



**TECHNIUM**  
SOCIAL SCIENCES JOURNAL

[www.techniumscience.com](http://www.techniumscience.com)



**Vol. 75/2025**  
**A New Decade for Social Changes**

**PLUS**  
**COMMUNICATION P**



**International**  
Communication & PR

## AI – aided translation: An eye on ChatGPT as a tool of Translating Literary Texts

Mai Abdel karim Mohammad Malkawi<sup>1</sup>, Ayman Altalihin<sup>2</sup>, Majd Saleem Hussein Abushunar<sup>3</sup>, Abdullah S. Aldelaa<sup>4</sup>

<sup>1 2 3 4</sup>Language Center, The Hashemite University, Zarqa, Jordan

[mmalkawi@hu.edu.jo](mailto:mmalkawi@hu.edu.jo)

**Abstract.** The present study aims at analyzing the output of ChatGPT translation of English literary texts into Arabic comparing it to that of other tools like Google Translate and Microsoft Bing. The paper is designed to explore both the efficiency and the pitfalls of the AI – aided ChatGPT when dealing with literary texts mainly poetry. The goal of this paper is to discover whether AI- tools offer precise translation of the examined texts. The approach used is a qualitative approach through which the results and the observations were studied and analyzed.

**Keywords.** ChatGPT, literary translation, google translate, Microsoft Bing, translation problems

### 1. Introduction

The rapid advancement of artificial intelligence nowadays has transformed various fields, including translation. Among the significant developments is ChatGPT, a sophisticated language model that offers deep learning techniques to facilitate learners' educational tasks. This paper tries to break down the efficacy of ChatGPT in translating literary texts, comparing it to other machine translation tools like Google Translate and Microsoft Bing. This area, literary translation, was chosen because it is a process through which unique and subtle challenges are inevitable due to the inherent complexity of literary language, including idiomatic expressions, cultural references, and poetic tone.

This study aims to evaluate the strength and weakness points of ChatGPT's translation through a series of comparative analyses against other machine translation tools.

### 2. Literature review:

The rise of ChatGPT as a machine learning model that can translate texts encouraged language researchers to test its efficiency and compare its translation with the translation of humans and other language models. This section provides a review of the most up-to-date research that evaluated the translation of ChatGPT:

Farrell, M. (2023) conducted two experiments: the first assessed the translation quality of the free-plan version of ChatGPT in comparison with the free versions of DeepL Translator and Google Translate through human evaluation, while the second examined the free-plan version of ChatGPT as an automatic post-editor of raw output obtained from the pay-for version

of DeepL Translator. The data of the study consisted of Wikipedia articles translated from English to Italian. The results of the first experiments showed that DeepL Translator performs better than Google Translate which in turn outperforms ChatGPT. The findings of the second experiment, on the other hand, revealed that the ChatGPT reaches average human translation (HT) levels of lexical variety in automatic monolingual machine translation post-editing (MTPE), but it surpasses average HT levels in automatic bilingual MTPE.

Khoshafah (2023) compared the translation of ChatGPT and human translations of different text types such as media, literary, scientific, religious, and legal texts. The study mainly examined the clarity and accuracy of Arabic and English translations generated by ChatGPT in comparison to human translations. The results showed that ChatGPT translation is not as good as human translation. Although ChatGPT and human translation may provide the same meaning to some extent, ChatGPT has some pitfalls regarding highly technical or specialized content, culturally sensitive outputs, Arabic grammar, idiomatic expressions, diacritical marks, and false cognates. The study recommended that human intervention is needed for accurate translation.

Valdivieso and López-Arroyo (2023) assessed the advancements of machine translation models such as ChatGPT and Google Translate when translating Spanish terminology into English. The researcher provided MQM-based, terminology-focused analysis to examine Spanish corpora of olive oil and wine tasting notes. The findings showed that ChatGPT outperformed Google Translate in error severity and the number of samples affected. Although the study exhibited great potential of machine translation systems, it indicated that people cannot rely solely on these systems especially when it comes to specialized languages.

Aghai (2024) evaluated the translation of ChatGPT and Google Translate of literary texts from Persian to English. To conduct the study, the researcher used a functional holistic model to analyze the translation of a Persian short story generated by the two platforms. He also carried out a critical error analysis to detect the areas of weaknesses in the translation of these tools. The results of the study revealed that both ChatGPT and Google Translate have some flaws regarding accuracy, equivalence, and text function. Additionally, the study revealed that ChatGPT frequently generates translations that are grammatically correct, and it may eliminate cultural nuances in order to achieve fluency. Conversely, Google Translate frequently leaves source text components as they are without change, which results in awkward translations or transliterations. The study concluded that machine translation models cannot fully substitute human translators who are needed to achieve high-quality translations that convey cultural nuances and idiomatic expressions.

Problems related to the translation of idiomatic expressions in the Arabic novels using neural machine translation are challenging to NMTs such as Google, Bing and Facebook translators. It is worth mentioning that Google and Bing translators significantly outperformed the Facebook translator in translating idioms. Generally, although NMTs had successfully rendered some idioms into the TL, but they faced some problems such as: unwanted literal translation, unjustified omission, culture bound expressions, wrong lexical prediction, Lack of cohesion, Ambiguity resulted from homonymy and polysemy, wrong choice of words, and colloquial idiomatic expressions (Aldelaa & Malkawi, 2024).

Alafnan (2024) examined the translation of ChatGPT and compared it with Google Translate. The researcher analyzed the translation of Arabic and English speeches of the King of Jordan, Abdullah II, delivered at significant international events in 2023. The study assessed the translations of Google Translate and ChatGPT based on meaning, functional and textual adequacy, target language mechanics, style, register, and idiomaticity. Regarding English-to-

Arabic translation, the study indicated that although Google Translate often gives satisfactory translations in terms of meaning, it revealed that there are some translation problems such as misunderstandings, inaccuracies, and biases. ChatGPT, However, gives a better translation, since its shortcomings are considered acceptable with minor adjustments. As for Arabic-to-English translation, Google Translate has some meaning and accuracy flaws that require major revisions, while ChatGPT provides translation errors that are acceptable with minor revisions. The study concluded that ChatGPT is a more preferable option for machine translation, but the human review is required to ensure an accurate and comprehensive translation of the original text.

Alosaimi and Alawad (2024) investigated the accuracy and clarity of ChatGPT translation of the separable phrasal verbs. The researchers carried out a qualitative analysis of the translation of 100 English phrasal verbs into the Arabic language. The results showed that ChatGPT can succeed in translating phrasal verbs clearly and accurately, yet it needs some training to improvement to generate high-quality human-like outputs and get rid of the minor translation errors that are found. The study also suggested that human editing is required to avoid any translation problems of ChatGPT.

### **3. Methodology**

For the sake of this study, six literary texts, mainly poems, were chosen. Those texts represented different literary eras and thus divergent linguistic styles. The works chosen were three Shakespearean sonnets (18, 130, 138) , Samuel Taylor Coleridge's (The Rime of the Ancient Mariner), Emily Dickinson's ( Hope is the Thing with Feathers) , and Robert Frost's (Stopping by Woods on a Snowy Evening).

These texts were translated by, ChatGpt, Google Translate, and Microsoft Bing. Then, the output of the three tools was compared, contrasted and evaluated. The question of the study, whether ChatGpt is the one that gives the most accurate translations, was handled in a qualitative approach rather than a quantitative through which the results of the observations were analyzed and explained on the basis scrutinizing and describing them.

Moreover, this study is one way, that is the translation of the texts were from English into Arabic. It is mentioning worthy that the black box strategy was used which means 'only the input and outputs of the NMT software are subject to assessment while the internal mechanisms of NMT process are ignored' (Aldelaa & Malkawi, 2024).

### **4. Results and discussion**

After the translation process was performed using the three tools, and after the assessment of their outputs for accuracy, word choice, and structure based on the grammar, syntax, and the semantics of Arabic, the following observations were found:

#### *4.1 .Gender distinction problems.*

Translating gendered language from English to Arabic involves several challenges due to differences in how each language handles gender. Translation challenges including gender differentiation can be intricate and multidimensional, particularly when dealing with language models like as ChatGPT. Below are some examples showing how the tools deal with gendered words.

An example of this problem is the rendition of “ into that silent sea” to

في تلك fi tlika al bahr al samet. It is clear that there is a confusion in the translation between the demonstrative pronoun “that” which refers to the masculine and the feminine singular. The above example is not rendered correctly, it has to be rendered as:

في ذلك البحر الصامت

simply because demonstrative pronouns in Arabic depend on the noun which it modifies.

Since the word البحر is masculine , it has to be preceded by the Arabic masculine demonstrative pronoun ذلك not تلك which is the feminine.

On the other hand both Google Translate and Microsoft Bing, opted for the accurate rendition using the correct demonstrative pronoun ذلك.

It is noticed that ChatGpt does not understand clearly the correct gender of the word breeze and its correct translation into Arabic. In this example as well, ChatGpt renders the phrase “down dropt the breeze as : انخفضت النسيم العليل inkhafadat al- naseem al- aleel .

Also here the noun النسيم which is masculine was accompanied with the verb انخفض to which the feminine marker ت was added. The accurate rendition must be :

و انخفض النسيم العليل

When it comes to Google Translate, it rendered it correctly into : while انخفض النسيم : أسفل إسقاط : Microsoft Bing gave a totally wrong , unnatural , and an understandable translation : أسفل إسقاط : which is merely a word-for- word translation of the original sentence which literally means (down the falling of the breeze).

#### 4.2. Syntactic structure shift.

Syntactic structure shift in translation refers to the changes in sentence structure that occur when translating from one language to another. English and Arabic have significantly different syntactic structures, which can lead to various shifts in translation. While English is SVO language, Arabic uses either SVO or VSO, but VSO is more common in Standard Arabic (Matthew,2013). In fact, when we want to focus on the noun (the subject) ,we should put it first i.e before the verb. Otherwise, we should start the sentence with the verb.

To illustrate this point, “the helmsman steered us through!” is rendered by ChatGpt into Arabic as:

الملاح يقودنا من خلاله. This translation is totally incorrect since there is a clear shift between the subject الملاح and the verb يقودنا. The correct order of the verbal sentence in Arabic is يقودنا الملاح من خلاله since there is no reason justifying putting the subject before the verb.

On the other hand, both Google translate and Microsoft Bing did not have this shift as they rendered it as: وقادنا قائد الدفة where the verb قاد is at the beginning of the sentence . However, they did not give the complete idea of the original sentence due to leaving the word (through) untranslated, thus creating a gap inside the whole picture presented in the two lines in the original text.

#### 4.3. Wrong translation of some dictions

Incorrect translations can significantly affect the meaning and tone of poetry. It’s essential to capture not just the literal meaning but also the communicative meaning and connotations of words. Each language has its own way of expressing these dictions, and an effective translation should preserve the original’s poetic essence and emotional depth. In fact, that what makes the translation of literary texts, mainly poetry, so challenging for these tools.

For example, “The bright-eyed Mariner” is unsuccessfully rendered by ChatGpt as البحار النابض بالحياة instead of البحار ذو العيون اللامعة. In fact, البحار النابض بالحياة could be seen as overly poetic or not quite capturing the literal sense of (bright-eyed). The translation for this expression might be العيون اللامعة which maintains the original meaning more accurately. Both Google translate and Microsoft Bing did not either present a precise rendition as they translated it as: البحار ذو العيون الساطعة. This rendition is not acceptable in Arabic because the word الساطعة usually collocates with the sun not the eyes.

Some translations of the collocations in English and their renditions into Arabic illustrate how some combinations sound more natural and are commonly used in Arabic. For instance, ChatGpt rendered “a good south wind sprung up behind” into Arabic as: نشأت ريح جنوبية. The phrasal verb (sprung up) has a used equivalent in Arabic which is: هبت and which always collocates with the word ريح. Google translate and Microsoft Bing were successful in rendering this English phrase verb to its Arabic precise equivalent.

In addition, the word wind would be better translated as رياح because it has a positive connotation whereas the word ريح has a negative one. Accordingly, “a good south wind sprung up” should be better translated as: هبت رياح جنوبية.

#### 4.4. Polysemy and homonymy problem

According to Longman Dictionary of Contemporary English Polysemy is “a word has two or more different meanings” and homonymy is “a word that is spelled the same and sounds the same as another, but is different in meaning or origin.” To clarify, the noun (lead) and the verb (lead) are homonyms whereas the word (bank) is polysemous as it has two different meanings. This area presented a challenge for machine translation tools as they lack the logic or the understanding of the context according to which they must choose the suitable equivalent.

A clear example of this problem was found when translating the title of the Coleridge’s poem. The word (rime) also (rhyme) is a polysemous word having different, though related, senses. According to Cambridge Dictionary RHYME | translate English to Arabic - Cambridge Dictionary “If a word rhymes with another word, the end part of the words sound the same.”

And according to Merriam Webster dictionary Rhyme Definition & Meaning - Merriam-Webster it is when “one of two or more words thus corresponding in sound”. In this sense it means:

in Arabic which is one of the elements of poetry. قافية

This word also has a different sense which is somehow related: “a composition in verse that rhymes” according to Merriam Webster. In this case it is translated as:

القصيدة المقفاة

in Arabic according to Cambridge dictionary which is describing the poem having rhymed lines.

ChatGpt was able to grasp this difference translating it to: انشودة which means a poem. While Google Translate and Microsoft Bing opted for the imprecise rendition which is قافية.

Another example is when translating Shakespeare’s line in his sonnet 18:

(and every fair from sometimes declines). In this line, the word (fair) is used to convey the same meaning i.e. beauty جمال in both times. ChatGpt was successful in its rendition of this as:

as it conveys the intended meaning of the original. كل ما هو جميل يتلاشى في بعض الأحيان من جماله

On the other hand , both Google translate and ChatGpt failed to convey the precise meaning due to their inability to choose the right meaning of the word fair . Google translate it as :

و كل معرض من المعرض يتراجع في وقت ما .

Here , incorrectly, (fair) was translated into the Arabic word المعرض (exhibition) which is another meaning for the word (fair).

More surprisingly, Microsoft Bing translated it into two different meanings of (fair) which are

(just and objective). Thus , the rendition was a total failure ; عادل (exhibition) and معرض و كل معرض من عادل في وقت ما يتراجع.

Which literary means: and every exhibition from a just person declines.

#### 4.5. *Wrong comparative structure*

One of the problems noticed in the resulting translation is the inability of the tools used to form the Arabic comparative structure. In Arabic the comparative is formed according to the Arabic morphological balance as (af'al) for example:

أكبر (akbar) bigger أكبر (big) (Kabeer) كبير  
أصغر (asghar) أصغر (small) (sagheer) صغير

There are rules to form the comparative (af'al) as not all the adjective use this formula to derive the comparative. According to (Abdul-Razzaq Hussein Ahmed:2023) :

"ألا تكون الصفة منه على وزن (أفعل) الذي مؤنثه (فعلاء)، وهي كل صفة مشبهة تدل على اللون مثل: أحمر حمراء، أو العيب مثل: أعور عوراء، أعرج عرجاء."

That is to say , any adjective that comes as (af'al) whose feminine is (fa'la) indicating colors or defects is irregular and as a result it should be formed by the using the word أكثر which means (more) and the verbal noun (al masdar) that reflects the color .

Thus, for example if we want to say that one apple is redder than another , we should say :

هذه التفاحة أكثر حمرة من تلك .

Red more

The three tools fail to formulate the accurate rendition for the comparative found in the Shakespearean line in sonnet 130:

" coral is far more red than her lips' red"

GhatGpt's translation ignored the previously discussed rule regarding color comparatives and translated it as:

المرجان أحمر بكثير من شفيتها.

In addition , Microsoft Bing and Google Translate also offered an inaccurate translation which is:

المرجان أحمر أكثر من شفيتها and that form the irregular color comparative as a regular one.

However, the accurate translation should be:

المرجان أكثر حمرة من شفيتها

## 5. Conclusion

The translation of literary texts , especially poetry is not an easy task for machine translation tools like Google Translate. For AI – aided platforms, though more developed and more advanced, the task is not any easier. This is due to “ the complexity of human language, which includes nuances and idiomatic expressions unique to each language. **Mohsen (2024)**

As can be noticed machine translation tools have their shortcomings in structure sometimes but their problem is found mainly on the semantic level when it comes to choosing the right equivalent or when dealing with polysemous words and homonyms. On the other hand, the AI – aided tool , ChatGpt, was more successful in dealing with the semantics and in choosing the accurate equivalences but it was less successful in the structural and the morphological level. Thus, it can be concluded that AI – aided tools do not produce accurate and natural translation that can be totally taken as is. The result appeared is that proofreading and revising the translation of these tools by humans is indispensable if we want to come up with reliable and perfect translation.

## References

- [1] Abdul-Razzaq Hussein Ahmed (2023). Comparative Form (af'al) in other Classification Stipulated in the Holy Quran. *مجلة تدبير. القرآن الكريم*. 15 صيغة التفضيل (أفعل) في غير بابها في القرآن الكريم.
- [2] Aghai, M. (2024). ChatGPT vs. Google Translate: Comparative Analysis of Translation Quality Citation. *Translation Studies*, 22 (85): 87—103. Retrieved from <https://journal.translationstudies.ir/ts/article/view/1156/195>
- [3] Alafnan, M. (2024). Large Language Models as Computational Linguistics Tools: A Comparative Analysis of ChatGPT and Google Machine Translations. *Journal of Artificial Intelligence and Technology*, 4. DOI:10.37965/jait.2024.0549.
- [4] Aldelaa, A & Malkawi, M (2024). Investigating Problems Related to the Translation of Idiomatic Expressions in the Arabic Novels Using Neural Machine Translation. *Theory and Practice in Language Studies* Vol. 14, No. 1, pp. 71-78. DOI: <https://doi.org/10.17507/tpls.1401.08>© 2024 ACADEMY PUBLICATION
- [5] Alosaimi A., B., & Alawad A., N. (2024). Evaluation of the Translation of Separable Phrasal Verbs Generated by ChatGPT. *Arab World English Journal*, 19: 282-291. DOI: <https://dx.doi.org/10.24093/awej/ChatGPT>
- [6] Farrell, M. (2023). Preliminary evaluation of ChatGPT as a machine translation engine and as an automatic post-editor of raw machine translation output from other machine translation engines. *Proceedings of the International Conference HiT-IT*, 108-113. DOI:10.26615/issn.2683-0078.2023\_009
- [7] Khoshafah, F. (2023). ChatGPT for Arabic-English Translation: Evaluating the Accuracy. *Research Square*, 1: 1–18. DOI: 10.21203/rs.3.rs-2814154/v2
- [8] Matthew S. Dryer. 2013. Order of Subject, Object and Verb. In: Dryer, Matthew S. & Haspelmath, Martin (eds.) *WALS Online* (v2020.3) [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.7385533>(Available online at <http://wals.info/chapter/81>, Accessed on 2024-09-14.)
- [9] Mohammed Ali Mohsen (2024). Artificial Intelligence in Academic Translation: A Comparative Study of Large Language Models and Google Translate. *Psycholinguistics*, 35 (2), 134–156 <https://doi.org/10.31470/2309-1797-2024-35-2-134-156>
- [10] Valdivieso S., & López-Arroyo, B. (2023). Google Translate vs. ChatGPT: Can non-language professionals trust them for specialized translation? Conference: International

Conference on Human-Informed Translation and Interpreting Technology,  
DOI:10.26615/issn.2683-0078.2023\_008.