A COMPARATIVE ANALYSIS OF ACCURACY BETWEEN GOOGLE TRANSLATE AND BARD IN TRANSLATING ABSTRACTS OF SCIENTIFIC JOURNALS

Rosita[1], Baharuddin[2], Lalu Jaswadi Putera[3], I Made Sujana[4]
English Education Program, Faculty of Teacher Training and Education,
University of Mataram
[1]piedpieper22@gmail.com, [2]bahar@unram.ac.id, [3]elputra@unram.ac.id
[4]car_msujana@unram.ac.id

ABSTRACT

The accuracy of machine translation is still debatable. This study aims to compare the accuracy of Google Translate and BARD in translating scientific journal abstracts. The main focus is to evaluate the extent to which these two platforms can preserve the meaning and quality of translation in the abstracts of Adabiyyāt journal. This research utilizes two main theories in evaluating translation accuracy, namely the Human-mediated Translation Edit Rate (HTER) by Snover (2006) to identify translation errors and the Translation Quality Index (TQI) by Schiaffino & Zearo (2005) to calculate overall accuracy results. The results of this study show that GT is proven to be more accurate than Bard in translating abstracts from scientific journals. This can be seen from the final result of GT's calculation, which is greater than that of Bard. GT scored 99,7% in total, while Bard only got 99,1%.

Keywords: translation, accuracy, google translate, bard

ABSTRAK

Keakuratan terjemahan mesin masih menjadi perdebatan. Penelitian ini bertujuan untuk membandingkan keakuratan Google Translate dan BARD dalam menerjemahkan abstrak jurnal ilmiah. Fokus utamanya adalah untuk mengevaluasi sejauh mana kedua platform ini dapat mempertahankan makna dan kualitas terjemahan dalam abstrak jurnal Adabiyyāt. Penelitian ini menggunakan dua teori utama dalam mengevaluasi keakuratan terjemahan, yaitu Human-mediated Translation Edit Rate (HTER) oleh Snover (2006) untuk mengidentifikasi kesalahan terjemahan dan Translation Quality Index (TQI) oleh Schiaffino dan Zearo (2005) untuk menghitung hasil keakuratan secara keseluruhan. Hasil dari penelitian ini menunjukkan bahwa GT terbukti lebih akurat dibandingkan Bard dalam menerjemahkan abstrak jurnal ilmiah. Hal ini terlihat dari hasil akhir perhitungan GT yang lebih besar dibandingkan Bard. GT memperoleh nilai total 99,7%, sedangkan Bard hanya memperoleh 99,1%.

Kata Kunci: penerjemahan, akurasi, google translate, bard

A. Introduction

Translation is essential in communication, serving as a means to convey knowledge and information, acting as a connecting link between individuals who speak diverse languages and come from various cultural backgrounds.

The act of translation in Indonesia is employed by various segments of society, including students in schools or universities, teachers/lecturers, and employees (Sumiati, et al., 2022). Baharuddin, et al. (2021) said that translation is the act of transferring a message from one language to another in an effort to re-express the message from the source language into the target language with the same content.

Translation is a means of communication that allows individuals to share information and society to gain knowledge or information through translation. Today, many scientific works, such as thesis, dissertations, papers, reports, journals and articles originating from English, have been translated into various languages, including Indonesian, and one of the important parts of these works is the abstract.

According to Fitria (2021), an abstract is a brief summary of the content of a scientific work. Furthermore, Fitria (2018) also revealed that abstract translators need to have a strong understanding of translation structures and techniques. Hence, some people choose to use machine translators as they have difficulty in translating texts.

In today's technologically advanced era, language translation has become an easy thing to do. With translation tools such as Google Translate and Google Bard, getting an accurate translation is now a quick and cost-effective process. However, careful consideration is needed in choosing the most appropriate translation engine for a particular source language (Sasmi, et al., 2023).

According to Wardana, L.A., et al. (2022) machine translation is a translation performed by a machine with a formula or formulas that have been entered into the program to assist the translation. Although they are both from Google, Google Translate and Bard have quite noticeable differences. Google Translate can translate text into numerous languages worldwide, it is user-friendly, efficient, and can be conveniently accessed through smartphones and similar technologies, saving time for users (Sumiati, et al., 2022). It is capable of translating words, phrases, and web pages (Brahmana, et al., 2020). While Bard is Google AI's large language model (LLM) trained with massive text and code datasets (SIAD, 2023).

Google Translate and Bard both offer various advantages when it

comes to translating. However, like any other machine, Google Translate and Bard certainly have weaknesses, including in terms of accuracy. Google Translate has limitations in translating complex sentences, sometimes still producing inaccurate translations (Arba, et al., 2023). Meanwhile, Bard's accuracy in translating text depends on various factors, including text complexity, language style, and context (Google's Bard, 2023). The accuracy aspect is the basic thing in evaluating the quality of translation. To ensure the accuracy of the translation, the original text must be conveyed correctly, and the target text must have the same meaning as the original text (Sanusi, 2019).

Due to this phenomenon, the researcher is interested to conduct research focused on comparing Google Translate and Bard in a term of accuracy when translating abstracts of scientific journals, namely the journal Adabiyyāt volume 6 number 2 in 2022, and volume 7 number 1 in 2023.

B. Research Method

Qualitative descriptive research was used in this study. Qualitative research is a method of scientific observation that collects non-numerical data (Maxfield, et al., 2014). This approach is used in this study to examine the translation accuracy of Google Translate and Bard when processing scientific journals from Adabiyyāt. Using Google Translate and Bard, the researcher will translate the abstract texts of the journal articles from Adabiyyāt: Journal of Language and Literature (https://ejournal.uin-suka.ac.id/adab/Adabiyyat) from Indonesian to English. Adabiyyāt is an accredited journal with a rating of 2 by Sinta (Science and Technology Index). Afterwards, the researcher examined the translated texts from the two translation engines. The data collection conducted by the researchers included 11 abstracts from the journal Adabiyyāt. The researchers used HTER and TQI to answer the research problem. To collect data, this study used purposive sampling method. To analyze the data, this study used several steps: translating samples from Indonesian to English using Google Translate and Bard, analyzing translation errors using HTER theory by Snover (2006). The equation for the HTER score, where SUB (*substitutions*), INS (*insertions*), DEL (*deletions*), and SHIFT (*shift*) are the number of substitutions, insertions, deletions and

shifts, respectively, and N (*number*) is the average number of reference words, is shown in the following equation.

$$HTER = \frac{SUB + INS + DEL + SHIFT}{N}$$

For example, if a translated text makes 10 errors in the average of 100 words reference text. It means that:

$$HTER = \frac{10}{100}$$

$$HTER = 0{,}1$$

Then, calculating accuracy results using TQI theory by Schiaffino and Zearo (2005). For example, if the calculated error made by Google Translate and Bard using the HTER formula is 0.1, then:

$$0{,}1 \times 100 = 10$$

$$100 - 10 = 90$$

Schiaffino and Zearo (2005) Translation Quality Index (TQI) ranging from Negative (0), Poor (1-49), Low (50-59), Improvable (60-69), Average (70-79), Good (80-89), to Excellent (90-100). This quality measurement is used to interpret HTER scores into descriptive translation quality.

So, the result of GT or Bard's accuracy in translating political science scientific text is 90, which falls within the Excellent category according to the TQI.

C. Research Finding

This study investigates the translation errors of Google Translate and Bard in translating scientific journal abstracts from Adabiyyat using HTER theory and TQI theory to measuring the accuracy. The researcher used the Indonesian version of the abstracts as the source text and the published version as the human translation. The samples themselves consist of the 2 most recent volumes of the journal Adabiyyat, namely volume 2 in 2022 and volume 1 in 2023. Volume 2 in 2022 has 6 articles and volume 1 in 2023 has 5 articles. Therefore, a total of 11 abstracts were analyzed.

Table 1 Types of Errors Made by GT and Bard

No.	Types of Error	GT	Bard	Quantity
1.	Deletion (DEL)	1	5	6
2.	Substitution (SUB)	3	9	12
3.	Insertion (INS)	3	4	7
4.	Shift (Shift)	-	-	-
Total		**7**	**18**	**25**

Based on the table above, there are 25 errors out of 4 types of translation errors found in the articles in Adabiyyat journal in volume 2 in 2022 and volume 1 in 2023. These

types are deletion, substitution, insertion, and shift. In GT's translation, there is 1 deletion type errors, 3 types of substitution type errors, and 3 types of insertion type errors. Meanwhile, in Bard's translation, there are 5 types of deletion errors, 9 types of substitution errors, and 4 insertion errors.

The most common error found was substitution which appeared 12 times out of 11 article abstracts. This error is characterized by changes in the translation that do not match the intent of the source text. In the GT translation, the total number of errors of this type was 3, and 9 in the Bard translation.

Here is the example of substitution errors found in the samples:

Appendix 3, text 3 (Vol. VI, No. 2, December 2022, p. 186-208):

> **ST:** *"Kedua, berdasarkan mitos baru, ḥumur 'keledai' dalam Surat al-Muddaṡṡir ayat ke 50-51 menggambarkan fanatisme kabilah dan kadar pertentangan orang Quraisy terhadap dakwah Nabi Muhammad saw, bukan pada simbol kebodohan yang ada pada mitos lama."* (7th sentence)
>
> **GT's translation:** "Second, based on the new myth, the age of the 'donkey' in Surah al-Muddaṡṡir verses 50-51 describes the fanaticism of the tribe and the level of opposition of the Quraish people to the preaching of the Prophet Muhammad, not the symbol of ignorance in the old myth." (7th sentence)

As shown in appendix 3, text 3 in the 7th sentence. There is a translation error marked by the bolded and underlined word, which is *"humur"* in ST and "the age of" in GT's translation. The word is categorized as an error because *"humur"* is translated as "the age of the". Whereas the *"humur"* referred to in the ST is the new myth in Surat al-Mudṡir verses 50-51.

The second most common errors found was insertion, with total of errors are 7. Insertion is characterized by the addition of words to the translation, while the added words are not found in the source text. In GT's translation, there are 2 errors found with this error. Meanwhile, in Bard's translation there are 4 error with this type.

Here is the example of insertion errors found in the samples:

Appendix 1, text 1 (Vol. VI, No. 2, December 2022, p. 140–164):

> **ST:** "Tulisan ini berpijak pada cerpen di Jawa Pos tahun 2021" (1st sentence)

Bard's translation: "This study is based on short stories **published** in Jawa Pos in 2021." (1st sentence)

As shown in the bolded and underlined word, the word "published" in appendix 1, text in the first sentence. The word is an insertion type translation error because there is no word related to *"diterbitkkan"* in the source text.

The third most common errors found was deletion, with total of error are 6. Deletion is characterized as an editing process by removing words that have incorrect meanings, class words, or redundant words. In GT's translation, 1 error were found. Meanwhile, 5 errors were found in Bard's translation.

Here is the example of insertion errors found in the samples:

Appendix 2, text 2 (Vol. VI, No. 2, December 2022, p. 165–185):

ST: *"Meskipun standar-standar itu begitu mainstream, nyatanya beberapa puisi penyair klasik tidak selalu sesuai dengan standar itu, bahkan puisi hasil gubahan penyair-penyair besar sekalipun."* (2nd sentence)

Bard's translation: "Although these standards are so mainstream, in reality some classical poems do not always conform to these standards, even poems composed by great poets." (2nd sentence)

As shown in appendix 2, text 2 in the second sentence. The word that is bolded and underlined in the ST, namely *"penyair",* is deleted in Bard's translation. Because the word is deleted, the meaning in ST also changes, whereas what is meant in ST is "*puisi penyair"* or "poet's poem", while Bard translates it just into "poems" or *"puisi".*

And lastly, there are no shift type errors in appendix 1 to 11.

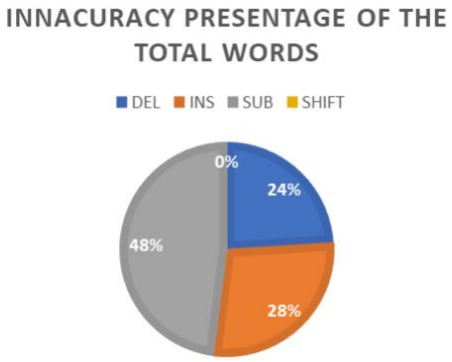

Chart 1 Inaccuracy Chart

A total of 1810 words were analyzed using Sover's theory. The translation errors were classified as follows: In the translations of GT and Bard, insertion errors were found in 7 of the samples studied, which is 28% of 100%. While

substitution type errors were found as many as 12 errors, which is 48%. The next error, deletion, is found as many as 6 errors, which is 24%. Meanwhile, shift type errors were not found in the samples or 0%.

D. Conclusion

This study found three types of translation errors in the abstracts, which means that not all types of errors are present in the abstracts based on the theory. The most common error is substitution which appears 3 times in GT and 9 times in Bard. The second most common error is insertion which appears 2 times in GT and 5 times in Bard. Meanwhile, the third most common error is deletion which appears 1 time in GT and 5 times in Bard. And finally, shift-type errors do not exist in GT and Bard translations.

In conclusion, the results of this study show that GT is proven to be more accurate than Bard in translating abstracts from scientific journals. This can be seen from the final result of GT's calculation, which is greater than that of Bard. GT scored 99,7% in total, while Bard only got 99,1%.

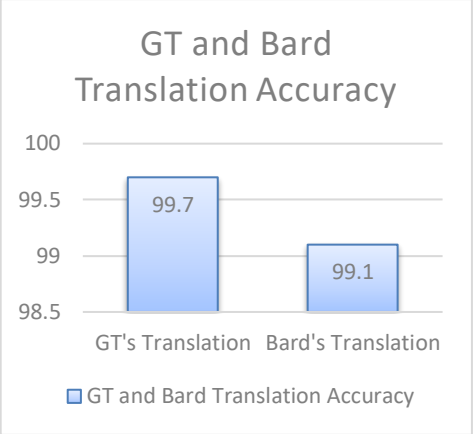

Chart 2 GT and BARD Translation Accuracy Chart

REFERENCES

Arba, N., Widyasari, W., Efendi, Y., & Syaputri, W. (2023). Analisa Hasil Terjemahan Google Translate Dalam Lirik Lagu “to the bone” Oleh Pamungkas. Jurnal Pembahsi (Pembelajaran Bahasa Dan Sastra Indonesia), 13(1), 55–67. https://doi.org/10.31851/pembahsi.v13i1.11874

Baharuddin, Baharuddin, et al. "Penerapan Teori Terjemahan pada Editing Hasil Terjemahan Google Translate pada Teks Akademik oleh Mahasiswa Universitas Mataram." Jurnal Ilmiah Profesi Pendidikan, vol. 6, no. 4, Dec. 2021, pp. 816-824, doi:10.29303/jipp.v6i4.390.

Baharuddin, B., Amin, M., Thohir, L., & Wardana, L. A. (2021). Penerapan Teori Terjemahan pada Editing Hasil Terjemahan Google Translate pada Teks Akademik oleh Mahasiswa Universitas Mataram. Jurnal Ilmiah Profesi Pendidikan, 6(4), 816-824.

Brahmana, C. R. S, Sofyan, R., & Putri, D. M. (2020). Problems in the

application of Google Translate as a learning media in translation. Language Literacy: Journal of Linguistics, Literature, and Language Teaching, 4(2), 384–389. https://doi.org/10.30743/ll.v4i2.2893

Brookshire, R. H., & Brundage, S. B. (2015). Writing Scientific Research in Communication Sciences and Disorder. Plural Publishing.

DAO, X.-Q. (2023). Which large language model should you use in Vietnamese education: Chatgpt, Bing Chat, or Bard? SSRN Electronic Journal. https://doi.org/10.2139/ssrn.4527476

Etikan, I., Musa, S. A., & Alkassim, R. S. (2016). Comparison of convenience sampling and purposive sampling. American journal of theoretical and applied statistics, 5(1), 1-4.

Fitria, T. N. (2018). Translation Techniques Found in English to Indonesian Abstract Translation of Journal Edunomika 2018. ELITE Journal Volume 05 Number 2

Fitria, T. N. (2020). Error Analysis of English Abstract in International Journal of Economics, Business and Accounting Research (IJEBAR). Al-Lisan: Jurnal Bahasa (e-Journal), 5.

Fitria, T. N. (2021). A review of machine translation tools: The translation's ability. Language Circle: Journal of Language and Literature, 16(1), 162–176. https://doi.org/10.15294/lc.v16i1.30961

Ghasemi, H., & Hashemian, M. (2016). A Comparative Study of" Google Translate" Translations: An Error Analysis of English-to-Persian and Persian-to-English Translations. English Language Teaching, 9(3), 13-17.

Google. (n.d.). Google. https://bard.google.com/?hl=in

Hasyim, M., Saleh, F., Yusuf, R., & Abbas, A. (2021). Artificial Intelligence: Machine Translation Accuracy in translating French-Indonesian culinary texts. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.3816594

Irfan, M. (2017). Machine Translation. Retrieved from https://www.researchgate.net/publication/320730405

Kadaoui, K., Magdy, S. M., Waheed, A., Khondaker, M. T. I., El-Shangiti, A. O., Nagoudi, E. M. B., & Abdul-Mageed, M. (2023). TARJAMAT: Evaluation of Bard and ChatGPT on Machine Translation of Ten Arabic Varieties. arXiv preprint arXiv:2308.03051.

Li, H., Graesser, A. C., & Cai, Z. (2014, May). Comparison of Google translation with human translation. In the twenty-seventh international flairs conference.

Maxfield, M. G., & Babbie, E. R. (2014). Research methods for criminal justice and criminology. Cengage Learning.

McDonald, S. V. (2020). Accuracy, readability, and acceptability in translation. Applied Translation. https://doi.org/10.51708/apptrans.v14n2.1238

Medvedev, Gennady. (2016). Google Translate in Teaching English. Journal of Teaching English for Specific and Academic Purposes, 4.1 (2016), 181–93

Mobaraki, M., & Aminzadeh, S. (2012). A study on different translation evaluation strategies to introduce an eclectic method. International Journal of English Linguistics, 2(6). https://doi.org/10.5539/ijel.v2n6p63

Munday, J. (2016). Introducing translation studies: theories and applications (4th ed.). London and New York: Routledge.

Munthe, I. B., Sipayung, K., & Lestari, F. D. (2023). Comparing The Translation Accuracy Between Google Translate and Professional Translator. Innovative: Journal Of Social Science Research, 3(6), 890-903.

Napitupulu, S. (2017). Analyzing Indonesian-English Abstracts Translation In View of Translation Errors By Google Translate. International Journal of English Language and Linguistics Research, 5, 15-23. European Center for Research Training and Development UK.

Peris, Á., Domingo, M., & Casacuberta, F. (2017). Interactive neural machine translation. Computer Speech & Language, 45, 201-220.

Sanusi, A. (2019). The quality translation analysis of student Qirāah text. Arabi: Journal of Arabic Studies, 4(1), 11. https://doi.org/10.24865/ajas.v4i1.119

Sari, A. N., & Zamzani, Z. (2020). An analysis of translation strategies of honorific term in the film “The boss baby.” Indonesian Journal of EFL and Linguistics, 5(2), 355. https://doi.org/10.21462/ijefl.v5i2.289

Sasmi, A.C., Baharuddin, Wardana, L.A. (2023). Google Nmt And Yandex Nmt Differences in Clarity and Naturalness Of Translating English Text Into Indonesian: A Comparative Study on Machine Translations (Doctoral dissertation, Universitas Mataram).

Schiaffino, R., & Zearo, F. (2005, November). Translation quality measurement in practice. In Proceedings of the 46th Annual Conference of the American Translation Association.

Siad, S. M. (2023). The Promise and Perils of Google's Bard for Scientific Research. HASTAC. Humanities Commons.

Snover, M., Dorr, B., Schwartz, R., Micciulla, L., & Makhoul, J. (2006). A Study of Translation Edit Rate with Targeted Human Annotation. In Proceedings of the 7th Conference of the Association for Machine Translation in the Americas: Technical Papers (pp. 223–231). Cambridge, Massachusetts, USA: Association for Machine Translation in the Americas.

Su, J., Zhang, X., Lin, Q., Qin, Y., Yao, J., & Liu, Y. (2019). Exploiting reverse target-side contexts for neural machine translation via asynchronous bidirectional decoding. Artificial Intelligence, 277, 103168.

Sugiyono, D. (2013). Metode penelitian pendidikan pendekatan kuantitatif, kualitatif dan R&D.

Sumiati, Baharuddin, & Saputra, A. (2022). The analysis of Google Translate Accuracy in translating procedural and narrative text. Journal of English Education Forum (JEEF), 2(1), 7–11. https://doi.org/10.29303/j.v2i1.270

Tanjung, Sufriati. (2018). Penialain Penerjemahan Jerman-Indonesia. Yogyakarta: Kanwa Publisher

UIN Sunan Kalijaga. (n.d.). Adabiyyat. E-Journal UIN Sunan Kalijaga Yogyakarta. https://ejournal.uin-suka.ac.id/adab/Adabiyyat

Vanjani, M., & Aiken, M. (2020). A comparison of free online machine language translators. J. Manag. Sci. Bus. Intell, 5, 26-31.

Wardana, L. A., Baharuddin, B., & Nurtaat, L. (2022). Kemampuan Mahasiswa melakukan post-editing terhadap Hasil Terjemahan Machine Translation. Jurnal Ilmiah Profesi Pendidikan, 7(1), 53-61.

Wu, Y., Schuster, M., Chen, Z., Le, Q. V., Norouzi, M., Macherey, W., ... & Dean, J. (2016). Google's neural machine translation system: Bridging the gap between human and machine translation. arXiv preprint arXiv:1609.08144.