Jurnal LINGUA SUSASTRA
e-ISSN:2746-704X
vol. 7, no. 1, 2026
p. 67-83
DOI: https://doi.org/10.24036/ls.v7i1.535

Translation Quality in Google Translate: A Syntactic Complexity Analysis of Indonesian–English Output

Naila Khoirunnisa[1,*] **Anam Sutopo**[2]
Universitas Muhammadiyah Surakarta[1,2]
*Corresponding Author. E-mail: naikhrns@gmail.com

Submitted: 9 May 2026 Revised: 21 May 2026 Accepted: 30 May 2026

Abstract. Machine translation has become increasingly prevalent in Indonesia, yet the effect of sentence complexity on translation quality remains underexplored in the context of the Indonesian–English language pair. This study investigates the accuracy, readability, and acceptability of Google Translate outputs across four sentence types (simple, compound, complex, and compound-complex) using Nababan's (2012) Translation Quality Assessment (TQA) framework. A descriptive qualitative method was employed, with 100 sentences purposively selected from the novel *Teruslah Bodoh Jangan Pintar* and evaluated by two expert raters. Results show that translation quality declines consistently as syntactic complexity increases: simple sentences achieved the highest scores across all three dimensions, while compound-complex sentences performed the lowest. Importantly, readability consistently exceeded accuracy across all sentence types, indicating that Google Translate tends to produce fluent-sounding output that nonetheless contains meaning distortions and unnatural lexical choices. These findings highlight the limitations of NMT systems in processing hierarchical clause relationships, subordination, and contextual nuance, and reinforce the need for human post-editing in complex translation tasks. The study provides a replicable syntactic classification model for future NMT quality research and offers practical guidance for language educators and translation practitioners in Indonesia.

Keywords: *Google Translate, accuracy, readability, acceptability, sentence complexity*

Abstrak. Terjemahan mesin semakin umum digunakan di Indonesia, namun pengaruh kompleksitas kalimat terhadap kualitas terjemahan masih belum banyak diteliti dalam konteks pasangan bahasa Indonesia–Inggris. Penelitian ini mengkaji akurasi, keterbacaan, dan penerimaan hasil terjemahan Google Translate pada empat jenis kalimat (sederhana, majemuk, kompleks, dan majemuk-kompleks) dengan menggunakan kerangka kerja Penilaian Kualitas Terjemahan (TQA) karya Nababan (2012). Metode kualitatif deskriptif digunakan, dengan 100 kalimat yang dipilih secara purposif dari novel *Teruslah Bodoh Jangan Pintar* dan dievaluasi oleh dua penilai ahli. Hasil menunjukkan bahwa kualitas terjemahan menurun secara konsisten seiring meningkatnya kompleksitas sintaksis: kalimat sederhana memperoleh skor tertinggi di ketiga dimensi, sedangkan kalimat majemuk-kompleks memperoleh skor terendah. Yang penting, keterbacaan secara konsisten melebihi akurasi di semua jenis kalimat, menunjukkan bahwa Google Translate cenderung menghasilkan keluaran yang terdengar lancar namun tetap mengandung distorsi makna dan pilihan leksikal yang tidak alami. Temuan ini menyoroti keterbatasan sistem NMT dalam memproses hubungan klausa hierarkis, subordinasi, dan nuansa kontekstual, serta memperkuat kebutuhan akan penyuntingan pasca-terjemahan oleh manusia dalam tugas terjemahan yang kompleks. Studi ini menyediakan model klasifikasi sintaksis yang dapat direplikasi untuk penelitian kualitas NMT di masa depan dan menawarkan panduan praktis bagi pendidik bahasa dan praktisi terjemahan di Indonesia.

Kata kunci: *Google Translate, akurasi, keterbacaan, keterimaan, kompleksitas kalimat*

INTRODUCTION

In the age of globalization, the advancement of digital technologies has revolutionized translation due to the emergence of NMT such as Google Translate that many students and scholars tend to use since they are convenient and fast to help understand a text written in a different language. The problem, however, emerges because users' unfamiliarity with syntactic structures leads to mistranslations that do not convey the meaning of complex sentences properly.

It seems that the major drawback of machine translation is associated with its inability to cope with syntactic complexity. Recent research suggests that although Google Translate is able to translate simple sentences, it fails to provide semantic and grammatical equivalence of complex ones. This suggestion is confirmed by the results of studies conducted by Suhono et al. (2020), which show that Google Translate demonstrates worse results with compound and complex sentences, thus translating complex sentences with poor accuracy (Azahrah & Aisyah, 2025).

This problem is important since the use of translation software in Indonesia is very widespread. According to the data obtained from the Global Digital Statistics Report (2025), about 50.5% of the Internet users in Indonesia employ translation software once a week (We Are Social, 2025), which is much higher than the worldwide average rate of 29%. However, this practice does not always result in enhanced language competency because people usually trust the output provided by online translation software without considering the structure and meanings behind it (Untara & Setiawan, 2020). Indeed, according to the findings made by A. Ismail, (2016), there are many grammatical, syntactic, and semantic errors committed in translating long and complicated sentences.

Despite the numerous works dedicated to the issues related to lexical and grammatical mistakes, while previous studies have examined translation errors or general quality in machine-translated texts, few have systematically compared translation quality across all four sentence types (simple, compound, complex, and compound-complex) using a multi-dimensional TQA framework applied to literary data in the Indonesian–English language pair. Indeed, there are only four major types of sentences: simple, compound, complex, and compound-complex.

This study focuses on Google Translate due to its popularity and accessibility among Indonesian students. A simple example shows that simple sentences can be translated well, but complex sentences often fail to maintain the relationship between clauses, resulting in errors in meaning. This indicates the limitations of machine translation in handling complex syntactic structures.

Theoretically, this study is grounded in Chomsky (1957) syntax theory, Poibeau, (2017) machine translation theory, and Catford (1965) translation shift theory, which explains how meaning and structure change during the translation process. Additionally, translation quality is measured using the Translation Quality Assessment (TQA) framework, which encompasses three main aspects: accuracy, readability, and acceptability (House, 1977; Larson, 1998; Nababan, 2012).

The method used is qualitative descriptive research employing content analysis. The data consists of 100 sentences from the novel *Teruslah Bodoh Jangan Pintar*, classified into four sentence types, then translated using Google Translate and analyzed based on the three quality aspects with the assistance of expert judgment.

Based on this background, the objective of this study is to analyze the quality of Google Translate’s translations of the four sentence types based on the aspects of accuracy, readability, and acceptability. This study is expected to provide theoretical

contributions to technology-based translation studies and offer practical implications for students, instructors, and researchers in using and evaluating machine translation.

Although previous studies have examined Google Translate's performance in translating Indonesian news texts (Ismail, 2016) and academic writing (Suhono et al., 2020; Winiharti et al., 2021), significant gaps remain in the literature. Most existing research identifies translation errors at the lexical or grammatical level without classifying source sentences according to syntactic type (simple, compound, complex, or compound-complex) as a primary analytical variable. Furthermore, few studies have simultaneously assessed accuracy, readability, and acceptability within a unified TQA framework applied to syntactically categorized data, leaving the relationship between sentence structure and multi-dimensional translation quality empirically uncharted. Literary texts, which exhibit a high density of complex syntactic structures, have also been largely overlooked as a data source in Indonesian machine translation studies. This study addresses these three gaps by combining syntactic classification with Nababan's (2012) multi-dimensional TQA framework, applied to a contemporary Indonesian literary text, to produce a more systematic and replicable evaluation of Google Translate's performance.

More recent studies have shown that although Google Translate has significantly improved with the implementation of Neural Machine Translation (NMT), problems still frequently occur in translating complex sentence structures. Several researchers found that simple sentences tend to produce more accurate translations, while compound, complex, and compound-complex sentences often lead to grammatical shifts, mistranslations, ambiguity, and loss of meaning. These findings indicate that sentence complexity strongly influences machine translation performance because longer and syntactically complicated structures require deeper contextual and semantic processing. Recent research has also shifted from merely identifying translation errors toward evaluating translation quality comprehensively, especially through the indicators of accuracy, acceptability, and readability.

Therefore, the present study differs from previous studies in both focus and analytical framework. This study does not only identify translation errors, but primarily analyzes translation quality in Google Translate outputs based on sentence structure complexity. The study classifies sentences into simple, compound, complex, and compound-complex types and evaluates their translation quality using the dimensions of accuracy, acceptability, and readability. By combining syntactic classification with translation quality assessment, the study offers a more systematic and comprehensive analysis of Google Translate performance in translating Indonesian literary texts into English.

METHODS

This study employed a descriptive qualitative design supported by frequency-based quantitative tabulation to systematically compare Google Translate's translation quality across four syntactic sentence types. The qualitative component enabled in-depth analysis of specific translation cases, while the frequency counts provided a basis for cross-category comparison of accuracy, readability, and acceptability scores.

For data collection, One hundred sentences were selected from the novel *Teruslah Bodoh Jangan Pintar* through purposive sampling, with 25 sentences each representing simple, compound, complex, and compound-complex structures based on syntactic classification. Data collection was carried out as follows: carefully reading the novel to

identify sentences with different syntactic structures; selecting sentences and grouping them into four categories; translating the sentences into English using Google Translate in a single session to ensure consistency; and organizing all data based on sentence type and translation results.Two experts in translation studies independently evaluated each translation using a three-point scale for accuracy, readability, and acceptability. Inter-rater agreement was checked, and disagreements were resolved through discussion. Their feedback was used to refine the interpretation and reduce subjectivity.

Syntactic classification was conducted by the researcher based on Quirk et al.'s (1985) framework for English sentence typology, adapted to Indonesian syntactic structures. Sentences were classified as follows: (1) simple sentences contain one independent clause with a single subject-predicate structure; (2) compound sentences contain two or more independent clauses joined by coordinating conjunctions; (3) complex sentences contain one independent clause and at least one subordinate clause connected by a subordinating conjunction; and (4) compound-complex sentences contain at least two independent clauses and one or more subordinate clauses. All classifications were verified by a second expert in Indonesian linguistics to ensure reliability of the syntactic categorization prior to translation.

Table 1. Nababan's Three-Dimensional Translation Quality Assessment Framework

Score	Accuracy	Readability	Acceptability
3	The message of the source text is fully transferred into the target text; no distortion of meaning	The translation is easily understood by the target reader without effort	The translation uses natural target-language norms, vocabulary, and expressions
2	Most of the source text message is transferred; minor distortion occurs	The translation is generally understandable but requires some effort in certain parts	Most of the translation uses natural expressions but some parts feel unnatural
1	The message is not transferred accurately; serious distortion of meaning is present	The translation is difficult to understand	The translation does not follow target-language norms; it feels like a translation

(*Adapted from* Nababan, 2012)

Scores were tabulated by sentence type to identify frequency distributions across quality categories. Representative sentence pairs (source and target) were selected for qualitative analysis to illustrate translation patterns, with reference to Catford's (1965) translation shift theory, Nida's (1964) dynamic equivalence, and Larson's (1998) meaning-based translation principles.

FINDINGS

Accuracy of Google Translate Output by Sentence Types

A total of 100 sentences were analyzed across four syntactic categories, with 25 sentences representing each type. Among the 25 simple sentences, 21 (84%) were rated as accurate and 4 (16%) as less accurate; none were classified as inaccurate. Among the 25 compound sentences, 17 (68%) were rated as accurate, 6 (24%) as less accurate, and 2 (8%) as inaccurate. Of the 25 complex sentences, 12 (48%) were rated as accurate, 11 (44%) as less accurate, and 2 (8%) as inaccurate. In the compound-complex category, 8 (32%) were rated as accurate, 11 (44%) as less accurate, and 6 (24%) as inaccurate.

Table 2. Translation Quality by Sentence Type

Sentence Type	Total	Accurate	Less Accurate	Inaccurate	Readable	Mod. Readable	Low Readable	Acceptable	Less Acceptable	Unacceptable
Simple	25	21 (84%)	4 (16%)	0 (0%)	23 (92%)	2 (8%)	0 (0%)	19 (76%)	6 (24%)	0 (0%)
Compound	25	17 (68%)	6 (24%)	2 (8%)	18 (72%)	7 (28%)	0 (0%)	15 (60%)	8 (32%)	2 (8%)
Complex	25	12 (48%)	11 (44%)	2 (8%)	16 (64%)	9 (36%)	0 (0%)	16 (64%)	7 (28%)	2 (8%)
Compound-Complex	25	8 (32%)	11 (44%)	6 (24%)	11 (44%)	12 (48%)	2 (8%)	13 (52%)	11 (44%)	2 (8%)
Total	100	58 (58%)	32 (32%)	10 (10%)	68 (68%)	30 (30%)	2 (2%)	63 (63%)	32 (32%)	6 (6%)

As shown in Table 2, translation quality consistently declines as sentence complexity increases across all three assessed dimensions. For accuracy, the proportion of fully accurate translations drops from 84% in simple sentences to 32% in compound-complex sentences, a reduction of 52 percentage points. A similar downward trend is observed in readability, declining from 92% to 44%, and in acceptability, from 76% to 52%. Notably, readability scores are consistently higher than accuracy scores across all four sentence types, indicating a systematic fluency-accuracy gap in Google Translate's output that becomes more pronounced as syntactic complexity increases. These patterns are discussed in detail in the following sub-sections.

a Simple Sentence

Example 1: Accurate translation
SL: *Semua kursi telah diisi.*
TL: All seats were taken.
Reason: The meaning is fully preserved, the passive structure is natural, and the sentence is acceptable in English.

Example 2: Less accurate translation
SL: *Mereka punya banyak tambang disini, tapi orang suruhannya yang mengurus.*
TL: They have a lot of mines here, but their people take care of them.
Reason: The phrase *orang suruhan* is too generally translated as *their people*, so the meaning becomes less precise.

b Compound Sentence

Example 1: Accurate translation
SL: *Dia mengikuti berita tapi apa hubungannya dengan dia?*
TL: He follows the news but what does it have to do with him?
Reason: Both clauses are accurately translated, and the contrastive relation is maintained well. The second clause sounds natural in English and preserves the original meaning without distortion.

Example 2: Less accurate translation
SL: *Mereka punya banyak tambang di sini, tapi orang suruhannya yang mengurus.*
TL: They have a lot of mines here, but their people take care of them.
Reason: The phrase orang *suruhan* is translated too generally as their people, causing a slight loss of meaning and reducing lexical precision.

Example 3: Inaccurate translation
SL: *Wajahnya pias, tapi dia tetap tidak mengerti.*
TL: His face was bright, but he still didn't understand.
Reason: *Pias* should be translated as *pale*, not *bright*; therefore the meaning is distorted.

c Complex Sentence
Example 1: Accurate translation
SL: *Saat wartawan sibuk dengan pengacara top itu, Ibu Sri membawa Ahmad keluar.*
TL: While reporters were busy with the top lawyer, Mrs. Sri took Ahmad outside.
Reason: The temporal relation is correctly conveyed through while, and both clauses are translated naturally and accurately without losing meaning.

Example 2: Less accurate translation
SL: *Jika tidak di ruangan terhormat ini, aku mungkin telah menangis sejak tadi.*
TL: If it weren't for this honorable room, I might have been crying a long time ago.
Reason: The phrase honorable room is not idiomatic in English. A more natural translation would be this honorable court or courtroom. The meaning is understandable but less precise.

Example 3: Inaccurate translation
SL: *Sejak tahun 1960-an tidak ada lagi tambang emas skala besar yang menggunakan merkuri karena itu merugikan perusahaan, pemisahan emas tidak efisien.*
TL: Since the 1960s, large-scale gold mines have stopped using mercury because it's detrimental to the company and inefficient gold separation.
Reason: The phrase inefficient gold separation is grammatically awkward and unclear, causing distortion of the intended cause-effect meaning.

d Compound Complex Sentence
Example 1: Accurate translation
SL: *Hotma Cornelius bahkan bisa menebak jika lawan akan membawa bawa masa kecil kliennya, dan dia telah menyiapkan kontra serangan.*
TL: Hotma Cornelius could even guess that the opponent would bring up his client's childhood, and he had prepared a counterattack.
Reason: The prediction and preparation expressed in both clauses are accurately translated. The phrase bring up appropriately represents *membawa-bawa,* and the meaning is fully preserved.

Example 2: Less accurate translation
SL: *Jika ahli ahli memutuskan tidak kita batalkan saja, tapi jika ahli ahli bilang teruskan, saya harap rakyat juga bisa memahami.*

TL: If the experts decide not to, we should cancel it, but if the experts say to continue, I hope the people can understand.
Reason: The phrase decide not to be ambiguous because it does not clearly explain what is being rejected. The overall message is understandable, but the meaning becomes less specific.

Example 3: Inaccurate translation
SL: *Maka gunakan pendekatan baru, aku tidak mau tahu proyek itu harus segera jalan, percuma pemerintah mengundang banyak investor jika setelah mereka berdatangan, memberikan komitmen investasi miliaran dolar ternyata di lapangan ada masalah.*
TL: So, use a new approach. I don't want the project to get underway immediately. It's pointless for the government to invite many investors if, after they arrive and commit billions of dollars in investment, there are problems on the ground.
Reason: The phrase *aku tidak mau tahu* expresses indifference, but it is translated as refusal (I don't want). This shifts the original meaning and results in inaccurate translation.

Overall, the accuracy data reveal a clear inverse relationship between syntactic complexity and translation precision, with fully accurate translations declining from 84% in simple sentences to 32% in compound-complex sentences (see Table 1). The primary sources of inaccuracy shift in nature across categories: in simple sentences, inaccuracies are predominantly lexical (e.g., imprecise equivalents for culturally specific terms), while in compound and complex sentences, they increasingly involve clause-relationship distortions and contextual misreadings. In compound-complex sentences, semantic shifts at the pragmatic level, such as the misrepresentation of *aku tidak mau tahu* as refusal rather than indifference, represent the most serious category of translation failure. The theoretical implications of these patterns are examined in the Discussion.

Readability of Google Translate Output by Sentence Types
There are a total of 100 sentences. Of these 25 are simple sentences, in which 23 are classified as readable and 2 as moderately readable, with none categorized as low readable. Among the 25 compound sentences, 18 are readable and 7 are moderately readable, with none categorized as low readable. Of the 25 complex sentences, 16 are readable and 9 are moderately readable, with none categorized as low readable. Meanwhile, for compound-complex sentences, 11 are readable, 12 are moderately readable, and 2 are low readable.

a Simple Sentence
Example 1: High readable translation
TL: All seats were taken.
Reason: The sentence is highly readable because it uses a simple passive structure and familiar vocabulary. The meaning can be understood immediately without ambiguity.

Example 2: Moderately readable translation
TL: One or two children elbowed each other.

Reason: The sentence structure is simple and generally clear, but the word elbowed may be less familiar to some readers, slightly reducing readability.

b Compound Sentence

Example 1: High readable translation
TL: They have a lot of mines here, but their people take care of them.
Reason: The sentence is easy to read because the conjunction but clearly connects the two clauses, and the vocabulary remains simple and familiar.

Example 2: Moderately readable translation
TL: He really didn't tell anyone, and other residents also knew exactly the risk.
Reason: The sentence is understandable, but the phrase knew exactly the risk sounds slightly unnatural. The combination of two clauses also requires more attention from readers.

c Complex Sentence

Example 1: High readable translation
TL: To avoid attracting attention, they gathered at night.
Reason: The sentence is concise and clear. The purpose relation is expressed naturally, and the vocabulary is easy to understand.

Example 2: Moderately readable translation
TL: Because the business permit required a location far from residential areas, according to her husband, their livestock pen was very far away.
Reason: The sentence is grammatically correct, but its long structure and additional information make it more difficult to process. Repetition of the word far also reduces clarity.

d Compound Complex Sentence

Example 1: Readable translation
TL: Hotma Cornelius could even guess that the opponent would bring up his client's childhood, and he had prepared a counterattack.
Reason: Despite containing multiple clauses, the connectors that and and clearly show the relationship between ideas, making the sentence understandable in one reading.

Example 2: Moderately readable translation
TL: If the experts decide not to, we should cancel it, but if the experts say to continue, I hope the people can understand.
Reason: The sentence contains several clauses connected by if and but, which increases complexity and requires more careful reading.

Example 3: Low readable translation
TL: The state has rights to all land, and the state has the right to grant it to designated individuals, organizations, and/or companies, so that it can be utilized and cultivated optimally for the welfare of the people.
Reason: The sentence is long and dense, containing multiple ideas, repetition, and formal vocabulary such as designated, utilized, and welfare, which makes it difficult to read efficiently.

The readability data confirm that Google Translate generally produces fluent surface-level output across all sentence types, with high readability scores even in categories where accuracy is significantly compromised. However, readability declines progressively from 92% in simple sentences to 44% in compound-complex sentences, with the two low-readability cases appearing exclusively in the compound-complex category. The co-occurrence of high readability and reduced accuracy, particularly evident in compound and complex sentences, constitutes a fluency-accuracy gap that has direct implications for uncritical reliance on machine translation output. This pattern is theorized further in the Discussion.

The Acceptability of Google Translate Output by Sentence Types

There are a total of 100 sentences. Of these, 25 are simple sentences, in which 19 are acceptable and 6 are less acceptable, with none categorized as unacceptable. Of the 25 compound sentences, 15 are acceptable, 8 are less acceptable, and 2 are unacceptable. Of the 25 complex sentences, 16 are acceptable, 7 are less acceptable, and 2 are unacceptable. Meanwhile, of the 25 compound-complex sentences, 13 are acceptable, 11 are less acceptable, and 2 are unacceptable.

a Simple Sentence

Example 1: Acceptable translation

TL: All seats were taken.

Reason: The sentence uses a natural passive structure and appropriate vocabulary. The translation sounds fluent and acceptable in English.

Example 2: Less acceptable translation

TL: One or two children elbowed each other.

Reason: The sentence is grammatically correct, but the combination of one or two and each other sounds slightly unnatural, reducing acceptability.

b Compound Sentence

Example 1: Acceptable translation

TL: There was no other choice, and that morning he still played badly.

Reason: The sentence uses a correct compound structure with and, and both clauses are naturally connected and easy to understand.

Example 2: Less acceptable translation

TL: They have a lot of mines here, but their people take care of them.

Reason: The phrase their people sounds too general and less natural in English, reducing the overall acceptability of the sentence.

Example 3: Unacceptable translation

TL: His face was bright, but he still didn't understand.

Reason: The word bright is an incorrect translation of pias, which should be pale. This creates unnatural meaning and makes the translation unacceptable.

c Complex Sentence

Example 1: Acceptable translation

TL: To avoid attracting attention, they gathered at night.

Reason: The sentence is grammatically correct and natural. The purpose relation is expressed clearly through the infinitive construction.

Example 2: Less acceptable translation
TL: Everyone also knew that they were engineering high-level transactions in the smelter export trade.
Reason: The sentence is understandable, but phrases such as engineering high-level transactions and smelter export trade sound awkward and overly formal, reducing naturalness.

Example 3: Unacceptable translation
TL: If it weren't for this honorable room, I might have been crying a long time ago.
Reason: The phrase honorable room is not idiomatic in English. More natural expressions would be courtroom or honorable court, so the translation becomes less acceptable

d Compound Complex Sentence
Example 1: Acceptable translation
TL: Hotma Cornelius could even guess that the opponent would bring up his client's childhood, and he had prepared a counterattack.
Reason: The sentence is well-structured and natural. The connectors that and and clearly show the relationship between clauses.

Example 2: Less acceptable translation
TL: If the experts decide not to, we should cancel it, but if the experts say to continue, I hope the people can understand.
Reason: The sentence is grammatically correct, but the phrase say to continue sounds slightly unnatural. A more natural form would be say we should continue.

Example 3: Unacceptable translation
TL: The state has rights to all land, and the state has the right to grant it to designated individuals, organizations, and/or companies, so that it can be utilized and cultivated optimally for the welfare of the people.
Reason: The sentence is overly long and uses formal, unnatural expressions such as designated individuals and utilized and cultivated. Repetition of the word state also reduces naturalness and acceptability.

Acceptability scores follow a pattern broadly parallel to accuracy, declining from 76% in simple sentences to 52% in compound-complex sentences, though with notable variation: complex sentences (64%) slightly outperform compound sentences (60%) in acceptability, suggesting that well-formed subordinate structures can sometimes produce more natural-sounding English than loosely coordinated compound clauses. Unacceptable translations, which is absent in simple sentences, appear consistently across compound, complex, and compound-complex categories (8% each), driven primarily by lexical errors that generate semantically opposite or non-idiomatic target expressions. A fuller linguistic explanation of these acceptability failures is provided in the Discussion.

Taken together, the findings across all three quality dimensions reveal two overarching patterns in Google Translate's Indonesian–English translation performance. First, a consistent complexity-quality gradient is evident: translation quality—measured by accuracy, readability, and acceptability—declines systematically as sentence

structure progresses from simple to compound-complex, with accuracy showing the steepest decline (84% to 32%) and readability the most gradual (92% to 44%). Second, a fluency-accuracy gap is present across all sentence types: readability scores consistently exceed accuracy scores, indicating that Google Translate prioritizes surface fluency over semantic precision. This gap widens as sentence complexity increases—most pronouncedly in compound-complex sentences, where translations frequently appear grammatically well-formed yet contain pragmatic distortions, lexical inaccuracies, and non-idiomatic expressions that undermine communicative equivalence. These two patterns form the basis of the theoretical discussion that follows.

DISCUSSION

The findings of this study reveal two systematic patterns in Google Translate's Indonesian–English translation performance that warrant theoretical explanation. The first is a complexity-quality gradient: accuracy, readability, and acceptability all decline as sentence structure progresses from simple to compound-complex, with accuracy registering the steepest decline from 84% to 32%. The second is a fluency-accuracy gap: readability scores consistently exceed accuracy scores across all four sentence types, indicating that the system prioritizes surface fluency over semantic precision. These patterns are not incidental but are directly attributable to the architectural and linguistic constraints of Neural Machine Translation (NMT) systems when confronted with hierarchically complex syntactic structures, a relationship that the following sub-sections examine through the lens of translation theory and computational linguistics.

This pattern can be explained linguistically. Simple sentences consist of a single independent clause, making them structurally straightforward and easier to process (Quirk et al., 1985). Because the grammatical relationship is direct and the number of semantic elements is limited, the machine can map source-language patterns into target-language structures more accurately. Simple sentences also contain fewer ambiguities in reference, tense, and clause dependency, reducing the possibility of mistranslation. As a result, meaning transfer occurs with minimal restructuring, supporting Larson (1998) view of translation. In line with Nida (1964) concept of dynamic equivalence, simple sentences tend to preserve the same communicative effect in the target language, which explains their high accuracy (around 84%).

In compound sentences, accuracy begins to decline (68% accurate). Although the overall meaning is usually preserved, problems arise in lexical choice and maintaining clause relationships. Linguistically, compound sentences require the system to process coordination between two or more independent clauses while maintaining logical relations such as contrast, addition, or consequence. Machine translation systems often prioritize literal lexical equivalence rather than contextual nuance. This reflects Catford (1965) concept of translation shifts and Baker (2011) idea of lexical asymmetry. For instance, translating “orang suruhan” as “their people” shows a loss of specific meaning because the Indonesian expression contains social and contextual nuance that does not directly correspond to a single English lexical item. As a result, semantic precision decreases even though the sentence remains understandable.

The decline is more noticeable in complex sentences, where only 48% are accurate. This is due to difficulties in handling subordinate relationships, as explained by (Chomsky, 1957). Complex sentences require the system to recognize hierarchical syntactic structures, including cause-effect, temporal, conditional, and purposive relations. When the system fails to identify which clause functions as the main idea and

which serves as supporting information, the logical meaning becomes distorted. This supports Nababan (2012) view that accuracy depends not only on word equivalence but also on the correct relationship between sentence elements. For example, awkward constructions such as “inefficient gold separation” occur because the machine fails to restructure the causal information naturally in English. From a functional perspective, House (1977) argues that failure to maintain these relationships disrupts communicative function.

Compound-complex sentences present the greatest challenge, with only 44% accurate translations. These sentences combine coordination and subordination simultaneously, requiring the machine to process multiple grammatical dependencies at once. Linguistically, this increases syntactic load because the system must track clause hierarchy, conjunctions, pronoun references, and contextual meaning across longer sentence spans.

Vinay and Darbelnet's (1958) taxonomy of translation techniques is also instructive here. Their distinction between direct translation (calque, literal translation) and oblique translation (transposition, modulation, equivalence, adaptation) highlights the specific failure mode of NMT systems: they default systematically to direct translation strategies, particularly calque and literal translation, even in contexts where oblique techniques, especially modulation and equivalence, are linguistically necessary to preserve naturalness in the target language. This default explains why compound-complex sentences in this study generate the highest proportion of unacceptable translations: their multi-clause structures demand multiple simultaneous oblique adjustments that NMT systems cannot reliably execute.

From a technological perspective, Google Translate relies on probabilistic patterns rather than full semantic understanding (Poibeau, 2019). Although it uses attention mechanisms Bahdanau et al., (2014), these systems mainly predict likely word sequences based on large datasets. They do not fully understand pragmatic intention, implied meaning, or cultural nuance. As sentence length increases, the attention mechanism must distribute focus across more linguistic units, making it harder to maintain coherence and accurate clause relationships. This limitation aligns with Halverson (2010) view that translation involves cognitive processes, such as inference and contextual interpretation, which machines cannot fully replicate.

These findings are consistent with previous studies Azahrah & Aisyah (2025); Suhono et al., (2020), which show that translation quality decreases as sentence complexity increases. Overall, Google Translate performs well for simple sentences but struggles with more complex structures, indicating that syntactic complexity is a key factor influencing accuracy.

The findings show that the readability of Google Translate outputs is generally more stable than their accuracy. However, readability still decreases as sentence complexity increases. Simple sentences have the highest readability (23 out of 25 sentences), followed by compound sentences, then complex sentences, and finally compound-complex sentences, which show the lowest readability, with some categorized as low (11 out of 25 highly readable). This indicates that while Google Translate can produce fluent text, readability is still influenced by syntactic complexity.

According to Nababan (2012), readability refers to how easily a translated text can be understood without significant effort. In this study, simple sentences are highly readable due to their straightforward structure. Linguistically, simple sentences reduce cognitive processing because readers only need to interpret one independent clause with limited grammatical embedding. This supports Larson (1998) view that readability

depends on effective restructuring in the target language. Because simple sentences require minimal restructuring, they tend to sound natural and clear. For example, sentences like "All seats were taken" show direct and unambiguous meaning. However, some cases are only moderately readable due to less familiar vocabulary, such as "elbowed," which requires additional processing despite the simple sentence structure.

Sweller (1988) explains that more complex structures increase cognitive load. This is evident in compound and complex sentences, where readers must process multiple clauses at once. Linguistically, coordination and subordination require readers to interpret logical relationships between clauses while simultaneously maintaining information in working memory. In compound sentences, readability begins to vary due to the need to understand relationships between clauses. Although many sentences remain readable, awkward phrasing and literal lexical choices reduce processing fluency.

In complex sentences, readability decreases further due to subordinate relationships and longer sentence structures. As Chomsky (1957) explains, such sentences require deeper syntactic processing. This is supported by Just and Carpenter (1992), who argue that readers must integrate information gradually in working memory. When sentences contain additional embedded information or repeated structures, comprehension becomes more difficult. This is reflected in examples like "know exactly the risk," which is grammatically understandable but pragmatically unnatural in English, forcing readers to spend additional effort interpreting intended meaning.

The decline is most significant in compound-complex sentences. These structures combine coordination and subordination, resulting in long and dense sentences with high information load. As Quirk et al., (1985) note, more clauses lead to more complex relationships. Long sentences containing repetition, formal diction, and multiple embedded phrases increase cognitive burden because readers must continuously connect ideas across clauses. In this study, only 44% of such sentences were highly readable. Sentences such as "the state has rights to all land…" demonstrate how repetition, nominalization, and formal vocabulary reduce readability and create processing difficulty.

From a machine translation perspective, this limitation is related to how systems process language. Poibeau (2019) explains that neural machine translation relies on probabilistic patterns, which may struggle with long and complex structures. Although attention mechanisms help Bahdanau et al., (2014), they are still limited in handling multiple clause relationships, often resulting in less coherent sentence flow.

One of the most significant findings of this study is the systematic divergence between readability and accuracy across all four sentence types, a pattern here termed the fluency-accuracy gap. Readability consistently outperformed accuracy by margins ranging from 8 percentage points in simple sentences (92% vs. 84%) to 12 percentage points in compound-complex sentences (44% vs. 32%), indicating that Google Translate regularly produces output that appears grammatically fluent yet contains semantic distortions, lexical inaccuracies, and pragmatic misrepresentations.

This pattern is not coincidental but is structurally embedded in how NMT systems are designed and trained. As Poibeau (2019) explains, NMT systems such as Google Translate are optimized through probabilistic sequence prediction, they are trained to generate the most statistically likely word sequence in the target language given the source input, rather than to reconstruct the source meaning with semantic precision. This training objective inherently prioritizes surface fluency over meaning fidelity, particularly when source structures fall outside the high-frequency patterns

dominating the training corpus. The result is output that reads naturally at the syntactic level while failing at the semantic and pragmatic levels, precisely the pattern observed in this study.

This finding aligns with McDonald's (2020) observation that machine translation systems prioritize fluency over precise meaning, and extends it by demonstrating that the gap is gradient rather than uniform, widening systematically as sentence complexity increases. It also supports Nord's (1997) functionalist argument that translation effectiveness must be measured not by surface form but by communicative function: a translation that reads fluently but misrepresents the source message fails its communicative purpose regardless of its grammatical correctness.

These findings are consistent with previous studies Suhono et al., (2020); Winiharti et al., (2021), which show that Google Translate produces generally readable text but struggles with complex structures. Overall, while Google Translate performs well in generating readable output, its readability decreases with increased sentence complexity due to clause density, embedded structures, repetition, and unnatural lexical combinations.

The practical implication is significant and directly relevant to Indonesian educational contexts, where students frequently accept Google Translate output at face value if it "sounds right" in English (Untara & Setiawan, 2020). The fluency-accuracy gap identified in this study provides empirical grounds for advocating critical post-editing literacy, the ability to evaluate machine translation output not merely for readability but for semantic and pragmatic fidelity, as a core competency in translation pedagogy and academic writing instruction.

The findings show that acceptability follows a pattern similar to accuracy: it decreases as sentence complexity increases. Simple sentences have the highest acceptability (19 out of 25 data), followed by compound and complex sentences with relatively similar results (15 and 16 acceptable data). Compound-complex sentences show the greatest variation, with some categorized as less acceptable or unacceptable. This indicates that more complex structures tend to produce less natural translations.

According to Nababan (2012), acceptability refers to how well a translation conforms to the norms of the target language, including grammar, vocabulary, and usage. In this study, simple sentences are highly acceptable because their structure is straightforward and aligns easily with English grammatical patterns. Linguistically, simple sentences allow direct transfer without extensive restructuring, reducing the risk of unnatural syntax. This supports Larson (1998) view that translations should follow natural target-language forms. As a result, simple sentences require minimal restructuring and tend to sound natural. Around 76% of the data fall into this category, as seen in examples like “All seats were taken,” which uses a natural and commonly accepted structure.

From the perspective of dynamic equivalence Nida (1964) acceptable translations produce a natural effect for readers. Simple sentences fulfill this principle because they are easy to read and do not feel like translations. Similarly, House (1977) emphasizes that communicative function must be maintained, which is generally achieved in simple sentences due to their clear and familiar structure.

In compound sentences, acceptability begins to vary. While many are still acceptable (60%), some sound less natural due to structural influence from the source language. Linguistically, Indonesian and English differ in collocation patterns and clause organization. The translation errors identified in this study can be further

theorized through Catford's (1965) concept of translation shifts and Baker's (2011) framework of collocational meaning.

Catford distinguishes between level shifts (where meaning is transferred at a different linguistic level) and category shifts (where grammatical categories change between source and target). Several compound and complex sentence errors in this dataset exemplify category shifts: for instance, the Indonesian causal structure karena itu merugikan perusahaan, pemisahan emas tidak efisien is rendered by Google Translate as "because it's detrimental to the company and inefficient gold separation", where the subordinate causal clause is collapsed into a coordinate structure, eliminating the cause-effect relationship and producing a grammatically awkward noun phrase (inefficient gold separation) in place of a full predicate.

Baker's (2011) concept of collocational meaning, the tendency of words to co-occur in predictable combinations within a given language, explains a separate but related category of acceptability failures in this dataset. Indonesian collocational norms frequently do not map onto English equivalents on a word-for-word basis, yet NMT systems trained on parallel corpora tend to reproduce source-language collocations literally. This produces expressions such as "knew exactly the risk" (for *tahu persis risikonya*) and "their people take care of them" (for *orang suruhannya yang mengurus*) both of which are grammatically parseable in English but violate English collocational conventions, reducing acceptability even where meaning is partially preserved.

In complex sentences, acceptability depends more on how well the system handles subordinate relationships. According to Nord (1997), translation should consider communicative purpose, which may fail when structures are too complex or unnatural. In this study, 64% of complex sentences are acceptable, but others show lexical awkwardness, such as “honorable room,” which is lexically understandable but not idiomatic in English. From a syntactic perspective, Chomsky (1957) explains that complex sentences require accurate representation of hierarchical relationships. If these are not properly conveyed, the result sounds unnatural. Quirk et al. (1985) also emphasize that clarity in English depends on proper clause arrangement and conjunction use.

The lowest acceptability is found in compound-complex sentences, where structures are long and dense. Only about 50% are acceptable, while the rest are less acceptable or unacceptable. Linguistically, long sentences containing multiple dependencies, repetition, and formal vocabulary reduce naturalness because English generally favors concise and efficient structures. This is due to the system’s limitations in handling multiple clause relationships simultaneously (Poibeau, 2019). Although attention mechanisms help (Bahdanau et al., 2014), they are not sufficient to maintain coherence in long sentences. As a result, translations may appear fragmented and less natural.

These findings are consistent with previous studies Winiharti et al., (2021), which show that Google Translate often produces understandable but less acceptable translations due to insufficient adaptation to target-language norms. Overall, acceptability is strongly influenced by syntactic complexity because linguistic features such as collocation, clause hierarchy, idiomatic usage, and information density determine how natural a translation sounds in the target language.

While Google Translate performs well in producing acceptable translations for simple and moderately complex sentences, it struggles with more complex structures.

Therefore, human revision is still necessary to ensure that translations are not only correct but also natural and appropriate in the target language.

CONCLUSION

The study demonstrates that syntactic complexity is a decisive factor in Google Translate's Indonesian–English translation quality. Across accuracy, readability, and acceptability, translation performance declines systematically as sentence structure shifts from simple to compound-complex, indicating that machine translation is most reliable when processing short, clause-limited constructions and least reliable when handling multi-clause sentences with dense semantic and pragmatic relations. More importantly, the findings show that surface fluency does not guarantee semantic fidelity: Google Translate frequently produces readable output that still contains lexical distortion, clause-relationship errors, or unnatural target-language expressions. This pattern confirms that syntactic complexity is not merely a formatting issue but a core determinant of machine translation reliability.

This study contributes a more systematic framework for evaluating machine translation quality by linking syntactic sentence type to Nababan's three-dimensional assessment model. Unlike studies that focus only on error identification or general translation quality, this research shows that sentence structure itself can serve as a predictive variable for translation performance. The integration of simple, compound, complex, and compound-complex sentence categories with accuracy, readability, and acceptability assessment offers a replicable model for future research in Indonesian–English machine translation evaluation. In this sense, the study does not merely confirm Google Translate's limitations; it clarifies where those limitations emerge and how they intensify across levels of syntactic complexity.

Several limitations should be acknowledged when interpreting these findings. First, the data were drawn from a single Indonesian novel, so the results may not fully represent other genres such as news, academic, legal, or technical texts. Second, the study evaluated only Google Translate, meaning that its performance cannot be assumed to reflect other machine translation systems. Third, the corpus was limited to 100 sentences, which provides useful comparative insight but does not support broad generalization across all Indonesian–English translation contexts. Finally, translation quality was assessed through expert judgment, which is appropriate for TQA studies but should ideally be complemented by quantitative reliability measures and additional evaluation metrics in future work.

Future research should extend this syntactic-complexity framework to other machine translation systems such as DeepL, ChatGPT, or other NMT platforms in order to determine whether the fluency-accuracy gap observed in Google Translate is system-specific or more general. Subsequent studies could also compare multiple text genres, including academic, journalistic, and legal texts, to examine how genre and sentence structure interact in shaping translation quality. In addition, future work would benefit from combining expert-based TQA with automated evaluation metrics and larger corpora so that syntactic complexity, lexical choice, and pragmatic fidelity can be analyzed more comprehensively. Such studies would strengthen the empirical basis for machine translation pedagogy and post-editing practice.

REFERENCES

A. Ismail, R. H. (2016). Errors Made in Google Translate in the Indonesian To English Translations of News Item Texts. *Journal of English Language Teaching*, *5*(2), 1–

6.
Bahdanau, D., Cho, K., & Bengio, Y. (2014). Neural machine translation by jointly learning to align and translate. *ArXiv Preprint*, *arXiv:1409*. https://doi.org/10.48550/arXiv.1409.0473
Baker, M. (2011). *In other words: A coursebook on translation* (2nd ed.). Routledge.
Catford, J. C. (1965). Language and Language Learning a Linguistic Theory of Translation. In *Oxford University Press*.
Chomsky, Noam. (1957). *Syntact Structures* (Second). Mouton & Co.
Halverson, S. L. (2010). Cognitive translation studies: Developments in theory and method. In G. M. Shreve & E. Angelone (Eds.), *Translation and cognition* (pp. 349–369). John Benjamins Publishing Company. https://doi.org/10.1075/ata.xv.16hal
House, J. (1977). *A Model for Translation Quality Assessment* (2nd ed.). Gunter Narr Verlag.
K. Azahrah, D. Aisyah, O. T. (2025). An Analysis of Google Translate Accuracy in Translating English To Indonesian. *Jurnal Pemikiran Pendidikan Dan Keguruan*, *1*(1), 84–88.
Larson, M. L. (1998). *Meaning-based Translation: A Guide to Cross-language Equivalence* (2nd ed.). University Press of America.
McDonald, S. V. (2020). Accuracy, readability, and acceptability in translation. *Applied Translation*. https://doi.org/10.51708/apptrans.v14n2.1238
Nababan, M. R. (2012). Pengembangan model penilaian kualitas terjemahan. *Kajian Linguistik Dan Sastra*, *24*(1), 39–57.
Newmark, P. (1988). *A textbook of translation*. Prentice Hall.
Nida, E. A. (1964). *Toward a Science of Translating* (2nd ed.). Brill.
Nord, C. (1997). *Translating as a purposeful activity: Functionalist approaches explained*. St. Jerome Publishing.
Poibeau, T. (2017). *Machine translation*. MIT Press. https://doi.org/10.7551/mitpress/9780262035899.001.0001
Poibeau, T. (2019). Machine Translation. In *3D Printing*. The MIT Press. https://doi.org/10.7551/mitpress/11800.003.0016
Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1985). *A comprehensive grammar of the English language*. Longman.
Suhono, S., Zuniati, M., Pratiwi, W., & Hasyim, U. (2020). *Clarifying Google Translate Problems Of Indonesia-English Translation Of Abstract Scientific Writing*. https://doi.org/10.4108/eai.24-1-2018.2292399
Untara, W., & Setiawan, T. (2020). Problema Mesin Penerjemah Berbasis AI dalam Proses Penerjemahan Buku Inggris-Indonesia dan Solusinya. *Jurnal Bahasa Dan Sastra*, *4*(1), 92–115.
Vinay, J.-P., & Darbelnet, J. (1958). *Stylistique comparée du français et de l'anglais: Méthode de traduction*. Didier.
We Are Social, & Meltwater. (2025). *Digital 2025: Global Overview Report.* wearesocial.com
Winiharti, M., Syihabuddin, & Sudana, D. (2021). The english Google translation of Indonesian lecturer's academic writing: A preliminary study. *Journal of Language and Linguistic Studies*, *17*(2), 706–719. https://doi.org/10.52462/jlls.49