translation aims to convey the meaning of the Source Language (SL) text in the receptor language's natural forms (Larson, 1998). The idiomatic translation is also similar to aesthetic-poetic translation. It refers to translation in which the translator considers the original version's affect, emotion, and feeling, the original author's aesthetic form, and any information contained in the message, for instance, a rhyme, a dramatic dialogue, or a novel (Choliludin, 2007).

Specifically, this research aims to find out the translation errors that mostly occurred with Google Translate and Bing Translator in Translating Indonesian folklores "Princess Tandampalik" and "Sigarlaki and Limbat" to English. This research is different from the previous researches on translation errors in terms of the genre of text translated and the aspect of language analyzed in the translation result. It is to add specific knowledge on the limitation of the two popular translator machines in translating a cultural text like folklore. This research is expected to contribute to the development of the translation field as an additional reference for translator machines' evaluation. The result of the research could be beneficial information for the users of translator machines.

## **METHOD**

This research applied a descriptive-qualitative method. The researchers collected, classified, analyzed, and drew conclusions without making any generalizations. The researchers became the main instrument and were helped by two validators. The main Source Text (ST) was two Indonesian folklores, "Princess Tandampalik" and "Sigarlaki and Limbat", published by Bhuana Ilmu Populer (Hu, 2019a; Hu, 2019b). Meanwhile, the Target Text (TT) is the translation result from two translator machines, Google Translate and Bing Translator. The researchers also used Target Text Key (TT-Key) which had been previously validated by the expert, Mrs. St. Hartina, S.Pd., M.Pd., the translation lecturer at IAIN Palopo.

The "document translation" feature was used in translating using Google Translate. Meanwhile, translation using Bing Translator was done by translating all sentences. The translated texts by Google Translate and Bing Translator were then read by the researchers.

The analysis of the performance of both translator machines was done by comparing their performance based on the number of errors in each sub-category. The analysis of the lexical errors produced by both translator machines was done using Vilar, et al.'s theory (2006). The data were tabulated in Ms. Excel to count and compare the translation errors caused by two translator machines. The data were described to the best understanding of the lexical errors problem.

## RESULTS

The data in this research consist of translations from the source language (SL) (Indonesia) to the target language (TL) (English) which is translated using Google Translate and Bing Translator. The data identified are phrases and sentences from two Indonesian folklores. There are 188 sentences translated by Google Translate and Bing Translator which consist of 96 sentences of "Princess Tandampalik" and 92 sentences of "Sigarlaki and Limbat".

The results of the data analysis have been validated by Dr. Masruddin, S.S., M.Hum., the lecturer of Error Analysis at IAIN Palopo. To show the comparison of the errors for each sub-category, the data were arranged in the table. Table 1. shows the errors identified in the performance of Google Translate and Bing Translator. It shows that the errors mainly occurred in lexical aspects by Google Translate and Bing Translator are incorrect words. A total of 152 errors in the form of incorrect word were found. The incorrect words mainly occurred in the translation of adjectives and adverbs in which both translator machines mostly translated them into noun form.

*Table 1. Total Errors Found in Google Translate and Bing Translator*

Types of Error	Google Translate	Bing Translator
Missing Word	12	5
Word Order	19	1
Incorrect Word	64	88
Unknown Word	8	1
Punctuation Error	0	0
<b>Total</b>	<b>103</b>	<b>95</b>

Out of 5 error categories according to Vilar, et al.'s theory, both Google Translate and Bing Translator made errors in 4 categories. Those errors are missing word, word order, incorrect word, and unknown word. Both translator machines did not make any punctuation error. In the folklore text translation, there are several variations of translation errors found. Due to a large amount of data, the researchers only give one example to represent each sub-category error. Each example consists of Source Text (ST) from Indonesian folklores and Target Text (TT) of the translation by translator machines.

### ***The Errors Found in English Translation by Google Translate***

#### *Missing Word*

The error happens when a word in the target sentence is missing, for instance in the text of Princess Tandampalik:

ST : *Tabib kerajaan **tak sanggup** mengobati*

TT : physician kingdom **was able** to treat

The example above shows a lexical error in the form of missing word. In the translation, the missing is the word "the" as the article before the word "physician." Here, the word "the" represents a definite article that is used to identify a particular person or object. The translation output by Google Translate also missed the word "not" after the word "was," in which "not" is used as an adverb to describe the verb "was." Thus, based on this analysis, the sentence should be translated into "The royal physician could not heal her."

#### *Word Order*

Word order error occurs to a word in both source and target language. It refers to syntactic ordering which can be distinguished as local or long-range orders, for instance in the text of Princess Tandampalik:

ST : ***Bencana buruk** terjadi*

TT : **The disaster worse** happened

The example above shows the lexical error made by Google Translate in the form of word order. Here, the word "worse" is an adjective that must be put before the word "disaster" which is a noun. Then, the correct sentence should be "The worse disaster happened."

#### *Incorrect Word*

This type of error occurs because the translator machine is unable to find the correct translation of a given word. Thus, Google Translate chooses an incorrect word in the target language, for instance in the text of Sigarlaki and Limbat:

ST : ***Lambat** yang sampai di dapur kebingungan*

TT : **The waste** that arrived in the kitchen was confused

The example above shows the lexical error made by Google Translate in the form of incorrect word. In the output sentence, Google Translate incorrectly translated the word "Lambat" which is the subject of the sentence into the word "the waste." Thus, the result of this translation is not appropriate. The sentence should be translated into "Lambat felt perplexed when he was in the kitchen."

#### *Unknown Word*

The unknown word occurs when the system misplaces the word into the source language. It can happen when the translator machine system cannot find the source language in its system. This type of error often occurs in the process of translating slang language, a traditional language, and the words that the translator machine system has not detected, for instance in the text of Sigarlaki and Limbat:

ST : *Namun, Limbat menuruti perintah dengan **tawakal***

TT : However, Limbat obeyed orders with **tawakal**

The example above shows that Google Translate made the unknown word error. The word "*tawakal*" is an adjective that in Indonesian is the same as the word "*sabar*" or "*berserah diri*." So the word "*tawakal*" can be translated into "patiently" or "helplessly." Thus, the sentence can be translated into "However, Limbat helplessly obeyed what his master instructed."

### ***The Errors Found in English Translation by Bing Translator***

#### *Missing Word*

Here is an example of a translation error in the form of missing word made by Bing Translator:

ST : *Hari yang buruk **membuat** Sigarlaki merasa kesal*

TT : A bad day upset Sigarlaki

It can be seen from the target language that Bing Translator omitted the word "*membuat*" as the verb of the sentence. The word "*membuat*" in the source language exists, wherein English is "made." If the word "made" does not exist in the target language, it will disrupt the meaning of the sentence because there is no verb in the sentence and the sentence will be unclear. Thus, the sentence can be translated into "The bad day ruined Sigarlaki's mood."

#### *Word Order*

Here is the example of a translation error in the form of word order made by Bing Translator:

ST : *Limbat segera muncul **dengan senang***

TT : Limbat **soon appeared happy**

In the example above, there are two types of errors found: incorrect word and word order. Bing Translator translated the phrase "*dengan senang*" into "happy." The use of the word "happy" is incorrect and should be changed to "happily." The word "*dengan senang*" functions as an adverb that should be placed after the word "*muncul*" which functions as the verb of the sentence. The word "*segera*" which was translated as "soon" also functions as an

adverb and should be placed after the verb “appeared”. So, the correct sentence should be “Limbata appeared soon happily.”

#### *Incorrect Word*

Here is the example of translation error in the form of incorrect word made by Bing Translator:

ST : ***Dia*** mengingat adat turun-temurun, jika Putri dilamar kerajaan seberang, negeri akan tertimpa musibah beruntun

TT : ***She*** remembers hereditary customs, if the Princess proposed to the opposite Kingdom, the country will be hit by successive disasters

Bing Translator translated the word "dia" with the wrong word choice. Before the sentence “*Dia mengingat adat turun-temurun, jika Putri dilamar kerajaan seberang, negeri akan tertimpa musibah beruntun*”, there is the sentence “King of Luwu is really worried”. So, the word “*dia*” in the source language refers to the word “King of Luwu” (*Raja Luwu*). The word “*dia*” as a subject pronoun must be translated as “he” in the target language. The source text (ST) should be translated into “He remembered the custom passed down from generation to generation, that if a prince from another kingdom proposed to the princess, consecutive disasters would hit the kingdom.”

#### *Unknown Word*

Here is the example of translation error in the form of unknown word made by Bing Translator:

ST : *Alkisah di Kerajaan Luwu yang **sentosa**, tinggallah Putri Tandampalik yang jelita*

TT : In the kingdom of Luwu **sentosa**, there lived the beautiful Princess Tandampalik

From the sentence, it can be seen that Bing Translator did not translate the word “*sentosa*” in the source language. The word “*sentosa*” could be translated into “prosperous.” So, the correct sentence could be “Once upon a time, in the prosperous Kingdom of Luwu, a beautiful princess named Princess Tandampalik lived.”

## **DISCUSSION**

Google Translate and Bing Translator are two tools that resulted from the development of technology in the translation field. Both translator machines provide quick translation services with multiple languages either through the web or application. Since all translation processes are done by Google Translate and Bing Translator, the human translator did not take a role in or during the translation process. The translation quality of Google Translate and Bing Translator is based on the system of the translator machine itself.

There are 188 sentences from two Indonesian folklores namely “Princess Tandampalik” and “Sigarlaki and Limbat” that have been analyzed. The errors were identified and then classified based on the hybrid taxonomy of error analysis by Vilar, et al.'s

namely missing word, word order, incorrect word, unknown word, and punctuation error. Based on the result of the research, the error mostly occurred in the translation made by Google Translate and Bing Translator was incorrect word with a total of 152 errors. Google Translate made 64 incorrect words in translating the two Indonesian folklores while Bing Translator made 88 incorrect words. The error with the smallest number found in the translation of Google translation and Bing Translator was unknown word. There were 17 unknown words in Google Translate output text and 2 unknown words in Bing Translator. Although Google Translate has a large amount of vocabulary, it does not guarantee that no errors will be found in the translation results. The same is true for Bing Translator that it is still confused in selecting words.

By looking at the total number of errors made by each translator machine, it can be said that Google Translate made more errors than Bing Translator in translating Indonesian folklores. However, both Google Translate and Bing Translator mostly make the translation error in the form of incorrect word. The incorrect word occurred more in Bing Translator than Google Translate. Thus, it can be implied that both translator machines have nearly identical performances. The differences between Google Translate and Bing Translator according to the results of this research are almost equal to their similarities. Both Google Translate and Bing Translator also show accuracy by not making any punctuation error.

The result of this research is similar to the previous study on the lexical errors made by the Instagram translator machine done by Susanti (2018). The result of Susanti's research shows that the Instagram translator produced so many errors. The incorrect word and missing word are the most frequent errors found in the translation of captions through the research (Susanti, 2018). She concluded that Instagram translator cannot represent genuine language. In a couple of data in this research, Google Translate and Bing Translator cannot represent the genuine language of Indonesian folklores into English.

Besides, the result of this research is different from that previous research on the translation analysis of Google Translate and Bing Translator conducted by Adiputra (2019). The result of Adiputra's research shows that Bing Translator made double errors than Google Translate. This is probably due to the difference in the theories used to categorize the errors of translation. In his research, Adiputra used Koponen's theory while, in this research, the researchers used Vilar, et al.'s theory. The folklores translated using translator machines are also different.

Furthermore, there are several advantages and disadvantages found when we use a translator machine. The researchers found specific advantages and disadvantages of using translator machines in this study. The advantage, of course, is in terms of translation speed, in which users can quickly get the translation results they want. Whilst, the most noticeable disadvantage is the lack of accuracy of translator machines like Google Translate and Bing Translator in translating literary texts such as folklore or other cultural texts.

Other benefits provided by the translator machine include it is affordable and does not cost as much as hiring human translator services, it is easy to access and user-friendly, also it can translate text into several languages reciprocally and simultaneously. On the other hand, another shortcoming of the translator machine is that it cannot understand the context or meaning. It only translates text literally with less or relatively low accuracy and credibility. This is evident in the example of idiom translation from the text of Indonesian folklores. What is more, the translator machine does not consider the cultural differences between languages.

Because translator machines focus on statistical rather than traditional rules-based analysis, they do not use grammatical rules in their systems. It is based on a technique known as statistical translator machine. Translation errors and oddities are made by translator machines because they use statistical matching to translate sentences. They often exchange common phrases for similar but unequal general terms in language and reverse the meaning of sentences. Moreover, from the data of this research, it can be implied that either Google Translate or Bing Translator is still weak in translating complex sentences. The limitation of the characters in the translator machine system is the cause.

When using a translator machine to translate a text of literature like folklore or other literary works, the user must first comprehend the source language comprehensively. As well, the user must be fluent in and understand the target language. It is also critical to double-check the translation results produced by the translator machine. Re-checking is required so that the source text's substance and meaning do not change and can be transferred comprehensively into the target language.

## **CONCLUSION**

Based on the results and discussion, the type of error that mostly occurred by Google Translate and Bing Translator in translating Indonesian folklores "Princess Tandampalik" and "Sigarlaki and Limbat" to English is incorrect word with a total of 152 errors. Google Translate made 64 incorrect words in translating the two Indonesian folklores while Bing Translator made 88 incorrect words. The incorrect words mainly occurred in the translation of adjectives and adverbs in which both translator machines translated them into noun form. The other types of errors made by Google Translate and Bing Translator in translating Indonesian folklores are missing word, word order error, and unknown word.

In the end, the translator machine still could not replace the role of human as an ideal translator. The researchers recommend to the translator machines' creators to evaluate and improve the quality of text translation in both websites and applications. It could be adding correction features and eliminating text translation limits. Next, the researchers suggest the users of translator machines, especially those who use it for academic purposes, to continually upgrade for better translator machines. Using the latest version of the translator machine application will give users the latest features that may affect the translation text results. The users should also pay attention to the limitation and the common errors made

by the translator machine that will help refine the translation results. Then, the researchers suggest that the study on translator machines should be continued and have a more comprehensive analysis. Since this research focuses on Indonesian folklore, the pattern may differ when other kinds of translation are analyzed.

## REFERENCES

- Adiputra, M. F. (2019). *Error Analysis in the Performance of Google Translate and Bing Translator in Translating Children's Story Book Pancuran Pangeran*. Undergraduate Thesis, UIN Sanata Dharma Yogyakarta. Retrieved from <http://repository.usd.ac.id/id/eprint/33179>
- Almahasees, Z. M. (2018). Assessment of Google and Microsoft Bing Translation of Journalistic Texts. *International Journal of Languages, Literature, and Linguistics*, 4(3) pp. 231–35. <https://doi.org/10.18178/ijlll.2018.4.3.178>
- Choliludin. (2007). *The Technique of Making Idiomatic Translation*. Bekasi: Kesaint Blanc.
- Fadilah, E. M. (2017). *Semantic Error Analysis of Instagram Machine Translation from Indonesian to English*. Undergraduate Thesis, UIN Syarif Hidayatullah Jakarta. Retrieved from <https://repository.uinjkt.ac.id/dspace/handle/123456789/35316>
- Hu, Lilis. (2019a). *Seri Dongeng 3D Nusantara: Putri Tandampalik*. Jakarta: Bhuana Ilmu Populer.
- Hu, Lilis. (2019b). *Seri Dongeng 3D Nusantara: Sigarlaki dan Limbat*. Jakarta: Bhuana Ilmu Populer.
- Larson, M. L. (1998). *Meaning-Based Translation: A Guide to Cross-Language Equivalence*. Maryland: University Press of America.
- Newmark, P. (1988). *A Textbook of Translation*. London: Prentice-Hall International.
- Sayogie, F. (2014). *Teori dan Praktik Penerjemahan*. Tangerang Selatan: Transpustaka.
- Sukma, K.R.P. (2019). *YouTube Auto-Generated Subtitle Performance in Translating Content in Vogue Magazine Channel*. Undergraduate Thesis, UIN Sanata Dharma Yogyakarta. Retrieved from <http://repository.usd.ac.id/id/eprint/36258>
- Susanti, E. (2018). *Lexical Errors Produced by Instagram Machine Translation*. Undergraduate Thesis, UIN Maulana Malik Ibrahim. Retrieved from <http://etheses.uin-malang.ac.id/id/eprint/14231>
- Veronika, S. (2018). *Instagram Translate's and Human Translation's Performance in Translating the Captions in @Basukibtp Instagram Account*. Undergraduate Thesis, UIN

Sanata                      Dharma                      Yogyakarta.                      Retrieved                      from  
<http://repository.usd.ac.id/id/eprint/18344>

Vilar, D., J. Xu, L. F. D'Haro, and H. Ney. (2006). Error Analysis of Statistical Machine Translation Output. *Proceedings of the Fifth International Conference on Language Resources and Evaluation*, pp. 697–702. Retrieved from <https://www.lrec-conf.org/proceedings/lrec2006/pdf/413>