

CHAPTER I

INTRODUCTION

1.1 Background of the Study

Translation is a complex process that involves converting a message from one language to another while ensuring the accurate conveyance of meaning, structure, and cultural nuances. It plays a vital role in today's globalized society, facilitating communication across diverse contexts, enhancing international business relations, and supporting language development in various fields (Saha, 2020). However, translation is not merely about replacing words; it necessitates a deep understanding of context, cultural significance, and the intricate relationship between language and meaning. Some words or phrases carry cultural significance that may not translate directly into another language. Ignoring these nuances can lead to translations that fail to convey the original content's depth, thus altering the audience's perception (Vilar et al., 2006). Lexical errors often arise when individual words or expressions are inaccurately translated, leading to misunderstandings, loss of intended meaning, and the potential for misinformation. Such inaccuracies can severely compromise the clarity of communication, particularly when critical concepts are distorted due to literal translations, lack of contextual awareness, or cultural disparities.

Lexical errors frequently occur in MT systems, stemming from the tools' inability to fully grasp the nuances of the source text or to appropriately adjust their

translations to fit the target language (Brglez & Vintar, 2022). This challenge is amplified in environments characterized by concise messages and limited context, such as Instagram captions.

Social media platforms like Instagram have emerged as pivotal spaces for information consumption and engagement with real-time events. Instagram captions serve a crucial function by delivering concise updates, opinions, and narratives to users. However, due to their brevity and the informal nature of social media language, including embedded cultural references, these captions present unique challenges for MT systems. For instance, the same word or phrase may possess multiple meanings depending on the situation, complicating the machine's ability to select the appropriate translation. Consequently, translation errors in social media can distort the intended message, contributing to the proliferation of misinformation.

Instagram's translation feature is widely utilized to enable non-English-speaking audiences to access global content, particularly during significant international events. Given the rapid pace of social media, where news spreads quickly, the importance of accurate translation cannot be overstated. Captions summarizing major events require precision, as even minor errors can significantly alter public perception and skew the interpretation of critical developments (Bielsa, 2009). Translation from the data source, which is The New York Times (NYTIMES) Instagram account, the situation observed reflects the global and highly sensitive nature of political events covered on the platform. The captions

posted on this account during December 2024, particularly regarding the Gaza-Palestine conflict and other political issues, offer a snapshot of how global events are presented and consumed on social media.

The phenomena that visible in the data source demonstrate the challenges of translating content that is both politically and socially charged can be found in The New York Times instagram account. The NYTIMES captions often summarize complex events in concise text, which, when machine-translated, may suffer from lexical errors such as miswordings or loss of nuance. This is especially concerning when discussing topics that have significant political ramifications, as any mistranslation could lead to misunderstandings or misinterpretations that affect public discourse. Given that the NYTIMES has millions of global followers, including many Indonesian speakers, ensuring the accuracy of these translations is crucial to maintaining trust and promoting informed, meaningful discussions on these issues.

The observed phenomenon in this context is the reliance on machine translation systems that may not fully capture the intricacies and sensitivities of such content. This highlights the need for more advanced translation tools that can handle the complexities of political and cultural language, ensuring that key messages remain intact across languages and are properly understood by audiences from different backgrounds. Misleading translations can distort the intended message, influencing public understanding and potentially affecting social and political opinions (Watts et al., 2021). For example, captions regarding sensitive

topics such as the Gaza conflict necessitate careful wording; translation errors could lead to biased or inaccurate narratives.

This research is crucial as it highlights the importance of linguistic precision in social media communications, especially within multilingual contexts. By identifying and analyzing translation errors from English to Indonesian, this study seeks to contribute to enhancing machine translation tools, ultimately fostering clearer and more accurate communication in global digital spaces.

1.2 Problems of the Study

Based on the background above, the problems of this study were:

1. What are the types of lexical error produced by Instagram's MT of captions from NYTIMES Instagram account?
2. What are the translation techniques used so that the translation of NYTIMES Instagram caption is more acceptable?

1.3 Objectives of the Study

The focus of the study was based on the research questions thus the research objectives were organized as follows:

1. To categorize the types of lexical errors produced by Instagram's MT of captions from NYTIMES Instagram account based on the theory from Vilar et al.
2. To analysing the translation technique used so the translation of caption from NYTIMES Instagram account is more acceptable.

1.4 Limitation of the Study

The scope of this study was specifically limited to analysing the lexical errors produced by Instagram's machine translation of captions from The New York Times (NYTIMES) account, focusing on the implications of these errors for communication and understanding using the theory proposed by Villar et al. (2006). The data source was confined to selected Instagram captions related to significant global political events occur in December 2024, such as the Gaza-Palestine conflict and Presidential. This temporal and contextual limitation ensures that the research remains relevant and manageable. The study also aimed to identify the translation techniques employed to improve the acceptability of these captions, concentrating on strategies that enhance the accuracy and clarity of machine translations using theory from Darbelnet (1989). By focusing on these elements, the research sought to highlight the importance of effective translation practices in social media contexts, providing valuable insights into the challenges posed by machine translation systems in conveying nuanced information.

1.5 Significance of the Study

This research was expected can be beneficial for various reasons such as for academics or community purposes. The categories of significance are having two expected, namely theoretical significance and practical significance.

1.5.1 Theoretical Significance

In theoretical terms, this study enhances the understanding of lexical errors in machine translation by applying a structured framework to analyse how

inaccuracies affect meaning and communication, particularly in digital contexts. It emphasizes the importance of contextual awareness and linguistic nuances in translation processes. By addressing the challenges of machine translation in social media environments, this research aims to provide valuable insights that can inform future developments in translation studies and computational linguistics.

1.5.2 Practical Significance

In practical terms, advances in machine error translation have significance for many sectors in the future. By delving into the intricacies of language, machine translation, and error categorization, this effort promises invaluable insights for educators, professionals, and researchers, especially for students at the Faculty of Foreign Languages, Universitas Mahasaraswati Denpasar. Specifically, the findings obtained from this exploration will provide a foundation for improving translation technologies, refining educational curricula, and fostering cross-disciplinary collaboration. Moreover, the results of this research will not only enrich the knowledge base among academia but also empower industries that depend on accurate translation services, paving the way for smoother communication across the diverse linguistic landscapes of an increasingly interconnected world.

CHAPTER II

REVIEW OF RELATED LITERATURE, CONCEPTS, AND THEORIES

2.1 Review of Related Literature

This section presents previous studies relevant to this research to provide a comparative view and support the analysis. It highlights how prior works align with or differ from the current research objectives. Five studies have been reviewed to obtain insights into the topic of lexical errors, machine translation, and caption analysis.

The first study was a thesis by Rohmani (2021), titled *Common Lexical Errors Made by Machine Translation in Cultural Texts*. This research focuses on identifying and analyzing lexical errors in machine translation using Carl's (1998) framework. By analyzing 553 pages from Arok Dedes selected through a simple random sampling technique, the study identified nine types of lexical errors from a total of 21 possible types. These errors include calque, misselection, false friends, inappropriate co-hyponyms, and semantically determined word choices. The results revealed that calque was the most dominant type of error. The study employs a qualitative descriptive method, making it relevant to the present research, as both explore lexical errors through qualitative analysis. However, while Rohmani's study concentrates on cultural texts, the present research extends its scope to Instagram captions related to current events.

The second study was a journal article by Nadilla (2023), titled *Lexical Errors Made by Instagram Machine Translation in Translating the Account of*

“CNN Indonesia” News Article. This research aims to identify the types of lexical errors produced by Instagram machine translation. Analyzing various captions, the study found four instances of missing words, ten incorrect words, and eight unknown words based on Vilar et al.'s (2006) error categories. The results indicate that Instagram machine translation frequently fails to accurately convey the target language. Employing a qualitative method, this study is relevant as it parallels the present research's focus on lexical errors in machine translations of social media captions. However, unlike the present study, which examines the NYTIMES's captions covering complex world events, Nadilla's research focuses on CNN Indonesia's captions, which are primarily news updates targeted at local Indonesian audiences.

The third was a thesis conducted by Carissa (2021), titled *The Analysis of Lexical Errors Produced by Fikinakii's Instagram Account Translated by Instagram Translation Machine*. This research investigates the lexical errors in translations from Bahasa Indonesia to English on the Fikinakii Instagram account, using Instagram's translation tool. The study applies Vilar et al.'s (2006) framework, which categorizes errors into incorrect words, missing words, word order issues, unknown words, and pronunciation errors. By employing a descriptive qualitative method, the researcher analyzes various posts to identify the types and forms of lexical errors present. The findings highlight the specific forms of errors generated by the translation tool, providing valuable insights into the limitations of machine translation in social media contexts. This study is relevant to the current research as it focuses on similar issues of lexical errors in Instagram translations, extending the

investigation to captions related to significant events. The findings highlight the specific forms of errors generated by the translation tool, providing valuable insights into the limitations of machine translation in social media contexts. While Carissa's research focuses on influencer-generated content with informal and conversational language, the current study investigates machine translation errors in journalistic captions from the NYTIMES Instagram Account, which typically use formal and precise language to convey nuanced information.

The fourth literature was an article by Prasetio (2023), titled *An Analysis of the Error Translation in Movie Trailers by YouTube Auto-Translate*. This research investigates the errors made by YouTube's Auto-Translate feature in translating movie trailers into Indonesian. The study compares these translations with those produced by professional translators and categorizes the errors using Vilar et al.'s error classification framework. The authors identified 14 distinct errors in the Auto-Translate output, with lexical errors occurring most frequently (63%). Other error types included disambiguation errors (27%), while word order and unknown words were identified only once each (5%). The findings emphasize that while machine translation can assist with basic word lookups, it often fails to capture context, highlighting the necessity for human post-editing when using MT tools. This study is relevant to the current research as it similarly examines the limitations of machine translation in different contexts, offering insights into the nature of lexical errors. Unlike Prasetio's focus on movie trailers, which are typically promotional and entertainment-oriented, the present study examines news captions from the New York Times. This shift in focus is significant because the accuracy of news-related

content is crucial for conveying complex events, making errors potentially more impactful than in entertainment contexts.

The fifth study was an article by Utami et al. (2021) titled Lexical Analysis of Semantic Errors Found in the Translation of Joko Widodo's Instagram Account. This research focuses on semantic errors in lexis resulting from Instagram Machine Translation of posts on Joko Widodo's Instagram account. The aim of the study is to identify and categorize the types of semantic errors present in these translations using James's (1998) framework for semantic error analysis. The researchers found that the most common issue was the confusion of sense relations, particularly in the selection of incorrect synonyms. The study employs a descriptive qualitative method, making it relevant to the present research, which also investigates the effectiveness of machine translation in rendering accurate lexical meanings in social media contexts. However, while Utami et al. concentrate on semantic errors, the current research extends to lexical errors, offering a broader perspective on the challenges faced by machine translation systems.

Finally, Vilar et al. (2006) offer a direct framework for error classification, which is extensively used in this study. Their work, *Lexical Errors in Machine Translation*, categorizes errors into omission, addition, and incorrect word use. The framework forms the theoretical foundation for identifying and classifying lexical errors in the current research.

The similarities between these studies and the present research lie in the focus on lexical error identification and the use of qualitative methods. However,

key differences also emerge. The current study uniquely focuses on Instagram captions from The New York Times account during major events, providing a novel context that has not been explored in previous research. Additionally, it offers a more specific exploration of translation techniques (borrowing, calque, adaptation) and their impact on meaning in machine-translated social media content.

2.2 Concepts

Concepts are provided to offer clarity about the key ideas used throughout the research. By defining these concepts, readers can align their understanding with that of the researcher, ensuring consistent interpretation. The definitions and explanations are drawn from relevant literature and expert sources, creating a solid foundation for the analysis presented in this study. Below are the core concepts applied in this research

2.2.1 Translation

Translation involves converting text or spoken words from one language to another while maintaining the original meaning and context. It requires a deep understanding of both the source and target languages, as well as cultural nuances. Effective translation is crucial in facilitating communication and promoting mutual understanding across different languages (Bassnett, 2014). Newmark (1988) explains that translation is not merely a linguistic conversion but also a process of transferring the intent and style of the source text. Nida (1964) highlights the concept of dynamic equivalence, which focuses on ensuring the translated text elicits the same response in the target audience as the original text did in the source

audience. Venuti (1995) discusses the translator's role in bridging cultural gaps and stresses the importance of balancing transparency and foreignization to align with audience expectations. These expert perspectives illustrate the complexity and significance of translation in fostering communication and mutual understanding across linguistic and cultural boundaries.

2.2.2 Machine Translation

Machine translation (MT) is the automated process of translating text or speech from one language to another using computer software. MT systems leverage algorithms and large databases of language data to produce translations quickly. While MT has made significant advancements, it often struggles with context, idiomatic expressions, and cultural references, which can lead to inaccuracies (Koehn, 2009).

2.2.3 Instagram Caption

Instagram captions are short texts accompanying images or videos on the platform, providing context, commentary, or additional information (Baker & Walsh, 2018). These captions often use informal language and may include slang, emojis, and cultural references, making them challenging for translation systems. The brevity and context-dependent nature of captions can result in translation errors, particularly when using machine translation tools.

2.3 Theories

This study employs established theories from the fields of translation studies and machine translation as analytical frameworks. These theories guide the identification, classification, and analysis of lexical errors encountered in the translation of Instagram captions. Below are the relevant theories applied in this research:

2.3.1 Translation Error

Translation errors are mistakes that occur during the translation process, which can arise from misunderstanding the source text, misapplying translation rules, or failing to consider the cultural context (House, 2015). Errors can lead to significant miscommunication and may distort the intended message. Vilar et al. (2006) do not provide a singular, explicit definition of "translation error" but instead approach the concept by categorizing errors in machine translation. Their framework breaks down errors into several types to facilitate analysis and improvement of translation systems. These include missing words, incorrect words, word order errors, and unknown words, among others. This categorization implies that a translation error is any deviation from the intended meaning, structure, or linguistic norms of the source language when rendered into the target language

2.3.2 Lexical Error

Lexical errors refer to inaccuracies in the translation of individual words or phrases, often resulting from a lack of understanding of their meanings or context.

Vilar et al. (2006) categorize lexical errors in machine translation into various types, including missing words, untranslated terms, and incorrect word choices. These errors can compromise the overall clarity and effectiveness of the translated text. From the study by Nadilla & Lubis (2023), the following examples illustrate common lexical errors identified in Instagram machine translation:

1. Missing Word

This happens when a necessary word or phrase is omitted during translation, which can alter the intended meaning of the original text. For example, the absence of an article or verb can lead to incomplete sentences that confuse the reader.

Table 2.3.2.1. Example of Missing Word

Source Language (SL)	Target Language (TL)
<i>Thailand menang telak 4-0 atas Timnas Indonesia di leg pertama final Piala AFF 2021, Rabu (29/12)</i>	Thailand won 4-0 against Indonesia National Team in the first leg of (the) AFF Cup 2021 final, Wednesday (29/12)

The definite article “the” before "AFF Cup 2021 final" is missing. While articles may seem minor, they play an essential role in English grammar to clarify specific nouns. The omission of “the” can confuse the reader, implying an incomplete or incorrect reference to the event. (Nadilla & Lubis, 2023:42)

2. Incorrect Word

Sometimes, words or phrases are left untranslated, particularly if the translator assumes the audience understands them or if the words are context-specific. This can create gaps in understanding, especially for readers unfamiliar with the source language. For the example:

Table 2.3.2.2. Example of Incorrect Word

Source Language (SL)	Target Language (TL)
<i>Mari, bantu saudara- saudara kita yang terdampak erupsi Gunung Semeru melalui platform berbuatbaik.id</i>	Come, help our brothers and sisters affected by the eruption of Mount Semeru through the platform of doinggood.id.

“berbuatbaik.id” is left untranslated as “doinggood.id.” Although it reflects the literal meaning of the original term, it disrupts the natural flow and may confuse readers (Nadilla & Lubis, 2023:39)

3. Unknown or Untranslated Word

This type of error occurs when a word is translated into an incorrect synonym or a word with a different connotation, leading to a misrepresentation of the original message. Such inaccuracies can arise from a lack of contextual awareness or cultural differences that affect meaning. For the example:

Table 2.3.2.3. Example of Untranslated Word

Source Language (SL)	Target Language (TL)
<i>Berdasarkan data dari komnas perempuan per 2015-2020, 27 kasus kekerasan seksual terjadi di perguruan tinggi.</i>	According to data from the women’s comnas for 2015-2020, 27 cases of sexual violence occurred in collages.

The word “komnas” was left untranslated and improperly rephrased as “comnas” without clarification that it refers to Komisi Nasional Perempuan (National Commission on Violence against Women). Additionally, “perguruan tinggi” was incorrectly translated as “collages” instead of the intended meaning, which is “universities” or “higher education institutions.” This mistake alters the meaning and risks confusing readers unfamiliar with the original context (Nadilla & Lubis, 2023:41)

2.3.3 Translation Techniques

Translation techniques refer to the strategies that translators use to convey meaning accurately and effectively in the target language. These techniques address the linguistic and cultural differences that arise during translation. According to Vinay & Darbelnet (1989), there are seven main translation methods: borrowing, calque, literal translation, transposition, modulation, equivalence, and adaptation. which are categorized into two groups: direct translation and oblique translation.

2.3.3.1 Direct translation techniques

1. Borrowing involves taking a word directly from the source language. For example, the word “format” remains unchanged in Indonesian as “format,” a strategy employed when there is no suitable equivalent (Ananda & Nababan, 2019:38).
2. Calque translates individual components of an expression literally. For example, the English term skyscraper is translated into French as *gratte-ciel*—a direct but adapted rendering (Vinay & Darbelnet, 1995: 32).
3. Literal Translation means translating word-for-word while following grammatical rules. In the study, the English phrase "I feel tired" is translated directly into "Saya merasa lelah" without losing grammatical accuracy (Ananda & Nababan, 2019:39).

2.3.3.2 Oblique translation

1. Transposition alters the grammatical structure but keeps the meaning intact. For instance, the English phrase he decided to go becomes *il a pris la*

décision d'y aller (literally: "he took the decision to go"), demonstrating a change from a verb to a noun phrase (Vinay & Darbelnet, 1995: 36).

2. Modulation changes the point of view or form of expression to better suit the target language's context. An example is "We're leaving at the first sign of trouble, and I'm not going to debate it with you," which is translated as "*Kita pergi begitu ada pertanda bahaya, tanpa banyak tanya.*" This modulation makes the expression more direct and clearer in Indonesian. (Ananda & Nababan, 2019:37).
3. Equivalence replaces an idiomatic expression with an equivalent phrase in the target language. An example is translating out of sight, out of mind into French as *loin des yeux, loin du cœur* ("far from the eyes, far from the heart")—capturing the same emotional resonance (Vinay & Darbelnet, 1995: 38).
4. Adaptation adjusts cultural references for better relevance. For example, a French reference to *jeu de boules* (a local sport) might be replaced with cricket when translated into English for a British audience (Vinay & Darbelnet, 1995: 39).
5. Deletion, technique involves omitting words or phrases from the source text that may be unnecessary or redundant in the target language. For example, in the source text "Really, you like the idea of Ohio?" the translation is simplified to "*Oh, ya? Kau suka Ohio?*" Here, the word "idea" is deleted for clarity and conciseness in the target language (Ananda & Nababan, 2019:36).