Journal of Language and Literature
Vol. 25 No. 2, October 2025, pp. 339 – 354
DOI: 10.24071/joll.v25i2.10468
Available at https://e-journal.usd.ac.id/index.php/JOLL/index

The Equivalence of Human Versus *Google* Translation: A Case Study of Translating Natural Science Writings

Aulia Addinillah Arum & Nadia Khumairo Ma'shumah
nadia.khumairo@mail.ugm.ac.id
Faculty of Cultural Sciences, Univeritas Gadjah Mada, Yogyakarta, INDONESIA

Abstract

Although recent advancements in machine translation have improved lexical and grammatical accuracy, assessing its effectiveness in rendering contextually and semantically accurate translations remains inadequate. This gap overlooks how systems like Google Translate handle specialized terms and subtle shifts in language style, particularly when working with field-specific texts. The objective of this study was to evaluate the performance of Google Translate in rendering natural science texts from English into Indonesian, with particular focus on its semantic accuracy when compared to human translation. The data consisted of 76 sentences drawn from six scientific texts, which were segmented into 41 single words as well as 35 multi-word phrases obtained from professional ProZ translator portfolios, and subsequently translated using Google Translate. The translations were analyzed for inaccuracies, classified as under-translation, over-translation, and mistranslation, drawing on Baker's (1992) theory of propositional meaning and also Nida's (1964) receptor-oriented framework. The analysis revealed that while Google Translate frequently produced grammatically correct structures, it often failed to generate contextually appropriate or domain-specific terms, resulting in semantic inequivalence with human translations. The findings indicate that these limitations are not inherent to the lexical output itself, but rather emerge from the system's inability to capture nuanced meanings, specialized registers, as well as situational contexts. Therefore, human translators remain indispensable in ensuring accuracy and reliability in field-specific translations, whereas machine translation is best positioned as a supportive tool for general comprehension.

Article information

Received: November 21, 2024

Revised: May 22, 2025

Accepted: June 2, 2025

***Keywords**: Human Translation; Google Translate; Propositional Meaning; Translation Equivalence*

Introduction

Equivalence has long been, and continues to be, a fundamental concern in both translation theory and practice, since "sameness" between two languages is neither linguistically nor culturally possible (Bassnett, 2005; Catford, 1965; Nida & Taber, 1974; Koller, 1979; Munday, 2009, 2008; Newmark, 1988). This results from the fact that each language embodies distinct

worldviews, value systems, as well as cultural references that may not have exact equivalents in another language. As a result, translators are often required to make interpretive choices, negotiating meaning across diverse linguistic and cultural boundaries to achieve functional or dynamic equivalence rather than mere formal correspondence, which prioritizes a one-to-one correspondence in linguistic form between the ST and TT (Nida & Taber, 1974). The challenge, therefore, resides not in reproducing identical forms, but in conveying meaning in a manner that is contextually appropriate and communicatively effective within the TL (Nida & Taber, 1974). This illustrates that translation entails much more than the mere substitution of lexical and grammatical elements from the SL to the TL. (Bassnett, 2002).

In his influential work, Jakobson (1959) outlined three modes of translation consisting of: intralingual, interlingual, and intersemiotic. He emphasizes the significance of interlingual translation, understood as translation between distinct written languages, and highlights the complex challenges it presents, particularly in relation to linguistic meaning and equivalence. Drawing on Saussure's theory of the linguistic sign, in which the signifier (*i.e.,* sound or written form) and the signified (*i.e.,* concept) are linked arbitrarily, he emphasized the intrinsic fluidity of meaning in translation. By addressing the challenge of achieving equivalence in different languages, such as English, Russian, Spanish, as well as German, he notes that "there is generally no complete equivalence between code-units." In interlingual translation, the translator tends to reconstruct and convey a message from the ST, mediating its linguistic form and semantic content within the TL. Translation, therefore, entails rendering equivalent messages across distinct linguistic codes (Jakobson, 1959).

The notion of 'equivalence' held sway as a key issue in translation throughout the 1970s and continued to exert significant influence in subsequent or later decades (Munday, 2008). Nonetheless, even though equivalence remains a cornerstone of translation studies, a definitive and universally accepted definition remains elusive among translation theorists. Nida and Taber (1974) define equivalence as the extent to which the receptors of a message in the TL respond to it in a manner that is substantially analogous to the response of recipients in the source language. This response, obviously, can never be identical, as the cultural and historical settings are too different; nevertheless, there should be a high degree of equivalence in the response, or the translation will have failed to accomplish its purpose (Nida & Taber, 1974).

By arguing that the SL and TL items rarely have the same meaning in the linguistic sense but they can function in the same or analogous situation, equivalence, based on Catford (1965), is generally established at the level of the text or sentence (hereinafter referred to as textual equivalence), while, the meaning, in this case, is subject to contextual conditions in the TT. In addition, Koller (1979) employed the concept of equivalence to denote equivalent elements within specific translation pairs and situational contexts. By employing Saussure's concept of *parole* as a main parameter, he emphasized that equivalence is neither universal nor absolute, but rather context-dependent, which is shaped by specific linguistic units as well as situational variables.

Larson (1984) emphasized the concept of meaning equivalence by defining translation as the transfer of meaning from the SL to the TL, and this is achieved by transitioning from the form of the first language to that of a second language through semantic structure. It is the 'meaning' that is transferred and must be held constant, whereas only the 'form' changes.

Otherwise, Newmark (1988) proposed the concept of equivalence in terms of achieving a comparable effect on the target readership as was obtained on the readership of the source text (or as close as possible to the ST/SL). This principle, without neglecting other methods, forms the foundation of his dual approach to translation, distinguishing between semantic and communicative translation. He then posits that achieving an equivalent effect constitutes the desirable outcome, rather than the aim, of any translation endeavor. Based on him, literal translation is considered the best approach in both semantic and communicative translation.

Hatim and Munday (2004), together with Baker (2018), recognized that equivalence is not absolute but rather relative, as it is influenced and determined by broader contextual

categories, including linguistic and cultural factors. Therefore, equivalence should not be perceived as a fixed or absolute concept but rather as a variable and negotiable outcome, shaped by the particularities of the SL and the TL, the purpose of the translation, and the expectations of the target audience. On the other hand, Pym (2004) added an important perspective to the debate by arguing that there is no such thing as perfect or total equivalence between languages. Although there are many different perspectives, it can be assumed that equivalence in translation is not impossible to obtain; however, full equivalence is not always attainable (Jacobson in Munday, 2008).

Due to its foundational significance, the concept of equivalence is frequently invoked in the assessment of translation quality. As Nida, cited in Angelelli and Jacobson (2009), argues translation quality can be understood through the principles or dichotomy of formal and also dynamic equivalence, which determine quality based on the response produced from target readers. This view is subsequently supported by Newmark (1988) and Hatim and Munday (2004). Collectively, while equivalence is inherently relative, it remains a useful metric for evaluating translation quality, especially when operationalized through empirical approaches such as translation acceptability and also error analysis. Across the three parameters used to evaluate translation quality, such as accuracy, readability, and acceptability (Castilho et al., 2018; Nababan, 2016; Nababan et al., 2012), acceptability focuses merely on whether the translation is perceived as natural as well as appropriate within the target language and culture (Castilho et al., 2018; Nababan et al., 2012; Nord, 1997), while error analysis in translation refers to the process of identifying, categorizing, and evaluating errors within a translated text in order to determine its level of accuracy, fidelity to the source content, and also the translator's competence in linguistic structures and discursive practice or language use (Baker, 2018; House, 2015; Nord, 1997; Wills, 1982).

While early research focused primarily on human translators as the sole agents of text rendering, recent decades have seen a marked increase in studies on machine translation by producing a wide array of findings as well as advancements. One interesting research that reveals the potential of *Google Translate* (GT) to produce near-accurate translation was carried out by Li, Graesser, and Cai (2014). Their study compared human translations with *Google*'s translations of 289 written as well as spoken passages from Chinese into English, and vice versa. By viewing the translation accuracy from the perspectives of formality and cohesion, the findings of their study indicated that *Google*'s English translations were strongly correlated with both human English translation and the original Chinese text.

In addition to the study of the accuracy of machine translation (MT), the study conducted by Läubli and Orrego-Carmona (2017) offers valuable insights into professional translators' evaluations of the machine translation quality. Based on their quantitative analysis of 13,000 tweets on machine translation quality posted by translators, the study revealed that negative perceptions outweigh positive ones by a ratio of 3:1 overall, and 5:1 in tweets referencing MT to human translation. They further advocate for enhanced collaboration among scientists, scholars, and professionals in order to advance the development of translation technologies.

Groves and Mundt (2015) conducted a preliminary study focusing on the grammatical accuracy of MT, with particular attention to the use of *Google Translate* in academic contexts. Specifically, their analysis centered on English translations by situating the grammatical evaluation within the linguistic norms and English syntactic conventions. Their findings indicate that while *Google Translate* is not entirely free from grammatical errors, its output satisfies the minimum language proficiency standards required for university admission. This underscore the potential of MT to support academic language use, while also highlighting the need for caution given their limitations. Thus, taken together, these aforementioned studies point to a gap in understanding how MT quality is perceived across different dimensions and by different users, highlighting the need for more detailed, domain-specific research.

With machine translation emerging as a prevalent tool for translating texts today, this study compares the equivalence achieved by human translators (HT) and *Google Translate*

(GT) when implemented in natural scientific writings. This study is important because of the distinctive features of this genre. Scientific writings, particularly in the natural sciences, are distinguished by specialized terminology, precise definitions, and also a formal register, which together distinguish them from other genres. This domain-specific complexity poses real challenges for maintaining translation equivalence, which underscores the urgency to evaluate how well MT tools handle technical terms and complex sentence structures.

This study focuses on evaluating machine translation quality, addressing the variability in findings across previous research that stem from differences in language pairings, quality metrics, and the method of analysis used by many researchers. In comparing machine with human translation, this study focuses on two categories: grammatical and lexico-semantic equivalence. Grammatical equivalence and/or inequivalence in this study is evaluated based on the word count as well as syntactic function of the translation (from English to Indonesian, and vice-versa), while semantic equivalence is evaluated based on the accuracy of meaning. This study then proposed the categorization of meaning inaccuracies at word and phrase level by adopting terminologies proposed by Baker (2018), consisting of over-translation, under-translation, and mistranslation, as outlined by Kim (2019) and Newmark (1988). This framework draws upon Nida's (1964) theory of the receptor's response by providing critical perspectives for enhancing the accuracy of translation evaluation. It also provides a structured method for identifying as well as addressing semantic errors in both human and machine translation.

Methodology

1. Data Sources and Data Collection Techniques

This study adopts a product- and function-oriented approach to translation as it aims to investigate existing translations and seeks to describe the function of translations in the target situation. Translation samples used in this study were collected from the portfolios of professional translators listed on *ProZ* (look at https://www.proz.com/), an online platform as well as community specifically designed for language professionals, including translators, interpreters, and also localization experts. The members were shortlisted using a filter function to screen for experts who specialize in translating scientific writings across various fields. Accordingly, only samples produced by highly rated translators are selected for further analysis. In this study, the names of the translators were not revealed as part of the research ethics.

The selection criteria were grounded in the translators' cumulative user ratings and professional credentials as indicated on the platform. Only those translators with a proven track record of client satisfaction, specialized competence in scientific translation, and also sustained active membership on the platform were included in the sample. Six translation samples on natural science theme were chosen as the primary material objects. Medical and environmental engineering were classified as subdomains of natural scientific writing, with the study specifically focusing on English-to-Indonesian translations. Both disciplines have a solid foundation in the natural sciences, drawing from chemistry, biology, and physical science, which justifies their inclusion within this category. In this context, "samples" refer to complete paragraphs containing a minimum of three sentences, which exemplify translated texts within the scientific domain.

Regarding this, Williams and Chestermen (2002) states that technical translation covers many kinds of specialized texts translation in the domains of science, medicine, and also in other disciplines. The translation of these texts needs a high level of subject knowledge and mastery of the relevant terminology. Style and clarity, text-type conventions, culture-specific, and also reader expectations become fixed or standardized, almost like patents. All of these samples were not selected randomly, but were purposefully chosen and identified based on relevance to the study's focus and the clarity of the SL and the TL relation. In this study, each sample was sourced from a distinct translator in order to ensure a broader representation of translation practices and also minimize bias.

After identifying the translation samples, the ST and TT were separated into sentences. 76 sentences from 6 different samples were manually input into *Google Translate* (GT). The three groups of texts, comprising of the ST, HT, and GT output were further systematically compiled in Excel. This enabled a systematic selection of the most significant words and phrases, which served as the primary units of analysis and constituted the final dataset for the study.

Word and phrase selection was based on the principal lexical class categories outlined by Givón (1993). According to Givón (1993) four main lexical categories frequently occur across languages, such as *nouns* (n.), *verbs* (v.), *adjectives* (adj.), and *adverbs* (adv.). However, the class of words examined in this study were limited to nouns, verbs, and adjectives, with nouns and verbs identified as the main lexical categories among the classes (Givón, 1993). Adjectives appeared with notable frequency in the dataset and were therefore included in the analysis to satisfy the statistical requirements. Moreover, phrases were also used as objects of analysis, by defining a phrase as the group of words that serves as a functional component of a longer utterance (Verhaar, 2016). Most of the identified phrases fall into categories of *Noun Phrases* (NP) and *Adjective Phrases* (Adj.P). In summary, a total of 41 words and 35 phrases were extracted from the English ST for further analysis.

Table 1 Samples of the data

No.	Source Text	Human Translation	*Google Translate*
1.	All hexavalent chromium compounds (including lead chromates) are **considered** to be suspect human carcinogens.	*Seluruh senyawa kromium heksavalen (termasuk timbal kromat)* ***diduga*** *karsinogen bagi manusia.*	*Semua senyawa kromium heksavalen (termasuk timbal kromat)* ***dianggap*** *sebagai karsinogen bagi manusia.*
2.	However, available epidemiological evidence does not **confirm** this **position.**	*Namun demikian, bukti epidemiologis yang ada tidak* ***menunjang pernyataan*** *ini.*	*Namun, bukti epidemiologi yang tersedia tidak* ***mengkonfirmasi posisi*** *ini.*
3.	In every case where excess lung cancer **incidences** have been reported, **exposure** was either to zinc chromate alone or involved mixed exposures to various combinations of zinc, lead, strontium and barium chromates.	*Dalam setiap kasus* ***terjadinya*** *kanker paru-paru berat yang telah dilaporkan,* ***paparan*** *diakibatkan oleh seng kromat saja atau campuran paparan oleh berbagai kombinasi dari seng, timbal, strontium, dan barium kromat.*	*Dalam setiap kasus di mana* ***insiden*** *kanker paru-paru berlebih telah dilaporkan, paparan baik seng kromat saja atau melibatkan* ***eksposur*** *campuran ke berbagai kombinasi seng, timbal, strontium dan barium kromat.*
4.	In **studies**, no increased incidence in lung cancer was observed.	*Dalam* ***kajian****, tidak diamati terjadinya peningkatan kanker paru-paru.*	*Dalam* ***penelitian****, tidak ada peningkatan insiden kanker paru-paru yang diamati.*
5.	As noted in the OSHA Lead Standard, repeated and prolonged exposures may cause delayed effects involving the blood, gastro-intestinal, nervous and reproductive systems.	*Sebagaimana tercatat pada Standar Arahan OSHA, paparan berulang dan lama dapat menyebabkan dampak tertunda pada sistem darah,* ***pencernaan****, syaraf, dan reproduksi.*	*Sebagaimana dicatat dalam Standar Timbal OSHA, pemaparan berulang dan berkepanjangan dapat menyebabkan efek tertunda yang melibatkan sistem darah, gastro-usus, saraf dan reproduksi.*
6.	Due to their extreme water insolubility these products are non-toxic to aquatic life.	*Karena sifatnya yang benar-benar tidak larut di dalam air, produk ini tidak beracun*	*Karena sifatnya yang sangat tidak larut dalam air, produk ini tidak beracun bagi kehidupan akuatik.*

No.	Source Text	Human Translation	*Google Translate*
		terhadap kehidupan perairan.	

** Bolds in Table 1 represent words, while underlines represent phrases.*

2. Data Analysis Techniques

In analyzing the final dataset, this study employed a descriptive qualitative approach, utilizing a comparative method to examine the outputs of machine translation against human translation. The researchers were positioned as the primary instruments in order to conduct a manual evaluation of selected words as well as phrases across source texts (ST), human-translated texts (TTs1), and also the machine-translated texts (TTs2). The assessment of the word- and phrase-level translations was based on two key criteria. The first criterion concerns equivalent lexicons and phrases, where both human (HT) and machine translations employ different words or phrases, but the overall meaning remains consistent and does not alter the intended message of the text. The second criterion involves inequivalent lexicons and phrases, characterized by the use of different words or phrases in the translations, with at least one instance causing a change in meaning at the textual level. A comprehensive review was subsequently conducted to assess which translation version was more similar to the original test. The evaluation primarily focused on detecting cases of inequivalence, where the translation output failed to convey the intended presupposed meaning accurately. as outlined by Baker (2018).

Results and Discussion

The nuance of surroundings determines language expression, especially when used in text writing. Thus, the equivalence of meaning becomes critical particularly when translated into another language. However, given that scientific writings mainly consist of declarative sentences, the concept of 'meaning' is defined as non-linguistic entity present in all sentences that express the same meaning. Drawing on the concepts of equivalence and translation quality, Baker (2018) further highlighted that when translation is described as 'inaccurate', it is often the propositional meaning that is being called into question since it provides the basis for judging an utterance as true or false (p. 12), which can be identified by the occurrences of errors like under-translation, over-translation, and mistranslation in the translation output. In this context, proposition is then defined as a structured grouping of concepts that conveys meaning as a single unit (Baker, 2018; Larson, 1984). It entails a semantic unit composed of multiple concepts, wherein one serves as the central element and the others are directly associated with it (Larson, 1984, p. 189).

Kim (2019) believes that the majority of translations tend to convey either more or less information. As a direct consequence, readers may encounter difficulties due to the semantic gap between the ST and TT. Newmark (1988) was one of the earliest scholars in translation studies to introduce the terms over-translate and under-translate as the tendency generated from semantic and communicative translation. When semantic translation looks back at the ST and seeks to maintain its characteristics as much as possible, communicative translation looks towards the demands of the addressees and tries to satisfy them as much as possible (Newmark, 1988; Panou, 2013).

However, as corresponding between ST and TT do not necessarily have precisely the same semantic range, translators tend to over- or under-translate most of the time, with the latter being more likely to occur (Newmark, 1988). Under-translation occurs when more general words (superordinate) are employed to enhance clarity, promote simplicity, and, in some cases, achieve brevity in the target text (Newmark, 1988). It happens when the loss of meaning is diminished through the use of more generalized expressions.

Contrary to Newmark's (1988) assertion, most occurrences of under-translation lead to ambiguity instead of clarity, as the TT fails to adequately convey the precise meaning of the ST, since they tend to provide less information. However, Newmark (1988) additionally stated

that under-translation can be warranted if an informative text is deficient in clarity. On the other hand, it may constitute an error when it is merely a literal translation.

Over-translation happens when the loss of meaning entails an increase in details (gain) (Kim, 2019). Between under-translation and over-translation, translators are more prone to do under- than over-translation (Newmark, 1988) since the use of more general words is often done as a strategy when the TL lacks of specific term (hyponym) as conveyed in the SL (Baker, 2018). One illustration of over-translation occurs when translators render words that would be better omitted than interpreted, such as modal connectives in German. When a translator freely translates ST by adding extra words or material that do not present in the original ST, he/she produces creative translation, which also falls under the category of over-translation (Kim, 2019).

Mistranslation of a single lexical item can significantly affect the coherence of a text (Baker, 2018). Mistranslation has long been used as an error indicator to assess translation quality. However, it is also deemed overused, conveying nothing beyond the basic assertion of 'error in translation' (Angelelli & Jacobson, 2009). Even so, meaning-oriented criteria can provide an assessor with the ability to explain which part of the meaning is mistranslated and why (Angelelli & Jacobson, 2009). Thus, it is worth noting that the more severe an error is, the more likely it is to negatively affect readers in some way (Lommel, 2018, p. 120).

Williams (2004) defines mistranslation as complete failure to render the meaning of a word or passage that contains a vital element of the message, resulting in a contradiction or substantial deviation from its essential meaning. He then noted that serious mistranslation can lead a definite lack in comprehension of the SL, nonsense, omission of a phrase or more. The keyword is essential. The determination of whether an essential element of the message was implicated rested with the quality controller or evaluator.

Human translators are more susceptible to errors involving under-translation or over-translation due to their reliance on intuition and contextual understanding, which, while enabling them to approximate meaning more effectively, also gives rise to subjectivity in the selection of the most appropriate equivalent. When it comes to machine translation, there is another level of mistranslation which results in what is called as 'globally error' (Sutrisno, 2020). This type of error emerges when the error affects the overall coherence, intent, or communicative function of the entire text. In contrast to local errors, which typically affects individual words or even short phrases, global errors disrupt the coherence and intended message of the entire text, posing a higher risk of misunderstanding among readers. This type of mistranslation could be triggered by how the machine generates a translation. Machine-generated translations, according to Sutrisno (2020), operate by segmenting the source text, detecting the match in the target language, and then determining the optimal distribution and match for the sentence.

The analysis of translation quality entails a detailed examination at how meaning is conveyed across languages, particularly when comparing human translation and machine-generated translation. As outlined in the theoretical framework, equivalence plays a central role in assessing accuracy, clarity, and appropriateness in the target text (TT). While human translators rely on contextual understanding and also intuition, *Google Translate* operates through automated segmentation and pattern recognition, which usually produces different levels of equivalence. Accordingly, the findings are presented in two parts: (1) the equivalence of HT and GT, and (2), the inequivalence that emerge between the two.

1. The Equivalence of Human Translation and Google Translation

To facilitate a clearer comprehension of the quantitative variations between HTs and MTs, the analysis centers on the differences in word counts. Findings demonstrated an average difference of 1.75 words between the outputs of humans and *Google* translations. This value was derived by dividing the total sum of word-count differences (123) by the number of data items analyzed (70). The subsequent Table 2 presents the distribution percentages for each

category of word-count difference(s) observed within the dataset.

Table 2 Frequencies and Percentages of Word-Count Differences Between Human Translation and *Google* Translation

No.	Word-count difference(s)	#Translation Pairs	Percentages (in %)
1.	0	16	22.9
2.	1	21	30.0
3.	2	19	27.1
4.	3	4	5.7
5.	4	2	2.9
6.	5	2	2.9
7.	6	3	4.3
8.	7	2	2.9
9.	8	1	1.4

Table 2 above illustrates the distribution of word-count differences between HT and MT generated by *Google Translate* throughout the examined translation pairs. Notably, the vast most observed differences are restricted to 0 to 2 words, accounting for approximately 80% of the total sample. Specifically, 22.9% of the translation pairs exhibit identical word counts, indicating exact parity in length between HT and MT outputs. Moreover, 30.0% and 27.1% of the pairs differ by only one and two words, respectively, underscoring a close alignment in text length in most cases.

In contrast, only a minority, 20% of the data, indicate word-count differences more than two words, with differences ranging from 3 to 8 words. This relatively small proportion suggests that significant divergences in length are uncommon. The predominance of minimal word-count variation between the HT and MT indicates considerable syntactic and structural congruence in the translations.

One plausible explanation for this form of convergence lies in the syntactic similarities shared by English and Indonesian, particularly their shared subject-verb-object (SVO) word order, which facilitates more straightforward correspondence in translation. This syntactic alignment contributes to GT's effectiveness in maintaining comparable sentence lengths to human translators, as supported by Sutrisno (2020). Consequently, the data indicate that MT can closely approximate HT in terms of length consistency in the context of English-Indonesian language pair.

Another significant finding regarding on word count reveals that human translations generally contain a higher number of words in comparison with *Google Translate* renditions. This finding indicates that human translators tend to produce more elaborated translations compared to *Google Translate*. Excerpt 1 below exemplifies how such elaboration generates a higher word count in human translations.

Excerpt 1.
English Version:
EPA's health assessment document [*for chromium*] states that [*animal cancer*] [*bioassay studies*] suggest that hexavalent chromium compounds (particularly [*soluble and sparingly soluble compounds*]) are probably the etiological agent in [*chromium related*] [*human cancer*].

Human Translation:
Dokumen penilaian kesehatan EPA dalam hal kromium menyatakan bahwa kajian uji hayati kanker pada hewan menyarankan bahwa senyawa kromium heksavalen (khususnya senyawa yang dapat larut dan yang sedikit dapat larut) kemungkinan merupakan bahan etiologis dalam kanker pada manusia yang terkait dengan kromium.

Google Translation:
Dokumen penilaian kesehatan EPA untuk kromium menyatakan bahwa studi bioassay kanker hewan menunjukkan bahwa senyawa kromium heksavalen (terutama

senyawa yang larut dan sedikit larut) mungkin merupakan agen etiologi dalam kromium terkait kanker manusia.

Excerpt 1 presents a data sample that demonstrates the greatest disparity in word count between human and machine-generated (*Google*) translations. Six translation units at the level of phrase were identified in order to differ between the human translation and the output generated by *Google Translate.* A salient example is the translation of the prepositional phrase [PP] *for chromium*, which was rendered as *dalam hal kromium* by human translation and as *untuk kromium* by *Google Translate*. The variation in the trestment of the preposition *for* reflects divergent semantic interpretations adopted by human and machine translation systems. A back-translation of the *Google* output yields a direct lexical equivalent, while the human translation aligns more closely with the English idiomatic expression for *chromium.* This contrast indicates that *Google Translate* employed a more literal translation strategy, as theorized by Vinay and Darbelnet (1955). According to their strategy, literal translation involves a direct, word-for-word rendering of the ST into the TT, maintaining the original lexical and syntactic structures as closely as possible. This strategy often prioritizes fidelity to the SL's surface form, which can lead to reduced naturalness or idiomatic accuracy in the TL.

Differences in preposition translation led to different interpretations and also inaccurate collocation. From the prescriptive grammar perspective, the *Google* translation version of *for chromium* is perceived as inaccurate as the proposition *untuk* in the Indonesian language is typically initiated with the verbal phrase, where the verb is generally an action verb. In addition, there is also a scientific term in ST, *bioassay studies*, which is translated differently by human and *Google,* where human tend to translate it as *kajian uji hayati,* while *Google* translate is as *study bioassay.*

When faced with unfamiliar terms in SL, human translators tend to adopt the approach of domestication, whereas *Google* translation appears to favor a foreignization approach, as evidenced in both outputs. When translating an unfamiliar or a more technical term such as *bioassay*, *Google* tends to employ a 'borrowing' strategy, wherein the ST is not adapted but retained unchanged in the target language. On the other hand, human implemented a calque procedure, whereby an expression from the SL is borrowed into the target language through the literal translation of each lexical unit, as proposed by Vinay and Darbelnet (1955). This also can be observed in the following Excerpts 2 and 3.

Excerpt 2.
English Version:
In every case where excess lung cancer incidences have been reported, **exposure** was either to zinc chromate alone or involved mixed exposures to various combinations of zinc, lead, strontium and barium chromates.

Human Translation:
Dalam setiap kasus terjadinya kanker paru-paru berat yang telah dilaporkan, ***paparan*** *diakibatkan oleh seng kromat saja atau campuran paparan oleh berbagai kombinasi dari seng, timbal, strontium, dan barium kromat.*

Google Translation:
Dalam setiap kasus di mana insiden kanker paru-paru berlebih telah dilaporkan, paparan baik seng kromat saja atau melibatkan ***eksposur*** *campuran ke berbagai kombinasi seng, timbal, strontium dan barium kromat.*

In Excerpt 2, the noun *exposure* serving as the main subject of the sentence is translated by human into Indonesian word *paparan.* In contrast, *Google* tends to render it as *exposur* by using naturalisation. This strategy involves directly importing the source language word into the TL, often with slight phonological or orthographic adjustments to fit the TL norms. While human translators prioritize semantic equivalence and naturalness, *Google Translate* relies on preserving the source language form, resulting in a less familiar or less idiomatic translation for native speakers of Indonesian.

Excerpt 3.
English Version:
Due to their extreme water insolubility these products are non-toxic to aquatic life.

Human Translation:
Karena sifatnya yang benar-benar tidak larut di dalam air, produk ini tidak beracun terhadap kehidupan perairan.

Google Translation:
Karena sifatnya yang sangat tidak larut dalam air, produk ini tidak beracun bagi kehidupan akuatik.

In the same manner, in Excerpt 3, adjective phrase with a prepositional phrase functioned as subject complement stating a property of the subject *non-toxic to aquatic life* is literally translated into *kehidupan perairan* by human and *kehidupan akuatik* by *Google* (machine), a more scientific-sounding equivalent, reflecting a tendency for formal equivalence or lexical borrowing from English. While both versions are grammatically correct, human translation delivers a clearer and stylistically appropriate version for typical Indonesian readers.

Despite the significant difference in word count between human translations and *Google* translations, the overall meaning of the source text still appears to be adequately preserved by *Google Translate*. This finding aligns with the results of previous studies by Li, Graesser, and Cai (2014), who demonstrated that *Google Translate* can produce translations that are generally understandable and readable, even when grammatical errors are present. It is also consistent with Groves and Mundt (2015), who revealed that *Google Translate*'s accuracy level is approaching the minimum threshold required for university admission in many institutions in comparison with international language proficiency standards.

A substantial degree of similarity between human and *Google* translations was identified, as demonstrated through the translation-unit pairs (TUP) analysis at word and phrase levels. Table 3, derived from the category of accurate words and phrases, illustrates that the content of the HT and GT is identical, which indicates translation equivalence.

Table 3 Accurate Lexicons and Phrases Samples' Data

Source Text	**Human Translation**	***Google* Translation**
Angioplasty and **stenting involve** some risks.	*'Tindakan angioplasti dan pemasangan stent* ***mengandung*** *risiko.'*	*'**Angioplasti** dan pemasangan stent* ***memiliki*** *beberapa risiko.'*

In the case of Table 3, scientific words such as *angioplasty, stenting,* and i*nvolve* are further identified as translation units. HT incorporates the additional word: *tindakan* [n.] to provide a more contextual clarity, while the GT adopts a literal strategy, choosing not to introduce any supplementary terms. The differentiation in the use of strategy then resulted in different forms in both versions. The HT transformed a single word or lexeme into a phrase, while the GT preserved the form of the ST. Nonetheless, both versions contain the same meaning and syntactic function in the sentence. As a result, both HT and GT can be deemed equivalent in terms of pragmatics, semantics, and grammar.

The word *stenting* is translated identically by both the human translation (HT) and the Google translation (GT), thereby producing an equivalent semantic effect in both versions. The translation of the word *involve* identifies many variations of the translated terms. To assess the equivalence of the two translations, the translation versions were interchanged between pairs. When the word *memiliki* was transferred into the human translation (HT) version, both the word-level and contextual meanings remained unchanged. This finding supports the conclusion that both translations are equivalent and effectively communicate the intended meaning to the reader.

Moreover, this interchangeability reflects a high level of semantic fidelity, suggesting that the HT and GT successfully preserved both denotative and connotative aspects of the ST. This consistency strengthens the credibility of the translation process and guarantees that the target audience obtains an accurate and faithful interpretation, despite of the strategy

Journal of Language and Literature
ISSN: 1410-5691 (print); 2580-5878 (online)

employed. This finding further emphasizes the critical role of lexical selection in preserving coherence as well as semantic integrity across various translation strategies.

2. The Inequivalence of Human Translation and *Google* Translation

Out of 76 translation-unit pairs analyzed, 40 instances of inequivalent translations were found at the words and phrases levels between humans' version and *Google Translate* outputs. The analysis revealed that GT accounted for 24 instances of inequivalent translations, while human translators contributed 16 such cases. Nevertheless, the main objective of this study is not to generate a quantitative account of inequivalence patterns, but rather to examine the various manifestations of inequivalence arising from both human and GT processes.

As previously outlined, this study adopts three categories of translation inaccuracies that lead to inequivalence: *under-translation, over-translation,* and *mistranslation*. These full categories fall under the broader framework of meaning deviation. Overall, both human and *Google Translate* exhibit instances of all three types of inaccuracies. Excerpt 4 illustrates how both translation methods engage in under-translation at the word and phrase levels.

Excerpt 4.
English Version:
An audible, rapid "clicking" or buzzing indicates airway obstruction. STOP! Take necessary steps to clear ***airway*****,** then re-position person to continue.

Human Translation:
Suara yang jelas terdengar berupa "klik" yang cepat atau dengungan menandakan ada hambatan jalan udara, BERHENTI! Lakukan yang diperlukan untuk membebaskan jalan udara, kemudian letak-kan kembali orang sakit untuk melanjut-kan.

Google **Translation:**
Suara "klik" yang cepat atau berdengung menandakan obstruksi jalan napas, BERHENTI!, ambil langkah yang diperlu-kan untuk membersihkan jalan napas, lalu ubah posisi orang untuk melanjutkan.

It is important to clarify that Excerpt 4 is taken from a text that provides instructions on the use of an emergency resuscitation kit for a patient. In this instance, the human translator rendered the term *airway* as *jalan udara*, while *Google Translate* produced *jalan nafas* for the same term. *Airway* is classified as a compound noun or noun composed of two independent lexical units: *air* and *way*. Accordingly, *jalan udara* represents a literal translation of the source term. Although the literal translation version may remain intelligible to the readers, *Google Translate* provided a more contextually appropriate alternative, *jalan nafas*, that more precisely captures the intended meaning of *airway* as a breathing passage in English.

Given that the context contained in the presented text is the use of a medical tool on a human patient, thus the translation of any specific term, such as *airway* must represent the context of situations and the participants participating in the conversation. For instance, if *jalan udara* is used as the translation of the given context, casual readers who have a lack of understanding of the resuscitation process might be confused about whether the referent of *jalan udara* is the *the airway system of a medical device and the trachea (or respiratory tract) of the patient*. Such inequivalence can result in ambiguity; however, in this instance, GT successfully avercomes this by employing a more specific and accurate term. The excerpt 5 below shows the under-translation done by GT.

Excerpt 5.
English Version:
To identify factors that influence the clinical response to 5-fluorouracil (5-FU), we ***studied*** the correlation between in vitro sensitivity to 5-FU and the expression of seven biological markers.

Human Translation:
Untuk mengidentifikasi faktor-faktor yang mempengaruhi respons klinis terhadap 5-florourasil (5-FU), kami ***meneliti*** *korelasi antara sensitifitas in vitro terhadap 5-FU dan kadar tujuh indikator biologis.*

Google **Translation:**
Untuk mengidentifikasi faktor-faktor yang mempengaruhi respon klinis terhadap 5-

fluorouracil (5-FU), kami ***mempelajari*** *korelasi antara sensitivitas in vitro terhadap 5-FU dan ekspresi tujuh penanda biologis.*

In Excerpt 5, the human translated verb *studied* as *meneliti*, whereas *Google Translate* produced *mempelajari*. When considered in isolation, there is nothing inherently incorrect about *Google*'s translation, as the verb *to study* and *mempelajari* convey a comparable lexical meaning. However, when the term is situated within a more specific context, the choice of *meneliti* becomes more appropriate as well as acceptable. This underscores the necessity of context-sensitive translation, given that lexical equivalence alone cannot guarantee semantic accuracy.

Excerpt 5 is situated within the context of a scientific research report. As a result, the use of verb *mempelajari* as the direct translation of *studied* is deemed overly general and does not effectively reflect the 'action' intended in this context. In contrast, the human translator's choice of *meneliti,* which corresponds to verb *to research* in English, is considered a more accurate and contextually fitting translation of the source text. This instance of inaccuracy obseverd by GT may stem from its algorithmic tendency to select superordinate terms and its constrained ability to employ hyponyms that align with the contextual meaning. However, it should be acknowledged that the inaccuracies found in the human and GT outputs, as shown in Excerpts 4 and 5, fall under the category of minor and tolerable errors.

Excerpt 6.
English Version:
An *incentive structure* that is attractive for one country but not others is likely to lead to international leakage.

Human Translation:
Sarana perangsang yang menarik bagi sebuah negara, tetapi tidak bagi negara-negara lainnya, boleh jadi menyebabkan kebocoran internasional.

***Google* Translation:**
Struktur insentif yang menarik untuk satu negara tetapi tidak untuk negara lain yang cenderung mengarah pada kebocoran internasional.

In Excerpt 6, the datum show that human tends to use localization strategy by rendering the source term *incentive structure* into native Indonesian words, in this context, *sarana* and *perangsang*. Although this approach illustrates an attempt to localize the phrase, the selected lexical items may not accurately convey the intended meaning. Indonesian includes loan-words *insentif* for *incentive* as well as *struktur* for *structure* that closely align with the source terms in both semantic (meaning) and morphological aspects (form). Moreover, the term *perangsang* is seldom used as equivalent to *incentive* in Indonesian, which compromises the accuracy of the translation and may cause confusion. However, *perangsang* possesses an equivalent lexical meaning of verb *to stimulus* in English. Therefore, infrequent use of such term makes it sound odd to Indonesian natives and effort to localize the terms with inaccurate translation demonstrates an example of over-translation that results in equivalence.

Interestingly, no cases of over-translation by GT were found in this study. This outcome aligns with the initial assumption that GT lacks the flexibility to select alternative expressions that provide refined specificity and contextual alignment (Li, Graesser, & Cai, 2014). Further, the findings indicate that the most common errors produced by GT at the levels of words and phrases are either under-translation and mistranslation.

Initially, one instance of mistranslation by a human translator was identified, where the term/noun *bare metal* was translated literally as *logam telanjang*, or referred to *naked metal* in English back-translation. Nevertheless, this initial identification was subsequently verified through *Google Boolean* searches. It was found that the term *logam telanjang* is frequently employed to describe *the material of a stent, a tubular device temporarily positioned within a blood vessel, canal, or duct to promote healing or relieve blockage*. Consequently, this sample was excluded from the category, resulting in the absence of human mistranslation cases in the dataset.

As predicted, numerous mistranslations produced by GT were found in this study. The following excerpt 7 illustrates an example of mistranslation by GT.

Excerpt 7.
English Version:
The unit will start cycling continually.

Human Translation:
Alat ini akan mulai siklusnya terus-menerus.

***Google* Translation:**
Unit akan mulai bersepeda terus menerus.

Similar to Excerpt 2 above, Excerpt 7 is drawn from the instructions for an emergency resuscitation kit (Oxylator FR-300). Therefore, the term *cycling* in the source text refers to the repeated action conducted by the resuscitation device. In this context, the human translator performed transposition by changing *cycling* from a verb into the noun s*iklusnya*. Whereas, GT tends to retain the original word class and syntactic function of *cycling* in the target text. However, the translation provided by GT in Excerpt 7 provides a clear example of a total mistranslation, rendering *cycling* as *bersepeda* 'riding a bicycle'. This error not only results in semantic inequivalence but is also regarded as humor among professional translators. (Läubli & Orrego-Carmona, 2017).

The goal of equivalent effect is crucial in natural science writings. In order to function coorectly, each text must stand for the same idea in each language and produce the same response. Otherwise, varied interpretations would give rise to confusion and potential loopholes. Therefore, it would be interesting to compare longer stretches of the different texts to see whether, despite careful transalation, any further segments present a different focus in meaning (Munday, 2008, p. 51).

In specialized domains such as natural science and medical translation, the demands for accuracy and equivalence are unnegotiable. The form of inequivalences found in this study highlight the limitations of *Google Translate* in selecting the most contextually accurate lexical items for specialized registers. This finding supports the results of Das et al. (2019), who contend that professionally translated medical guidelines remain indispensable in multiple languages, despite the availability of machine translation tools like *Google Translate*.

Furthermore, this finding suggests that *Google Translate* (GT) is capable of producing grammatically accurate translations, despite not entirely error-free, particularly when the source and target languages exhibit similar syntactic structures. This finding concurs with the results of Groves and Mundt (2015) and may help explain why grammatical errors tend to increase when *Google Translate* is used for language pairs with distinct syntactic patterns, such as English and Chinese, and vice-versa (Li, Graesser, & Cai, 2014; Groves & Mundt, 2015).

The comparative analysis of human and *Google Translate* uncovers both notable areas of convergence and substantial divergences in translational performance. On the one hand, equivalence between HT and GT is particularly pronounced at the grammatical level, notably in sentence structure as well as basic syntactic patterns. In this study, GT demonstrated an ability to reproduce structurally coherent TT when working between languages that share similar word order and grammatical systems, such as English and Indonesian. This finding supports the initial research (*e.g.,* Groves and Mundt (2015) and also Li, Graesser, and Cai (2014)) suggesting that machine translation systems are able to generate acceptable output in language pairs with syntactic compatibility. Moreover, GT often parallels HT in terms of word count and sentence length, indicating a degree of structural equivalence that makes MT suitable for capturing the gist of general or non-specialized content. As proposed by Koller (1995):

> "*The result of a text-processing activity, by means of which a SL text is transposed into a TL text. Between the resultant text in L2 (the TL text) and the source text in L1 (the SL text) there exists a relationship that can be designated as a translational, or equi-valence relation.*" (p. 196)

However, the analysis also underscores critical inequivalences, notably at the lexical and semantic levels. GT frequently struggles with context-sensitive and domain-specific terms, especially within specialized fields such as natural science. These over-, under-, and mistranslation compromise the propositional meaning of the source text and often result in target language outputs that are inaccurate or even misleading. In contrast to human, who

are capable of interpreting nuance, register, and communicative intent, GT lacks pragmatic sensitivity and the contextual awareness vital to accurately convey meaning within context-specific and discipline-oriented content. Thus, while GT may function as a useful aid for initial comprehension, particularly for L2 users, HT remains indispensable.

Conclusion

Although machine translation (MT) is not yet achieved full reliability for producing high-quality translations, it still serves as effective resource to facilitate rapid access to important information, with particular relevance for L2 users. The all findings of this study affirm that *Google Translate*, while occasionally prone to lexical inaccuracies and semantic mismatches, shows proficiency in producing grammatically precise translations, primarily when dealing with languages that have comparable syntactic orders, including English and Indonesian.

This finding directly responds to research problem regarding about how effectively GT maintains equivalence in terms of lexical and grammatical dimensions compared to human translators. Notably, the analysis revealed that GT's output closely mirrors human translation in terms of word count and sentence structure. The evidence supports the view that, although imperfect, machine translation demonstrates promising potential in delivering acceptable translation quality, particularly in structurally comparable language pairs.

Nevertheless, it is posited that the most critical challenge in machine translation lies in semantic inaccuracies arising from improper lexical choices. This limitation chiefly arises from the system's incapacity for determining contextually suitable alternative expressions or domain-specific terms embedded in source text. As such, the nuanced judgment and expert knowledge of human translators is crucial for producing accurate translations. The evolution of machine translation may enable translators to access the essential meaning of the source text, yet it remains incapable of generating a fully polished and publishable translation.

One significant constraint of this research lies in its modest dataset, which may limit the depth and scope of analysis into machine translation (MT) performance. Since machine translation systems are trained on large-scale corpora, rigorous performance assessments must be grounded in reliable quantitative data. Future research should expand and diversify datasets across varied domains and language pairings to underpin more robust performance studies. Ongoing comparative studies of MT releases, supported by a comprehensive class of under-translation, over-translation, as well as mistranslation errors, will reveal trends in performance improvements and also highlight enduring shortcomings. Ultimately, studies of hybrid translation models, which integrating human expertise with machine automation, and reader-centric assessments will help for informing the development of more effective strategies for optimizing translation precision and user satisfaction.

Acknowledgment

We express our heartfelt appreciation to Dr. Adi Sutrisno, M. A., from Universitas Gadjah Mada [UGM), Yogyakarta for his professional advice and expertise in the fields of multimedia translation as well as translation of computer assistance. His great support and motivations have contributed to our comprehension and also scholarly growth in these domains.

References

Angelelli, C. V., & Jacobson, H. E. (Eds.). (2009). *Testing and Assessment in Translation and Interpreting Studies: A Call for Dialogue Between Research and Practice: Vol. XIV*. John Benjamins Publishing Company. https://doi.org/10.1075/ata.xiv

Baker, M. (2018). *In Other Words: A Coursebook on Translation* (3rd ed.). Routledge.

Bassnett, S. (2002). *Translation Studies* (3rd ed.). Routledge.

Bassnett, S. (2005). *Translation Studies* (4th ed.). Routledge.

Castilho, S., Doherty, S., Gaspari, F., & Moorkens, J. (2018). Approaches to Human and Machine Translation Quality

Assessment. In J. Moorkens, S. Castilho, F. Gaspari, & S. Doherty (Eds.), *Translation Quality Assessment: From Principles to Practice* (Vol. 1, pp. 9–38). Springer International Publishing AG. https://doi.org/10.1007/978- 3- 319-91241-7

Catford, J. C. (1965). *A Linguistic Theory of Translation*. Oxford University Press.

Das, P., Kuznetsova, A., Zhu, M., & Milanaik, R. (2019). Dangers of machine translation: the need for professionally translated anticipatory guidance resources for limited English proficiency caregivers. *Clinical Pediatrics*, *58*(2), 247–249. https://doi.org/10.1177/0009922818809494

Givón, T. (1993). *English Grammar: A Function-Based Introduction* (Vol. 1). John Benjamins Publishing Company. https://doi.org/10.1075/z.engram1

Groves, M., & Mundt, K. (2015). Friend or foe? Google Translate in language for academic purposes. *English for Specific Purposes*, *37*, 112–121. https://doi.org/10.1016/j.esp.2014.09.001

Hatim, B., & Munday, J. (2004). *Translation: An Advanced Resource Book*. Routledge.

House, J. (2015). *Translation Quality Assessment: Past and Present*. Routledge.

Jakobson, R. (1959). On linguistic aspects of translation. In Brower R. (Ed.), *On Translation* (pp. 232-239). Harvard University Press. https://doi.org/10.4159/harvard.9780674731615.c18

Kim, W. D. (2019). *Translations in Korea: Theory and Practice*. Springer.

Koller, W. (1979). *Einführung in die Übersetzungswissenschaft*. Quelle and Meyer.

Koller, W. (1995). The concept of equivalence and the object of translation studies. *Target: International Journal of Translation Studies*, 7(2), 191–222. https://doi.org/10.1075/target.7.2.02kol

Larson, M. L. (1984). *Meaning-Based Translation: A Guide to Cross-language Equivalence*. University Press of America, Inc.

Läubli, S., & Orrego-Carmona, D. (2017). *When Google Translate Is Better Than Some Human Colleagues, Those People Are No Longer Colleagues*. https://doi.org/10.5167/UZH-147260

Li, H., Graesser, A. C., & Cai, Z. (2014). Comparison of Google translation with human translation. *Proceedings of the Twenty-Seventh International Florida Artificial Intelligence Research Society Conference*, *6*, 190–195.

Lommel, A. (2018). Metrics for translation quality assessment: a case for standardising error typologies. In J. Moorkens, S. Castilho, F. Gaspari, & Doherty, Stephen (Eds.), *Translation Quality Assessment: From Principles to Practice* (Vol. 1, pp. 109–128). Springer International Publishing AG. https://doi.org/10.1007/978- 3- 319-91241-7

Munday, J. (2009). *The Routledge Companion to Translation Studies* (J. Munday, Ed.; p. 252). Routledge.

Munday, J. (2008). *Introducing Translation Studies: Theories and Applications* (2nd ed., p. 236). Routledge.

Nababan, M. R. (2016). *Teori Menerjemah Bahasa Inggris* (4th ed.). Pustaka Pelajar.

Nababan, M. R., Nuraeni, A., & Sumardiono. (2012). Pengembangan model penilaian kualitas terjemahan. *Kajian Linguistik Dan Sastra*, *24*(1), 39–57. https://doi.org/10.23917/kls.v24i1.101

Newmark, P. (1988). *A Textbook of Translation*. Prentice Hall International vUIO Ltd.

Nida, E. A. (1964). *Toward a Science of Translating: With Special Reference to Principles and Procedures Involved in Bible Translating*. E. J. Brill.

Nida, E. A., & Taber, C. R. (1974). *The Theory and Practice of Translation*. E. J. Brill.

Nida, E., & Taber, C. (1974). *The Theory and Practice of Translation*. E. J. Brill.

Nord, C. (1997). *Translation as a Purposeful Activity, Functionalist Approaches Explained*. St. Jerome Publishing.

Panou, D. (2013). Equivalence in translation theories: A critical evaluation. *Theory and Practice in Language Studies*, *3*(1). https://doi.org/10.4304/tpls.3.1.1-6

Pym, A. (2004). *The Moving Text: Localization, Translation, and Distribution*. John Benjamins.

Sutrisno, A. (2020). The accuracy and shortcomings of Google Translate

translating English sentences to Indonesian. *Education Quarterly Reviews*, *3*(4). https://doi.org/10.31014/aior.1993.03.04.161

Verhaar, J. W. M. (2016). *Asas-Asas Linguistik Umum* (9th ed.). Gadjah Mada University Press.

Vinay, J. P., & Darbelnet, J. (1955). *Comparative Stylistics of French and English: A Methodology for Translation*. John Benjamins.

Williams, J., & Chestermen, A. (2002). *The Map: A Beginner's Guide to Doing Research in Translation Studies* (2nd ed.). St. Jerome Publishing

Williams, M. (2004). T*ranslation Quality Assessment: An Argumentation-Centered Approach.* University of Ottawa Press.

Wills, W. (1982). *The Science of Translation: Problems and Methods*. Gunter.