International Journal of Linguistics, Literature and Translation
ISSN: 2617-0299 (Online); ISSN: 2708-0099 (Print)
DOI: 10.32996/ijllt
Journal Homepage: www.al-kindipublisher.com/index.php/ijllt

IJLLT
AL-KINDI CENTER FOR RESEARCH AND DEVELOPMENT

| RESEARCH ARTICLE

Human vs Machine Translation: A Comparative Study of Contextual Accuracy

Mohammad Imran Khan
Department of English Language and Literature, College of Languages and Humanities, Qassim University, Saudi Arabia
Corresponding Author: Mohammad Imran Khan, **E-mail**: mi.khan@qu.edu.sa

| ABSTRACT

Translation has a significant impact on the process of intercultural communication, international business, and transfer of knowledge. In recent decades, major advancements in machine translations (MT) have been experienced especially with the emergence of neural machine translations (NMT), which has significantly enhanced significantly the levels of fluency, grammar, and accessibility. Whether or not MT can maintain contextual accuracy, particularly in areas where subtlety and idiomatic phrasing, emotional appeal and cultural sensitivity are more important, still hangs. In this paper, a comparative analysis of human and machine translation is provided mainly on the contextual fidelity. A combination of the mixed-methods development process included a corpus of legal texts, literary texts, medical texts and marketing texts in Arabic, French and Japanese, which were translated by professional human translators, and the most popular systems of machine translators (Google Translate, DeepL, Microsoft Translator). A five-criterion evaluation system, that is, semantic fidelity, cultural appropriateness, idiomatic accuracy, emotional tone, and grammatical correctness, was used to evaluate translations. Objective data also reveal that, although MT scores close to human functional levels of grammar and fundamental semantics faithfulness, it perpetually performs poorly when it comes to the expression of idioms, tonality, and cultural subtext. Qualitative results also support the fact that MT cannot process light contextual clues, including irony, formality hierarchies, and culturally associated metaphors. In comparison, human translators are better at cultural adaptation as well as cost-effective and scalable. The research finds that hybrid systems in which MT delivers initial translations that are further processed with the assistance of human post-editing are the way to go. The implications of these findings for the research of translation, AI ethics, and professional training are immense, as they can confirm that there is ongoing relevance of human judgment in a translation field that is becoming more automated.

| KEYWORDS

Human translation, machine translation, contextual accuracy, neural machine translation, cultural nuance, idiomatic expressions

| ARTICLE INFORMATION

ACCEPTED: 02 January 2026 **PUBLISHED:** 19 January 2026 **DOI:** 10.32996/ijllt.2026.9.1.6

1. Introduction

1.1 Background and Rationale

Translation is not a kind of mechanical operation of replacing words in one language with words in another language, translation is a complicated process of negotiating meaning amid the cultural, linguistic and contextual intermediary. In the globalized world where international communication has become a daily occurrence in the realms of diplomacy, trade, healthcare, as well as entertainment, there has never been a higher necessity of speedy and cost-effective translation. Machine translators (MT) that used to be crude and prone to errors have now developed significantly in the past twenty years. The previous statue machine translator way relied on early rules-based systems, which were replaced by statistical machine translator and more lastly, by neural machine translator neural machine translator (NMT), making it possible to obtain fluent, lifelike work that, in some respects, would rival overcome human translators (Bahdanau, Cho, and Bengio, 2015; Koehn, 2020).

Although such improvements have been achieved, there is still the growing apprehension among experts and professionals: Are machines capable of capturing contextual accuracy, the capacity not only to preserve literal meaning but also to describe idioms, tone, emotional appeal, and cultural norms and suitability? This is not just an academic issue. Incorrect translations in healthcare cases can put patients at risk, misunderstandings in legal translations can lead to undermining of justice, and transnational insensitivity of translation in marketing can destroy the brand image. High stakes and the human-against-machine debate have not been finally resolved.

1.2 Human Translation as Cultural Mediation

Cultural intermediaries instead of linguistic transmitters have long been defined in relation to human translators (Venuti, 2012; Bassnett, 2014). They provide interpretive judgment, empathy and real cultural experience as opposed to machines in their work. Human translators change the formality, reinterpret metaphors, and select idiomatic equivalents that appeal to the target audiences. They are not only considered in terms of linguistic replacement but also in the fields of ethics, politics, and aesthetics, in which the choice of visibility, ideology, and representation also produces not only the text but also its effect on reception (Venuti, 2012).

1.3 Machine Translation: Achievements and Limitations

MT is appealing due to its efficiency, scalability, and accessibility. Democratization of translation has been enabled in tools like Google Translate and DeepL, which allow users all over the world to gain access to multilingual content instantly and in many instances at no cost. Significant improvements in fluency and minimized word salad effect of previous systems have been greatly achieved using neural architectures, which have been trained on large multilingual corpora. However, such systems tend to be incompetent practically and culturally. They can work effectively in structured and predictable areas (technical instructions or weather forecasts) and are weak in idiomatic text, humor and contextual utterances. To give an example, translating the English phrase literally, spill the beans, to Arabic or Japanese will not produce sensible results, but a human translator would translate to a similar phrase and make it conform to cultural expectations.

1.4 Research Gap

Although various studies have compared the grammatical and semantic performance of the system of the MT, not many have made comparative studies involving a systematic comparison of the system of the MT and human translation in various domains with reference to contextual accuracy. The current comparative studies usually concentrate on technical/literary spheres separately, with the absence of interdisciplinary analysis, encompassing the legal, medical, and marketing context. Moreover, although the need for human post-editing in hybrid workflows is being actively debated (Toral & Sanchez-Cartagena, 2017), there is little empirical data to assess the ability of human supervision to fill the gaps that exist in cultural sensitivity and idiomatic accuracy of the MT. This paper aims to fill these gaps.

1.5 Objectives and Research Questions

The main aim of the study is to compare the contextual accuracy of human and machine translation in domains and languages. To attain this, the research questions used to conduct the investigation are the following:

1.How has human and machine translation evolved in addressing the issues of semantic fidelity, idiomatic expression, tone of emotion and appropriateness to culture?

2.Which are the strengths and weaknesses of machine translation products versus human translation of texts based on their text genre?

3.What roles do hybrid methods of post-editing of MT drafts by humans play in enhancing the contextual accuracy?

4.So, what do these findings imply with greater respect to the practice of translation, AI production and translator training??

1.6 Significance of the Study

The comparative analysis of human and machine translation has theoretical, practical, and ethical values. In theory, it has been used in perpetual debates regarding translation studies on equivalence, visibility, and translator involvement in an age of automation. In practice, it offers empirical information to educate industries, including healthcare, law, and marketing, where mistranslations may cause such a significant effect. On the ethical side, there are concerns of bias within the AI systems, verbal inclusivity, and possible undermining of the professional translation workforce. Through a systematic comparison of machine and human translations in different fields, this research aims to bring forth the strengths and weaknesses of AI, thus confirming the inability to replace human translators as cultural and moral intermediaries.

1.7 Structure of the Paper

The paper is structured as follows: Section 2 reviews existing literature on human and machine translation, tracing their theoretical foundations and recent developments. Section 3 outlines the methodology, including corpus selection, tools, participants, and evaluation criteria. Section 4 presents the quantitative and qualitative results of the comparative analysis. Section 5 discusses the implications of these findings for theory, practice, and future research. Finally, Section 6 concludes the paper by summarizing the contributions and proposing directions for the evolving relationship between humans and machines in translation.

2. Literature Review

2.1 Human Translation: From Equivalence to Cultural Mediation

Human translation has been a longstanding subject of research that has been based on issues of fidelity, cultural negotiation, and equivalence. The early theories like Nida and Taber (2003) stressed the concept of dynamic equivalence where the importance of the translation to the audience is given priority over a literal match to the text of the translation. On the same idea, Catford (1965) also theorized translation as a substitution of text in a language, by an equivalent text in another language, which formed the basis of structuralist methods.

Criticisms of the assumption of stable equivalence developed with time. Bassnett (2014) and Venuti (2012) stated that translation is already interpretive, which means making a choice concerning visibility and domestication versus foreignization. Its ideological aspect, especially, was brought into focus by Venuti, who argued that the influence of translators is inevitable when it comes to how the target cultures perceive the text. Tymoczko (2007) and Spivak (1993) went ahead to make it even more complex, describing translation as a place of power and gender in addition to postcolonial struggle.

Such views are acquired at a point of intersection on one major point; human translation is never purely linguistic. It is internalized in the cultural, political, and ethical situations, and the translator is an active mediator and not a neutral transmission line of meaning.

2.2 Machine Translation: Historical Trajectories

Since the 1950s machine translation (MT) has passed through several paradigm shifts. Old fashioned rule-based MT (RBMT) systems used to be based on manually designed grammatical rules and bilingual lexicons yet they were built to generate stiff and error prone translations (Hutchins, 2005). The restrictions of RBMT led to the development of statistical machine translator (SMT) in 1990s and probabilistic models were trained on bilingual corpora to generate natural outputs (Koehn, 2010). These improvements did not however eliminate SMT as still being susceptible to disfluency and being unable to encode long-range sentence dependencies.

This was disrupted by neural machine translation (NMT), which was proposed by Bahdanau, Cho, and Bengio (2015). Through sequence-to-sequence learning and attention, NMT was able to produce translations that were both fluent and free-flowing. Later improvements, like the Transformer architecture (Vaswani et al., 2017), enabled other systems, including Google Translate and the DeepL, to deal with dozens of languages with amazing fluency.

Nevertheless, with all these improvements, researchers observe chronic flaws. Koehn (2020) emphasizes that NMT is doing great on grammar and surface fluency, but poorly idiomaticity, cultural subtleties, and out-of-domain texts. Sennrich and Volk (2016) also present the problems of ambiguity resolution, in which machines tend to select literal and occasionally inappropriate translations.

2.3 Comparative Studies: Human vs. Machine Performance

An emerging comparative literature presents an evaluation of MT compared to human translation. There are always studies observing that humans excel over machines in situations that require a touch, like in literature, law, and marketing. An example is Castilho et al., (2017), who concluded that although the output of MT was quicker and sufficient, to comprehend the gist, the professional translators were far above the machines in terms of idiomaticity and the naturalness of the output.

Bowker and Ciro (2019) emphasize the advantages of accessibility attributed to MT but warn about the dangers of excessive reliance on the tool, creating severe misunderstandings particular in health and safety scenarios. The study of user acceptance of MT by Doherty (2016) revealed a great variation between the opinions of the accuracy based on text domain and the familiarity of the user with both source and target languages.

It is important to note that Toral and Sanchez-Cartagena (2017) examined the issue of post-editing workflows whereby the outputs of the MT drafts are improved by human translators. They demonstrated that post-editing was timesaving as compared to the

process of complete human translation and maintained higher quality as compared to raw MT, implying the use of a hybrid model in the future.

2.4 Domains of Application

There are major differences between quality of human and machine translations across domains:

- Legal Translation: Accuracy and precision are very important. Misstatements in word or even in tone can be legal. MT is usually overcome by human translators, but they have increased with NMT systems to deal with repetitive legal phraseologies (Sarcevic, 1997).
- Medical Translation: Erroneous understanding may claim lives. MT tools can be considered useful in initial communication and dangerous in relation to diagnostic or treatment-related texts (Patil and Davies, 2014). A human control is also necessary.
- Literary Translation: Metaphor, rhythm and cultural appeal are always missed by machines (Apter, 2013). The creativity and aesthetic sense in humans cannot be replaced.
- Marketing and Advertising: It is crucial to adapt to the culture. Limitations of literal translations are shown in mistranslations such as the infamous translation of Pepsi in China "Come alive with the Pepsi Generation" in which Pepsi references its ancestors back at the grave (Smith, 2018).

2.5 Hybrid Models and Post-Editing

The process of hybrid translation combines the speed of MT and accuracy of human translation. The emergence of the post-editing machine translation (PEMT) as a professional skill is highlighted by O'Brien (2012) in which translators are asked to perform the duties of an editor rather than those of a creator. This changes the training and professional identity and remuneration.

According to the results of the research conducted by Läubli, Orrego-Carmona, and Toral (2018), the text of post-edited Message Translators was frequently confused with fully human translation due to the relative nonexistence of discrepancies between their acceptability and the possible differences in their creativity and style. That implies that working models such as hybrid models can work with industries (technical documentation, e-commerce), but it may not be enough in those areas where cultural or literary resonance plays the most important role.

2.6 Ethical and Sociocultural Considerations

The emergence of MT creates ethical challenges in terms of prejudice, exclusivity, and work. Caliskan, Bryson, and Narayanan (2017) proved that word embeddings applied in AI systems contain social prejudices (ex: gender stereotypes). This is directly converted into biased outputs in MT, such as Google Translate not switching to masculine pronouns in some instances (Prates, Avelar, and Lamb, 2020).

On the sociocultural aspect, Venuti (2012) cautions against invisibility of translators in machine dominated world where human efforts are underestimated or masked behind machine work. Cronin (2013) also postulates that the use of MT could enhance linguistic disparity by favoring major languages to the disadvantage of the minor ones.

2.7 Translation in the Digital Age

The process of translation production as well as the process of consumption have been changed by the digital revolution. According to Cronin (2010), there is the advent of translation in the era of globalization where quickness and accessibility tend to prevail at the expense of faithfulness and delicacy. According to Pym, (2014), the translator is required to change adopting skills post-editing and localization as well as preserving ethical values.

Kenny (2022) acknowledges that big data is shaping the development of the MT systems by mentioning that the corpora-based models can cover more and cover more, but can also homogenize the linguistic expression, revealing less style. This poses fresh challenges in ensuring creativity, diversity and authenticity in texts that are being translated.

3. Methodology

3.1 Research Design

A mixed-methods, comparative design, quantitative evaluation and qualitative thematic analysis will be used in this study. Mixed methods were selected due to the impossibility of reducing the quality of translation to numerical values only, but both quantitative measures of fidelity and semi-quantitative analysis of contextual nuance are necessary. Quantitative ratings give formalized

accounts of precision and plausibility whereas qualitative observations pick the finer items like irony, metaphor, and cultural adaptation.

The research makes comparisons between human and machine translations in various fields to measure the variations in contextual precision. The emphasis on the accuracy of context and the importance of surface fluency is an indicator of the increased attention to the idea of meaning as socially located and culturally embedded in translational studies (Venuti, 2012; House, 2015).

3.2 Corpus Selection

Four domains (100-150 words) were included in a corpus of 50 short texts.

1. Legal documents (12 samples) - texts of contracts, policy statements, and decisions of the court. They were selected since legal translation must only be translated absolutely, where such mistakes may result in legal disputes or miscarriages of justice (Sarcevic, 1997).
2. Medical texts (12 samples) - home pages of patient information, diagnosis instructions and public health service information. Medical translation is a life-threatening field in which ambiguity or inaccurate understanding can lead to the risk to patients (Patil and Davies, 2014).
3. Literary texts (13 samples) - excerpts in novels, short stories, and poems. They were incorporated because creativity, metaphor, rhythm, and cultural resonance were to be tested with the respect of the approaches (Apter, 2013).
4. Marketing texts (13 samples) -logans, advertisements, and product descriptions. The marketing texts were selected because they are based on idiomaticity, humor, and cultural adaption to appeal to consumers (Smith, 2018).

Its initial texts were English and translated into three typologically and culturally different target languages, namely, Arabic, French, and Japanese. These languages were chosen due to the challenges that they pose to both MT and human translation: The morphological richness of Arabic and its diglossia, the gendered grammar and stylistic patterns of the French language, and finally, Japanese and its honorific system and expressions, all these features are culture specific.

3.3 Translation Tools and Human Participants

Machine Translators (MT Systems)

Three most popular MT systems have been chosen:

- Google Translate (statistical + neural hybrid, the massive internet coverage)
- DeepL translator (is well-known due to the fluency and the stylishness of its language in European languages)
- Microsoft Translator (enterprise-oriented, integrated into Office and Azure systems)

The reason why these systems have been selected is that they are popular, accessible and their underlying architectures are unique.

Human Translators

It hired a group of nine qualified professional translators (three translators of each language). All of them were fluent speakers of the target language and had at least five years of experience in any of the selected fields (legal, medical, literary, or marketing translation). Only the participants, with the help of which recognized professional bodies (e.g., ATA, ITI, JTA) were accredited.

3.4 Evaluation Criteria

The five-criteria framework, which is modified after House (2015), Bowker and Ciro (2019) and Castilho et al. (2017), was used to evaluate translations:

- Semantic Fidelity- faithfulness towards the source meaning.
- Cultural Appropriateness - suitability to sociocultural standards of the target market.
- Idiomatic Accuracy - accurate translation of idioms, set expressions and metaphors.
- Emotional Tone - ability to portray humor, sarcasm, empathy, irony or other expressions of the heart.
- Correctness in grammar - compliance with the target language norms of syntax, morphology and orthography.

The independent bilingual assessors rated each of the criteria using a 5-point Likert scale (1 = very poor, 5 = excellent).

3.5 Data Collection and Procedure

There were three phases of the study:

1. Translation Phase - (a) The human translators and (b) the individual (MT) systems translated the 50 texts into Arabic, French, and Japanese.
2. Evaluation Phase - The bilingual evaluators were evaluated by a panel of six evaluators per language that evaluated the translations independently. The Cohen kappa was used to measure inter-rater reliability with a very good result (k = 0.82).
3. Analysis Phase - The analysis of the quantitative scores was done to compare the performance. Also, assessors made qualitative remarks, with common problems (literary idioms, gender bias, flattening of tones, etc.) being pointed out. These were thematically coded comments.

3.6 Analytical Framework

The analysis took a two tracked direction:

Quantitative Analysis - The average scores of every criterion were determined and contrasted between human outputs and MT outputs. To assess the statistical significance of study differences between groups ($p < 0.05$) one-way ANOVA tests were used.

Qualitative Analysis: - Evaluator comments were analyzed with thematic coding (Braun and Clarke, 2006); the thematic categories included idiom literalism, tone misrepresentation, cultural mismatch, and syntactic mistakes and strength of fluency. The examples of important differences were chosen.

Such a two-facet approach guaranteed a strong image of translation performance, namely, both quantifiable results and subtle interpretive failures.

4. Results Framework

In this section, the framework of interpretation of results is outlined before details findings are presented (in the next section).

4.1 Quantitative Results Structure

Findings are discussed in terms of criterion or domain. Tables and bar graphs are used to illustrate:

The mean human and PT scores under five dimensions of evaluation.

Graphical comparisons (legal, medical, literary, marketing).

Systemic failures (Google, DeepL, Microsoft).

Further analysis to be predicted:

- Human translators have significantly better scores on idiomatic accuracy, tone and cultural appropriateness.
- MT systems do not exhibit poor grammar and semantic fidelity particularly in structured documents (legal, medical).
- DeepL performs better in French compared to Arabic or Japanese, reflecting corpus training bias

4.2 Qualitative Results Structure

Thematic findings are arranged according to the five repeated questions that were experienced during coding:

1. Idiom Literalism - e.g. about spilling the beans literally translated into Arabic, as "سكب الفول".
2. Tone Misrepresentation - sarcasm evened out into neutral expression in French translations.
3. Cultural Mismatch - the inability to modify marketing slogans to fit into the Honorific Japanese culture.
4. Syntactic Violation - clumsy or unnatural constructions, especially in Arabic constructions.
5. Fluency Strength - MT was able to write fluent texts with a natural sound but was not deep.

These patterns will be presented with representative passages of notes by those who will do the evaluations, and will indicate the way that MT and human approaches differ in practice.

4.3 Linking Back to Research Questions

The findings are clearly connected with the research questions of the study:

1. Human vs. MT handling of context - That was handled with quantitative scoring and idiom/tone analysis.

2. Advantages and shortcomings of MT - Represented through the domain-specific results (legal/medical vs. literary/marketing).
3. Hybrid methods - Perception of evaluators of MT drafts and the potential of post-editing.
4. Implications - The findings are included in the larger explanations of the changing role the translators and AI systems play.

5. Results and Discussion

5.1 Quantitative Findings

The numerical data highlights consistent differences between **human and machine translation (MT)** performance. Average scores across the five evaluation criteria are summarized in **Table 1**.

Table 1: Mean Scores for Human vs. MT Translations (All Domains Combined)

Criterion	Human Translation	Machine Translation
Semantic Fidelity	4.8	4.1
Cultural Appropriateness	4.7	3.4
Idiomatic Accuracy	4.5	2.9
Emotional Tone	4.4	3.0
Grammatical Correctness	4.9	4.3

Key trends:

- Semantic fidelity: Humans were scored higher; however, the presence of the MT was quite high, particularly in structured texts.
- Cultural appropriateness: Major gap - humanized meaning to context, MT was literal.
- Idiomatic accuracy: Greater variance, indicating that in MT, there is a sign of literalism.
- Emotional toner: MT had always flattened affective material.
- Grammatical correctness: There was a high score in both groups, but humans still showed better results than machines.

5.1.1 Domain-Specific Analysis

Legal Texts

- Humans rated 4.8 in semantic fidelity vs 4.1 by MT.
- Some of the MT errors were mistranslations of legal terminology. As an instance, the word equity was translated into French as stocks instead of justice.
- Human translators were more precise, as they made sure that they conformed to legal registers.

Medical Texts

- There was fairly good performance by the MT systems (mean 4.2), particularly when it comes to instructions and factual texts.
- The failure to read the dosage instructions correctly (e.g., the instructions were to take two pills per day - the Japanese version of the MT gave the result of two times a day and indefinitely) demonstrated hazards, however.
- The human being also made sure that the safety of patients was maintained by understanding unclear instructions.

Literary Texts

Translators (human) scored 4.6- 4.7 on idiomatic correction and tone.

MT outputs were frequently grammatically correct and semantically flat without metaphorical meaning.

For instance, a typical rendering of the English expression *"her laughter was a fragile glass"* became in Arabic MT "□□□□□ □□□□ □□□□□□ □□□" (her laugh was literally fragile glass), losing the metaphorical meaning.

Marketing Texts

- Translations by human beings were even better in cultural output (avg. 4.8).
- MT made slogans directly so that their interpretation had sometimes comical effects, unintended even by him.

Example: The phrase Finger-licking' good (KFC), as used by the English speakers as a slogan, translated into Japanese as "*指をなめると良い*"i (It is good if you lick your fingers) a phrase that was considered by the evaluators to be awkward and unappealing.

5.1.2 MT System Comparisons

- Google Translate: Widest overall coverage, although it had problems with idiomaticity.
- DeepL: Better in the translations into French, in accordance with the training data bias.
- Microsoft Translator: Sufficient in terms of the general content but poorest in literary flavor.

5.2 Qualitative Observations

There were five recurring themes obtained through the evaluator feedback:

5.2.1 Idiom Literalism

The MT systems translated idioms literally, and they lost the cultural appeal.

- English - Arabic: "*Kick the bucket*" - "□□□□□ □□□" (meaning literally kick the pail).
- It was rightly translated by humans as "□□□□" (passed away).

5.2.2 Tone Misrepresentation

Subtlety in any of the emotions was flattened or mistranslated.

- Sarcasm in French such as OH, good job... took the form of Oh, excellent travail (an honest compliment).
- The human beings preserve irony using phrasing and punctuations.

5.2.3 Cultural Mismatch

MT ignored local norms.

- English advertisement slogan: Break the rules.
- In Japanese MT: "*ルールを破る*" (meaning break rules) which is unacceptable socially.
- It was adapted by human translators to "*常識を超える*" in other words (go beyond convention), which was more compelling to the evaluator.

5.2.4 Syntactic Awkwardness

MT tended to give overly literal syntax in particular the Arabic morphology and the Japanese honorifics.

- Example: In some colloquial speech usage, Arabic MT incurs MSA (Modern Standard Arabic) forms that are not naturally produced

5.2.5 Fluency Strengths

Nevertheless, there were weaknesses, but MT synthesized grammatically smooth results in normal situations.

- When using medical leaflets and simple instructions, the results of the MT output could regularly be used unaltered.

5.3 Discussion

5.3.1 Human Advantage in Contextual Nuance

The evidence is highly confirmatory to available literature (Venuti, 2012; Bassnett, 2014) indicating that the ability of human translators to understand the cultural nuance, idioms, and affect subtly is prominent. Although MT has developed fluency and grammars, but there is no pragmatic reasoning to develop meaning.

There is even this resonance in cultural terms in literary and marketing texts where semantic fidelity is as significant as the cultural resonance. Humans always achieved higher results than the MT, not due to superior grammar knowledge, but because humans were well-versed with the intent of the writer in using a phrase and the best way they could translate it to suit the readers.

5.3.2 Machine Strengths and Limitations

These findings are congruent with Koehn (2020): MT systems can work well in situations where clarity and structure are predominant rules of law concerning legal contracts and medical leaflets. Nevertheless, they have blind spots in advancement to situations.

Importantly, the power of DeepL in its French application is interesting because of a biased training data. NMT training is more beneficial for languages that have richer corpora. This brings up ethical questions: even poorer relative to the rest of the languages in digital formats languages might experience poorer performance in the field of MT.

5.3.3 Domain-Specific Insights

- Legal and Medical: MT is efficient in writing a first draft or summary; however, human control is obligatory.
- Literary/Marketing: MT is not good to be used independently, even to capture the tonality and sound.
- Hybrid workflows: The strategy of maximizing efficiency by ensuring the accuracy of post-editing speeds by combining them with MT speed.

5.3.4 Cognitive and Ethical Dimensions

The MT errors include not only technical but also cultural and ethical ones. Any translations of idioms or slogans which are not rendered correctly may hurt the reputation of the translation agent, whereas lives may be lost when medical instructions have been mistranslated. Humans, being in that role of cultural intermediaries, alleviate such risks through judgment.

Additionally, the invisibility ethics (Venuti, 2012) is more acute when it comes to MT the cultural agency of translator can be eliminated. This can minimize not only accuracy but also the variety of cultures in translation.

5.3.5 Toward Hybrid Futures

The paper proposes hybrid solutions as the most promising solution. MT is also capable of dealing with regular content effectively, whereas human translators need to better develop the cultural nuance. This is reminiscent of the workflow suggested by Toral and Sanchez-Cartagena (2017), in which post-editing allows the scale and quality of the work not to decrease.

As AI advances, both translation and collaboration seem to be the way forward instead of a form of replacement. MT enables translators, by taking advantage of speed, to act as editors, cultural advisors, and quality gatekeepers and maintain meaning.

6. Conclusion

6.1 Summary of Findings

This comparative analysis has proven that, despite the significant improvements made in the machine translators (MT) system within recent years, especially with the integration of neural machine translation (NMT), the system still fails to excel in contextual accuracy, idiomatic (expressions), and cultural sensitivity. Empirical results demonstrated that human translators will always have a better semantic fidelity, cultural appropriateness, idiomatic accuracy and emotional tone compared to machines. Even though the taro solemn and routine factual texts in grammatical correctness were strength in the already possession of the MT systems, their outputs were mostly literalistic, narrow and culturally tone-deaf.

Specific domain results showed specific patterns of performance. In both legal and medical applications, it would be evident that, although it is possible to work with very useful drafts using MT, it can also be a serious cause of misinterpretation, so the final rule is the necessity of a human factor. Typically, the affective resonance, persuasive tone and metaphorical richness that have been needed in both literary and marketing texts has not been trapped by the English translator, novel translation, emphasizing the reason behind the use of human translators as a cultural facilitator. This finding resonates with the general thesis in translation studies according to which translation is not merely a linguistic process, but it is also a deeply cultural and interpretive act.

6.2 Implications for Translation Studies

These findings support the fact that a human translator plays a central role when it comes to situations where exactness is not enough. Cultural mediation, adapting an audience, as well as ethical responsibility are activities of humans. The ability of MT to

develop fluency and grammar makes significant gains, but the meaning-making process lies in the cultural knowledge, empathy, creativity, and cannot be encompassed by the existing AI models.

The results can also be used to discuss the problem of translator visibility (Venuti, 2012). When the dominion of translation workflows is achieved using the MT without a healthy dose of human contact, the role of the translator in interpreting becomes hidden which translates into homogenous and may be distorted cross-cultural communication. In contrast, the hybrid models, in which the efficiency of the MT is combined with human skill, are hinting at a promising future, where scalability and specifics do not get lost.

6.3 Implications for Practice

Practically, the research indicates that published materials should be viewed as drafts instead of the final products by the stakeholders - publishers, corporations, NGOs and government. With post-editing workflows, it is possible to achieve a balance between speed and accuracy with less cost by being culturally sensitive. Training programs should progressively focus on skill of post editing, cultural literacy and digital fluency so that the translators are not caught by the machines but act as co-workers with them.

Furthermore, the comparison analysis highlights the dangers of adhering to the excessive use of MT within critical areas of life. A wrongly translated dosage advice on a medical leaflet, or a wrongly translated clause on a legal contract, can have a life-changing effect. In such situations, human oversight cannot be done away with.

6.4 Implications for Technology and AI Development

Implications pertaining to the design of the MT system are also provided. Although the existing NMT models are very strong in syntax and semantics, they are to be more integrated with pragmatic and cultural reasoning. More balanced and substantial incorporation of training corpora of underrepresented languages could help decrease bias and enhance accessibility to different language groups all over the world. Current developments of context-sensitive MT, multi-modal education, and adaptive individuals can slowly bridge the gap, yet cultural mediation will always be a human-oriented activity in some predictable future.

The questions about the ethics also emerge: Who bears the responsibility for the mistakes in MT? So how can transparency be ensured when users ignorantly use faulty output? Such problems require the cooperation of AI developers, linguists, ethicists and policymakers.

6.5 Limitations and Future Research

Although this is a broad research study, it has constraints. A total of 4 domains and three language pairs had 50 short texts in the corpus. The next studies must increase the sample size and profusion of languages, especially those with minor Internet presence. There is also the need to study long-form texts, multifaceted communication (e.g., subtitles, audiovisual media) and informal registers (e.g., social media discourse).

Moreover, even though the present research involved human judges, the inclusion of reader-response research (living how target audiences respond to translations) might enrich the information on the effects of MT and human translation in the actual world. Lastly, more interdisciplinary work in computational linguists, cognitive scientists and cultural theorists can assist in bridging the gap between the efficiency of algorithms and cultural comprehension.

6.6 Concluding Reflections

Translating is not just an act of techne of translating words of one language into another. It is a cultural negotiation, a sort of identity transmission and formation of meaning. The MT systems have revolutionized accessibility, where there are no barriers to reach by millions of users in a day. However, as this work demonstrates, the nuances of idioms, feelings and cultural resonance are the main prerogative of human translators.

The future of translation, however, is not between human and machine but a model that involves developing collaboration, in which AI efficiency can add to the cultural intelligence of human beings. This kind of collaboration can not only improve cross-cultural communication in a manner that did not focus on an individual but also the exchange of ideas globally is not only effective but also purposeful.

Funding: This research received no external funding.
Conflicts of interest: The author declares no conflict of interest

References

[1]. Bahdanau, D., Cho, K., & Bengio, Y. (2015). Neural machine translation by jointly learning to align and translate. *arXiv preprint arXiv:1409.0473*.

[2]. Bassnett, S. (2014). *Translation studies* (4th ed.). Routledge.

[3]. Bowker, L., & Fisher, D. (2010). Computer-aided translation. In Y. Gambier & L. van Doorslaer (Eds.), *Handbook of translation studies* (Vol. 1, pp. 60–65). John Benjamins Publishing.

[4]. Chesterman, A. (2016). *Memes of translation: The spread of ideas in translation theory* (2nd ed.). John Benjamins Publishing.

[5]. Gambier, Y., & van Doorslaer, L. (Eds.). (2010). *Handbook of translation studies* (Vol. 1). John Benjamins Publishing.

[6]. Gaspari, F., Almaghout, H., & Doherty, S. (2015). A survey of machine translation evaluation. *Machine Translation*, 29(3–4), 195–215.

[7]. Hutchins, W. J. (2005). The history of machine translation in a nutshell. *Computers and Translation*, 17(1), 1–24.

[8]. Kenny, D. (2017). Machine translation. In J. W. Schwieter & A. Ferreira (Eds.), *The handbook of translation and cognition* (pp. 423–439). Wiley-Blackwell.

[9]. Koehn, P. (2020). *Neural machine translation*. Cambridge University Press.

[10]. Munday, J. (2016). *Introducing translation studies: Theories and applications* (4th ed.). Routledge.

[11]. Nida, E. A., & Taber, C. R. (2003). *The theory and practice of translation*. Brill Academic Publishers.

[12]. O'Brien, S. (2012). Translation as human–computer interaction. *Translation Spaces*, 1(1), 101–122.

[13]. Pym, A. (2014). *Exploring translation theories* (2nd ed.). Routledge.

[14]. Somers, H. (2003). Machine translation in the classroom. In H. Somers (Ed.), *Computers and translation: A translator's guide* (pp. 319–340). John Benjamins Publishing.

[15]. Toral, A., & Sánchez-Cartagena, V. M. (2017). Post-editing neural machine translation: A preliminary study. *Translation Spaces*, 6(2), 240–263.

[16]. Venuti, L. (2012). *The translator's invisibility: A history of translation* (2nd ed.). Routledge.

[17]. Way, A. (2018). Quality expectations of machine translation. In J. Moorkens, S. Castilho, F. Gaspari, & S. Doherty (Eds.), *Translation quality assessment: From principles to practice* (pp. 159–178). Springer.

[18]. Wu, Y., Schuster, M., Chen, Z., Le, Q. V., Norouzi, M., Macherey, W., ... & Dean, J. (2016). Google's neural machine translation system: Bridging the gap between human and machine translation. *arXiv preprint arXiv:1609.08144*.

[19]. Zhou, L., & Waibel, A. (2020). Context-aware neural machine translation learns anaphora resolution. *ACL Anthology*, 38(2), 1–12.