IDEAS
Journal of Language Teaching and Learning, Linguistics and Literature
ISSN 2338-4778 (Print)
ISSN 2548-4192 (Online)
Volume 14, Number 1, June 2026
pp. 4067 - 4084

Issued by English study program of UIN Palopo

Exploring The Quality of AI Translation in Roblox Chat

Siti Nabila Amira Wangsi[1], Erfan Muhamad Fauzi[2], Fourus Huznatul Abqoriyyah[3]
[1,2,3]English Literature, Adab and Humanities, Islamic State University of Sunan Gunung Djati Bandung

Article Info

Received: 2026-06-25
Revised: 2026-06-29
Accepted: 2026-06-29

Keywords:
translation quality assessment,
machine translation,
neural machine translation,
Roblox,
translation

DOI:
10.24256/ideas. v14i1. 11458

Corresponding Author:
Siti Nabila Amira Wangsi
nabilawangsi@gmail.com
English Literature, Adab and Humanities, Islamic State University of Sunan Gunung Djati Bandung

Abstract

This study evaluates the translation quality of Roblox's AI translation system by combining translation quality assessment (TQA) with syntactic clause analysis in a multilingual online gaming environment. The study aims to examine how the system translates different clause types based on Nababan's theory, which is assessed by parameters of accuracy, acceptability, and readability. A qualitative approach was employed using 15 representative Indonesian–English translation pairs, selected from an initial dataset of 34 chat conversations collected from real-time Roblox gameplay. The data were classified into declarative, interrogative, and imperative clauses based on the syntactic framework of Huddleston and Pullum before being evaluated using the TQA model. The results show that the Roblox AI translation system performed best on interrogative clauses, while imperative and informal declarative clauses produced lower translation quality due to frequent lexical, syntactic, and semantic errors. The overall translation quality score was 1.98 on a 1–3 scale, indicating moderate-to-low translation quality. These findings demonstrate that combining TQA with syntactic clause analysis provides a more comprehensive approach to gaming communication and highlights the need for translation models that better accommodate informal language, gaming-specific expressions, and context-dependent interactions.

1. Introduction

Gaming has become a primary area to build connection, expression, and influence with people, especially Gen Z and Gen Alpha. Roblox is an online platform and large-scale game creation system that allows users to design, share, and play custom-made experiences using tools called Roblox Studio. The platform serves as a virtual space that facilitates social interaction, collaboration, and creativity through features such as live chat and group activities(Wicaksono, Anwar, & Asmara, 2025). In this situation, Roblox provides a tool called AI translation that facilitates interactions among players going well.

This dependency has evolved in recent years as the platform has grown quickly across the globe. The global expansion is forcing Roblox to have and demonstrate an accurate translation, which is crucial to ensuring a smooth social interaction, and maintain coordinated communication to reduce misunderstandings among players. To achieve effective communication in a dynamic gaming environment, the translation process must go beyond literal substitution by preserving the syntactic structure and communicative function of the source text.

Syntax is a branch of grammar that examines how words are organized into clauses and sentences, with the clause serving as the fundamental grammatical unit consisting of a subject and predicate. According to Huddleston & Pullum (2002), clauses are classified into four major types: declarative, interrogative, imperative, and exclamative, each distinguished by specific syntactic features. This classification provides a systematic framework for analyzing whether machine translation preserves the structural relationships of the source language while producing grammatically appropriate constructions in the target language.

Since translation quality is evaluated in terms of accuracy, acceptability, and readability, the successful mapping of syntactic constituents is essential to preserving meaning, producing natural target-language expressions, and ensuring ease of comprehension. Consequently, clause structure analysis serves as an objective basis for examining the relationship between syntactic patterns and the quality of machine translations.

Newmark (1988)states that translation not only involves transferring word-for-word and lexical matter, but also conveying meaning for communicative and pragmatic nuances. Effective translation requires a semantic equivalent, acceptability, and exact context. Unlike human translation, a machine translation system works by statistical shape and probability, often producing fails understanding of implicit content, and non-standard language used in online communication. Classic translation contains understanding text on many levels, including textual, referential, naturalness, and cohesive levels. Hence, translation can be read as an interpretative process which forcing to equivalent in literal and communicative terms depending on the text.

ISSN 2338-4778 (Print)
ISSN 2548-4192 (Online)

Historically, evolved machine translation focusing on scientific text and substances has a structure and formal terminology. In this case, rather than processing casual conversational expressions, translation algorithms were built to handle highly organised language. This historical limitation has caused Machine Translation to frequently encounter difficulties when applied to informal and spontaneous forms of communication, such as online conversations (Dostert, 1963).Otherwise, the transition to Neural Machine Translation (NMT) has increased the fluency of machine-generated language, but it has not addressed many fundamental semantic and pragmatic concerns.

According to Poibeau (2017), NMT is a coding language that uses complex linear representations to generate smoother, more natural sentences. Nevertheless, this increase in fluency does not translate into an increase in meaning accuracy. As a result, the system often produces mistranslations or semantically inaccurate outputs when processing informal chat expressions. These limitations are particularly evident in Indonesian–English translation, as Indonesian lacks grammatical tense and gender distinctions while relying on flexible word order and pragmatic particles such as *lah, kok, and kan*, which have no direct English equivalents.

Current developments in translation technology have shifted from conventional NMT systems to more adaptive approaches by integrating Large Language Models (LLMs). Although LLMs improve translation performance through techniques such as in-context learning and chain-of-thought prompting, they still face challenges in maintaining accuracy and consistency when translating syntactically complex sentences or texts requiring pragmatic interpretation, particularly in low-resource languages (Wang et al., 2024). Therefore, translation quality assessment (TQA) is crucial for evaluating effectiveness. Translation quality can be assessed by three aspects: accuracy, acceptability, and readability (Nababan, Nuraeni, & Sumardiono, 2012). This model presents a practical framework to evaluate non-formal translation.

Nababan et al. (2012) state that a good quality translation is easy to understand. Many translation scholars agree that translation is considered good quality if: (1) translation is accurate in content, (2) translation is expressed in accordance with the rules of the target language and does not conflict with the norms and culture of the target language, and (3) translation can be easily understood by the target readers. He proposes a translation quality assessment model based on three parameters: accuracy, acceptability, and readability. Accuracy is the first parameter of assessing translation quality because it is directly connected to correctly conveying the meaning.

Acceptability related to the reasonable translation level. Within this parameter, translation was assessed based on grammatical rules, lexical choices, and linguistic norms that apply in the target language. Acceptable translations must feel natural and not show the influence of the source language. Readability is

assessed by the extent to which the translation is easy to understand for the reader without needing much effort. This parameter focuses on clear sentences, fluency information, and the complexity level of the language used (Nababan et al., 2012).

Previous studies highlight weaknesses in AI-based automatic translation across digital platforms. Research on Twitter and Facebook shows that their translation systems frequently rely on literal strategies, leading to inaccurate meanings, grammatical errors, and poor handling of context, idioms, and complex sentence structures (Cahyaningrum, 2021; Kanoko & Sutrisno, 2024; Umbar, Aisyah Adilah, Rahma Raihana Romadona, & Muhammad Rafli, 2023). Similar limitations are found in YouTube's AI-generated subtitles for song lyrics, where pragmatic, linguistic, and textual errors emerge due to difficulties in translating figurative and poetic language (Haikal & Ali, 2025). Other studies indicate that ChatGPT produces Arabic–Indonesian translations of classical Islamic texts with generally high accuracy, acceptability, and readability, though expert supervision is still required (Ma'arif, Salma, & Afyudiin, 2025).

Additionally, this research aims to explore the quality, including accuracy, acceptability, and readability, of Indonesian–English AI translation in Roblox chat. This study assumes that different clause types exhibit varying levels of syntactic complexity, which may influence the performance of Roblox's built-in AI translation system. Consequently, translation quality is examined by relating clause type and syntactic complexity to the TQA parameters of accuracy, acceptability, and readability.

2. Method

Research Design

This research uses a qualitative approach based on the framework of Creswell, with a purposive sampling technique to analyse AI quality translation in Roblox chat. Creswell & Creswell (2018) explain that qualitative research emphasises detailed description and interpretation of meaning within real contexts.

Source of Data

The source of data consists of naturalistic textual data extracted from real-time chat conversations on the Roblox platform. A total of 34 Indonesian–English translated chat pairs were initially collected from gameplay sessions. After applying the predetermined sampling criteria, 15 representative sentence pairs were selected as the final dataset for analysis. These chat messages represent authentic user interactions and serve as the primary data for examining the syntactic performance and translation quality of the Roblox NMT system.

ISSN 2338-4778 (Print)
ISSN 2548-4192 (Online)

Technique of Collecting Data

Primary data were collected through participatory observation while playing games on the Roblox platform with friends, then manually transcribed into the data table to facilitate the classification and the identification of linguistic errors produced by the translation machine (Creswell & Creswell, 2018). The inclusion criteria consisted of: (1) chat messages originally written in Indonesian and automatically translated into English by the Roblox AI translation system; (2) complete sentence pairs with clear source and target texts; (3) representative examples of declarative, interrogative, and imperative sentence types; and (4) utterances containing sufficient linguistic features for syntactic and translation quality analysis. Messages containing only emojis, symbols, duplicated expressions, or incomplete utterances were excluded.

Before the translation quality assessment was conducted, each sentence was classified according to its clause type based on English syntactic characteristics, namely declarative, interrogative, and imperative structures. The classification was performed before the analysis to ensure that every translation example represented the intended sentence category consistently. The data were obtained from naturally occurring in-game conversations without collecting personal information such as usernames, player IDs, or other identifiable data.

The translation quality assessment was conducted by the researcher using the translation quality assessment (TQA) framework proposed by Nababan, which evaluates translation quality based on three parameters: accuracy, acceptability, and readability.

Table 1.Parameter of Translation Quality

Parameter	Score	Indicators
Accuracy	3	• Identical meaning
		• No distortion
		• No additions
		• No omissions
		• Correct terms
		• Proper context
	2	• Missing parts
		• Slight distortion
		• Ambiguous phrasing
		• Mistranslated terms
	1	• Wrong translation
		• Total omission
		• Fatal additions
Acceptability	3	• Standard grammar
		• Natural flow
		• Common word choice

		• No SL interference
		• Correct tenses/agreements
		• Proper register/context
	2	• Minor grammar errors
		• Rigid/awkward terms
		• Slight SL interference
	1	• Literal/word-for-word
		• Violates grammar rules
		• Violates cultural norms
Readability	3	• Understandable in a single read
		• Not confusing
		• Clear wording
		• Familiar vocabulary
		• Smooth sentence flow
		• Understandable without SL text
	2	• Requires re-reading
		• Slightly complex/wordy
		• Minor parts cause confusion
	1	• Completely unclear
		• Nonsensical phrasing

Technique of Analysis Data

The collected data are analyzed using a translation quality assessment formula. Each parameter is assessed on a scale of 1 to 3 and calculated using the Nababan weighted assessment formula, namely:

$$Total\ Score = \frac{[(Accuracy\ x\ 3) + (Acceptance\ x\ 2) + (Readability\ x\ 1)]}{6}$$

This formula is used to compute the final translation quality score. Accuracy received the highest weight because preserving the meaning of the source text is considered the primary objective of translation, followed by acceptability, which assesses conformity with the rules and the naturalness of the translation in the target language, while readability assesses whether the text can be understood after semantic equivalence has been maintained. Dividing by 6 produces a weighted average, so the final score reflects the quality of the translation proportionally (Nababan et al., 2012).

3. Results

This section presents the results of an analysis of the quality of machine translation in player conversations in the Roblox chat. The data presented includes evaluations of accuracy, acceptability, and readability, which are then classified by clause type to more comprehensively identify system performance patterns.

ISSN 2338-4778 (Print)
ISSN 2548-4192 (Online)

Table 2.Average Score

Parameter	Average Score
Accuracy	1,97
Acceptability	1,91
Readability	2,12
Final Score	**1,98**

Figure 1.Score Based on Clause Type

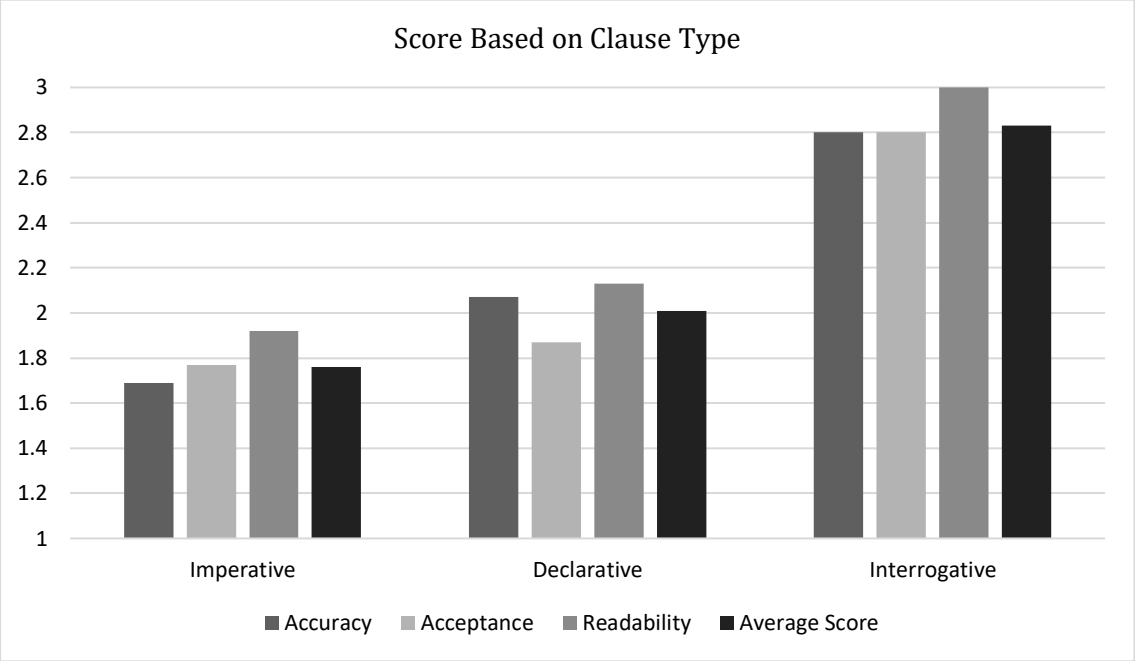

Figure 1 shows that the analysis included three clause types: imperative, declarative, and interrogative. No exclamative clauses were identified in the collected dataset. This is because the naturally occurring Roblox conversations observed during data collection predominantly consisted of commands, statements, and questions. Emotional expressions were generally conveyed through declarative sentences, interjections, abbreviations, or sentence fragments rather than canonical exclamative clause structures. Therefore, the absence of exclamative clauses reflects the linguistic characteristics of the collected corpus and does not affect the analysis of the clause types that occurred naturally during gameplay.

Table 3. Translation Quality Assessment in Imperative Simple Sentences

Data	Source Language	Target Language	Accuracy	Acceptability	Readability	Final Score
3	ngumpet ada hantunya	dummy, there's a ghost	2	2	2	2
9	cari lensa yang satunya	looking for the other one	3	3	3	3
14	sanaan, aku ga bisa liat	there, i cant see	2	2	2	2
19	gelap banget, hidupin lampunya	so dark, livein his lamp	1	1	1	1
26	ayo jalan lagi	come on road again	1	1	1	1

Table 4.Translation Quality Assessment in Declarative Simple Sentences

Data	Source Language	Target Language	Accuracy	Acceptability	Readability	Final Score
1	jalannya jongkok, dia bisa denger kita	his way is rumbling; he can avenge us	1	1	1	1
15	kipasnya nyala	its nyala fan	1	1	1	1
24	aku di depan pintu, ga ada kamu tuh	i'm at the door, no you tuh	2	1	2	1.6
31	aku lihat hantunya, dia di depan aku tadi	i saw her ghost, he in front of me that	1	1	1	1
33	kuncinya ga bisa diambil	key ga can be taken	1	1	1	1

ISSN 2338-4778 (Print)
ISSN 2548-4192 (Online)

Table 5.Translation Quality Assessment in Interrogative Simple Sentences

Data	Source Language	Target Language	Accuracy	Acceptability	Readability	Final Score
11	kita harus kemana?	where should we go?	3	3	3	3
17	tadi itu apa?	what is that?	3	3	3	2
28	kunci ini buat apa sih?	what the hell is this key for?	2	2	3	2.16
29	kamu liat apa?	what are you seeing?	3	3	3	3
30	kenapa dua-duanya kebuka?	why are both open?	3	3	3	3

Table 6. Frequency of Error Translation

Error Category	Description	Data	Frequency
Lexical Error	Incorrect word choice or untranslated lexical items that alter the intended meaning.	1, 3, 15, 19, 26, 28, 31	**7**
Syntactic Error	Incorrect sentence structure, word order, or grammatical arrangement.	14, 15, 19, 24, 31, 33,	**6**
Semantic Error	Distortion or loss of the original meaning.	1, 3, 9, 19, 26	**5**
Source Language Interference	Indonesian words or structures remain in the English translation.	15, 24, 33	**3**
Omission	Source-language information is omitted in the translation.	3	**1**
Addition	Information not present in the source text is added in the translation.	19, 31	**2**
Pragmatic Error	Failure to preserve contextual or interpersonal meaning.	28	**1**

4. Discussion

This section discusses the translation quality of AI translation on the Roblox platform, which is divided into three types of sentences: imperative, declarative, and interrogative. Assessment using instruments, TQA with weightings: Accuracy

(3), Acceptability (2), and Readability (1). The 15 data points were chosen because they are unique and representative.

Quality Translation in Imperative Clauses

Data 3:

SL: *"ngumpet ada hantunya"*

TL: "dummy, there's a ghost"

This simple sentence is classified as an imperative form that serves as an urgent warning for teammates to seek protection. Still, Roblox's AI translates "ngumpet" which means to hide as dummy, which shows the existence of lexical simplification that influences accuracy scores, as confirmed in the framework of the systematic evaluation of Saehu & Hkikmat (2025a). The quality of translations shows a low score with a 2 on each parameter. On aspect accuracy, the translation was deemed inappropriate lexical error in which the word "ngumpet" (hiding) was translated as "dummy" (stupid), so that the message of survival strategy becomes distorted; on aspect acceptance, the use of this irrelevant diction makes the sentence only to inform or statement instead of imperative speech; while in aspects readability, although the sentence structure is technically legible, word selection "dummy" which has no semantic connection to the game's horror context causes ambiguity.

Data 9:

SL: *"cari lensa yang satunya"*

TL: "look for the other lens"

This data is classified as an imperative form, serving as a clear instruction to teammates to search for specific objects (lenses) to advance the game's progress. Based on the assessment instrument, this data obtained a perfect score of 3 on all three parameters: because the machine manages to capture the functional meaning of instructions precisely without any missing information or distortion; acceptance, because the translation results in the target language feel natural and in accordance with internal communication norms games, making it easier for players to understand the objective; as well as readability, because the sentence structure is straightforward, so messages can be processed quickly by players in intense game situations.

Data 14:

SL: *"sanaan, I can't see it"*

TL: "there, i cant see"

ISSN 2338-4778 (Print)
ISSN 2548-4192 (Online)

This data is classified into imperative form, which asks teammates to shift so that the player's view is not obstructed. Based on the assessment instrument, this data obtained a score of 2 for all parameters: accuracy, because even though the functional message is conveyed, there is a reduction in meaning where the word “sanaan” (instruction to shift) is only translated as a pointing word “there”, so that the nuance of the command to move is lost; acceptance, because the translation results in the target language feel less natural and lose the urgency of the command that should be felt by the person you are talking to; as well as readability, because although the sentence is short enough to understand, missing the specific imperative aspect makes the instruction ambiguous for the player of the target language. Such partial meaning transfer may affect overall translation quality despite the sentence remaining understandable (Ma’arif et al., 2025).

Data 19:

SL: *“gelap banget, hidupin lampunya”*

TL: “so dark, livein his lamp”

This data is classified into imperative forms that command to turn on the light amidst dark game situations. Based on the assessment instrument, this data obtained the lowest score of 1 on all parameters because of a significant shift in meaning due to the inaccuracy of selecting lexical equivalents, which can obscure the original intent (Karyani, Nababan, & Marmanto, 2020). Due to a fatal lexical processing failure in the verb “hidupin” (turn on) is translated literally be "live-in", which results in the instruction message becoming completely meaningless; does not reflect communicative norms in the target language and appears illogical to English players; as well as using wrong diction creates sentences that completely unclear and confusing the target player from understanding the need for necessary action. Roblox’s AI is still stuck in single-word-based processing without the ability to understand technical context or electronic devices, which, in context, can cause significant coordination failures between players.

Data 26:

SL: *“ayo jalan lagi”*

TL: “come on road again”

This data is classified into imperative form that asks to keep walking and continue exploration in the game. The translation quality evaluation shows the lowest score (1) across all parameters. It shows weaknesses in grammatical aspects, where automatic evaluation metrics, as suggested by Wu, Wang, & Li (2022)can help identify these distortions of meaning more objectively.

Accuracy, because the Roblox’s AI system fails to identify the word class “jalan” as a verb (move/walk), and instead translates it literally to “road”, so that the purpose of ask to move becomes unconveyed and the meaning of the instruction

changes completely; acceptance, because of the translation results "come on road again" sounds very unnatural and does not fit the context of the conversation games ongoing; as well as readability, because the structure of the target sentence becomes unclear and confusing, which makes it difficult for English players to understand the intended action expected by the speaker.

Quality Translation in Declarative Clauses

Data 1:

SL: *"jalannya jongkok, dia bisa denger kita"*
TL: "his way is rumbling, he can avenge us"

This data is classified as a declarative form that reports the situation or behavior of other players in the game. Evaluation of translation quality in this data shows the lowest score of 1 on all parameters. The NMT system experienced bad semantic failure by translating the word *"jongkok"* (squat) into "rumbling" and *"denger"* (hear) becomes "avenge", so the original meaning of the statement is completely lost and turned into an irrelevant context; acceptance, because the translation results in the target language sound illogical and confusing, creating a sentence structure that is completely inconsistent with internal communication norms games; as well readability, because extreme distortions of meaning make the message difficult for target language players to understand, which directly hinders the exchange of accurate information regarding the situation on the ground. Structure is a considerable deviation in meaning, so according to Koka et al. (2023), it is necessary to standardize AI-based evaluation to improve the quality of translation in complex sentences.

Data 15:

SL: *"kipasnya nyala"*

TL: "its nyala fan"

Translation quality evaluation shows the lowest score of 1 on all parameters. The system fails to translate the word *"nyala"* and instead leaves it in the source language (*transliteration failed*), so that the equivalent words for the active status are not conveyed, and cause incomplete information. The translation results "its nyala fan" feel very strange and unnatural for target language players who do not understand Indonesian vocabulary, as well as because mixing the target sentence with the source language hinders quick understanding, forcing players to guess the meaning of the statement. This failure proves that the NMT system does not yet have an adequate lexical database for generic, informal descriptive terms in the game.

ISSN 2338-4778 (Print)
ISSN 2548-4192 (Online)

Data 24:

SL: *"aku di depan pintu, ga ada kamu tuh"*

TL: "i'm at the door, no you tuh"

This statement conveys player position information to teammates. The results of the translation quality evaluation show the score accuracy 2, because even though the information on the location and where is the player is conveyed, there is a distortion of meaning due to the system's inability to process the word "tuh" particle which actually leaves a void of meaning; score acceptance 1, due to significant syntactic error in existential loss *to-be* which should be translated as "there is no you" but it just becomes "no you"—as well as the remaining source language particles *"tuh"* make sentences not grammatical and unnatural to the target language player; as well as score readability 1, because even if the sentence structure is grammatically damaged, the main message still be guessed or understood by the reader even if it requires extra effort in processing the sentence.

Data 31:

SL: *"aku lihat hantunya, dia di depan aku tadi*

TL: "i saw her ghost, he in front of me that"

This data serves to state information that someone saw the ghost. The results of the translation quality evaluation show a score of 1 in all parameters. The system fails to translate the stressing *"-nya"* correctly and instead translates to the pronoun "her". In Indonesian, the suffix *"-nya"* can be used for stressing or as a pronoun of the third person. Also, in the second clause, the system removed to-be and mistranslated *"tadi"* (before) to "that". So, it can be seen that either the accuracy, acceptability, and readability are fail.

Data 33:

SL: *"kuncinya ga bisa diambil"*

TL: "key ga can be taken"

This data is classified as a declarative form that reports technical obstacles or the status of objects in the game. Evaluation of translation quality in this data shows the lowest score of 1 on all parameters. The NMT system fails to translate the negation *"ga"* (no) and leaves it untranslated. Moreover, the translation results are very unnatural due to language mixing and chaotic use of grammar, which violates the communication norms of the target language. The word "can" in translation makes the message very misleading for the player of the target language, which has the potential to lead to fatal misunderstandings where the player thinks the object can be retrieved, when in reality it is the opposite.

Quality Translation in Interrogative Clauses

Data 11:

SL: "*kita harus kemana?*"

TL: "where should we go?"

This data is classified as an interrogative form to ask teammates for directions or decisions regarding the purpose of the trip in the game. Evaluation of translation quality shows a perfect score of 3 on all parameters. The system manages to capture the speaker's communication intentions and translate them precisely into English interrogative form. The result is very natural, grammatical, and in accordance with communication norms in digital games; in addition, the message can be understood instantly by target language players without cognitive barriers.

Data 17:

SL: "*tadi itu apa?*"

TL: "what is that?"

This data is classified as an interrogative form to ask for information or clarification regarding an event or object that has just occurred/appeared in the game. Evaluation of translation quality shows a perfect score of 3 on all parameters. The NMT system handles the question structure of the source language in a semantically correct and accurate manner. The translation is acceptable because the result is very natural, grammatical, and in accordance with communication norms in digital games without any distortion of meaning, as well as readability, because the message can be understood instantly and clearly by the target language player.

Data 28:

SL: "*kunci ini buat apa sih?*"

TL: "what the hell is this key for?"

Evaluation of translation quality in this data shows an accuracy and acceptability score of 2, due to a significant addition phrase in which the affirming particle *"sih"*— which in this context functions as a marker of stress in the sentence or relaxed curiosity is inserted into the swear phrase "the hell" which gives an aggressive or rude feel that is not present in the original speech. This violates the norm of politeness in the communication of the target language, so that the translation feels pragmatically imprecise.

However, the readability score is 3 given because even though there is a shift in tone, the main meaning of the question regarding key functions is still conveyed clearly to the target language player without obstacles to understanding. The inability of the NMT system to translate Indonesian modal particles contextually is

ISSN 2338-4778 (Print)
ISSN 2548-4192 (Online)

in line with the findings of Karyani et al. (2020), which emphasizes that failure to capture pragmatic and cultural nuances is often the main weak point in the automatic translation process, even when the message appears lexically legible.

Data 29:

TL: "*kamu liat apa?*"

SL: "what are you seeing?"

Translation quality in this data shows a perfect score (3) on all parameters: accuracy, because the NMT system manages to capture the nuances of progressive aspects in the source language and translate them into the right present continuous tense in English, so that the meaning of ongoing actions is conveyed accurately; acceptability, because the translation results are very natural, grammatical, and in accordance with communication norms in digital games; as well as readability, because the message can be understood instantly and clearly by the player of the target language without any obstacle to understanding.

Data 30:

SL: "*kenapa dua-duanya kebuka?"*

TL: "why are both open?"

The results of the evaluation of translation quality in this data show 3 in all parameters. The system succeeded in identifying and translating *"dua-duanya"* precisely into the target language as "both", so that the original meaning of the message is conveyed completely. Moreover, results feel very natural, grammatical, and in harmony with communication norms in digital games, and the message can be understood clearly and quickly by the target language player. This success shows that the system can handle specific terms referring to the quantity of objects well, which is in line with the arguments in Wu et al. (2022) that accuracy in language model-based translation is highly dependent on the system's ability to process lexical entities contextually.

This study revealed significant differences in translation quality across clause types. Interrogative clauses achieved the highest performance because their standardized syntactic patterns, such as *wh*-questions and auxiliary verbs, provide explicit grammatical cues that facilitate accurate translation. In contrast, imperative clauses produced the lowest performance due to the frequent omission of subjects in informal game chat, which obscures the intended directive meaning. These findings indicate that NMT performs more reliably when processing structurally explicit clauses but remains vulnerable to informal constructions characterized by ellipsis and reduced grammatical marking.

These findings are in line with Mukande, Dinkov, Superbo, & O'Connor (2025), which states that NMT still has difficulty translating video game dialogue that contains humor, creative expression, and context that is highly dependent on the situation. These limitations have the potential to reduce pragmatic accuracy so that

the speaker's intentions are not always conveyed completely to the player. Along the same lines, Shahmerdanova (2025) also asserts that NMT still has weaknesses in dealing with cultural nuances, idioms, and contextual meanings. In contrast to the game localization scenario, which still allows for a post-editing process by humans, communication on the Roblox platform takes place in real time, so human translator intervention is impractical to implement.

Compared with other AI-powered translation systems such as Google Translate, DeepL, and ChatGPT, the Roblox NMT system demonstrates comparable strengths in producing fluent translations for structurally predictable sentences but shows lower performance when translating highly contextual, elliptical, and interaction-driven utterances that characterize real-time gaming communication.

This finding is consistent with Shujie (2025), who argues that ChatGPT performs effectively on standardized and technical texts but still struggles with complex contextual dependency, cultural adaptation, and implicit meaning, resulting in semantic deviations when the discourse context is limited. Likewise, Saehu & Hkikmat (2025)explain that although modern NMT-based systems, including ChatGPT, Google Translate, and DeepL, have substantially improved translation accuracy and readability, they remain challenged by cultural expressions, idioms, and highly contextual language. modeling dynamic conversational context during real-time multilingual interactions.

Therefore, the findings of this study show the importance of developing an NMT model that has better conversational context comprehension, slang, and online community expression capabilities so that translation quality can be improved. Extend previous translation studies research where messages are often fragmented and rely heavily on shared situational knowledge. From a computational linguistics perspective, the results also suggest that current NMT architectures remain limited in modeling dynamic conversational context during real-time multilingual interactions.

However, the findings of this research need to be interpreted according to the scope of the research because the analysis is only carried out on the AI translation on the Roblox platform with the English–Indonesian language pair. Therefore, the translation error patterns found do not represent system performance on other gaming platforms or different language pairs. Future improvement to AI translation models for multilingual gaming communities needs to be more focused on the ability to understand conversational contexts through a more specific corpus of game domains as well as better context modeling mechanisms. This recommendation is in line with Shujie (2025), who emphasizes the importance of improving contextual adaptation and dynamic feedback mechanisms to improve the quality of the NMT system.

ISSN 2338-4778 (Print)
ISSN 2548-4192 (Online)

5. Conclusion

Based on the results of the analysis of conversation data from Roblox, it can be concluded that the quality of the translations produced by the system NMT still varies. The system shows very stable performance in translating interrogative sentences, where question syntax patterns are successfully mapped accurately into the target language structure. However, this research also shows that there are data that obtain low scores due to inaccuracies in mapping sentence structure, distortions in syntactic construction, and the system's inability to process informal linguistic elements that deviate from standard structure. The inability of the system to handle advanced syntactic variations and affirming particles is a major obstacle to achieving ideal translation equivalence.

Therefore, assessing translation quality does not only require considering the commensurability of sentence structure, but also the aspects of acceptability and readability are in accordance with the communicative context. Finally, this research confirms that human involvement is still needed to evaluate and refine machine translation results, especially in informal texts that have high syntactic complexity beyond the routine patterns that the current system is capable of reaching.

6. References

Cahyaningrum, I. O. (2021). *Quality of Automatic Translation for Post on Facebook*. Retrieved from http://publikasi.dinus.ac.id/index.php/estructural

Creswell, John. W., & Creswell, J. D. (2018). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*.

Dostert, L. E. (1963). *MACHINE TRANSLATION AND AUTOMATIC LANGUAGE DATA PROCESSING*.

Haikal, F., & Ali, M. (2025). Translation Error in Auto Translation Feature in Song "Higuchi Ai-Akuma No Ko" The First Take. In *International Journal of Computer in Humanities* (Vol. 5).

Huddleston, R., & Pullum, G. K. (2002). *THE CAMBRIDGE GRAMMAR OF THE ENGLISH LANGUAGE*.

Kanoko, N. P., & Sutrisno, A. (2024). Translation Evaluation of Social Media Auto-Translate Feature toward News Headline. *Sasdaya: Gadjah Mada Journal of Humanities*, *8*(2), 235–255. https://doi.org/10.22146/sasdaya.13180

Karyani, L., Nababan, M. R., & Marmanto, S. (2020). Translation Analysis on Dayak Cultural Terms From Dayak Ngaju to Indonesian and English. *Langkawi: Journal of The Association for Arabic and English*, *6*(1), 41. https://doi.org/10.31332/lkw.v6i1.1676

Koka, N., Akan, Md. F., Kana'n, B. H. I., Khan, M. R., Zulfiquar, F., & Jan, N. (2023). IMPACT OF ARTIFICIAL INTELLIGENCE (AI) ON TRANSLATION QUALITY: ASSESSMENT AND EVALUATION. *Journal of Southwest Jiaotong University*, *58*(4). https://doi.org/10.35741/issn.0258-2724.58.4.72

Ma'arif, S., Salma, A., & Afyudiin, M. S. (2025). The Quality of ChatGPT's Arabic-Indonesian Translation of Matan al-Ghayah wa at-Taqrib: Implications for

Arabic Language Learning in Pesantren. *ATHLA : Journal of Arabic Teaching, Linguistic and Literature*, *6*(2), 81–97. https://doi.org/10.22515/athla.v6i2.12555

Mukande, T., Dinkov, D., Superbo, R., & O'Connor, N. (2025). Accelerating Workflows in Video Game Translation: A Recommender System for Review and Post-Edit Assignments. *ACM Transactions on Recommender Systems*. https://doi.org/10.1145/3778861

Nababan, M., Nuraeni, A., & Sumardiono. (2012). Pengembangan Model Penilaian Kualitas Terjemahan. *Kajian Linguistik dan Sastra*.

Newmark, P. (1988). *A TEXTBOOK OF TRANSLATION W *MRtt SHANGHAI FOREIGN LANGUAGE EDUCATION PRESS*.

Poibeau, T. (2017). *The MIT Press Essential Knowledge Series*.

Saehu, A., & Hkikmat, M. M. (2025a). The quality and accuracy of AI-generated translation in translating communication based topics: Bringing translation quality assessments into practices. In *Role of AI in Translation and Interpretation* (pp. 237–265). IGI Global. https://doi.org/10.4018/979-8-3373-0060-3.ch009

Saehu, A., & Hkikmat, M. M. (2025b). The Quality and Accuracy of AI-Generated Translation in Translating Communication-Based Topics. *Advances in Computational Intelligence and Robotics*, 237–265. Retrieved from https://doi.org/10.4018/979-8-3373-0060-3.ch009

Shahmerdanova, R. (2025). Artificial Intelligence in Translation: Challenges and Opportunities. *Acta Globalis Humanitatis et Linguarum*, *2*(1). https://doi.org/10.69760/aghel.02500108

Shujie, Z. (2025). Applications and Challenges of ChatGPT in Translation: A Multidimensional Exploration of Theory and Practice. *Sino-US English Teaching*, *22*(2), 35–41.

Umbar, K., Aisyah Adilah, Rahma Raihana Romadona, & Muhammad Rafli. (2023). Analisis Akurasi Penerjemahan Bahasa Arab Lewat Fitur Auto Translate pada Aplikasi Twitter. *Jurnal Naskhi: Jurnal Kajian Pendidikan Dan Bahasa Arab*, *5*(1), 21–33. https://doi.org/10.47435/naskhi.v5i1.1460

Wang, Y., Zhang, J., Shi, T., Deng, D., Tian, Y., & Matsumoto, T. (2024). Recent Advances in Interactive Machine Translation with Large Language Models. *IEEE Access*, *12*. https://doi.org/10.1109/ACCESS.2024.3487352

Wicaksono, R. D., Anwar, K., & Asmara, C. H. (2025). Students Perception of Roblox and Language Barrier in ESL. *Indonesian Teaching English to Speakers of Other Languages Journal*, *2*(2). https://doi.org/10.30587/inatesol.v2i2.9139

Wu, D., Wang, M., & Li, X. (2022). Automatic Scoring for Translations Based on Language Models. *Computational Intelligence and Neuroscience*, *2022*. https://doi.org/10.1155/2022/2171206