Error Analysis of Abstract Translation in Scientific Writing by Using Google Translate

Nyayu Yayu Suryani¹, Tira Nur Fitria²
¹Sekolah Tinggi Ilmu Kesehatan Siti Khadijah, Palembang
Email: nyayu.yv@stikes-sitikhadijah.ac.id
²Institut Teknologi Bisnis AAS Indonesia
Email: tiranurfitria@gmail.com

ABSTRACT

Google Translate is a multilingual neural machine translation service developed by Google to translate text, documents, and websites from one language into another. The purpose of this research is to find out Google Translate errors when translating bachelor's paper abstracts from Indonesian into English. Six abstracts of Non-English major undergraduate student papers from diverse faculties Sekolah Tinggi Ilmu Kesehatan Siti Khadijah Palembang were chosen at random. Following that, the data is compared for each sentence segment, as well as any words or phrases with problems are investigated. The result shows that the primary errors of abstract translations include lexico-semantic, tense, preposition, word order, distribution, and use of verb group, as well as active and passive voice errors. Abstract translation outcomes deviate significantly from expectations. Both translating English writings into Indonesian and Indonesian texts into English takes advanced translation abilities. Regarding translation, several factors should be considered; they include the linguistic system utilized for both languages and the cultural understanding of the two languages. In the case of scientific writing translation, involve a translator in improving the quality of translation. In terms of academics, it is recommended that further research be done on comparable challenges in abstract translation, but with more variety from diverse sources.

Keywords: abstract translation, scientific writing, google translate

Introduction

The context of machine translation implementation and study reveals a wide range of translation machines that have been the subject of a lot of machine translation quality testing. Research, example-based, open-source, pragmatic-based, rule-based, and statistical machine translation are all examples of machine translation (Elliott, 2006). Collecting samples, detecting errors, classifying, and analyzing them are all examples of error analysis. It also divided errors into two categories: syntactical-morphological errors and lexical-semantic errors. It classified incorrect Syntactical-morphological errors including the misuse of prepositions, documents, plural morphemes, qualifiers and intensifiers, and the use of standard Persian construction in English, while lexical-semantic errors include cross association and language switch. Linguistic errors are classified as (a) orthographic errors, (b) phonological errors, (c) lexico-semantic errors, and (d) morphological-syntactic errors (Keshavarz, 1999). Machine translation assessment research has exploded in popularity in recent years, and some professionals have expressed an interest in making use of corrective feedback to evaluate machine translations (Eftekhar & Nouraey, 2013). To evaluate the writings of adult Persian English students, models of error analysis were used. In his paper, mistakes in foreign language learning are a normal occurrence, and the importance of L1 is unavoidable. Error analysis is critical for students because it identifies
creating problems in their writing (Omidipour, 2014).

Translate via Google Translate is a free tool that allows us to change written documents from one text to another. It currently supports 90 languages. It is aware not only to translate a single word, but also a sentence, a paragraph of text, or a webpage. A website to translate a text, Google Translate looks at various documentaries to find the best translation pattern between human-translated texts. Statistical Machine Translation (SMT) stands for pattern searching. As a result, the accuracy of Google Translate is determined by the amount of human-translated texts that translate via google searches (Karami, 2014a).

Translate via Google began as a machine translation system based on a rules system (Fitria, 2018b). In 2006, it was followed by an SMT that used a statistical model to evaluate the translation of an expression. SMT employs a bilingual text corpus, which is a database of sentences written in both the source and target languages. For the computer to measure the likelihood of the terms, a large number of sentences will be given, for example, from English to Persian. If a term has a 75% probability of being translated into Y, for example, it would select Y as the translation of X. For machine translations of languages with simple linguistics and laws, It is simpler to use rule-based models and more effective.

The standard of translations is determined by the data given to the computer and a language pair used in the translation process for a machine translation service like Google Translate, which supports 90 languages and uses statistical models. (Karami, 2014a). For the first time, it conducted research and evaluated Google Translate’s translation accuracy across 50 languages, rather than just two. The writer concluded the study by pointing out that Google Translate is much better at translating converting one European language into another than it is at translating Asian languages into European languages (Aiken & Balan, 2015).

In terms of subject-verb agreement, when comparing Google Translate to human translators, Google Translate does not handle subject-verb agreement as well as human translators when translating English sentences into Persian (Bozorgian & Azadmanesh, 2015). The results of automated machine translation metrics are not only insufficient and unclear in defining machine translation efficiency, but they are also approximate and uncertain. As a result, they fall short of offering sufficient insight for error analysis (Chris Callison-Burch Miles Osborne Philipp Koehn, 2006). Google Translate is used to translate 100 methodically selected sentences from the Motarjem Hamrah translator software, which is then examined and organized into distinct tables based on the requirements given in the model. The results of evaluating and tabulating the frequency of errors, as well as a chi-square test, revealed no major variations in Google Translate’s quality from English to Persian and Persian to English. Furthermore, the most and least common errors were lexico-semantic and active/passive voice errors, respectively. To develop the framework, the paper identifies potential research directions (Ghasemi & Hashemian, 2016).

Several previous studies have been conducted related to abstract translation. (Fitria, 2018a) states that when abstractions are translated, information is transferred from the source language (SL) to the target language (TL). It should generate comparable text. As a result, readers can comprehend and decide whether or not to read the document in its entirety. It is supported by Alam (2020) that Google Translate may be used to translate text on a word-by-word, phrase-by-phrase, clause-by-clause, sentence-by-sentence, or even discourse-by-discourse basis. However, inaccuracies and anomalies in the translation of the supplied text are frequently discovered. While, Lawa et al (2022) state that the results of the Google Translate translation were “correct,” 14.28 % “very accurate,” and 5.72 % “inaccurate.”

A linguistic-textual activity that involves re-contextualizing a text in one language into another is referred to as translation. However, translation, as a linguistic-textual activity, is subject to and impacted by a range of extra-linguistic elements and contexts (House, 2015). A model of translation-oriented text analysis like this could be useful not only for translator students and teachers but also for experienced translators (Nord, 2005).

The translation is not only about tying our shoes or brushing our teeth; it is always a deliberate act. The translation is a complex procedure that necessitates rapid multilayered evaluations of semantic fields, idiomatic expressions, reader or listener response to sociology and psychology, and cultural differences (Robinson, 2003).

Machine Translation (MT) is the new
traditional and accepted name for computerized systems that produce translations from one natural language to another, with or without human intervention (Hutchins, W., & Somers, 1992). Google Translate is one of the many services offered by Google. Nearly 60 languages are supported by Google Translate. Google translator made and continues to make numerous blunders and errors, however in other circumstances, Google translator improved its translations by providing its users with tools such as editing, substituting, and deleting the incorrect alternatives it supplied. Google machine-translated sentences to compare and contrast with human reference translations that follow subject-verb agreement norms. After that, students studying translation will learn how to analyze translated sentences using both manual and automatic evaluation criteria (Bozorgian & Azadmanesh, 2015).

Machine Translation (MT) has recently progressed for language pairs with large amounts of transmitting data, but translation quality has not yet reached acceptable levels, especially for resource-poor languages with little or no parallel text to train statistical or example-based MT systems (Llitjós et al., 2005). Because the amount of translated texts fluctuates from user to user, the quality of Google Translate is determined by the amount of human-translated texts that Google Translate searches (Karami, 2014b). Google Translate is one of several machine translations that people use to translate texts into over 90 different languages all over the world. It can translate not only words, but also phrases, text sections, and whole pages. To translate a text, Google Translate looks through a variety of documentaries to discover the greatest translation pattern between human-translated texts. Statistical Machine Translation is the term for this pattern-finding process (Napitupulu, 2017).

Google Translate is certainly one of the most simple and accessible tools for anyone to use when they need to translate something. Students have found the benefits of this application and are using it increasingly frequently both inside and outside the classroom because it provides quick and rather accurate dual translation services in a variety of languages. The translation is based on several patterns that can be found in a variety of books. The authors of Google Translate are well aware of the difficulties that come with accurate translations, particularly in circumstances of exceptions and exceptions to exceptions. As a result, Google has chosen to examine a large number of documents that have already been translated by humans. These documents were compiled from several sources, including books, and various websites (Medvedev, 2016).

The translation will frequently be dependent on a specific language or language pair. Furthermore, there are certain clear advantages to using Google Translate: It is free and instant; it supports a wide range of languages for input and output; it supports voice recognition, web page translation, and the uploading of complete files for instant translation.

Error analysis is used to (1) characterize the tactics employed by learners in language education, (2) identify the causes of errors, and (3) gather information on frequent language learning challenges to produce resources and tactics to help learners avoid their mistakes (Longman Dictionary of Language Teaching & Applied Linguistics, 2010). Error analysis is a sort of linguistic analysis that examines the mistakes that students make. It entails comparing the errors made in the Target Language (TL) and the TL itself (Corder, 1974). Error analysis can be used to: a) determine how well someone knows a language, b) determine how a person learns a language, and c) learn about common language acquisition difficulties.”

Furthermore, error analysis examines the actual errors made by foreign language learners and attempts to describe the causes of such errors (Richards, 1992). With a range of methodologies, error analysis aids in the identification, classification, and systematic interpretation of language learners’ errors (Khodabandeh, 2007). When it comes to errors committed by English students, errors occur when the learner’s mastery of the target language's rules is lacking. Errors are thought to be systematic, guided by rules, and also rule-governed when they adhere to the learner's interlanguage rules (Keshavzar, 1999).

A translation error is viewed as some form of non-equivalence between the source text and target text or non-adequacy of the destination text (Koller, 1979). Translation error is multifaceted; it is an unavoidable fact resulting, among other things, from the tension between translation error dimensions—broad categories of translation errors like those related to meanings, sign up, or they create being measured—and this same fact that languages define having to adjust (Solano-Flores, G., 2014b).

In Indonesia, students must translate undergraduate thesis abstracts into English to complete their final year academic papers. Some colleges require students to have the abstracts of their scientific papers translated by an official institutional language center, while others do not need students to have their abstracts translated into English by an official translation agency (Fitria, 2021b). As a result, students are free to use whatever resources they have. Their abstracts are translated as quickly as possible without regard for the accuracy of the translated content. Students’ last resort is to use Google Translate to translate students’ abstracts because the service is free and gives real-time results in seconds.

The majority of students, on the other hand, are oblivious of the implications it has on them, such as incorrect word choice, prepositions, word order, and so on. They have been accused of using Google Translate to translate their bachelor's paper abstracts into English. This phenomenon has been observed at Sekolah Tinggi Ilmu Kesehatan Siti Khadijah Palembang. The college administration does not demand final-year students to use a professional translator to appropriately translate the abstracts of their bachelor’s papers.

Therefore, this research finds out Google Translate made mistakes when translating bachelor's paper abstracts from Indonesian to English in the translation of abstracts from Indonesian to English, how common are lexicosemantic, tense, preposition, word order, distribution, and use of verb group, as well as active and passive voice errors.

Method

A qualitative technique was used in this investigation. Qualitative methods emphasize more on observing phenomena and researching more on the substance of the meaning of these phenomena. The analysis and sharpness of qualitative research are greatly affected by the strength of the words and sentences used.

The document is used for collecting data in this research. A large number of facts and data are stored in materials in the form of documentation. Most of the available data are in the form of letters, diaries, souvenirs, reports, artifacts, photos, and so on. The document used in this research is taken from the abstract journal. A total of fifteen abstracts were chosen at random from several faculties to collect the data, including the faculties of nursing, midwifery students, and pharmacy students.

Machine translation, i.e. Google Translate, was used to translate all of the abstracts. A total of 6 phrase fragments were obtained from the fifteen abstracts.

The initial step in assessing the data was to tabulate and compare the rates of various types of Indonesian to English translation errors. The frequency of accurate and erroneous translated taken for several sorts of translation faults, such as lexicosemantic errors, tense errors, incorrect use of prepositions, word order errors, flaws in the distribution and usage of verb groups, and active and passive voice errors (Keshavarz, 1999).

Findings and Discussion

Some of the research's findings demonstrate Google Translate's capacity to translate Indonesian articles from English to Indonesian in terms of clarity and correctness. The following are some examples of clarity and correctness in Google Translate translation results.

Error in Lexical Semantic

Lexical semantics is the study meaning of a word. The task of segmenting a sentence into its lexical expressions and applying semantic labels to those expressions is known as lexical-semantic analysis (LSA). A word or set of words with a "basic" meaning or function is referred to as lexical expression (Schneider, 2012). The following table is error translation in lexical-semantic:

<table>
<thead>
<tr>
<th>Source Text</th>
<th>Google Text</th>
<th>Suggested</th>
</tr>
</thead>
<tbody>
<tr>
<td>Berdasarkan</td>
<td>Based on</td>
<td>Based on the</td>
</tr>
<tr>
<td>evaluasi</td>
<td>evaluations</td>
<td>evaluation a</td>
</tr>
<tr>
<td>dilakukan</td>
<td>carried out for</td>
<td>maximum period</td>
</tr>
<tr>
<td>selama 3x24</td>
<td>3x24 hours</td>
<td>of 3 x 24 hours</td>
</tr>
<tr>
<td>jam</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Based on the table above, shows the translation of text in Bahasa “dilakukan” translated by Google translate “carried out for 3x24 hours”, but it suggested “a maximum period of 3x24 hours”, this sentencing error in lexical-semantics.

Tenses

In English, we utilize the concept of tense to refer to time in three ways: past, present, and future. The language uses tenses and aspects to express time—how an event occurs in the past, present, or future. They're important for comprehending the grammar and structure of all languages, and they're popular among linguists (Binnick, 2011). The following table is error translation in tenses:

Table 2
Based on the table above, shows the translation of the text in Bahasa “ialah” translated by Google translate “is”, but it suggested “was”, this sentencing error found in tenses.

**Preposition**

A preposition is a tiny, common word that indicates direction (to in "a letter to you"), location (at in "at the door"), or time (by in "by noon"), or that introduces an item (of in "a basket of apples"). The object of a preposition is usually a noun (noon), a noun phrase (the door), or a pronoun (you). The following table is an error in preposition:

<table>
<thead>
<tr>
<th>Source text</th>
<th>Google text</th>
<th>Suggested</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sumber</td>
<td>Sumber</td>
<td>Sumber</td>
</tr>
<tr>
<td>kolak</td>
<td>kolak</td>
<td>kolak</td>
</tr>
</tbody>
</table>

Based on the table above, shows the translation of the text in Bahasa “pada” translated by Google translate “for”, but it suggested “of”, this sentencing error in a preposition.

Prepositions come in three different types: single preposition, phrase preposition, and prepositional phrase (Fitria, 2022). The purpose of a preposition in discovered is to indicate location and time. Some prepositions were not translated into the target language (TL) and were translated as di, ke, dalam, pada, selama, dari Ni wayan 2021. The usage of the prepositions in and at was found to be more common than the use of the preposition on. Misformation mistakes (substitution of "in" for "at," "in" for "on," "in" for "from," "at" for "to," and "at" for "in") and addition errors (adding "in" and "at") are the two types of errors (Suzanne, 2017). Wrong prepositions were the most common error, with L 1 interference being the leading source of preposition errors (Akarapisit, 2009).

**Word Order**

The study of word order is crucial to linguistics since it is one of the key properties on which languages are compared (Song, 2012). The following table is an error in word order:

<table>
<thead>
<tr>
<th>Source text</th>
<th>Google text</th>
<th>Suggested</th>
</tr>
</thead>
</table>

Based on the table above, shows the translation of text in Bahasa “didapatkan” translated by Google translate “in the study found that the patient’s main complaint”, but it suggested “was found”, this sentencing error in word order.

The English language system uses articles, set word order, and tenses to describe the verb aspect, whereas the Spanish language system uses perfective and imperfective aspects. The usage of a formal foreign language is also linked to appropriate punctuation, which causes challenges for students (Klimova, 2012).

**Distribution and Use of Verb Group**

Tense, aspect, mood, and modality are among the most difficult and fascinating problems in the English language (Leech, 2004). Though the verb group's structure is simpler than the noun group's, the verb's elements are more diverse. This group consists of three components: the auxiliary, the verb, and the verb's extension (Blake, 2002). The following table is an error in the distribution and use of verb group:

<table>
<thead>
<tr>
<th>Source text</th>
<th>Google text</th>
<th>Suggested</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hasil penelitian yang didapatkan yaitu pasien mengeluh sesak nafas terus-menerus.</td>
<td>Hasil penelitian yang didapatkan yaitu pasien mengeluh sesak nafas terus-menerus.</td>
<td>Hasil penelitian yang didapatkan yaitu pasien mengeluh sesak nafas terus-menerus.</td>
</tr>
</tbody>
</table>

Based on the table above, shows the translation of text in Bahasa “yaitu” translated by Google translate “was”, but it suggested “obtained”, this sentencing error in distribution and use of verb group.

**Active and Passive Voice**

The subject of a verb in this form is the person or thing who acts. The subject acts on the object in an active voice sentence. The performer is emphasized by using the active voice. The passive voice is a verb form in which the subject is impacted by the verb's action. In
other words, the form of a verb in which someone or something else acts on the sentence's subject (Joshi, 2014). The following table is an error in active and passive voice:

Table 6

<table>
<thead>
<tr>
<th>Source Text</th>
<th>Google Text</th>
<th>Suggested</th>
</tr>
</thead>
<tbody>
<tr>
<td>Telah dilakukan observasi oleh dokter di rumah sakit.</td>
<td>Has observation by doctor in the hospital</td>
<td>The observations have been made by doctor in the hospital</td>
</tr>
</tbody>
</table>

Based on the table above, shows the translation of text in Bahasa “telah dilakukan observasi” translated by Google translate “has observation”, but it suggested “observations have been made”, this sentencing error in active and passive. Omissions, additions, misformations, and misordering are the four types of errors found in this document. Errors are more likely to occur due to misformation. The errors can be evaluated to see whether teaching items require additional explanation. It can also provide important information on the many types of errors so that people are aware of the risks of making those mistakes (Abdulqodir, 2006).

Discussion

Occasionally, translational abstract outcomes deviate significantly from expectations. Both translating English writings into Indonesian and Indonesian texts into English takes advanced translation abilities. These abilities are necessary to communicate the information and knowledge included in the texts (Fitria, 2018). Regarding translation, several factors should be considered; they include the linguistic system utilized for both languages and the cultural understanding of the two languages.

Many authors now believe that the passive voice degrades the quality of scientific writing, contrary to the previous consensus. Studies utilizing scientific articles, on the other hand, are uncommon. This study looked at the proportion of passives utilized, as well as the circumstances and forms in which they appeared, using a corpus of 60 scientific research articles from six publications.

The findings found that passive clauses made up around 30% of all clauses. The canonical form was the most common, followed by the naked passive, accounting for about a quarter of all clauses studied. In major sentences, passives were most commonly utilized, followed by relative and adverbial clauses. Around 29% of all passives were found in this area (Alvin, 2014).

Those statements are similar to Lawa et al. (2022) state that the results of the Google Translate translation were very accurate, but also inaccurate. Google Translate's results were deemed accurate because Google Translate increased the quality of translation by transitioning from statistical machine translation to neural mechanic translation. It was rather accurate, as the result of translating a lengthy statement and several paragraphs, Google Translate produced a reasonable outcome. While it was deemed erroneous due to the numerous errors in the translation results from Google Translate, the translation results were referred to as pre-translation that needed to be rechecked by users.

Students had difficulties due to the structural or pattern discrepancies between Indonesian and English, as well as the preciseness of words in context. Alam (2020) also states that Google Translate may be used to translate text on a word-by-word, phrase-by-phrase, clause-by-clause, sentence-by-sentence, or even discourse-by-discourse basis. However, inaccuracies and anomalies in the translation of the supplied text are frequently discovered. Errors can be caused by a lack of language skills such as the learner's lack of knowledge of the target language's correct rules or an error in translating using specific translation tools (Fitria, 2020).

The findings are also supported by (Fitria, 2021a) that Google Translate is a machine translator, but the final translation result will always contain some degree of ambiguity and error, as is the case with Indonesian articles translated into English. Due to the difficulty of mastering English grammar, users cannot fairly expect a machine translator to grasp every aspect of human communication. That is why, when it comes to the clarity and accuracy of Google Translate, the response is that it has a long way to go before it can consistently, clearly, and properly translate the language without errors.

Google Translate's English translation remains murky in terms of intelligibility, although it translated the language word for word. In terms of correctness, it refers to the
grammatical, punctuation, and spelling norms that regulate writing. Grammatical and punctuation errors are examples of non-correctness. While machine translators have made significant progress in a short period, many characteristics, such as grammar and punctuation, remain untranslated.

Machine translation technology cannot be detached from human effort (Fitria, 2021). In other words, customers should never rely on machine translation but rather on expert translation. No machine translator, on the other hand, can create translation outputs as precisely as human skills. Translation, in particular, is concerned with machine translation as a means of assisting translators in evaluating the dictionaries used as language elements.

In this relation, it is necessary to emphasize that the existence of machine translation is a supplement to the world of translation, not a replacement because the sophistication of the machine will never match the cognitive flexibility of the human brain in adapting the translation results to the current context. Accurate translation is usually subjective and frequently a one-way street. As a result, many machine translators may be used to translate.

**Conclusion**

This study looked into the types of problems that were discovered in the translation of Indonesian papers into English by students at STIK Siti Khadijah as a result of using Google Translate to translate the text. The primary fault identified in abstract translations, lexicosemantic, has influenced the content of the abstract. Some terms in the original text have been mistranslated and do not fit the context. As a result, it is recommended that human translators adapt students’ words by considering the context as a whole. In terms of academics, it is recommended that further research be done on comparable challenges in abstract translation, but with more variety from diverse sources.

**References**


