Journal of English Language Teaching, Linguistics and Literature

Journal homepage: journal.iainlhokseumawe.ac.id

The Analysis of Errors on Translating Informative Texts by Google Translate

Titik Ismailia[1]

[1]*Department of Language, Communication, and Tourism, State Polytechnic of Jember, Jawa Timur, Indonesia, 68121.*

ARTICEL INFO

Keywords:
Error Analysis
Translation
Informative Text
Google Translates

History:

Received (6 July 2022)

Revised (8 August 2022)
Accepted (10 August 2022)

ABSTRACT

This research aims to describe and classify the errors that was made by students in translating informative text in English into "Bahasa Indonesia" with Google translate (GT) application. This research applies errors analysis method with content analysis approach. The procedure to do the research are: data collection, identification of errors, classification of errors, and statement of errors frequency. The result shows that the errors mostly found in lexis level such as omission, adding, and untranslated both in content and function. While in the grammar level, there were some errors found like the word class changing, tenses, and agreement. In semantic level, there were errors on confusion of senses, wrong choice, and idioms. Lastly, in discourse level there were errors on style, variety, and should not be translated. To conclude, the lexical errors that found in the result of translating informative text by Google is for about 98 %. The other errors also found in grammar for about 0,007%, semantic for about 0.008% and discourse for about 0.005%. It means that Google translate made many errors in translating informative text, but it can be as an alternative choice to get the rapid translation result from English text into Bahasa Indonesia.

INTRODUCTION

The need of translation is for spreading knowledge, information, and concept to all over the world. It can help people to communicate effectively in different language and country because it is not only sharing information but also culture. This condition makes translation as the way to make communication run smoothly. Students can read many books or e-books and articles in any language, sharing ideas with friend from another country, or sending emails. Businessman can also do their job from home by using internet and gadgets. The rapid growth of online communication connects each other to apply translation in daily lives. Basically, translation could come to the usage of national language by using technology in our lives from computer programs to product labels (Imre, 2012).

The use of different language with their own rules leads to the problems on understanding meaning, grammar, and culture. For example: students need to get the meaning of an English word through bilingual dictionary. They also learn grammar by writing a sentence (Chandra & Yuyun, 2018). By reading articles, students can learn cultures in everyday living, doing business, or religion and cuisine of different country. This strategy helps the students to be more confidence when they are able to comprehend the meaning or a word or sentence in their own language (Abbasian & Biria, 2017). Translation is the way to facilitate the communication process by answering the problems on interpreting ideas from L1 (source language) to L2 (target language). The primary function of translation is to address appropriate meaning of a word or sentence linguistically, semantically, and pragmatically (Al-Musawi, 2014).

Since the important function of translation in communication, the quality of translated text in target language rely on the competences of translator in language, text, and culture. Translation competence is to

identify as expecting expertise in various fields includes linguistic competence, cultural competence, textual competence, domain/subject specific competence, research competence, and transfer competence (Schafner, 2005). Furthermore, the translation activity in 21st century is not only done by translation industry such as translation company, translation agency, or translation departments in company but also translating being done by people and machine translation (MT).

Machine translation (MT) is a process that is a computer automatically translates text from one source language to a different target language. Two considerations of using MT are for fast and accurate translations. Some examples are Google translate, Microsoft Translator, and Bing Translator. Google Translate is a translation service by Google that is directly translates words, phrases, web pages between English and over 100 other languages. As a statistical MT service since 2006, this machine is translating text in English first then continue to the target language with a few exceptions including Catalan-Spanish. In November 2016, this machine changes into neural machine translation that is translate whole sentences at a time. Since May 2022, this machine used in all 133 languages. Just as April 2016, the total user is over 500 million with more than 100 billion words translated daily.

The accuracy of the result form this MT can be vary. It will be beneficial for the person that need a fast translation process with the minimum result of accurate, acceptable, and readable because this MT does not have competencies of translator (Purwaningsih, 2016; Tsai, 2019). The result of translating discussion text, exposition text, and narrative text by using GT shows that the accurate happening is 49.1 % and 37.1 %. It means that in the inaccuracy indicators like omission, addition, different meaning and zero meaning can be found in most sentences (Nadhianti, 2016; Stapleton & Kin, 2019). When GT is used for translating 80 English sentences into *Bahasa Indonesia*, the accuracy by using a table adapted from Memsource criteria is only valid with words and phrases and the accurately translating 60.37% (Sutrisno, 2020).

Further, the use of GT for the students at Universitas Negeri Jakarta as an alternative tool for doing references book translation is favorable (Herlina et al., 2019). The result of translation using GT for procedural text is far from perfect (Halimah, 2018). There are 25 errors in Kompas.com news text on grammar and selection of inappropriate words text from *Bahasa Indonesia* into English (Rahmania & Triyono, 2019). Google translate failed to recognize idiomatic expressions, and illogical sentence at the target text caused by cultural difference (Amilia & Yuwono, 2020). Google translate has errors on morphology, and syntactical aspects on translating narrative text (Anggaira, 2017; Jannah et al., 2022). From the result show that the accuracy of GT for translating English Text into *Bahasa Indonesia* can be positive and negative sides.

As a tool for translating, GT has its limitation in translating text from English into *Bahasa Indonesia*. This limitation can be from linguistics like grammar and lexical, the number of words, kinds of language to be translated, and the accuracy. This limitation lead to the errors on linguistics and meaning. It makes the result can be good for European languages but not for Asian languages (Aiken & Balan, 2011). The classification of error for machine translation are orthography, lexis, grammar, semantic, and discourse (Costa et al., 2015).

Orthography level is related to punctuation, capitalization, and spelling. The improper use of punctuation means a punctuation error. Capitalization refers to the use of capital letters. For instance, the use of the first letter in the beginning of a sentence. A spelling error concerns the substitution, addition, or deletion of one or more letters to the orthography of a word. Lexis errors contain following types of error: omission, addition, and untranslated. Omission and addition errors analyzed regarding the type of words like: a. content words (lexical words) that is words that carry the meaning of a sentence such as noun (John, room), adjectives (pleasant, ancient), b. function words (grammatical words) that is words that serve to express grammatical relationship with other words in a sentence. For example: preposition (from, at), pronouns (she, I, anybody).

This article is licensed under CC Attribution-NonCommercial 4.0
DOI: 10.47766/jetlee.v2i2.498

When the translation of a word existing in the source text is disappeared in the target text, it called omission. While addition errors show the opposite aspect which is the target text present the added words that was not stated in the source text. All examples provided below in English is abbreviated as EN, while *Bahasa Indonesia* is abbreviated as IDN.

1. Omission error (content words)
EN: **There are** many people….
IDN: **Ada** banyak orang….
2. Omission (function words)
EN:In many countries **the legal system**…..
IDN: Di banyak negara **sistem hukum**

In the first example two English words translated into one word in *Bahasa Indonesia*. In the second example the English words "the" is not translated into *Bahasa Indonesia*.

3. Addition error (content words)
EN: **In** 2013,….
IDN: **Pada tahun** 2013,….
4. Addition (function words)
EN: Parents **of** young children
IDN: Orang tua **dari** anak kecil
5. Untranslated
EN: non-organic
IDN: non-organik

In the third example the two English words translated into three words *Bahasa Indonesia*. In the fourth example there is addition of the word "dari" in *Bahasa Indonesia*. In the fifth example he English words translated into the same words and the only change the alphabet c into k.

Grammar errors appear in two types such as misselection and misordering. Misselection errors arise at word class level for example the source text is adjective then the target text change into noun, verbal level (tense and person), errors of agreement (gender, number, person), and in contraction between preposition and articles. It can be more than one of these problems called blend.

1. The example: Misselection : word class error
EN: both….
IDN: keduanya…..
In the example show the change of word class from adjective into pronoun.
2. The example:Misselection : tense
EN: A farmer in Kenya participated…..
IDN: Seorang petani berpartisipasi……
In the example show the tense from past tense into present tense.

Misordering errors are concerning with syntactical problems that the sentences may demonstrate. The good result of translation is not only related to form use in the right context but also to arrange them in the right order.

EN: the world's population
IDN: populasi dunia

This article is licensed under CC Attribution-NonCommercial 4.0 DOI: 10.47766/jetlee.v2i2.498

In the example the change of form use in English into *Bahasa Indonesia*.

Semantic errors classify into confusion of senses, wrong choice, collocational errors, and idioms. Confusion of senses happen when a word in source text translated into something representing one of its possible meanings in target text, but it is not correct.

EN: judges and magistrates
IDN: hakim dan hakim

Wrong choice means that the source word translated into the wrong target word without any apparent relation.

EN: the most cycling event
IDN: acara bersepeda terbanyak

Collocational error means the wrong natural combination of words. Idioms is the opposite of collocation The example of wrong collocational errors are seriously, badly, horribly. The examples of the wrong idioms are I am care less, by purpose, one in the same. Discourse level errors classified into style, variety, and should not be translated. Style errors is a bad stylistic choice of words when translating a sentence like a repetition of a word in a near context. Variety errors happen when the result of translation is not representing the target language. Should not be translated means the source text don't need any translation in a target text for example the title of book or film.

For example:
EN: malware
IDN: perangkat lunak perusak.

The errors type can be seen on the figure below:

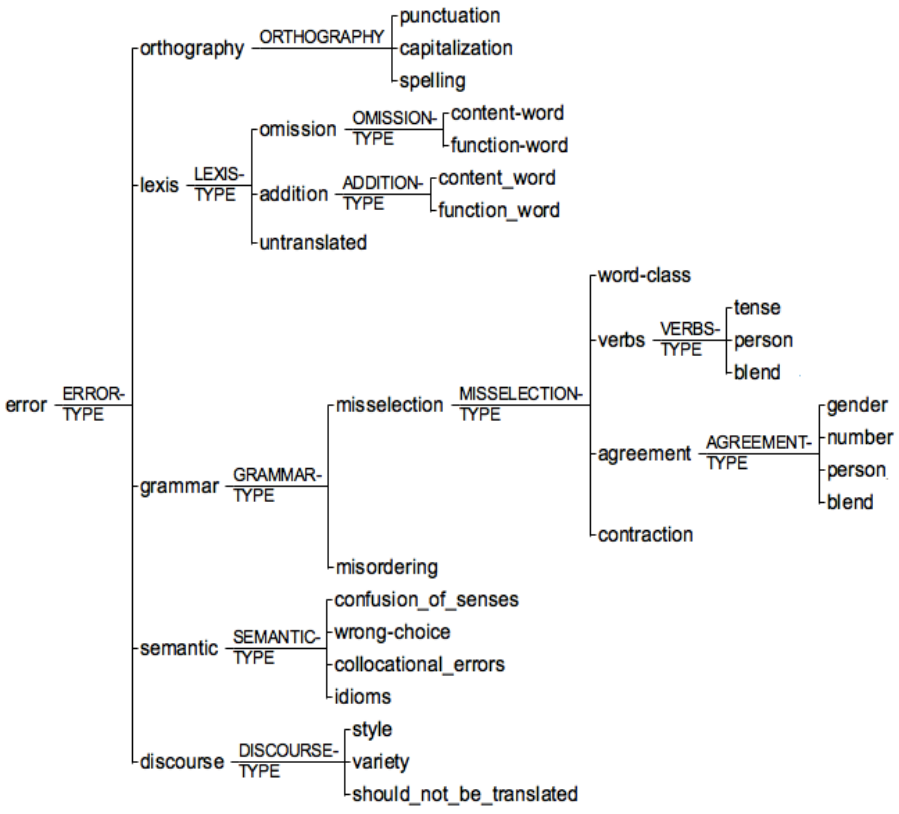

Figure 1. The Taxonomy of Errors (Costa et al., 2015)

This article is licensed under CC Attribution-NonCommercial 4.0 DOI: 10.47766/jetlee.v2i2.498

METHODS

Error analysis research method refers to language studies concerning on the linguistic errors made by second language learners. This kind of research aimed to study about linguistic deviation makes by second language learners. Error analysis is usually planned to recognize the kinds of learners' errors in second language learners. The procedure of the error analysis includes the following four steps: First is data collection, usually from students' compositions. Second is identification of errors through the process of coding. Third is classification of errors into error types which might be grouped in accordance with the grammatical area (e.g. Errors in articles, errors in nouns, errors in verbs, etc.), the causing factors (L1 interference, overgeneralization, transfer of training, or communication strategy), or the characteristics of errors (global vs. local). Fourth is a statement of error frequency (Huang in Artowardoyo, 2018).

The method of this research is error analysis. This kind of research aimed to find the errors made by the second language learner. The procedure that the writer adapted from Huang in Artowardoyo (2018) are data collection, classification or errors types, and a statement of error frequency. First is data collection that is collected from the result of students' translation by using GT. The type of chosen text is informative text with selected topics: legal system, information technology, organic food, end of mining boom in Australia, the human heart, languages, Tour de France, and Jeans: the history of great cloth. Every text consists of 400 to 500 words. Second is identification of errors through the process of coding. In this step the writer identify errors in every text based on the lexis, grammatical, semantic, and discourse level. Third is classification of errors into error types which might be grouped in line with the lexis, grammatical (mis selection: word class, verbs, agreement, contraction), semantic (confusion of senses, wrong choices, collocational errors, idioms), discourse i.e. style, variety, and should not be translated (Costa et al., 2015). Fourth is a statement of error frequency applied to make a conclusion about the frequency of errors made by students with GT. The last is a statement of error frequency. It is done by calculating all the errors of every category.

RESULTS AND DISCUSSION

This research was done through error analysis procedure: data collecting, identification of errors through the process of coding, classification or errors, and a statement about the frequency of errors. Data collecting from the result of the final test semester of Translation subject. There are 62 students that was submitted the translation result as the subject of the research. The selected topic was informative text entitled legal system, protection vs evolving malware, organic food, end of mining boom in Australia, the human heart, languages, despite high-tech, Japan sticks to the fax machine, *Tour de France*, and jeans: the history of great cloth. The length of the text was 382- 448 words. The students had to choose one title from the selected texts. Then, they had to translate the text from English into Bahasa Indonesia. They can apply machine translation to get the result of translation. Most of them use GT to help them translating the text given. They can use any references and dictionary from internet as well. They were free to choose any dictionary and articles to support the result of translation accurate, readable, and understandable.

The second step is identification of error through the process of coding. There are four types of errors includes lexis, grammar semantic, and discourse. The identification done by analyze the result of translation for all selected topic. Below is the result of identification of errors that is taken from the target language. Errors in lexis consist of omission, addition, and untranslated. There were 576 errors appear in all selected text. The analysis was done by grouping the three types of errors in every text through a table. The errors were

This article is licensed under CC Attribution-NonCommercial 4.0
DOI: 10.47766/jetlee.v2i2.498

symbolized with different color: yellow for omission, green for adding, blue for untranslated. The example result of analysis is as follows:

No	Source text	Target text
1.	Legal system In many countries **the** legal systems **are** very complex. This is because they have been developed over very long times. Every time **that a** new case is decided it can have **an** effect on future cases. **This** means **that** the people who work **in** legal situations have **to be** clever. They have **to** know and understand all **the** rules and precedents. **A** precedent **is** when **a** previous decision will have **an** effect on future cases as it was related to an important point.	Sistem hukum Di banyak negara, sistem hukumnya sangat kompleks. Hal ini karena mereka telah dikembangkan dalam waktu yang sangat lama. Setiap kali kasus baru diputuskan, hal itu dapat berdampak pada kasus-kasus di masa depan. Artinya orang yang bekerja di bidang hukum harus pandai. Mereka harus mengetahui dan memahami semua aturan dan preseden. Preseden adalah ketika keputusan sebelumnya akan berdampak pada kasus-kasus di masa depan karena itu terkait dengan poin penting.
3.	Protection vs Evolving Malware **In 2013**, **the** world's population **of** humans hit 7 billion. In **the** same year, **the** number **of** mobile phones in **the** world was 6.8 billion. **This** means **that there were that** many entry points to the internet. Accessibility to **mobile phones** has helped the farming and agricultural community **all over the** world.	Perlindungan vs. Malware yang Berkembang Pada tahun 2013, populasi manusia di dunia mencapai 7 miliar. Pada tahun yang sama, jumlah ponsel di dunia adalah 6,8 miliar. Artinya, ada banyak titik masuk ke internet. Aksesibilitas ke ponsel telah membantu komunitas pertanian dan pertanian di seluruh dunia.

Table 1. The example results of errors analysis in lexis level

From the table shows that the words the, of, are, that, a, an, are not translated. It is because in *Bahasa Indonesia* does not need article (a, an), or determiner (the) to form singular and plural nouns form in a sentence but in English both of those elements are requirement to say about nouns. Then, the words mobile phones, allover are translated into *ponsel*, and *di seluruh*. In this form, there is an omission from English words mobile phones, and all over the world into *ponsel* and *di seluruh dunia*. The total errors in the legal system text are omission 36, adding 0, untranslated 0. The total errors in the protection vs evolving malware text are omission (content 6, form 42), adding (content 3, form 0), untranslated 0. The total errors in the end of mining boom in western Australia text are omission (content 11, form 41), adding (content 2, form 0), untranslated 0. The total errors in the languages text are omission (content 7, form 31), adding (content 11, form 0), untranslated 0. The total errors in the organic foods text are omission (content 21, form 10), adding (content 12, form 5), untranslated 3. The total errors in despite high-tech Japan sticks to the fax machine text are omission (content 7, form 41), adding (content 1, form 0), untranslated 0. The total errors in the human heart text are omission (content 22, form 67), adding (content 2, form 0), untranslated 0. The total errors in the tour de France text are omission (content 27, form 62), adding (content 8, form 0), untranslated 0. The total errors in the jeans: the history of great clothes text is omission (content 30, form 38), adding (content 10, form 0), untranslated 0.

The two types of errors in grammar are misselection and mis ordering. Misselection consists of four sub types: word class, verbs, agreement, and contraction. While mis ordering means syntactical problems that may happen in sentence or made a sentence in the right order. There in not error in mis ordering because the result of GT in *Bahasa Indonesia* only changes some parts of a sentence such as adjective and noun. In English, adjective come first then noun but in Bahasa Indonesia noun come first then adjective. For example: in the legal system translated into *sistem hukum*, the organic food translated into *makanan organik*. The example result of errors in grammar shows in the table.

This article is licensed under CC Attribution-NonCommercial 4.0 DOI: 10.47766/jetlee.v2i2.498

No	Title	Misselection			Misordering	
		Wordclass	Tense	Agreement	contraction	
1.	Organic food	both = keduanya			-	-
2.	Protection vs evolving malware		A farmer **participated**= Seorang petani **berpartisipasi**		-	-
3.	Jeans			students = siswa schools = sekolah (number)	-	-
4.				artist= artis (person)	-	-

Table 2. The example result of grammar errors

In table shows that there are misselection of the meaning like the meaning both is *keduanya.* There is a change of meaning from adjectives into pronoun. In the second example shows the change of meaning from past tense into present tense. In agreement, there are two examples of number categorization from students into *siswa* and schools into *sekolah.* In English, there is a must to form plural version of noun by using suffix -s or -es but it doesn't need in *Bahasa Indonesia*. In gender categorization, there is an example of meaning from artist in English into artis in *Bahasa Indonesia*. It means that there is no specific person shown in the meaning in *Bahasa Indonesia*. It can be all performers in arts such as actor, painter, or singer. But in English there is different definition for artist that is a person who practices any of the various arts such as a sculptor, novelist, poet or filmmaker and actor means a person whose profession is acting on the stage, in movies, or television.

Errors in semantic comes in four types: confusion of senses, wrong choice, collocational error, and idioms. The example provides in the table.

No	Title	Confusion of senses	Wrong choice	Collocational error	Idioms
1.	Legal system	judges and magistrates= hakim dan hakim		-	-
2.	Tour de France		The most cycling event=acara bersepeda terbanyak	-	-

Table 3. The example errors on semantic

There is one example from the confusion of senses error in the legal system text that is judges and magistrates translated into *hakim* and *hakim*. The correct meaning should be *juri* dan *hakim*. It happened because in Indonesia there is not judges (*juri*) in a court to make a decision determined by a team of magistrates. Thus, the meaning of judges translated into *hakim*. There is one example of wrong choice error. The meaning of the most cycling translated into *acara bersepeda terbanyak.* The correct meaning should be *acara bersepeda terbesar*.

This article is licensed under CC Attribution-NonCommercial 4.0 DOI: 10.47766/jetlee.v2i2.498

The errors in discourse consist of style, variety, and should not be translated. The example of style is in the legal system text. The words "a solicitor is a lawyer…" translated into "*seorang pengacara adalah pengacara*". There is repetition meaning on the different words. Next example about variety error in protection vs evolving malware text. The word "….led to an increase of population" translated into "*menuju peningkatan populasi* ". There is no translation for article "an" in *Bahasa Indonesia* but there should be an article in English grammar. Then, the example of should not be translated error is the word "organic" translated into "*organik*", the word "malware" translated into "malware" in Bahasa Indonesia. The examples provide in the table.

No	Title	Style	Variety	Should not be translated
1.	Legal system	A solicitor is a lawyer= seorang pengacara adalah pengacara		
2.	End of mining boom		…led to an increase in population = menuju peningkatan populasi	
3.	Organic food			organic=organik
4.	Protection vs evolving malware			Malware=malware

Table 4. The example result of discourse error

The third step was done through grouping the errors in every text into lexis, grammar, semantic, and discourse. There are 9 titles with 3656 words, and 229 sentences. The total errors for all categories is 588. There are 576 errors in lexis, 4 errors in grammar, 5 errors in semantic, and 3 errors in discourse.

No.	Errors level	Numbers	Percentage
1.	Lexis	576	97.9%
2.	Grammar	4	0.007%
3.	Semantic	5	0.008 %
4.	Discourse	3	0.005%

Table 5. The total frequency of errors in lexis, grammar, semantic, and discourse level

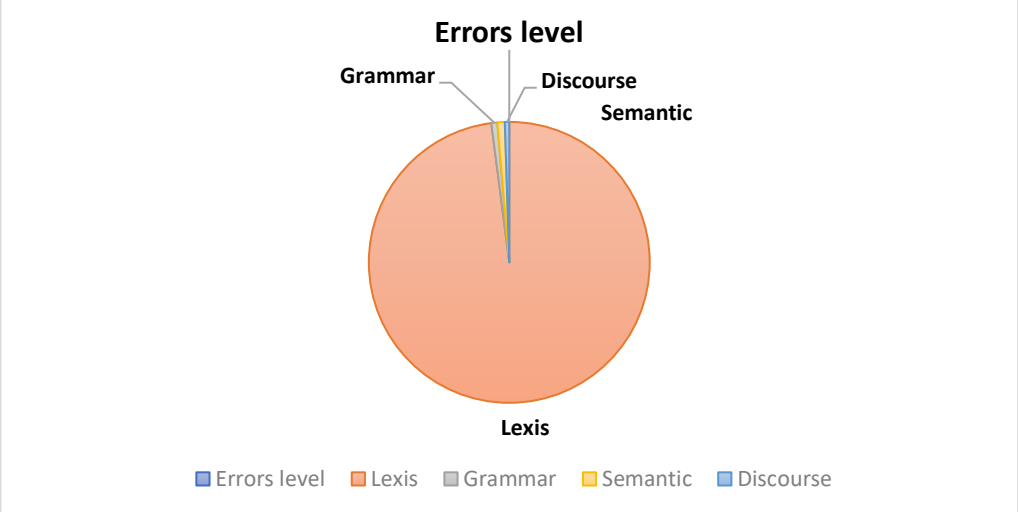

Figure 2. The frequency of errors for all level

This article is licensed under CC Attribution-NonCommercial 4.0
DOI: 10.47766/jetlee.v2i2.498

The last step was a statement of errors frequency. In this part, there are 97.9 % errors in lexis level, 0,007 % errors in grammar level, 0.008 % errors in semantic level, and 0.005 % errors in discourse level. It means that the most errors are in the level of lexis or lexical element in a language. While the other level such as grammar, semantic, and discourse is only a few elements. The frequency or errors provide in the figure.

CONCLUSION

In the rapid changing of technology, the use of GT is crucial to get the quick meaning of text. The use of MT especially GT help the user to understand the meaning of source text easily. In line with the information on errors in translation result from GT it occurs that there is limitation of words for translating, many errors in lexis, grammar, semantic, and discourse. These aspects draw to the result that GT had disadvantages for translating. It also need adjustment to get readable result of translation. The user has to pay attention more on basic element of language, choice of words, and exact meaning of words. The percentage of errors show that mainly errors appear on lexis 98 %, grammar 0.007%, semantic 0.008%, and discourse 0.005%. Accordingly, the user has to find the exact meaning from English into Bahasa Indonesia. The basic element of language as the most errors demonstrate the usage of GT as an alternative way to translate a text quickly and understand a text universally.

REFERENCES

Abbasian, R., & Biria, R. (2017). English Language Textbooks in EFL Education: Do Improve Students' National, International and Target Culture Familiarity?.

Aiken, M., & Balan, S. (2011). An analysis of Google Translate accuracy. *Translation journal*, *16*(2), 1-3.

Al-Musawi, N. M. (2014). Strategic use of translation in learning English as a Foreign Language (EFL) among Bahrain university students. *Comprehensive Psychology*, *3*, 10-03.

Amilia, I. K., & Yuwono, D. E. (2020). A STUDY OF THE TRANSLATION OF GOOGLE TRANSLATE. *Lingua: Jurnal Ilmiah*, *16*(2), 1-21.

Anggaira, A. S. (2017). Linguistic Errors on Narrative Text Translation Using Google Translate. *Pedagogy: Journal of English Language Teaching*, *5*(1), 1-14.

Atmowardoyo, H. (2018). Research methods in TEFL studies: Descriptive research, case study, error analysis, and R & D. *Journal of Language Teaching and Research*, *9*(1), 197-204.

Chandra, S. O., & Yuyun, I. (2018). The use of google translate in EFL essay writing. *LLT Journal: A Journal on Language and Language Teaching*, *21*(2), 228-238.

Costa, Â., Ling, W., Luís, T., Correia, R., & Coheur, L. (2015). A linguistically motivated taxonomy for Machine Translation error analysis. *Machine Translation*, *29*(2), 127-161.

Halimah, H. (2018). ERROR ANALYSIS IN ENGLISH-INDONESIAN MACHINE TRANSLATE. In *PROSIDING SEMINAR NASIONAL & INTERNASIONAL* (Vol. 1, No. 1).

Herlina, N., Dewanti, R., & Lustiyantie, N. (2019). Google translate as an alternative tool for assisting students in doing translation: A case study at Universitas Negeri Jakarta, Indonesia. *BAHTERA: Jurnal Pendidikan Bahasa Dan Sastra*, *18*(1), 70-78.

Imre, Attila (2012). Communication through translation. *Proceeding of the Communication, Context, Interdiciplinarity*, 2. Editura.

This article is licensed under CC Attribution-NonCommercial 4.0 DOI: 10.47766/jetlee.v2i2.498

Jannah, I. R., Yenni, C., Rahmi, T., & Alimnur, A. (2022). An Error Analysis of Students' Translation In Narrative Text. *JETLEE: Journal of English Language Teaching, Linguistics, and Literature*, *2*(1).

Nadhianti, M. (2016). An analysis of accuracy level of google translate in English-bahasa Indonesia and bahasa Indonesia-English translations. *Sastra Inggris-Quill*, *5*(4).

Purwaningsih, D. (2016). Comparing Translation Produced by Google Translation Tool to Translation Produced by Translator. *Journal of English Language Studies*, *1*(1).

Rahmannia, M., & Triyono, S. (2019). A Study of Google Translate Translations: An Error Analysis of Indonesian-to-English Texts. *International Journal of Linguistics, Literature and Translation (IJLLT)*, *2*(3), 196-200.

Schäffner, C. (2005). Preparing students of translation for the real world: Needs, methods, constraints. In *On the relationships between translation theory and translation practice* (pp. 237-248). Peter Lang publishing group.

Stapleton, P., & Kin, B. L. K. (2019). Assessing the accuracy and teachers' impressions of Google Translate: A study of primary L2 writers in Hong Kong. *English for Specific Purposes*, *56*, 18-34.

Sutrisno, A. (2020). The Accuracy and Shortcomings of Google Translate Translating English Sentences to Indonesia. *Education Quarterly Reviews*, *3*(4).

Tsai, S. C. (2019). Using google translate in EFL drafts: a preliminary investigation. *Computer Assisted Language Learning*, *32*(5-6), 510-526.

Copyright: @2022 The Author(s).
Published by *JETLEE* (Journal of English Language Teaching, Linguistics, and Literature).

This article is licensed under CC Attribution-NonCommercial 4.0 DOI: 10.47766/jetlee.v2i2.498