

A comparative analysis of Google Translate and ChatGPT in translating Borneo Bulletin News into standard Malay

Amieziezaitul Syazlien Ezzeq Ezrynah Binti Amirul Syafiee¹, Lalu Nurul Yaqin^{2*}

^{1,2}Malay Language and Linguistics Programme, Faculty of Arts & Social Sciences,
Universiti Brunei Darussalam, Brunei Darussalam

*) Corresponding Author, email: lalu.yaqin@ubd.edu.bn

ABSTRACT

In the digital era, machine translation tools like Google Translate and ChatGPT have become essential for cross-linguistic communication, particularly in translating formal texts such as news articles. This study aims to evaluate the effectiveness of these tools in translating English news articles from *Borneo Bulletin* into Standard Malay, focusing on lexical accuracy, syntactic structure, and contextual appropriateness. A qualitative comparative analysis was conducted by translating selected articles using both Google Translate and ChatGPT-4. The outputs were manually analyzed using a qualitative rubric guided by meaning-based translation theory, focusing on word choice, sentence structure, the original tone, and cultural nuances. The results indicate that ChatGPT-4 outperforms Google Translate in maintaining lexical precision, grammatical coherence, and contextual relevance. ChatGPT-4 demonstrated superior handling of idiomatic expressions and complex sentence constructions, producing translations that adhered closely to journalistic norms in Standard Malay. However, both tools exhibited limitations in accurately translating culturally specific references and specialized terminology. Google Translate, despite improvements through neural machine translation, tended to produce overly literal translations, leading to a loss of tone and clarity in formal contexts. These findings highlight the potential and limitations of machine translation in journalism, where accuracy and tone are crucial for public trust and information dissemination. The study emphasizes the necessity for enhanced machine translation algorithms and the integration of human feedback to elevate translation quality in low-resource languages such as Malay.

Keywords: ChatGPT-4; Google Translate; machine translation; news translation; standard Malay

First Received:	Revised:	Accepted:	Published:
February 1 st , 2025	March 28 th , 2025	March 28 th , 2025	March 30 th , 2025

How to cite (in APA style):

Syafiee, A. S. E. E. B. A. & Yaqin, L. N. (2025). A comparative analysis of Google Translate and ChatGPT in translating Borneo Bulletin News into standard Malay. *Journal of Research on English and Language Learning (J-ReaLL)*, 6(1), 115–126.
<https://doi.org/10.33474/j-reall.v6i1.23252>

INTRODUCTION

In the digital era, the rapid evolution of machine translation (MT) tools has significantly influenced multilingual communication across various fields, including academia, business, and media. Among the most prominent MT tools are Google Translate and ChatGPT, both of which leverage advanced artificial intelligence (AI) to facilitate cross-language understanding. Despite their widespread use,

these tools often struggle with accurately translating between languages with distinct syntactic and cultural structures, such as English and standard Malay. This study evaluates the translation performance of Google Translate and ChatGPT, focusing on their ability to maintain linguistic accuracy, contextual appropriateness, and formal tone in news articles from the Borneo Bulletin.

Recent studies have highlighted persistent challenges in MT accuracy, particularly in translating complex grammatical structures and idiomatic expressions. Jiang and Hyland (2024) noted that ChatGPT, while contextually adept, encounters difficulties with complex sentence structures and passive constructions. Similarly, Naveen and Trojovský (2024) reported that MT frequently misinterprets complex sentence structures and grammatical transformations. Further research by Alosaimi (2024) evaluated ChatGPT's limitations with separable phrasal verbs, while Jiao et al. (2023) emphasized the improvements of ChatGPT-4 in translation tasks. Lai et al. (2023) explored the multilingual capabilities of large language models, revealing mixed results in non-English translations. The primary research problem addressed in this study is the inconsistent accuracy and contextual appropriateness of translations produced by Google Translate and ChatGPT, especially when applied to formal news content. These inconsistencies pose risks of misinformation and misinterpretation, which are critical concerns in journalistic contexts.

To address this, the study employs Larson's (1984) meaning-based translation theory to evaluate the tools' effectiveness in preserving original meanings and cultural contexts. Previous research offers general solutions to MT challenges by enhancing neural machine translation (NMT) algorithms and incorporating context-aware AI models. Chang and Hsu (2022) investigated the application of Chinese-English MT in academic texts, demonstrating the potential for improved accuracy through refined linguistic models. Salleh et al. (2024) focused on Malay-to-Arabic news translation, identifying specific strategies to bridge resource gaps and enhance translation quality. These studies underscore the importance of adapting MT tools to the unique linguistic features of target languages.

However, despite these advancements, there remains a significant research gap in evaluating the performance of MT tools in translating English to Standard Malay, particularly in the formal context of news reporting. This gap highlights the need for targeted studies that assess the capabilities and limitations of current MT technologies in handling Malay's complex syntactic structures and cultural nuances. This study aims to fill this gap by systematically analyzing translations of English news texts from the Borneo Bulletin using Google Translate and ChatGPT. The research focuses on three key aspects: the accuracy of word choice, the syntactic structure of translated texts, and the contextual alignment with the original content. By comparing these factors, the study provides a comprehensive evaluation of the strengths and weaknesses of each tool. The novelty of this research lies in its focus on formal news translation, a context that demands high linguistic precision and cultural sensitivity. By integrating theoretical frameworks with empirical analysis, the study contributes to the broader discourse on MT effectiveness in multilingual communication. The findings are expected to inform future improvements in MT algorithms and offer practical insights for human translators and AI developers alike. The scope of this study is limited to translations of news articles from the Borneo Bulletin, a leading newspaper in Brunei, providing a relevant and consistent dataset for analysis. The study's outcomes will have broader implications for enhancing MT tools' performance in other formal and professional contexts, thereby supporting more accurate and reliable cross-language communication.

LITERATURE REVIEW

Evolution of machine translation technologies

The evolution of machine translation (MT) technologies has significantly transformed cross-linguistic communication, particularly in formal settings such as journalism. Initially, MT systems were primarily rule-based, relying on fixed linguistic rules and direct word-for-word translations. While effective for basic tasks, these systems struggled to handle complex linguistic structures, idiomatic expressions, and cultural subtleties, which often led to inaccurate or awkward translations (Alkhawaja et al., 2020). This

limitation prompted the development of more sophisticated approaches, culminating in the advent of neural machine translation (NMT).

NMT marked a revolutionary shift in MT by leveraging artificial neural networks to model the relationships between words and phrases in a more dynamic and context-sensitive manner. Google Translate, one of the most prominent NMT tools, demonstrated significant improvements in fluency and contextual coherence, enabling more natural and accurate translations (Chang & Hsu, 2022). However, despite these advancements, NMT systems are not without their flaws. They often encounter difficulties in translating idiomatic expressions, nuanced cultural references, and domain-specific terminology, which are critical for maintaining the integrity of formal texts like news articles. Building on the foundation laid by NMT, ChatGPT, powered by large language models (LLMs), represents the next frontier in machine translation technology. LLMs like ChatGPT are designed to understand and generate human-like text by analyzing vast amounts of linguistic data, allowing for greater contextual understanding, syntactic accuracy, and adaptability to various linguistic styles (Jiao et al., 2023). Alkhawaja (2024) explored the capabilities of ChatGPT-3 in translating Arabic-English language pairs, highlighting its potential to maintain contextual relevance and grammatical coherence. The study found that ChatGPT-3 excelled in preserving the tone and style of the source text, especially in formal writing, thus offering a promising alternative to traditional NMT systems.

Despite these advancements, both Google Translate and ChatGPT face persistent challenges when translating languages with distinct syntactic and cultural frameworks, such as English to Standard Malay. The structural differences between these languages, including variations in sentence construction, verb conjugation, and cultural idioms, pose significant hurdles for machine translation tools. For instance, Standard Malay often employs different syntactic patterns and expressions that do not have direct equivalents in English, complicating the translation process. As a result, even advanced MT systems may produce translations that are technically correct but fail to capture the intended meaning or cultural nuance of the original text (Yaqin et al., 2025; Yaqin, 2025).

Moreover, the effectiveness of MT tools in handling formal contexts, such as news articles, is further complicated by the need to preserve journalistic standards of accuracy, tone, and objectivity. Inaccurate translations or misinterpretations can lead to misinformation, which is particularly problematic in journalism, where the credibility of the source is paramount. Therefore, while ChatGPT and Google Translate have made remarkable strides in improving machine translation quality, there remains a critical need for further refinement, particularly in addressing the unique linguistic challenges posed by languages like Standard Malay.

Challenges in translating English to Standard Malay

Translating between English and Standard Malay presents a unique set of challenges, primarily due to differences in syntactic structures, lexical choices, and cultural contexts. Jiang and Hyland (2024) emphasized that while ChatGPT excels in contextual understanding, it struggles with complex sentence structures, passive constructions, and maintaining a formal tone. Similarly, ALAqad and Al-Saggap (2021) reported that Google Translate often misinterprets idiomatic expressions and technical jargon, resulting in translations that lack clarity and fail to preserve the intended tone of the original text.

The application of Larson's (1984) meaning-based translation theory underscores the importance of preserving original meanings and cultural nuances in translation. This theory posits that effective translation transcends literal word substitution, focusing instead on conveying the source text's intended message within the cultural and linguistic norms of the target language. However, both Google Translate and ChatGPT struggle to consistently meet these standards, particularly in news translation, where accuracy and tone are crucial for maintaining public trust. Feng et al. (2024) explored the application of transformer-based models in improving machine translation and post-editing for Chinese tourism texts, identifying similar challenges in handling cultural nuances and idiomatic expressions. These findings are relevant to English-Malay translation, where cultural context plays a significant role in the accuracy and appropriateness of translations.

Comparative studies on machine translation tools

Comparative studies of machine translation (MT) tools have shed light on their strengths and limitations, offering crucial insights into their applicability across different linguistic contexts. [Alosaimi and Alawad \(2024\)](#) conducted a comprehensive evaluation of ChatGPT's translation capabilities, particularly focusing on its handling of separable phrasal verbs and complex syntactic structures. While the findings indicated notable improvements in these areas, the study highlighted persistent gaps in maintaining contextual appropriateness, especially in the translation of formal texts like news articles. This shortcoming underscores the need for further refinement in MT models to enhance their ability to accurately convey the tone and intent of the original text. Similarly, [Salleh et al. \(2024\)](#) investigated the translation of Malay-to-Arabic news content, identifying specific strategies to bridge resource gaps and improve translation quality. These strategies included the use of specialized glossaries and contextual training, emphasizing the importance of tailoring MT tools to the unique linguistic features of target languages. The study's findings underline the necessity for MT systems to incorporate language-specific adaptations to enhance translation fidelity and contextual accuracy.

[Zayyanu \(2024\)](#) further emphasized the critical role of expanding training datasets to encompass a broader spectrum of linguistic and cultural contexts. This is particularly vital for low-resource languages like Malay, where limited training data can hinder the MT system's ability to accurately translate cultural references, idiomatic expressions, and nuanced language forms essential for formal writing. The inclusion of diverse linguistic data not only improves the system's lexical choices but also enhances its ability to maintain the stylistic and rhetorical elements of the source text.

[Karabayeva and Kalizhanova \(2024\)](#) explored the evaluation of machine translation through rhetorical analysis, demonstrating that while MT tools have made significant strides in handling straightforward translations, they often falter in preserving rhetorical devices and stylistic nuances. The study highlighted those elements such as metaphors, analogies, and tone are frequently lost or misrepresented in machine-generated translations, which can significantly alter the intended message of the source text. This limitation is particularly critical in formal contexts like journalism, where the precise conveyance of tone and style is essential for maintaining the credibility and integrity of the information. Innovative approaches to enhancing MT quality have also been explored through the integration of human feedback. [Santy et al. \(2019\)](#) introduced the concept of interactive neural machine translation (INMT) prediction, which incorporates real-time human input to refine and correct machine-generated translations. This interactive model allows for dynamic adjustments during the translation process, significantly improving the overall quality and contextual appropriateness of the output. Similarly, [Şahin and Gürses \(2021\)](#) examined English-Turkish literary translation through human-machine interaction, emphasizing the benefits of combining human expertise with machine efficiency. Their findings suggest that such hybrid approaches can lead to more accurate and culturally sensitive translations, bridging the gap between machine capabilities and human linguistic intuition. Despite these advancements, a significant research gap persists in evaluating the performance of MT tools in translating formal news content from English to Standard Malay. This gap highlights the need for targeted research that combines theoretical frameworks, empirical analysis, and the integration of human feedback to refine MT algorithms. By addressing these gaps, future studies can contribute to the development of more sophisticated and contextually aware MT systems capable of meeting the high standards required for formal and professional translations.

METHOD

This research adopts a qualitative approach to evaluate the effectiveness of artificial intelligence (AI) technologies, specifically Google Translate and ChatGPT-4, in translating English news articles from the Borneo Bulletin into Standard Malay. The primary objective is to assess the translation accuracy, syntactic fidelity, and contextual appropriateness of outputs generated by these two machine translation tools.

Data collection and preparation

The data for this study were sourced from *Borneo Bulletin*, a leading newspaper in Brunei known for its formal and structured language use. The dataset comprises three news, each covering a different topic. The news was selected randomly to ensure a diverse representation of subject matter and linguistic complexity, with each article containing approximately 200 to 300 words. The selected English texts were then input into both Google Translate and ChatGPT-4 to generate translations into Standard Malay. This process provided a robust dataset for comparative analysis. The translations obtained from both tools were collected manually. This ensured the integrity of the data and allowed for precise alignment of source texts with their translated counterparts. The focus on news is critical, as this genre demands a high degree of linguistic precision and contextual accuracy, providing an ideal testing ground for evaluating machine translation performance.

Analytical framework and evaluation criteria

The evaluation of translation accuracy is guided by [Larson's \(1984\)](#) meaning-based translation theory, which emphasizes the preservation of original meaning and cultural context. According to [Larson \(1984\)](#), effective translation transcends mere word substitution, aiming to convey the source text's intended message in a manner that resonates with the linguistic and cultural norms of the target audience. This theoretical framework provides a comprehensive basis for assessing the fidelity of machine-generated translations. Three primary criteria guide the analysis:

1. **Word choice accuracy:** This criterion assesses whether the translated text accurately reflects the meanings of the original words and phrases. It involves evaluating the appropriateness of lexical choices and their alignment with Standard Malay usage. *High:* Lexical choices are accurate, contextually appropriate, and aligned with standard journalistic usage. *Moderate:* Mostly accurate, with minor inconsistencies or generalizations. *Low:* Literal or incorrect word choices that distort meaning or tone.
2. **Syntactic structure:** This involves examining the grammatical construction of the translated texts. The focus is on ensuring that sentence structures conform to the syntactic norms of Standard Malay, particularly in handling complex grammatical forms and idiomatic expressions. *High:* Grammatically coherent and adheres to Standard Malay syntax. *Moderate:* Minor structural issues but still comprehensible. *Low:* Awkward or fragmented syntax that disrupts meaning.
3. **Contextual appropriateness:** This criterion evaluates whether the translation maintains the appropriate tone, style, and contextual relevance required for formal news reporting. It assesses how well the translation preserves the source text's communicative intent and cultural nuances. *High:* Maintains formal tone, style and accurately reflects cultural and contextual meaning. *Moderate:* Partially effective, with some loss in tone or nuance. *Low:* Fails to preserve tone or context, leading to misinterpretation.

To enhance the rigor of the linguistic evaluation, *Kamus Dewan Edisi Keempat*, a comprehensive Malay language dictionary, is used as a reference. This resource provides authoritative definitions and usage examples, ensuring that the translations adhere to standardized Malay linguistic conventions.

Data analysis procedure

The analysis process involves a detailed comparison of the original English texts and their translated versions, with translation accuracy guided by [Larson's \(1984\)](#) framework. This comparison focuses on identifying and categorizing translation errors, such as incorrect word choices, syntactic inconsistencies, and contextual misalignments. By systematically examining these elements, the study aims to uncover patterns and trends in the performance of Google Translate and ChatGPT-4. The comparative analysis extends across different types of content and linguistic challenges, allowing for a nuanced understanding of each tool's strengths and limitations. This approach facilitates a comprehensive evaluation of how well each tool handles various linguistic complexities inherent in English-Malay translation tasks. While this study provides valuable insights into the capabilities of

machine translation tools, it is important to acknowledge its limitations. The focus on a specific dataset news from the *Borneo Bulletin* may limit the generalizability of the findings to other text genres or translation contexts. Future research could expand the scope to include a broader range of text types and languages, thereby offering a more comprehensive evaluation of machine translation technologies. By systematically analyzing the translations produced by Google Translate and ChatGPT-4, this study aims to contribute to the broader understanding of AI's role in multilingual communication. The findings are expected to inform the ongoing development of more accurate and reliable machine translation tools, enhancing their applicability in diverse linguistic and professional contexts.

FINDINGS

This section presents the outcomes of the comparative analysis between Google Translate and ChatGPT-4 in translating English news articles from *Borneo Bulletin* into Standard Malay. The analysis focuses on three primary aspects: lexical accuracy, syntactic structure, and contextual appropriateness. By evaluating these dimensions, the study identifies the strengths and limitations of each translation tool in handling formal journalistic texts.

Lexical accuracy

Lexical accuracy is crucial in ensuring the intended meaning of the original text is preserved. Both Google Translate and ChatGPT-4 demonstrated varying levels of effectiveness in selecting appropriate words, with noticeable differences in handling idiomatic expressions and technical terms. ChatGPT-4 generally outperformed Google Translate in selecting contextually suitable words, especially when dealing with idiomatic phrases. For example, the phrase "*blood donation drive*" was translated as "*Kempen derma darah*" by ChatGPT-4, accurately reflecting the notion of an organized campaign. In contrast, Google Translate rendered it as "*Sumbangan derma darah*," which incorrectly implies a single act of donation rather than an organized event.

Table 1

Lexical comparisons between Google Translate and ChatGPT-4

Original Text (English)	Google Translate (Malay)	ChatGPT-4 (Malay)
"Blood donation drive"	"Sumbangan derma darah"	"Kempen derma darah"
"Uniform and Service Division"	"Bahagian Pakaian Seragam dan Perkhidmatan"	"Bahagian Beruniform dan Khidmat"
"Ministry of Education"	"KPM"	"Kementerian Pendidikan"

Despite ChatGPT-4's superior performance, both tools struggled with specialized terminology, particularly in political and technical news contexts. Google Translate frequently misinterpreted technical terms, while ChatGPT-4 occasionally generalized terms, reflecting the limitations of current AI in domain-specific translations (Wang, 2024).

Syntactic structure

The syntactic structure of translated texts plays a critical role in maintaining readability and coherence. This study evaluated the grammatical constructions of sentences produced by both tools, focusing on their ability to handle complex sentence structures, such as subordinate clauses and passive voice. ChatGPT-4 consistently demonstrated superior syntactic coherence. It handled passive constructions and complex clauses effectively, producing translations that adhered closely to Standard Malay grammar. For instance, the sentence "*The index revealed that 25% of Malaysians are interested in buying an EV*" was translated by ChatGPT-4 as "*Indeks tersebut mendedahkan bahawa 25 peratus rakyat Malaysia berminat untuk membeli kenderaan elektrik (EV).*" This translation maintained both clarity and grammatical correctness. On the other hand, Google Translate produced "*Indeks itu mendedahkan bahawa 25 peratus rakyat Malaysia berminat untuk membeli EV,*" which, while correct, lacked the formal tone necessary for journalistic writing.

Table 2
Syntactic structure comparisons between Google Translate and ChatGPT-4

Original Text (English)	Google Translate (Malay)	ChatGPT-4 (Malay)
"The index revealed that 25% of Malaysians are interested in buying an EV."	"Indeks itu mendedahkan bahawa 25 peratus rakyat Malaysia berminat untuk membeli EV."	"Indeks tersebut mendedahkan bahawa 25 peratus rakyat Malaysia berminat untuk membeli kenderaan elektrik (EV)."
"Despite slow global demand."	"Walaupun permintaan global perlahan."	"Walaupun permintaan global yang perlahan."

Google Translates often exhibited fragmented or awkward sentence constructions, deviating from Standard Malay norms, especially in sentences with multiple clauses (Zaki & Ahmed, 2024).

Contextual appropriateness

Preserving context, tone, and cultural nuances is vital in news translation. The formal tone characteristic of journalistic writing must be retained to ensure the credibility of the translated content. The analysis highlighted that ChatGPT-4 outperformed Google Translate in this area, producing translations that better captured the formal and factual style of *Borneo Bulletin* articles. For example, the sentence “Lee, born in Perth to Korean parents, beat Alison Lee in a playoff” was translated by ChatGPT-4 as “Lee, yang dilahirkan di Perth oleh ibu bapa berketurunan Korea, mengalahkan Alison Lee dalam perlawanan playoff.” This version maintained both the formal tone and the clarity expected in news reporting. Conversely, Google Translate’s version, “Lee, yang dilahirkan di Perth daripada ibu bapa Korea, menewaskan Alison Lee dalam playoff,” used less precise phrasing, and the term “menewaskan” added a more aggressive tone, which was not present in the original text.

Table 3
Contextual appropriateness in Translations by Google Translate and ChatGPT-4

Original Text (English)	Google Translate (Malay)	ChatGPT-4 (Malay)
"Lee, born in Perth to Korean parents, beat Alison Lee in a playoff."	"Lee, yang dilahirkan di Perth daripada ibu bapa Korea, menewaskan Alison Lee dalam playoff."	"Lee, yang dilahirkan di Perth oleh ibu bapa berketurunan Korea, mengalahkan Alison Lee dalam perlawanan playoff."
"Held at the MoE’s new premises."	"Yang diadakan di premis baharu KPM."	"Yang diadakan di premis baharu Kementerian Pendidikan."

ChatGPT-4 demonstrated a better grasp of cultural references and formal expressions, aligning more closely with the expectations of Malay-speaking audiences. However, both tools faced challenges in translating idiomatic expressions and culturally specific terms, underscoring the difficulty of encoding cultural nuance in AI translation systems (Soysal, 2023; Zaki & Ahmed, 2024).

Common translation errors

Several recurring errors were identified in both Google Translate and ChatGPT-4 outputs, categorized as lexical, syntactic, and contextual errors.

Table 4
Common translation errors in Google Translate and ChatGPT-4 Outputs

Error Type	Google Translate	ChatGPT-4
Lexical Errors	Literal translations, distorted meanings	Overgeneralization, occasional imprecision
Syntactic Errors	Fragmented sentences, awkward structures	Minor issues with complex constructions
Contextual Errors	Misinterpretation of idiomatic expressions	Occasional loss of cultural nuances

These errors highlight ongoing challenges in machine translation, particularly in maintaining the balance between linguistic accuracy and contextual appropriateness (Broussard et al., 2019). Errors in machine translation can have profound implications, especially in formal contexts such as journalism. Inaccuracies in translating key terms, idiomatic expressions, or nuanced phrases may inadvertently

lead to the dissemination of misinformation. This misrepresentation can distort public understanding of critical events or issues and contribute to the erosion of trust in media sources. This concern is particularly acute in areas like political and economic reporting, where precision and clarity are paramount. Misinterpretations in these domains may influence public opinion, policy decisions, and even international relations (Li, 2024).

Another important factor is the ability of machine translation systems to preserve the original tone, style, and voice of journalistic content. The failure to maintain these elements can compromise the authenticity and perceived credibility of translated articles. While advanced AI systems like ChatGPT-4 have shown notable improvements in preserving tone and context, they still exhibit occasional lapses, particularly in cultural sensitivity. These lapses can result in translations that feel awkward, inappropriate, or even offensive in certain cultural contexts. Such issues underscore the need for continuous refinement in AI translation technologies to better align with the ethical standards and cultural expectations of journalism (Soysal, 2023).

Furthermore, the reliance on AI translations without adequate human oversight poses ethical concerns related to media bias and representation. Automated systems, trained on specific datasets, may inadvertently reflect biases present in their training material, leading to skewed reporting. In journalism, where impartiality and objectivity are foundational, these biases can have far-reaching consequences, influencing not only the accuracy of information but also public trust in the media ecosystem.

DISCUSSION

The comparative analysis of Google Translate and ChatGPT in translating English news articles from *Borneo Bulletin* into Standard Malay reveals both significant advancements and persistent challenges within machine translation (MT) technologies. While both tools demonstrate impressive capabilities, particularly in terms of fluency and basic contextual understanding, their limitations become evident when tasked with more nuanced linguistic and cultural translations. This discussion synthesizes these findings, focusing on lexical accuracy, syntactic structure, and contextual appropriateness while addressing broader implications for journalism and public information dissemination.

Both Google Translate and ChatGPT have undergone substantial technological transformations, leading to marked improvements in their translation outputs. Google Translate's shift from Statistical Machine Translation (SMT) to Neural Machine Translation (NMT) has notably enhanced its fluency and contextual coherence (Chang & Hsu, 2022). Similarly, ChatGPT, leveraging large language models (LLMs), has demonstrated the ability to generate coherent and contextually relevant translations. However, the comparative results indicate that while these tools are proficient in straightforward translations, they struggle with preserving the subtleties inherent in formal writing, particularly in journalism where accuracy and tone are critical.

In terms of lexical accuracy, ChatGPT generally outperformed Google Translate by selecting more contextually appropriate words, especially in idiomatic expressions and technical terms. For instance, in translating the phrase "*blood donation drive*," ChatGPT rendered it as "*Kempen derma darah*," accurately reflecting the notion of an organized campaign. In contrast, Google Translate's output, "*Sumbangan derma darah*," suggested an individual act of donation, misrepresenting the source text's meaning. This discrepancy underscores the importance of contextual sensitivity in translation, especially in formal texts where slight lexical misinterpretations can significantly alter the intended message. However, both tools struggled with specialized terminology, particularly in political and technical news contexts, highlighting ongoing challenges in domain-specific translations (Wang, 2024).

Syntactically, ChatGPT demonstrated superior performance in handling complex sentence structures. Its translations maintained grammatical coherence, particularly in sentences with subordinate clauses and passive voice constructions. For example, the sentence "*The index revealed that 25% of Malaysians are interested in buying an EV*" was translated by ChatGPT as "*Indeks tersebut mendedahkan bahawa 25 peratus rakyat Malaysia berminat untuk membeli kenderaan elektrik (EV)*," preserving both the formal tone and grammatical accuracy expected in news reporting. Conversely, Google

Translate produced “*Indeks itu mendedahkan bahawa 25 peratus rakyat Malaysia berminat untuk membeli EV,*” which, while grammatically correct, lacked the necessary formal tone and completeness, omitting the full term for EV, which is crucial in formal journalism.

Despite these strengths, both tools displayed notable limitations in preserving the contextual appropriateness of translations. Contextual appropriateness is vital in news translation as it affects the tone, style, and cultural relevance of the content. ChatGPT excelled in maintaining the formal tone characteristic of *Borneo Bulletin* articles, producing translations that adhered closely to journalistic norms. However, it occasionally failed to capture subtle cultural references, a limitation exacerbated in low-resource languages like Malay, where training data may be insufficient to fully inform the model (Zaki & Ahmed, 2024). Google Translate, while improved, often produced translations that were overly literal, leading to a loss of the original article's tone and potential misinterpretations of key messages (Salleh, 2024).

The implications of these translation errors are particularly significant in journalism. Inaccuracies in translating key terms or phrases can lead to misinterpretations of news articles, thereby affecting public understanding and trust in media sources. For example, in political reporting, the mistranslation of a critical term could alter the perceived stance of a news article, potentially leading to misinformation. Broussard et al. (2019) highlight how even minor inaccuracies in translation can distort public perception, a risk that is amplified in sensitive contexts such as political or health reporting. Furthermore, Li (2024) argues that the loss of the journalist's voice through the failure to maintain tone and stylistic nuances can undermine the credibility of news outlets, compromising the integrity of the information conveyed.

Previous literature has extensively discussed the challenges and limitations of MT tools in preserving the original tone and meaning of translated texts. Critics have noted that MT tools often produce overly literal translations, leading to the loss of emotional depth and cultural nuances inherent in the source material (Adnan et al., 2018). This literalness, while useful in technical translations, can be detrimental in journalistic writing, where tone, style, and cultural context are crucial for conveying the intended message. Moreover, the reliance on algorithms can lead to inconsistencies, particularly when translating texts that require a nuanced understanding of cultural subtleties and complex syntactic structures.

Addressing these challenges requires a multifaceted approach. One key recommendation is the incorporation of human feedback into the translation process. Soysal (2023) suggests that integrating human evaluators can refine AI algorithms, enhancing their ability to interpret context and cultural nuances accurately. This approach not only improves the immediate output of translations but also contributes to the long-term development of more sophisticated and contextually aware MT systems.

Additionally, expanding and diversifying training datasets is critical for improving the adaptability of MT tools. Zaki and Ahmed (2024) emphasize the need for comprehensive datasets that encompass a wide range of linguistic and cultural contexts, particularly for low-resource languages like Malay. This expansion would enable MT tools to better understand and translate cultural references, idiomatic expressions, and nuanced language forms that are essential in formal writing.

Implementing hybrid models that combine rule-based and statistical approaches is another promising avenue for improving translation accuracy. Siregar et al. (2024) advocate for this approach, arguing that the integration of structured linguistic rules with statistical learning can enhance the flexibility and precision of MT outputs. This combination allows for a more nuanced handling of complex syntactic structures and context-dependent translations, which are often challenging for purely algorithmic models. Fostering collaboration between linguists and AI developers is essential to ensure that MT tools are not only technically proficient but also culturally sensitive. highlights the importance of interdisciplinary collaboration in developing translation tools that can accurately capture both the linguistic and cultural dimensions of language. Such partnerships can bridge the gap between technological capability and linguistic accuracy, leading to more reliable and contextually appropriate translations.

CONCLUSION

This study offers a comparative analysis of Google Translate and ChatGPT-4 in translating English news articles from *Borneo Bulletin* into Standard Malay. The findings highlight ChatGPT-4's superior performance in lexical accuracy, syntactic structure, and contextual appropriateness. ChatGPT-4 demonstrated better handling of idiomatic expressions, formal tone, and complex sentence structures, producing translations that closely adhered to Standard Malay conventions. However, it still exhibited occasional lapses in cultural sensitivity and domain-specific terminology. Google Translate, while benefiting from advancements in NMT, struggled with maintaining the formal tone and accurately translating culturally specific references. Its tendency toward literal translations often resulted in a loss of nuance, affecting the clarity and credibility of the translated texts. Both tools faced challenges in fully capturing the subtleties of Malay syntax and cultural nuances, especially given the limitations of training data for low-resource languages. The implications of these findings are significant, particularly in journalism, where translation accuracy is critical for maintaining public trust and ensuring the correct dissemination of information. Translation errors can lead to misinterpretation, potentially causing misinformation and undermining the credibility of media outlets. This study focuses solely on translating English news from *Borneo Bulletin* into Standard Malay, which may not reflect other text genres or language combinations. Moreover, the reliance on qualitative evaluation limits the generalizability of the findings to broader contexts, including informal or conversational language. This study contributes to the broader understanding of AI-powered translation tools' strengths and limitations in formal, context-sensitive translations. It underscores the need for continuous refinement of machine translation algorithms, incorporation of human feedback, and the development of more comprehensive training datasets. Future research should explore hybrid models that combine AI with human input and examine machine translation performance across other low-resource languages and formal domains to further enhance accuracy and cultural sensitivity.

REFERENCES

- Adnan, T. H., Apandi, M. M., Kamaruddin, H., Salowi, M. A., Law, K. B., Haniff, J., & Goh, P. P. (2018). Catquest-9SF questionnaire: Validation of Malay and Chinese-language versions using Rasch analysis. *Health and Quality of Life Outcomes*, 16(1). <https://doi.org/10.1186/s12955-017-0833-3>
- AlAqad, M. H., & Al-Saggaf, M. A. (2021). Issues in translating cultural terms between English and Malay: A comparative analysis. *Pedagogical Research*, 6(4). <https://doi.org/10.29333/pr/11272>
- Alkhawaja, L. (2024). Unveiling the new frontier: ChatGPT-3 powered translation for Arabic-English language pairs. *Theory and Practice in Language Studies*, 14(2), 347-357. <https://doi.org/10.17507/tpls.1402.05>
- Alkhawaja, L., Ibrahim, H., Ghnaim, F., & Awwad, S. (2020). Neural machine translation: Fine-grained evaluation of Google Translate output for English-to-Arabic translation. *International Journal of English Linguistics*, 10(4), 43. <https://doi.org/10.5539/ijel.v10n4p43>
- Alosaimi, B. A., & Alawad, N. A. (2024). Evaluation of the translation of separable phrasal verbs generated by ChatGPT. *Open Science Framework*. <https://doi.org/10.31235/osf.io/y6mca>
- Broussard, M., Diakopoulos, N., Guzman, A. L., Abebe, R., Dupagne, M., & Chuan, C. H. (2019). Artificial intelligence and journalism. *Journalism & mass communication quarterly*, 96(3), 673-695. <https://doi.org/10.1177/1077699019859901>
- Chang, P. L., & Hsu, C. H. (2022). The feasibility of Chinese–English machine translation applied to academic texts: Using thesis abstracts from National Digital Library of Theses and Dissertations (NDLTD) in Taiwan. *Translogos Translation Studies Journal*, 5(2), 60–83. <https://doi.org/10.29228/translogos.47>
- Feng, Y., Hu, Q., & Yang, S. (2024). Improving machine translation and post-editing for Chinese tourism texts using transformer-based models. *Applied and Computational Engineering*, 34(1), 7-12. <https://doi.org/10.54254/2755-2721/34/20230286>

- Jiang, F., & Hyland, K. (2024). Does ChatGPT argue like students? Bundles in argumentative essays. *Applied Linguistics*, amae052. <https://doi.org/10.1093/applin/amae052>
- Jiao, W., Wang, W., Huang, J. T., Wang, X., & Tu, Z. (2023). Is ChatGPT a good translator? Yes, with GPT-4 as the engine. *arXiv Preprint*. <https://doi.org/10.48550/arxiv.2301.08745>
- Karabayeva, I., & Kalizhanova, A. (2024). Evaluating machine translation of literature through rhetorical analysis. *Journal of Translation and Language Studies*, 5(1), 1-9. <https://doi.org/10.48185/jtls.v5i1.962>
- Lai, V. D., Ngo Trung, N., & Veyseh, A. (2023). ChatGPT beyond English: Towards a comprehensive evaluation of large language models in multilingual learning. *Findings of the Association for Computational Linguistics: EMNLP 2023*. <https://doi.org/10.18653/v1/2023.findings-emnlp.878>
- Larson, M. L. (1984). *Meaning-based translation: A guide to cross-language equivalence*. Lanham, MD: University Press of America.
- Li, N. (2024). Analysis of the global literary significance of foreign novel translations in late Qing China: A case study of 'Novels Translated by Lin Shu' (1899–1919). *Critical Survey*, 36(3), 29-47. <https://doi.org/10.3167/cs.2024.360304>
- Naveen, P., & Trojovský, P. (2024). Overview and challenges of machine translation for contextually appropriate translations. *Iscience*, 27(10).
- Salleh, P. A. H. B. M., Ton, D. M. R., Sabdin, M. M., & Kadhim, A. D. K. A. P. (2024). Bridging the resource gap for Malay-to-Arabic translation: Evaluating machine translation of news headlines. *International Journal of Academic Research in Progressive Education and Development*, 13(3). <https://doi.org/10.6007/ijarped/v13-i3/22148>
- Santy, S., Dandapat, S., Choudhury, M., & Bali, K. (2019). INMT: Interactive neural machine translation prediction. <https://doi.org/10.18653/v1/d19-3018>
- Şahin, M., & Gürses, S. (2021). English-Turkish literary translation through human-machine interaction. *Tradumàtica Tecnologies De La Traducció*, 19, 171-203. <https://doi.org/10.5565/rev/tradumatica.284>
- Siregar, S. D. P., Al Farisi, M. Z., & Naser, A. D. M. (2024). Mapping the landscape of machine translation error research: A bibliometric review (1980-2023). *Buletin Al-Turas*, 30(1). <https://doi.org/10.15408/bat.v30i1.32655>
- Soysal, F. (2023). Enhancing translation studies with Artificial Intelligence (AI): Challenges, opportunities, and proposals. *Karamanoğlu Mehmetbey Üniversitesi Uluslararası Filoloji ve Çeviribilim Dergisi*, 5(2), 177-191. <https://doi.org/10.55036/ufced.1402649>
- Wang, J. (2024). Bibliometric insights into translation technology: A CiteSpace analysis of Web of Science core collection publications (2000-2024). *Translating the Future: Exploring the Impact of Technology and AI on Modern Translation Studies*, 85-136. <https://doi.org/10.46679/9788196780593ch05>
- Yaqin, L. N., Yusof, B., Yusof, N., & Damit, A. R. (2025). Students' perception of using ChatGPT as an AI-integrated tool in the Malay Language. *Jurnal Penelitian dan Pengkajian Ilmu Pendidikan: e-Saintika*, 9(1), 13-31. <https://doi.org/10.36312/e-saintika.v9i1.2584>
- Yaqin, L. N. (2025). Emerging research in AI-assisted language learning: A systematic literature review. *Multi-Industry Digitalization and Technological Governance in the AI Era*, 29-50. <https://doi.org/10.4018/979-8-3373-1681-9.ch002>
- Zaki, M. Z., & Ahmed, U. (2024). Bridging linguistic divides: The impact of AI-powered translation systems on communication equity and inclusion. *Journal of translation and language studies*, 5(2), 20-30. <https://doi.org/10.48185/jtls.v5i2.1065>
- Zayyanu, Z. M. (2024). Transforming worlds: The intersection of translation technology and transformers in the digital age. *Computer Science & Engineering: An International Journal*. 14(3), 1-14. <https://doi.org/10.5121/cseij.2024.14301>

CONFLICT OF INTEREST STATEMENT:

The authors of the manuscript affirm that they have no conflicting interests.

Copyright © 2025 Amieziezaitul Syazlien Ezzeq Ezrynah Binti Amirul Syafiee, Lalu Nurul Yaqin