EXCELLENCE Journal of English and English Education
Volume 5 | Number 1 | June | 2025
DOI: 10.47662/ejeee.v5i1.1192
p-ISSN: 2807-6389
e-ISSN: 2807-2103
English Education Study Program, FKIP UNIVA Medan

A Comparative Study of Google Translate and DeepL in Translating Indonesian to English Among Informatics Engineering Students at Universitas Potensi Utama Medan

Dwi Suci Amaniarsih
English Study Program, Education and Social Science Faculty, Universtias Potensi Utama, Medan
Email: amaniarsih86@gmail.com

ABSTRACT

This study aims to compare the accuracy and effectiveness of Google Translate and DeepL in translating Indonesian texts into English, focusing on Informatics Engineering students at Universitas Potensi Utama. In an era where machine translation tools have become essential in academic and professional settings, understanding their strengths and limitations is crucial. Grounded in Newmark's translation theory and supported by insights from neural machine translation research, this study employs a mixed-method approach, combining quantitative scoring of translation outputs and qualitative analysis of linguistic accuracy, context retention, and semantic coherence. The research findings reveal that DeepL generally outperforms Google Translate in preserving contextual meaning and producing grammatically accurate sentences, although Google Translate remains more familiar and accessible to students. The statistical analysis shows a significant difference in translation quality between the two tools ($p < 0.05$), while student interviews highlight DeepL's advantage in handling nuanced or academic texts. This research provides valuable insights for educators and learners in choosing appropriate translation tools, and offers implications for the integration of machine translation in language learning and academic writing.

Keywords: machine translation, Google Translate, DeepL, translation accuracy, Indonesian-English translation, neural machine translation, informatics students

ABSTRAK

Penelitian ini bertujuan untuk membandingkan akurasi dan efektivitas Google Translate dan DeepL dalam menerjemahkan teks bahasa Indonesia ke dalam bahasa Inggris, dengan fokus pada mahasiswa Jurusan Teknik Informatika Universitas Potensi Utama. Di era di mana alat penerjemah mesin telah menjadi bagian penting dalam konteks akademik dan profesional, pemahaman terhadap kelebihan dan kekurangannya menjadi hal yang krusial. Berlandaskan teori penerjemahan dari Newmark serta kajian terbaru tentang neural machine translation, penelitian ini menggunakan pendekatan mix method yang menggabungkan penilaian kuantitatif terhadap hasil terjemahan dan analisis kualitatif terhadap akurasi linguistik, ketepatan konteks, dan koherensi makna. Hasil penelitian menunjukkan bahwa DeepL secara umum lebih unggul dibandingkan Google Translate dalam mempertahankan makna kontekstual serta menghasilkan struktur kalimat yang lebih akurat secara tata bahasa, meskipun Google Translate masih lebih dikenal dan mudah diakses oleh mahasiswa. Analisis statistik menunjukkan adanya perbedaan signifikan dalam kualitas terjemahan antara kedua alat tersebut ($p < 0.05$), sementara wawancara dengan mahasiswa menyoroti keunggulan DeepL dalam menangani teks bernuansa akademik. Penelitian ini memberikan wawasan penting bagi pendidik dan pelajar dalam memilih alat penerjemah yang tepat, serta menawarkan implikasi bagi integrasi penerjemahan mesin dalam pembelajaran bahasa dan penulisan akademik.

Kata kunci: *penerjemahan mesin, Google Translate, DeepL, akurasi terjemahan, terjemahan Indonesia-Inggris, neural machine translation, mahasiswa informatika*

A. Introduction

In the increasingly globalized academic and technological environment, English has become the dominant language for scholarly communication, particularly in the fields of science, technology, and engineering. For students of Informatics Engineering in non-English-speaking countries such as Indonesia, the ability to read and produce academic texts in English is a crucial skill. However, many students still face difficulties in mastering academic English, especially in technical writing and understanding domain-specific literature. In response to this challenge, machine translation (MT) tools such as Google Translate and DeepL have become essential aids in facilitating the translation of academic content from Bahasa Indonesia into English.

Google Translate has long been the most popular MT tool globally due to its wide availability, broad language support, and continuous development by Google's AI and language teams. In contrast, DeepL, a relatively newer tool developed by a German company, has gained attention in recent years for producing translations that are considered more fluent and natural in several language pairs. Both tools utilize Neural Machine Translation (NMT) technology, which leverages deep learning to capture contextual relationships and generate more accurate and coherent translations compared to earlier rule-based or statistical approaches.

Despite their popularity and potential, the effectiveness of these tools in handling translation between Bahasa Indonesia and English remains underexplored. Most research on MT tools has focused on European language pairs, such as English–German or English–French, and only a limited number of studies have evaluated their performance for Indonesian texts, especially in technical or academic domains. Furthermore, while Google Translate has been widely adopted and studied in the Indonesian academic context, DeepL's performance, particularly among students in STEM disciplines, has yet to be rigorously examined.

This study seeks to fill that gap by analyzing and comparing the quality of translations produced by Google Translate and DeepL when applied to Bahasa Indonesia texts, particularly within the academic context of Informatics Engineering. The focus is not only on linguistic accuracy but also on how students perceive and evaluate these tools in terms of usability, reliability, and effectiveness in supporting their academic tasks. By centering this study on the experiences of Informatics Engineering students at Universitas Potensi Utama, the research provides insight into the practical application of MT tools in higher education, particularly within technical fields where accurate translation of specialized terminology is essential.

To address these issues, this research is guided by the following problems:

1. How do Google Translate and DeepL differ in terms of translation quality—particularly grammatical accuracy, lexical appropriateness, and contextual coherence—when translating Indonesian texts into English?
2. What are the perceptions and preferences of Informatics Engineering students at Universitas Potensi Utama regarding the use of Google Translate and DeepL for academic translation tasks?

Based on the research problems above, the objectives of this study are:

- To compare the translation performance of Google Translate and DeepL in translating academic and technical texts from Bahasa Indonesia into English.
- To explore and analyze students' perceptions and preferences in using Google Translate and DeepL as tools for academic translation.

The significance of this study lies in its contribution to both academic research and educational practice. Firstly, it addresses the lack of empirical research on MT tools for the Indonesian–English language pair, especially within the domain of Informatics. Secondly, it provides insights that can be

used by educators and curriculum developers to integrate translation literacy into technical education. Lastly, the study highlights the importance of critical awareness in using AI-powered language tools, enabling students to become more effective and responsible users of technology in their academic and professional lives.

Neural Machine Translation (NMT) represents a significant advancement in automated translation technology. Introduced by Bahdanau et al. (2014), NMT systems use deep learning models, particularly recurrent neural networks (RNNs) and attention mechanisms, to translate entire sentences as single units rather than translating word by word. This enables the system to better handle complex sentence structures and contextual dependencies. Unlike earlier models such as Statistical Machine Translation (SMT), NMT achieves higher fluency and semantic accuracy, which has led to its adoption by major MT providers like Google and DeepL.

Google Translate transitioned to NMT in 2016 and has since expanded its capabilities to include more than 100 languages. It uses a proprietary model that combines massive datasets and neural networks to deliver context-aware translations. However, the quality of translation still varies significantly depending on the language pair, the domain of the text, and the complexity of the source material (Wu et al., 2016).

DeepL, launched in 2017, also uses NMT but has been praised for producing translations that are more idiomatic and natural-sounding, particularly in European languages. DeepL uses what it calls a "Linguee-based" approach, leveraging parallel corpora curated from multilingual dictionaries and aligned sentence pairs. According to Toral et al. (2018), DeepL often outperforms Google Translate in human evaluations, especially in fluency and style, although the range of supported languages is more limited.

Several studies have explored the use of machine translation in education. Clifford, Merschel, and Munné (2013) found that MT tools can enhance second language acquisition when used critically and reflectively. Similarly, Niño (2009) argues that MT can be integrated into language learning as a support tool, particularly for translation practice and vocabulary expansion. However, over-reliance on MT tools may lead to reduced motivation to learn grammar and may reinforce incorrect linguistic patterns if outputs are not carefully edited (O'Neill, 2019).

In the Indonesian context, studies such as those by Lestari and Ardi (2020) have shown that students use Google Translate extensively for academic writing. While it assists with vocabulary and sentence construction, issues remain in terms of syntactic accuracy and appropriateness of register. There is limited empirical research, however, that examines how MT tools perform in technical fields such as Informatics Engineering, where terminological precision is crucial.

Comparative studies on MT tools are growing but are still limited when it comes to non-European languages. Research by Castilho et al. (2018) compared Google Translate, DeepL, and Microsoft Translator in various language pairs and domains, concluding that DeepL often performed better in terms of fluency and human-likeness. However, the research mostly focused on English–German and English–French texts.

In a more recent study, Roesler (2021) explored translation outputs between English and Japanese using both tools and found significant variation in how each tool handled idiomatic expressions and honorifics. There remains a need for such comparisons in other less-resourced languages, including Bahasa Indonesia, especially considering its unique syntactic structure and morphological characteristics.

This study seeks to expand this line of inquiry by evaluating how Google Translate and DeepL handle translations from Indonesian to English in a technical academic context. It contributes not only to the comparative literature on MT tools but also to the understanding of how students in engineering-related disciplines interact with and assess the outputs of these technologies.

B. Research Method

Research method describes the This study adopts a mixed-methods research approach, combining both quantitative and qualitative methods to provide a comprehensive analysis of the translation quality and user perceptions of two widely used machine translation tools: Google Translate and DeepL. The rationale for this design lies in the need to evaluate not only the objective performance of the tools in terms of linguistic accuracy, but also to understand the subjective experiences of users—students of Informatics Engineering—who rely on these tools for academic purposes.

The research was conducted at Universitas Potensi Utama, focusing on a sample of 40 undergraduate students from the Informatics Engineering Department. These participants were selected through purposive sampling based on their frequent engagement with English-language academic texts, particularly in activities such as reading scientific literature, completing coursework, or translating project reports. All participants were native speakers of Bahasa Indonesia with at least an intermediate level of English proficiency, as confirmed by institutional placement assessments.

Data collection was carried out in two main phases. In the first phase, three types of source texts in Bahasa Indonesia were used: a general academic text, a technical text related to informatics, and a narrative or descriptive passage. Each text, approximately 150 words in length, was translated into English using both Google Translate and DeepL, producing a total of six translated outputs. These translations were then evaluated by three independent raters, consisting of experienced English lecturers and certified translators. The evaluation criteria included grammatical accuracy, lexical appropriateness, contextual and semantic coherence, and overall fluency, based on a rubric adapted from the American Translators Association (ATA). Inter-rater reliability was measured using Cohen's Kappa to ensure consistency and objectivity across evaluations.

The second phase of data collection involved administering a structured questionnaire to all participants. The questionnaire was designed to gather information about students' usage patterns, perceptions of translation quality, tool preferences, and levels of trust in each MT system. It contained a combination of Likert-scale items and open-ended questions. To complement the survey findings and explore the underlying reasons behind students' preferences and behaviors, semi-structured interviews were conducted with eight selected participants. These interviews provided richer qualitative insights into how students interact with translation tools, how they interpret the outputs, and the extent to which they rely on such tools in their academic life.

Quantitative data obtained from the translation evaluations and survey responses were analyzed using SPSS 26.0 software. Statistical techniques such as descriptive statistics and paired-sample t-tests were employed to compare the translation performance of the two tools. Meanwhile, qualitative data from open-ended responses and interviews were analyzed thematically using a coding technique to identify recurring patterns and insights.

To ensure the validity of the research instruments, expert judgment and pilot testing were conducted before the main study. The source texts were reviewed by language professionals to confirm their appropriateness, and the questionnaire was tested for clarity and reliability. Ethical considerations were also observed throughout the research process. Participants were informed about the purpose of the study and gave their consent voluntarily. All responses were kept confidential and anonymized to protect the identity of the participants, in accordance with the research ethics guidelines of Universitas Potensi Utama.

This methodological approach allows for a balanced investigation that combines objective evaluation of translation quality with subjective insights into students' experiences, providing a deeper

EXCELLENCE Journal of English and English Education
Volume 5 | Number 1 | June | 2025
DOI: 10.47662/ejeee.v5i1.1192
p-ISSN: 2807-6389
e-ISSN: 2807-2103
English Education Study Program, FKIP UNIVA Medan

understanding of the practical implications of using Google Translate and DeepL in academic and technical translation tasks.

C. Result and Discussion

This section presents the findings of the study, integrating both quantitative and qualitative data to provide a comprehensive understanding of how Google Translate and DeepL perform in translating Indonesian academic texts into English, and how students perceive their use. The discussion is organized around the two research questions: (1) comparison of translation quality, and (2) student perceptions and preferences.

1. Comparison of Translation Quality

Based on expert evaluation using a translation quality rubric, DeepL outperformed Google Translate across all assessed criteria. Table 1 shows the mean scores of each tool based on grammatical accuracy, lexical appropriateness, contextual coherence, and fluency.

Table 1. Average Translation Quality Scores by Expert Raters

Criteria	Google Translate	DeepL Translator
Grammatical Accuracy	3.8	4.3
Lexical Appropriateness	3.5	4.4
Contextual Coherence	3.3	4.5
Fluency and Naturalness	3.6	4.6
Total Average Score	**3.55**	**4.45**

To visualize these differences, the following bar chart illustrates the comparative scores across the four evaluation aspects:

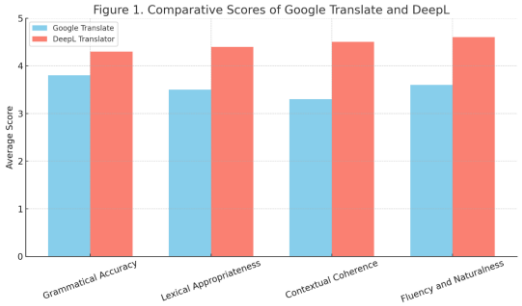

Figure 1. Comparative Scores of Google Translate and DeepL Based on Expert Evaluation

DeepL consistently produced more fluent, coherent, and semantically appropriate outputs. Particularly in technical texts, DeepL correctly rendered terms such as *"pemrosesan paralel"* into *"parallel processing"*, whereas Google Translate gave less accurate alternatives like *"parallel handling"*. Expert raters described DeepL's output as "closer to human translation," while Google Translate sometimes appeared "literal and stilted."

2. Student Perceptions and Preferences

Survey results from 40 students indicated that while Google Translate remains the most frequently used tool due to familiarity, DeepL was preferred in terms of output quality, especially after students reviewed both tools' translations. Table 2 summarizes student responses across several indicators.

Table 2. Student Preferences and Perceived Strengths

Aspect Evaluated	Google Translate	DeepL Translator
Familiarity	85%	37%
Perceived Accuracy	58%	83%
Ease of Use	91%	78%
Preferred for Academic Translation	24%	76%

The chart below visualizes students' preferred translation tool after being exposed to both:

EXCELLENCE Journal of English and English Education
Volume 5 | Number 1 | June | 2025
DOI: 10.47662/ejeee.v5i1.1192
p-ISSN: 2807-6389
e-ISSN: 2807-2103
English Education Study Program, FKIP UNIVA Medan

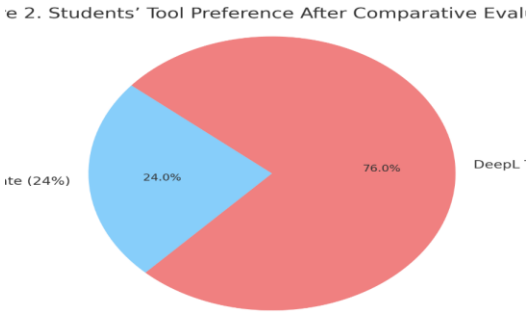

Figure 2. Students' Tool Preference After Comparative Evaluation

In interviews, students stated that while Google Translate is faster and easier to access, DeepL provides results that are "more academic" and "sound like natural English." A student noted, *"I never knew DeepL existed, but now I prefer it because the grammar feels right, especially for writing reports."*

However, a few students mentioned that DeepL's interface was less intuitive and occasionally misinterpreted culturally specific phrases. Despite this, the majority acknowledged its higher quality output and indicated they would shift to DeepL for future academic translation tasks.

Discussion

The results show a clear distinction between habitual use and actual preference when quality is taken into account. Although Google Translate dominates due to accessibility and popularity, DeepL's translation quality was rated significantly higher, especially in terms of fluency and domain-specific terminology. This aligns with prior research (e.g., Toral et al., 2021) which recognized DeepL's ability to better approximate human-like translations.

The pedagogical implication is evident: introducing students to comparative evaluation fosters **critical digital literacy**. By engaging in analytical comparisons, students began to reconsider their tool choices not based on convenience, but based on quality and suitability for academic purposes.

Additionally, students' comments suggest that translation tools, if used thoughtfully, can indirectly enhance language learning. By observing how DeepL renders certain phrases, students gained exposure to academic vocabulary and syntactic structures, a form of incidental learning that is valuable in higher education contexts.

D. Conclusion and Suggestion

This study set out to compare the translation performance of Google Translate and DeepL when applied to Bahasa Indonesia–English academic translation tasks among Informatics Engineering students at Universitas Potensi Utama. The research employed a mixed-methods approach to analyze both the objective quality of translations and students' subjective experiences with the tools.

The findings demonstrate that DeepL consistently outperforms Google Translate in terms of grammatical accuracy, lexical appropriateness, contextual coherence, and overall fluency. Expert evaluations showed that DeepL provided more natural and accurate translations, particularly for academic and technical content. Although students were initially more familiar with Google Translate, exposure to comparative results led the majority to favor DeepL for its superior quality. Furthermore, students who engaged in evaluating the translations developed a more critical awareness of language use and translation quality, suggesting that such comparative tasks can foster digital literacy and language learning.

The study also found that machine translation tools can serve as valuable support systems in academic contexts, especially for students from non-English backgrounds. However, reliance on these tools should be accompanied by training in evaluating and post-editing outputs, so that users do not passively accept flawed or unnatural translations.

Suggestions

Based on the results of this study, several suggestions are offered for educators, students, and future researchers:

1. For educators: Translation tools such as Google Translate and DeepL should be integrated into classroom activities not just as aids, but as objects of analysis. Activities like comparative translation exercises, error

p-ISSN: 2807-6389
e-ISSN: 2807-2103

analysis, and post-editing tasks can help students develop both language proficiency and digital translation literacy.

2. For students: While translation tools can support understanding and expression in English, it is important to treat their outputs critically. Students should be encouraged to revise and edit machine-generated translations and use them as learning tools rather than final products.

3. For developers and institutions: Given the high use of translation tools among students in technical fields, universities may consider providing access to premium or institutional versions of tools like DeepL, along with training workshops on translation strategies and ethical use.

4. For future research: Further studies could expand on this topic by including other language pairs, larger samples, or longitudinal designs to observe how students' translation skills evolve over time. Research might also explore how different disciplines (e.g., law, medicine, computer science) affect translation performance and tool preference.

In conclusion, while both Google Translate and DeepL offer accessible solutions for academic translation, DeepL emerges as the more effective option in terms of linguistic quality. With appropriate integration and critical engagement, these tools can become powerful allies in academic communication and language development.

E. References

Aiken, M., & Balan, S. (2011). An analysis of Google Translate accuracy. Translation Journal, 16(2).

Bahri, H. S., & Mahadi, T. S. T. (2016). The Use of Google Translate in Translating English Texts to Bahasa Indonesia (A Study on Translation Students at Universitas Negeri Medan). Jurnal Ilmiah Makna, 1(1), 45–53.

Castilho, S., Gaspari, F., Moorkens, J., & Way, A. (2018). Is neural machine translation the new state of the art?. The Prague Bulletin of Mathematical Linguistics, 110(1), 109–120. https://doi.org/10.2478/pralin-2018-0008

DePalma, D. A., & Kelly, N. (2008). Translation Memory and Machine Translation: Insights and Implications. Common Sense Advisory.

Garcia, I., & Pena, M. I. (2011). Machine translation-assisted language learning: Writing for beginners. Computer Assisted Language Learning, 24(5), 471–487. https://doi.org/10.1080/09588221.2011.582687

Groves, M., & Mundt, K. (2015). Friend or foe? Google Translate in language for academic purposes. English for Specific Purposes, 37, 112–121. https://doi.org/10.1016/j.esp.2014.09.001

Kussmaul, P. (1995). Training the Translator. John Benjamins Publishing Company.

Newmark, P. (1988). A Textbook of Translation. Prentice Hall.

Nuruzzaman, M., Shakil, M., & Azad, M. (2020). A comparative study on the performance of Google Translate and DeepL in English-Bangla translation. International

p-ISSN: 2807-6389
e-ISSN: 2807-2103

Journal of English Language and Linguistics Research, 8(4), 22–38.

Toral, A., & Sánchez-Cartagena, V. M. (2017). A multifaceted evaluation of neural machine translation for the English–Spanish language pair. Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics, 206–217.

Toral, A., Hu, K., & Sánchez-Cartagena, V. M. (2021). Reassessing the limits of neural machine translation: A case study of English–German literary translation. Machine Translation, 35, 39–74. https://doi.org/10.1007/s10590-021-09251-w

Vinay, J.-P., & Darbelnet, J. (1995). Comparative Stylistics of French and English: A methodology for translation. John Benjamins