

CHAPTER I: INTRODUCTION

1.1 Background of the Study

In the contemporary digital era, the proliferation of information and communication technologies has revolutionized various aspects of human life, including language communication. As a result, the demand for quick and accessible translation tools has significantly increased. One of the most widely used and easily accessible machine translation tools is Google Translate. It offers instant translation services for a vast number of languages, including English and Bahasa Indonesia. The convenience and speed it provides have made it a go-to tool for students, researchers, travelers, and general users who need to understand or produce text in a foreign language.

However, most existing studies tend to evaluate GT's performance within a single genre, such as academic or narrative texts only, without comparing how linguistic variation across genres influences translation accuracy. In fact, each genre carries distinct linguistic structures and communicative functions, which may significantly affect machine translation performance.

In the contemporary digital era, the advancement of information technology has transformed how humans communicate and access multilingual content. Among various tools available, Google Translate (GT) has become a dominant Neural Machine Translation (NMT) platform due to its convenience and accessibility. However, its performance across different text types remains inconsistent. While GT performs reasonably well in literal and straightforward sentences, it often fails to maintain contextual accuracy, stylistic equivalence, and pragmatic meaning in more complex or genre-specific texts.

Such linguistic divergence among genres is expected to trigger different types and frequencies of translation errors in GT outputs, ranging from lexical mistranslations to syntactic distortion or loss of pragmatic meaning. Moreover, translation performance may also differ depending on translation direction (English–Indonesian vs. Indonesian–English), as previous research suggests that NMT systems are often biased toward one dominant language pair.

Several prior studies (Hidayah & Setiyono, 2020; Indrayani & Indriani, 2021; Wulandari et al., 2022) have shown that GT's accuracy varies significantly depending on the genre and linguistic complexity of the text. However, few studies have compared the translation accuracy across multiple text types particularly academic, narrative, and journalistic between English–Bahasa Indonesia and Bahasa Indonesia English directions. This constitutes a research gap that the current study aims to fill.

These three genres were deliberately selected because each present distinct linguistic and functional characteristic:

1. Academic texts emphasize formality, technical terminology, and syntactic precision.
2. Narrative texts rely heavily on figurative expressions, cohesion, and literary style.
3. Journalistic texts prioritize conciseness, factuality, and socio-cultural tone.

These differences are expected to affect GT's translation accuracy differently. Therefore, this study conducts a comparative analysis to examine how genre variation influences translation accuracy, clarity, and naturalness. By doing so, this research contributes to a deeper understanding of NMT limitations and linguistic variation in translation performance.

This study aims to investigate the accuracy of Google Translate in translating texts between English and Bahasa Indonesia, and vice versa. By analyzing its performance on different types of texts, this research seeks to provide a comprehensive understanding of its

strengths and weaknesses. The findings of this study will be beneficial for language learners, educators, and anyone who uses this tool, helping them to be more critical and aware of its limitations.

1.2 Identification and Scope of the Problems

Based on the background outlined above, this study aims to answer the following research questions:

- 1.2.1 How accurate is Google Translate in translating academic, narrative, and journalistic texts between English and Bahasa Indonesia?
- 1.2.2 What types of translation errors occur most frequently in each genre?
- 1.2.3 How do error types and accuracy levels differ among the three genres across translation directions (EN-ID vs. ID-EN)?

1.3 Objective of the Study

This study has the following objectives:

- 1.3.1 To evaluate the accuracy level of Google Translate in translating different genres of texts between English and Bahasa Indonesia.
- 1.3.2 To identify and classify common types of errors in the translations.
- 1.3.3 To compare and analyze how linguistic and stylistic variations among text types influence the translation accuracy, clarity, and naturalness of Google Translate output.

1.4 Significance of the Study

The findings of this research are expected to have both theoretical and practical significance:

1.4.1 Theoretical Significance

This study enriches translation studies by applying error taxonomy (Nababan et al., 2012; House, 1997; Larson, 1984) and translation quality assessment (TQA) to evaluate NMT performance. It also contributes to discussions on equivalence, accuracy, and genre-sensitive translation, offering empirical data for the development of machine translation evaluation frameworks.

1.4.2 Practical Significance

- a. Language learners and educators can use the findings to better understand GT's strengths and weaknesses in specific text genres.
- b. Researchers can use the framework as a model for comparative translation

analysis.

- c. General users can be more aware of GT's limitations in genre-specific contexts.

1.5 Scope and Limitation of the Study

This research focuses on analyzing the accuracy of Google Translate's performance in two specific language pairs: English-Bahasa Indonesia and Bahasa Indonesia-English. The study will utilize a limited number of texts from different genres (e.g., academic, narrative, and journalistic) as a sample. The analysis of accuracy will be conducted by a human rater (the researcher) based on a pre-defined rubric or scoring criteria. This study does not aim to compare Google Translate with other machine translation tools or to delve into the technical mechanisms or algorithms behind its operation.

CHAPTER II: LITERATURE REVIEW

2.1 Theoretical Background

1. Definition of Translation

Theoretical debates on translation have evolved from a mere replacement of structurally based textual material (Catford, 1965) toward the reconstruction of meaning that takes style and context into account (Nida & Taber, 1969; Larson, 1984). Critically, while structural approaches offer clear linguistic parameters, Larson's functional approach is far more relevant to this study because it operationalizes evaluation into three concrete dimensions: accuracy, clarity, and fluency.

An interdisciplinary synthesis of perspectives (Baker, 2001) underscores that translation is not merely a mechanical process but a complex interaction between culture and genre. Focusing on written translation in this study enables a controlled analysis of Google Translate's performance across academic, narrative, and journalistic genres. Through this comparison, it becomes possible to assess the extent to which translation quality transcends lexical equivalence toward reader-oriented functional adequacy.

2. Machine Translation

The evolution of machine translation (MT) systems from rule-based approaches to Statistical Machine Translation (SMT), and more recently to Neural Machine Translation (NMT), signifies a shift toward holistic sentence-level modeling (Koehn, 2021; Popović, 2020). However, from a critical perspective, the sophistication of NMT does not necessarily guarantee stylistic quality comparable to that of human translation. Although NMT excels in contextual fluency at the sentence level, it often produces outputs that are emotionally flat or fails to capture genre-specific registers.

These limitations in handling idiomatic expressions and cultural nuances reinforce the urgency of employing Larson's multidimensional evaluation framework to assess NMT performance objectively. By dissecting the operational mechanisms of NMT, this study seeks to demonstrate that high lexical-level accuracy does not necessarily correspond to stylistic naturalness. This premise provides the foundation for examining the extent to which Google Translate is capable of bridging structural differences between English and Indonesian across diverse communicative contexts.

2.1. Google Translate

As a leading NMT system, Google Translate has demonstrated significant improvements in contextual fluency since its transition to deep learning in 2016. Nevertheless, critical analyses of this system frequently reveal a dichotomy between semantic adequacy and

syntactic naturalness. Its reliance on massive corpus data occasionally results in awkward structures in genres that demand specific tonal qualities, such as journalistic objectivity or academic formality.

This phenomenon illustrates that Google Translate continues to operate within the constraints of statistical probability, which may fail to reproduce the communicative effects intended by the original author. By evaluating its performance across three distinct genres, this study aims to determine whether the strengths of contemporary NMT are universal in nature or highly contingent upon the rhetorical characteristics of the source text. The findings are expected to offer recommendations regarding the limits of machine translation effectiveness in context-sensitive applications.

2.2. Accuracy in Translation

Accuracy constitutes a primary indicator of meaning fidelity between the source text and the target text, encompassing lexical, grammatical, and semantic dimensions. From an analytical standpoint, accuracy must be systematically examined to identify the points at which meaning distortion occurs, particularly in domain-specific texts. It is important to note that high informational accuracy does not necessarily guarantee functional quality if the target text fails to achieve readability or acceptability.

Therefore, accuracy evaluation in this study is not treated as an isolated element but is understood to interact dynamically with the parameters of clarity and fluency. The use of an evaluation rubric adapted from Larson (1984) and Nababan et al. (2012) enables a more comprehensive assessment of Google Translate's output. This approach ensures that the analysis focuses not only on what is translated, but also on how the message is conveyed within the cultural and linguistic context of the target language.

3. Translation Quality Assessment (TQA)

The TQA model adopted in this study synthesizes parameters from Larson (1984) and Nababan et al. (2012) to construct an objective assessment instrument. The use of a multidimensional approach is crucial, as minor errors in a single aspect—such as lexical choice—can significantly alter reader perception and overall message interpretation. TQA thus functions as a bridge that transforms descriptive qualitative analysis into measurable and scientifically accountable data.

Analytically, the application of TQA enables researchers to identify patterns of machine translation failure beyond mere grammatical errors. The emphasis on accuracy as a key indicator of success within the TQA framework provides insight into the extent to which message integrity is preserved amid the complexity of differing linguistic structures. Through this model, the study offers a balanced evaluation of the strengths and limitations of NMT technology in handling genre variation.

4. Error Analysis Models

Error analysis functions as a diagnostic instrument for systematically categorizing translation failures based on the taxonomy proposed by Dulay, Burt, and Krashen (1982). By classifying errors into omission, addition, misformation, and misordering, researchers are able to map recurring patterns of systemic failure within NMT systems. The integration of this model into the TQA framework provides a

deeper analytical layer regarding the technical nature of translation quality degradation.

Overall, an understanding of error types and their implications facilitates the provision of specific feedback for the advancement of automatic translation technologies. This approach not only identifies the presence of errors but also explores the underlying reasons behind the machine’s failure to process particular linguistic units. Consequently, the integration of error analysis models strengthens rigorous evaluation practices and supports the improvement of cross-linguistic communication quality in digital media.

2.2 Previous Related Studies

Table 2. 1 Table of Previous Related Studies

Researcher	year	focus	Findings
Hidayah &Setiyono	2020	Accuracy of Google Translate in academic and narrative texts	Google Translate is more accurate in translating academic texts but encounters difficulties with narrative texts containing idiomatic expressions.
Indrayani &Indriani	2021	Errors in Google Translate for journalistic texts	Google Translate maintains lexical and syntactic accuracy but fails to preserve objectivity and tone in journalistic texts.
Wulandari et al.	2022	Use of Google Translate in translating literary texts	Google Translate faces challenges related to cultural nuances and idiomatic expressions, resulting in unnatural translations.
Koehn	2021	Comparison between NMT and SMT	NMT improves contextual fluency but continues to struggle with idiomatic expressions and domain-specific terminology
Popovic	2020	Performance of NMT systems such as Google Translate	NMT performs better in direct and technical translations but has difficulties with complex genres such as narrative and literary texts.

Therefore, this study differs from previous research in two key aspects. First, it uses a multi-genre comparative approach (academic, narrative, and journalistic), allowing for cross-genre evaluation rather than separate analyses. Second, it employs a bidirectional assessment (EN–ID and ID–EN) to examine whether translation direction impacts accuracy differently depending on genre. By combining Translation Quality Assessment (TQA) and the Error Analysis taxonomy within a genre-based comparative framework, this study presents a more detailed, linguistically grounded evaluation model, assessing both functional equivalence and error typology

simultaneously.

While previous studies have evaluated Google Translate's performance, most have been limited to a single genre, such as academic writing (Hidayah & Setiyono, 2020; Indrayani & Indriani, 2021), narrative texts (Lestari & Kurniawan, 2022), or journalistic reports (Sari & Putra, 2020). Some studies have even examined only English-to-Indonesian translations, without assessing the reverse direction, which is crucial for academic publications and international communication.

Furthermore, no research has conducted a systematic comparison across genres and translation directions using the combined framework of TQA (House, 1997; Nababan et al., 2012) and Error Analysis (Dulay et al., 1982). This creates a methodological gap, as genre-based linguistic variation for example, nominalization in academic discourse, metaphorical expressions in narrative writing, or formulaic collocations in journalistic texts has not been thoroughly examined in relation to machine translation accuracy.

2.3 Theoretical Framework

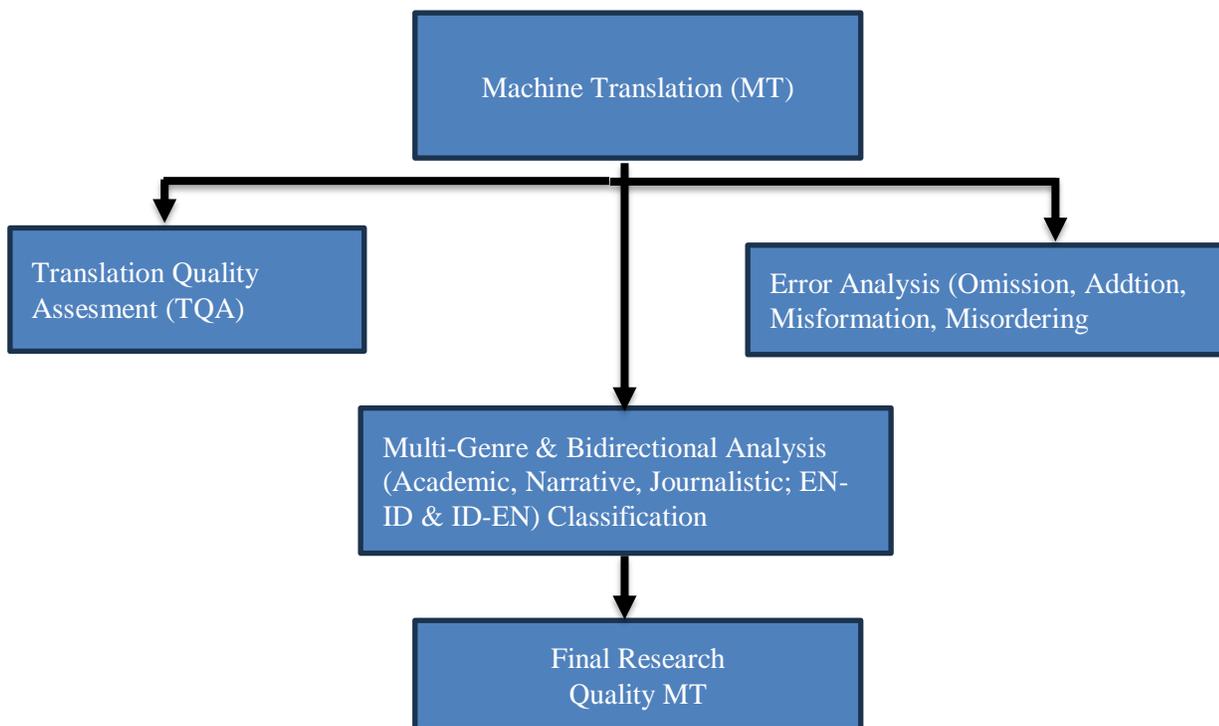


Figure 3.1 Theoretical Framework