

Translation Quality Of Capcut Pro's Auto-Translated Captions In Indonesian Audiovisual Content (A Study of Instagram Reels from @fadiljaidi)

^{1*} Muhammad Abdee Praja Mukti, ² Ruminda, ³ Fours Huznatul Abqoriyyah

^{1,2,3} English Literature, UIN Sunan Gunung Djati, Bandung, Indonesia

*Corresponding Author e-mail: muhammadabdeeprajamukti@gmail.com

Received: May 2026; Revised: June 2026; Published: July 2026

Abstract

The development of artificial intelligence and machine translation technology has significantly transformed audiovisual communication within digital media platforms. One of the most widely used applications is CapCut Pro, which provides an auto-translated captions feature that automatically generates multilingual subtitles for audiovisual content. Despite its practical advantages, the translation quality of machine-generated subtitles remains problematic, particularly in informal conversational discourse commonly found on social media platforms. This study aims to analyze the translation quality of CapCut Pro's auto-translated captions in Indonesian audiovisual content, specifically Instagram Reels from @fadiljaidi. The study analyzes 53 subtitle data derived from one Instagram Reel video. The study employs a qualitative descriptive method using Translation Quality Assessment (TQA) principles, audiovisual translation theory, and machine translation analysis. The data consist of Indonesian spoken utterances and English subtitles automatically generated by CapCut Pro. The findings reveal that the generated subtitles demonstrate varying levels of translation quality categorized into accurate, moderately accurate, and inaccurate translations. Specifically, 34% of subtitles are categorized as accurate, 45% as moderately accurate, and 21% as inaccurate. Most subtitle outputs fall into the moderately accurate category, indicating that the system generally succeeds in transferring basic semantic meaning but still experiences contextual and pragmatic limitations. The dominant translation issues identified in the study include pragmatic shifts, lexical shifts, literal translation, grammatical awkwardness, and untranslated expressions. The findings further indicate that CapCut Pro's machine translation system struggles to interpret conversational Indonesian discourse particles, colloquial expressions, and culturally contextual utterances naturally in English. Although the feature improves multilingual accessibility in digital audiovisual communication, machine-generated subtitles still require human linguistic evaluation to ensure contextual accuracy and communicative naturalness.

Keywords: Translation Quality, Audiovisual Translation, Machine Translation, CapCut Pro, Subtitle Translation.

How to Cite: Mukti, M. A. P., Ruminda, & Abqoriyyah, F. H. (2026). An analysis of the translation quality of CapCut Pro's auto-translated captions in Indonesian audiovisual content: A study of Instagram Reels from @fadiljaidi. *Journal of Authentic Research*, 5(2), 4058–4074. <https://doi.org/10.36312/7qnv0q06>



<https://doi.org/10.36312/7qnv0q06>

Copyright© 2026, Mukti et al.

This is an open-access article under the [CC-BY-SA](#) License.



INTRODUCTION

The rapid development of digital communication technology has significantly transformed the production and consumption of audiovisual content across social media platforms (Nicolaou, 2021). Applications such as Instagram, TikTok, and YouTube Shorts have become dominant spaces for short-form audiovisual communication that reaches multilingual audiences worldwide (Dasola & Olayimika, 2026). In this digital environment, subtitles play an essential role in helping viewers understand audiovisual messages delivered in different languages. Audiovisual translation therefore becomes increasingly important because it functions not only as

linguistic transfer but also as a bridge for intercultural communication in multimedia environments. Chaume (2018) explains that audiovisual translation is a multidisciplinary field involving language, visuals, sound, and communicative interaction within audiovisual media. Similarly, subtitle translation has become one of the most influential forms of audiovisual translation in contemporary digital communication because subtitles improve accessibility and multilingual communication in online media (Chaume, 2018; Díaz-Cintas, 2020).

The increasing demand for multilingual audiovisual communication has encouraged the rapid development of machine translation and automatic subtitling technologies. Artificial Intelligence (AI) systems are now widely integrated into digital platforms to generate subtitles and translate audiovisual content automatically. Translation technologies continue to evolve alongside artificial intelligence development, enabling users to access multilingual communication more efficiently in digital environments (Bowker, 2023). In addition, localization and digital translation technologies increasingly influence global communication practices because online audiovisual content requires faster and broader multilingual accessibility (Jiménez-Crespo, 2024).

One of the most widely used applications integrating AI-based subtitle technology is CapCut Pro, a video editing application particularly popular among Indonesian content creators due to its free access, user-friendly interface, and advanced automatic subtitle features. The application provides an auto-translated captions feature that automatically generates subtitles and translates spoken language into different target languages, including English. This feature has become increasingly popular because it allows audiovisual content to reach international audiences more efficiently. Automatic subtitle systems are developed to combine speech recognition and machine translation technology in order to improve multilingual accessibility within audiovisual communication (Kristjansson et al., 2015). Furthermore, automatic subtitling technology has become increasingly important in digital audiovisual media because subtitle automation significantly accelerates subtitle production processes for online video communication (Papi et al., 2023).

Despite the advancement of machine translation technology, subtitle translation quality remains a major issue in audiovisual communication. Machine-generated subtitles frequently experience lexical inaccuracies, contextual misunderstanding, and pragmatic shifts because conversational meaning cannot always be translated literally. Translation quality depends on equivalence between source text and target text at lexical, grammatical, textual, and pragmatic levels (Baker, 2018). Similarly, translation quality assessment must evaluate whether communicative function and contextual meaning are successfully transferred into the target language (House, 2015).

The challenge becomes more complex in subtitle translation because audiovisual subtitles operate under technical constraints such as timing, synchronization, readability, and limited textual space. Subtitle translation therefore requires condensation and adaptation strategies in order to maintain communicative clarity within audiovisual media (Díaz-Cintas & Remael, 2020). Likewise, subtitle quality assessment should consider readability, synchronization, and communicative

effectiveness because subtitle meaning is strongly influenced by audiovisual interaction (Romero-Fresco & Pöhhacker, 2017).

Recent developments in Neural Machine Translation (NMT) have improved machine translation fluency and contextual prediction significantly. Neural Machine Translation systems use deep learning architectures to generate more natural and fluent translation outputs compared to earlier rule-based translation systems (Koehn, 2020). However, although NMT systems improve translation efficiency through contextual probability prediction, contextual interpretation and pragmatic understanding remain major limitations in machine-generated translation (Stahlberg, 2020).

The limitations of machine translation systems become particularly visible in social media communication because digital audiovisual discourse frequently contains informal expressions, slang, discourse particles, and culturally contextual utterances. Contextual appropriateness remains one of the most difficult challenges in machine translation because conversational meaning depends heavily on pragmatic interpretation and sociocultural context (Naveen & Trojovský, 2024). Similarly, contextual processing remains a significant weakness in Neural Machine Translation systems despite recent technological advancements (Castilho & Knowles, 2025).

Several previous studies have also identified translation problems within social media auto-translation systems. Instagram auto-translation frequently experiences lexical and contextual inaccuracies despite improving multilingual accessibility (Serinastity et al., 2025). Likewise, lexical errors, grammatical distortion, and contextual misunderstanding frequently appear in Indonesian-English TikTok auto-translation content (Shafira & Nugroho, 2023). These findings indicate that machine-generated subtitle translation still experiences significant limitations when translating informal audiovisual discourse.

The linguistic characteristics of the selected content creator, @fadiljaidi, further strengthen the rationale for this research. His Instagram Reels are known for their informal conversational style, frequent use of Indonesian discourse particles such as "deh," "lah," "dong," "nah," and "kok," as well as humor, sarcasm, and culturally contextual references. These features pose particular challenges for machine translation systems, which typically rely on formal and standardized language patterns.

The influence of AI on translation practice has also become increasingly significant within digital communication environments. AI-based translation systems are widely used because they improve communication efficiency and accessibility among digital users (Mukti et al., 2024). In addition, AI technologies contribute to lexical evolution and language adaptation within multilingual communication (Mukti & Fauzi, 2025). Translation quality assessment in machine translation systems therefore remains essential because automated translation frequently produces contextual inconsistencies despite improving translation speed (Moorkens et al., 2018). Furthermore, automated translation technologies should be critically evaluated because machine-generated outputs often require human linguistic revision to maintain communicative accuracy (Moorkens et al., 2024).

Machine translation quality assessment should also involve detailed error classification because translation errors often occur at lexical, grammatical, semantic, and pragmatic levels simultaneously (Popović, 2018). Similarly, machine translation

evaluation requires comparison with human-translated references in order to examine communicative accuracy and contextual appropriateness within social media communication (Al Amer et al., 2025). These findings demonstrate that machine-generated subtitle translation remains an important area of research within Translation Studies and audiovisual communication.

Furthermore, digital social media communication increasingly relies on multimodal interaction involving audiovisual elements, online engagement, and cross-cultural communication. Instagram and TikTok communication patterns are strongly influenced by digital linguistic strategies and audiovisual interaction, making translation quality increasingly important in global digital communication (Amure et al., 2026). Consequently, subtitle translation quality in social media environments should be evaluated not only linguistically but also communicatively and contextually.

Based on these issues, this study is guided by the following research questions: (1) What is the quality of CapCut Pro's auto-translated captions in translating Indonesian audiovisual content into English, categorized as accurate, moderately accurate, or inaccurate? (2) What types of translation issues appear in the machine-generated subtitles? (3) What are the strengths and limitations of CapCut Pro's auto-translated captions in translating informal Indonesian conversational discourse?

This study focuses specifically on analyzing translations from Indonesian as the source language into English as the target language. Other language pairs available in CapCut Pro are not examined in this research.

METHOD

Research Design

This study employs a qualitative descriptive research design to examine the translation quality of CapCut Pro's auto-translated captions in Indonesian audiovisual content. The qualitative approach is considered appropriate because the study focuses on analyzing meaning transfer, contextual equivalence, and translation issues in machine-generated subtitles rather than measuring variables statistically. The research applies Translation Quality Assessment (TQA) as the main analytical framework to evaluate whether the generated subtitles successfully preserve the semantic meaning and communicative intention of the original Indonesian utterances.

In this study, the analysis is not limited to identify whether the subtitles are correct or incorrect. Instead, it also examines the types of translation issues that emerge in the output of CapCut Pro. Therefore, the study combines issue identification and quality evaluation in a single analytical process. The subtitle data are first examined to identify the dominant translation issue and then assessed based on translation quality criteria. This approach allows the researcher to describe not only the quality level of the subtitles but also the linguistic reasons behind the evaluation.

The research focuses on Instagram Reels content from @fadiljaidi because the videos contain conversational Indonesian expressions, compound sentences, and complex utterances that are relevant for translation quality analysis. From the available subtitle data, 15 representative segments are selected purposively for close analysis. The selected data are expected to provide sufficient variation in lexical,

grammatical, and pragmatic phenomena, which makes them appropriate for a detailed qualitative study.

Subject of the Research

The subject of this study is one Instagram Reel uploaded by the Instagram account @fadiljaidi. The selected content was chosen because it contains conversational Indonesian language with informal expressions, compound sentences, and context-dependent utterances. These linguistic characteristics make the video suitable for translation quality analysis, especially in relation to machine-generated subtitles produced by CapCut Pro.

The study uses 15 representative subtitle segments taken from the selected Instagram Reel. The segments were chosen purposively based on their relevance to translation analysis. The researcher prioritizes subtitles that contain compound or complex structures, contextual meaning, and translation phenomena such as lexical, grammatical, and pragmatic issues. Short utterances that do not contribute significantly to translation analysis are not prioritized as research data.

The selected Instagram Reel has high audience engagement and contains sufficient subtitle data for translation quality analysis. Social media content was chosen because machine translation systems often experience contextual difficulties when translating informal and conversational language. Serinastity et al. (2025) and Shafira and Nugroho (2023) both found that automatic subtitle translation in social media frequently produces lexical inaccuracies, contextual misunderstanding, and grammatical awkwardness due to informal language usage (Serinastity et al., 2025; Shafira & Nugroho, 2023).

Research Instrument

The primary instrument of this study is the researcher, who serves as the main analyst in identifying translation issues and evaluating subtitle quality. Since this research uses a qualitative descriptive approach, the researcher plays an important role in interpreting meaning, classifying translation problems, and determining the quality of the subtitles.

In addition to the researcher, the study uses an analytical framework consisting of two main instruments. The first instrument is a translation issue classification table, which is used to identify and categorize the types of translation issues found in the subtitle data. The second instrument is a translation quality evaluation table, which is used to assess the subtitles according to three categories: Accurate, Moderately Accurate, and Inaccurate.

The translation issue categories are adapted from translation error classification principles proposed by (Popović, 2018) and supported by related studies on auto-translation error analysis such as Shafira and (Nugroho, 2023). The translation quality evaluation criteria are adapted from (Nababan et al., 2012), whose translation assessment model is widely used in translation studies.

Table 1. Translation Issue Categories

Category	Description	Reference
----------	-------------	-----------

Lexical-Semantic	Translation problems related to word choice, semantic distortion, literal lexical transfer, or failure to preserve the intended meaning of the source text.	Baker (2018); Popović (2018); Shafira & Nugroho (2023)
Grammatical	Translation problems involving sentence structure, tense, word order, or syntactic naturalness in the target language.	Baker (2018)
Pragmatic	Translation problems related to contextual meaning, communicative intention, discourse markers, and interpersonal meaning.	Baker (2018); House (2015)

Table 2. Translation Quality Evaluation Criteria

Evaluation	Criteria
Accurate	The subtitle successfully transfers the meaning, communicative intention, and contextual function of the source text without major distortion.
Moderately Accurate	The subtitle generally transfers the intended meaning, but minor lexical, grammatical, or pragmatic shifts are found. These shifts do not significantly affect overall comprehension.
Inaccurate	The subtitle fails to transfer the intended meaning because of semantic distortion, contextual misunderstanding, or major grammatical errors.

The category “Moderately Accurate” is used as a stylistic adaptation of “Less Accurate” in (Nababan et al., 2012), while maintaining the original evaluation principle of partial meaning transfer.

Research Procedure

The research procedure is carried out in several stages.

First, the researcher selects one Instagram Reel from @fadiljaidi that contains sufficient subtitle data for translation analysis. The selected video is then processed using CapCut Pro’s auto-translated captions feature to generate English subtitles automatically.

Second, the Indonesian spoken utterances and the English subtitles are transcribed and arranged into parallel data units. The researcher then selects 15 representative subtitle segments that contain compound or complex sentence structures and clear translation phenomena.

Third, each subtitle segment is examined to identify the type of translation issue. The identified issues are then classified into one of the three main categories: Lexical-Semantic, Grammatical, or Pragmatic.

Fourth, each subtitle is evaluated using the translation quality criteria adapted from Nababan et al. (2012). The subtitles are classified as Accurate, Moderately Accurate, or Inaccurate based on how well the target text preserves the meaning of the source text.

Fifth, the researcher calculates the frequency of each translation issue and each translation quality category in order to identify dominant patterns in the subtitle output. The results are then interpreted descriptively to explain the strengths and

limitations of CapCut Pro's auto-translated captions in Indonesian audiovisual content.

Data Analysis

The data are analyzed using a qualitative descriptive method combined with Translation Quality Assessment. The analysis begins by comparing the source text and the target text to identify the type of translation issue found in each subtitle segment. Each segment is first examined to determine whether the problem is lexical-semantic, grammatical, or pragmatic in nature.

After the type of issue is identified, the subtitle is classified into the appropriate evaluation category. The evaluation is based on the following logic:

1. If the subtitle successfully preserves the intended meaning and communicative function, it is classified as Accurate.
2. If the subtitle preserves the general meaning but still contains minor lexical, grammatical, or pragmatic shifts, it is classified as Moderately Accurate.
3. If the subtitle fails to preserve the meaning because of semantic distortion or contextual misunderstanding, it is classified as Inaccurate.

To make the analytical process more systematic, the study uses the following hierarchical framework: Source Text → Target Text → Type of Issue → Category → Evaluation → Frequency Calculation → Interpretation

Table 3. Translation Issue Framework

Type of Issue	Operational Definition	Reference
Semantic Distortion	The intended meaning of the source text is distorted because the translation fails to preserve semantic equivalence, resulting in a different or misleading message in the target language.	Popović (2018); Baker (2018)
Lexical Shift	The translation uses an inappropriate lexical equivalent that weakens, changes, or partially conveys the intended meaning without completely distorting the message.	Baker (2018); Popović (2018)
Literal Translation	The source text is translated word-for-word without adapting it to the contextual, idiomatic, or natural expression of the target language.	Shafira & Nugroho (2023); Popović (2018)
Grammatical Awkwardness	The translation generally preserves meaning but contains unnatural grammatical structures, incorrect tense, inappropriate word order, or other syntactic problems in the target language.	Baker (2018); Koehn (2020)
Pragmatic Shift	The translation fails to preserve contextual meaning, communicative intention, interpersonal tone, discourse markers, or implied meaning of the source utterance.	House (2015); Baker (2018)

The classification of translation issues presented in Table 3 serves as the analytical framework for identifying translation problems in the subtitle data. Each subtitle segment was first examined to determine the dominant translation issue based on its linguistic characteristics before being grouped into one of the broader analytical categories, namely lexical-semantic, grammatical, or pragmatic. This framework enables a systematic distinction between the specific translation problem

(*Type of Issue*) and its broader linguistic dimension (*Category*), thereby increasing the transparency and consistency of the analysis (Popović, 2018; Baker, 2018; House, 2015).

RESULTS AND DISCUSSION

Analytical Findings

The analysis of CapCut Pro's auto-translated captions in the Instagram Reels content from @fadiljaidi reveals that the generated subtitles demonstrate varying levels of translation quality. The findings show that CapCut Pro can transfer basic semantic meaning in several utterances, but it still struggles with contextual equivalence, grammatical naturalness, and pragmatic appropriateness when the source text contains compound or complex conversational structures. In line with the revised methodological focus, the representative data below are selected from subtitles that contain richer translation phenomena and are therefore more suitable for detailed analysis. (Baker, 2018; House, 2015; Popović, 2018).

Table 4. Representative Translation Quality Analysis

No	Source Text	Target Text	Type of Issue	Category	Evaluation
1	day I want this my husband	Today I want to make my husband	Semantic distortion	Lexical-Semantic	Inaccurate
2	The first We enter Formerly the sweet potato	the first one we put the yam in first	Grammatical Awkwardness	Grammatical	Moderately Accurate
3	half spoon straight salt tea entered Just yes	half a teaspoon of salt just put it in	Grammatical Awkwardness	Grammatical	Moderately Accurate
4	Then prepare 6 sheets pandan leaves after That concluded Keep going come on in	then prepare 6 pieces of pandanus leaves after it was knotted keep entering	Literal Translation	Lexical-Semantic	Inaccurate
5	I do not I use sugar I use soy sauce	I don't use sugar I used soy sauce	Grammatical Awkwardness	Grammatical	Moderately Accurate
6	stay entered Just soy sauce approximately 200 milli lah	just put the soy sauce in it's about 200 milli	Lexical shift	Lexical	Moderately Accurate
7	well now We enter the bananas if No there are bananas, right? banana compote, please	well now we put the bananas in if there are no bananas not banana compote	Pragmatic shift	Pragmatic	Inaccurate
8	after That We direct add the coconut milk	after that we immediately enter the coconut	Grammatical awkwardness	Grammatical	Inaccurate

No	Source Text	Target Text	Type of Issue	Category	Evaluation
	too Keep going what to do Again yes	milk too keep doing it again yeah			
9	stay stirred like this stirred stir-stir until evenly must until evenly yes	stay stirred dong like this stirred stirred until evenly distributed have to get evenly distributed yeah	Pragmatic shift	Pragmatic	Inaccurate
10	uh uh my hand Ouch understand Again fast	uh uh my hand oops you know fasting	Pragmatic shift	Pragmatic	Moderately Accurate
11	yes forgive me Darling You what to cook for me This Thanks Lots	oh my gosh baby you cook me what is this thank you very much	Pragmatic shift	Pragmatic	Moderately Accurate
12	I made it use material special lol	I made it with special ingredients	No Issue	Lexical	Accurate
13	oh Yes use material special What This	oh yes what special material use this	lexical shift	Lexical-Semantic	Inaccurate
14	Honestly, I didn't expect it. Darling You use soy sauce with coconut sugar as the composition organic So more low sugar yeah Darling yes	Honestly I don't expect that baby you use soy sauce composition organic coconut sugar so lower sugar yeah baby yeah	Pragmatic shift	Pragmatic	Moderately Accurate
15	right Like someone has an allergy The same soya bean like that yes Darling yes	some people are allergic to soy, right baby?	No Issue	Pragmatic	Accurate

The table shows that the translation outputs are not uniformly problematic. Some subtitles successfully preserve the intended meaning, while others experience visible shifts in meaning and form. The overall pattern suggests that CapCut Pro performs better when translating simple and direct utterances, but its output becomes less stable when the source text includes conversational Indonesian expressions, discourse markers, or culturally bound phrasing. This supports the view that translation quality must be assessed not only at the lexical level but also in terms of communicative function and contextual equivalence. (Baker, 2018; House, 2015)

The first clear pattern appears in lexical-semantic problems. In Data 1, the subtitle *"Today I want to make my husband"* shows semantic distortion because the source utterance *"bikinin suami"* does not mean creating the husband himself, but making food for him. This indicates that the machine translation system processes the utterance literally without adequately interpreting the benefactive function of the verb phrase. A similar problem appears in Data 4 and Data 13, where *"keep entering deh"* and *"what special material use this"* demonstrate that the system still struggles to choose lexical equivalents that fit the context. These findings are consistent with the idea that machine translation often prioritizes lexical matching over contextual meaning, which can produce semantic distortion in conversational language. (Popović, 2018; Baker, 2018)

A second pattern is grammatical inconsistency. In Data 2, the translation *"the first one we put the yam in first"* is understandable, but the word order sounds unnatural in English. Data 3 also shows a minor grammatical shift in *"half a teaspoon of salt just put it in,"* where the intended imperative structure is not fully realized in the target text. Data 5 and Data 8 show more visible grammatical awkwardness through tense inconsistency and incomplete sentence structure. These examples suggest that CapCut Pro can preserve the core propositional meaning of some utterances, but the grammatical realization in English is still imperfect. This aligns with the view that grammatical equivalence is not only about preserving meaning but also about producing natural and structurally acceptable target-language sentences. (Baker, 2018; Koehn, 2020)

The third and most dominant pattern involves pragmatic issues. In Data 7, the expression *"if there are no bananas not banana compote"* does not convey the persuasive function of the Indonesian discourse particle *dong*. The same problem appears even more clearly in Data 9, where *"stay stirred dong"* preserves the lexical form of the utterance but fails to transfer its interpersonal and conversational force. Data 11 and Data 14 also show pragmatic shift because the emotional tone, intimacy markers, and conversational nuance of the Indonesian utterances are only partly rendered in English. These findings demonstrate that the system struggles to interpret discourse markers, implied meaning, and relational tone, which are central features of informal audiovisual conversation. In this sense, the translation may remain semantically understandable, but it is pragmatically weakened. (House, 2015; Naveen & Trojovský, 2024; Castilho & Knowles, 2025)

At the same time, several subtitles are translated successfully. Data 12 (*"I made it with special ingredients"*) and Data 15 (*"some people are allergic to soy, right baby?"*) are relatively accurate because they preserve both meaning and communicative intent. These examples show that CapCut Pro performs better when the source utterance is straightforward and the contextual meaning is not highly dependent on discourse nuance. However, once the source text becomes more conversational, playful, or culturally loaded, the quality of the translation tends to decline. This pattern reflects the broader limitation of neural machine translation systems, which are generally strong in processing predictable structures but weaker in handling pragmatic and sociolinguistic complexity. (Koehn, 2020; Stahlberg, 2020; Castilho & Knowles, 2025)

Overall, the analytical findings indicate that CapCut Pro's auto-translated captions are capable of preserving basic meaning in some cases, but they still face substantial difficulty in translating conversational Indonesian discourse naturally and

accurately. The most problematic areas are lexical-semantic distortion and pragmatic shift, especially in sentences that contain idiomatic expressions, discourse particles, and culturally specific conversational markers. Therefore, the data suggest that machine-generated subtitles remain useful for multilingual accessibility, but they still require human revision when used for informal audiovisual communication. (Baker, 2018; House, 2015; Romer o-Fresco & Pöchhacker, 2017)

Translation Quality Classification

To provide a clearer overview of the translation quality produced by CapCut Pro's auto-translated captions, the subtitle outputs were classified into three evaluation categories: **accurate**, **moderately accurate**, and **inaccurate**. The classification was carried out using the evaluation criteria adapted from Nababan et al. (2012) and interpreted through the lens of translation quality assessment, equivalence, and audiovisual translation principles (Nababan et al., 2012; Baker, 2018; House, 2015).

Table 5. Translation Quality Classification

Evaluation Category	Total Data	Percentage	Interpretation
Accurate	18	34%	The subtitles successfully preserve semantic meaning and communicative intention.
Moderately Accurate	24	45%	The subtitles generally preserve meaning but contain minor lexical, grammatical, or pragmatic shifts.
Inaccurate	11	21%	The subtitles fail to preserve contextual meaning due to mistranslation or semantic distortion.
Total	53		-

Based on the table above, the majority of subtitle outputs generated by CapCut Pro fall into the moderately accurate category, representing 45% of the total analyzed data. This percentage indicates that the machine translation system is generally capable of conveying the primary semantic meaning of Indonesian audiovisual utterances into English. However, the subtitles still contain contextual, grammatical, and pragmatic inconsistencies that reduce communicative naturalness within conversational English discourse. This result suggests that CapCut Pro performs adequately at the level of general meaning transfer, but it is not yet fully reliable for maintaining natural and context-sensitive subtitle translation. (Nababan et al., 2012; Baker, 2018)

The moderately accurate category frequently appears in subtitle outputs that preserve the main message of the source text but still contain awkward lexical choices, unnatural sentence structures, or minor pragmatic shifts. Expressions such as *"I used soy sauce"* and *"tasty baby"* remain understandable semantically; however, they sound less natural in conversational English due to tense inconsistency or contextual awkwardness. In some cases, the target text is communicatively intelligible, but it does not fully reflect the interpersonal tone or discourse style of the source utterance. This finding supports Baker's view that translation quality should not be judged only by lexical correspondence, but also by pragmatic and contextual equivalence (Baker, 2018; House, 2015).

The accurate category represents 34% of the analyzed subtitles. These subtitles successfully preserve semantic meaning and communicative intention with relatively natural English structures. Accurate translations mostly appear in short and direct

utterances containing simple lexical expressions, such as *"thank you very much"* and *"this is soy sauce."* These examples demonstrate that machine translation systems generally perform effectively when processing straightforward language structures with clear contextual correspondence. This pattern is consistent with Koehn's explanation that Neural Machine Translation systems are strongest in handling predictable and structurally simple utterances, especially when the source text does not contain high pragmatic complexity (Koehn, 2020; Stahlberg, 2020).

Meanwhile, the inaccurate category accounts for 21% of the total data. Although this percentage is lower than the other categories, the inaccurate subtitles reveal significant contextual and pragmatic problems within machine-generated translation. Most inaccurate translations occur in conversational Indonesian expressions, colloquial discourse, and context-dependent utterances. For example, the subtitle *"Today I want to make my husband"* demonstrates semantic distortion because the system translates the Indonesian expression literally without understanding its contextual meaning. Similarly, subtitles such as *"stay stirred dong"* illustrate pragmatic failure because the discourse particle *dong* is translated structurally rather than communicatively. These findings indicate that CapCut Pro's machine translation system still struggles with sociolinguistic interpretation and conversational discourse processing, especially when the source utterance depends on implied meaning or informal interpersonal tone. (Naveen & Trojovský, 2024; Castilho & Knowles, 2025)

Overall, the percentage distribution demonstrates that CapCut Pro's auto-translated captions possess moderate translation effectiveness within Indonesian audiovisual social media content. The system successfully transfers basic semantic information in many cases; however, conversational Indonesian expressions, discourse particles, and culturally contextual utterances remain major challenges for machine-generated subtitle translation. From a translation quality perspective, this means that CapCut Pro is useful for improving multilingual accessibility, but it cannot yet replace human revision when the goal is to produce subtitles that are fully accurate, natural, and contextually appropriate. This finding aligns with the broader discussion in translation technology studies that automated systems remain efficient tools, yet still require human evaluation for communicative reliability (Bowker, 2023; Moorkens et al., 2024; Al Amer et al., 2025).

Types of Translation Issues

The analysis identified five dominant types of translation issues in the auto-translated captions generated by CapCut Pro, namely semantic distortion, lexical shift, literal translation, grammatical awkwardness, and pragmatic shift. These issues represent different linguistic problems that affect the overall quality of the translated subtitles. The frequency of each issue is presented in Table 6.

Table 6. Frequency of Translation Issues

Type of Issue	Frequency
Pragmatic Shift	17
Lexical Shift	14
Literal Translation	10
Grammatical Awkwardness	7
Semantic Distortion	5

The frequency distribution indicates that pragmatic shift is the most dominant issue found in the subtitle data, followed by lexical shift, literal translation, grammatical awkwardness, and semantic distortion. This pattern suggests that CapCut Pro generally succeeds in transferring the surface meaning of the source text but still experiences considerable difficulty in preserving contextual meaning and communicative intention, particularly in conversational Indonesian discourse. These findings reinforce the notion that subtitle quality should be evaluated not only through linguistic accuracy but also through communicative effectiveness in audiovisual contexts (Chaume, 2018; Díaz-Cintas & Remael, 2020).

Semantic Distortion

Semantic distortion occurs when the translated subtitle changes or misrepresents the intended meaning of the source utterance. Unlike lexical shift, semantic distortion results in a substantial alteration of meaning that may lead the audience to misunderstand the original message.

One representative example is the subtitle "Today I want to make my husband", translated from "*hari ini Aku mau bikin suami Aku.*" The Indonesian verb *bikin* expresses a benefactive action ("to make something for someone"), whereas the English translation incorrectly implies "creating the husband." This mistranslation changes the intended meaning and therefore represents semantic distortion rather than a simple lexical inaccuracy.

This finding demonstrates that CapCut Pro occasionally prioritizes direct lexical correspondence over semantic equivalence, leading to a loss of communicative meaning. Similar phenomena have been identified in machine translation studies, where lexical prediction does not always reflect contextual interpretation (Popović, 2018; Baker, 2018).

Lexical Shift

Lexical shift refers to the selection of a target-language word that is semantically related to the source expression but is not the most appropriate choice in the given context. Although the general meaning may still be understood, the translation becomes less precise or less natural.

For example, the subtitle "what special material use this" translates the Indonesian word *bahan* as *material* instead of *ingredient*. Both words are semantically related; however, *ingredient* is the appropriate lexical choice within the context of cooking. A similar phenomenon appears in several subtitle segments where the system preserves the overall meaning but weakens communicative naturalness because of inappropriate lexical selection.

These findings indicate that CapCut Pro is generally capable of identifying semantic relationships between words but still struggles to select contextually appropriate lexical equivalents in informal audiovisual discourse (Baker, 2018; Popović, 2018).

Literal Translation

Literal translation occurs when the subtitle is translated word-for-word according to the source-language structure without adapting it to the natural conventions of the target language.

A representative example is "we just jump right in", translated from "*kita langsung masukin.*" In the cooking context, *masukin* means "to add an ingredient,"

rather than "jump right in." Although the translation follows a literal interpretation of the source expression, it fails to communicate the intended procedural meaning.

This finding is consistent with previous research showing that automatic translation systems frequently rely on structural correspondence instead of contextual adaptation when translating conversational language. Consequently, literal translation often produces subtitles that are grammatically acceptable but pragmatically inappropriate (Shafira & Nugroho, 2023; Popović, 2018).

Grammatical Awkwardness

Grammatical awkwardness refers to subtitle outputs that generally preserve meaning but contain unnatural grammatical structures, inappropriate tense usage, or unusual word order.

For instance, the subtitle "the first one we put the yam in first" communicates the intended message but contains redundant wording and unnatural sentence structure. Similarly, "half a teaspoon of salt just put it in" lacks the imperative structure expected in natural English.

Although these translations remain understandable, their grammatical quality reduces subtitle readability and naturalness. This finding supports previous studies indicating that neural machine translation systems often preserve semantic content while still generating syntactically awkward sentences, particularly when translating spoken language (Koehn, 2020; Stahlberg, 2020).

Pragmatic Shift

Pragmatic shift is the most dominant issue identified in this study. It occurs when the subtitle fails to preserve communicative intention, interpersonal tone, discourse markers, or implied meaning contained in the source utterance. One representative example is the subtitle "stay stirred dong." The Indonesian discourse particle *dong* functions as a marker of persuasion and familiarity rather than carrying lexical meaning. Since English does not have a direct equivalent, CapCut Pro retains the particle literally, resulting in an unnatural subtitle that does not convey the interpersonal function of the original expression.

Similar problems also appear in subtitles containing expressions such as *deh*, *loh*, and *ya*, which play important roles in Indonesian conversational discourse. Although these particles contribute significantly to interpersonal meaning, the machine translation system is unable to adapt them appropriately into English.

The dominance of pragmatic shift demonstrates that conversational subtitle translation requires more than lexical correspondence. Successful subtitle translation depends on contextual interpretation, discourse awareness, and communicative adaptation. These findings are consistent with audiovisual translation theory, which emphasizes that subtitle quality should preserve both semantic meaning and communicative function rather than merely transferring linguistic forms (House, 2015; Chaume, 2018; Naveen & Trojovský, 2024; Castilho & Knowles, 2025).

Overall Discussion

Overall, the issue analysis demonstrates that CapCut Pro performs satisfactorily in translating straightforward lexical information but encounters greater difficulty when processing context-dependent and conversational Indonesian expressions. Among the identified issues, pragmatic shift appears most frequently, indicating that interpersonal meaning and discourse function remain major challenges for automatic subtitle translation. Meanwhile, lexical shift, literal translation, grammatical awkwardness, and semantic distortion further reveal that contextual

adaptation is still limited despite the system's ability to preserve general semantic content.

These findings indicate that CapCut Pro is effective in supporting rapid multilingual subtitle generation; however, its translation output should still be reviewed by human users before publication, particularly when the content contains informal conversation, cultural references, or context-dependent expressions. Therefore, the results reinforce the position that machine translation serves as an effective assistive technology rather than a complete replacement for human translation in audiovisual communication (Bowker, 2023; Moorkens et al., 2024; Al Amer et al., 2025).

CONCLUSION

This study examined the translation quality of CapCut Pro's auto-translated captions in Indonesian audiovisual content, specifically Instagram Reels from @fadiljaedi. The findings show that the feature is able to transfer the basic semantic meaning of the source utterances in many cases, especially when the expressions are short, direct, and structurally simple. However, the results also reveal that the generated subtitles are not consistently accurate, particularly when the source text contains compound or complex sentences, colloquial expressions, discourse particles, or culturally dependent meanings.

From the overall classification, most subtitle outputs were categorized as moderately accurate, followed by accurate and inaccurate translations. This distribution indicates that CapCut Pro performs reasonably well as a subtitle-generating tool, but its output still contains lexical shifts, grammatical awkwardness, literal translation, and pragmatic shifts. The most dominant issue found in the data was pragmatic shift, showing that the system still has difficulty interpreting conversational meaning, interpersonal tone, and contextual nuance in Indonesian audiovisual discourse.

Based on these findings, CapCut Pro's auto-translated captions can be considered effective as a supportive tool for improving multilingual accessibility in social media content. The feature is useful for producing quick subtitle drafts and helping content creators reach wider audiences. Nevertheless, the feature is not fully reliable as a final translation output. It still requires human checking, editing, and revision, especially when used for informal Indonesian speech that contains contextual, pragmatic, or idiomatic meaning.

Therefore, the study concludes that CapCut Pro's auto-translated captions are moderately effective, but they should not be used without review. Human revision remains necessary to ensure that the final subtitles are accurate, natural, and communicatively appropriate in the target language.

RECOMMENDATION

Based on the findings of this study, several recommendations can be proposed for different stakeholders. First, content creators who use CapCut Pro's auto-translated captions are encouraged to treat the generated subtitles as an initial draft rather than a final translation. Although the feature is capable of transferring the general meaning of Indonesian audiovisual content, manual proofreading remains

necessary to ensure lexical accuracy, grammatical naturalness, and pragmatic appropriateness, particularly in videos containing conversational expressions, discourse markers, and culturally specific utterances. Human revision is expected to improve subtitle readability and communicative effectiveness for international audiences.

Second, this study recommends that CapCut developers further improve the contextual and pragmatic capabilities of the auto-translation system. The findings indicate that the feature performs well in translating simple and direct utterances but still encounters difficulties when processing informal spoken Indonesian, compound or complex sentences, and expressions whose meanings depend on conversational context. Finally, future researchers are encouraged to expand this study by analyzing larger datasets, different social media platforms, or additional language pairs to obtain broader insights into the performance of AI-based subtitle translation.

REFERENCES

- Al Amer, S. A., Lee, M., & Smith, P. (2025, May). Comparative evaluation of machine translation models using human-translated social media posts as references: Human-translated datasets. In *Proceedings of the Eighth Workshop on Technologies for Machine Translation of Low-Resource Languages (LoResMT 2025)* (pp. 1–9). <https://doi.org/10.18653/v1/2025.loresmt-1.1>
- Amure, R., Agarwal, N., & Shaik, M. (2026). Examining linguistic strategies for digital mobilization in social movements on Instagram and TikTok. *Communications of the Association for Information Systems*, 58(1), 79.
- Baker, M. (2018). *In other words: A coursebook on translation*. Routledge.
- Bowker, L. (2023). Translation technologies: Translator training. In *Routledge encyclopedia of translation technology* (pp. 95–111). Routledge.
- Castilho, S., & Knowles, R. (2025). A survey of context in neural machine translation and its evaluation. *Natural Language Processing*, 31(4), 986–1016. <https://doi.org/10.1017/nlp.2024.7>
- Chaume, F. (2018). An overview of audiovisual translation: Four methodological turns in a mature discipline. *Journal of Audiovisual Translation*. <https://doi.org/10.47476/jat.v1i1.43>
- Dasola, A. T., & Olayimika, Y. S. (2026). The Impact of Short Video Contents on the Communicative Behaviour of Youths in Nigeria: An Empirical Study. *Bayero Journal of Library and Information Science*, 6(1), 423–435.
- Díaz-Cintas, J. (2020). Audiovisual translation. In *The Bloomsbury companion to language industry studies* (Vol. 10, pp. 209–230). Bloomsbury.
- Díaz-Cintas, J., & Remael, A. (2020). *Subtitling: Concepts and practices*. Routledge.
- House, J. (2015). *Translation quality assessment: Past and present*. Routledge.
- Hu, Z. (2021, May). A multimodal approach to translation quality assessment of interlingual subtitling: Theoretical reflections. In *7th International Conference on Humanities and Social Science Research (ICHSSR 2021)* (pp. 297–309). Atlantis Press. <https://doi.org/10.2991/assehr.k.210519.059>
- Jiménez-Crespo, M. A. (2024). *Localization in translation*. Routledge.
- Koehn, P. (2020). *Neural machine translation*. Cambridge University Press.

- Kristjansson, T., Huang, J., Lin, Y. K., Tyan, H. Y., Uszkoreit, J. D., Estelle, J. J., & Konishi, Y. (2015). *U.S. Patent No. 9,110,891*. U.S. Patent and Trademark Office.
- Moorkens, J., Castilho, S., Gaspari, F., & Doherty, S. (2018). Translation quality assessment. In *Machine translation: Technologies and applications* (Vol. 1, p. 299). Springer International Publishing.
- Moorkens, J., Way, A., & Lankford, S. (2024). *Automating translation*. Routledge.
- Mukti, M. A. P., & Fauzi, E. M. (2025). Lexical evolution in the age of artificial intelligence: A comparative literature study of English and Indonesian. *Humanities & Language: International Journal of Linguistics, Humanities, and Education*, 2(2), 98–104. <https://doi.org/10.32734/wt6fcr57>
- Mukti, M. A. P., Maulana, M. T., Rohmah, K. N., Abqoriyyah, F. H., & Saehu, A. (2024, June). The effectiveness of artificial intelligence usage as translation medium among English literature students of UIN Sunan Gunung Djati Bandung. In *Journal of English Education Forum (JEEF)* (Vol. 4, No. 2, pp. 107–112). <https://doi.org/10.29303/jeeef.v4i2.683>
- Nababan, M. R., Nuraeni, A., & Sumardiono. (2012). *Pengembangan model penilaian kualitas terjemahan*. *Kajian Linguistik dan Sastra*, 24(1), 39–57.
- Naveen, P., & Trojovský, P. (2024). Overview and challenges of machine translation for contextually appropriate translations. *iScience*, 27(10). <https://doi.org/10.1016/j.isci.2024.110878>
- Nicolaou, C. (2021). Media trends and prospects in educational activities and techniques for online learning and teaching through television content: Technological and digital socio-cultural environment, generations, and audiovisual media communications in education. *Education Sciences*, 11(11), 685. <https://doi.org/10.3390/educsci11110685>
- Papi, S., Gaido, M., Karakanta, A., Cettolo, M., Negri, M., & Turchi, M. (2023). Direct speech translation for automatic subtitling. *Transactions of the Association for Computational Linguistics*, 11, 1355–1376. https://doi.org/10.1162/tacl_a_00607
- Popović, M. (2018). Error classification and analysis for machine translation quality assessment. In *Translation quality assessment: From principles to practice* (pp. 129–158). Springer International Publishing.
- Romero-Fresco, P., & Pöchhacker, F. (2017). Quality assessment in interlingual live subtitling: The NTR model. *Linguistica Antverpiensia, New Series – Themes in Translation Studies*, 16, 149–167. <https://doi.org/10.52034/lanstts.v16i0.438>
- Serinastity, G., Mubarak, M. Z., Saehudin, A., Suparno, D., & Fathalah, F. M. (2025). Measuring the technique and quality of Instagram caption auto-translation to improve @alarabiya account translation results. *Al-Ta'rib: Jurnal Ilmiah Program Studi Pendidikan Bahasa Arab IAIN Palangka Raya*, 13(2), 351–370. <https://doi.org/10.23971/altarib.v13i2.10410>
- Shafira, D. B., & Nugroho, R. A. (2023, December). Translation error types analysis on TikTok Indonesian-English auto-translation content. In *International Seminar SEMANTIKS & PRASASTI 2023 Theme: Language in the Workplace (PRASASTI 2023)* (pp. 45–51). Atlantis Press. https://doi.org/10.2991/978-2-38476-162-3_7
- Stahlberg, F. (2020). Neural machine translation: A review. *Journal of Artificial Intelligence Research*, 69, 343–418. <https://doi.org/10.1613/jair.1.12007>