Translation Quality of Artificial Intelligence and Machine Translation Vs. Human Translation Utilizing MTPE Skills (An Empirical Study on Allusion Translation)

Ibrahim Jibreel [1,*]

Received: 06 September 2024
Revised: 07 September 2024
Accepted: 23 September 2024

[1] Associate Professor of English & Translation Studies, Department of English & Translation, University of Science & Technology, Hodeidah. ORCID: https://orcid.org/0000-0002-6231-3059
* Corresponding author. E-mail: ibjib80@gmail.com

Translation Quality of Artificial Intelligence and Machine Translation Vs. Human Translation Utilizing MTPE Skills (An Empirical Study on Allusion Translation)

Abstract:

Allusion is one of the culture-bound expressions that need careful consideration while translating. Machine translation (MT) and human translators (HTs) encounter difficulties in dealing with them. This study compares Translation Quality (TQ) of MT and Artificial Intelligence (AI) to HTs utilizing MTPE focusing on identifying the MTPE skills to keep HT in favor of MT and AI.

A quantitative and qualitative mixed method was adopted using a test of 30-item in-context English-to-Arabic allusions translated by Google Translate and ChatGPT and then given to a random sample of 40 HTs. The TQ of AI, MT and HT target texts were assessed following O'Brien's (2012) model. The participants wrote reports on MTPE skills and were involved in a focus group discussion to determine the MTPE skills used. *One-Sample t-Test*, *One-Way ANOVA* and POST HOC Test were used. Results show HTs utilizing MTPE are of *Moderate Quality* (60%), and MT and AI-based translations are of *Low Quality* (44.44% & 42.22%). HTs employ some MTPE skills and strategies that resulted in statistically significant differences between HTs of allusions compared to MT and AI in favor of HTs. The study recommends enhancing MTPE skills among translation students and implementing training for further developing translators.

Keywords: *Allusion translation, Artificial Intelligence (AI), Human Translation (HT), Machine Translation (MT), MTPE & Translation Quality (TQ).*

https://doi.org/10.20428/jss.v30i3.2545

مقارنة جودة ترجمة الذكاء الاصطناعي ومحركات الترجمة الآلية مع ترجمة الإنسان: الحاجة إلى مهارات التحرير اللاحقة (دراسة تطبيقية على ترجمة الإشارات الضمنية (التلميح)

د. إبراهيم جبريل [1،*]

الملخص:

تعد الإشارات الضمنية (التلميحات) أحد التعبيرات المرتبطة بالثقافة والتي تحتاج إلى مزيد من الاهتمام أثناء الترجمة. إذ تواجه الترجمة الآلية والمترجمون صعوبات في التعامل معها. تقارن الدراسة جودة ترجمات الترجمة الآلية والذكاء الاصطناعي مع ترجمة الإنسان مستفيداً من مهارات التحرير اللاحقة مع التركيز على تحديد هذه المهارات لإبقاء ترجمة الإنسان أعلى جودة من ترجمة الآلة والذكاء الاصطناعي. اتبعت الدراسة المنهج المختلط مستخدمة اختباراً مكونا من 30 إشارة ضمنية في جمل سياقية لترجمتها من الإنجليزية إلى العربية أولاً بواسطة Google Translate و ChatGPT ثم تقديمها لعينة من 40 مشاركًا لترجمتها مستفيدين من مهارات التحرير اللاحق مع كتابة تقرير عن المهارات والاستراتيجيات والمصادر التي استفادوا منها، كما شاركوا في مجموعة نقاش مركزة. تم تصحيح النصوص الناتجة عن ترجمة الذكاء الاصطناعي و الآلة وتحليلها لتقييم جودة الترجمة وفقًا لنموذج O'Brien (2012). تم استخدام اختبار t لعينة واحدة، والتباين الأحادي، واختبار المقارنات البعدية المتعددة. تظهر النتائج أن ترجمة الانسان مستفيدا من مهارات التحرير اللاحق جاءت بجودة متوسطة (60٪)؛ بينما جاءت الترجمات الآلية والذكاء الاصطناعي ذات جودة منخفضة بنسبة (44.44٪ و42.22٪). استخدم المترجمون مهارات تحرير لاحقة أدت إلى فروق دالة إحصائية بين ترجمة الانسان مقارنة بترجمة الآلة والذكاء الاصطناعي. توصي الدراسة بتعزيز مهارات التحرير اللاحقة بين طلبة الترجمة وتدريب المترجمين عليها للمزيد من التطوير.

الكلمات المفتاحية: ***ترجمة الإشارات الضمنية، الذكاء الاصطناعي، ترجمة الإنسان، ترجمة الآلة، مهارات التحرير اللاحق، جودة الترجمة.***

---

[1] أستاذ اللغة الإنجليزية والترجمة المشارك، قسم اللغة الإنجليزية والترجمة، جامعة العلوم والتكنولوجيا، الحديدة

[*] عنوان المراسلة ibjib80@gmail.com

Introduction

In the rapidly evolving translation industry, the integration of technology has sparked considerable debate regarding the quality of translations produced by artificial intelligence (AI) and machine translation (MT) systems compared to those produced by human translators, particularly when enhanced by machine translation post-editing (MTPE) skills. Machine Translation (MT) utilizes algorithms to render the source text from one language to another, leveraging vast databases of linguistic patterns and structures. While MT has significantly improved in recent years, its limitations become pronounced when tasked with translating non-technical texts, especially those rich in cultural references such as allusions. For Hutchins & Somers (1992), it is nearly impossible for MT systems to manage and predict all the necessary contextual information and background knowledge to accurately determine the meaning (and translation!) in every situation. Considering these issues, there were calls suggesting to use MT with human assistance. For them, human involvement can move between machine-aided human translation (MAHT) and human-aided machine translation (HAMT). In the former, the translator implements the translation and refers to e-sources to revise, edit, correct or improve. In the latter, the translator uses MT to perform the translation and helps to improve it. These two concepts were referred to as (CAT) tools.

Human translation (HT), on the other hand, inherently possesses the capacity for contextual understanding and cultural sensitivity, traits that are often critical in producing high-quality translations, Pym (2013). Recently, the incorporation of MTPE allows human translators to utilize MT as a preliminary tool, combining machine output speed with the nuanced understanding of human cognition. Despite the promising potential of this hybrid approach, the literature review may have overlooked a systematic comparison of translation quality across these methods, particularly in the realm of allusion translation.

Previous studies have often focused on the technical aspects of MT and AI, evaluating their performance in straightforward translation tasks without adequately addressing their effectiveness in conveying culturally embedded meanings. This gap highlights the need for a focused empirical study to specifically examine the intricacies involved in allusion translation to critically analyze the translation quality of AI and MT against that of human translators utilizing MTPE skills, focusing on the MTPE skills required to render allusions effectively.

Aims of the Study

Mainly, this study aims to compare the TQ of MT, AI-based translations with human translations utilizing MTPE with regard to the translation of Allusions.

Specifically, the following are sub-aims:

https://doi.org/10.20428/jss.v30i3.2545

1. to determine whether there are significant differences among the mean scores of the translation quality resulting from MT, AI & HT utilizing MTPE for allusion translation.
2. to identify error types and the strategies used to render allusions from English into Arabic.
3. to specify the MTPE skills needed for translating allusions effectively.

Literature Review

Allusion Definition, Types & Translation

Generally speaking, Huges (2009) defined allusion as "something said or written that mentions a subject, person etc indirectly". Precisely, an allusion, "culture bump" as described by Leppihalme (1997), is a figure of speech that indicates intertextuality. According to Abrams (1957), an allusion is a brief reference to a person, place, event, or another literary work without explicitly identifying it. Leppihalme argues that understanding allusions requires more than just explicit knowledge; it requires familiarity with the specific culture. Allusions are considered culture-specific elements and can be challenging to translate successfully without the translator's knowledge of their references.

Academically speaking, allusions enhance the quality of writing serving as literary expressions that add ambiguity or exaggeration. Sometimes, as in social or political constraint situations, they help writers as hedging devices when direct speech is not feasible. Allusions hold a significant influence in convincing readers to accept the author's viewpoints, particularly when referencing religious texts or well-known literary works.

One of the prominent divisions of allusion is that of Leppihalme (1997). She divided them thematically into four types including religious allusions, historical allusions, literary allusions, mythological allusions and popular cultural allusions. The first type comprises images and passages from religious texts; the second includes historical events and historical periods; the third deals with figures, events and images from popular myths; and the fourth discusses recent historical and popular cultural moments. Bahrami (2011) adopted the classification of Leppihalme (1994) who attempted to systematize the strategies of key-phrase (KP) allusions and Proper noun allusions (PN) following specific translation strategies, in the form of a hierarchical decision process. For him, adopting a more creative and reader-oriented transliteral role would offer more variety of translation strategies and decrease 'culture bumps' in translations.

Several studies have investigated the translation of allusions such as Leppihalme (1997), Bahrami (2012), Tao (2013) and Samir & Moallemi (2023). They conclude that the translation of allusions needs literal or direct translation with rephrasing to explain the annotative meaning and the translator may need some strategies that help to refer to the exact meaning. In addition, Wakelin (2007) compares two anonymous classical translations of De Consulatu Stilichonis by Claudian and Knyghthode and Bataile by Vegetius. Both poems seem to reflect the dynastic troubles and 'Wars of the Roses' of the mid-15th century. In both translations, there is an intellectual vagueness of the translation of allusion and how it leads to unpredictability of the reading process, essential to translation; complicate or transform 'propaganda'.

Concerning translation strategies of allusions, we have to admit that we are translating literary-specific-cultural expressions. Thus, literal translation cannot help rendering their meaning into the TL because of cultural relational connotation. Both linguistic and extralinguistic factors that affect the text to be reproduced into another language have to be taken into account, (Al-haj et al, 2021). In general, Lawrence Venuti (1998) proposes two main approaches to translation strategies: domestication and foreignization. Domestication involves translating a text in a way that aligns closely with the cultural values of the target language, aiming for fluency, transparency, comprehension, and readability. On the other hand, foreignization focuses on preserving the cultural values and characteristics of the source language, highlighting the foreignness of the source text in the translated text. It seeks to incorporate the linguistic, stylistic, and cultural aspects of the source text into the target language.

Translation strategies still have no borders and sometimes they can be used one after another. Different terms were coined by different scholars. For Jibreel et al (2016) scholars divide translation strategies into two main types regardless of their terminological controversy. When this concept is discussed, we remember scholars like Vinay and Darblnet (1958), Newmark (1988) Venuti (1998), Pedersen (2007), Baker (1992) and Ghazala (1995).

Precisely speaking, Leppihalme (1997) proposes several strategies to translate the allusions that come in the form of proper nouns (PN) or key phrases (KP). In the former, these strategies include name retention, name replacement and omission of the name. Under each strategy, there are several sub-strategies or procedures. For the latter i.e. (KP), she suggests some strategies such as using standard translation; literal translation (minimum change); adding extra-allusive guidance to the text; providing additional information via footnotes and endnotes; introducing textual features that indicate the presence of borrowed words; replacing it with a performed TL item; rephrasing the allusion with a clear expression of its meaning; re-creating the allusion by creatively constructing a passage that reproduces its effects; and omitting the allusion completely.

MT & AI Translation

The idea of machine translation (MT) goes back to "(1949) when Weaver introduced Americans to the idea of using computers for translation", Hutchins (2003). From that time on, there have been increasing developments in MT that generate the impression that translators will be out of their jobs. This impression is now being

https://doi.org/10.20428/jss.v30i3.2545

strengthened by AI translation technology, such as ChatGPT service. It backs into mind the idea of "the extremist concept of Fully Automatic High-Quality Translation (FAHQT)" which was dominant in "the first period of MT development", Austermuhl (2014:157). AI translation is one advanced form of the machine translation process that works utilizing intelligent behavior. As a result, it can analyze, understand and render an ST into another TL. In both services, developers try to minimize human involvement degree. However, "MT still needs to be supervised by human translators. So, what does the future hold for AI translation?", García (2022). Recently, using MT and AI in translation is characterized by its easiness, efficiency and reliability, in some wide cases.

ChatGPT is an AI model that uses natural language processing (NLP) techniques to translate text between various languages. It has evolved from early rule-based systems that relied on predefined grammatical rules and dictionaries to statistical machine translation (SMT) in the 1990s. However, it still struggles with context and fluency. The introduction of neural machine translation (NMT) marked a turning point, as it allowed models to learn to translate whole sentences rather than individual words, improving coherence and context. ChatGPT's role is to understand and generate text in multiple languages effectively, maintain context over longer conversations, and refine its translation quality through continuous interaction with users. It offers multilingual support, contextual awareness, and adaptability, making it a versatile tool for global communication. However, it faces challenges such as ambiguity and cultural nuances, which could lead to potential inaccuracies in translation. Despite these challenges, ChatGPT continues to advance in AI translation capabilities.

Machine translation and AI have become one of the threats as well as opportunities for human translators. Some voices claim that MT & AI may carry out some translation of ST in an excellent way. Recently, many scholars have focused on MT and MTPE, among them are Cholewska (2021) and Ginovart and Colominas (2020), Hutchins (2001), Hutchins (2003), Jibreel (2023), Kocmi, et al (2022), Lee and Liao (2011) and Zou (2022). Lyu, et al (2023), Qiu, et al. (2019) compared MT to HT and found a great development of MT that should be taken into consideration. For Shin and Kim (2017), the emergence of an artificial intelligence translation system with its interactive advantages that you may explain the context to it, and it can create another direction of improvement in which a human is still the dominant. In addition, Zhao, et al (2023) have attempted an evaluation of the translation quality of ChatGPT and got significant results. Moreover, Xiao (2021) compares AI translation to manual translation. They found that AI translation is challenging to quickly improve due to its reliance on sound input, making it difficult to fully comprehend and accurately translate information. In contrast, manual translation benefits from using multiple senses to comprehensively judge and analyze information, resulting in more precise and accurate translations. However, in formal occasions with high professional requirements, manual translation is still the preferred choice, and AI translation is less applicable. Therefore, human translation remains essential.

TQA Concept and Models

Translation theorists and practitioners apply different methods based on different concepts to assess a piece of translation, but they agree that whatever the method is, appropriate standards are required to measure the degree of translation quality. That is why House (2015: 1) reports that "Translation quality assessment can thus said to be at the heart of any theory of translation". Theoretically, "Quality assessment means a check of selected parts (or perhaps the whole) of a translation, or by someone standards, as well as the standards of the translating organization and the client, have been met with respect to one or more parameters.", Mossop (2013).

Almost, either those who depend on a holistic method or those who depend on analytics agree that errors are the vital element in the quality translation judgment. Therefore, Séguin (1990: 98) argues that errors obviously give us information about translation quality and also are windows into the translating process. Based on an experimental study, Christopher (2003: 424) finds that "…mistakes are doubtless the main factor in influencing our judgment of translation quality, and I would like to a plea for a more positive attitude towards them."

The concept of "errors" and, sometimes, the matter of overlap between "errors" and "mistakes" among several scholars, about the types of errors and their impact on translation result in different descriptions of error types. Newmark (1988) distinguishes referential errors from linguistic errors. By error-referenced type, he means information errors that are related to real-world knowledge that is related to facts, and fiction knowledge has to be considered here. On the other extreme, linguistic errors are related to linguistic areas. Otherwise, Pym (1992) distinguishes between binary and non-binary errors. Clearly wrong errors are binary but non-binary errors refer to the varying degrees of (in)adequacy of a piece of translation. Related to the effect of errors, Sager (1983) mentions three types of effects caused by errors in translation: linguistic effect, semantic effect and pragmatic effect. Recently, O'Brien (2012: 62) differentiates between three types of errors: minor, major and critical. While minor errors are technical and have no negative effect on meaning, major errors affect the meaning negatively but the whole message is still understandable. Critical errors, on the other hand, have a negative impact on meaning and product usability.

- **Human Assessment**

  Several models of translation quality assessment have been proposed by specialists and scholars of translation studies. They have made significant efforts to develop a model for assessing and evaluating the quality of translated texts. Their efforts have resulted in various models, such as House (1981), Nord (1997), Al-Qinai (1999), Reiss (2000) and Williams (2009). While Reiss's (2000) model

https://doi.org/10.20428/jss.v30i3.2545

focuses on qualitative assessments, others like Williams's (2009) propose quantitative evaluation approaches. However, Nord (1997) and House (1981, 1997, and 2015) combine both qualitative and quantitative aspects. O'Brien (2012) offers a practical model also. Despite the theoretical differences among these models, they all aim to judge the quality of the TT translation.

Mateo (2014) reviews the metrics of translation quality assessment including SICAL (Système Canadien d'appréciation de la Qualité Linguistique4), LISA (developed by Localization Industry Standards Association) QA model, SAE J2450, the Quality Assessment Tool (QAT) and TAUS (Dynamic Quality Evaluation Model). He identifies the pros and cons of these TQA metrics concluding with the viewpoint of Hönig (1998) that "Metrics bestow systematicity and reproducibility on a process that necessarily requires human intervention". In addition, he proposes an outline for a TQM that attempts to take into account the following important two components: a quantitative tool or metric and a qualitative grading rubric. In the rubric, there is an error category along with a credit point and deduction point that helps the rater to conclude with a final decision of the translation quality.

- **Automatic Assessment**

  Human editors who are proficient in both SL and TL are the best to evaluate the quality of machine translation (MT). However, automatic evaluation systems such as (BLEU, NIST, METEOR, TER, and WER) "are often used because they are faster, cheaper, and language Independent", Maurya, Ravindran, Anirudh, & Murthy (2020). Even though there are challenges and developments in the fields of Statistical Machine Translation (SMT) and Neural Machine Translation (NMT), the evaluation of MT and AI outputs will combine automatic metrics with manual evaluation.

Machine Translation Post Editing (MTPE)

Machine Translation Post Editing (MTPE) increasingly emerges as a solution of MT problems and challenges. (ISO 18587 (2017) and scholars such as (Koby, 2001 & TAUS, 2010 & Zhang & Torres-Hostench 2022) agree that MTPE is a human editing of machine translation. It helps qualify productivity and increases productivity, Federico, et al (2012), Läubli et al., (2013), Zampieri & Vela, (2014) and Zhechev, 2014). Zaretskaya (2017) differentiates between two types of MTPE quality, namely Light and Full post-editing. Light PE involves only rectifying major errors made by machine translation to ensure the translated text is understandable. The quality of the translation after Light PE is not as good as expected in a regular translation project. It may come across as too literal or unnatural and may have minor objective errors. On the other hand, Full PE has higher quality requirements and the translation after Full PE must be free from any errors or stylistic flaws.

Post-editing of a translation can be carried out for human or non-human. Human translation comes in the form of a translation product of any text type; non-human output can be in the form of MT or AI translation or translation memories matches. For each type of PE, the focus may differ as explained by Mossop (2013). In both types, the main focus is on identifying errors and correcting them; or in some cases, seeking an improvement for a translation with correct equivalents but poor style. Cid and Ventura (2020) study MTPE skills in course descriptions and educators' viewpoints. They find that the three most important PE skills are identifying MT output errors, decision-making about editing or discarding MT results and applying PE guidelines.

Methods

Having taken the above literature review, this study utilizes a mixed-method analysis. The researcher collected data using MT translation & AI for 30 allusions in context and from translators allowed to edit the MT or AI translation of the allusion sentences. The collected data of both were corrected, analyzed and evaluated to decide on their quality adopting O'Brien's (2012) model of TQA with slight modifications to suit allusion translation. Errors from MT and AI in addition to humans in the MTPE process were identified, categorized and rated with a focus on allusion translation mistakes. To categorize the translation strategies used by HTs, Leppihalme's (1997) strategies of translating allusions were adopted.

To specify the post-editing skills associated with MT-AI translated allusions, the researcher analyzed the participants' reports in addition to following the focus group method to let the 40-translators talk about the skills practiced during their MTPE. Focus group or group discussion is used to obtain qualitative data from a group of people about the topic being discussed, Dawson (2002). The participants were invited to discuss the skills of MTPE. A colleague specialist in translation was the moderator. Participants' discussions were recorded, summarized and analyzed thematically.

ST Text Selection

The (30) selected English sentences including allusions were carefully and randomly selected from different sources to test MT & AI translation compared to HAMT in which the translator's role focuses on Post-Editing (PE) skills. The researcher was circumspect to include all 5 types of allusions viz. religious, historical, literary, mythological and popular culture allusions. The researcher prepared a checklist of the 30 along with their standard translations and sent them to 5 validators to decide on their relevance, clarity, appropriateness and translation correctness. Based, on their review and evaluation, modifications were carried out.

MT & AI Selection

MTs are mostly AI-based. However, with the emergence of AI translation-based services, the idea of separating testing MT from AI translation came. For this reason, Cady, Tsou, and Lee (2023) compared Chinese-English MT performance involving ChatGPT and MT providers and the Efficacy of AI-mediated post-editing. Many fear

https://doi.org/10.20428/jss.v30i3.2545

AI translations in a way they think they are better or will be better than that of HT. What makes AI translation different from that of MT is the possibility you can ask the former to give alternative translation or more possible choices.

Among several MTs such as Bing Microsoft, Google Translate, Reverso Context, Yandex, Systran, …etc), Google Translate is very famous and has been improving. For this study, the researcher used Google Translate for its efficiency, easy access and availability (Jibreel (2023) and Mahdi (2023). For AI selection, it was based on the fame of AI (ChatGPT) translation in addition to its rapid development.

Human Translation Sampling

Initially, 50 graduate students, from the Department of English & Translation at the University of Science & Technology, Hodeidah established in 2012-2013, were randomly selected. The graduation lists have been taken from the Archive of Graduation Affairs Section for the first five batches with their contact details. The researcher randomly selected a proportional sample using a computer algorithm, 10 graduate students for each batch. After selection, the researcher contacted all of them. 5 of them have apologized justifying that they are no longer working as translators and the other 5 asked for permission justifying that they have no experience. Therefore, the actual sample of this study was 40 participants.

Training on MTPE

MT followed a neurological automatic-based system that generated the TT and the error-risk of the outcomes cannot be ignored. Therefore, studies evaluated the effectiveness of MTPE training reported that "PE training may be an effective way of helping identify and correct MT errors", Zhang and Torres-Hostench (2022: 12). Besides, researchers such as Sycz-Opoń and Galuskina (2017) and Rico et al. (2017) have suggested the idea of specific MTPE training. In this respect, On 17-18 February 2024, the 40 participants were invited to an explanatory lecture about MTPE skills and allusions. On the first day, the researcher declared the research purpose and assured the participants about following research ethics and asked them to sign an informed consent. All of them have proved and signed. Then, the researcher explained the concept of MTPE and how to perform its skills. The lecture continued for two hours.

On the second day, the sample was divided into two groups of 20 students. The researcher has arranged a practical session in the Computer Lab. Two-hour training for each group from 8 am to 1 pm, with a one-hour break in between. Ten sentences including allusions (later excluded from the main Test) were practiced with a focus on MTPE utilizing any resources of data to assure the meaning of the ST, attest the proper nouns' indication and deciding on the TT.

The Setting Test

Two days after training, the 40 graduate students were given the test (30 allusion sentences) and asked to translate them from English into Arabic providing them with

Google Translate and ChatGPT translations asking them to apply MTPE skills to reach the correct equivalent. They have access to the internet to make use of all resources.

Adopted TQA Model

Two different translators participated in this study: MT including (Google Translate & ChatGPT) and Human translators. Therefore, qualifying the TT outputs was not that easy task. After reviewing the TQA models as outlined in section 3.2, the researcher decided to adopt a model that will be suitable for the translation quality of allusion translation in the context of sentences. The researcher followed O'Brien's (2012) model with some adaptations.

Correction

Determining Errors

To determine the TQ, errors were identified, corrected, classified and counted to qualify and quantify the problem to reach fair judgment on the TQ of both MT and AI translation and on HT utilizing MTPE skills on the other hand. Table 1 below illustrates the details.

Table 1 Error Type and TQ Assessment

	Error Severity	Definition	Error Type	Fixed Penalization Schema / Cut Marks	Deserved Marks	Translation Quality
Sharon's Error Classification[1]	Minor	are technical and have no negative effect on meaning,	**Language:** Grammar, Spelling, Syntax and Punctuation	1	2	Moderate
	Major	affect the meaning negatively but the whole message is still understandable.	**Accuracy**: Inaccurate references, style or meaning but the general meaning still understood.	2	1	Low
	Critical	have negative impact on meaning and the product usability.	Mistranslation of the allusion word/phrase that negatively affects the	3	0	Poor

[1] It is to note that just error description and definitions are adopted from O'Brien (2012). The other evaluation model components are suggested by the researcher based on other published studies in this regard as well as his own experience in translation quality assessment.

https://doi.org/10.20428/jss.v30i3.2545

			meaning of the TT.			
Added by the researcher	Error Free	Where no minor, major or critical errors are found	No Error	0	3	High

There were 30 "Allusions" in the context of a full sentence being given to both HTs and AI & MT. The outputs of the latter and the translations of the former were corrected and given marks. Three marks were given to the effective High-Quality translation or the translation without any error; 2 marks for the Moderate-Quality translation or the sentences with minor errors; 1 mark for the Low-Quality translation or the translation with major error and 0 for the very Low-Quality translation or the translation with critical error. Accordingly, the total marks for the AI & MT were calculated as well as that of human translators. The total marks expected for each is 90. Therefore, the grading rubric for quality translation is shown in Table 2 below[2].

Table 2a Quality Description Values and the Grading Rubric

Quality Description	Poor	Low	Moderate	High
Scores	0 ≤ 22.25	22.26 ≤ 44.51	44.52 ≤ 66.77	66.78 ≤ 90

The total number of scores is out of 90, so the final assessment for each translator, AI or MT equals the sum of their obtained scores in the thirty translations of allusions. Likewise, four descriptions are given to each translator's performance gained quality based on the scores obtained: high quality, moderate, low or poor quality.

Table 2b Grading Rubric for the Means of HT

Quality Description	Poor	Low	Moderate	High
Mean mark for each sentence	1	2	3	4
Scores obtained out of (120)	0 ≤ 30	31 ≤ 61	62 ≤ 92	93 ≤ 120

To illustrate, the number (120) comes from (90x40/30) where 90 is the total score that each HT may obtain, 40 is the number of HTs and 30 is the number of the allusion sentences. Thus, 120 is the resulting score of this formula if the total TQ of HT is evaluated as High.

- ***Statistical Analysis***

Comparing the results, frequencies and percentages of the corrected translations were calculated to assess the overall evaluation of TQ. The *ONE-WAY ANOVA* was performed to compare the results and see whether there are significant differences among participants (AI, MT & HT) regarding translating allusions. Within the same group viz. HT utilizing MTPE, the One-Sample *t-Test* were used for significant differences and Post Hoc Test of multiple comparisons to determine for which translation the difference is in favor.

Validity & Reliability

For validity, the selected allusions and their standard translations, accompanied by the research title and aims were given to 4 experts in a prepared checklist to decide on their relevance, clarity and correctness of their translations. Then, the list of sentences is modified accordingly.

Reliability was checked by two experts in literary translation using Rater Agreement Procedure by correcting to double-check the translation test after providing them with the TQA model adopted- the rubric and metrics.

Table 3 Reliability Testing of Translation Evaluation

		AllusionRater1	AllusionRater2
AllusionRater1	Pearson Correlation	1	.938(**)
	Sig. (2-tailed)		.000
	N	90	90
AllusionRater2	Pearson Correlation	.938(**)	1
	Sig. (2-tailed)	.000	
	N	90	90

*** Correlation is significant at the 0.01 level (2-tailed).*

As evident from the Pearson Correlation test, the correlation is positive (.938) and the **p-value* shows statistical significance (.000).

Results

Translation Quality

Over all Comparison

Table 4 Translation Quality of AI, MT & HT

Translator	Obtained Marks (out of 90)	Percentage	Translation Quality
AI (ChatGPT)	40	44.44%	Low
MT (Google Translate)	38	42.22%	Low
HT (40 Translators) Rater2	60	66.67%	Moderate

*At first, it is better to note that the mark shown in Table 4 for HT is the total means for the translations of the 30 allusions carried out by the 40 translators who

https://doi.org/10.20428/jss.v30i3.2545

participated in this study.

As Table 4 shows, the obtained marks of the translation resulted from ChatGPT is 40, 38 is for Google Translate and 60 is for human translators. To put it another way, MT has got 42.22% (N=38); AI 44.44% (N=40) and HT has got 66.67% (N=60). Based on the grading rubrics of translation quality outlined in Tables 2[a] and 2[b], HT is considered of Moderate Quality followed by ChatGPT and Google Translate which have got (Low Quality) respectively.

HT utilizing MTPE

One-Sample t-Test

Table 5 One-Sample t-Test

	Test Value = 79.23					
	T	*Df*	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
Marks of HT of (30) Allusions	.001	29	.999	.003	-6.28	6.29

The general result of the *One-Sample t-Test* shows no statistically significant differences among the HTs' mean scores at the significant level (.05) where t=.001, *df*=29, m=.003 and the (*P*-value=.999).

MT & AI Vs. Human Translation

Normality & Homogeneity Tests

To compare the translations of MT, AI and HT utilizing MTPE in addition to testing the differences among the three, *One-Way ANOVA* was used. Before deciding on this type of statistics both normality and homogeneity tests were statistically satisfied.

Differences Among the Three Translations

Table 6 Mean Differences

	N	**Mean**	**Std. Deviation**	**Std. Error**	**95% Confidence Interval for Mean**		**Minimum**	**Maximum**
					Lower Bound	**Upper Bound**		
AI	30	1.33	1.028	.188	.95	1.72	0	3
MT	30	1.27	1.048	.191	.88	1.66	0	3
Human Translator	30	2.00	.455	.083	1.83	2.17	1	3
Total	90	1.53	.939	.099	1.34	1.73	0	3

Table 6 reveals that the mean of HT scores of allusions is significantly higher (2.00, Std. Deviation = 0.455) than that of AI (1.33, Std. Deviation = 1.028) and MT (1.27, Std. Deviation = 1.048). The results indicate–a better translation quality of HT compared to MT and AI. This result is supportive for the result of Table 2.

Analysis of Variance

Table 7 ONE-WAY ANOVA

	Sum of Squares	Df	Mean Square	F	Sig.
Between Groups	9.867	2	4.933	6.263	.003
Within Groups	68.533	87	.788		
Total	78.400	89			

The test of variance (One-Way ANOVA) indicates a statistically significant difference in mean scores among the three translations of allusions by Google, ChatGPT and human translators utilizing MTPE at the level of (> .05) where *F*=2, 87=6.263, *P-value*= (.003). This result leads to enquiring whether this significant difference is in favor of MT, AI or HT utilizing MTPE. Therefore, to get an exact answer, the Post Hoc Test was carried out.

Multiple Comparisons: Post Hoc Test

Table 8 Multiple Comparisons Post Hoc Test

	(I) AI, MT & Human	(J) AI, MT & Human	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Tukey	AI	MT	.067	.229	.954	-.48	.61
		Human Translator	-.667(*)	.229	.013	-1.21	-.12
	MT	AI	-.067	.229	.954	-.61	.48
		Human Translator	-.733(*)	.229	.005	-1.28	-.19
	Human Translator	AI	.667(*)	.229	.013	.12	1.21
		MT	.733(*)	.229	.005	.19	1.28

* The mean difference is significant at the .05 level.

The results from the one-way ANOVA do not indicate which of the three groups differ from one another. Therefore, in many cases, it is of interest to follow the analysis with a Post Hoc test or a planned comparison among particular means. In this study, *Tukey* post hoc test was performed. As in Table 8, there was a statistically significant difference in the translations of HT utilizing MTPE compared to AI (*p*-value = .013) outcomes for the 30 allusions and for HT vs MT (*p*-value = .005). In both cases, the comparison was in favor of HT using MTPE. However, there were no significant differences between the outcomes of the MT and AI (*p-value*= .954).

Error Type

Table 9 AI, MT & Human*Crosstabulation

Error Type		AI, MT & Human			Total Freq.	%
		AI	**MT**	**HT**		
Critical	**Freq.**	8	10	0	18	20.0
	%	44.4	55.6	0	100	
Major	**Freq.**	8	5	3	16	17.8
	%	50	31.3	18.7	100	
Minor	**Freq.**	10	12	24	46	51.1
	%	21.7	26.1	52.2	100	
Error Free	**Freq.**	4	3	3	10	11.1
	%	40	30	30	100	
					Total	100

Table 9 explains that Minor errors are the most frequent among other types 51.1% (N=46), followed by Critical 20% (N=18), Major errors are 17.8%(N=16) and Error-free with the least percentage 11.1% (N=10).

MTPE Skills Utilized

To identify the MTPE skills applied in translating allusions, the researcher depended on two tools: an analysis of the translator's report about his/her translation of allusions and the focus group conducted for this purpose.

Analysis of Translators' Reports & Focus Group

In this study, the focus group method was used to collect data for information that the analysis of participants' translations and their sentence-aligned report may not provide.

a. Translators' Reports Analysis

Reporting about what the translator has made while MTPE may help to absorb some significant issues. Unfortunately, participants in this study showed less information in this regard. Their reports namely mentioned that they have gone through the Web to ensure the meaning of the ST allusion using Google search, translation cafes and language forums, online dictionaries, blogs, corpora... etc. Figure 1 specifies the participants' sources to determine the ST exact meaning.

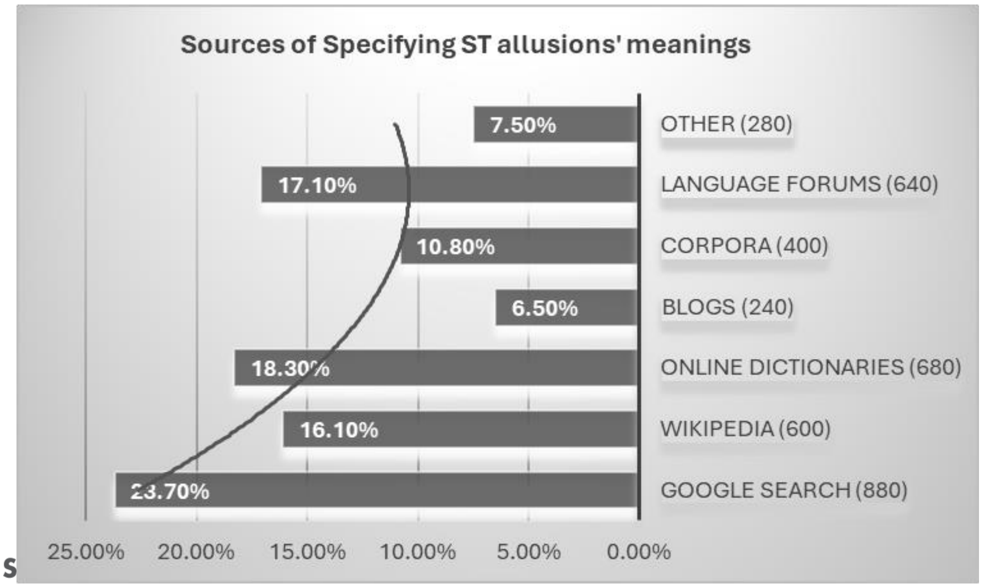

Figure 1 Sources of Specifying ST Allusions

Sources of Identifying Allusion Meaning

During the process of identifying the exact meaning of the allude words or phrases, as evident in Figure 1, Google search comes at first among the other sources with a percentage of 23.7% (N=880) followed by online dictionaries 18.70% (N=680) and Language Forums with 17.10 (N=640), and Wikipedia with 16.10 (N=600) respectively. Searching a corpus and other sources (like direct contact of a translator, consulting people in a WatsUp language or translation group or referring back to a translation memory) comes with low rates of 10.80%(N=400) and 7.50%(N=280) and is blogs 6.6% (N=240).

Utilizing Translation Strategies

For Hatim & Mason (1990), it is important to solve translation problems to identify who is translating what, for whom, when, where, why and in what circumstances. Applying that, the researcher depended on Leppihalme's (1997) strategies of translating allusions, Table 10 summarizes the strategies utilized during the MTPE of HT in this study.

Table 10 Strategies of Translating Allusions

Allusion Type	**The Strategy**	**Freq.**	**%**			**Rank**
Key Phrase (KP)	▪ Retention of the allusion	149	12.52 %			6
	▪ Changing it somehow	36	3.03%			8
	▪ Omitting the allusion altogether from the basis of both lists	158	13.27%			4
Proper Noun (PN)	1. **Retention of the name**: *a.* Retention of the name as such	190	15.97%			1
	b. Retention of the name with some additional guidance	170	14.29%	364	42.77%	3
	c. Retention of the name with detailed explanations (footnotes etc.)	4	0.34%			10
	2. **Replacement of the name by another name:** *a.* Replacement of the name with another source-language name	26	2.18%	155	18.21%	9

b. Replacement of the name with a target-language name Juliet and Romeo	125	10.5%			7
3. Omission of the name: *a.* Omission of the name, but the sense conveyed through a common noun	151	12.69%	332	39.01%	5
b. Omission of the name and allusion completely.	181	15.21%			2
TOTAL	**1190**	**100%**	**851**	**100%**	

Table 10 shows that the most 5 strategies utilized while translating allusions are "*Retention of the name as such*" with a percentage of 15.97%(N=190) followed by "*Omission of the name and allusion completely*" with 15.21%(N=181), "Retention of the name with some additions" with 14.29% (N=170), "*Omitting the allusion* altogether from the basis of both lists " in KP with 13.27%(N=158) and "omission of the name and translating the sense" with 12.69%(N=151) respectively.

As nearly most of the ST allusions in this study are PNs, one could notice that using the same name mentioned in the allusion comes in the first rank with a percentage of 42.77% (N=364). In contrast, omitting the PN of the allusion obtains the second rank with a percentage of 39.01% (N=332). Substituting the PN allusion by another PN either in the ST or TT gets the lowest rank with 18.21% (N=155).

b. Focus Group Analysis

In their reports against each allusion translation, most of the translators have mentioned some of the MTPE used to edit the AI and MT outcomes. No one of them states the detection of errors as a step in MTPE, but in the focus group most of them, if not all, agree that error detection is the first step through reading the output of AI and MT.

Identifying MT/AI output errors

Through the discussion, there were different ideas about the way of finding the errors, but most of the translators indicated that this step was the first to be performed. It entails the following process:

- Understanding the ST
- Identifying the meaning of the allusion word in the ST
- Comparing the ST to the TT

Identifying the ST meaning and Equivalent for the Allusion

During the discussion, most of the participants agree that identifying the meaning of the allusion in the ST leads to success or failure in completing the translation task.

Without full understanding, the target equivalence cannot be determined.

Since the translators have access to the internet, most of them state that they search for the allusion that comes in bold, either in the form of a PN or in the form of KP to understand their meaning in English and identify the equivalent in Arabic.

In some few cases, they succeed in getting a cultural substitution such as in the case of Allusion [21].

*I'll be your **Romeo** if you'll be my **Juliet**.*

Romeo and ***Juliet*** in English literature is a famous play of Shakespeare in which love is the dominant theme. The participants varied in their translations. Some of them go to foreignize the meaning and wrote the same names in the ST:

كوني لي كروميو أكن لك جوليت. They did not modify the output of the MT and AI.

On the other hand, some participants domesticated the translation providing some Arabic PN in famous love stories in Arabic literature such as: Qais and Laila, Gameel and Buthainah, Kuthayer and Azza, Antar and Ablah etc. The following are some examples of the participants' translations transcripted in English:

كوني لي ليلى أكن لك قيسا	*Kooni lee Laila akun laki Qaisa*
كوني لي كبثينه أكن لك كجميل	*Kooni lee ka Buthainah Akun akun laki ka Gameel*
كوني لي عبلة أكن لك عنتر	*Kooni lee Ablah Akun laki Antar*

Participants differed in their thoughts and beliefs; they provided a religious concept of love in this regard. Consider this example:

كوني لي كـ خديجة أكن لك كـ محمد	Kooni Lee Khadeejah Akun Laki Ka Mohammed

They provided the PN of the prophet *Mohammed* and his first wife *Khadijah* as a symbol of love although it is not famous in Arabic literature.

In some cases, the PNs of Romeo and Juliet were omitted and, instead, the meaning

كوني لي عشيقةً أكن لك عشيقاً	*Kooni lee Asheeqatn akun laki Asheeqn*
سأحبك كثيراً إذا أحببتني كثيراً	*Sa Ohebbki Katheern Etha Ahbbtinee Kaheernl*

Some of the participants discussed their searching the web for the word/phrase/proper noun written in bold in the ST sentence, where necessary until they discovered the exact alluded meaning. Then, they rephrased the nearest translation produced by AI or

https://doi.org/10.20428/jss.v30i3.2545

MT.

Some sentences in the test were clear and direct. Some of the translators said they depended on their understanding of the ST and noticed the literal translation provided by AI and MT. For them, there was no need to go through the web. Instead, they did their best to modify the outcomes of AI and MT. [8, 21, ……etc.]
Where the alluded words were not easily found, translators may consult encyclopedias, as said in the discussion, especially *Wikipedia* to get the specific meaning. Encyclopedia is one of the important CAT tools that help the translator where dictionaries and MT engines may fail. [1, 29 and 30].

Decision-making about editing

Pragmatically, Levý (1967) views translation as a decision-making process in which the translator has to choose from several alternatives. In this regard, some indicated that while performing their PE of the MT and AI outputs, they were faced with the dilemma of making the correct decision. For example, as some stated " in sentence No: 21, **Romeo** & **Juliet** are symbols of love, romance and sacrifice. Both MT and AI have transliterated the two names. The PN equivalents that may come into mind in Arabic literature are PN like Qais and Laila, Gameel and Buthainah, Kuthayer and Azza, etc. So, it was very difficult to decide on an alternative. Instead, some participants said that they kept the source text PN to avoid such a problem.

Rewriting the final draft

Once the meaning becomes clear and the target equivalent is determined, the translator rewrites the final draft of the MT or AI outputs whatever is nearer to the best TT. During the discussion, mostly all the participants claim that they, after all, rewrite the final sentence in Arabic paying attention to the TL grammar, structure and style. Analyzing the TTs, it is noticed that most of the participants depend on the MT and AI outputs' structure and style and their modification focuses on the "word or phrase" allusion.

Discussion of the Results

Translation Quality of AI, MT Vs. HT

The findings showed that HT of allusions is considered of Moderate Quality with a percentage of 66.67% (N=60) followed by ChatGPT with 44.44% (N=40) and Google Translate with 42.22% (N=38); both of Low Quality respectively. Besides, ANOVA test indicates a statistically significant difference in mean score among the three translations of allusions by Google Translate, ChatGPT and HTs utilizing MTPE at the level of significance (.05), *P-value*= (.003). In addition, *Tukey* test proved a statistically significant difference in the translations of HT utilizing MTPE compared to AI (*p*-value = .013) outcomes for the 30 allusions and for HT vs MT (*p*-value = .005). In both cases, the comparison was in favor of HT using MTPE. However, there were no significant differences between the outcomes of the MT and AI (*p-value*= .954). These results are in line with Lee and Liao (2011) in terms of scores obtained, there was a significant difference between the *No MT* and *with MT* sets in favor of *with MT* Groups. Moreover, studies like Samman (2022), proved the MTPE efficiency in translation quality via error analysis. For Vasiliauskienė (2023) remarkably, post-editors show more competence with instructional manuals translation.

Error Type and Strategies

In this respect, HT utilizing MTPE has got the lion's share of the minor errors 52.2%(N=24) compared to MT (26.1%(N=12) and AI 21.7% (N=10) respectively. Regarding Major errors, HT obtained 18.7%(N=3) followed by MT 31.3%(N=5) and AI 50% (N=8). Critical errors have been distributed between AI and MT with a higher proportion for MT outcomes 55.6%(N=10) over AI 44.4%(N=8). However, no critical error has been registered for HT in the total mean. Their translation is either free of error, minor or major. These findings are in line with Lee and Liao (2011) who found MT translations very useful in reducing translation errors when followed by human editing. Based on Pym's (1992) binary and non-binary errors, they found, their results showed that the Group *with MT* set performed better than the Group *Without MT*, and the number of binary errors was nearly halved in the with MT sets.

As for translation strategies utilized, participants mostly retained *the allusion name as it is* 15.97%(N=190) or sometimes retained *the name with some additions* 14.29% (N=170). In contrast, they omitted *the name and allusion completely* 15.21%(N=181) or in KP they *omitted the allusion altogether from the basis of both lists* 13.27%(N=158). In other cases, participants *omitted the name and translated the sense* 12.69%(N=151). These results are similar to Bahrami (2011) where the "retention of PN allusions without any guidance" came in the first rank and 'literal translation with minimum change' was the first in KP allusions. For Hamidreza (2023) the 'retention of names' and the 'replacement of names by another' were the main strategies for PN allusions and 'literal/minimum change translation' strategy for KP allusions.

MTPE Skills needed

Both translators' reports and the focus group discussion revealed that translators applied some integrated skills to improve MT and AI outcomes of allusions. They include but not limited to:

a) Identifying MT/AI output errors,
b) Identifying the ST allusion meaning,
c) Decision-making about editing and
d) Rewriting the final draft

Broadly, Fradana (2023) theoretically investigated the abilities needed by translators utilizing MT and producing high-quality TT. He concluded that the competencies needed by the translator in the digital era are language proficiency, digital literacy, context and cultural competencies, research skills, specialization expertise,

https://doi.org/10.20428/jss.v30i3.2545

collaboration competence and continuous learning competence. Cid, Colominas, and Oliver (2020) studied the profile of the post-editor with a translation industry view. They suggested *six skills of PE viz.* 'the capacity (1) to post-edit up to human quality (full PE); (2) to post-edit according to PE guidelines; (3) to decide when to work on a segment or discard it; (4) to identify MT output errors; (5) to post-edit to a good enough quality (light PE); and (6) to apply the right correction strategy", pp (17-18).

Conclusion

This study attempted to compare MT and AI translations to HTs utilizing MTPE skills regarding allusions from English into Arabic. It attested whether there were significant differences between the mean scores of the three. It mainly focused on MTPE skills to significantly contribute to the literature of translation studies in this crucial state-of-the-art issue. A test of (30) allusion sentences were translated by MT and AI and the TQ was assessed. Then, the MT and AI-based translations were given to 40 HTs to post-edit them utilizing MTPE skills. Their TQ was evaluated and compared to the TQ of MT and AI. To determine the MTPE skills utilized, HTs were asked to report the PE skills, translation resources and strategies used to accurately find the TL equivalent allusion in addition to a focus group discussion to get in-depth information about the MTPE skills utilized. After a statistical analysis, results have led to a conclusion that MT and AI translation quality regarding allusions is ***Low*** compared to a *Moderate Quality* of HT utilized MTPE skills. Although the results indicated statistically significant differences at the level of significance (.05) between MT, AI and HTs in favor of HT, it seems that the findings are not satisfactory. It is expected that HT takes an advanced level of quality and translators utilizing MTPE show better performance. This may be attributed to the fact that translation courses over the past years have included a little focus on how HTs can make use of MT and AI tools represented in one course for electronic tools or translation technology. For MT and AI, no significant differences were recorded in their TQ of allusions, which may be because most of the MT and AI translation services depend on similar databases and generated algorithms and systems. Nevertheless, AI-based translation showed a better level than MT (44.4% over 46.4%). These findings are compatible with Zong (2018), Samman (2022) and Abdelali and Bennoudi (2023) although different text types were investigated. Concerning the error type, it is worth mentioning, that the most frequent error types were Minor, Critical and Major respectively. However, HTs recorded no critical errors in allusion translation while MT and AI recorded critical, major and minor errors. Thus, PE may be effective and beneficial in qualifying HT. In addition, HTs' reports about the resources and strategies used while translating allusions revealed the huge effort devoted where Google search comes at first followed by online dictionaries, Language Forums and Wikipedia respectively while searching a corpus and blogs comes with low rates along with direct contact of a translator, consulting people in a WatsUp language or translation group or referring back to a translation memory. Furthermore, HTs exploited strategies to pick up the TL exact equivalent of the ST allusion. Most important are "*Retention of the name as such*, "*Omission of the name and allusion completely*", "Retention of the name with some additions", "*Omitting the allusion altogether from the basis of both lists* " in KP and "*omission of the name and translating the sense*". As nearly most of the ST allusions in this study are PNs, one could notice that using the same name mentioned in the allusion came in the first rank. In contrast, omitting the PN of the allusion obtained the second rank and substituting the PN allusion by another PN either in the ST or TT has got the lowest rank. Moreover, the Focus Group analysis showed that the MTPE skills used by HTs include identifying the ST allusion meaning and equivalent, identifying MT/AI output errors, and decision-making about editing and writing the final draft. For each PE skill, there were several procedures enhanced. For example, to identify the Allusion ST meaning and get an acceptable equivalent, HTs agree that they try understanding the ST, identifying the meaning of the allusion word in the ST, comparing the ST to the TT, using several sources and deciding on one available option.

Based on these findings, it is recommended to revisit the current courses concerning MTPE skills in light of the accelerating development of MT and AI translation tools focusing on how these results might impact the future of translation practices. Translation of culturally bound structures such as allusions should have more practice among translator trainees as well as translation programs and HT shall benefit from all translation technology resources increasingly available.

A foresight consideration suggests a study that tests the TQ regarding allusion translation using several MTS and AI-based translation websites instead of depending on ChatGPT and Google Translate only. It is also recommended to expand the exploration of the MTPE skills in a separate study that may provide a handy guide for translation students as well as translators in this subject-matter state-of-the-art issue in translation studies.

Funding

This research received grant No. (23/2023) from the Arab Observatory for Translation (An affiliate of ALECSO), which is supported by the Literature, Publication and Translation Commission in Saudi Arabia.

References

Abdelali, Z. A. I. D., & Bennoudi, H. (2023). AI vs. Human Translators: Navigating the Complex World of Religious Texts and Cultural Sensitivity. *International Journal of Linguistics, Literature and Translation*, *6*(11), 173-182. https://doi.org/10.32996/ijllt.2023.6.11.21

Abrams, M., (1957). *A Glossary of Literary Terms: MH.*: Holt, Rinehart and winston.

Al-Haj, N., Jibreel, I., Sharafuddin, M., & Al-Shameri, H. H.. (2021) *Translation of twenty yemeni short stories: evaluation of the problems and the strategies employed.* Language & Translation,. *9*(2), 19-30.

Austermuhl, F., (2014). *Electronic tools for translators*. Routledge.

Baker, M. (1992). *In Other Words,* London & New York: Routledge.

Bahrami, N., (2012). Strategies used in the translation of allusions in Hafiz Shirazi's poetry. *Journal of Language and Culture, 3*(1), 1-9.

Cholewska, D. (2021, August). Machine translation post-editing (MTPE) from the perspective of translation trainees: Implications for translation pedagogy. In *Proceedings of machine translation summit XVIII: Users and providers track* . 200-210.

Cid, C. G., Colominas, C., & Oliver, A. (2020). Language industry views on the profile of the post-editor. *Translation Spaces*, *9*(2), 283-313.

Cid, C. G., & Ventura, C. C. (2020). The MT post-editing skill set: course descriptions and educators' thoughts. In *Translation Revision and Post-editing* (226-246). Routledge.

Dawson, C. (2002). *Practical research methods: A user-friendly guide to mastering research*. How to book

Federico, M., Cattelan, A., & Trombetti, M. (2012). Measuring user productivity in machine translation enhanced computer assisted translation. In *Proceedings of the 10th Conference of the Association for Machine Translation in the Americas: Research papers*.

Fradana, H. (2023). Translator Competencies Utilizing Translation Machines In The Digital Era. In *Proceeding of International Seminar on Adab and Humanities* ,*5*(1), 201-212.

García, M.Á. (2022). *AI Translation: The Future of Language Learning*. PANGEANIC.

Ghazala, H. (1995). Translation as problems and solutions. *Syria: Dar El-Kalem El-Arabi. Quarterly*, *17*(3), 121-139.

Cid, C. G., & Ventura, C. C. (2020). The MT post-editing skill set: course descriptions and educators' thoughts. In *Translation Revision and Post-Editing* (pp. 226-246). Routledge.

Hamidreza, A. B. D. I. (2023) Translating Allusions as Complex Cultural Resources for Translators: The Case of Virginia Woolf's Mrs. Dalloway. *transLogos Translation Studies Journal*, *6*(2), 85-102.

Hutchins, J. (2003). Machine translation and computer-based translation tools: What's available and how it's used. *A new spectrum of translation studies*, 13-48.

Hutchins, W. J., & Somers, H. L. (1992). *An Introduction fo Machine Translation.* Academic Press Limited, London.

Hutchins, W.J.(2001). *Machine translation over fifty years. Histoire epistémologie langage, 23*(1): p. 7-31.

Huges, G. (2009*). Longman dictionary of contemporary English. Lexikos, 6.*, Pearson Education Limited.

Jibreel, I., Al-Abbasi, A., & Al-Maqaleh, A. (2016). Towards a proposed refined classification of translation strategies with reference to written discourse. *International Journal of English Language, Literature & Translation Studies*, *3*(3), 437-443.

Jibreel, I. (2023) *Online Machine Translation Efficiency in Translating Fixed Expressions Between English and Arabic (Proverbs as a Case-in-Point). Theory and Practice in Language Studies, 13*(5), 1148-1158.

Kocmi, T., Bawden, R., Bojar, O., Dvorkovich, A., Federmann, C., Fishel, M., ... & Popović, M. (2022, December). Findings of the 2022 conference on machine translation (WMT22). In *Proceedings of the Seventh Conference on Machine Translation (WMT)* .1-45.

Koponen M, Mossop B, Robert IS, Scocchera G, (eds.) (2020). *Translation revision and post-editing: industry practices and cognitive processes*. London: Routledge.

Läubli, S., Fishel, M., Massey, G., Ehrensberger-Dow, M., Volk, M., O'Brien, S., ... & Specia, L. (2013). Assessing post-editing efficiency in a realistic translation environment.

Lee, J., & Liao, P. (2011). A Comparative Study of Human Translation and Machine Translation with Post-editing. *Compilation & Translation Review*, *4*(2).

Leppihalme, R. (1997). *Culture bumps: An empirical approach to the translation of allusions* (Vol. 10). Multilingual Matters.

https://doi.org/10.20428/jss.v30i3.2545

Leppihalme, R. (1994). Translating allusions: When minimum change is not enough. *Target*, *6*(2), 177-193.

Levý, J. (1967). Translation as a decision process. *To honor Roman Jakobson: Essays on the occasion of his seventieth birthday II*, 1171-1182.

Lyu, C., Du, Z., Xu, J., Duan, Y., Wu, M., Lynn, T., ... & Wang, L. (2023). A Paradigm Shift: The Future of Machine Translation Lies with Large Language Models. *arXiv preprint arXiv:2305.01181*.

Mateo, R. M. (2014). A deeper look into metrics for translation quality assessment (TQA): a case study. *Miscelánea: A Journal of English and American Studies*, (49), 73-94.

Maurya, K. K., Ravindran, R. P., Anirudh, C. R., & Murthy, K. N. (2020). Machine translation evaluation: Manual versus automatic—a comparative study. In *Data Engineering and Communication Technology: Proceedings of 3rd ICDECT-2K19* .541-553. Springer Singapore.

Newmark, P. (1988). Pragmatic translation and literalism. *TTR: traduction, terminologie, rédaction*, *1*(2), 133-145.

O'brien, S. (2012). Towards a dynamic quality evaluation model for translation. *The Journal of Specialised Translation*, *17*(1), 55-77.

Pedersen, J. (2007). Cultural interchangeability: The effects of substituting cultural references in subtitling. *Perspectives: studies in translatology*, *15*(1), 30-48.

Zhang, H., & Torres-Hostench, O. (2022). Training in machine translation post-editing for foreign language students. *Language Learning & Technology*, *26*(1), 1–17. http://hdl.handle.net/10125/73466

Zong, Z. (2018). Research on the relations between machine translation and human translation. In *Journal of Physics: Conference Series* , *1087*(6) p. 062046). IOP Publishing.

Zou, S. (2022). Analysis of Machine Translation and Post-Translation Editing Ability Using Semantic Information Entropy Technology*. Journal of Environmental and Public Health.*