International Journal of English Language Education
ISSN 2325-0887
2025, Vol. 13, No. 2

A Comparative Study of the Linguistic Features of Human Versus Machine Translation: A Formal, Syntactic and Lexical Study

Othman Aljohani

English Language Institute, University of Jeddah, KSA

2871 University Road, Jeddah 23890, Saudi Arabia

Received: July 18, 2025 Accepted: August 5, 2025 Published: August 6, 2025

doi:10.5296/ijele.v13i2.23074 URL: https://doi.org/10.5296/ijele.v13i2.23074

Abstract

The study provides a comparative analysis of two articles that have been translated using machine translation (Google Translate), a human translation of the same article by two bilingual translators, and the adjusted translation of the machine translation by the two translators. The objective of this study is to expedite the progression of artificial translation and enhance its capabilities to streamline the translation of professional writings and the algorithms involved in this process, regarding word count, lexical items, and syntactic sentence reconstruction. The study employs a quantitative analysis of word count, syntactic change structures, and lexical item alteration to demonstrate the difference between the two translations. The findings reveal despite the advancements in machine translation, particularly after the incorporation of the benefits of artificial intelligence (AI) into translation algorithms, the personal translation still possesses its aesthetic aspects, which account for the context and eliminate any ambiguity that may exist in the text. The study may interest translators who further propose to conduct comparative study of the human and machine translation systems and especially to those who plan to adopt of hybrid approach to translation.

Keywords: human translation, machine translation, lexis, source text. syntax, target text

1. Introduction

Artificial intelligence has revolutionized the way in which the perspectives of the world are viewed. However, a great number of translators are utilizing artificial intelligence translation machines programs. This matter leads to inappropriate translated texts without realizing the cultural and social dimensions of the text. Advancements in machine translation technology have led to generalized and oversimplified comparisons with human performance. So, it is essential to identify the actual differences between the two modes of translation accurately. The use of computers to mutually transform two languages is generally referred to as machine translation (MT) (Baker, 2007). In the wake of neural machine translation systems (NMT), widespread research activity has begun to identify similarities and differences between machine and human translation (HT). Two major approaches emerged in regard to the comparative analysis of the two translation systems. Some scholars (L¨aubli et al., 2018; Popel et al., 2020) appraise a holistic perception of quality, while other researchers (e.g. Vanmassenhove et al., 2019) target more specific features such as the ranges of lexical choices and text cohesion in their attempt to ascertain the qualitative differences between machine and human translations.

With regards to formal characteristics of translation, HT often demonstrates exceptional proficiency in preserving the formal tone and style of the source text (ST). Human translators possess the ability to adjust to various registers and styles, therefore guaranteeing that the translation produced is suitable for the given context. On the other side, MT frequently has difficulties in preserving a formal tone, particularly in intricate texts. Although significant progress has been achieved, MT may sometimes be too literal or overlook subtleties. On the syntactic characteristics level, HT possesses superior ability in managing intricate syntactic structures, therefore guaranteeing the grammatical accuracy and natural-sounding nature of the translated text. However, MT may have difficulties with syntax, particularly in languages characterized by adaptable word ordering or intricate grammatical rules. Sentence construction errors are more prevalent. On the linguistic characteristics level, human translation has exceptional proficiency in selecting appropriate vocabulary and expressions, taking into account the surrounding circumstances, implications, and cultural subtleties. In addition, they show enhanced proficiency in handling colloquial idioms and slang. MT, however, often has difficulties in selecting appropriate vocabulary, particularly when dealing with idiomatic phrases and terms that have several interpreted meanings. Nevertheless, their fast improvement is facilitated by the progress made in artificial intelligence (AI) and machine learning.

2. Literature Review

Translation helps people from diverse linguistic and cultural backgrounds to engage in meaningful and effective communication. In the wake of innovations in MT through AI, global communication among people having different language backgrounds has eased. MT has also appeared as a very strong competitor for the HT which used to be the only way of deciphering a ST. Despite being fast and efficient, MT produced target texts (TT) have been frequently questioned for issues in regard to syntactic accuracy, lexical appropriateness, semantic depth,

coherence, and pragmatics. The present study aims to compare MT with HT dilating upon formal, syntactic, lexical features that determine the quality of translation. The study is expected to unfold some viable insights on the effect of technological advancements on the process and quality of translation.

The foremost point of departure between MT and HT is the variation in syntax and grammatical accuracy. HTs are coupled with an intuitive awareness of cultural sensitivities and contextual features, and thereby imbibe syntactic features more robustly than MT (House, 2015). Human translators can adequately reformulate a variety of sentence construction not only to maintain grammatical accuracy but also to ensure text cohesion and coherence. Conversely, as observed by Kohen (2020), rule-based and statistical frameworks of translation find it challenging to handle morphological ordering, subject-verb agreement and formulaic expressions such as the idioms, phrasal verbs, and proverbs. Owing to advances in NMT that uses deep learning algorithms, syntactic quality of MT has remarkably refined, but still there are issues with complex syntactic structures (Bahdanau et al., 2014).

The quality of translation is also determined by the extent to which the TT maintains semanticity of the ST. Venuti (2018) considers HT semantically more aligned with the ST because the translators understand the contextual nuances inside out and can employ a range of lexical variety to produce translation that is compatible with the pragmatics and other discourse features of the target language. Their handling of the formulaic language, word connotations, and polysemy, for instance, in producing translation equivalents in the ST is far more refined than one finds in translation done through machine. According to Papineni et al. (2002), this happens because of the reliance on pattern recognition and probabilistic frameworks in attempt for lexical choices by MT. This distorts the understanding of the polysemous words, thereby creating faulty equivalents. Although recent advancements have ensured that NMT achieves semantic finesse through contextual embeddings, producing comprehensibility for words with different meanings is still needed (Vaswani et al., 2017).

Another essential descriptor of authentic translation is the extent to which the translator or the translation device showcase pragmatic competence and a cognizance of the cultural sensibilities that formulate both the ST and the TT. Baker (2011) opines that human translators possess the knack to incorporate pragmatics variables such as pun, humour, politeness expressions, and idiomatic language more competently than the MT versions. This view was also earlier echoed by Pym (2010) who found MT tools lacking in exploiting cultural sensibilities and contextual parameters to provide for the pragmatic equivalence. Although substantial leaps have been made in the output of MT, especially after the induction of AI tools, MT devises still struggle to effectively handle matters relating to formality, variations in sociolects and vernaculars, and appellatives (Doherty, 2016).

The quality of translation is also determined by fluency and readability. Following Munday (2016). HT generally maintains the natural flow of reading by retaining the stylistic features and rhetorical functions, and the TT reads as the natural equivalent of the ST. Typically, such translations effectively use the text forming sources, for instance, the conjunctive and the referential cohesion devices not only to make the text readable but also to create texture that is

the primary feature of any piece of discourse. On the other hand, MT received, especially from the older versions were more or less literal in content devoid of real fluency. Presently, as observed by Luong et al. (2015), NMT have displayed the capacity to generate higher levels of text coherence and text cohesion with the result that their translations appear more natural. Regardless of the AI potential and advancements, Toral and Way (2018) still find MT facing challenges to produce naturally sounding phrases, lexical variety, and consistency, especially while translating gendered language.

One major leap in the domain of NMT is its ability to create systems which effectively synchronize contextual associations between the lexical items. Contrary to the rule-based and statistical frameworks which processed individual words for producing translation equivalents, NMT has the potential to use even the complex structures (Sutskever et al., 2014). This comprehensibility has facilitated the MT to substantially enhance natural readability and contextual knowledge (Johnson et al., 2017). Despite these developments, NMT faces problems while translating low frequency words, languages with limited linguistic data, and subject specific vocabulary (Koehn & Knowles, 2017).

In view of the weaknesses inherent in the MT systems, post-editing is being considered an alternative to increase the quality of translation. Following, O'Brien (2012), MT generated texts can be operated upon by the human translators for any amendments in regard to levels of formality, syntactic accuracy, lexical appropriateness so that the TT may appear natural and readable. Blending AI produced translations with human interventions have started to gain acceptance in workplace contexts, thereby promoting the availability of translation facility that can be fluent, contextually comprehensible, and linguistically accurate (Carl et al., 2011; Toral & Way, 2018).

In translation research, naturally occurring lexical collocations or chunks have been extensively researched, especially in regard to their frequent incidence and occurrence in specific combinations (Bestgen, 2021). Baldwin and Kim (2010) consider these lexical constellations statistically significant representative samples of the language as they are analyzed in a more statistically robust method compared with individual occurrences of the lexical items in the same sequence. Researchers implement CollGram method to investigate the presence of these formulaic lexical associations. This method applies mutual information (MI) and t-score to find out the rankings as are unfolded by frequencies in the target corpus (Bernardini, 2017); Bestgen and Granger, 2014; Durrant and Schmitt, 2009).

Reliant on MI and t-score indices, the CollGram algorithm quantitatively measures the extent to which the words in a bigram are associated with each other. Incidence and frequency of the word associations are the predominantly focused indicators which are mutually collaborative in both MI and t-score analysis. However, both these indices have distinct preference for the frequency type. While MI reveals more inclination towards lower frequency such as the "self-fulfilling prophesies" or thinly distributed word combinations, t-score seems to have preference for higher frequency collocation, for instance, "such as" or "greater than". Research has established that both advanced and intermediate level non-native speakers (NNS) in relation to the native speakers (NS) display minimal use of the word combinations with high MI scores,

and relatively high frequency use of the collocations with high t-scores (Bestgen & Granger, 2014; Durrant & Schimtt, 2009). The outcomes are compatible with the usage-based theories of language learning, and entail that both the frequency and the collocational patterns in the target language have statistically significant impact on the acquisition of formulaic expressions (Durrant & Schimtt, 2009). More interestingly, these incongruities were also visible between the ST and the TT. One viable reason for the disparity could be because of the translators' inclination towards the desired standardization (Bestgen & Granger, 2019, Bestgen, 2018).

Following the proclaimed effect of the use of frequency on neural frameworks (Koehn & Knowles, 2017, Li et.al., 2020), the present study proposes to validate the hypothesis that identical effects may be observed while analyzing HTs and NMTs. More specifically, it has been hypothesized that NMTs might reveal an underuse of the MI frequency score and an overuse of the t-score frequency score. As an instance, the researcher refers to Bestgen (2018), whose study conducted an extensive analysis of two automated processes that partly use this framework. The results of his analysis unfolded that the t-score index using NMT produced densely occurring collocational sequences which were formed of high frequency lexical items such as "you know," "out of," "more than" etc. In contrast, he found quite sparsely occurring collocational strings generated by highly low frequency lexical items when MI index was applied. Formulaic expressions such as self-fulfilling prophesy, sparsely populated, or sunnier climes are a few instances of the findings.

These outcomes can be corroborated with those studies which have shown comparable diversification while foreign language acquisition is in process (Bestgen and Granger, 2014 ; Durrant and Schmitt, 2009). The findings of these studies entail that these variations need to be appreciated on the framework of usage-based theories of language learning. These theories postulate that the incidence and frequency of language features in the input can have statistically significant effect on the use of formulaic expressions (Durrant and Schmitt, 2009). Following Koehn and Knowles (2017) and Li et al. (2020), varieties in translation can be explained in the same way as the impact of brain models on the frequency of use.

There may be another perspective about these approaches, however. Bestgen (2021) samples were all published newspaper articles of high standards in French which were later translated into English for one of its international editions. However, as pointed out by Ponomarenko (2019) translation of journalistic writing follows distinctly different processes of discourse production compared with other translated text types as it involves frequent resort to reviewing, retelling and rewriting. Moreover, it was also observed that domestication of information takes preference over translation accuracy and quality as far as the publication of international news is concerned. Assuming that the translated news articles possess significant relevance and interest for the global readers who are not the NS of the source language, the translators typically apply lexico-grammatical choices either through deletions or additions. The intended outcomes are likely to be relatively less literal translation when compared with the MT of the same text, and this explains variations in the use of formulaic expressions and lexical and grammatical collocations.

Seen from this perspective, legal or legislative texts may be considered an accurate corpus for

comparison. Since the ultimate goal of all translation endeavours is to achieve accurate translation equivalence, it becomes imperative that the target text is showcased in its aggregation devoid of any deletions or amendments, thereby ensuring that the target text is the real, authentic and reconcilable version of the source text (Sosoni, 2011).

2.1 Research Questions

The researcher developed the following two research questions for the study:

1. Does machine translation vary from human translation in terms of syntax, vocabulary, and formal conventions?
2. If there are any disparities between machine translation and human translation in terms of syntax, vocabulary, and formalisation, how may they be modified?

3. Method

3.1 Translation Corpus

The sample texts for the study were derived from an article titled "Reading Japan" by David Remnick, which was published on January 29, 1995. It was translated and published in the Syrian General Organization for Books' affiliated Journal of Cultural Bridges No. 25-26 dated 2021. The well-known NMT system, Google Translate (translate.google.com), was used to translate all original texts into Arabic. The translation was completed on June 24, 2023. Then, translation experts affiliated with the Syrian General Authority for Books evaluated and approved the researcher's translation and the MT of the same article. A comparison was conducted.

When the original English-language text was translated into Arabic, the number of lexical items dropped from 5,000 words to 3,566. By contrast to the machine translation, the human translation had lexical items count of 4,329. This was an increase of roughly 25%.

3.2 Procedure

The translator's responsibilities were restricted to reviewing the original material, refining the MT of that text, and modifying the MT to align with the professional translation of the specific content. This approach revealed several elements, such as the practice of translating frequently used words in the automated text while disregarding the pertinent context in the target text. This led the translator to revise the existing translation in order to incorporate a human element that considers the context.

The original text, the machine translated text, and the modified translated text were categorized based on paragraph style, sentence structure, word count, and symbol count. Next, the language distinctions were categorized into syntactic and lexical variations. Regarding syntax, there was a modification in sentence placement and the incorporation of grammatical tools. On the lexical aspect, it was further separated into alterations in lexical items and additions of lexical items. The aim of the alteration was to align with the context, the intended significance of the original author, and the aesthetic conventions of the Arabic language. Subsequently, a comprehensive

analysis was conducted to identify any significant differences between the data acquired by the machine and human translation.

3.3 Data Analysis

The data were collected and then processed utilizing SPSS Statistics Version 25, in which descriptive statistics were used in presenting the results of the data collection.

The two research topics will be explicitly debated and evaluated following the presentation of the data statistics and their implications. The results will be presented at a later stage, and implications regarding the effective use of WhatsApp will be derived. In addition to the recommendations for future research, a conclusion will be obtained.

According to Microsoft Word 10's, word count, the original text, machine translated text, and human translation of the text were classified into multiple categories, as shown in Table 1:

Table 1. Data Categorization

	Source text in English	**Machine translation in Arabic**	**Human Adjusted translation in Arabic**
Paragraphs	42	14	38
Sentences	248	218	237
Word count	5011	3566	4329
Symbol count	23472	16632	20436

Subsequently, the sentences were segmented by either underlining the word or phrase that underwent processing, or by highlighting it in bold, or both, in order to differentiate it from the remaining vocabulary. Next, a comprehensive analysis was provided for each of the previously proposed alternatives.

Table 2. Percentage of Whole Article Modifications

Linguistic Field	**Modification Type**	**Items Number**	**Total**	**Percentages**	**Total**
Syntactic field	Changing the position of the sentence	75	170	44.1	35.7
	Adding grammatical tools	95		55.9	
Lexical field	Lexical item change	180	305	60	64.3
	Lexical item addition	125		40	
Total			**475**		**100**

4. Results and Discussion

The results were obtained from the analysis of the sample to answer the two research questions that were generated. The first question aimed at finding out if there was any difference between machine translation and human translation at the syntactic, lexical and formal levels; while the second question sought to explore if there was any difference between machine translation and human translation at the syntactic, lexical and formal levels, how it can be adjusted. Presented below are the translation samples from the ST and TT:

عندما كان الروائي الياباني كنزابورو أوي في ستوكهولم للحصول على جائزة نوبل في الأداب في كانون الأول، أخبر كل الحاضرين أنه سينسحب الآن من ساحة نجاحه. وما فتئ يخبر جمهوره المهذب والمحاورين: "سأتوقف عن كتابة الروايات". لقد كان تصريحاً غريباً قام به بمنتهى الهدوء والبهجة حتّى أن أحداً لم يصدّقه على ما يبدو. إذ يبلغ أوي من العمر ستين عاماً فحسب ويتمتع بصحة جيدة ويُعرف بأنه كاتب النثر الياباني الرائد. ساهمت الجائزة في زيادة مبيعات كتبه في اليابان ودفعت الناشرين الأجانب للتهافت لترجمة المزيد من رواياته وقصصه. لم يكن أوي منهكاً ولا مكتئباً، ولم يشعر قط بالسلام الداخلي كما هو الأمر في حينه. كان قراره ينبع من رغبته إغلاق ملفّ ما وليس تعبيراً عن أزمة ما. يقول إنه وبرحابة صدر يتخلى عن الأدب لأن المهمة التي حددها لنفسه قبل 31 عاماً لم تعد ضرورية – ألا وهي التحدث بطريقة ما عن ابنه هيكاري الذي أصيب بأضرار بالغة في الدماغ -. فقد اعتبر أن هيكاري الذي من النادر أن يتحدث والذي يعاني من نوبات وينبغي الاعتناء به في جميع الأوقات قد وجد من يتحدّث عنه. لقد قام بتأليف بعض الموسيقى الرائعة للبيانو والناي، التي يصفها والده قائلاً : إنها "كالندى المتلألئ على أوراق العشب" - وقد أصدر مؤخراً ألبومه الثاني المعنون "موسيقا هيكاري أوي، 2".

4.1. Sample translation instance 1

- **Original English Text**
  When the Japanese novelist Kenzaburo Oe was in Stockholm to collect the Nobel Prize in Literature in December, he told all *who would listen* that he would now withdraw from the arena of his success. Over and over, he told his polite audiences and interviewers, “I am going to stop writing **fiction**.”
- **Machine Translation into Arabic**
  عندما كان الروائي الياباني كينزابورو أو في ستوكهولم لجمع جائزة نوبل في الأدب في ديسمبر، أخبر كل *من سيستمع* بأنه سينسحب الآن من ساحة نجاحه. ***لقد انتهى الأمر مراراً وتكراراً*** جماهيره المهذبة ومقابلاته، "سأقوم بتوقف عن كتابة **الخيال**".
- **Human Modified Translation into Arabic**
  عندما كان الروائي الياباني كنزابورو أوي في ستوكهولم للحصول على جائزة نوبل في الأداب في كانون الأول، أخبر كل *الحاضرين* أنه سينسحب الآن من مضمار نجاحه. ***وما فتئ*** يخبر جمهوره المهذب ومحاوريه: "سأقلع عن كتابة **الروايات**".

The word “to collect” means "to gather", but in this context it means that the writer "Oe" is in Sweden to "get" the Nobel Prize. The phrase “*who would listen*” means “who listen”. This is the direct machine translation while what is meant here by the original writer is “who are present during the press conference”.

4.2. Sample translation instance 2

- **Original Text in English**

It was such a queer statement, **and delivered with such calm and good cheer**, that no one seemed to believe it.

- **Machine Translation into Arabic**
  كان هذا البيان الغامض، **وتسليم بهذا الهدوء والبهجة الجيدة**، أنه لا يبدو أن أحد يصدق ذلك
- **Human Modified Translation into Arabic**
  لقد كان تصريحاً غريباً **قام به بمنتهى الهدوء والبهجة** حتّى أن أحداً لم يصدّقه على ما يبدو.

4.3. Sample translation instance 3

- **Original Text in English**
  The prize has accelerated sales of his books in Japan, and sent his foreign publishers racing to translate more of his novels and stories.
- **Machine Translation into Arabic**

لقد سجلت الجائزة مبيعات كتبه في اليابان، وأرسلت ناشري أجنبيين لترجمة المزيد من رواياته وقصص

- **Human Modified Translation into Arabic**
  ساهمت الجائزة في زيادة مبيعات كتبه في اليابان ودفعت الناشرين الأجانب للتهافت لترجمة المزيد من رواياته وقصصه.

4.4. Sample translation instance 4

- **Original Text in English**
  Oe is neither tired nor depressed. He has never felt more at peace. *His decision is more one of closure than of crisis.*
- **Machine Translation into Arabic**
  أوي ليس متعبا ولا الاكتئاب. لم يشعر أبدا بالسلام. *قراره هو أكثر من إغلاق من الأزمة.*
- **Human Modified Translation into Arabic**
  لم يكن أوي منهكاً ولا مكتئباً، ولم يشعر قط بالسلام الداخلي كما هو الأمر في حينه. *كان قراره ينبع من رغبته إغلاق ملفّ ما وليس تعبيراً عن أزمة ما.*

Following Rayson (1991), the researcher tokenized each translated text and tagged POS with CLAWS7 thereby leading to the extraction of all bigrams. Bigram extraction was, however, interrupted due to variation in punctuation marks and also due to any character sequences that were not consistent with a lexical item/s. The hundred-million-word British National Corpus (BNC2) was utilized in those cases where a bigram was identified without a specific name or number. However, with the identified bigrams, the MI and t-scores which were in correlation were used to establish the collocational status of the bigram. Following Bestgen (2018) and Durrant and Schmitt (2009) strongly collocational incidences were considered to be bigrams with >5 for MI and >6 for t-scores. Finally, the percentage that determined the status of the bigrams as being strongly collocational was calculated for each instance of the text and the association index while comparing these with the overall presence of the bigrams in the text.

This research assists in the contributions conducted around the world about the future of artificial intelligence. This research shows the actual differences between human and machine translations regarding syntactic, lexical and formal levels.

This research attempts to establish a comparative study between an article translated by

machine translation (MT) (Google Translate) and the researcher's translation of the same article. Despite the development in machine translation, especially after the introduction of the advantages of artificial intelligence in translation algorithms[1], this research shows that although machine translation contributes to forming a clear picture of the content in the article, it requires the personal adjustment of humans to provide adequate and comprehensible meaning of sentences and context. The human translator captures the aesthetic dimension, intricacies and cultural subtleties of the text, which takes into account the context and remove the ambiguity that may surround the text. Thus, human translators modify the course of translation, alter the locations of words or sentences, and add a vocabulary or more to provide the text with the indispensable spirit of the target language. They might add indispensable explanations to the reader in the target language in order to fathom the depths of the text. The previous matter is still not present in machine translation at the present time.

This research compares machine translation of the identical source material from a contemporary system against a human translation using a brief case study. There were two reasons for doing this. The first is about how to characterize the differences and what ideas and resources are helpful in giving such a comparison context. The second is to compare the pros and cons of machine translation with human translation, keeping in mind that no one translation article can really capture the vast scope of the translation industry.

The aforementioned distinctions are further categorized into grammatical and lexical disparities. Grammatically, there was a modification in the placement and alignment of the sentences. With regards to vocabulary, the words were modified to align with the context, intended meaning of the original author, and the aesthetics of the Arabic language.

The findings of the study reveal that Google Translate has made amazing progress in the past six years compared to what it used to be. However, MT though it is time saving and safe to store due to the typing facility is still far from HT in terms of overall quality. With the large number of documents and books that need translation, the work of the translator will move in the current and next stage from direct translation to revision and linguistic coordination. Similarly, the field of vocabulary is the field that is still far from being realized by MT, and it is a field that takes into account the context and the aesthetics of the language and the addition of necessary vocabulary in order for the sentence to reach the mind and soul of the reader. So are the cultural aspects which are still far from being accurately understood by MT. Nevertheless, the grammatical arrangement of the sentence is an area that MT was able to approach in most of the sentences that were studied. Since Google Translate was looking for the corresponding words or phrases, it was good for short sentences or phrases. But things tended to get inaccurate when longer text with complex grammar rules were attempted to be synchronized. Because when context and sentence structure came into play, Google Translate was unable to process and generate meaningful translations.

This paper illustrates that, no matter the progress made in MT, especially with the integration of AI into translation algorithms, HT still retains its intricacies, cultural subtleties and aesthetic elements. More specifically, machine translation contributes to the formation of a clear

understanding of the article's content. Consequently, the translators alter the course of the translation, alter the placement of words or sentences, and incorporate a word to integrate the essential essence of the target language into the text. In order for the reader in the target language to fully understand the text, translators may include essential explanations machine translation is currently far from accomplishing this objective. The translator's responsibilities were restricted to the following: reading the original text, re-reading the machine translation, and modifying the machine translation to align with the professional translation of the text. In this approach, a number of aspects were identified, such as the translation of common words in the automated text and the disregard of the relevant context in the target text. This prompted the translators to rephrase the current translation to provide a human character that considers the context.

5. Conclusion

In view of the rapid advancements in the domain of AI, it is anticipated that more hybrid models involving NMT and human input will be an integral part of the future translation paraphernalia. While NMT has the potential to speedily handle extensive data, human collaboration will be required to make essential tweaks as far as linguistic accuracy, cultural sensibility, and critical maneuvering are concerned. However, as more and more use of AI becomes the norm, ethical issues in regard to training data and human interventions will have to be addressed. Nevertheless, a judicious application of the hybrid model will have to be made tapping, for instance, the potential of MT for increased performance in syntax, semantics, and readability of the TT. Similarly, human translators can lend their expertise in refining the linguistic weaknesses, cultural variations, and discourse features so that the TT best represents the nuances of the ST.

References

Bahdanau, D., Cho, K., & Bengio, Y. (2014). Neural Machine Translation by Jointly Learning to Alignand Translate..arXivpreprintarXiv: https://doi.org/10.48550/arXiv.1409.0473

Baldwin, T., & Kim, S.N. (2010). Multiword expressions. In: Indurkhya, N., Damerau, F.J. (eds.) *Handbook of Natural Language Processing, 2*67-292. CRC Press.

Baker, M. (2007). Patterns of idiomaticity in translated vs. non-translated text. *Belgian Journal of Linguistics, 21*, 11-21.

Baker, M. (2011). *In Other Words: A Coursebook on Translation.* Routledge.

Bernardini, S. (2017). Collocations in translated language. combining parallel, comparable and reference corpora. In: Proceedings of the Corpus Linguistics Conference. 1, 16. Lancaster University.

Bestgen, Y. (2018). Normalisation en traduction : analyse automatique des collocations dans des corpus. *Des mots aux actes 7*, 459-469.

Bestgen, Y. (2021). LAST at CMCL 2021 Shared Task: Predicting Gaze Data During Reading with a Gradient Boosting Decision Tree Approach. In *Proceedings of the Workshop on Cognitive Modeling and Computational Linguistics*, pages 90–96, Online. Association for Computational Linguistics.

Bestgen, Y., & Granger, S. (2014). Quantifying the development of phraseological competence in L2 English writing: An automated approach. *Journal of Second Language Writing, 26*, 28-41.

Bestgen, Y., & Granger, S. (2019). Collocation et traduction. analyse automatique au moyen d'indices d'association. In: Kau_er, M., Keromnes, Y. (eds.) Theorie und Empirie in der Phraseologie / Approches th_eoriques et empiriques en phras_eologie, 101-113. Stau_enburg Verlag.

Carl, M., Dragsted, B., & Jakobsen, A. L. (2011). A Taxonomy of Human Translation Styles. *Translation and Interpreting Studies, 6*(2), 93-116.

Doherty, S. (2016). Evaluating the Impact of Controlled Language on Translation Quality and Post-editing Effort. *Journal of Specialised Translation, 25*, 144-163.

Durrant, P., & Schmitt, N. (2009). To what extent do native and non-native writers make use of collocations? *International Review of Applied Linguistics in Language Teaching, 47*, 157-177.

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

Johnson, M., Schuster, M., Le, Q. V., Krikun, M., Wu, Y., Chen, Z., & Dean, J. (2017). Google's Multilingual Neural Machine Translation System: Enabling Zero-Shot Translation. arXiv preprint arXiv: https://doi.org/10.48550/arXiv.1611.04558

Koehn, P. (2020). *Neural Machine Translation.* Cambridge University Press.

Koehn, P., & Knowles, R. (2017). Six Challenges for Neural Machine Translation. arXiv preprintarXiv: https://doi.org/10.48550/arXiv.1706.03872

Läubli, S., Sennrich, R., & Volk, M. (2018). Has Machine Translation Achieved Human Parity? A Case for Document-level Evaluation. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 4791–4796, Brussels, Belgium. Association for Computational Linguistics.

Li, M., Roller, S., Kulikov, I., Welleck, S., Boureau, Y. L., Cho, K., & Weston, J. (2020). Don't say that! making inconsistent dialogue unlikely with unlikelihood training. In: Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. pp. 4715-4728. Association for Computational Linguistics.

Luong, M.-T., Pham, H., & Manning, C. D. (2015). Effective Approaches to Attention-based Neural Machine Translation. arXiv preprint arXiv: https://doi.org/10.48550/arXiv.1508.04025

Munday, J. (2016). *Introducing Translation Studies: Theories and Applications*. Routledge.

O'Brien, S. (2012). *Post-Editing Machine Translation.* John Benjamins Publishing Company.

Papineni,K., Roukos, S. Ward, T. Henderson, J., & Reeder, F. (2002). Corpus-based comprehensive and diagnostic MT evaluation: Initial Arabic, Chinese, French, and Spanish results. In *Proceedings of Human Language Technology 2002*, San Diego, CA.

Ponomarenko, L. (2019). Translating identities in multilingual news. Ph.D. thesis, Universitat Autonoma de Barcelona.

Popel, M., Tomkova, M., Tomek, J., Kaiser, L., Uszkoreit, J., Bojar, O., & Zabokrtsky, Z. (2020). Transforming machine translation: A deep learning system reaches news translation quality comparable to human professionals. *Nature Communications 11, 1*(15). https://doi.org/10.1038/s41467-020-18073-9.

Pym, A. (2010). Translation and text transfer. An essay on the principles of cross-cultural communication. Intercultural Studies Group.

Rayson, P. (1991). Matrix: A statistical method and software tool for linguistic analysis through corpus comparison. Ph.D. thesis, Lancaster University.

Sosoni, V. (2012). A Hybrid Translation Theory for EU Texts. Vertimo Studijos. 5. 76-89. https://doi.org/10.15388/VertStud.2012.5.10561.

Sutskever, I., Vinyals, O., & Le, Q. (2014). Sequence to Sequence Learning with Neural Networks. Advances in Neural Information Processing Systems. 4.

Toral, A., & Way, A. (2018). What Level of Quality Can Neural Machine Translation Attain on Literary Text? *Translation Studies, 11*(2), 131-150.

Vanmassenhove, E., Shterionov, D., & Way, A. (2019). Lost in Translation: Loss and Decay of Linguistic Richness in Machine Translation. https://doi.org/10.48550/arXiv.1906.12068.

Vaswani, A. Shazeer, N. Parmar, N. Uszkoreit, J. Jones, L. Gomez, AN., & Kaiser, P. I. (2017). Attention is all you need. In: Advances in neural information processing systems. Red Hook: Curran Associates, 6000-6010.

Venuti, L. (2018). *The Translator's Invisibility: A History of Translation*. Routledge.

APPENDIX

Source Text in English

When Japanese novelist Kenzaburo Oe was in Stockholm collecting for the Nobel Prize in Literature in December, he told all who would listen that he would now withdraw from his success. He repeatedly told his polite audiences and interviewers, "I'm going to stop." And he just kept telling his polite audiences and interviewers, "I'm going to stop writing fiction." It was such an ambiguous pronouncement, and given that a strange pronouncement he made with such calmness and good cheer, that no one even seemed to believe it. OE is nobody believed him for sixty sixty years, it seems. Oe is only sixty years old and in good health, recognized as an outstanding writer in Japan and recognized as a Japanese prose writer. I registered Major. The award contributed to increasing the sales of his books in Japan, sent foreign publishers and prompted foreign publishers to scramble to translate more of his novels and stories. OE is not tired, not exhausted, nor depressed. He was never depressed, and he never felt at peace. His decision is more of a closure than a crisis. He says, internal as he says, because the mission is the matter in time. His decision stemmed from his desire to close a file, and not as an expression of a crisis. He says that he gladly abandons literature because the task he set himself one year out of thirty years ago—to speak 31 years ago is no longer necessary—that is, to speak one way or another for the sake of his brain-damaged son, is no longer necessary. - Hikari, whose brain was badly damaged. He considered that Hikari, who rarely speaks and who suffers from seizures, seldom speaks, and must and should be taken care of at all times, has found his voice. He has composed some great music for piano and vine – “like the dew shining on the flute, which his father describes as ‘like the dew sparkling on the leaves of the grass’ is how his father describes it – and he recently released his own CD – and he has recently released his second album, Music Entitled Music by Hikari Oe Aoi , 2." ".

Acknowledgments

Not Applicable.

Funding

Not Applicable.

Competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Informed consent

Obtained.

Ethics approval

The Publication Ethics Committee of the Macrothink Institute.

The journal's policies adhere to the Core Practices established by the Committee on Publication Ethics (COPE).

Provenance and peer review

Not commissioned; externally double-blind peer reviewed.

Data availability statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

Data sharing statement

No additional data are available.

Open access

Copyrights