

Understanding Google Translate's Limitations in English – Chinese Translation: A Linguistic Error Analysis Approach

R. D. A. E. Jayasundara¹ , R. P. Yaddhige²

NSBM Green University

ashani.j@nsbm.ac.lk¹ , rivinduy@masholdings.com²

Abstract

Machine Translation (MT) tools such as Google Translate (GT) are widely used by language learners for translating texts between Chinese and English languages, underscoring the need for systematic evaluation of MT output accuracy. This study examines the quality of GT by conducting an error analysis of Chinese-to-English and English-to-Chinese translations across texts of varying genres and lengths. The research was motivated by the growing perception that GT has improved in recent years, raising the question of whether learners can rely on GT for translation accuracy without the guidance of a language instructor. To contextualise learner engagement with GT, a preliminary questionnaire was administered to beginner-level undergraduates studying Chinese as a Foreign Language at a private higher education institution in Sri Lanka to examine usage patterns and reliance on the tool. Using a qualitative thematic analysis approach, the study identifies several linguistic errors produced by GT. The findings reveal persistent limitations in GT's handling of linguistic, semantic, and contextual features, thereby offering insights into its reliability and pedagogical implications for Chinese–English translation in language learning settings.

Keywords: *Error Analysis, Google Translate, Language Error, Machine Translation, Source Language, Target Language.*

1. Introduction

This study originates from the observation that Machine Translation (MT) systems, notably Google Translate (GT), have advanced considerably. Due to globalization and the increasing need for cross-language communication, the demand for accurate Chinese–English translation has grown significantly. Translation between Chinese and English requires more than the simple transfer of words; it involves conveying contextual meaning, idiomatic expressions, and pragmatic nuances unique to both languages. Achieving an accurate translation requires time, competence, and strong linguistic knowledge in both the Source Language and the Target Language. Chinese and English differ greatly in grammar, structure, and cultural expressions, and therefore producing an accurate translation is challenging, and linguistic errors are likely to occur in the process. With the rapid advancement of technology, many learners and users now rely on Machine Translation tools such as Google Translate to handle large volumes of Chinese–English translation quickly. While MT provides efficiency and convenience, its output must still be examined for linguistic accuracy to ensure that meaning is preserved without distortion.

2. Literature Review

Machine Translation refers to the use of rule-based systems or machine learning algorithms to translate text or speech from one language to another. A number of MT engines are widely used today, including Google Translate, DeepL, Baidu Translate, Microsoft Translator, and Amazon Translate. Among these, Google Translate has become one of the most commonly used tools for bridging communication gaps, especially in translating between English and Chinese for academic, social, and professional purposes. In the Sri Lankan context, Google Translate is frequently used by undergraduates who rely on it to understand unfamiliar English or Chinese words and phrases, particularly those studying Chinese as a foreign language. Those who want to know the translation of some phrases, or students who are learning foreign language, probably use Google translate to fulfil their needs of information (Hilma, 2011).

Research by Aslerasouli & Abbasian (2015) on undergraduates' translation of texts using GT compare both HT and MT to investigate their qualities where the findings show that there is a significant difference in the quality of HT and MT and the HT's quality was still higher than that of MT. Study by Anggaira (2017) on high school students' translations from English to Indonesian analyze the aspects of language errors that appear on GT of which the results indicate that GT translates word by word, but overlooks sentence context, resulting in morphological, syntactic, and semantic errors. Anggaira (2017) asserts that human improvements to the GT output are required to get translation accuracy.

2.1. Translation

Translation is basically change of form. The original form of the content being translated is generally called the Source Language, and the form into which it is to be changed is referred as the Receptor Language or Target Language. Translation, by dictionary definition means entails converting content from one form to another, to turn into another language (The Merriam-Webster Dictionary. (1974). Massachusetts: Merriam-Webster). Translation, at its core, is the process of rendering written or spoken content from one language (the Source Language) into another language (the Target Language or Receptor Language) while maintaining the meaning, tone, and intent of the original message.

Translation is a process of written communication between the writer and the possible audience through the translator's conveying meaning of a source language into a target language considering the elements of lexicon, grammar, and context to meet the requirement of naturalness of a target text; a procedure with four levels: the SL text level, the level of language; the referential level, the level of objects and events, real or imaginary; the cohesive level; and the level of naturalness (Truong, 2023).

Although several techniques exist, BLEU (Bilingual Evaluation Understudy) is perhaps the most common MT translation evaluation tool used by most researchers. (Papineni et al., 2002), and some studies have shown that it achieves a high correlation with human judgments of quality (Couglin, 2003). Using this technique, scores ranging from 0 to 100 are calculated for a sample of translated text by comparing it to a reference translation, and the method takes into account the number of words that are the same in the two passages as well as the word order.

Kembaren et al. (2023) in their research analyzed the grammatical errors produced when using GT in writing narrative texts from Indonesian language to English language in writing narrative text among Culinary Art students where they used GT to complete assignments or translate tasks. Kembaren et al. used Surface Strategy Taxonomy (SST) theory of analysis is used to categorize the data, which included errors in omission, addition, misinformation, and misordering.

Machine Translation is the mechanization and automation of the process of translating from one natural language into another. Translation is a task which needs to tackle the 'semantic barriers' between languages using real world encyclopaedic knowledge and requires a full understanding of natural language (Sin-Wai, 1988).

Hutchins & Somers (1992) define MT as the traditional and standard computerised system responsible to produce translations from one natural language into another, with or without human assistance, earlier introduced in English as 'mechanical translation' and 'automatic translation' and in other languages as French- traduction automatique, Russian- avtomatičeskii perevod. Hutchins & Somers further elaborate that boundary between Machine- Aided Human Translation and Human- Aided Machine Translation (HAMT) are often uncertain and the term Computer-Aided (or Computer-

Assisted) Translation (both CAT) can sometimes cover both. But the central core of MT itself is the automation of the full translation process.

2.2. Error Analysis

Error Analysis, a fundamental branch of applied linguistics, emerged in the sixties to address students' performance (Shrestha, 1979). Keshavarz (2011) defined EA as collecting samples of errors, identifying errors, classifying, and evaluating them, further considering errors to be systematic, and regarded as rule-governed, when they follow the rules of the learner's interlanguage. According to Hutchins & Somers (1992), quality assessment appraises the 'goodness' of a translation, while error analysis measures the 'badness' of the MT output by identifying errors and assessing how much is there to correct to achieve the standard, acceptable status as a translation.

According to (Ubol, 1988), EA is a methodical description, systematic examination and interpretation of errors made by learners of the TL. EA focuses on identifying errors, clarifying of the occurrence and the production of their oral or written expression distinguishes from that of native speaker or TL standard. EA plays a crucial role in advancing the theoretical understanding within applied linguistics and among language professionals. According to Gass & Selinker (1994), EA makes a comparative analysis between the errors a learner makes in producing the TL and the TL form itself. It is similar to the weak version of contrastive analysis in that both start from learner production data; however, in contrastive analysis the comparison is made with the native language, whereas in error analysis it is made with the TL.

2.3. Errors generated by Google Translate

MT is evaluated as a piece of software in terms of general parameters such as speed and number of supported file formats, subject to existing standards and criteria. Evaluation of MT becomes a matter of text quality assessment because MT outputs are essentially in the form of text. The major approaches of MT evaluation, including both manual and automatic, will be presented with a discussion of their strengths and weaknesses (Sin-Wai, 1988). Google Translate's performance is often linked to the amount of human-translated text it has access to. The system relies heavily on vast corpora of bilingual texts, which it uses to "learn" how to translate between languages. Google Translate depends on the number of human translated texts searched by Google Translate (Karami, 2014).

3. Methodology

This study was conducted to examine the accuracy and limitations of Google Translate (GT) in the context of English to Chinese and Chinese to English translation. A qualitative research design was employed to analyse GT's performance across a range of authentic texts, including articles, brochures,

manuals, and website excerpts. Both the original source texts and their GT-generated translations were examined to identify translation errors and determine areas where GT fails to maintain linguistic and semantic accuracy.

In this study, qualitative analysis was used to locate and classify translation errors according to major linguistic categories, including syntax, morphology, lexis, and semantics. This approach enabled a detailed exploration of the linguistic mismatches that occur in Chinese and English machine translation.

As part of the preliminary phase, a questionnaire was administered to a group of undergraduates studying Chinese as a Foreign Language at a private university in Sri Lanka. The aim was to obtain background information on their use of Google Translate for language learning and translation tasks. The sample consisted of 20 learners enrolled in business-related degree programmes, of whom 10 were taking Chinese as an elective foreign language at the beginner level. These students attended language classes for three hours per week over a period of five months. Translation tasks were assigned toward the end of the course once students had acquired basic sentence construction skills, allowing for a comparison between student-generated translations and GT output. The insights from this preliminary data supported the need for evaluating GT's accuracy and informed the qualitative analysis undertaken in this research.

Based on the responses given through the questionnaire on the use of Google Translate in their translation work, a translation task was given for each language group separately. The data collected from the participants gave background information about them, their tendency to use GT and the causes. At the first stage, the students were not allowed to use GT to do the translation task. Once they completed the translation task on their own, the students were asked to use GT on their mobile devices and compare and contrast their writing with the TT produced by GT.

Based on the findings from the preliminary phase, a set of authentic Chinese and English texts of varying genres and lengths were selected for detailed analysis. These included materials such as articles, brochures, manuals, and website excerpts, chosen to represent a broad range of real-world communication contexts relevant to the learners' academic and professional needs. Each text was translated using Google Translate, and the machine-generated target texts were then compared with the original source texts to identify translation inaccuracies and determine patterns of linguistic errors. The selection of authentic documents ensured that the study examined GT's performance in practical and contextually diverse situations. Additionally, a simple translation task was assigned to the undergraduates studying Chinese as a Foreign Language at the beginner level, allowing for a supplementary comparison between learner-generated translations and GT output.

4. Data Analysis

The analysis revealed that the 20 participants were beginner-level learners of Chinese, which was an important factor in interpreting their reliance on Google Translate despite their limited language competence. Usage patterns indicated that nine students used Google Translate daily, while another eight reported using it at least once a week, demonstrating frequent dependence on MT tools for basic translation needs. Before completing the translation task, a majority of the participants believed that Google Translate produced translations that were “accurate” or “very accurate.” However, after engaging with the translation activity and comparing machine output with human translation, fewer students maintained this perception. This shift highlights the learners’ revised understanding of the limitations of GT in producing linguistically and contextually accurate translations between Chinese and English.

Responses to the open-ended and close-ended questions provided rich qualitative insights into learners’ experiences, expectations, and challenges when relying on GT. Their reflections offered nuanced perspectives on how effectively GT supported—or hindered—their translation processes. To evaluate GT’s translation accuracy, the study used authentic Chinese and English texts for comparison between machine-generated translations and reference translations. This range of authentic materials allowed the analysis to assess GT’s performance across diverse linguistic contexts, including both technical and non-technical content. Particular attention was given to contextual meaning, as Chinese and English differ significantly in syntactic structure, cultural references, and semantic expression. The analysis therefore focused on how well GT preserved meaning and nuance when translating between these two linguistically distinct languages.

5. Discussion

The analysis of Google Translate (GT) output reveals that, despite technological advancements, the tool continues to produce translations that appear unnatural, awkward, or semantically inaccurate when translating between English and Chinese. Many of these errors result from literal, word-for-word translation strategies that fail to capture meaning or contextual nuance—an issue particularly pronounced due to the significant structural, syntactic, and cultural differences between the two languages.

Across the dataset, error patterns varied depending on text type and length. Shorter texts tended to yield disproportionately higher error rates, indicating that GT struggles when limited contextual information is available. Although GT performed better with longer texts, where additional context helps inform translation choices, errors still persisted. These findings suggest that users should exercise caution when relying solely on GT, especially in tasks requiring precision or cultural and semantic accuracy.

Learner feedback further supports these observations. Several prominent themes emerged from the participants' reflections, revealing the practical challenges they encountered when using GT for English–Chinese translation tasks.

1. **Inaccurate or misleading translations:** Participants frequently noted that GT often produced incorrect or incoherent output, making it difficult for them to rely on the tool for accurate comprehension or correction of translation tasks.
2. **Confusion and frustration:** Many learners expressed feeling confused or discouraged by inconsistent or nonsensical translations. Instead of serving as a supportive learning aid, GT often introduced uncertainty and undermined their confidence.
3. **Negative impact on language learning:** Participants reported that dependency on GT hindered rather than supported their language acquisition. Inaccurate translations led to misunderstandings and prevented learners from developing reliable awareness of Chinese sentence structure and meaning.
4. **Overall dissatisfaction:** Learners' reflections conveyed a strong sense of disappointment, as GT did not meet their expectations as a translation support tool. Their experiences highlight the limitations of MT tools for beginner learners who lack the linguistic knowledge needed to identify and correct machine-generated errors.

Overall, the findings point to a shared perception among learners that GT is not sufficiently reliable for correcting translation tasks at the beginner level. The recurring issues of inaccuracy, conceptual misunderstanding, and learner frustration underscore the need for more guided pedagogical approaches that combine human instruction with technology, rather than depending solely on automated translation systems.

These results reinforce the importance of critically evaluating MT performance across diverse text types and contexts, and they highlight the ongoing need for improved MT accuracy in English–Chinese translation—a particularly complex language pair due to its profound linguistic and cultural contrasts.

Error Category	Description	Typical Example (Generic)	Impact on Translation Quality
Lexical Errors	Incorrect word choice, mistranslated vocabulary, wrong synonyms, incorrect polysemy interpretation.	“Bank” translated as 岸 instead of 银行.	Alters meaning; creates inaccurate or misleading translations.

Semantic Errors	Misrepresentation of meaning, missing implied meaning, misunderstanding idioms or culturally bound expressions.	Idiom translated literally: “Break the ice” → 打破冰块.	Produces unnatural, nonsensical, or misleading expressions.
Syntactic Errors	Incorrect sentence structure, word order, or grammar patterns misaligned with Chinese or English syntax.	English SVO → incorrect Chinese word order.	Lowers readability and disrupts natural flow.
Morphological Errors	Errors involving plural markers, tense, aspect, classifiers, or missing morphological markers.	Missing classifier: “three books” → 三书 instead of 三本书.	Causes grammatical inaccuracy; affects clarity.
Pragmatic Errors	Context or register not maintained; inappropriate level of formality; culturally unsuitable phrases.	Formal English translated into overly casual Chinese.	Results in socially inappropriate or contextually incorrect output.
Literal Translation Errors	Direct, word-for-word translation that fails to capture meaning or nuance.	“Heavy rain of problems” → 问题的暴雨.	Produces unnatural and sometimes nonsensical text.
Omission Errors	Missing key words, phrases, or grammar components.	Omitting “not” or modal verbs.	Distorts intended meaning.
Addition Errors	Extra words or meanings added by GT that do not exist in the source text.	Adding unnecessary intensifiers.	Alters original message; affects fidelity.
Named Entity Errors	Incorrect translation of names, places, organisations, or culturally specific terms.	Translating “Shanghai Disneyland” incorrectly.	Breaks accuracy and cultural meaning.

Summary of Linguistic Error Types Identified in English–Chinese and Chinese–English GT Output

Error Category	Frequency	% of Errors	Total	Notes
Lexical	12	30%		High frequency due to polysemy and word choice issues.
Semantic	9	22.5%		Literal translations of idioms are common.
Syntactic	8	20%		Word-order mismatches between English and Chinese.
Morphological	5	12.5%		Classifier and aspect errors common.
Pragmatic	4	10%		Register mismatch; inappropriate tone.
Named Entity	2	5%		Mostly due to incorrect transliteration.

Error Frequency by Category (Generic Example Structure)

Source (ST)	Text	GT Output	Correct Translation	Error Type	Explanation
ST example in English		GT Chinese output	Human translation	Lexical	Wrong word choice changes meaning.
Chinese ST example		GT English output	Human translation	Syntax	Incorrect word order.
Chinese ST		GT English	Human translation	Morphology	Missing classifier.
English ST		GT Chinese	Human translation	Semantics	Literal idiom translation.

Examples of Errors Extracted from GT

Theme	Description	Representative Learner Insight
Accuracy Concerns	Learners noted frequent errors and inconsistencies.	“The translation looked correct, but the meaning was wrong.”
Confusion & Frustration	Inaccurate output caused uncertainty.	“GT made me more confused than before.”

Learning Setbacks	GT hindered learning instead of supporting it.	“I couldn’t understand which one was correct—GT or my work.”
Discouragement	Expectations of reliability were unmet.	“Using GT made me doubt my progress.”

Summary of Participant Feedback Themes Related to GT Use

6. Conclusion

This study examined the accuracy and limitations of Google Translate (GT) in translating between English and Chinese, highlighting both its potential and its constraints in a language-learning context. The findings reveal that while GT can serve as a useful support tool, particularly for beginners seeking quick meaning access, its output remains prone to errors arising from literal translation, structural mismatches, and contextual misunderstandings. These issues are especially evident when translating between languages with significant linguistic and cultural differences, such as English and Chinese.

Given these limitations, it is essential that learners first attempt their own translations before relying on GT for comparison or correction. Teachers play a critical role in guiding students to critically evaluate GT output rather than accepting machine translations as authoritative. Students must be encouraged to verify meaning independently and develop an awareness of GT’s common error patterns.

Although the analysis provided meaningful insights into GT’s performance across different text types and lengths, it does not represent the full spectrum of translation challenges that may arise. Nonetheless, the study reinforces the need for careful review of GT-generated translations and underlines the importance of human oversight in the learning process.

Ultimately, this research supports the informed and strategic use of GT as a supplementary tool rather than a primary translation method. By understanding its strengths and weaknesses, learners can harness GT effectively while simultaneously strengthening their English–Chinese translation skills and overall linguistic proficiency.

References

- Anggaira, A. S. (2017). Linguistic Errors on Narrative Text Translation Using Google Translate. *Pedagogy*, 5(1), 1–14.
- Aslerasouli, P., & Abbasian, G.-R. (2015). Comparison of Google Online Translation and Human Translation with Regard to Soft vs. Hard Science Texts . *Journal of Applied Linguistics and Language Research*, 2(3), 169–184.
- Couglin, D. (2003). Correlating automated and human assessments of machine translation quality. *Proceedings of Machine Translation Summit IX: Papers*.
- Gass, M. S., & Selinker, L. (1994). *SECOND LANGUAGE ACQUISITION: An introductory course*. <https://doi.org/https://bpb-us-e2.wpmucdn.com/websites.umass.edu/dist/c/2494/files/2015/08/Gass.Second-Language-Acquisition.pdf>
- Hilma, R. (2011). LITERAL TRANSLATION USING GOOGLE TRANSLATE IN TRANSLATING THE TEXT FROM FRENCH TO ENGLISH IN DIGITAL TOURISM BROCHURE “BIENVENUE À PARIS.” In *BINUS BUSINESS REVIEW* (Vol. 2, Issue Mei).
- Hutchins, J., & Somers, H. (1992). *An Introduction to Machine Translation*. Academic Press.
- Karami, O. (2014). The brief view on Google Translate machine. . *Seminar in Artificial Intelligence on Natural Language, German*.
- Kembaren, F. R. W., Daulay, S. H., & Srikandi, A. (2023). Grammatical Errors of Google Translate in Translating Indonesian Narrative Text to English. *Journal of English Language Studies*, 5(2).
- Keshavarz, M. H. (2011). *Contrastive Analysis and Error Analysis*.
- Papineni, K., Roukos, S., Ward, T., & Zhu, W.-J. (2002). BLEU: a Method for Automatic Evaluation of Machine Translation. *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL)*, 311–318.
- Shrestha, A. (1979). *Error Analysis: A Pragmatic Approach*. CNAS Journal, 1–9.
- Sin-Wai, C. (1988). *The Routledge encyclopedia of translation technology*. Routledge. <http://dspace.vnbrims.org:13000/xmlui/bitstream/handle/123456789/4402/Routledge%20Encyclopedia%20of%20Translation%20Technology.pdf?sequence=1&isAllowed=y>
- The Merriam-Webster Dictionary. (1974). Massachusetts: Merriam-Webster.(2024).
- Ubol, C. (1988). *An Error Analysis of English Compositions by Singapore Students*.