



Research Article

© 2023 Adiel et al.
This is an open access article licensed under the Creative Commons
Attribution-NonCommercial 4.0 International License
(<https://creativecommons.org/licenses/by-nc/4.0/>)

Received: 10 March 2023 / Accepted: 16 June 2023 / Published: 5 July 2023

Accuracy and Problems of Machine-Based Translation in Contrast to Human-Based Translation when Rendering Health Awareness Texts Versus Poetry Texts

Mousab Alata Elseddig Adiel¹

Muawia Abdelmagid Elsadig²

Abdelrahman Altigani³

Yasir Abdelgadir Mohamed⁴

Babikir Eltigani Siddig Ahmed⁵

Salah Mohammed Osman Elhassan⁶

¹Assistant Professor, Translation Studies (English & Arabic Languages),
Imam Abdulrahman Bin Faisal University (Iau),
Dammam 31441, Saudi Arabia

²Assistant Professor, Computer Sciences and Information Security,
IMAM Abdulrahman Bin Faisal University (Iau), Dammam 31441, Saudi Arabia

³Department of Computer Science, College of Computer Science and Information Technology,
Imam Abdulrahman Bin Faisal University, Dammam 31441, Saudi Arabia

⁴College of Business Administration, MIS Department,
A'Sharqiyah University, Ibra, Oman

⁵Assistant Professor, Department of English Language,
Buraimi University College, Alburaimi, Sultanate of Oman

⁶Translation Program Coordinator, Lecturer, Department of English Language,
Buraimi University College, Alburaimi, Sultanate of Oman

DOI: <https://doi.org/10.36941/ajis-2023-0108>

Abstract

The aim of this article is to examine the accuracy of machine-based translation in contrast to human-based translation from Arabic into English only and not vice-versa. Through their own experiments, the authors compared the findings to findings of some other researchers in this field. The article deals with machine-based translation between these two distant languages in the context of different types of topics, mainly medical language (Health Awareness Texts) as a scientific translation publicly published texts, and literature language (Ancient Arabic Poetry Stanzas) as a literary translation non-publicly published texts according to the authors' experiments and data analysis. The data of the first type group were collected from various materials used in translation projects at King Fahad Hospital of Imam Abdulrahman Bin Faisal University while the second type data were selected from an article previously published by the first author but not available publicly or translated. Per the methodology, the authors used words errors rate equation (WER) after the translation has been thoroughly investigated by two external professional translators. The results of the study show that machine-based translation has its advantages but still does not totally replace human role in translating from Arabic to English because the outcome still requires deep editing.

Keywords: Machine Translation, Health Awareness, Human-based Translation, Poetry, Accuracy, Semantic, Literary, Social, Culture, Machine Learning

1. Introduction

The definition of machine translation has been discussed by many scholars, such as (Olohan, 2011). Such translation depends on the memory of the machine and its algorithms in the process of recalling the stored segments. Among the advantages enumerated by scholars regarding this subject are the elimination of redundant work, the facilitation of ensuring consistency, and an increase in productivity for freelance translators. In this article, we have searched the machine accuracy of translating various types of texts and compared the machine behaviour in dealing with publicly published texts and non-publicly published texts along with discussing problems encountered by the machine in this course. The following parts of this article include question of the study, significance of the study, limitations of the study, literature review, scenarios from the authors work environment, data sample along with machine used for the experiment, methodology, and findings of the study. The method has been adopted contained reviewing the outcome of translation by two external translators for having a better linguistic analysis, and a mathematical equation was used for measuring the accuracy, in addition to interviewing some computer professionals regarding some remarks about the experiment.

2. Question of the Study

To which extent does the machine-based translation require editing through words substitution, insertion, or deletion of the produced texts considering the far-distance between the types of the texts selected as well as the dissimilarity between the two languages, mainly from Arabic into English?

3. Significance of the Study

This study gains its importance from answering the question above along with considering the trend of using machines which became phenomenal where many persons think it may translate texts perfectly. Accordingly, this study was set to explain that it is very early for the machine to render language in perfection and to correct this misconception, particularly when the linguistic traits or families of two languages are different. The study also focuses on the voids result from use of the machine in translating various texts through linguistic analysis and mathematical measure for the accuracy of the translated text outcome which signifies the importance of the study for software developers, linguists, and alarms those tend to use machine-based translation without deep editing or strict follow-up.

4. Limitations of the Study

This study deals only with translation from Arabic into English. It compares machine-based translation to human-based translation in this context. Thus, it considers only two very different types of texts; health awareness texts as a scientific translation (called here the first type), and ancient Arabic poetry as a literary translation (called here the second type).

5. Literature Review

For the purposes of this study, it is fruitful to review the advantages and disadvantages of using machine-based translation and how these attributes influence translation from Arabic into English. (Austermuhl, 2001; Alcina, 2008) ran tests about the benefits of using machine-based translation. (Dillon et al, 2006) and (Fulford et al, 2005) encouraged the use of machine translation but only after becoming aware of the shortcomings of software and not totally rejecting it. They recommended training on software, although this was found to be time-consuming for the translators involved.

Thus, there were disadvantages to adopting machine translation.

Taking the above-mentioned in consideration, the experiment confirms that machine translation may sometimes boost productivity, but at other times, it adds time for translation from Arabic into English because of the need to fix inconsistencies, which requires additional time for editing. Therefore, it is too early to depend mainly on machines in performing translations between these two spaced out languages, especially for texts not stored in machines or not widely published on internet websites.

(Bowker, 2005) has conducted studies on the advantages mentioned above. (Bowker, 2007) checked the software effect on textuality reporting and found that it affects text cohesion. (Kenny, 1999) claimed that machines could save time in completing translator tasks, but he advised avoiding using them for segments with difficult syntactic structures. This view confirms that machine still cannot make large breakthroughs. (O'Brien's, 2006) research targeted cognitive loads in terms of whether the software is good for repetitive segments but found mixed results for other segments.

One of the important things we must take into consideration are the correlations of translation beyond linguistic effect like the effect of translation on other cultural or social contexts for texts like poetry where machine translation can't render language in such contexts perfectly. The result of translation will be just equivalent words for the source language without considering the effect on the context, and in this case, we can't call it a translation.

(Cadwell et al. 2008) conducted an experiment using focus groups to investigate issues related to adopting or uncritically using machine translation software.

In this article, following the analysis of the above literature, the authors selected to run an experiment regarding whether machine-based translation can make some breakthroughs and may one day make notable advance, particularly on aspects of the morphology and syntactic structures of the Arabic language. It is true that the results of some previous research were contradictory, but there were also some meaningful developments from the current experiment. However, the differences between human power and machine power still require more research to tackle linguistic problems result from translation for perfect treatment to translation outcomes in the language contexts referred to above. Otherwise, machine translation could not yet tackle this issue, or in other words, one can say that some unacceptable errors still occur.

(Vieira et al, 2020) conducted research on perspectives and issues related to management and production of translation software. In their analysis they explained how software can add doubt to services provided and trigger miscommunication. They also discussed how the increasingly common use of such software may affect language and translation projects. Regarding their research, the authors consider that the affection on language may be negative even if it seems positive in some sides, for example, if some software types save the translator time positively, they may affect the produced text quality negatively and this is unpleasant issue with no doubt.

(Brynjolfsson et al, 2014) argued that machine learning is part of a new technological era called the second machine age mentioning that opinions about the possible effect on human labour are different where some believe that as a positive opportunity for interaction between people and the machine while others think expectation related to the machine are exaggerated (e.g., Gordon, 2014). Irrelevant to translators' position from the second era of the machine and its potential negative effects, there have been some striking employment and economic trends observed over the past few decades, including the polarization of jobs (Autor, 2015; Goos et al, 2007) and a reversal of the demand for cognitive labour (Beaudry et al., 2013). Through the experiment of this study, the authors totally agree with these points of views.

A report released by the Translation Automation User Society (TAUS) stressed that creative tasks are a key to translators' sustainability. Translators' future roles include being writers and advisors because of technology impact or effect (Meer, 2017). This opinion may be applicable on other languages, but regarding translation into Arabic, the authors see it is very early to be applicable as the following scenarios confirm.

5.1 Scenarios from the authors' work environment

The subject of machine-based translation versus human-based translation has been abundantly discussed by many researchers and bloggers. Different views have been provided. In this article, the authors investigated these various views in relation to translation between the two languages (from Arabic into English). The previous researchers' different opinions vary according to their respective research interests. Some discuss the issue from the side of the effect on translators' livelihoods, and some discuss the perspective of translated text quality. Some are opponents of machine translation while others are proponents, but the authors of this current study intend to measure to which extent does the machine-based translation require editing through words substitution, insertion, or deletion of the produced texts considering the far-distance between the types of the texts selected as well as the dissimilarity between two languages of different families?

In general, machines began to be used in various fields for the purpose of improving products and saving time, allowing people to complete their work more quickly. This is not a forbidden issue because improvement and promotion via technology is preferable, but not in the case of language since language is like a human being without caring it properly will lose it and affect it negatively making many distortions.

Every writer has a unique style that distinguishes him or her but imagine what would happen if all languages became computerized languages. The authors still remember when two of our non-Arab colleagues translated (وحدة التصنيف) into their native language using a machine. The result in English was (وحدة الترتيب). Later, they called our unit translator to check it. The translator's answer explained that the outcome was (Training Unit) while they were targeting (Ranking Unit). They laughed loudly, and directly were asked, "Was it a machine?!" This small incident shows why the authors believe that machines are still very far from being about to complete translations independently, and this reality is known even among the public.

The authors came across another experiment when checking the translation of some ancient Arabic poetry selecting a variety of stanzas from poems about "love" because it is a common subject. The machine translation was too erroneous in many cases presented at the conclusion of this research. Additionally, the authors also checked health awareness texts. Here, the result was different. The machine translation from Arabic into English was somehow good to the extent that the authors were taken aback, particularly when dealing with medical terms. The authors then traced this difference in results between the machine-based translations of the literary and medical texts. By researching this matter further, the authors found that the medical passages were taken from different websites. This means that if a source text is commonly used and has previously been published on the internet, it will be easier for the machine to identify it and translate. Another result is that the medical texts were previously translated by other parties from all over the Arabic world.

In an interview, a computer-science professional confirmed that this results because many medical students use machines to repeatedly check their own translations from English into Arabic. Another explanation arising from the authors investigation is that medical language is a purely academic language and is widely used on the internet.

Below, the authors show how machine translations are a type of memory translation. It is important to note that the amount of time required for editing translations of such text is low because, as mentioned, the content is already available on the internet; in other words, the language is repeated.

However, the translations of poetry texts and other literary genres were so bad that one cannot call them true "translations". This result appeared because the language used in poems is not only ancient but also unavailable to the machine translation tools via the internet. Accordingly, the authors trust that if someone translates an old Arabic poem and put it on different internet websites with its Arabic source text, the machine may be able to grab it directly and provide it to the user when attempting to translate. Otherwise, translations of poetry, prose, and other literature genres will directly lose their beauty and finesse, or there will be a need to store many translations for the

reader's enjoyment. Additionally, as a matter of taste, this is not a desirable outcome. Imagine if all people thought in the same way, talked the same language, and used the same computerized language, then stereotypes problem will occur. In dealing with the translation of literature genres, one should still account for the importance of human involvement because the translator is requested to convey the spirit of the text's meaning, just as the poet intended, and to carefully manipulate the texts in a way that respects the cultural and sociological dimensions.

Another experiment involved one of our customers wanting to notify—in the context of translating technical terms relevant to quality management studies—us that the Arabic equivalent for the word criteria is (محكات). One of the authors directly understood that the customer had used an online dictionary where (محكات) was -at that time- one of the lexical equivalents provided by the machine. He asked, "What is your suggested translation for it?". One of the authors said it would be (معايير فرعية) or (ضوابط) in Arabic. In this case, the machine translator did not have the correct Arabic word in its dictionary. That author searched the word in English-Arabic dictionaries and did not find the proper equivalent, but by using own translation skills, he looked up the word in a French-Arabic dictionary and found an equivalent that matched the general understanding of Arabic culture.

For the information of the readers, the current research general experiment dealt only with health awareness and poetry texts to check the required editing for the produced texts. There were no technical or other religious types, or any text-types used beyond the narrated two types. Therefore, the authors did not judge machine-based translation for other text types in this article.

In the following part of this study, the authors state the methodology they followed in running this research along with the method used for calculation of the machine linguistic errors as well as the equation used for checking the machine accuracy.

6. Data, Sample Size, and Machine Used

The sample included two groups; one for the texts related to health awareness, and the other included stanzas from Arabic classic poetry (old poetry) as literature genres translation.

The sample of the first group is greater than the sample of the second group because the authors early noted the machine disability to translate such genres. The reason for selecting these two groups of samples because the first group texts are publicly published and available on the internet different websites while the second group texts are not publicly available. Table (1) shows the sample size for both groups.

The data of the first group is internally taken from brochures of KFHU (King Fahd Hospital of the University affiliated to Imam Abdulrahman Bin Faisal University), but the topics are publicly available on the intranets of other institutions with simple development by KFHU staff. The data of the second group is publicly published on Arabic books of poetry and available on internet different websites, but their translation is not publicly published in English language.

Regarding the topics of the first group, it included variety of topics related to health awareness as a part of medical texts. Its topics were about Alzheimer, Psychological disorders, Obsessive-compulsive disorder, Osteoporosis, Empyema, Slipped disc, Rickets, Lupus, etc....

Regarding the topics of the second group, it included stanzas about divine love, asceticism, and flirtation poetry.

Regarding the machine used, the authors depended on (Google Translate) for it is widely commonly used and free.

7. Methodology

For this study purpose, the authors translated the two samples of the two different text types using Google Translate for checking the ability of the machine in rendering two different types of texts and delivered it to two external translators for assessment. The authors provided the two translators with the original English texts and their translation into Arabic, then requested them to send their

remarks about the development of translation knowing that it is a machine-based translation. Their remarks about the words produced were to substitute some words, delete some words, and to insert other words.

According to the words highlighted by the two translators, the author decided to use the Word Error Rate Equation shown below to determine the accuracy of machine-based translation compared to human-based translation.

Word Error Rate might be more applicable if it involves the transcription of paragraphs and sentences of words with meaning (Kenneth Leung,2021).

$$WER = (S+D+I) / N$$

Whereas,

WER stands for Word Error Rate

S: Number of the words that have been substituted to give the intended meaning.

D: Number of the words has been deleted. This represents redundant words or meaningless words.

I: number of insertions or number of the words has been inserted to correctly reflect the intended meaning.

N: number of the words in the reference or the exact number of the words being translated from the original text.

8. Calculations Results

Table (1) shows the equation inputs for both datasets, health awareness texts dataset, and Arabic classic poetry dataset which includes the numbers of words suggested for deletions, substitutions, and insertions per the recommendations of the two external translators.

Table 1: Shows the equation inputs for both datasets

	S: Substituted Words	D: Deleted words	I: Inserted words	N: Words of the original text
Health education texts	13356	21420	20412	252000
Arabic classic poetry	1101	720	803	3000

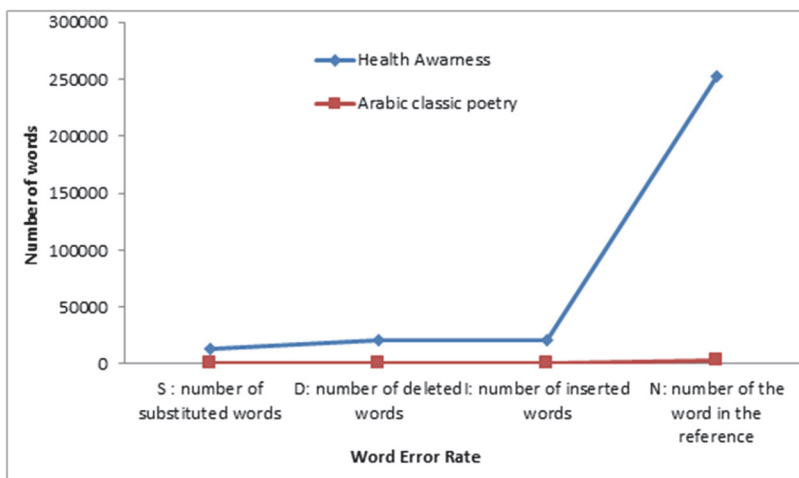


Figure 1. Shows the variance in accuracy between the types of the data according to the machine-based translation

According to the computed numbers of substitutions, deletions and insertions words shown in table 1, the error rates for both datasets are shown below:

WER of Health awareness texts translated from Arabic into English = 21.9%

WER of Arabic classic poetry translated from Arabic into English = 87.47%

These results reflect somehow not bad accuracy when using machine-based translation for the first type of texts which are medical texts where the ratio reached 78.1%. While, for the second type of texts related to Arabic classic poetry, a poor performance has been reported for the machine-based translation as it scored high rate of errors that reached 87.47% which indicates a poor accuracy level.

9. Findings of the Study

- Per the calculations of the data given in the table (1), and as indicated by the graph below it, the results indicate that using of machine-based translation for the first type of health awareness texts require tiny editing which is not the case for the second type types of literature genres texts where the rate of errors is considered higher. In other words, machine-based translation for the first type requires minor changes and major changes for the second type.
- Per the comments of the two external translators, the changes required for the first type are not linguistically crucial (just words and the text will be of normal understood language), but so essential for the second type because of the cultural dimension (changes of words required but still the gap will be big because of linguistic considerations will be clarified in the following below points).
- The first type texts are more circulated publicly and this is the reason behind its better accuracy when translated by a machine. Here the authors also interviewed two machine learning experts to see their opinion about this point and they agreed that it is because of the data publicly published on other website of the internet in English language. In contrary to the first type texts, the second type texts have no translation available publicly on the internet websites, just newly selected by the author for the purpose of this experiment.
- Another reason for machine-based inaccuracy of Arabic classic poetry texts (the second type), and as expected, is because the Arabic classic poetry texts, in all cases, do not have the direct meaning of the words. The target meaning can be understood by considering many semantic factors and different situations influenced by the full understanding of rhetoric of Arabic language and other aspects include but not limited to the below linguistic points:
 - Diacritics, or phonetic signs of the Arabic words because still the machine can not differentiate it at this stage.
 - The hidden pronoun in Arabic language which is not applicable for English language and up to now it is not guessed well by the machine.
 - Polysemous Arabic words.
 - Textual cohesion.
 - Euphemism
 - Arabic words of wishing which have no equivalents in English (Adiel, 2021).

Note: All the above linguistic points are per the limitation of the study checked texts per the sample used. Other texts may include more limitations for the machine in the context of linguistic limitations.

- Accordingly, it is too early to depend mainly on machines in performing translations between these two different languages, especially for texts not stored in machines or not widely published on internet websites.
- The machine is not able to tackle beyond-linguistic contexts like emotional and cultural texts.
- If some software types save the translator time positively but they may affect the produced-

text quality negatively, which is unpleasant issue with no doubt, that means there is still a need to the translator role for editing.

- The assumption says that translators' future roles will only include being writers and advisors regarding cultural issues is not true for translation from Arabic language into English language because the role of the translator still more required. This may be applicable in other languages for limited contexts.

In the following part, the authors present a summary of the study per the experiments encountered, the studies reviewed, and the results inferred in a short conclusion.

10. Conclusion

As per the discussion of the abovementioned experiments, the machine may fail to translate phrases or words. This issue confirms that there is a fundamental risk when using machines for translations up to date of this study.

In translating literary genres, such as ancient Arabic poetry, machine-based translation is a questionable choice for various reasons, including linguistic, cultural, and social dimensions as well as the specific traits of poetry.

Machines can translate health awareness texts to some extent adequately from Arabic into English – with minor postediting needed – for the following reasons:

- **There are often similar texts widely available on the internet.**
- **Medical language used in the field is fixed and does not require any creativity, unlike in literature genres.**
- **It seems that the audience of these texts already broadly use machine translations.**

Many previous studies have focused on speed and production with machines. This is an irrelevant issue when talking about translation from Arabic into English or vice versa, as accuracy and saving the language are more important than speed or production. On the ground, one can see that many translation companies hire cheap translators with little experience to decrease their financial costs, exploiting the translators' need for money. These companies allege that speed is necessary to satisfy their demanding customers. Such practices affect the quality of the translations. Evidence of this is clear to Arabic linguists through the poor texts they review daily from different websites and publications.

One of previous studies hypothesis claims that machine agency will replace human agency and affect translators by taking their jobs; however, this is currently impossible and far from true for Arabic-English translators. If it occurs in the future for any technological developments, linguists and translators are still required as editors. The machine agency will only help with production and speed, not accuracy, according to our experiment.

References

- Adiel, M. (2021). Automatic Translation of Arabic Classic Poetry; Case-Study of Google Translate. *International Journal of Humanities Social Sciences and Education*, 8(8):81-95.
- Austermuhl, F. (2001). *Electronic Tools for Translators*. Routledge. Doi.org/10.4324/978135760353.
- Autor, H. (2015) Why Are There Still so Many Jobs? The History and Future of Workplace Automation. *Journal of Economic Perspectives* 29(3):3-30.
- Beaudry, P., David A., & Benjamin, M. (2013). *The Great Reversal in the Demand for Skill and Cognitive Tasks*. National Bureau of Economic Research Working Paper Series No. 18901. Doi: 10.3386/w18901.
- Bowker, L. (2007). Translation memory and "text". In *Lexicography, terminology, and translation. Text-based studies in honour of Ingrid Meyer*, Lynne Bowker. Ottawa Press of Ottawa University. 175 -187
- Bowker, L. (2005). Productivity vs quality? A pilot study on the impact of translation memory systems. *Localisation Focus*, 4(1):13-20.
- Brynjolfsson, E., & Andrew, M. (2014). *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies*. New York and London: W. W. Norton & Company.

- Cadwell, P., O'Brien, & Carlos, S. (2018). Resistance and Accommodation: Factors for the (Non-) Adoption of Machine Translation among Professional Translators. *Perspectives*, 26(3): 301-321. Doi: 10.1080/0907676X.2017.1337210
- Dillon, S. & Fraser, J. (2006). Translators and TM: An investigation of translators' perceptions of translation memory adoption. *Machine Translation*, 20(2), 67-79.
- Fulford, H. & Joaquin, G. (2005). Translation and technology: A study of UK freelance translators. *JoSTrans – The Journal of Specialised Translation*, 04, http://www.jostrans.org/issue04/art_fulford_zafra.php.
- Goos, M. & Manning, A. (2007). Lousy and Lovely Jobs: The Rising Polarization of Work in Britain." *The Review of Economics and Statistics* 89(1):118-133. Doi: 10.1162/rest.89.1.118.
- Gordon, J. (2014). *The Demise of U.S. Economic Growth: Restatement, Rebuttal, and Reflections*. National Bureau of Economic Research Working Paper Series No. 19895. Doi: 10.3386/w19895.
- Kenny, D. (1999). CAT tools in an academic environment: What are they good for? *Target*, 11(1): 65-82.
- Kenneth, L.(2021). Evaluate OCR Output Quality with Character Error Rate and Word Error Rate. Published in *Towards Data Science*. June,24,2021. Found on <https://kennethleungty.medium.com>.
- Vieira,L. & Alonso, E. (2020).Translating perceptions and managing expectations: an analysis of management and production performance perspective on machine translation.*Perspectives*, 28(2): 163184, Doi: 10.1080/0907676X.2019.1646776
- Meer, J. (2017). The Translation Industry in 2022: A Report from the *TAUS Industry Summit*. *Translation Automation User Society*. Accessed April 18, 2017. <https://www.taus.net/resources/blog/the-story-of-the-translation-industry-in-22>
- O'Brien, S. (2006). Eye tracking and translation memory matches, *Studies in Technology*. *Perspectives*,14(3): 185-205.
- Olohan, M. 2011. Translators and Translation Technology: The Dance of Agency. *Translation Studies*, 4 (3): 342-357. Doi: 10.1080/14781700.2011.589656.