

A COMPARATIVE ANALYSIS OF THE ACCURACY OF MACHINE TRANSLATION IN TRANSLATING ENGLISH TO INDONESIAN

Serli Kalo Raben¹, Normalia Sirande², Judith Ratu Tandi Arrang³,
Muhammad Ilham⁴

¹SMP Negeri 3 Makale, ^{2,3}Universitas Kristen Indonesia Toraja, ⁴Universitas
Ichsan Sidenreng Rappang

serlykara24@gmail.com

ABSTRACT

Machine translation has become increasingly popular, but its accuracy is still a subject of debate. This study aims to compare the accuracy of two popular machine translation tools, Google Translate and DeepL Translate, in translating English to Indonesian. This research applied a descriptive quantitative research methodology. The research design involves comparing the accuracy of Google Translate and DeepL Translate by analyzing the output translation texts and using narrative texts as the sample. The data collection procedure involves conducting a written test using data cards and a data sheet. The accuracy of the translations is calculated based on the percentage of accurately translated meanings. The findings of this research indicate that both Google Translate and DeepL Translate provide translations that are considered less accurate. This statement is based on the finding of accuracy occurrence. DeepL Translate outperforms Google Translate with a higher percentage. Specifically, DeepL Translate achieves an accuracy rate of 73.1%, while Google Translate lags slightly behind at 63.4%. This suggests that DeepL Translate demonstrates a better ability to produce accurate translations across the board. In conclusion, this research highlights the need for improvement in machine translation tools for translating English to Indonesian. Both Google Translate and DeepL Translate show limitations in accuracy, indicating the importance of continuing research and development in the field of machine translation.

Keywords: Comparative Study, Translation Quality, Machine Translation Platforms, Accuracy

INTRODUCTION

Translation is the process of converting written or spoken content from one language to another while preserving its meaning and intent. According to Catford (1965), translation is the replacement of a text in one language (the source language) with an equivalent text in another language (the target language). It encompasses not only the transfer of words but also cultural nuances and contextual meanings. In the realm of EFL learning, machine translation has emerged as a valuable tool, especially when used with human assistance, but novice learners should have good background knowledge to evaluate the translation result (Wibawa, 2020). While machine translation offers quick access to translations and assists in expanding vocabulary, its limitations in accurately capturing language complexity, contextual nuances, and idiomatic expressions persist (Bhabha, 1994; Chesterman, 1997;

TEFL Overseas Journal

Teaching English as a Foreign Language Journal

ISSN 2461-0240 (Print), 2828-9544 (Online)

Volume 13 Number 1 April 2025

Saunders & Toury, 1997). Addressing these challenges necessitates ongoing research to enhance machine translation accuracy through context-aware algorithms and linguistic theories. Despite its utility, machine translation should supplement, not replace, active language engagement to foster robust language skills development.

In the evolving landscape of language education, translation stands as a crucial conduit for navigating linguistic and cultural divides. As technology continues to advance, machine translation has emerged as a prominent tool, particularly in supporting language learners. This research is situated within the sphere of English language learning, with a specific focus on learners in Indonesia. Given the prevalence of machine translation platforms like Google Translate and DeepL Translate, there exists a pressing need to evaluate their accuracy and efficacy in translating English texts into Indonesian, a language of significant importance in the region. This study aims to delve into the accuracy of machine translations, shedding light on their strengths and limitations in accurately conveying meaning from English to Indonesian.

Several studies have contributed to the discourse surrounding machine translation accuracy and its implications for language education. Nasution (2022) explored the translation quality of Google Translate in website localization, focusing on its relevance for language learning. Kasperaviciene et al. (2020) conducted a study assessing the quality of machine translation output, comparing the performance of Google and Microsoft translation applications. Xie (2021) examined the quality of English-Chinese machine translation systems, shedding light on their strengths and weaknesses. Sumiati et al. (2022) analyzed Google Translate's accuracy in translating different text types, offering insights into its performance across diverse linguistic contexts. (Hasyim et al., 2021) investigated Google Translate's accuracy in translating culinary texts, highlighting its limitations in conveying cultural nuances. Farahani (2020) and Farahsani et al. (2021) conducted studies evaluating Google Translate's accuracy in translating specialized language, particularly mechanical engineering terminology. These studies collectively contribute to our understanding of machine translation accuracy and its implications for language learning, providing valuable insights into its performance across various contexts and language pairs.

Existing studies have highlighted several challenges and limitations in machine translation, including issues with idiomatic expressions, cultural nuances, context awareness, and accuracy in specialized domains such as mechanical engineering and culinary texts. Additionally, the reliance on machine translation in educational settings, such as English language learning courses, raises questions about its effectiveness and the need for human assistance in interpreting translation results. Therefore, there is a need for comprehensive research to assess the accuracy of machine translation, particularly in translating English to Indonesian, and to identify areas for improvement.

This research conduct a comparative analysis of the accuracy of machine translation, specifically Google Translate and DeepL Translate, in translating English texts to Indonesian. The research aims to assess the level of accuracy in terms of translation fidelity, including indicators such as addition, omission,

different meaning, and zero meaning. By identifying the strengths and weaknesses of each machine translation platform, the research seeks to provide valuable insights for language learners, educators, and users of machine translation services.

This research contribute to enhancing understanding of machine translation accuracy, particularly in the context of English to Indonesian translation. By comparing the performance of Google Translate and DeepL Translate, the research will identify areas for improvement in machine translation algorithms and techniques. Language learners and educators can benefit from insights into the limitations of machine translation and the importance of human involvement in the translation process. Furthermore, the research findings can inform the development of more effective machine translation tools and strategies for improving translation quality in educational and professional settings.

METHODS

The research design employed a quantitative descriptive approach, focusing on describing and summarizing data without manipulating variables (Morgan et al., 1999). Conducted at the Christian University of Indonesia Toraja, the study aimed to compare the accuracy levels of Google Translate and DeepL Translate in translating English texts into Indonesian. Purposive sampling was used to select diverse English texts for translation, reflecting various linguistic complexities and contextual nuances. A written test served as the research instrument, evaluating translations based on predefined criteria such as addition, omission, different meaning, and zero meaning (Rivera-Trigueros, 2022). The data collection procedure involved selecting test samples, translating texts using both translation systems, evaluating translations, recording results, and quantitatively analyzing data to compare accuracy levels. Through systematic analysis of specific indicators of inaccuracy, including addition, omission, different meaning, and zero meaning, the study aimed to provide insights into the strengths and limitations of machine translation systems in the context of English to Indonesian translation (Mangatur Nababan et al., 2012).

FINDINGS AND DISCUSSION

In translation, accuracy means that the message of the source text is preserved in the target texts. Here, accuracy is related to the accurate message. The meaning of the source text as the main concern in translation should be transferred accurately in the target text. A text can be regarded as accurate if there are more than 50% occurrences of accurate sentences in the text. On the contrary, it is regarded as inaccurate if there are less than 50% occurrences of accurate sentences in the text. To determine whether a text is accurate or not, there are four indicators to use: omission, addition, different meaning, and zero meaning.

As described in the methodology chapter, the data for this study came from analyzing machine translations of the short story "The Little Match Girl" produced by Google Translate and DeepL Translate from English to Indonesian. The analysis focused on evaluating the accuracy level based on the four inaccurate translation indicators proposed by Larson (1984): addition, omission, different meaning, and

zero meaning. A total of 41 sentences from the story were input into both machine translation tools and their outputs were carefully analyzed and compared to the human translation.

Table 4. 1 The Occurrence of Accurate and Inaccurate Sentences

Accuracy		Google Translate	Percentage	DeepL Translate	Percentage
Acc		26	63.4 %	30	73.1 %
Inaccurate	Add	0	0 %	0	0 %
	Om	6	14.6 %	1	2.43 %
	Dm	9	21.9 %	10	24.3 %
	zm	0	0 %	0	0 %
Total		41	100 %	41	100 %

In terms of overall accuracy, DeepL Translate outperforms Google Translate with a higher percentage. Specifically, DeepL Translate achieves an accuracy rate of 73.1%, while Google Translate lags slightly behind at 63.4%. This suggests that DeepL Translate demonstrates a better ability to produce accurate translations across the board. However, it's essential to delve into the specific categories to understand where each translation engine excels or falls short.

In the evaluation presented in the table, both Google Translate and DeepL Translate show zero occurrences of additions. This suggests that both translation engines maintain fidelity to the original text in terms of content, without introducing any extraneous information during the translation process. Thus, users can rely on the translations provided by both Google Translate and DeepL Translate to accurately convey the intended meaning without unnecessary additions.

The omission category indicates instances where the translated text lacks certain elements present in the original text. In the evaluation, Google Translate demonstrates six instances of omissions, accounting for 14.6% of the total evaluated translations. On the other hand, DeepL Translate only exhibits one omission, representing a mere 2.43% of the total. This suggests that DeepL Translate is more adept at preserving the completeness of the original text during translation compared to Google Translate, as evidenced by its lower percentage of omissions.

Data 8

SL: Nobody had bought anything of her the whole livelong day; no one had given her a single farthing.

GT: *Tidak ada yang membelikannya apa pun sepanjang hari; tidak ada seorang pun yang memberinya satu sen pun.*

DT: *Tidak ada yang membelikannya sesuatu sepanjang hari itu; tidak ada yang memberinya uang sepeser pun.*

In data 8, both translations aim to convey the main essence of the source sentence, which describes nobody buying anything from her the whole day and not receiving a single farthing. But GT and DT translates the phrase "Nobody had bought anything of her" as "*Tidak ada yang membelikannya sesuatu/apa pun*", which there is a slight deviation in the phrasing, as the translation does not explicitly mention "of her" as in the source sentence that indicate omission in translation. The phrase "Nobody had bought anything of her" in the original text can indeed be more accurately translated as "*Tidak seorang pun yang membeli sesuatu/apapun darinya*" in Indonesian. This translation better reflects the intended meaning of the original text, indicating that nobody had purchased anything from the person being referred to.

Different meaning refers to cases where the translated text conveys a meaning different from that of the original text. Google Translate and DeepL Translate both exhibit instances of different meanings in their translations. Google Translate shows nine instances of different meanings, representing 21.9% of the total evaluated translations. DeepL Translate, meanwhile, demonstrates ten instances of different meanings, accounting for 24.3% of the total. While both translation engines produce translations with altered meanings in some cases, DeepL Translate has a slightly higher percentage of such instances compared to Google Translate.

Data 5

SL: One slipper was nowhere to be found; the other had been laid hold of by an urchin, and off he ran with it; he thought it would do capitally for a cradle when he someday or other should have children himself.

GT: *Satu sandal tidak ditemukan; yang lainnya telah ditangkap oleh seekor bulu babi, dan dia pun lari bersamanya; dia pikir itu akan sangat bermanfaat bagi sebuah buaian ketika dia suatu hari nanti akan memiliki anak sendiri.*

DT: *Satu sandal tidak dapat ditemukan; sandal yang satunya lagi dipegang oleh seekor landak, dan dia berlari dengan sandal itu; dia pikir sandal itu akan sangat berguna sebagai buaian ketika dia suatu hari nanti memiliki anak.*

In data 5, both Google Translate and DeepL Translate not accurately capture the meaning of the source sentence, the translations of "*bulu babi*" by Google Translate and "*landak*" by DeepL Translate are inaccurate in this context. However, the term "urchin" in the original text actually refers to a mischievous child, not an animal. Therefore, more accurate translations would use terms such as "*anak nakal*" or "*anak jalanan*" to convey the intended meaning of "urchin".

Zero meaning pertains to instances where the translated text lacks any meaningful content or fails to convey the intended meaning of the original text. In the evaluation provided in the table, both Google Translate and DeepL Translate show no occurrences of zero meaning, indicating that neither translation engine produces entirely nonsensical or meaningless translations. This suggests that users

can generally expect both Google Translate and DeepL Translate to provide translations that maintain some degree of coherence and relevance to the original text, even though they may contain inaccuracies or differences in meaning.

The results of this study highlight significant insights into the accuracy of machine translation tools—Google Translate (GT) and DeepL Translate (DT)—in rendering literary texts from English to Indonesian. By evaluating the translation outputs of 41 sentences from *The Little Match Girl* using the accuracy framework proposed by Larson (1984), which includes the indicators of omission, addition, different meaning, and zero meaning, the study provides a nuanced understanding of how each tool performs.

CONCLUSION

This research study aims at finding out the most accurate between Google Translate and DeepL Translate in translating English to Indonesian. The result shows that both DeepL Translate in English Indonesia are considered as less accurate translation. This statement is based on the finding of accuracy occurrence DeepL Translate outperforms Google Translate with a higher percentage. Specifically, DeepL Translate achieves an accuracy rate of 73.1%, while Google Translate lags slightly behind at 63.4%. This suggests that DeepL Translate demonstrates a better ability to produce accurate translations across the board. The research found that while machine translation offers a convenient and accessible tool for language learning and translation purposes, it still faces challenges in ensuring translation accuracy. The inherent complexity of language, lack of context awareness, and difficulties in capturing idiomatic expressions and cultural nuances pose hurdles for machine translation systems.

REFERENCE

- Bhabha, H. K. (1994). *The Location of Culture* (Routledge Classics). In *The location of culture*.
- Catford, J. C. (1965). *Language and Language Learning a Linguistic Theory of Translation*. Oxford University Press, 110.
- Chesterman, A. (1997). *Memes of Translation: The Spread of Ideas in Translation Theory* (Benjamins Translation Library). Amsterdam & Philadelphia: Benjamins, 22.
- Farahani, M. V. (2020). Adequacy in Machine vs. Human Translation: A Comparative Study of English and Persian Languages. *Applied Linguistics Research Journal*. <https://doi.org/10.14744/alrj.2020.98700>
- Farahsani, Y., Rini, I. P., & Jaya, P. H. (2021). Google Translate Accuracy in Translating Specialized Language From English to Bahasa Indonesia: A Case Study on Mechanical Engineering Terminology. *International Conference on Sustainable Innovation 2020–Social, Humanity, and Education*, 518(1).
- Hasyim, M., Saleh, F., Yusuf, R., & Abbas, A. (2021). Artificial Intelligence: Machine Translation Accuracy in Translating French-Indonesian Culinary

TEFL Overseas Journal

Teaching English as a Foreign Language Journal

ISSN 2461-0240 (Print), 2828-9544 (Online)

Volume 13 Number 1 April 2025

Texts. *International Journal of Advanced Computer Science and Applications*, 12(3). <https://doi.org/10.14569/IJACSA.2021.0120323>

Kasperaviciene, R., Motiejuniene, J., & Patašiene, I. (2020). Quality assessment of machine translation output: Cognitive evaluation approach in an eye tracking experiment. *Texto Livre*, 13(2). <https://doi.org/10.35699/1983-3652.2020.24399>

Larson, M. L. (1984). Meaning-based translation: A guide to cross-language equivalence. *University Press of America*.

Mangatur Nababan, R., Ardiana Nuraeni, S., & Sumardiono. (2012). The translation of cultural terms in the novel “The Da Vinci” into Indonesian. *International Journal of Linguistics*, 4(2), 1–16.

Morgan, G. A., Gliner, J. A., & Harmon, R. J. (1999). Quantitative research approaches. *Journal of the American Academy of Child and Adolescent Psychiatry*, 38(12). <https://doi.org/10.1097/00004583-199912000-00022>

Nasution, D. K. (2022). Machine Translation in Website Localization: Assessing its Translation Quality for Language Learning. *AL-ISHLAH: Jurnal Pendidikan*, 14(2). <https://doi.org/10.35445/alishlah.v14i2.1308>

Rivera-Trigueros, I. (2022). Machine translation systems and quality assessment: a systematic review. *Language Resources and Evaluation*, 56(2). <https://doi.org/10.1007/s10579-021-09537-5>

Saunders, G. E., & Toury, G. (1997). Descriptive Translation Studies and beyond. *Language*, 73(2). <https://doi.org/10.2307/416056>

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). <https://doi.org/10.29303/j.v2i1.270>

Wibawa, A. P. (2020). *Invited Speech 2 : An Overview Of Machine Translation*. <https://doi.org/10.1109/ic2ie47452.2019.8940876>

Xie, S. (2021). A Comparative Study on the Quality of English-Chinese Machine Translation in the Era of Artificial Intelligence. *ACM International Conference Proceeding Series*. <https://doi.org/10.1145/3495018.3495378>