
TRANSLATION ERRORS IN UNDERGRADUATE SCIENTIFIC WRITINGS: A CORPUS-BASED STUDY

Roswani Siregar¹, Heni Subagiharti², Diah Syafitri Handayani³,
Sutarno⁴, Farida Hariani⁵, Liesna Andriany⁶

^{1,5} Universitas Al Azhar, Medan, Indonesia

² Universitas Asahan, Sumatera Utara, Indonesia

³ Universitas Sumatera Utara, Medan, Indonesia

⁴ Cendana Polytechnic, Medan, Indonesia

⁶ Universitas Islam Sumatera Utara, Medan, Indonesia

E-mail: roses_air@yahoo.com

Received: 2024-12-02

Accepted: 2025-01-29

Published: 2025-06-26

Abstract

This study examines translation errors in 100 undergraduate abstracts from five Indonesian universities, focusing on syntactic, semantic, and miscellaneous errors. The corpus of the study was collected from the online repository. All the information related to the authors was eliminated from the corpus, and only the main body of the abstracts was used as samples in this study. Most errors are categorized as syntactic errors, consisting of errors in tenses, prepositions, and punctuation. Another category of errors is semantic errors, which consist of literal translations, deviations in meaning, and additions. Those errors were generally due to students' lack of knowledge of translation. By identifying errors, these findings can be a basis for consideration when introducing translation materials in English courses at the university level.

Keywords: *errors; scientific; students; translation writing; undergraduate*

1. Introduction

Translation acts as a bridge to connect different languages and cultures and makes a significant contribution to cross-cultural communication (Jahanshahi et al., 2015). However, it is not an easy job to do because it is not merely replacing words in one language with another, but conveying the meaning as naturally as possible. Translation examines the processes and relationships between translation and other scientific disciplines.

Scientific language has unique characteristics that make it completely different from other genres. Scientific works can be recognized, both in terms of wording and structure. Scientific language is regarded as a conceptual map, where experienced readers can follow a written line of reasoning full of general concepts and a set of specific terms used, which vary depending on the target audience. Semantic and grammatical features are characteristic of scientific language, where a very compact syntactic structure, elimination of articles or prepositions to obtain conciseness, avoidance of relative clauses and subordination, pre-modification, and complex nominalization (Gotti, 2003).

In Indonesia, most study programs require that abstracts of scientific papers should be written both in Indonesian and in English. Abstract translation serves as a measure of linguistic competence and is a critical component of scientific communication. However, translation errors remain prevalent, underscoring the need for systematic analysis and pedagogical interventions. Thus, translation should be taught at the university level (Amin, 2019), and it is necessary to carefully train prospective translators to produce efficient translators (Cúc, 2017).

Obstacles in the translation process often cause students difficulties in translating. These barriers include the lack of understanding of the text and resources available to search for new vocabulary, expressions, and idioms. Other barriers are the inability to identify problems in translation and find possible compensatory strategies, or difficulties in producing the target language, for example, finding the equivalent vocabulary, idioms, structures, grammar, or substitutions (Amin, 2019). Thus, it is necessary to investigate further the types of translation errors made by students in their scientific writing abstracts.

2. Literature Review

Scientific writing often relies on precision, and small differences in words can carry significant meaning. Errors in scientific writing can lead to miscommunication or misunderstanding of complex ideas, especially when nuances of language are lost or misinterpreted. Sentence structures in the target language might differ significantly, leading to awkward phrasing or changes in meaning. Moreover, the changing of the meaning or improper translation is detrimental to the readers.

2.1 Errors in Translation

Translation errors refer to the production of incorrect speech and written forms by non-native speakers of the target language, due to incomplete knowledge of the rules of the target language. According to (Pym, 2018), errors originate from different levels of language, pragmatics, or culture. Another statement from (Séguinot, 1989) argues that errors are a result of misunderstanding the source text or the translator's inability to produce the target text (as cited in (Jahanshahi et al., 2015)).

Errors are significant in three ways, as they tell the teacher what needs to be emphasized, how language learning progresses, and what prerequisites have to be achieved (James, 2013). Therefore, some previous studies in translation have been conducted in this respect. For instance, (Na, 2017) identifies the errors in translating topic-comment structures. (Károly, 2012) reveals the recurring patterns of lexical, syntactic, and textual errors in translations from English into Hungarian. (Zhao, 2013) investigates the reasons behind some errors in students' English-Chinese translation by analyzing their mistakes in idiomatic usage, and the lack of knowledge in the cultural background. (Ardeshiri & Zarafshan, 2014) found that understanding the pragmatic aspects was the most frequent problem in translating from English into Persian. (Abdelmajid & Karazoun, 2016) concludes that EFL students' linguistic errors in translation are grammatical, discourse, and lexical. (Utomo, 2016) classifies students' grammatical translation errors as those of omission, addition, selection, and ordering. (Wonggranu, 2017) reports that students have syntactical and semantic errors in translation. All these studies agree on the importance of identifying the reasons behind students' errors in translation. They suggest teaching foreign learners

some strategies or techniques to correct their errors. Moreover, it is recommended to raise students' awareness of their frequent errors in translation to avoid such types of errors.

Hurtado in Waddington (2001) put forward a translation quality assessment model that is based on error analysis. This model is categorized into three types of errors. Translation error: Inappropriate translation that affects the understanding of the source text. These errors are divided into eight categories: contresens (misinterpretation), faux sens (wrong meaning), nonsense (no meaning), additions, omissions, unresolved extra-linguistic references, loss of meaning, and inappropriate linguistic variations (register, style, dialect, etc.). Major errors in translation: Inappropriate translation that affects the expression in the target language. These major errors are divided into five categories: spelling, grammar, lexical items, text and style. Minor translation errors: Inadequate translation that affects the delivery of the main function or secondary function of the source text.

2.2 Model of Analysis

There is no fixed model for categorizing translation errors. Translation errors start from the lack of understanding, incorrect use of words, and so on. According to (Pym, 2018), there are two types of translation errors, namely binary errors and non-binary errors. Binary errors refer to any errors that are considered incorrect translations. In contrast, non-binary errors refer to totally wrong translations but may be inappropriate and correctable. In the classroom, this categorization of errors can reveal students' language or translation competence. Binary errors imply that language competence needs to be improved. On the other hand, non-binary errors reflect translation competence, which Pym defines as the ability to create a group of target texts, and then select the best one that suits the purpose and reader (Wongranu, 2017).

Cordero (1984) and West (1994) studied error analysis to understand the behavior of language users. The analysis aimed to understand the linguistic aspects of error production. After several decades, there is now a resurgence of interest in the concept of linguistic relativity regarding error production. One of them is an analysis of language competence through analysis of errors in the translation process.

Error analysis has helped in understanding errors not only as an undesirable phenomenon in a language but as a source of information that is applicable to improving production in a second language. Errors discovered in writing can shed light on the writing process and help us understand the mechanisms used by non-native speakers. As a result, by understanding these error patterns, several strategies can be plotted to improve writing skills in a second language and the cognitive processes of language production are considered when analyzing errors.

Understanding another language does not depend on identifying structures equivalent to those of the mother tongue but on the equivalence between concepts that arise from reality and then identifying the appropriate way to express them. Students who use English as a second language rely on word-to-word links in the early stages, but as their second language acquisition progresses, they link the L2 directly to concepts (Wu et al., 2016). For example, nouns help us communicate and express thoughts, and consequently change along with it, thus preventing total equality between languages. In this sense, to uncover the mechanisms behind language, cognitive interpretation of language becomes important.

The first issue in this paper is error identification. Correct identification of errors is based on determining the causes and processes followed in language production. In analyzing errors, it should not merely focus on their nature, but also identify and interpret errors in a second language (Tisgam, 2009) (Carrió-Pastor & Mestre-Mestre, 2014). Therefore, in this paper, it is important to discuss interlingual errors which are caused by the interference of the first language with the second language, and intralingual errors that occur regardless of the mother tongue and are caused by an inadequate learning process (Larsen-Freeman, 2013; Wongranu, 2017). Errors caused by interference with the mother tongue or deficiencies in second language acquisition that refer to linguistic and socio-communicative aspects. A third cause of errors may be due to cognitive aspects of language production, namely conceptual disturbance. This third cause can explain some errors caused by misunderstanding the relationship between images, concepts, and terms. Second-language speakers must be aware of the fact that words are simply representations of concepts and that the ability to associate one concept with several terms is required if they speak several languages (Wongranu, 2017). Reported three other common errors in translation: syntactic error, semantic error, and miscellaneous error.

3. Research Method

From the outset, the type of corpus used in the study and the conclusions of the lexical errors made in scientific English were established. The corpus of the study was collected from the online repository of five universities. It consists of one hundred abstracts in English and Indonesian as part of the final task written by undergraduate students of the Economics Faculty for the period 2018 to 2022. All the information related to the authors was eliminated from the corpus and only the main body of the abstracts were used as samples in this study. Those samples were filed into the word processor (Appendix 1).

A list of the tables is provided in Microsoft Excel that consists of columns for the coding section of scientific writing title, types of error, and translation quality assessment scoring. An evaluator analyzed the sample with scrutiny based on the type of translation errors and put a mark with different colors for each type. The evaluator took notes for each error taken from the samples and put it forward in the data tabulation. Furthermore, the errors and assessment categories were scored and analyzed with statistics for the reliability test.

4. Results and Discussion

The results of this research start from corpus statistics, which explains corpus statistical data as a research object. Then, after the analysis was carried out, it was found that the frequency of errors was grouped into 3 categories, namely syntactic, semantic, and other errors. After that, descriptive statistics are presented from the assessment of translation quality.

Table 1 Corpus Statistics

Description	Total
Total corpus	100
Total Words in corpus	22,142
Total Sentences in corpus	864
The average word in the corpus	221

Average sentences in the corpus	8
Average paragraph	2 to 3

Table 1 shows that the total corpus studied was 100 abstracts with a total of 22,142 words, while the total number of sentences was 864. It is known that the average number of words in the abstract was 221 words contained in 2 to 3 paragraphs.

4.1 Errors in Abstract Translation

The corpus used in this paper for the detection and categorization of errors displays the features shown in Tables 2 and 3 below. It can be seen that the types of error, frequency, and percentage of subcategories. The evaluator's decision to translate the assessment result is also summarized in Table 3.

Table 2 Frequency of translation errors

No	Type of error	Frequency	Percentage
1	Syntactic	266	66,3%
2	Semantic	127	31,7%
3	Miscellaneous	8	2,0 %
Total		401	

As shown in Table 2, out of a hundred abstract translations, there are 401 errors found that are categorized into syntactic, semantic, and miscellaneous errors. Syntactic errors had the highest frequency (66.3%). This type of error could be divided into 6 sub-categories. In the second place, the semantic error had 127 errors that were divided into 5 subcategories. Translation errors that did not fall into the syntactic or semantic categories were classified as miscellaneous errors. It had 2.0% of total errors, as provided in Table 3.

Table 3 Categories of Error

No	Categories	Occurences	Percentage
A	Syntactic error		
1	Countability	18	3,7%
2	Determiner	28	7,0%
3	Tenses	99	24,7%
4	Preposition	73	18,0%
5	Punctuation	45	11,2%
6	Passive voice	7	1,7%
B	Semantic Error		
7	Too Literal	96	23,7%
8	Addition	9	2,2%
9	Omission	3	0,7%
10	Deviation	14	3,5%
11	Upper/Lower Case	6	1,5%
C	Miscellaneous		
12	Inconsistency	8	2,0%
Total		401	100%

Table 4 Descriptive statistic of Errors

Types of Errors	N Corpus	Min.	Max.	Mean	Std. Deviation
Syntactic Error	100	0	6	2,70	1,396
Semantic Error	100	0	4	1,36	,823
Valid N (listwise)	100				

As seen in Table 4 and Figure 1, most errors are categorized as syntactic errors, namely 66.3%, consisting of errors in tenses (24.7%), prepositions (18%), punctuation (11.2%), determiner (7%), countability (3.7%), and passive (1.7%). The second error is categorized as a semantic error, namely 31.7%, consisting of too literal errors (23.7%), deviation (3.5%), addition (2.2%), upper/lower case (1.5%), and omission (0.7%). Another type of error is inconsistency, which is categorized as miscellaneous error (2.0%).

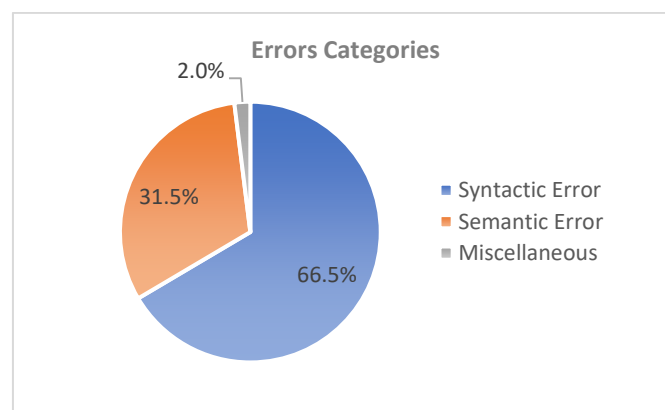


Figure 1 Categories of Errors

4.2 The Syntactic Errors

Approximately all of the students received negative points for the tenses because of incorrect or unclear paragraphing. There are eight to ten sentences in an abstract, though they translate the whole text into one to three paragraphs. Other prominent syntactic errors are prepositions, punctuation, and determiner. In some cases, problems in the use of countability and passive voice were observed as well. The sources of such an error type might refer to students' carelessness and lack of knowledge of the conventions of the Indonesian language regarding punctuation. As noted earlier, the most frequent error type found in the translations of the play text is tenses.

Countability refers to errors involving countable, uncountable, singular, and plural nouns. This concept does not exist in Indonesia. Indonesian nouns always take the same form regardless of the quantity and whether we can count the nouns is not of importance. In contrast, English nouns are divided into countable and uncountable nouns. In addition, countable nouns can take a singular or plural form, depending on the quantifier. Therefore, students often make mistakes if they ignore this aspect. Students are often found committing mistakes such as not putting an "s" after countable plural nouns, or using the plural form with.

Errors involving determiners were another problem. This type of error includes articles and quantifiers. In fact, articles seem to be the more prominent problem, often causing problems for students. Despite the apparent simplicity with only three choices, "a", "an" and "the", the students found this to be one of the most difficult grammatical aspects of English.

The students were often confused between definite and non-definite nouns; therefore, they are frequently uncertain whether to use “a/an” or “the”. In the case of quantifiers, they usually use them correctly. However, when it comes to the word most and non-specific numbers, such as “sample” and “population”, the students often use them incorrectly. For example, the students often wrote “total sample is 45 people”.

Students' translation procedures refer to processes the students used to translate the source text. Most of the students read the source text, translate it in their minds, and write down the target text. This procedure is referred to as the “read-and-translate approach” and it is often used in translation classrooms (Davies, 2004 as cited in (Zainudin & Awal, 2012). This approach, without thorough revision, often results in a word-for-word or a literal translation.

4.3 The Semantic Errors

It has been found in various studies that the most frequent type of error in the translation of this type of text is terminology. This finding is in line with (Duklim, 2022) statement about non-fiction texts being difficult to understand because of the terminology they contain. In parallel, these findings support Soltani et al. (2020) who argued that translating special texts such as abstracts is difficult because the way ideas are expressed in academic language is different from everyday language.

The second translation error is a deviation of meaning. In this study, the percentage of meaning deviation errors was 3.5%. The frequency is 14 out of 401 errors. Translation errors, and deviations in meaning, are vague translations. Readers can easily obtain the purpose of the writing by reading clear passages so that vague translations prevent readers from understanding the meaning properly.

Word-for-word translation is concerned with translation close to the source text, where words are translated in the same sense as they appear in the source text and the context is ignored. The translation still approaches the source text in a literal translation, and the grammatical construction is adjusted as closely as possible to the target text in the process. However, the meaning of the translated words is still far from the context, because in the end the source text is read and translated word for word. As a result, this often leads to grammatical errors in the output.

4.4 Translation Quality Assessment

Once errors are identified, an assessment of the quality of the translation is carried out. The assessment scores were tabulated into a table and descriptive analysis was carried out as shown in Table 5 below.

Table 5 Descriptive Statistic Translation Quality

Components	Descriptive Statistics						
	N	Mean			Std. Deviation	Variance	
	Corpus	Min	Max	Statistic	Std. Error	Statistic	Statistic
Accuracy	100	3	4	3,20	,040	,402	,162
Acceptability	100	3	5	3,60	,051	,512	,263
Readability	100	3	5	3,95	,039	,386	,149

The results of the descriptive analysis show that the average score for the level of translation accuracy is 3.20 on a scale of 5. The level of acceptability is with an average score of 3.60 on a scale of 5. Meanwhile, the level of readability is with an average score of 3.95 on a scale of 5. Apart from that, the minimum score for all assessments is 3 and the maximum score is 5. It can be seen that the highest score in the translation results is readability.

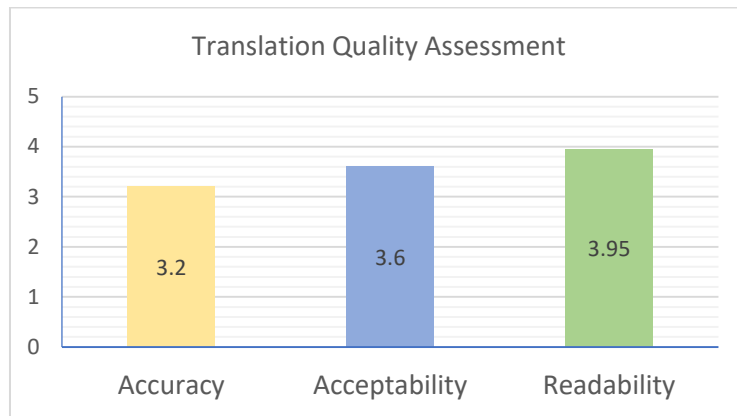


Figure 2 Translation Quality Assessment

From the results of descriptive statistics on translation quality, it is known that the translation accuracy level is 3.20, based on a scale of 5. This means that the translation accuracy is categorized as quite good. The average level of acceptability is 3.6 on a scale of 5. This means that the acceptability of the translation is also considered quite good. In terms of translation readability, an average of 3.95 on a scale of 5 was obtained, which means the translation readability level is categorized as good. Overall, it can be said that the quality of the abstract translation in English in this study is quite good.

4.5 The Cause of Errors

More class time should be given to the topics that are the most challenging for students. Some grammatical points, such as nouns, articles, and tenses, are problematic for students, so two to three class periods instead of one should be spent on these topics. Among the factors causing translation errors, the majority of students do not understand the text of the source language.

The cause of error is also driven by the lack of vocabulary and poor understanding on the source language context. Meanwhile, the factor of error in the least translation is the level of the quality of the source language using incorrect grammatical, the sentence is taxa (blurred), the use of the idea is not coherent, and many punctuations. The factor causing this error is that the source language structure is not grammatical (Silalahi et al., 2018).

5. Conclusion

This research investigates the types of translation errors frequently made by students at 5 universities in Indonesia and identifies the causes. Most errors are categorized as syntactic errors, which consist of errors in tenses, prepositions, and punctuation. Another category of errors is semantic errors, which consist of literal translations, deviations in meaning, and additions. By identifying errors in translating abstracts into English, these findings can be a basis for consideration for introducing translation materials in English

courses at the university level. Moreover, translation is related to other scientific disciplines, including sociology, psychology, philosophy, economics, etc. Further studies should explore longitudinal data to assess the impact of translation training on error reduction.

From the results of this research, it was concluded that the errors found in translation were generally due to students' lack of knowledge in translation. The absence of translation material means there is no translation practice for departments other than the language faculty. Thus, translation-focused training across disciplines should be established to enhance students' linguistic and cultural competence.

References

- Abdelmajid, G., & Karazoun, A. (2016). A Linguistic Analysis on Errors Committed by Jordanian EFL Undergraduate Students: A Case of News Headlines in Jordanian Newspapers. *English Language Teaching*, 9(8), p170. <https://doi.org/10.5539/ELT.V9N8P170>
- Amin, E. A.-R. (2019). Using Awareness Raising in Syntactic and Semantic Errors to Foster Translation Performance among Majmaah University EFL Students. *Arab World English Journal*, 10(2), 196–212. <https://doi.org/10.24093/AWEJ/VOL10NO2.16>
- Ardehshiri, M., & Zarafshan, M. (2014). Students' Causes of Errors in Translating Pragmatic Senses. *International Journal of English and Education*. Vol.3 (4).
- Carrió-Pastor, M. L., & Mestre-Mestre, E. M. (2014). Lexical Errors in Second Language Scientific Writing: Some Conceptual Implications. *International Journal of English Studies*, 14(1), 97–108. <https://doi.org/10.6018/IJES/14/1/154361>
- Cordero, A. D. (1984). *The Role of Translation in Second Language Acquisition*. American Association of Teachers of French. <https://www.jstor.org/stable/392747>
- Cúc, P. T. K. (2017). An Analysis of Translation Errors: A Case Study of Vietnamese EFL Students. *International Journal of English Linguistics*, 8(1), p22. <https://doi.org/10.5539/IJEL.V8N1P22>
- Duklim, B. (2022). Translation Errors Made by Thai University Students: A Study on Types and Causes. *REFlections*, 29(2), 344–360. <https://doi.org/10.61508/REFL.V29I2.259858>
- Gotti, M. (2003). *Specialized Discourse : linguistic Features And Changing Conventions*. 351. Retrieved from <https://tesl-ej.org/ej29/r10.html>
- Jahanshahi, M., Jahanshahi, M., & Kafipour, R. (2015). Error Analysis of English Translation of Islamic Texts by Iranian Translators. *Journal of Applied Linguistics and Language Research*, 2(3), 238–252. <http://www.jallr.com/index.php/JALLR/article/view/57>
- James, C. (2013). Errors in Language Learning and Use : Exploring Error Analysis. *Errors in Language Learning and Use*. <https://doi.org/10.4324/9781315842912>
- Károly, A. (2012). Translation competence and translation performance: Lexical, syntactic and textual patterns in student translations of a specialized EU genre. *English for Specific Purposes*, 31(1), 36–46. <https://doi.org/10.1016/J.ESP.2011.05.005>
- Larsen-Freeman, D. (2013). *Teaching and Principles of Language Teaching* (3rd edition). Oxford handbooks for language teachers. In *Oxford, UK: Oxford University Press*.
- Na, P. P. Q. (2017). Errors In The Translation Of Topic-Comment Structures Of Vietnamese Into English - Asian EFL Journal: Asian EFL Journal. *The Asian EFL Journal*, 7(3), 1–30.

<https://www.asian-efl-journal.com/main-editions-new/errors-in-the-translation-of-topic-comment-structures-of-vietnamese-into-english/index.htm>

- Pym, A. (2018). Where Translation Studies lost the plot. *Translation and Translanguaging in Multilingual Contexts*, 4(2), 203–222. <https://doi.org/10.1075/TTMC.00010.PYM>
- Séguinot, C. (1989). Christina Schäffner (ed.). The Role of Discourse Analysis for Translation and in Translator Training. *La Traduction Au Canada: Tendances et Traditions*, 15(1), 243–244. <https://doi.org/https://doi.org/10.7202/006809ar>
- Silalahi, M., Rafli, Z., & Rasyid, Y. (2018). The Analysis of Errors in Translation of Scientific Text from English to Indonesian Language. *JETL (Journal of Education, Teaching and Learning)*, 3(1), 23–27. <https://doi.org/10.26737/jetl.v1i1.456>
- Soltani, F., Nemati, A., & Yamini, M. (2020). An Analysis Of Translation Errors In 5 Literary Genres based on American Translation Association (ATA) framework. *Cogent Arts & Humanities*, 7(1), 1799732. <https://doi.org/10.1080/23311983.2020.1799732>
- Tisgam, Khalida H. (2009). Is It Possible to Teach Translation in the Classroom. *Journal of College of Education For Women*. Vol.20(2), 528–551.
- Utomo, G. S. F. (2016). *An Analysis of Grammatical Errors in Translation Texts of Sixth Semester English Department Students Of Fbba In The Academic Year Of 2012/2013*.
- Waddington, C. (2001). Different Methods of Evaluating Student Translations: The Question of Validity. *Meta Journal Des Traducteurs Translators' Journal*, 46(2), 310–325. <https://www.erudit.org/en/journals/meta/2001-v46-n2-meta159/004583ar.pdf>
- West, R. (1994). Needs Analysis In Language Teaching. *Language Teaching*, 27(1), 1–19. <https://doi.org/10.1017/S0261444800007527>
- Wongranu, P. (2017). Errors in Translation Made by English Major Students: A Study on Types And Causes. *Kasetsart Journal of Social Sciences*, 38(2), 117–122. <https://doi.org/10.1016/J.KJSS.2016.11.003>
- Wu, Y., et al. (2016). *Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation*. <https://arxiv.org/abs/1609.08144v2>
- Zainudin, I. S., & Awal, N. M. (2012). Translation Techniques: Problems and Solutions. *Procedia - Social and Behavioral Sciences*, 59, 328–334. <https://doi.org/10.1016/J.SBSPRO.2012.09.282>
- Zhao, C. (2013). Analysis of Mongolian Students' Common Translation Errors and Its Solutions. *English Language Teaching*, 6(3), p78. <https://doi.org/10.5539/ELT.V6N3P78>