LLT Journal, e-ISSN 2579-9533, p-ISSN 1410-7201, Vol. 28, No. 1, April 2025, pp. 495-517

LLT Journal: A Journal on Language and Language Teaching
http://e-journal.usd.ac.id/index.php/LLT
Sanata Dharma University, Yogyakarta, Indonesia

PERFORMANCE AND ACCURACY OF CHATGPT IN GENERATING MALAY ACADEMIC TEXTS: A COMPARATIVE STUDY WITH EXPERT CORRECTIONS

Lalu Nurul Yaqin[1*], Hasmidar Hassan[2], and Badriyah Yusof[3]
[1,2,3]Universiti Brunei Darussalam, Brunei Darussalam
lalu.yaqin@ubd.edu.bn[1], hasmidar.hassan@ubd.edu.bn[2],
and badriyah.yusof@ubd.edu.bn[3]
*correspondence: lalu.yaqin@ubd.edu.bn
https://doi.org/10.24071/llt.v28i1.11698
received 10 February 2025; accepted 26 April 2025

Abstract

The increasing use of artificial intelligence in academic writing has raised concerns about the accuracy and coherence of AI-generated texts, particularly in underrepresented languages like Malay. This study evaluates the performance of ChatGPT in generating Malay academic texts by comparing AI-generated outputs with expert corrected versions, focusing on grammatical errors, structural inconsistencies, and lexical inaccuracies. A comparative analysis was conducted on two datasets: Trained Dataset (TD), where prompts included detailed context, and Untrained Dataset (UTD). ChatGPT-generated texts were reviewed by Malay linguistics and translation experts, who identified and corrected grammatical errors. A quantitative and qualitative analysis assessed error frequency and categorized linguistic challenges. Findings reveal that UTD contained significantly more grammatical errors (87 errors) than TD (18 errors), demonstrating the role of structured prompts in enhancing text quality. Common errors in UTD included incorrect sentence structure (27.59%), omission of names (14.94%), and inappropriate word choices (11.49%). While TD showed improved grammatical accuracy, errors in phrase structure, conjunction usage, and affixation persisted. The study concludes that AI-generated Malay texts lack syntactic stability, requiring expert intervention and model refinement. These findings highlight the need for linguistic adaptation, expanded training datasets and the integration of expert to enhance AI-generated Malay academic writing. Ultimately, this study presents a case study that provides empirical evidence that context-aware prompt engineering and expert-in-the-loop approaches are essential for enhancing the quality of AI outputs, especially in non-English settings. It also advocates for the development of AI models that can capture nuances and linguistic diversity, vital for inclusive education for all.

Keywords: AI-generated text, ChatGPT, expert correction, linguistic accuracy Malay academic writing

Introduction

The use of artificial intelligence (AI) in academic writing has significantly altered the processes of text generation and refinement. Among these advancements, ChatGPT, developed by OpenAI, is a powerful tool that supports multiple languages and various writing tasks (Adıgüzel et al., 2023; Xu et al., 2024). Recent studies highlight its effectiveness in proofreading and drafting. For instance, Lechien et al. (2024) examined ChatGPT-4's capability in identifying and correcting manuscript errors, demonstrating its value in academic contexts. However, these tools encounter challenges in maintaining the integrity, accuracy, and contextual appropriateness of the generated texts, especially in non-English academic writing (Daungsupawong & Wiwanitkit, 2023; Dubey, 2024; Rojas, 2024; Syafiee & Yaqin, 2025; Yaqin et al., 2025).

While ChatGPT's capabilities have been extensively studied in English contexts (Dergaa et al., 2023; Kostka & Toncelli, 2023; Li, 2020; Nhan, 2024; Ruksakulpiwat, 2024), its use in generating academic texts in Malay has been underexplored (Nomoto, 2023; Nomoto et al., 2024). Malay poses unique linguistic challenges for AI tools due to its extensive use of affixation (Nik Safiah, 2010), context-dependent words such as functional "*pemeri*" *ialah* vs *adalah* (Hirobumi, 2011; Omar, 1993), and complex sentence structures that demand precise grammatical agreement (Ramli, 1995). These challenges highlight the need for tailored approaches to effectively generate and refine Malay academic texts. Furthermore, tools predominantly trained on English datasets often struggle to handle cultural nuances and localized expressions specific to Malay.

Gao et al. (2024) and Hu et al. (2024) highlight that ChatGPT's performance varies with task complexity, highlighting the necessity of expert oversight. This variability is particularly true in Malay, where cultural and linguistic nuances are integral to academic writing (Seong, 2014; Singh, 2015; Singh & Kaur, 2016). An in-depth evaluation of ChatGPT's ability to navigate these intricacies is essential for grasping its strengths and weaknesses in this language. Despite advances in AI text generation, challenges remain in achieving accuracy, coherence, and contextual relevance in underrepresented languages like Malay**.**

ChatGPT's reliance on large datasets for training presents risks of errors, particularly when dealing with complex linguistic structures or culturally specific contexts. Rojas (2024) emphasized the ethical concerns and the need for expert oversight when applying ChatGPT to academic tasks. Similarly, Niloy et al. (2024) cautioned against an over-reliance on AI, which may impact students' creative writing skills. These issues are even more pronounced in Malay, where context and cultural references play a significantly role in shaping meaning**.** These challenges underscore the importance of developing AI tools that can effectively address grammatical and contextual complexities.

To address these challenges, this study conducts a comparative analysis of ChatGPT-generated texts and their expert-corrected versions. Research by Ebrahimian et al. (2023) and Safrai and Orwig(2024) suggest that integrating AI tools with expert enhances the quality and reliability of academic writing. This collaborative approach maximizes the strengths of AI and expert judgment, mitigating the errors inherent in AI-generated texts. Furthermore, this research also contributes to the advancement of AI technologies for underrepresented languages worldwide by focusing on the Malay language.

This study contributes to developing of more effective AI-driven tools and instructional strategies for Malay academic writing by highlighting specific error categories and their implications. It investigates ChatGPT's accuracy and linguistic performance in generating Malay academic texts by comparing the AI outputs to expert-corrected versions. The findings aim to inform future advancements in AI linguistic models, enhance grammar instruction for Malay learners, and benefit other domains (e.g., content creation, and public communication) that require accurate Malay text generation**.** Ultimately, the study's focus on Malay's unique linguistic features offers a pathway to refine AI in various non-English contexts.

Method

This study employed a comparative research design (Mills et. al, 2008), broadly defined as "the evaluation of the similarities, differences, and associations between entities" to evaluate ChatGPT's performance to generate Malay academic texts. This approach is strategically chosen to analyze differences between texts generated under varying conditions, i.e. comparing outputs from structured prompts with input documents (TD) against those without input documents (UTD). Unlike experimental designs where variables can be controlled and manipulated, comparative research observe phenomena as it naturally occurs (in this case it is without artificial manipulation of the AI's internal parameters), suitable for analysis that requires researcher to look at nuances and detailed comparison across conditions.

This study has three main inclusion criteria. The first is articles in Malay academics from verifiable sources. Second, an expert reviewer in Malay linguistics should rely on recognized grammar references during evaluation. Third, the correction scope was limited to grammatical, syntactic, and lexical errors. Therefore, texts written in non-standard Malay and drawn from non-academic genres were excluded. Based on these criteria, 40 passages from peer-reviewed journal articles in linguistics, education, and the social sciences were selected, thereby capturing a wide range of linguistic complexity and cultural references. Each passage was converted into a prompt and submitted to ChatGPT in Malay, producing two datasets: one with structured, document-based prompts (Trained Dataset) and one with general prompts lacking supporting documents (Untrained Dataset).

1. Trained Dataset (TD): Texts generated using input documents and structured prompts.
   Example Prompt: *Carefully read the attached document and summarize its key sections. Write one paragraph each for the introduction, research purposes, results of the study, and conclusion. Ensure that the writing is in high-quality standard Malay, maintaining clarity, coherence, and an appropriate academic tone.*
2. Untrained Dataset (UTD): Texts generated using prompts without an input document.
   Example Prompt: *What are the main linguistic differences and similarities among Malay dialects from various regions? Discuss the implications of these variations for mutual intelligibility among speakers. Ensure that the writing is in high-quality standard Malay, maintaining clarity, coherence, and an appropriate academic tone.*

As the analysis focused on how prompt specificity affects ChatGPT's linguistic accuracy and grammatical consistency**,** three experts in Malay linguistics and academic writing reviewed the generated outputs, utilizing their expertise in linguistic research and publications. Their role included identifying and correcting errors in grammatical accuracy, contextual appropriateness, coherence and readability according to predefined linguistic guidelines. Corrections were made without altering the original meaning of the AI-generated text**,** serving as a benchmark for evaluating ChatGPT's grammatical accuracy. Instead of using a formal correction model, draws on the established tradition in linguistics, where native speakers' intuitive judgments about morphosyntactic well-formedness are considered valuable evidence for grammatical theories. With more than two decades of experience and academic training, these experts qua grammarian consulted established Malay grammar references when discrepancies arose to ensure consistency throughout assessments. This approach is also deemed to be suitable and appropriate, particularly given the importance especially, considering the importance of context and cultural relevance that may not be easily captured in formal correction models.

Procedures and data analysis

A systematic comparison of ChatGPT-generated texts and their expert-corrected versions was conducted through quantitative and qualitative analyses**.** The quantitative aspect measured grammatical error frequency and distribution, categorized them into predefined grammatical types and applied the descriptive statistics to compare the Trained Dataset (TD) and Untrained Dataset (UTD)**.** Inter-rater agreement measures ensured the reliability of expert corrections. The qualitative aspect assessed errors' impacts on clarity, coherence, and contextual relevance**,** identifying challenges such as affixation errors, noun omissions, and syntactic inconsistencies.

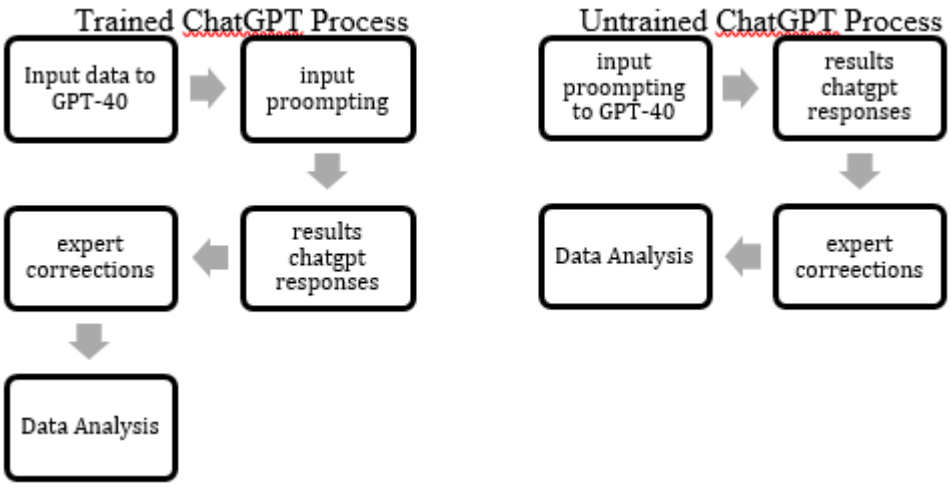

Figure 1. Procedures of data analysis (trained ChatGPT process vs. untrained ChatGPT process)

Figure 1 compares the processes of generating and analyzing Malay academic texts using trained and untrained versions of ChatGPT. The left panel illustrates the

trained approach, where structured input data is provided to GPT-40, along with detailed prompt engineering. The AI-generated responses are then reviewed and refined by experts with professional background in linguistics, Malay grammar, translation, and cultural studies using insights from error-analysis, descriptive grammar, and equivalence studies before undergoing final data analysis. In contrast, the right panel represents the untrained process, which involves providing minimal prompts to GPT-40. This results in raw responses that lack structured input. Expert subsequently review these responses before data analysis.

In the data analysis, these 40 texts were analyzed quantitative and qualitatively. The quantitative analysis involved tallying and classifying the errors identified in the TD and UTD into grammatical errors, incorrect use of conjunctions, 'pemeri' words of *adalah/ialah*, *iaitu* or which is/that is, word choices, names (nouns), and affixation, among others (refer to Table 1 for the complete summary) before applying descriptive statistics (for tabulation and comparison). To follow, qualitative analysis evaluated the impact of errors on the texts' clarity, coherence, and contextual relevance of the texts, with particular attention to recurring linguistic challenges unique to Malay academic writing. Tables and figures were used to present error frequencies and illustrate comparative findings.

Findings and Discussion

This section presents this study's findings, focusing on the linguistic errors observed in ChatGPT-generated Malay academic texts when compared to expert-corrected versions. Key areas include grammatical errors, incorrect sentence structures, misuse of words and phrases, and issues with '*pemeri*' words and affixation. The results are organized by error categories to maintain clarity and alignment with the study objectives.

Grammatical errors

The evaluation of ChatGPT's responses against expert corrections in addressing grammatical errors demonstrates notable differences in accuracy and clarity. One notable issue is the AI's occasional omission of explicit subjects such as *"Hal ini"* occasionally hinders the clarity of the context. This aligns with Rojas, (2024), who underscores the importance of a contextualized approach in AI text generation. Another critical factor is adverb placement. The transformation of phrase *"pembaca dapat dengan mudah memproses dan memahami maklumat tersebut"* to *"pembaca dapat memproses dan memahami maklumat tersebut dengan mudah"* illustrates how word order can greatly readability. Moreover, the insertion of *"bahawa"* in complementary clauses reflects ChatGPT's inconsistent handling of function words, a challenge similarly noted by Gao et al. (2024) for non-English languages. The language model also struggles with verb usage, often leading to omissions or misapplications that obscure meaning. Niloy et al. (2024) suggests that these gaps may arise from a lack of training data for lesser-studied languages, where correct verb forms are essential for conveying nuanced meanings. Lack of explicit verbs in examples 4 and 5 demonstrates how contextual cues are often missed by AI models primarily trained on English corpora.

#Example 1: Trained data (TD)

Category		Content
Grammatical Error	:	Subject Omission
ChatGPT Response	:	*Akhbar X menggunakan frasa seperti 'manipulasi GST' dan 'tipu rakyat' yang menonjol dan bermaksud memperdayakan rakyat untuk meraih undi. [ ] Ini menunjukkan bahawa akhbar X cenderung mempengaruhi persepsi pembaca terhadap Pakatan Harapan.*
Expert Correction	:	*Hal ini menunjukkan bahawa...*

#Example 2: Trained data (TD)

Category		Content
Grammatical Error	:	Phrase Structure
ChatGPT Response	:	*Penggunaan bentuk linguistik yang sesuai dapat menggambarkan kecenderungan penulis terhadap isu yang dilaporkan, dan dengan pemahaman latar belakang yang sama, pembaca dapat dengan mudah memproses dan memahami maklumat tersebut* ***[ ]***
Expert Correction	:	*Pembaca dapat memproses dan memahami maklumat tersebut* ***dengan mudah.***

#Example 3: Trained data (TD)

Category		Content
Grammatical Error	:	Incorrect complement sentences
ChatGPT Response	:	*Hasil kajian menunjukkan* ***[ ]*** *terdapat peningkatan yang signifikan dalam jumlah penerbitan mengenai BMB dari tahun 2013 hingga 2022, dengan sejumlah 482 artikel yang ditemui dalam pangkalan data Scopus pada 8 Mei 2023.*
Expert Correction	:	*Hasil kajian menunjukkan* ***bahawa*** *terdapat peningkatan yang signifikan dalam jumlah penerbitan mengenai BMB dari tahun 2013 hingga 2022.*

#Example 4: Trained data (TD)

Category		Content
Grammatical Error	:	Omission of verbs
ChatGPT Response	:	*Dengan* ***[ ]*** *pendekatan yang betul, korpus dapat terus memberikan sumbangan yang signifikan dalam penyelidikan linguistik Bahasa Melayu.*
Expert Correction	:	*Dengan* ***menggunakan*** *pendekatan yang betul, korpus dapat terus memberikan sumbangan...*

#Example 5: Untrained data (UTD)

Category		Content
Grammatical Error	:	Incorrect sentence structure

Category		Content
ChatGPT Response	:	*Selain itu, bahasa Melayu juga digunakan ~~dalam upacara adat seperti perkahwinan, kenduri, dan majlis kesyukuran, di mana pantun dan doa-doa tradisional dibacakan.~~*
Expert Correction	:	*Selain itu, bahasa Melayu juga digunakan **untuk berpantun dan membaca doa-doa tradisional dalam upacara adat seperti perkahwinan, kenduri, dan majlis kesyukuran.***

#Example 6: Untrained data (UTD)

Category		Content
Grammatical Error	:	Use of transitive verb phrases + prepositional phrase
ChatGPT Response	:	*Penggunaan bahasa ini dalam pelbagai aspek kehidupan sosial dan politik **menyumbang[ ]** kepada kohesi sosial dan mengurangkan ketegangan etnik dengan menyediakan satu platform komunikasi yang diterima umum.*
Expert Correction	:	*Penggunaan bahasa ini dalam pelbagai aspek kehidupan sosial dan politik **menyumbangkan** kohesi sosial dan mengurangkan ketegangan etnik dengan...*

#Example 7: Untrained data (UTD)

Category		Content
Grammatical Error	:	Sentence structure
ChatGPT Response	:	*Untuk memastikan kelestarian bahasa Melayu, adalah penting **[ ]** untuk mengimbangi **~~antara menerima~~** pengaruh asing **~~dan mengekalkan~~** identiti bahasa asli **[ ]**dengan mempromosikan penggunaan istilah Melayu yang autentik dan menggalakkan pembelajaran bahasa secara mendalam.*
Expert Correction	:	***usaha untuk mengimbangi penerimaan** pengaruh asing **dan pengekalan** identiti bahasa asli **adalah penting** dengan mempromosi penggunaan istilah Melayu...*

#Example 8: Untrained data (UTD)

Category		Contents
Grammatical Error	:	Sentence structure
ChatGPT Response	:	*Walaupun kaedah-kaedah ini menunjukkan keberkesanan dalam meningkatkan minat dan pemahaman pelajar, **~~masih terdapat keperluan untuk~~** penambahbaikan berterusan bagi memastikan semua jurang pendidikan dapat diatasi dengan lebih berkesan, terutamanya di kawasan luar bandar yang mungkin kurang akses kepada teknologi moden**…[ ]***
Expert Correction	:	*penambahbaikan berterusan bagi memastikan semua jurang pendidikan dapat diatasi dengan lebih berkesan, terutamanya di kawasan luar bandar yang mungkin kurang akses pada teknologi moden **masih perlu dilakukan**.*

From a computational linguistics perspective, these recurring errors underscore ChatGPT's limitations in handling specific Malay grammatical features**,** especially concerning subject omission, transitivity and verb forms. While the

Trained Dataset (TD) exhibited fewer errors of affixation and subject omission compared to the Untrained Dataset (UTD), improvements were consistent across all error types. This inconsistency suggests that while providing more context in prompts can mitigate certain error patterns it does not resolve deeper linguistic complexities (Gao et al., 2024a; Rojas, 2024). Improving Malay language models will require expanded datasets incorporating diverse sentence structures, culturally nuanced examples, and explicit grammatical rules. Integrating curated Malay linguistic corpora could significantly enhance AI performance in academic writing.

Incorrect sentence structure

The document analysis reveals multiple preposition errors that affect sentence clarity, such as redundant usage, incorrect substitutions, and omissions. Rojas, (2024) underscores how precise language structures enhance AI-generated text. For instance, removing an unnecessary "*kepada*" sharpens meaning, reflecting the tailored syntactic approach**.** As Niloy et al. (2024) suggest, non-English AI outputs often require refined syntactic rules. In Malay, preposition errors obscure conceptual links, as seen when *"terhadap"* should replace *"dalam"* to accurately reflect the object of impact.

#Example 9: Trained data (TD)

Category		Content
Incorrect Sentence Structure	:	Redundant Preposition
ChatGPT Response	:	*Selain itu, hasil kajian ini juga* ~~***memiliki***~~ *implikasi* ***[ ]*** *dalam pengembangan hala tuju dasar untuk menyokong penyelidikan lebih lanjut dalam bidang ini.*
Expert Correction	:	*Selain itu, hasil kajian ini juga* ***memberikan*** *implikasi* ***penting*** *terhadap pengembangan hala tuju dasar untuk menyokong penyelidikan lebih lanjut dalam bidang ini.*

#Example 10: Trained data (TD)

Category		Content
Incorrect Sentence Structure	:	Incorrect Substitution
ChatGPT Response	:	*Dengan* ***[ ]*** *pendekatan yang betul, korpus dapat terus memberikan sumbangan yang signifikan* ~~***dalam***~~ *penyelidikan linguistik Bahasa Melayu.*
Expert Correction	:	*Dengan* ***menggunakan*** *pendekatan yang betul, Korpus dapat terus memberikan sumbangan yang signifikan* ***terhadap*** *penyelidikan linguistik Bahasa Melayu.*

#Example 11: Untrained data (TD)

Category		Content
Incorrect Sentence Structure	:	Incorrect Substitution
ChatGPT Response	:	*Contohnya, dialek Kelantan mempunyai sebutan yang lebih nasal dan berbeza* ~~***dalam***~~ *penggunaan kata ganti nama diri*

Category		Content
		seperti "demo" (kamu) ***~~berbanding~~ [ ]*** *"awak" dalam dialek Johor.*
Expert Correction	:	*Contohnya, dialek Kelantan mempunyai sebutan yang lebih nasal dan berbeza* ***dari segi*** *penggunaan kata ganti nama diri seperti "demo" (kamu)* ***dibandingkan dengan*** *"awak" dalam dialek Johor.*

#Example 12: Untrained data (TD)

Category		Content
Incorrect Sentence Structure	:	Incorrect Substitution
ChatGPT Response	:	*Di Indonesia, Bahasa Indonesia, yang merupakan varian Bahasa Melayu, digunakan sebagai bahasa rasmi dan lingua franca yang menghubungkan lebih* ***dari[ ]*** *700 etnik yang berlainan, sekali gus membina identiti nasional yang kukuh.*
Expert Correction	:	*Digunakan sebagai bahasa rasmi dan lingua franca yang menghubungkan lebih dari****pada*** *700 etnik yang berlainan…*

#Example 13: Untrained data (TD)

Category		Content
Incorrect Sentence Structure	:	Redundant Preposition
ChatGPT Response	:	*Yang menjadi asas* ***~~kepada~~*** *identiti etnik Brunei. Penggunaan bahasa ini dalam pelbagai aspek kehidupan sosial dan politik* ***~~menyumbang kepada~~*** *kohesi sosial dan mengurangkan ketegangan etnik dengan menyediakan satu platform komunikasi yang diterima umum.*
Expert Correction	:	*Yang menjadi asas* ***bagi*** *identiti etnik Brunei. Penggunaan bahasa ini dalam pelbagai aspek kehidupan sosial dan politik* ***menyumbangkan*** *kohesi sosial…*

#Example 14: Untrained data (TD)

Category		Content
Incorrect Sentence Structure	:	Incorrect Substitution
ChatGPT Response	:	*Globalisasi telah memberikan kesan yang mendalam terhadap bahasa Melayu, terutamanya melalui kemasukan banyak kata pinjaman* ***dari[ ]*** *bahasa asing seperti* ***[ ]****Inggeris,* ***[ ]*** *Arab,* ***[ ]****Portugis, dan* ***[ ]*** *Belanda.*
Expert Correction	:	*...terutamanya melalui kemasukan banyak kata pinjaman dari****pada*** *bahasa asing seperti* ***bahasa*** *Inggeris,* ***bahasa*** *Arab,* ***bahasa*** *Portugis, dan* ***bahasa*** *Belanda.*

#Example 15: Untrained data (TD)

Category		Content
Incorrect Sentence Structure	:	Omitted Preposition

Category		Content
ChatGPT Response	:	*mempromosikan bahasa Melayu* ~~***di***~~ *peringkat antarabangsa melalui media digital, pendidikan, dan pertukaran budaya.*
Expert Correction	:	*mempromosikan bahasa Melayu* ***pada*** *peringkat antarabangsa melalui media digital, pendidikan, dan pertukaran budaya.*

The incorrect use of conjunctions

The examples above illustrate specific corrections, emphasizing the importance of precise conjunction usage in Malay syntax for enhancing clarity and academic tone (Gao et al., 2024b; Rojas, 2024; Sr, 2024). For instance**,** *"malahan"* adds emphasis, creating a stronger additive relationship than *"tetapi juga,"* which aligns with the academic style often expected in formal texts. The omission of conjunctions in examples 18, 20, and 26 highlights missing words such as *"untuk"* or *"dan,"* which can make the sentences incomplete or ambiguous. The misuse of conjunctions in Examples 16, 17, 19, 21, and 23 involves incorrect or vague connector choices, resulting in weakened emphasis or unclear additive relationships. This issue often stems from constructions influenced by English (Hu et al., 2024). In examples 22 and 24, redundant or incorrectly placed conjunctions show how extra words or misplaced connectors can disrupt sentence flow and academic rigor.

Omitting a critical conjunction can obscure the logical flow between clauses. For instance, use of *"untuk"* in Example 18 clarifies the purpose of the study. Similarly, example 25 shows how *"malahan"* strengthens the additive function, emphasizing the extended role of Malay in preserving history and introducing new expressions. Without these markers, readers may struggle to grasp the nuances of academic arguments.

#Example 16: Trained data (TD)

Category		Content
Incorrect use of conjunctions	:	Incorrect use of conjunctions
ChatGPT Response	:	*Frasa ini bukan sahaja berfungsi sebagai klu untuk memudahkan pembaca menginterpretasi maklumat,* ***[ ]*** *mencerminkan kecenderungan nilai dan pandangan penulis terhadap isu tersebut****.[ ]***
Expert Correction	:	*Frasa ini bukan sahaja berfungsi sebagai klu untuk memudahkan pembaca menginterpretasi maklumat,* ***malahan*** *mencerminkan kecenderungan nilai dan pandangan penulis terhadap isu tersebut* ***juga.***

#Example 17: Trained data (TD)

Category		Content
Incorrect use of conjunctions	:	Incorrect use of conjunctions
ChatGPT Response	:	*Pada zaman ini, akhbar bukan sahaja menyampaikan berita dan maklumat kepada masyarakat,* ~~***tetapi***~~ *juga membantu masyarakat menyesuaikan diri dengan perubahan dan membimbing mereka ke arah kemajuan dan pembaikan.*
Expert Correction	:	*Pada zaman ini, akhbar bukan sahaja menyampaikan berita dan maklumat kepada masyarakat,* ***malahan*** *membantu masyarakat*

Category	Content
	menyesuaikan diri dengan perubahan dan membimbing mereka ke arah kemajuan dan pembaikan juga.

#Example 18: Trained data (TD)

Category		Content
Incorrect use of conjunctions	:	Omission of conjunction
ChatGPT Response	:	*Kajian ini berusaha **[ ]** menjelaskan cara input linguistik pada peringkat semantik dapat diinterpretasikan untuk menjelaskan maksud sebenar yang ingin disampaikan oleh laporan akhbar tersebut.*
Expert Correction	:	*Kajian ini berusaha **untuk** menjelaskan cara input linguistik pada peringkat semantik dapat diinterpretasi bagi menjelaskan maksud sebenar yang ingin disampaikan oleh laporan akhbar tersebut.*

#Example 19: Trained data (TD)

Category		Content
Incorrect use of conjunctions	:	Incorrect use of conjunctions
ChatGPT Response	:	*Dapat dianalisis**[ ]** menggunakan Pendekatan Pragmatik, iaitu Teori Relevans.*
Expert Correction	:	*Dapat dianalisis **dengan** menggunakan pendekatan Pragmatik, iaitu Teori Relevans.*

#Example 20: Trained data (TD)

Category		Content
Incorrect use of conjunctions	:	Omission of prepositions
ChatGPT Response	:	*Tulisan ini menjadi asas **[ ]** evolusi sistem tulisan Melayu kuno.*
Expert Correction	:	*Tulisan ini menjadi asas **bagi** evolusi sistem tulisan Melayu kuno.*

#Example 21: Untrained data (UTD)

Category		Content
Incorrect use of conjunctions	:	Incorrect use of conjunctions
ChatGPT Response	:	*Inskripsi kuno ini bukan sahaja menunjukkan kekayaan sejarah bahasa Melayu **[ ]** mempengaruhi bahasa Melayu moden dari segi kosa kata, struktur, dan tulisan **[ ]***
Expert Correction	:	*Inskripsi kuno ini bukan sahaja menunjukkan kekayaan sejarah bahasa Melayu, **malahan** mempengaruhi bahasa Melayu moden dari segi kosa kata, struktur, dan tulisan **juga**.*

#Example 22: Untrained data (TD)

Category		Content
Incorrect sentence structure	:	Incorrect sentence structure

Category		Content
ChatGPT Response	:	*Selain itu, bahasa Melayu juga digunakan* ***~~dalam upacara adat seperti perkahwinan, kenduri, dan majlis kesyukuran, di mana pantun dan doa-doa tradisional dibacakan.~~***
Expert Correction	:	*Selain itu, bahasa Melayu juga digunakan* ***untuk berpantun dan membaca doa-doa tradisional dalam*** *upacara adat seperti perkahwinan, kenduri, dan majlis kesyukuran.*

#Example 23: Untrained data (TD)

Category		Content
Incorrect use of conjunctions	:	Incorrect use of conjunctions
ChatGPT Response	:	*Bahasa Melayu* ***~~tidak~~*** *hanya menjadi medium komunikasi* ***[ ]*** *simbol identiti dan alat perpaduan nasional* ***[ ]***
Expert Correction	:	*Bahasa Melayu* ***bukan*** *hanya menjadi medium komunikasi,* ***malahan menjadi*** *simbol identiti dan alat perpaduan nasional* ***juga.***

#Example 24: Untrained data (TD)

Category		Content
Incorrect sentence structure	:	Incorrect sentence structure
ChatGPT Response	:	*Di Indonesia, Bahasa Indonesia berfungsi sebagai lingua franca,* ***[ ]*** *penduduk juga fasih dalam bahasa daerah seperti Jawa, Sunda.*
Expert Correction	:	*Di Indonesia, Bahasa Indonesia berfungsi sebagai lingua franca,* ***namun ramai*** *penduduk juga fasih dalam bahasa daerah seperti Jawa, Sunda.*

#Example 25: Untrained Data (TD)

Category		Content
Grammatical Error	:	Incorrect use of conjunctions
ChatGPT Response	:	*Bukan sahaja memelihara sejarah dan mitos,* ***[ ]*** *memperkenalkan ungkapan, istilah, dan struktur Bahasa* ***[ ].***
Expert Correction	:	*Bukan sahaja memelihara sejarah dan mitos,* ***malahan*** *memperkenalkan ungkapan, istilah, dan struktur bahasa* ***juga.***

#Example 26: Untrained Data (TD)

Category		Content
Grammatical Error	:	Incorrect use of conjunctions
ChatGPT Response	:	*Yang sederhana dan sering diakhiri dengan vokal,* ***[ ]*** *menjadikannya mudah diucapkan.*
Expert Correction	:	*Yang sederhana dan sering diakhiri dengan vokal,* ***dan*** *menjadikannya mudah diucapkan.*

Incorrect use of 'pemeri' words

In *Bahasa Melayu, "adalah"* is typically used for non-definitive or descriptive clauses, while *"ialah"* links the subject to a definitive or specified predicate. Rojas (2024), notes that omitting unnecessary *pemeri* words can enhance fluency in academic texts. For instance, example 27 shows how removing *"adalah"*

eliminates redundancy**,** making the sentence more concise. Gao et al. (2024b) suggest that English-influenced models may overuse forms akin to "is," leading to awkward formality. Misuse of "*Adalah*" vs *"Ialah*" in Example 28, substituting *"bahasa Melayu ialah bahasa aglutinatif"* for *"bahasa Melayu adalah bahasa aglutinatif"* emphasizes definitive identification. Failing to use *"ialah"* can obscure the logical relationship between subject and predicate, which identifies as a common AI shortcoming in lesser-studied languages. While omitting "*adalah*" can improve clarity when it adds no semantic value, excessive omission may risk losing important nuances.

#Example 27: Trained data (TD)

Category		Content
Incorrect use of 'pemeri' words	:	Incorrect use of 'pemeri' words
ChatGPT Response	:	*Dengan memastikan berita yang disampaikan* ~~***adalah***~~ *berasaskan fakta, baharu, menarik minat, dan tidak berat sebelah untuk mengelakkan kemusykilan dalam kalangan rakyat.*
Expert Correction	:	*Dengan memastikan berita yang disampaikan berasaskan fakta, baharu, menarik minat, dan tidak berat sebelah…*

#Example 28: Untrained data (TD)

Category		Content
Incorrect use of 'pemeri' words	:	Incorrect use of 'pemeri' words
ChatGPT Response	:	*Dari segi tatabahasa, bahasa Melayu* ~~***adalah***~~ *bahasa aglutinatif yang menggunakan imbuhan seperti awalan, akhiran, apitan…*
Expert Correction	:	*Dari segi tatabahasa, bahasa Melayu* ***ialah*** *bahasa aglutinatif yang menggunakan imbuhan seperti awalan…*

The usage of which is/that is (iaitu)

The connector *"iaitu"* in Bahasa Melayu serves two main purposes: (1) providing specific clarification about a subject or timeframe, and (2) ensuring smooth sentence flow by preventing ambiguity. In example 29**,** inserting "*iaitu*" explicitly defines the period 2013–2022, aligning with Gao et al. (2024b) emphasis on precise connectors in AI-generated texts. Employing *"iaitu"* not only clarifies the time frame but also reinforces a formal academic. Omitting it may cause readers to assume that "*2013 hingga 2022*" is merely an appositive phrase rather than a defining clause. By explicitly stating *"iaitu,"* the sentence structure becomes more coherent, supporting the goal of concise yet unambiguous communication. In TD, contextual prompts often reduce omissions of *"iaitu."* However, UTD may omit it entirely, mirroring English-based simplifications.

#Example 29: Trained data (TD)

Category		Content
The usage of "which is/that is" (iaitu)	:	The usage of "which is/that is" (iaitu)
ChatGPT Response	:	*Bahasa Melayu Brunei (BMB) sepanjang dekad yang lalu,* ***[ ]*** *dari tahun 2013 hingga 2022.*

Category		Content
Expert Correction	:	*Bahasa Melayu Brunei (BMB) sepanjang dekad yang lalu,* ***iaitu*** *dari tahun 2013 hingga 2022.*

Inappropriate word choices

Choosing precise and concise words is essential for clarity and precision in Malay academic writing. The redundancy and imprecise word usage can weaken the academic tone, and AI models often default to generic terms when prompt lacks context. In examples 30 and 31, replacing *"titik rujukan"* with *"rujukan"* and *"memiliki implikasi*" with *"memberikan implikasi"* eliminates unnecessary or less precise terms, aligning with Malay norms for succinct communication. The shift from *"relevansi"* to *"kerelevanan"* in example 32 underscores the grammatical expectations in Bahasa Melayu. Niloy et al. (2024) note that AI-driven outputs may overlook morphological forms when trained primarily on English-centric data. As Hu et al. (2024), highlight, lesser-studied languages face persistent lexical mismatches, reinforcing the need for explicit Malay usage rules. Meanwhile, substituting "Influencer" and "content creator" in example 34 with "*Pempengaruh*" and "*pencipta kandungan"* maintains lexical consistency in formal Malay contexts.

#Example 30: Trained data (TD)

Category		Content
Grammatical Error	:	Inappropriate word choices
ChatGPT Response	:	*Hasil kajian ini dapat dijadikan* ~~***titik***~~ *rujukan bagi penyelidikan lanjut dan...*
Expert Correction	:	*Hasil kajian ini dapat dijadikan rujukan bagi penyelidikan lanjut dan...*

#Example 31: Trained data (TD)

Category		Content
Inappropriate word choices	:	Inappropriate word choices
ChatGPT Response	:	*Selain itu, hasil kajian ini juga* ~~***memiliki***~~ *implikasi penting dalam pengembangan hala...*
Expert Correction	:	*Selain itu, hasil kajian ini juga* ***memberikan*** *implikasi penting terhadap perkembangan...*

#Example 32: Trained data (TD)

Category		Content
Inappropriate word choices	:	Inappropriate word choices
ChatGPT Response	:	*Korpus harus sentiasa dikemas kini dan dijana daripada sumber yang* ~~***baru***~~ *untuk memastikan* ~~***relevansi***~~ *dan...*
Expert Correction	:	*Korpus harus sentiasa dikemas kini dan dijana daripada sumber yang* ***baharu*** *untuk memastikan* ***kerelevanan*** *dan...*

#Example 33: Untrained data (TD)

Category	Content
Inappropriate word choices	Inappropriate word choices

Category	Content
ChatGPT Response	*Penggunaan kata ganti nama diri seperti "demo" (kamu) **~~berbanding~~** "awak" dalam dialek Johor.*
Expert Correction	*Penggunaan kata ganti nama diri seperti "demo" (kamu) **manakala** "awak" dalam dialek Johor.*

#Example 34: Untrained data (TD)

Category		Content
Inappropriate word choices	:	Inappropriate word choices
ChatGPT Response	:	***~~Influencer~~*** *dan* ***~~content creator~~*** *juga menggunakan Bahasa Melayu untuk menghasilkan kandungan yang menarik dan...*
Expert Correction	:	***Pempengaruh*** *dan* ***pencipta kandungan*** *juga menggunakan bahasa Melayu untuk menghasilkan kandungan yang menarik dan...*

#Example 35: Untrained data (TD)

Category		Content
Grammatical Error	:	Inappropriate word choices
ChatGPT Response	:	*Penggunaan bahasa Melayu di platform digital ini* ***[ ] ~~membantu~~ ~~mempromosikan~~*** *dan mengekalkan* ***~~relevansi~~*** *bahasa dalam era teknologi...*
Expert Correction	:	*Penggunaan Bahasa Melayu di platform digital ini* ***dapat meningkatkan usaha untuk*** *mempromosi dan mengekalkan* ***kerelevanan*** *bahasa dalam era teknologi...*

Omission of name (nouns)

The omission of specific nouns (e.g., *"dialek," "bahasa,"* or *"hal")* can lead to ambiguity and reduced clarity in Malay academic writing. The omitting critical subjects confuse readers, especially if the context is not explicit. Gao et al. (2024b) suggest that English-based simplifications in AI training might overlook the importance of repeated noun references, leading to truncated or unclear phrases in Malay.

#Example 36: Untrained data (TD)

Category		Content
Omission of name (nouns)	:	Omission of name (nouns)
ChatGPT Response	:	*Bahasa Melayu mempunyai pelbagai dialek yang berbeza mengikut kawasan, seperti dialek Kelantan,* ***[ ]****Terengganu,* ***[ ]*** *Kedah, dan* ***[ ]*** *Johor di Malaysia,* ***~~serta~~*** *dialek Riau,* ***[ ]*** *Palembang,* ***dan [ ]*** *Minangkabau di Indonesia.*
Expert Correction	:	*...mengikut kawasan, seperti dialek Kelantan,* ***dialek*** *Terengganu,* ***dialek*** *Kedah,* ***dialek*** *Johor di Malaysia,* ***dan*** *dialek Riau,* ***dialek*** *Palembang,* ***serta dialek*** *Minangkabau di Indonesia.*

#Example 37: Untrained data (TD)

Category		Content
Omission of name (nouns)	:	Omission of name (nouns)
ChatGPT Response	:	*Kemasukan banyak kata pinjaman dari bahasa asing seperti* ***[ ]*** *Inggeris,* ***[ ]*** *Arab,* ***[ ]*** *Portugis, dan* ***[ ]*** *Belanda.*

Category		Content
Expert Correction	:	*Kemasukan banyak kata pinjaman daripada bahasa asing seperti **bahasa** Inggeris, **bahasa** Arab, **bahasa** Portugis, dan **bahasa** Belanda.*

#Example 38: Untrained data (TD)

Category		Content
Grammatical Error	:	Omission of name (nouns)
ChatGPT Response	:	*Walaupun **[ ]** ini memperkayakan bahasa dari segi perbendaharaan kata, ia juga menimbulkan cabaran dalam mengekalkan keaslian.*
Expert Correction	:	*Walaupun **hal** ini memperkaya bahasa ……*

Incorrect affixation

#Example 39: Trained data (TD)

Category		Content
Grammatical Error	:	Incorrect affixation
ChatGPT Response	:	*Pada peringkat semantik dapat ~~**diinterpretasikan**~~ untuk menjelaskan maksud sebenar yang ingin disampaikan oleh laporan akhbar tersebut.*
Expert Correction	:	*Pada peringkat semantik, **xxxxx** dapat **diinterpretasi** untuk menjelaskan maksud sebenar yang ingin disampaikan oleh laporan akhbar tersebut.*

#Example 40: Untrained data (TD)

Category		Content
Incorrect affixation	:	Incorrect affixation
ChatGPT Response	:	***~~Hikayat, seperti "Hikayat Hang Tuah" dan "Hikayat Raja Pasai," mencatatkan~~*** *sejarah dan mitos yang mengandungi pelajaran moral ….*
Expert Correction	:	***"Hikayat Hang Tuah" dan "Hikayat Raja Pasai," mencatat sejarah** dan mitos...*

In Bahasa Melayu, affixes such as *"-kan," "me-," "di-"* are crucial for grammatical clarity. However, excessive use or incorrect application can lead to confusion. The AI models primarily trained on English data might default to literal morphological structures, leading to unnecessary complexity. In example 39**,** removing the affix "*-kan*" from "*diinterpretasikan*" to "*diinterpretasi*" aligns with Malay norms for brevity and clarity. Complex verbs, like "*mencatatkan*" in example 40, can hinder reading flow. By reverting to *"mencatat*" (causative), the sentence becomes direct and easier to comprehend. That models lacking diverse training data often struggle with morphological errors, highlighting the importance of simplification in formal Malay contexts.

Table 1 consolidates all the grammatical error types identified across the 40 examples, categorized into Trained Data (TD) and Untrained Data (UTD). Each row indicates the number of times a particular error type appears in TD and UTD,

followed by the total and corresponding percentages. The "Expert Correction" (shown in red in the source data) was used to confirm each instance of error.

Table 1. Grammatical error types in trained vs. untrained datasets

Grammatical Error	Freq.Rev. words (TD)	% (TD)	Freq.Rev. words (UTD)	% (UTD)
Subject Omission	1	5.56%		
Phrase Structure	2	11.11%		
Incorrect Complement Sentences	1	5.56%		
Omission of Verbs	1	5.56%		
Incorrect Sentence Structure			24	27.59%
Use of Transitive Verb Phrases + Prepositional Phrase			1	1.15%
Sentence structure			11	12.64%
Redundant Preposition	2	11.11%	7	8.05%
Incorrect Substitution	2	11.11%	3	3.45%
Omitted Preposition			1	1.15%
Incorrect Use of Conjunctions	2	11.11%	7	8.05%
Incorrect use of 'pemeri' words	1	5.56%	1	1.15%
The usage of 'which is/that is' (iaitu)	1	5.56%		
Inappropriate word choices	4	22.22%	10	11.49%
Omission of name (nouns)			13	14.94%
Incorrect affixation	1	5.56%	9	10.34%
Total	18	100.00%	87	100.00%

Table 1 presents a comparative analysis of grammatical errors in the TD and UTD**,** focusing on their frequency and revision word counts. The total number of grammatical errors in the UTD dataset (87 errors) is significantly higher than in the TD dataset (18 errors)**,** indicating that the trained dataset exhibits greater grammatical accuracy. Among the errors found in the UTD, the most prevalent issue is Incorrect Sentence Structure (27.59%)**,** followed by Omission of Name (Nouns) (14.94%) and Inappropriate Word Choices (11.49%)**.** In contrast, in the TD, the highest proportion of errors is attributed to Inappropriate Word Choices (22.22%)**,** followed by Phrase Structure (11.11%)**,** Incorrect Substitution (11.11%)**,** Redundant Preposition (11.11%)**,** and Incorrect Use of Conjunctions (11.11%)**.** Notably, certain error categories, such as Use of Transitive Verb Phrases + Prepositional Phrase (1.15%)**,** Omitted Preposition (1.15%)**,** and Incorrect Use of *'Pemeri'* Words (1.15%)**,** occur at much lower frequencies in the UTD, suggesting that these errors may not be as pervasive in untrained AI-generated text.

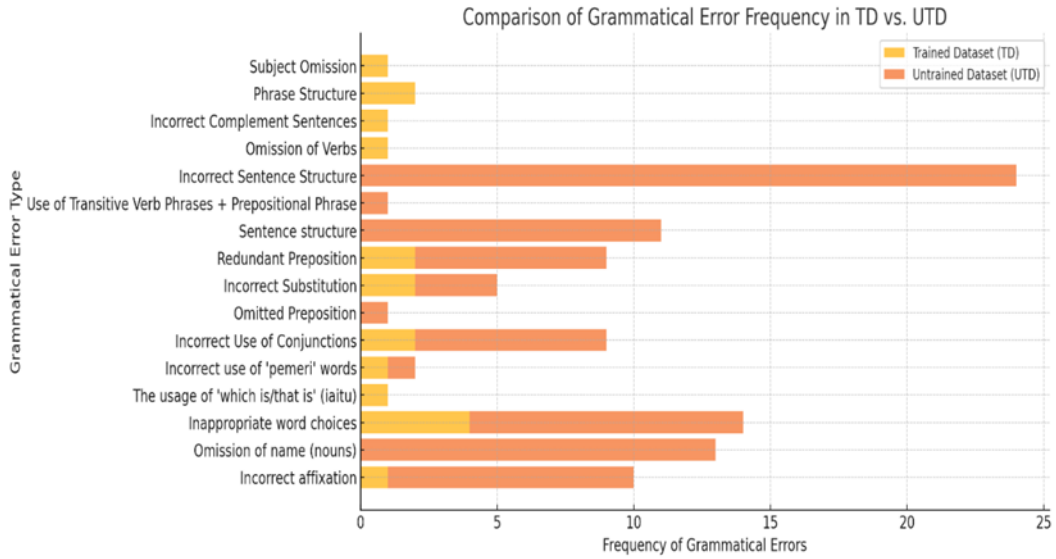

Figure 2. Comparison of grammatical error frequency in TD vs. UTD

The comparison between trained (TD) and untrained (UTD) outputs from ChatGPT indicates significant advancements in grammatical accuracy for AI-generated Malay text. While the TD dataset contains fewer errors, it still contains issues related to word choice, phrase structure, and conjunction usage**.** In contrast, UTD reveals more severe problems, including poor sentence structure and noun omissions, underscoring the need for further refinement in AI-generated content. The high standard deviation in the UTD dataset points to inconsistent error distribution**,** with certain categories showing disproportionately higher error rates. This inconsistency implies that AI-generated Malay texts, when untrained, lack grammatical stability and coherence**,** reinforcing the necessity of further training and integration of linguistic rules.

These findings have important implications for research on the quality of AI-generated text especially in low-resource languages like Malay. The results confirm that ChatGPT performs significantly better when trained and fine-tuned**,** but will still encounter challenges in navigating the intricacies of Malay grammar. The high frequency of Incorrect Sentence Structure errors in the UTD dataset suggests a critical need for improved syntactic rule integration in AI-generated texts**.** Moreover, this study can serve as a valuable reference for future AI language model training efforts**,** particularly for low-resource languages that require additional linguistic adaptation to improve accuracy and coherence. By addressing these identified grammatical weaknesses, future iterations of AI-driven text generators can enhance their linguistic performance**,** ensuring they produce higher-quality academic writing in Malay.

Discussion

This study demonstrates a notable disparity in the grammatical accuracy and linguistic quality of ChatGPT-generated Malay academic texts depending on the specificity of prompts and presence of trained data. The Untrained Dataset (UTD) showed a higher frequency of grammatical errors, with prominent issues in incorrect sentence structure, omission of names, and inappropriate word choices. In

contrast, the Trained Dataset (TD) primarily had errors related to word choice, phrase structure, and incorrect substitution**.** The UTD also showed issues such as misuse of transitive verb, omitted prepositions, and incorrect *'pemeri'* words. These findings highlight the critical role of prompt engineering and targeted training in enhancing AI-generated text quality for underrepresented languages such as Malay (Nomoto, 2023). These findings also align with Gao et al. (2024)**,** who emphasize that ChatGPT's performance in non-English contexts is influenced by linguistic complexity and training data quality**.**

While trained prompts can significantly improve ChatGPT's ability to generate coherent and accurate academic texts in Malay (Dergaa et al., 2023; Kostka & Toncelli, 2023) grammatical challenges such as problems with phrase structure, affixation, and conjunction are still persistent -- indicating that prompt engineering alone is insufficient to meet academic standards (Dubey, 2024; Glahn, 2024; Rojas, 2024; Safrai & Orwig, 2024). In fact, studies conducted by Rojas (2024) and Gao et al. (2024) on low-resource language contexts have shown that there is a high rate of incorrect sentence structures and noun omissions in UTD, reinforcing the need for AI models to be fine-tuned with language-specific rules and datasets. Similarly, studies recorded by Daungsupawong and Wiwanitkit (2023) and Li (2024) have also shown a much higher error frequency and greater variation in linguistic accuracy**,** reinforcing the necessity for clear and structured prompt design to enhance AI-generated outputs in non-English academic contexts.

The study highlights unique linguistic challenges in AI-generated Malay texts, demonstrating the complexity of adapting AI models to underrepresented languages (Nomoto et al., 2024) especially in maintaining proper sentence flow and logical structure. Despite issues with lexical choices, others include Malay's structural complexity, particularly affixation, '*pemeri*' words (*ialah* vs. *adalah*), and transitive verb constructions, which ChatGPT often misuses or omits entirely. ChatGPT frequently misapplies conjunctions and prepositions**,** leading to ambiguity and awkward phrasing**,** particularly in UTD. These challenges are compounded by the model's tendency to default to English-influenced sentence structures, leading to inappropriate word choices and overgeneralization of Malay grammatical rules—a phenomenon also reported by Niloy et al. (2024). Comparatively, research on major languages including English have demonstrated that large language models benefit from extensive, high-quality training corpora, which are often lacking for Malay and similar languages.

To this, expert is required for refining AI-generated Malay texts to ensure grammatical correctness, clarity, and coherence. Expert intervention was particularly valuable in refining word choice, sentence structure, and affixation, confirming Ebrahimian et al. (2023) and Safrai and Orwig (2024)'s argument that a collaborative approach between AI and expert expertise yields the most reliable academic writing. Furthermore, the presence of systematic weaknesses in ChatGPT's processing of Malay grammar suggests that AI models must be trained specifically on Malay academic texts to enhance linguistic accuracy. Without such linguistic fine-tuning, AI-generated texts will continue to struggle with Malay-specific grammatical structures and cultural nuances**.**

Conclusion

This study examines the effects of prompt engineering on the quality of ChatGPT-generated academic texts in Malay. Results indicate that the Trained Dataset (TD) outperformed the Untrained Dataset (UTD) in grammatical accuracy, with the UTD facing issues such as incorrect sentence structures, omission of names, and inappropriate word choices**.** While context-rich prompts improved the TD, persistent linguistic challenges such as affixation errors, misuse of conjunctions, and structural inconsistencies indicate that ChatGPT's current language model struggles with Malay's grammatical complexity. While this study reinforces the presence of these challenges, it enhances our theoretical understanding of AI limitations when working with morphologically rich languages emphasizing the necessity for expert oversight in refining AI-generated texts and the importance of training AI models with diverse, context-rich Malay datasets**.** Despite limited by size of dataset and domain focus, insights from this study also contributes to the growing research on AI-driven text generation in non-English academic contexts, highlighting key areas where Malay linguistic adaptation remains underdeveloped**.** Future studies should address these limitations by scaling dataset diversity and broadening domain coverage. At the same time, it also emphasizes that linguistic diversity in training data is essential for achieving robust multilingual AI performance. Thus, future research should focus on fine-tuning AI models with syntactic rule integration, expanding Malay linguistic corpora, and exploring AI-assisted writing tools to support academic authorship in Malay.

Additionally, this study reinforces the importance of prompt engineering, indicating that prompt specificity and contextual awareness are crucial for eliciting higher-quality outputs from larger language models. Ultimately, this study advocates for hybrid AI advancements and expert expertise in improving the accuracy and usability of AI-generated texts for underrepresented languages. Several key areas for future research emerge from these findings. First, refining AI models for Malay academic writing is crucial, particularly in addressing grammatical accuracy, affixation, and syntactic complexity, so that it can be utilized as guidance for educators and instructors alike, especially in designing curricula and assessments. Second, expanding Malay linguistic datasets by incorporating formal academic writing, research papers, and high-quality translations will benefit AI training. Third, conducting comparative studies with other low-resource languages sheds light on AI's adaptability to various grammatical structures motivates institutions and policymakers to invest in resources that support underrepresented languages. Finally, investigating the effectiveness of AI-assisted writing tools in academic settings will support researchers, students, and educators in producing quality work.

References

Adıgüzel, T., Kaya, M. H., & Cansu, F. K. (2023). Revolutionizing education with AI: Exploring the transformative potential of ChatGPT. *Contemporary Educational Technology*, *15*(3), ep429. https://doi.org/10.30935/cedtech/13152

Daungsupawong, H., & Wiwanitkit, V. (2023). Comment on the use of artificial intelligence in writing scientific papers. *Brain Communications*, *6*(1), fcad354. https://doi.org/10.1093/braincomms/fcad354

Dergaa, I., Chamari, K., Żmijewski, P., & Saad, H. B. (2023). From expert writing to artificial intelligence generated text: Examining the prospects and potential threats of ChatGPT in academic writing. *Biology of Sport*, *40*(2), 615–622. https://doi.org/10.5114/biolsport.2023.125623

Dubey, S. (2024). Redefining cognitive domains in the era of ChatGPT: A comprehensive analysis of artificial intelligence's influence and future implications. *Medical Research Archives*, *12*(6), 1-7. https://doi.org/10.18103/mra.v12i6.5383

Ebrahimian, M., Behnam, B., Ghayebi, N., & Sobhrakhshankhah, E. (2023). ChatGPT in Iranian medical licensing examination: Evaluating the diagnostic accuracy and decision-making capabilities of an AI-based model. *BMJ Health & Care Informatics*, *30*(1), 1-6. http://dx.doi.org/10.1136/bmjhci-2023-100815

Gao, Y., Wang, R., & Hou, F. (2024). How to design translation prompts for ChatGPT: An empirical study. *Proceedings of the 6th ACM International Conference on Multimedia in Asia Workshops*, 1–7. http://dx.doi.org/10.1145/3700410.3702123

Glahn, K. (2024). Using ChatGPT to teach English for academic purposes. *Jeicom*, *5*(2), 63–75. https://doi.org/10.34097/jeicom-5-2-december-23-5

Hirobumi, S. R. (2011). Kata pemeri: Satu takrifan nahuan dalam kerangka ayat bahasa Melayu. *Jurnal Bahasa*, *11*(2), 163-195.

Hu, X., Liu, A., & Dai, Y. (2024). Combining ChatGPT and knowledge graph for explainable machine learning-driven design: A case study. *Journal of Engineering Design*, 1–23. https://doi.org/10.1080/09544828.2024.2355758

Kostka, I., & Toncelli, R. (2023). Exploring applications of ChatGPT to English language teaching: Opportunities, challenges, and recommendations. *Teaching English as a Second or Foreign Language—TESL-EJ*, *27*(3). https://doi.org/10.55593/ej.27107int

Lechien, J. R., Gorton, A., Robertson, J., & Vaira, L. A. (2024). Is ChatGPT-4 accurate in proofreading a manuscript in otolaryngology–head and neck surgery? *Otolaryngology–Head and Neck Surgery*, *170*(6), 1527–1530. https://doi.org/10.1002/ohn.526

Li, J. (2024). Exploring the potential of artificial intelligence to enhance the writing of English academic papers by non-native English-speaking medical students - The educational application of ChatGPT. *BMC Medical Education*, *24*(1), 736. https://doi.org/10.1186/s12909-024-05738-y

Li, R. (2020). Using artificial intelligence in learning English as a foreign language: An examination of IELTS LIULISHUO as an online platform. *Journal of Higher Education Research*, *1*(2), 85-89. https://doi.org/10.32629/jher.v1i2.178

Li, X. (2023). Empowering Chinese language learners from low-income families to improve their Chinese writing with ChatGPT's assistance afterschool. *Languages*, *8*(4), 238. https://doi.org/10.3390/languages8040238

Mills, M., Collier, D., Mahon, J., Ebbinghaus, B., Eggan, F., Gravlee, C., & Maxwell, J. (2008). Comparative research. In *The SAGE encyclopedia of qualitative research methods* (Vol. 0, pp. -). SAGE Publications, Inc., https://doi.org/10.4135/9781412963909.n55

Nhan, L. (2024). Vietnamese university students' perceptions in learning English using ChatGPT. *International Journal of Science and Management Studies (IJSMS), 7*(1), 142–148. https://doi.org/10.51386/25815946/ijsms-v7i1p121

Niloy, A. C., Akter, S., Sultana, N., Sultana, J., & Rahman, S. I. U. (2024). Is ChatGPT a menace for creative writing ability? An experiment. *Journal of Computer Assisted Learning*, *40*(2), 919–930. https://doi.org/10.1111/jcal.12929

Nomoto, H. (2023). Issues surrounding the use of ChatGPT in similar languages: The case of Malay and Indonesian. *Proceedings of the 13th International Joint Conference on Natural Language Processing and the 3rd Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics*, 76-82. Bali: Association for Computational Linguistics. https://doi.org/10.18653/v1/2023.ijcnlp-short.9

Nomoto, H., Moeljadi, D., & Razak, F. A. A. (2024). Masalah teknologi dan isu sosial berkaitan penggunaan ChatGPT dalam Bahasa Melayu: Technological and social issues related to using ChatGPT in Malay. *RENTAS: Jurnal Bahasa, Sastera Dan Budaya*, *3*(1), 1–22. https://doi.org/10.32890/rentas2024.3.1

Omar, A. H. (1993). *Nahu Melayu mutakhir* (4th ed.). Kuala Lumpur: Dewan Bahasa dan Pustaka.

Rojas, A. J. (2024). An investigation into ChatGPT's application for a scientific writing assignment. *Journal of* Chemical *Education*, *101*(5), 1959–1965. https://doi.org/10.1021/acs.jchemed.4c00034

Ruksakulpiwat, S. (2024). Assessing the efficacy of ChatGPT versus expert researchers in identifying relevant studies on mHealth interventions for improving medication adherence in patients with ischemic stroke when conducting systematic reviews: Comparative analysis. *JMIR MHealth and UHealth*, *12*, e51526. https://doi.org/10.2196/51526

Safiah, K. N., Farid, M. O., Hashim, M., & Hamid, M. A. (2010). *Tatabahasa dewan* (3rd ed.). Kuala Lumpur: Dewan Bahasa dan Pustaka.

Safrai, M., & Orwig, K. E. (2024). Utilizing artificial intelligence in academic writing: An in-depth evaluation of a scientific review on fertility preservation written by ChatGPT-4. *Journal of Assisted Reproduction and Genetics*, *41*(7), 1871-1880. https://doi.org/10.1007/s10815-024-03089-7

Seong, Y. P. (2014). Strategic competence and L2 speaking assessment. *Teachers College, Columbia University Working Papers in TESOL & Applied Linguistics*, *14*(1), 13-24

Singh, M. K. M. (2015). International graduate students' academic writing practices in Malaysia: Challenges and solutions. *Journal of International Students*, *5*(1), 12-22. https://doi.org/10.32674/jis.v5i1.439

Singh, M. K. M., & Kaur, M. (2016). An emic perspective on academic writing difficulties among international graduate students in Malaysia. *GEMA Online® Journal of Language Studies*, *16*(3), 83-97. http://dx.doi.org/10.17576/gema-2016-1603-06

Sr, H. İ. S. (2024). *Evaluation of ChatGPT's responses to Google searches about dry eye (preprint)*. https://doi.org/10.2196/preprints.60357

Syafiee, A. S. E. E. 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

Xu, X., Chen, Y., & Miao, J. (2024). Opportunities, challenges, and future directions of large language models, including ChatGPT in medical education: A systematic scoping review. *Journal of Educational Evaluation for Health Professions*, *21*, 6. https://doi.org/10.3352/jeehp.2024.21.6

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