Al-Noor Journal for Digital Media Studies
October, 2024; v.1, n.0, en.2 (32-57)
ISSN: 3007-9853, https://doi.org/10.69513/jnfdms.v.1.i.0.en.2

The effect of using AI and Google Translate on translating BBC media texts into Arabic

Ph.D. Kais A. Kadhim
University Malaysia Terengganu
Kais.kadhim@umt.edu.my
Orcid : 0000-0003-1759-248X

Received: 7/ 9/2024 Accepted: 11/10/2024 Published: 21/ 10/2024

Abstract

Recent post-human translation studies increasingly explore the intricate relationship between artificial intelligence (AI) and translation. This paper examines the complexities of AI, with a focus on machine translation (MT), and its transformative effect on translation in light of technological advancements. Such transformation has led to new roles and skills for translators. Our research primarily scrutinizes the style and messaging quality of English-Arabic media text translations. We aim to illuminate how AI and MT challenge the translation industry, discussing both potential benefits and risks for language professionals, and presenting a theoretical framework for the changing roles of human translators. This study specifically identifies common errors in MT for Arabic-English news texts and vice versa, assessing translation quality, fluency, semantic accuracy, and the need for human intervention in corrections.

Al-Noor Journal for Digital Media Studies
October, 2024; v.1, n.0, en.2 (32-57)
ISSN: 3007-9853, https://doi.org/10.69513/jnfdms.v.1.i.0.en.2

We selected purposeful examples from online newspapers and used a qualitative method for data analysis, identifying errors based on Hsu's (2014) machine translation error classification.

Keywords: Artificial Intelligence (AI), errors, meaning, machine translation, message

Introduction

To begin with, it is no exaggeration to say that the rising penetration of artificial intelligence (AI) into all aspects of life is profoundly altering the functioning of today's over-computerized world. Borrowing from Aldous Huxley, consider how smart homes, web search, online shopping, cyber security, and even self-driving automobiles are reshaping our 'brave new world' in an extra-linguistic reality. AI is now regarded as critical to the digital transition of modern society and has become a priority for the EU. The embryonic and ever-evolving translation sector, with its language agents - translators - appears to be on the path to 'industrialization' and is not immune to modern technological developments (Abdelaal and Alazzawie, 2020). In this context, the connection between AI and translation, and the repercussions of this complex relationship, have become hot topics in today's post-human translation studies (see, for example, Thomas, 2017; Olohan, 2017; Cronin, 2020; Jiménez-Crespo, 2020). The latest, and arguably worrisome, data figures demonstrate the omnipresence of machine translation (MT) as a result of the rising penetration of AI and deep learning (DL) into the modern translation sector. Some studies report that machines are now responsible for 99% of all translations, while Google Translate, the world's most popular translation tool, translates 143 billion words

per day, or 20 words per person, since its launch in 2016 (Liu, 2021; Way, 2021). Moreover, neural machine translation (NMT) engines are trained on massive volumes of data. These systems learn from textual patterns, discover word relationships, and operate by estimating the likelihood of specific words. They are more accurate and human-like, with some researchers believing they have attained human parity, particularly with the introduction of transformer models (Hassan et al., 2018; Laki and Yang, 2022). Although this claim may seem overstated, the quality of NMT input is constantly improving in various fields such as business, industry, and entertainment (Abdelaal and Alazzawie, 2020). These systems are often not controlled or trained by translators, making translators a subset of users. Even from the perspective of the engineers who trained the engines, the output is frequently difficult to estimate and analyze. Google Translate, the most widely used and freely accessible online machine translation engine, switched to an NMT engine in November 2016.

Literature View

Ping (2009) postulates that machine translation (MT) is a discipline of translation employing computer software to translate submitted texts. This system was developed in the 1950s (Hutchins, 1995; Balkan, 1998) and was initially used to produce a raw output, providing a preliminary understanding of the submitted source text (ST) based on statistical analysis. The original purpose of MT systems was to replace human translators, but this proved unsuccessful. The high error rate in translations, resulting from lexical and syntactic ambiguities, led to a decline in enthusiasm for using MT systems (Juan, 1994). This failure led to the development of computer-assisted translation (CAT), which integrates human

translators in refining translations provided by computer-based software.

There are three types of machine translation systems: neural machine translation (NMT), rule-based machine translation (RBMT), and statistical machine translation (SMT). Google Translate (GT) has been transitioning from SMT to NMT since 2016. NMT predicts the likelihood of word sequences using an artificial neural network, generally modeling full sentences in a single integrated model, and requires less memory space (Abdelaal and Alazzawie, 2020). Deep learning and representation learning techniques are employed in NMT models.

GT is a free, accessible MT system (Ali, 2018) with the advantage of accessing a large number of translation databases, primarily comprising translations done by humans. Many scholars assert that GT is the most popular MT system, evidenced by its 200 million users (e.g., Aziz, Sousa, & Specia, 2012; Jia et al., 2012; Karami, 2014; Komeili, Farughi, & Rahimi, 2011). Currently, GT utilizes NMT technology, which, according to Castilho, Gaspari, Moorkens, and Way (2017), offers significantly superior translation output in several language pairs compared to SMT systems. NMT's advantage is its use of artificial intelligence to simultaneously represent all sentences in the target language, rather than dividing them into smaller segments like phrases or words. In this context, Handschuh (2013) conducted a study on German-English translation using four distinct internet services.

Studies on MT systems such as GT, SYSTRAN, Bing, and Babylon have found that MT output is typically erroneous and imprecise. According to Keshavarz (1999), common faults include

erroneous tense usage, errors in verb group usage, incorrect word order, inaccurate preposition use, inappropriate use of active and passive voice, and issues related to article usage. Researchers like DeCamp (2009) and Vilar, Xu, D'Haro, and Ney (2006) have identified omission as the most frequent error in MT outputs. Others, such as Gaspari, Toral, and Naskar (2011), and Valotkaite and Asadullah (2012), have found mistranslations to be the most common errors.

Lotz and Rensburg (2016) conducted a study on two types of texts: those from slide-show presentations and those from online news reports. These texts were sent through MT systems annually for four years (2010 to 2013) to assess improvements in MT output quality. The study revealed that the most common types of errors were omission, mistranslation, non-translation, and grammatical mistakes. Additionally, Lotz and Rensburg discovered that while there seemed to be an increase in translation quality up until 2012, this improvement plateaued, with some 2013 outputs exhibiting more errors than those from the previous year (Ali, 2018).

Data

The information was purposefully gathered from one main source, the BBC, because it is considered a professional news website, and it also has an Arabic version. MT was chosen since it is freely available and frequently used (Drugan, 2013). Furthermore, AI (ChatGPT) is used because it represents the latest in artificial intelligence technology, which may provide better translation. The researchers employed a mixed-methods approach in this study, analyzing the source texts qualitatively. The researcher, who is also a translator, conducted a descriptive analysis as a form of qualitative inquiry. The study's dataset consists of five selections

Al-Noor Journal for Digital Media Studies
October, 2024; v.1, n.0, en.2 (32-57)
ISSN: 3007-9853, https://doi.org/10.69513/jnfdms.v.1.i.0.en.2

from various news texts. The use of news is justified as it is readable by laypeople and accessible to both educated and uneducated individuals.

Hsu's (2014) model

The qualitative analysis of the gathered data utilized Hsu's (2014) model, which provides an in-depth typology of errors, as depicted in Figure 1 below. Hsu (2014) categorizes errors into orthographic, morphological, lexical, semantic, and syntactic types, offering a comprehensive framework for their classification.

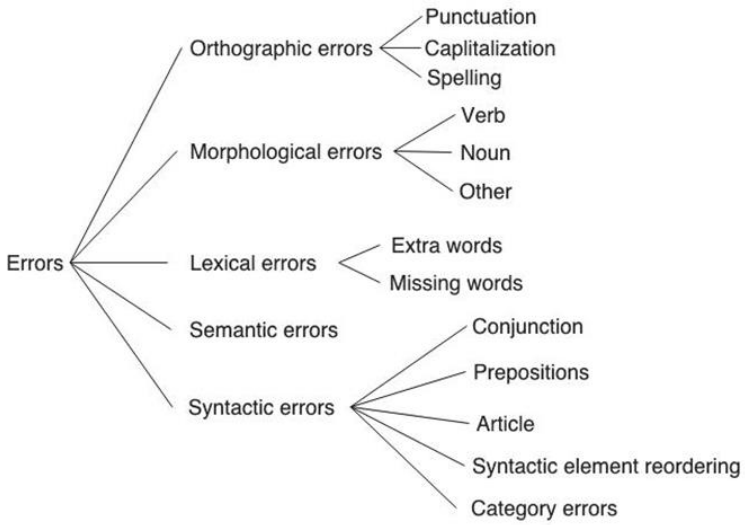


Figure 1: Classification of MT Errors (adopted from Hsu, 2014)

According to her classification, errors are categorized into orthographic, morphological, lexical, semantic, and syntactic errors. Furthermore, each of these main categories is further subdivided. For example, orthographic errors are subdivided into punctuation, capitalization, and spelling. Morphological errors are divided into verbs, nouns, and others. Lexical errors are split into

Al-Noor Journal for Digital Media Studies
October, 2024; v.1, n.0, en.2 (32-57)
ISSN: 3007-9853, https://doi.org/10.69513/jnfdms.v.1.i.0.en.2

extra words and missing words. Finally, syntactic errors are subdivided into conjunctions, prepositions, articles, syntactic elements, and category errors.

Data Analysis

Extract 1:

SpaceX has been accused of unlawfully firing eight workers who were critical of its multi-billionaire chief executive Elon Musk. (BBC : https://www.bbc.com/news/business-67878940)

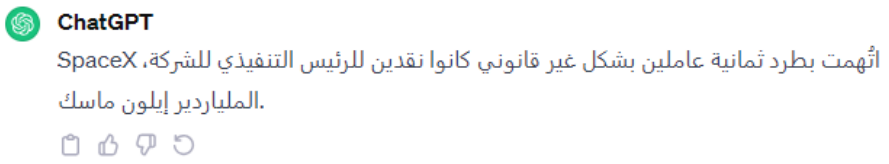
Arabic version: تواجه شركة استكشاف الفضاء، سبيس أكس، تهما بتسريح 8 موظفين، بطريقة غير قانونية، لأنهم انتقدوا رئيسها التنفيذي، الملياردير إيلون ماسك.

https://www.bbc.com/arabic/articles/c72ykje8nkxo.

Machine Translation (MT):

اتُهمت شركة **SpaceX** بطرد ثمانية موظفين بشكل غير قانوني بسبب انتقادهم لرئيسها التنفيذي الملياردير إيلون ماسك**.**

AI (ChatGPT):



The Arabic version, translated from the BBC source text by MT and AI, displays various translation issues, notably in orthographic and semantic errors. Reviewing the Arabic translation, it is clear that at the sentence's start, the translator uses a comma (،) after the Arabic phrase '**تواجه شركة استكشاف الفضاء**' to denote the company name 'SpaceX.' Additionally, another comma follows 'SpaceX' in

the human translation. However, the MT and AI output omits the comma associated with 'SpaceX,' leading to incorrect punctuation between sentences and obscuring the intended meaning in both translations. Moreover, the Arabic translation for 'firing' in the source text is inaccurately rendered as '**طرد**' in both MT and AI. Semantically, there is a difference between '**تسريح**' (laying off) and '**طرد**' (dismissal) in Arabic. '**تسريح**' suggests a necessary action due to market shifts or economic challenges, while '**طـرد**' indicates a discretionary action, where employers may reconsider, especially if the employee shows signs of improvement.

This semantic discrepancy might create ambiguity in the original text, leading to different meanings compared to the original message in the BBC source text. Additionally, 'criticized' in the BBC source text is incorrectly translated as '**نـقـديـن**' in the AI, a completely incorrect spelling that results in a confusing and meaningless interpretation. Such orthographic errors significantly undermine the translation's quality. Furthermore, the sentence structure in the AI translation is incorrect, with a lack of agreement between verbs and nouns and improper use of syntactic elements, as seen when a verb is used without referring to the subject 'SpaceX,' distorting the overall meaning.

Another morphological error in the Arabic translation involves the AI sentence starting with '**اتـهمت**' without a clear subject, suggesting the omission of 'SpaceX,' leading to an entirely incorrect meaning. Additionally, a syntactic error occurs with the unnecessary use of the article 'ا' attached to '**لرئيسها**,' which could affect the grammatical structure of the sentence.

Al-Noor Journal for Digital Media Studies
October, 2024; v.1, n.0, en.2 (32-57)
ISSN: 3007-9853, https://doi.org/10.69513/jnfdms.v.1.i.0.en.2

Extract 2:

ST: A complaint by a US labour agency says the employees sent an open letter to the firm's executives in 2022, detailing workplace concerns.

(BBC : https://www.bbc.com/news/business-67878940**)**

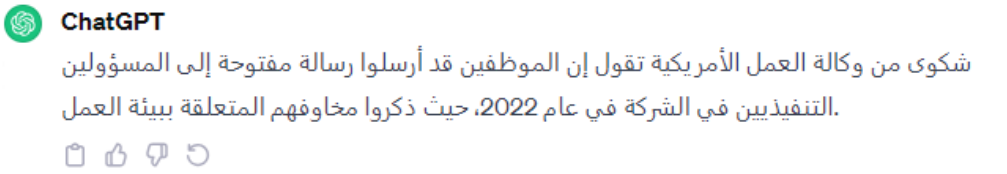
TT:

وجاء في الشكوى، التي تقدمت بها وكالة تشغيل أمريكية، أن الموظفين بعثوا رسالة مفتوحة إلى المديرين في الشركة في 2022 يذكرون فيها تفاصيل ظروف العمل الصعبة

https://www.bbc.com/arabic/articles/c72ykje8nkxo.

Machine Translation (MT):

تقول شكوى مقدمة من وكالة عمل أمريكية إن الموظفين أرسلوا خطابًا مفتوحًا إلى المديرين التنفيذيين للشركة في عام 2022، يتضمن تفاصيل المخاوف المتعلقة بمكان العمل.



The semantic error pertains to the verbal sentence structure in which both MT and AL start with "**شكوى**," leading to a syntactic error related to the use of the definite article '**ال**' with the noun '**شكوى**.' This article implies that '**شكوى**' has been mentioned before in the sentence. However, in the source text (ST), 'complaint' is only mentioned for the first time. A similar issue arises in the AI translation, where '**شـــكوى**' begins at the sentence's beginning as indefinite, creating a semantic gap in the overall meaning of the

sentence. Hence, there is no equivalence between the ST message and the AI Arabic translation.

Furthermore, there is a semantic error in the way both MT and AI translate the noun phrase 'labour agency' into Arabic. The noun phrase 'labour agency' in the source text (ST) implies an agency related to strategic policies for job placement. However, the translation '**وكالة عمل**' in Arabic refers to a type of activity within the workplace. This direct translation of the lexical word 'labour' from the ST into Arabic has resulted in a lexical gap or semantic error, deviating from the intended meaning of the context.

Moreover, orthographic errors are evident in both MT and AI translations. Punctuation errors, specifically the use of the comma (،), are present in both Arabic versions of MT and AI. The punctuation error is related to the use of the comma (،) in both Arabic versions of MT and AI. It is observed that the comma is omitted from "**وكالة عمل أمريكية**" in both MT and AI. The function of the comma in Arabic is to indicate that the sentence is short and to signal that the speech continues after.

Furthermore, there is a slight semantic gap and error between the Arabic lexical words "**ارســـلو**" and "**بعثوا**." The source text (ST) refers to the lexical word '**بعثوا**' because the intended meaning in the ST context pertains to the action that needs to be implemented along with the message. Therefore, the ST does not only refer to sending a message to remind the firm's executive team about actions in the workplace, but it also emphasizes the need for implementation. In contrast, in both the MT and AI versions, the translated text uses '**ارســـلو**,' which in Arabic only indicates the delivery of a message without a direct indication of the action to be taken in the workplace.

Al-Noor Journal for Digital Media Studies
October, 2024; v.1, n.0, en.2 (32-57)
ISSN: 3007-9853, https://doi.org/10.69513/jnfdms.v.1.i.0.en.2

Another semantic error is related to the use of '**رسالة و خطاب**' in the ST, where the ST implies sending a specific message containing detailed information to the firm's executive team regarding the workplace. In contrast, in both the MT and AI versions, the term '**خطاب**' is used, indicating that the sent message is not specific but rather contains general information for the firm's executives.

Finally, there are syntactic errors related to the syntactic element ordering in the Arabic version of AI, where the phrase " **المتعلقة ببيئة عمل**" is syntactically arranged incorrectly due to the missing "**ب**" before the word "**بيئة**." Furthermore, the article '**ال**' is missing from the lexical word '**العمل**,' resulting in a wrong meaning for the entire structure.

Extract 3:

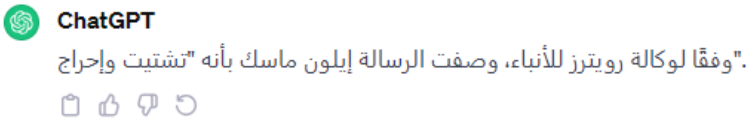
ST: The letter called him a "distraction and embarrassment", according to Reuters news agency. (BBC : https://www.bbc.com/news/business-67878940)

TT:

"وحسب وكالة رويترز، فإن المديرين وصفوا الرسالة بأنها "إلهاء وخزي."

Machine Translation (MT):

ووصفته الرسالة بأنه "مصدر إلهاء وإحراج"، بحسب وكالة رويترز للأنباء.



We notice that both MT and AI exhibit semantic confusion in translating the intended message of the ST phrase "distraction and embarrassment." Regarding the content of MT, we observe that 'embarrassment' is translated as '**احراج**,' which is unsuitable in this

context. The correct translation should be '**خزي**,' a more appropriate and direct translation for the word 'embarrassment' in this specific context. MT seems to use '**احـراج**" as a direct translation without considering the intended meaning and its relevance to the context.

A similar issue occurs in the AI translation, where 'distraction' is also translated into '**تشـــتيت**' using a direct method. Furthermore, there is a syntactic error when translating the ST into Arabic through AI. The second phrase in the Arabic version suddenly starts with the conjunction '**و**.'

The function of a comma in Arabic is to connect ideas between sentences, and because the first sentence in Arabic is not semantically connected, the presence of the conjunction '**و**' is not applied correctly. These semantic and syntactic errors create ambiguous meanings. Another semantic error is the omission of '**المديرين**' from both MT and AI, leading to a semantic gap in the translation of the ST into Arabic. The MT version begins with "**ووصفته الرسالة**," and it is unclear who the letter is describing. The subject of the sentence seems to be missing, indicating a syntactic error in this sentence.

Extract 4:

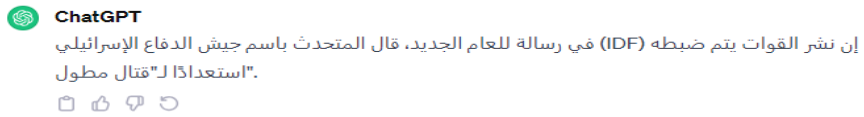
ST: In a new year's message, the Israel Defense Forces (IDF) spokesman said troop deployments were being adjusted to prepare for "prolonged fighting".
https://www.bbc.com/news/world-middle-east-67855117

وفي رسالة بمناسبة العام الجديد، قال المتحدث باسم الجيش الإسرائيلي، دانييل هاغاري، إنه يجري تعديل انتشار القوات للاستعداد "لقتال طويل الأمد."

Al-Noor Journal for Digital Media Studies
October, 2024; v.1, n.0, en.2 (32-57)
ISSN: 3007-9853, https://doi.org/10.69513/jnfdms.v.1.i.0.en.2

Machine Translation (MT):

وفي رسالة بمناسبة العام الجديد، قال المتحدث باسم قوات الدفاع الإسرائيلية إنه يتم تعديل انتشار القوات استعدادًا "لقتال طويل الأمد."



The translation by MT can be considered somewhat acceptable since the general message is clear and aligns with the intended meaning in the ST. However, it is worth noting the way the noun phrase 'the Israel Defence Forces' is translated in MT using a direct approach. In MT, we observed that the word 'force' in the ST is translated as **'قوات'** (troops). In contrast, the content generated by AI is not acceptable due to a semantic error in the translation of lexical words into Arabic. Furthermore, the AI translation contains numerous missing words and an overall ungrammatical sentence structure. For instance, one of the morphological errors relates to sentence arrangement; neither verbal nor nominal sentences are structured correctly, resulting in a clear error in organizing the verb, noun, and object within the entire sentence**.**

Additionally, there is an orthographic error related to punctuation, such as the use and placement of the comma in the Arabic version. Consequently, the sentence arrangement in AI is entirely unacceptable due to semantic and syntactic errors that persist throughout the sentence**.**

Extract 5:

ST: . Daniel Hagari said some troops - especially reservists - would be withdrawn to allow them to regroup.

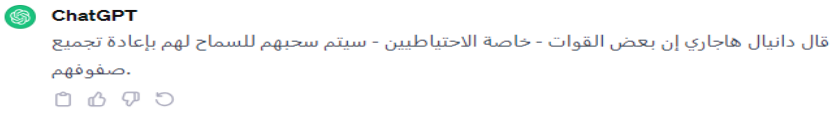
Al-Noor Journal for Digital Media Studies
October, 2024; v.1, n.0, en.2 (32-57)
ISSN: 3007-9853, https://doi.org/10.69513/jnfdms.v.1.i.0.en.2

TT:

وأضاف أن بعض القوات، وخاصة جنود الاحتياط، ستُسحب للسماح لها بإعادة تجميع صفوفها.

Machine Translation (MT):

وقال دانييل هاغاري إن بعض القوات، وخاصة جنود الاحتياط، سيتم سحبها للسماح لهم بإعادة تجميع صفوفهم.



The translation of the ST through both MT and AI is somewhat acceptable, but there are a few semantic and syntactic errors affecting the quality of the translation output. In the MT output, the overall message of the ST is translated almost equally to the TT using a direct translation. However, the semantic issue arises from the use of '**سماح**,' as regrouping the army does not necessarily require permission; it is a matter of following rules and orders to allow the army to regroup. A similar case occurs in the AI Arabic version where the term '**سماح**' is also used in the same context.

Furthermore, in the AI translation, '**خاصة الاحتياطيين**' is used to translate 'especially reservists' into Arabic. It appears that AI employed a direct translation to establish a form of semantic equivalence between the BBC ST and AI, but the word choice does not fully capture the intended meaning. It would be better translated as '**جنود الاحتياط**' to convey a similar meaning as indicated in the ST.

Extract 6:

ST : "These adaptations are intended to ensure the planning and preparation for continuing the war in 2024," he said.

وأوضح هاغاري أن "هذه التعديلات تهدف إلى ضمان التخطيط والإعداد لمواصلة الحرب في العام 2024."

Al-Noor Journal for Digital Media Studies
October, 2024; v.1, n.0, en.2 (32-57)
ISSN: 3007-9853, https://doi.org/10.69513/jnfdms.v.1.i.0.en.2

Machine Translation (MT):

وأضاف أن "هذه التعديلات تهدف إلى ضمان التخطيط والإعداد لمواصلة الحرب في عام 2024."

There is a semantic error in the translation of 'adaptations' from the ST into Arabic through AI. We noticed that the lexical noun 'adaptations' is translated as '**التعديلات**' in both the TT and the MT. In these translations, the translator used a functional method to convey a message similar to that in the BBC source text. However, in the AI translation, the same lexical word is rendered as '**التكيفات**'. This type of translation results in a semantic error as it does not align with the context of the sentence in the ST.

AI employed a direct translation procedure to translate the lexical word 'adaptations' into Arabic. Another syntactic error is related to the use of the article '**ال**', which should be attached to the noun '**العام**'. We noticed that this article is missing from both translations, both MT and AI. In the Arabic original translation, the article has been added to '**العام**' as it refers to a specific year, which is 2024.

Furthermore, the lexical verb 'said' in the ST is translated with a different meaning that does not convey the intended meaning of the ST. In MT, the verb 'said' is translated as '**وأضـــاف**', introducing a clear difference between 'said' and 'add' in terms of meaning. This semantic error creates another issue in conveying the intended meaning of the ST from BBC. As for AI, the verb 'said' is translated directly, but without consideration of the grammatical case.

Al-Noor Journal for Digital Media Studies
October, 2024; v.1, n.0, en.2 (32-57)
ISSN: 3007-9853, https://doi.org/10.69513/jnfdms.v.1.i.0.en.2

Extract 7:

Saudi Arabia's ambassador to the UK says it is interested in normalising relations with Israel after the war in Gaza, but that any deal must lead to the creation of a Palestinian state.
https://www.bbc.com/news/world-middle-east-67922238

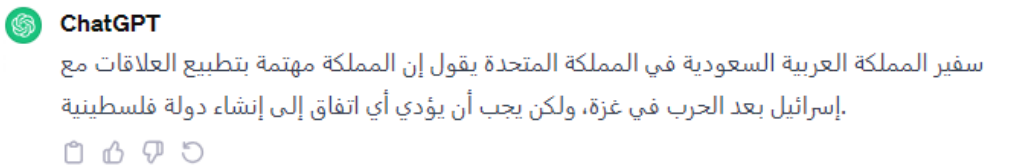
TT:

قال سفير السعودية في بريطانيا، الأمير خالد بن بندر، إن بلاده مهتمة بتطبيع العلاقات مع إسرائيل بعد حرب غزة، ولكن أي اتفاق للتطبيع "لابد أن يؤدي إلى إنشاء دولة فلسطينية."

https://www.bbc.com/arabic/articles/crg8pz45vwko

Machine Translation (MT):

قال سفير المملكة العربية السعودية لدى المملكة المتحدة إنه مهتم بتطبيع العلاقات مع إسرائيل بعد الحرب في غزة، لكن أي اتفاق يجب أن يؤدي إلى إنشاء دولة فلسطينية.



The output for both translations, MT and AI, appears to be generally acceptable since they effectively convey the meaning or message of the ST using a direct translation procedure. There are no significant issues with semantic adequacy in both translations. However, there is a semantic error related to the orthographic aspect.

We noticed that the punctuation in the TT is correctly distributed, indicating the continuation of ideas in the TT message. In contrast, in both MT and AI, the punctuation marks, specifically the comma,

are omitted from both translations. This omission affects the flow of the text and should be corrected for proper punctuation.

A syntactic error occurred in the translation of the BBC source text into Arabic through MT. The error is related to the deletion of the conjunction '**و**' from the lexical word '**لكن**.' The function of '**و**' is to indicate the degree to which ideas are connected in the text, and deleting it creates a syntactic error as there is no indication of the connection of ideas, as shown in the ST. However, it seems that AI managed to maintain this conjunction '**و**' when translating the ST text message into Arabic.

There is also a semantic ambiguity in the use of '**ان المملكة**...' in the AI translation. The phrase '**ان المملكـة**...' could refer to either the Kingdom of Saudi Arabia or the United Kingdom in the Arabic version, as the lexical word can be applicable to both countries. However, the message in the BBC text is intended for the Kingdom of Saudi Arabia. To avoid ambiguity, the translation should specify '**المملكة العربية السعودية**' to clearly indicate the intended country.

Extract 8:

ST: Prince Khalid bin Bandar told the BBC a pact was "close" when the kingdom paused US-brokered talks after Hamas's deadly attacks on Israel on 7 October.

TT:

وكشـف الأمير خالد بن بندر، في تصـريح لبي بي سـي، أن الاتفاق كان "وشـيكا"، عندما علقت السعودية المحادثات بوساطة أمريكية، إثر الحرب الإسرائيلية على غزة بعد هجمات حركة المقاومة الإسـلامية حماس، على مواقع إسـرائيلية، يوم 7 أكتوبر تشرين الأول.

Al-Noor Journal for Digital Media Studies
October, 2024; v.1, n.0, en.2 (32-57)
ISSN: 3007-9853, https://doi.org/10.69513/jnfdms.v.1.i.0.en.2

Machine Translation (MT):
قال الأمير خالد بن بندر لبي بي ســـي إن الاتفاق كان قريبا ‘قريبا’; عندما أوقفت المملكة المحادثات التي جرت بوســـاطة أمريكية بعد هجمات حماس القاتلة على إسرائيل في 7 أكتوبر.

AI:
وقال الأمير خالد بن بندر لبي بي ســـي إن الاتفاق كان "قريبا" عندما أوقفت المملكة المحادثات التي جرت بوساطة أمريكية بعد هجمات حماس القاتلة على إسرائيل في 7 أكتوبر.

We observed that both MT and AI successfully translated the messages from the BBC source text (ST). However, there is a deletion in the translation through MT, specifically the omission of 'told the BBC.' Furthermore, numerous punctuation errors related to the use of commas can be found in the translated text. All the commas in the ST have been removed in the MT version, resulting in Arabic grammar that is less reader-friendly and hinders the understanding of the intended function of punctuation marks as indicated in the ST.

Similar issues with errors related to the use of punctuation marks can be found in the AI version, where all commas have been omitted.

Additionally, there is a semantic error related to the translation of the lexical word 'told' in the ST into Arabic through AI. AI translated 'told' as '**كشف**,' which introduces a difference in meaning between '**قال**' and '**كشــف**' in Arabic. '**كشــف**' is typically used when revealing secret or hidden information to the public, unlike '**قال**,' which can be used in various conversational contexts without necessarily implying a secret situation. This semantic error has created a significant gap in the translated message into Arabic, as

Al-Noor Journal for Digital Media Studies
October, 2024; v.1, n.0, en.2 (32-57)
ISSN: 3007-9853, https://doi.org/10.69513/jnfdms.v.1.i.0.en.2

there is no direct equivalence or relationship between "**كشـف**" and "**قال**" in Arabic.

Another issue is the way the Arabic version starts in the MT translation, which leads to a syntactic error. For example, the phrase "**عنـدمـا أوقفـت المملكـة المحـادثـات التي جرت**" is incomplete and unrelated to the second sentence that starts with "**قريباً**...". There is no semantic connection between both phrases, resulting in syntactic and semantic errors in the overall meaning of the translation.

Furthermore, there is a semantic error in the translation of 'close' into Arabic by both MT and AI. There is a distinction between '**وشيكا**' and '**قريبا**' in Arabic. '**وشيكا**' refers to the future, whereas '**قريبا**' can be used for both the past and the future. Therefore, in the context of the BBC ST, which implies a future event, it is more appropriate to use '**وشيكا**' in the Arabic version.

Extract 9:

ST: Saudi Arabia still believed in establishing ties with Israel despite the "deplorable" casualty figures in Gaza, he said, but it would not "come at the cost of the Palestinian people".

https://www.bbc.com/news/world-middle-east-67922238

TT:

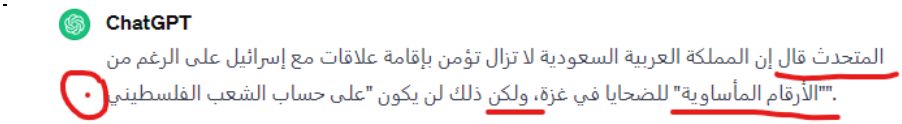
وأضاف أن السعودية لا تزال تؤمن بإقامة علاقات مع إسرائيل، على الرغم من الخسائر الفادحة في الأرواح في غزة، ولكنه قال إن ذلك لا ينبغي أن يكون "على حساب الشعب الفلسطيني."

https://www.bbc.com/arabic/articles/crg8pz45vwko

Machine Translation (MT):

وأضـاف أن السـعودية لا تزال تؤمن بإقامة علاقات مع إسـرائيل رغم أرقام الضـحايا "المؤسفة" في غزة، لكن ذلك "لن يأتي على حساب الشعب الفلسطيني."

AI:



Al-Noor Journal for Digital Media Studies
October, 2024; v.1, n.0, en.2 (32-57)
ISSN: 3007-9853, https://doi.org/10.69513/jnfdms.v.1.i.0.en.2

There are several types of semantic errors, including orthographic, lexical, and syntactic errors, when translating the BBC ST into Arabic. Both MT and AI exhibit similar errors, such as the incorrect use of punctuation marks, specifically the 'comma,' after the lexical noun 'اسرائيل' in both Arabic versions.

Furthermore, there is a syntactic error in the MT Arabic version, where the conjunction 'و' is omitted from the lexical conjunction 'لـكـن'. This omission affects the proper translation of syntactic elements in MT and AI, as the function of the conjunction in the Arabic sentence is to create continuity with the preceding information.

In addition, the function of punctuation, such as the comma '،' is to separate short sentences and create continuity among the sentences.

Another error is related to the use of "على الرغم من" in the MT Arabic version. In this case, the word "غير" is used incorrectly to convey a similar meaning to the original phrase "على الرغم من." Such a semantic error cannot be resolved by translating the cohesive marker 'despite' into the word "غير" alone, and it impacts the quality of meaning.

Another syntactic error is related to the use of "لـن" in the MT Arabic version. This adverb is not suitable in this context, as its function is to negate the future. It would be better to use 'ينبغي' in the Arabic version instead of 'لـن' to convey a stronger meaning aligned with the intended message of the ST.

Furthermore, the way the Arabic translation by AI begins creates a semantic error and ambiguous meaning related to the use of 'المتحدث قال'. This is a syntactic error that occurred when translating the message from the BBC text into Arabic. The semantic error is related to lexical errors in the Arabic version, where 'المتحدث قال'

does not exist in the source text (ST) and has been added to the beginning of the sentence.

Additionally, there is a semantic error in the translation of 'deplorable' into Arabic through AI translation. AI uses '**أرقام مؤسفة**' to translate the lexical item 'deplorable' into Arabic, which creates a semantic error and affects the quality of the message in the ST.

Extract 10:

ST: The ambassador also warned that there was a "failure of humanity"

over Gaza, with the international community not doing enough to end the fighting.

https://www.bbc.com/news/world-middle-east-67922238

TT:

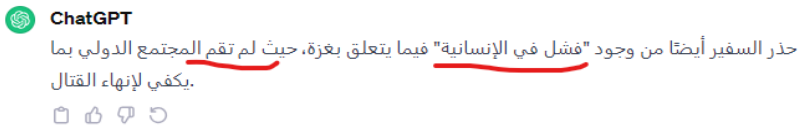
ونبه السفير السعودي أيضا إلى "خذلان الإنسانية" في غزة، لأن المجتمع الدولي لم يفعل ما فيه الكفاية لإنهاء القتال

https://www.bbc.com/arabic/articles/crg8pz45vwko

Machine Translation (MT):

وحذر السفير أيضا من أن هناك "فشلا إنسانيا" بشأن غزة، مع عدم قيام المجتمع الدولي بما يكفي لإنهاء القتال.

AI:



We have observed that both GT and AI Arabic versions exhibit similar semantic and syntactic errors. In the case of GT, the semantic error is related to the translation of "failure of humanity" into Arabic. The choice of "**فشلا**" for "failure of humanity" in GT is not suitable, as it represents a direct translation procedure that may

not be appropriate for standard language, especially in a media context. It would be more suitable to use phrases like ' **خيبة أمل في الإنسانية**" or '**فشل في الإنسانية**,' which better convey the meaning of "failure of humanity."

A similar semantic error is present in the AI version, where the Arabic translation is '**فشل في الإنسانية**,' which is also a direct translation and not a suitable choice for the intended meaning of the text.

Furthermore, there is a syntactic error related to the use of '**تقم**' in the Arabic version of AI. The correct word should be '**يقم**' as the noun phrase 'the international community' is masculine, while '**تقم**' is used only when referring to a feminine subject.

In conclusion, both GT and AI versions exhibit similar semantic and syntactic errors in the translation of "failure of humanity" into Arabic. These errors highlight the challenges in achieving high-quality translations, especially when dealing with nuanced or context-specific phrases.

Conclusion

The purpose of this research is to investigate the translation of news texts from English into Arabic using AI and MT. The study has revealed that both translations fall short of achieving high-quality translations of BBC news from English into Arabic, as various semantic and syntactic errors have occurred during the translation process. The research aims to identify common errors in the raw output of GT and AI, assessing the quality of translation errors, semantic adequacy, and the fluency of the translated text. To accomplish this, 10 articles from BBC news texts were randomly selected, exposing various types of errors, including semantic, morphological, lexical, and syntactic errors in the translated output.

Regarding the quality of translation errors, most were identified as major errors in both MT and AI, although there were instances where the errors in AI or even MT were minor. In terms of fluency, it was found that the translations of MT and AI exhibited some fluency but with errors related to the semantic aspect. In conclusion, this research underscores the need for further improvements in GT and AI to achieve higher-quality translation output.

References

1. Abselaal, N. M., & Alazzawi, A. (2020). Machine translation: The case of Arabic-English translation of news texts. *Theory and Practice in Language Studies, 10* (4), 408–418. https://doi.org/10.17507/tpls.1004.12

2. Ali, A. M. (2020). Quality and machine translation: An evaluation of online machine translation of English into Arabic texts. *Open Journal of Modern Linguistics, 10*, 524–548. https://doi.org/10.4236/ojml.2020.104036

3. Aiken, M., & Balan, S. (2001). An analysis of Google Translate accuracy. *Translation Journal, 6*(12). Retrieved from http://www.translationdirectory.com/articles/article2320.php

4. Arnold, D. J., Balkan, L., Humphreys, R. L., Meijer, S., & Sadler, L. (1994). *Machine translation: An introductory guide*. Cambridge, MA/Oxford: Blackwell.

5. Aziz, W., Sousa, S., & Specia, L. (2012). PET: A tool for post-editing and assessing machine translation. In *Proceedings of the*

International Conference on Language Resources and Evaluation (pp. 3982–3987). Istanbul, Turkey. Retrieved from http://www.mt-archive.info/Aslib-2012-Aziz.pdf

6. Balkan, L. (1998). Machine-aided translation. In M. Baker (Ed.), *Routledge Encyclopedia of Translation Studies* (pp. 134–148). New York: Routledge.

7. Castilho, S., Gaspari, F., Moorkens, J., & Way, A. (2017). Integrating machine translation into MOOCs. In *Proceedings of the EDULEARN17 Conference* (pp. 1–10). Barcelona, Spain.

8. Cronin, M. (2020). Translation and posthumanism. In N. Pokorn & K. Koskinen (Eds.), *Routledge Handbook of Translation and Ethics* (pp. 279–293). London: Routledge.

9. DeCamp, J. (2009). What is missing in user-centric MT? In *Proceedings of the MT Summit*. Ottawa, Canada.

10. Gaspari, F., Toral, A., & Naskar, S. (2011). User-focused task-oriented MT evaluation for wikis: A case study. In *Proceedings of the Third Joint EM+/CNGL Workshop: Bringing MT to the User: Research Meets Translators*.

11. Hutchins, J. (2007). Machine translation: A concise history. In *Computer Aided Translation: Theory and Practice*, *13*(29-70), 11.

12. Karami, O. (2014). The brief view on Google Translate machine. Retrieved from http://ceur-ws.org/Vol-867/Paper17.pdf

13. Keshavarz, M. H. (1999). *Contrastive analysis and error analysis*. Tehran: Rahnama Press. Retrieved from http://bookyar.net/?paperno=142803

14. Komeili, Z., Jalil, F. H., & Ali, R. (2011). An investigation of the translation problems incurred by English-to-Persian machine translation: Padideh, Pars, and Google software. *Procedia - Social and Behavioral Sciences, 28*, 1079–1082. https://doi.org/10.1016/j.sbspro.2011.11.202

15. Laki, L., & Yang, Z. G. (2022). Neural machine translation for Hungarian. *Acta Linguistica Academica, 69*(4), 501–520. https://doi.org/10.1556/2062.2022

16. Lotz, S., & van Rensburg, A. (2014). Translation technology explored: Has a three-year maturation period done Google Translate any good? *Stellenbosch Papers in Linguistics Plus, 43*(1), 235–259. https://doi.org/10.5842/43-0-205

17. Liu, H. (2021). Foreword. In E. Angelone, M. Ehrensberger-Dow, & G. Massey (Eds.), *The Bloomsbury Companion to Language Industry Studies*
(pp. viii–xii). London: Bloomsbury Academic.

18. Olohan, M. (2017). Technology, translation, and society. *Target, 29*(2), 264–283. https://doi.org/10.1075/target.29.2.04olo

19. Ping, K. (2009). Machine translation. In M. Baker & G. Saldanha (Eds.), *Encyclopedia of Translation Studies* (2nd ed.). New York: Routledge.

20. Valotkaite, J., & Asadullah, M. (2012). Error detection for post-editing rule-based machine translation. In *Proceedings of the Association for Machine Translation in the Americas, Workshop on Post-editing Technology and Practice*.

21. Vilar, D., Xu, J., D'Haro, L., & Ney, H. (2006). Error analysis of statistical machine translation output. In *Proceedings of the LREC* (pp. 697–702).

Biodata:

Dr. Kais A. Kadhim is a Associate Professor at Universiti Malaysia Terengganu, Malaysia. His M.A and Phd In Translation from University of Science Malaysia (USM). His Main Research Interest is Cultural Linguistics, Media Discourse, Stylistics and Translation studies. He Has Published 50 Papers in Cultural Linguistics, Translation Theory and Media Discourse in Thomson Reuters, Scopus And Refereed Journals. Furthermore, He Has Written 3 Books in Linguistics and Translation. His Latest Book Entitled ' A Discourse Analytical Approach to Stylistic and Semantic Variations Employed in Arabic Translations of English News Stories', (LINCOM/In Press).