Common Lexical Error

by Common Lexical Error Common Lexical Error

COMMON LEXICAL ERRORS MADE BY MACHINE TRANSLATION ON CULTURAL TEXT

Nanda Fitri Mar'athus Sholikhah[1], Rohmani Nur Indah[2]

[1]Department of English Language Education, Faculty of Education and Teacher Training, State Islamic Institute of Kediri

nandafitrims@gmail.com

[2]Department of English Literature, Faculty of Humanities, Universitas Islam Negeri Maulana Malik Ibrahim Malang

indah@bsi.uin-malang.ac.id

ABSTRACT

Machine translation is one tool of Google that presents various languages to translate. As a translator machine, the results of Google Translate are not always perfectly correct which is still needed to be revised. Arok Dedes story is one of the Javanese stories that contain elements of culture. Translating texts which contain elements of a culture is not easy because one region to another have different cultures, so that it is difficult to look for parallel words that contain elements of culture. This study is aimed at two main purposes: (1) finding out the types of lexical errors made by machine translation in translating cultural text and (2) knowing the most dominant type of lexical errors made by machine translation in translating cultural text. This study was carried out in a population of 553 pages of Arok Dedes story. A simple random sampling technique was done to select samples. The study results are that there are only 9 types of the total 21 types of lexical errors, namely calque, misselection, consonant-based type, false friend, vowel-based type, inappropriate co-hyponym statistically weighted preferences, semantically determined word selection, and preposition partners. The most dominant error of lexical errors is calque.

Keywords: *Lexical Errors; Machine Translation; Cultural Text*

Internet technology today allows everyone to access information from all over the world anytime and anywhere. One of the tools that help internet users find information effectively is Google. Google has a wide variety of applications and features that its users can take advantage of. In education, Google is one of the media that is often visited by students and teachers. Google developers recognize that the information presented on web pages on the internet can be in multiple languages. Therefore, many web visitors use machine translation assistance to help them translate from one language to another.

Various types of machine translation can be accessed through Google, such as Google Translate, Yandex Translate, Translate.com, and Bing Microsoft Translator (Kumparan.com Tekno & Sains,

2020). The four machine translations are the most frequently visited by Google users in Indonesia. Google Translate is a translation service that has been successfully recognized by the world. This translation, which is shaded by Google, has a database of up to 103 languages globally, including regional languages in Indonesia such as Javanese and Sundanese (Kumparan.com Tekno & Sains, 2020).

As a machine translator, the results from Google Translate are not entirely correct. The translation is generally defined as transferring a message from the first language to the target language (Amini & Bayesteh, 2020). The translation process, according to Brown, includes several steps: analyzing the structure of the text, how to transfer a language into the target language, and review the result of translation (in Bojar, 2011). The translation is a complex process. Translators translate a vocabulary or sentence and pay attention to context. Translation errors can be caused by a misunderstanding of the translation results that do not accurately translate the meaning so that the translated text is not structured contextually. It happens due to the incorrect choice of words in a sentence.

The results of this translation can be called a pre-translation that still needs to be revised (House, 2015). Translations from machine translators need to be corrected to identify language errors because many machine translators doubt their quality. Brown (in Wijayanto, 2020) continues that although errors cannot be corrected by themselves, errors can be observed, analyzed, and classified, and this process is called error analysis. Likewise, Corder (in Jamilah, 2012) defined error analysis as the main process in getting second language learning where students face or make mistakes. Analytical error theory applies not only to language learning learners but also to errors found in machine translation. Error analysis is the identification and classification of individual errors in the use of machine translation. It helps evaluate the target language produced by machine-assisted translation (Keshavarz, 1999). This theory can be applied in finding and observing machine translation errors, which are then edited by human translators.

This article analyzes the errors translated by the machine translator, namely Google. Error analysis is a technique for identifying, classifying, and interpreting errors systematically made by students who are learning the language by using linguistic-based theories and procedures (Pateda, 2001). In this article, the authors only focus on analyzing lexical errors in cultural texts. In this case, the lexical error is the improper use of lexical items in a certain context due to confusion between two words (Llach, 2005), due to formal or semantic similarities of L1 or L2 influences (Maheswari et al., 2020). In general, lexical errors will only affect lexical words, whereas grammatical errors will only affect grammatical words (Hemchua & Schmitt, 2006).

This research's data source was taken from the dialogue containing the cultural text in the story of Arok Dedes by Pramoedya Ananta Toer. Cultural text is a unique language. Google cannot translate cultural text terms easily. Cultural texts can contain terms in local languages, for example, Javanese. Cultural texts are objects, actions, and behaviors that express Hoed's cultural meaning (Hoed, 2006). One of the books that have many cultural texts in it is the storybook Arok Dedes. This book is one of the books that tells about Javanese stories and contains cultural words. Besides, one of the sources used in this book is also from the Pararaton book, which is a book that contains the story of Javanese kings. The authors chose Pramoedya Ananta Toer's work because he is considered one of the most prolific writers in Indonesian literature history. Pramoedya has produced more than 50 works and translated into more than 41 foreign languages.

Translating text containing elements of a culture is not easy because in one area to another region has a different culture (Supendi, 2017). It is not easy to find parallel words that contain elements of culture,

religion, social, customs, social organization, procedures, sign language, and ecology. Machine translators cannot translate cultural terms easily. Examples of cultural text in Javanese stories are terms of an object, action, and behavior or technical terms that express cultural meanings or are written in pure regional languages.

The purpose of this study is to identify the types of lexical errors made by machine translators in translating cultural texts in Pramoedya Ananta Toer's Arok Dedes dialogue and to determine the most dominant types of lexical errors. This study can provide useful information on how machine translation can be used effectively when translating various texts, such as texts containing cultural elements. This study's findings can clarify the types of lexical errors found in the Arok Dedes dialogue created by Google Translate. Furthermore, the results can be used as a reference in the future to train machine translation users to be more careful and have to draw on re-examining the results of their translations so that they do not necessarily adopt or take the translation results. It will help them get a fast translation even if they still have to make some edits and retouch the translation. Thus, this study's findings are expected to provide guidelines for various learning fields or language skills and those interested in translation.

RESEARCH METHOD

This section discusses the research design, data sources, research instruments, data collection, and data analysis. This study used a descriptive analysis method to describe the lexical error analysis results in Arok Dedes' story dialogue. Researchers also use qualitative descriptive because it tries to describe the errors found in the object. This study discusses machine translators' lexical errors in translating Arok Dedes' story dialogue, which contains cultural texts.

The data collected is in the form of words and sentences. This study describes the types of lexical errors in translating machines in translating dialogue in Arok Dedes' stories. In this research, the data

source is a book entitled "Arok Dedes." This book was written by Pramoedya Ananta Toer, who is one of the most prolific writers in the history of Indonesian literature. Pramoedya has produced numorous works and translated into several foreign languages.

The instrument used in qualitative research is a human instrument. Accordingly, this study used a human instrument, namely the researchers, who act as an instrument to collect and analyze the data. Moreover, this research uses cultural texts that have been translated by machine translators to collect the required data. Data collection techniques also cover tracing relevant informations from books, the internet, journals, articles, and others. According to Sudaryanto, there are five data collection strategies in linguistic research: recording techniques, recording techniques, separating data techniques, data transfer techniques, and replacing techniques (Sudaryanto, 1992). In this study, the authors only used three data collection strategies: 1) recording techniques. This technique is used to collect data by recording it using a notebook. 2) Separation technique is a strategy to separate data from other data to find similarities and the distribution between them: 3) Transfer technique, namely transferring data to other media.

The data collection procedures of this study are as follows: first, the authors read the entire book Arok Dedes. Second, the writers gave a sign to the text to be analyzed, namely the cultural text contained in Arok Dedes. Third, the writers listed the text that has been selected, which is a dialogue that contains cultural texts. Fourth, the writers translated the text into a translation machine, namely Google, Translate. Fifth, the authors rewrote the cultural text with the machine translator's translation to a new page as the first data. Sixth, the authors identified the machine translators' lexical errors in translating Arok Dedes' story dialogue. Finally, the authors conducted data checking with the experts so that the data are valid.

In this research, the descriptive qualitative method employed focuses on words, phrases, and sentences rather than

numbers. There are several procedures in analyzing data based on Keshavarz and Brown (in Hana Amanah, 2017), who explain that error analysis refers to collecting samples, identifying errors and classifying them into which category they belong to, and finally evaluating the errors. Carl has the same idea about the error analysis framework. He agrees that the first step involves selecting the target language, followed by identifying the errors made (Carl, 1998). Furthermore, errors are classified and explained.

The steps are as follows: first, identify the errors found in the object. The second is to classify errors based on each scope. After the authors identified lexical errors, they classified the errors based on the error category. The third is explaining each of the mistakes. The fourth is calculating the percentage of error. At this stage, the authors used tabulation to describe the frequency and percentage of errors. By calculating the frequency of errors per item, the most frequent and least frequent errors are identified. Furthermore, the authors compared the errors that occur between the two translation machines. The last is to conclude. In this study, the authors drew conclusion and also provided some suggestions.

Besides, the authors used triangulation to obtain valid data (Arifin, 2011). This study used theoretical triangulation because the authors realized that their knowledge was limited. The authors used James's theory and supports it. The authors also used the theory of Kezhavars, which holds the same thinking as James, to examine the results of the identified errors. The authors used more than one theoretical position in interpreting the data. Surface structure theory is based on the taxonomy of lexical errors. Overall, the current classification of lexical errors falls into two main categories: formal and semantic errors summarized in figure 1.

A. Formal errors
  1. Formal Misselection
    a. Suffix type
    b. Prefixing type
    c. Vowel-based type
    d. Consonant-based type
    e. False type
  2. Misformations
    a. Borrowing
    b. Coinage
    c. Calque
  3. Distortion
    a. Ommision
    b. Overinclusion
    c. Misselection
    d. Misordering
    e. Blending
B. Semantic errors
  1. Confusion of sense relation
    a. A superonym for a hyponym
    b. A hyponym for a superonym
    c. Inappropriat co-hyponyms
    d. Wrong near synonyms
  2. Collocation errors
    a. Semantically determined word selection
    b. Statistically weighted preferences
    c. Arbitrary combination
    d. Preposition partners

Figure 1. Summary of Lexical Error Taxonomy

RESEARCH FINDINGS

This section discusses research findings and discusses data analysis on lexical errors made by Google Translate in translating cultural texts. The authors will explain and describe the research findings in Google Translate in translating the cultural text in the dialogue in the Ken Arok Ken Dedes story written by Pramoedya Ananta Toer. These study's research findings were verbal, but the authors turned them into numerical by calculating the number, percentage, and frequency of each error category. Because

this research is a qualitative descriptive study, the data results will be described qualitatively by presenting the frequency and percentage of errors. The frequency and percentage of each error category are the results of calculations using certain formulas that have been mentioned in the previous chapter.

The Types and The Dominant Type of Lexical Errosmade by Google Translate in Translating Cultural Text

As shown in table 1, there are thirty-four lexical errors found from fourteen pages in Arok Dedes' story dialogue that contains cultural texts. Some errors are common (for example, calque, statistically weighted preferences, and semantically determined word selection). Others are relatively rare (e.g., false type, misselection, consonant based type, vowel based type, inappropriate co-hyponyms, and preposition partners). There were no errors in any other categories (for example, suffix types, prefixing types, borrowing, coinage, omission, overinclusion, misordering, blending, a supermom for hyponym, hyponym for superonym, wrong near-synonym, and arbitrary).

Table 1. The Frequency and Percentage of Lexical Errors

Error Types	Number of Errors (34)	Percent (%)
A. Formal errors		
1. Formal Misselection		
a. Vowel-based type	1	3.00%
b. Consonant-based type	1	3.00%
c. False type		
2. Misformations	5	15.00%
a. Calque		
3. Distortion	9	27.00%
a. Misselection		
B. Semantic errors	4	12.00%
1. Confusion of sense relation		
a. Inappropriate co hyponyms	1	3.00%
2. Collocation errors		
a. Semantically determined word selection	6	18.00%
b. Statistically weighted preferences	6	18.00%
c. Preposition partners	1	3.00%

Based on the table above, the authors provide further explanation as follows:

a. Formal Errors

In formal errors, the most problematic error category in the data is a calque (there are 27.00% of all errors), followed by false type (15.00%), then misselection (12.00%). The errors that rarely appear in formal errors are consonant based type and vowel based type, only 3.00% of the total. These findings indicate that the Indonesian language, especially the word structure of culture, is quite influential in making English

sentences. The five formal errors that arise in this study will be discussed below by providing some examples and explanations of how they occur and how they should occur.

1) Calque

Calque is a literal translation error in L1. The following are some of the calque errors that appeared in Arok Dedes' dialogue:

Source Text: Dan hancur kau dan kalian di Sanggarana sana

Google Translate: And destroyed you and you in Sanngarana there

The translation above indicates that the two sentences in English above are the result of the translation of L1 (Indonesian), and in this case, the two sentences are not quite right.

2) False Friend

A false friend can be caused by divergent polysemic, overlapping partial semantics, or loanwords taken from English words and sometimes have overlapping meanings. This is an example of a false friend.:

Source Text: Inilah sahaya, ya, Durga

Google Translate: This is Sarah, yes, Durga

The use of the word 'Sarah' to translate the word 'sahaya' is not correct. Obviously, the two have different meanings, sahaya means servant, or it can be said to be 'I,' so if the word is translated as 'Sarah,' it is incorrect because the meaning is different. The suggested translation is "This is me, Durga".

3) Misselection

Actually, in this case, the words are not in L2. However, those errors result from incorrect implementation of the target language without glitches or misspellings in L1. Here is an example:

Source Text: Semua yang jahat berasal dari orang-orang Syiwa yang memuliakan kama tanpa batas itu.

Google Translate: All evil comes from Shaiwa, who glorify the infinite kama.

In this case, there was some confusion to choose the word <Syiva>. The use of the word * shaiwa in this sentence is wrong because it is not in L2.

4) Consonant Based Type

This error case is almost the same as the vowel based type case. In this case, they are almost the same shape but have different consonants. Let's take an example:

Source Text: Tak pernah yang mulia melakukan wadad kecuali hanya untukmu

Google Translate: Never a noble do wadhad except only for you

In L2, the word 'wadhad' is not available. The correct one is 'wadad' because this word has its own meaning. Between 'wadhad' and 'wadad' have almost the same form but different consonants.

5) Vowel Based Type

The vowel based type mistake is that they have almost the same shape but have different vowels. The example is:

Source Text: Mereka mencoba untuk melawan pria itu dengan pisau

Google Translate: They try to go against the grain to that men with knife.

'Man' and 'men' have the same meaning but differ in their spelling and usage. The first uses the vowel 'a'. The second uses the vowel 'e.' The first is singular, and the second is plural. The use of the word 'men' in the sentence above is not correct because it must be singular, not plural. That is right, by using the word 'man.'.

b. Semantic Errors

Of the two main categories of semantic errors, collocation errors are the most frequently found in the Arok Dedes story translated by Google Translate, with a total of 38.00%, while confusion of sense relations is in second place, 3.00% of the total. This shows that there is a big difference between the confusion of sense relation and collocation errors. Collocation errors occur in more than half of these cases. Of the eight categories of semantic errors, only four errors appeared in the Arok Dedes dialog, which was translated by Google Translate. The four mistakes will be discussed below.

1) Statistically Weighted Preferences

Statistically weighted preferences are words or phrases often used in conjunction with other words or phrases and sound natural and appropriate for native speakers. An incorrect phrase may not be completely wrong; it is imprecise.

There are six statistically weighted preferences found from thirty-four samples. We will discuss one of them. Here is a mistake:

Source Text: Bukankah Yang Mulia Akuwu sudah cukup memuliakan kau, Dedes? Pramesywari Tumapel? Telah mengangkat naik kau dalam perkawinan kebesaran ini?

Google Translate: Is not your Honor Akuwu enough to glorify you, Dedes? Pramesywari Tumapel? Has raised you in this marriage of greatness?

In the case of these statistically weighted preferences, the phrase 'your Honor' is the mistake. The word 'honor' is not appropriate to attach to the word Akuwu, a person who can glorify someone. This word is inaccurate because it is not an appropriate adjective or combination for 'Akuwu.' Even though we change the speech part of the word 'honor' to an adjective, it still does not make the phrase correct. The suggested translation is 'glorious.'

2) Semantically Determined Word Selection

This is an example of an error found in the Arok Dedes dialog:

Source Text:
a) *Apakah Hyang Wisynu menitahkan agar kalian memelihara elang itu dengan tekun juga?*
b) *Itulah akibat pahit peninggalan Sri Erlangga.Apakah gurumu yang lama, Tantripala, yang membisikan itu pada kupingmu, maka kau berani bicara seperti itu?*

Google Translate:
a) *Did Hyang Wisynu tell you to keep the eagle diligently too?*
b) *That's the bitter consequence of Sri Erlangga. What is your old teacher, Tantripala, who whispers it to your*

ears, and then you dare to speak like that?

The words 'tell' and 'bitter' in the two sentences above are examples of semantically word determined selection. In the first sentence (the word 'tell'), let's see how the first language was used before translation. "Did Hyang Wisynu command you to empower that eagle diligently too?" The word 'commission' means that someone is ordered to do something. Command here means action. If we use the word 'tell,' it is not wrong, but it is inappropriate. The word 'tell' here is only a command word, not an action. So the suggested word to use is 'command/ask'.

In the second sentences (the word 'bitter'), it can be seen that several words may have the same meaning. For example, the words 'bitter' and 'bad.' have negative meanings. Bitter is a word to express a negative that can be felt most clearly, such as butter's bitterness. However, if we look at the context of "That is the bitter result of Sri Erlangga's legacy.", It means that it is not something that can be immediately felt clearly by the human senses. However, the word is against bad circumstances or bad conditions. So the suggested translation is "That is the bad consequence of Sri Erlangga"

3) Preposition Partners

Preposition adalah kata yang Reveal relationships. Preposition partners relate to prepositions that follow before or after a verb, adjective, or noun. This is an example of the preposition partners found in the Arok Dedes dialog by Google Translate:

Source Text: Setidak-tidaknya dari Hyang Bthara Guru aku tahu, dua hari lagi kalian akan mendapat perintah untuk mengangkut upeti ke Kediri. Dari Hyang Wisynu aku tahu, kalian akan lakukan itu dengan patuh.

Google Translate: At least from Hyang Bthara Guru I know, in two days you will get an order to bring tribute to Kediri. From Hyang Wisynu I know you will do it obediently.

The preposition 'in' in a sentence is incorrect because the preposition 'in' is not a preposition to indicate time. The preposition 'in' is a preposition that indicates direction. Actually, in this sentence, the word above means 'for two days.' There is a time preposition most commonly used in English is 'on.' However, if we look at the sentence, it is appropriate that we use the 'for' to denote the specified time period.

4) Inappropiate Co-hyponym

There is ample neurolinguistic evidence to suggest that humans store words in such mental relationships of understanding. The meaning of vocabulary usually involves concepts and their relationships in the lexical field. Therefore, the category of lexical errors is reasonable concerning this system.

Inappropriate co-hyponym is an error that occurs due to incorrect choice of words or words that imply the wrong meaning or do not fit the actual context of the sentence. The example:

*Source Text: Demi Hyang Wisynu, pada hari penutupan **brahmancarya** ini, kami umumkan pada semua yang mendengar, pengantin kami ini, Dedes, kami angkat jadi Pramesywari, untuk menurunkan anak yang kelak menggantikan kami.*

*Google Translate: "For the sake of Hyang Wisynu, on the day of the closing of this **brahman**, we announce to all who hear, our bride, Dedes, we shall be Pramesywari, to bring down the child who will succeed us."*

Google translates the word 'brahmancarya' as 'brahman.' However, both have very different meanings of the word. The word 'brahmancaraya' is a self-discovery state, but the word 'brahman' implies a saint. So the word 'brahman' implies a meaning that is wrong or not right in the context of the actual sentence.

The table results show that formal errors occur more frequently (59.00%) than semantic errors (41.00%). These results also indicate that morphological knowledge is more difficult than semantic knowledge. As shown in table 1, there were a total of 34 lexical errors found in Arok Dedes Pramoedya Ananta Toer's story dialogue, which was translated by Google Translate. The most dominant type of lexical error is calque.

DISCUSSION

The translation is translating the meaning of the source text into another target text that can be accepted by the reader in the way intended by the author or translator. Likewise, translation is when the target text is transferred or transformed into the target text's surface structure, which involves three stages (analysis, transfer, and re-structure) (Nida & Taber, 1982). The use of machine translation (Google Translate) is an alternative way for most students or language users who need to translate their work from one language to another without paying any fees.

Hence, the best option is to use Google Translate, which is available free online. However, the question that arises is how Google Translate helps produce a good translation from the source text to the target text. Based on the findings of all texts used in this study, it can be said that the results of Google Translate still need to be revised, especially since this text uses Javanese stories and focuses on dialogue that contains cultural words or phrases. Based on some people, they agree that the use of Google Translate helps people to generate output from the target text even though in reality, sometimes the results do not make sense because Google Translate does not understand the context of the source text.

From the example "This is my son, yes, Durga" it translates to "This is Sarah, yes, Durga" and not "This is me, Durga." Herein lies the machine translation weakness, which connects language word for word but by no means with its meaning. Using the error analysis framework proposed by (Carl, 1998), the sample of errors made by Google Translate can be checked. Errors are categorized as formal errors and semantic errors, and each error is checked in detail.

For example, the translation, "By Hyang Wisynu, on the closing day of this Brahmancarya ..." is translated as "For the sake of Hyang Wisynu, on the day of closing of this Brahman ...". This is an interesting mistake made by Google Translate for failing to recognize that the word "Brahmancarya" refers to "Brahman." In fact, the two mean very different words. Thus, the source text, which is literally translated or mirrored, results in the target text's wrong text structure, which also affects its meaning. Based on these findings, formal errors were more frequent than semantic errors. This shows that knowledge of morphology is more difficult than semantic knowledge. Besides, it proves that Google Translate cannot translate cultural elements easily.

Lexical errors have been proven to be major errors in the dialogue of the translation of Arok Dedes Pramoedya Ananta Toer's story, which was translated by Google Translate. Lexical errors contribute to nearly half of all errors made. A closer look at the data shows that the second problem type of lexical error is semantic. This type of error can be attributed to underdeveloped vocabulary knowledge. Machine translation uses words in the semantically correct field, but the connotative meaning of the words used does not fit the context. This error is considered an error in the connotation of the word culture. There are two possible reasons for this. The first reason may be that machine translations do not have the same word count to cover the semantic field. The second reason that makes sense is that the machine translator does not fully know the word. It means that they do not know the appropriate collocates.

An important part of knowing a word is when, where, and how to use it. Formal errors make the most common lexical errors. The reasons for this error can be myriad. The authors may have accessed the wrong word in their mental lexicon because of how it was stored. Perhaps the word is included in their receptive vocabulary. The target language and the resulting language are morphologically similar, and as a result, machine translators think they can correctly produce the target language. Another common error that occurs is an 'incorrect collocation.' There could be several reasons for this collocation error. In the first place, there may be L1 interference, and the resulting collocation is the result of direct translation. Another possible reason is underdeveloped knowledge of the word. The word may be in their productive vocabulary, but machine translation may not understand the different connotation coloring, thus placing the word in the wrong context.

There are several mistakes in the 'word-formation.' This type of error frequently occurs for two reasons. The first is a misapplication of the L2 derived rule, or sometimes, machine translation applies the L1 derived rule to produce an L2 target word. This is outside the scope of this study but will be an interesting research topic for a different study. In summary, half of all errors made are classified as lexical errors. The most common type of lexical error encountered has to deal with a limited understanding of the semantic range of the words and how they intersect with other words. It was also found that the most significant mistake Google Translate made was in the wrong word order. In translation units (in dialogs containing cultural text), Google Translate cannot recognize the source text's root text due to the order of the elements that appear completely independent or are used for different lexical purposes.

This section also presents the results of previous research in the form of Farah Hana Amanah, Error Made by Google Translate and It is Rectification by Human Translators, Faculty of Languages and Linguistics University of Malaya Kuala Lumpur, 2017 where three human translators rectify or edit Google Translation raw outputs with access to the source text. Error corrections made by human translators are made based on a taxonomy of error categorization. The results show that human translators produce accurate output compared to Google translations. The reason

is because human translators know more about the context of the text than Google Translate.

Error analysis is collecting samples, identifying errors, and classifying errors, and finally evaluating the errors (which means they are edited by human translation) if necessary (Keshavarz, 1999). Following the framework of Keshavarz and Vilar et al. (in Hana Amanah, 2017), the output of Google Translation is given to three human translations. However, their translations may have different approaches to correcting errors or translation units. The criteria that need to be considered by human translation are the target language user, the function and adequacy of the text, and the ability to transfer content that is not specific or specific. The concern is the similarity of the two texts (source text and target text) (Carl, 1998). The equality must-see semantic and textual aspects as well as syntactic and lexical aspects. They cannot be seen alone because each language has different linguistic items and is sometimes ambiguous in its usage. This is because both the source and target texts must match each other in their function because each text itself has certain functions such as expressive, informative, or vocative. With regards to the quality of the final product produced by the three human translators for the correction of errors made by Google Translate, the human translator has the most acceptable and accurate translation.

The results also have implications for lexical teaching. If this problem is truly universal for all countries studying English as a Foreign or Second Language, a greater focus on collocations and word families is needed. This new research has provided confirmatory evidence to support the hypothesis that students with the same background at the same developmental stage but from different nationalities can make similar lexical errors in terms of type and number (Hemchua & Schmitt, 2006). This verifies their claim that these findings will 'appeal to the broader context of English as a Second Language (ESL)/English as a Foreign Language (EFL).' This research demonstrates the importance of understanding how the lexis is acquired and identifying where learning has not taken place and therefore the areas of teaching and/or remedial correction, hope that this paper helps fill gaps in future research and will help revive interest by encouraging practicing teachers to act on your own.

In other words, the research findings indicate that L1 plays an important role in the acquisition of L2 lexeme. It may also be helpful to encourage native English speaking EFL teachers to learn about the language of the host country they teach to understand the reasons behind some of the mistakes their students make. Besides, students must be trained on how to use machine translation effectively. Furthermore, all new lexical items must be taught in context. Students can also practice using Google to translate several types of text in vocabulary learning because of the confusion of binary terms and close synonyms shows the importance of this particular training.

For example, more direct teaching of words' morphological structure and the associations and collocations of words is required. Alternatively, as Zughoul (in Hamdi, 2016) tested alternatively, problematic word lists can be created and given to students, however. However, for many lists to bear fruit, problem words must be taught in their context and encouraged to use new words in their speech. in writing class. As in previous research, a lexical error analysis of advanced language learners' writings analyzed the misrepresentations intending to provide instructional advice for advanced language learners through lexical error analysis in their writing (Wells, 2013). The lexical error areas in advanced language learners (ALLs) have been studied very little. The study looked at the lexical errors made by advanced language learners in a university setting. It aimed to determine what types of lexical errors ALLs commit, the effect of direct first language translation on lexical errors, the effect of separate category cases on lexical errors, and pedagogical

implications. It was found that a large number of lexical errors were made. More than 50% of lexical errors relate to learners not understanding the semantic ranges of words and not understanding the corresponding word sets. Based on these findings, several approaches and activities are provided for use with ALL. The focus of this activity is to create individualized and differentiated instruction through the use of student writing and goal setting. This activity also provides deeper vocabulary knowledge to ALLs using semantic mapping, studying collocations, and using concordances.

As found in this study, teachers can use exercises to help students differentiate between minimal pairs and increase their morphological awareness when teaching vocabulary and spelling. To deal with collocation errors, students can be informed about the corporation's value and can be encouraged to access the corporation online and use these facilities when studying collocation. Also, students at beginner and lower intermediate levels can initially memorize word pieces in the learning collocation. With this in mind, we suggest that teachers use a taxonomy of lexical error or develop their own in their vocabulary teaching. We firmly believe that this taxonomy serves not only as a research tool but, more importantly, as a learning tool that teachers should use. When used effectively, this lexical error taxonomy can help students improve their metacognitive skills in recognizing and perhaps even correcting their own mistakes. This could be a way to help minimize lexical error fossilization.

CONCLUSION

After conducting a study of lexical errors, the authors came to several conclusions. This conclusion is to answer the research problems of this study. Of the twenty-one subtypes of lexical errors, only nine subtypes appeared in the Arok Dedes dialogue, and the rest did not appear at all. Some of Arok Dedes' dialogue's common mistakes are calque, statistical weight preferences, and semantically determined word selection.

Less common lexical error subtypes are false friend types, misselection, consonant based types, vowel based types, inappropriate co-hyponyms, and preposition partners. The other categories do not appear at all. They are the suffix type, prefix type, borrowing, coinage, omission, overinclusion, misordering, blending, a superonym for hyponyms, a hyponym for a supermom, wrong near-synonyms, and arbitrary combination. Of the two main types of lexical errors, formal errors are more problematic than semantic errors found in the Arok Dedes dialog.

Google translate does not easily translate morphological knowledge rather than semantics. Google is a great source of information and knowledge. From Google, we can know and learn all the sciences. One of the Google tools that helps internet users and has common uses is Google translate. This is because Google translate presents various languages. However, as a machine translator, Google Translate's results are not entirely correct. Therefore, it is suggested to readers, especially users, to be more corrective and careful in adopting them.

During this research, the researcher found another problem in Arok Dedes' dialogue, which was translated by Google Translate. Another problem that arises is how to revise the results of Google Translate, which have many errors and still need correction. Therefore, the authors suggest that other researchers conduct further research on this matter. Furthermore, Google Translate has difficulty with certain types of errors, as this study shows. It would be interesting to test other machine translations to compare accuracy rates to see which of them can produce the translation output with the fewest errors. Further research may also explore other text type genres and use larger volumes of text.
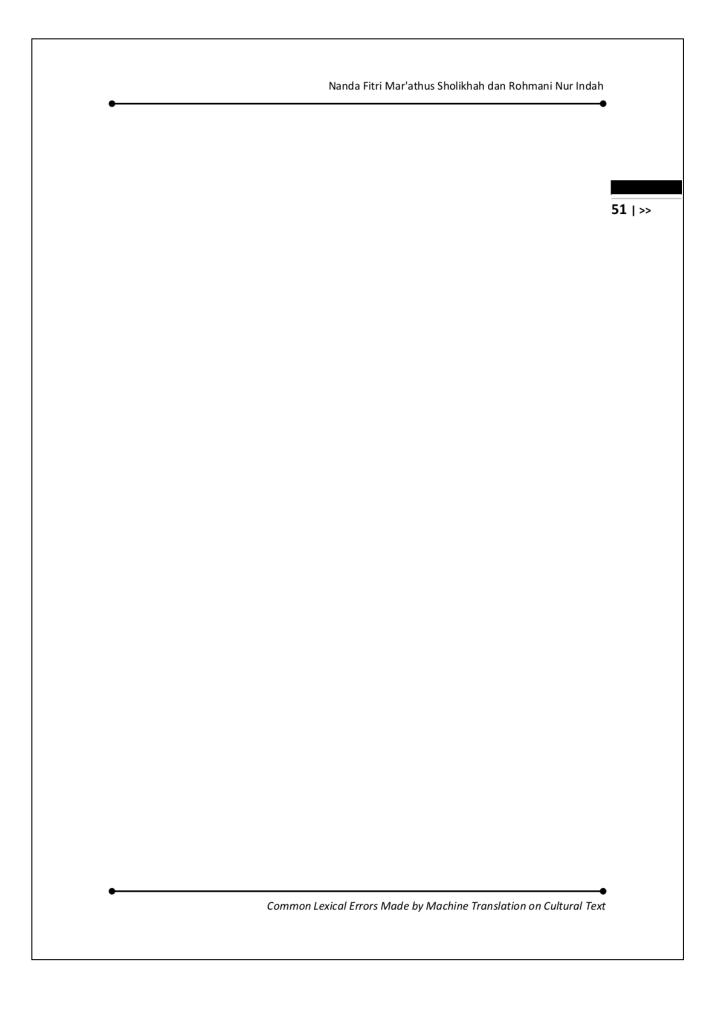
ACKNOWLEDGEMENT

appreciation to the advisors, Mr Bahruddin,   who always guide and give suggestions. S.S, M.Pd and Mr Chothibul Umam, M.Pd

REFERENCES

Amini, R., & Bayesteh, J. (2020). *Investigating Error Analysis in Interlinguistic English Language Teaching and The Consequences of Contradictory Analysis on Teaching Quality*. *4*(1), 41–44. https://doi.org/10.22034/jhi.2020.109840.

Arifin. (2011). *Metode Penelitian Kualitatif, Kuantitatif, dan R & D*. Alfabeta.

Bojar, O. (2011). Analyzing Error Types in English-Czech Machine Translation. *The Prague Bulletin of Mathematical Linguistics*, *95*(1), 63–76. https://doi.org/10.2478/v10108-011-0005-2

Carl, J. (1998). *Errors in Language Learning and Use: Exploring Error Analysis*. Longman.

Hamdi, S. (2016). An Analysis of Lexical Errors in The English Compositions of Tunisian Learners. *International Journal of Humanities and Cultural Studies*, *2*(4), 643–652.

Hana Amanah, F. (2017). *Errors Made by Google Translate and Its Rectification by Human Translators*. (Doctoral dissertation, University of Malaya).

Hemchua, S., & Schmitt, N. (2006). An Analysis of Lexical Errors in The English Compositions of Thai Learners. *Prospect : An Australian Journal of TESOL*, *21*(3), 3–25.

Hoed. (2006). *Teori dan Masalah Penerjemahan: Sebuah Pengantar Ringkas*. Prodc.

House, J. (2015). *Translation Quality Assessments*. Routledge.

Jamilah, N. (2012). Semantic Translation Errors as A Result of Google Translate (A Case of Errors of Translation on Homonymous and Polysemous Words in Bahasa Indonesia). *Diglossia: Jurnal Kajian Ilmiah Kebahasaan Dan Kesusastraan 4*(1).

Keshavarz, M. H. (1999). *Contrastive Analysis and Error Analysis*. Rahnama Press.

Kumparan.Com Tekno & Sains. (2020). Kumparan.Com Tekno & Sains.

Llach, M. P. A. (2005). A Critical Review of The Terminology and Taxonomies Used in The Literature on Lexical Errors. *Miscelanea: A Journal of English and American Studies* (31), 11–24.

Maheswari, P. A. M., Adnyani, N. L. P. S., & Suwastini, N. K. A. (2020). Interlanguage Analysis on Indonesian EFL Learners' Compositions. *Lingua Didaktika: Jurnal Bahasa dan Pembelajaran Bahasa, 14*(2), 121-131. https://doi.org/10.24036/ld.v14i2.108643

Nida, E. A. & Taber, C. R. (Eds.). (1982). *The Theory and Practice of Translation*. (vol.8). E. J. Brill Archieve.

Pateda, M. (2001). *Semantik Leksikal*. Rineka Cipta.

Sudaryanto. (1992). *Metode Linguistik*. Gajah Mada University Press.

Supendi, N. (2017). Analyzing Indonesian-English Abstracts Translation in View of Translation Errors by Google Translate. *Ekp*, *13*(3), 1576–1580.

Wells, N. (2013). *A Lexical Error Analysis of Advanced Language Learners' Writings*. (Doctoral dissertation). 108.

Wijayanto, P. W. (2020). *The Translation Analysis of English Imperatives into Javanese by Google Translate*. *021*, 210–217. https://doi.org/10.5220/0009018902100217.

Common Lexical Error