Available online at www.jlls.org

JOURNAL OF LANGUAGE AND LINGUISTIC STUDIES

ISSN: 1305-578X

Journal of Language and Linguistic Studies, 17(4), 2065-2080; 2021

Evaluation of google translate in rendering English COVID-19 texts into Arabic

Zakaryia Almahasees [a 1], Samah Meqdadi [b], Yousef Albudairi [c]

[a, b] *Applied Science Private University, Amman, Jordan*
[c] *University of Western Australia, Perth, Australia*

APA Citation:
Almahasees, Z., Meqdadi, S., & Albudairi, Y. (2021). Evaluation of google translate in rendering English COVID-19 texts into Arabic. *Journal of Language and Linguistic Studies, 17*(4), 2065-2080. Doi: 10.52462/jlls.149
Submission Date:20/05/2021
Acceptance Date:21/07/2021

Abstract

Machine Translation (MT) has the potential to provide instant translation in times of crisis. MT provides real solutions that can remove borders between people and COVID-19 information. The widespread of MT system makes it worthy of scrutinizing the capacity of the most prominent MT system, Google Translate, to deal with COVID-19 texts into Arabic. The study adopted (Costa et al., 2015a) framework in analysing the output of Google Translate output service in terms of orography, grammar, lexis, and semantics. The study's corpus was extracted from World Health Organization (WHO), United Nations Children's Emergency Fund (UNICEF), U.S. Food and Drug Administration (FDA), the Foreign, Commonwealth & Development Office (FCDO), and European Centre for Disease Prevention and Control (ECDC). The paper reveals that Google Translate committed a set of errors: semantic, grammatical, lexical, and punctuation. Such errors inhibit the intelligibility of the translated texts. It also indicates that MT might work as an aid to translate general information about COVID-19, but it is still incapable of dealing with critical information about COVID-19. The paper concludes that MT can be an effective tool, but it can never replace human translators.

Keywords: Machine Translation during COVID-19; English-Arabic Translation; Error Analysis; Google Translate; Machine Translation during crises

1. Introduction

On March 11st, 2020, the general director of WHO, Tedros Adhanom, declared that COVID-19 had become a global pandemic since it spreads rapidly worldwide (WHO, 2020). This statement has led world governments to impose restrictions on people's movements (S Haider & Al-Salman, 2020). Some countries enforced complete lockdown, where people were not allowed to go outside their homes except doing shopping and emergency cases (Al-Salman & Haider, 2021b). Translators' jobs have been affected due to COVID-19 restrictions, but the flow of COVID-19 information is unprecedented (Al-Salman & Haider, 2021b). Such flow of information is beyond the capacity of human translators, and therefore people use Machine Translation (MT) services to render English COVID-19 content into their languages. Almahasees (2020) shows that MT could help prevent the outbreak by rendering the available content into world languages.

[1] Corresponding author.
E-mail address: zmhases@hotmail.com

MT works as a tool to fight COVID-19 (Haider & Al-Salman, 2020). The importance of MT in providing a translation of COVID-19 content into Arabic makes it worth investigating to scrutinize the capacity of Google Translate in rendering English COVID-19 texts into Arabic in terms of Error Analysis.

1.1. Machine Translation (MT)

MT is the automatic translation from one language into another using computers. Theoretically, MT is a branch of computational linguistics, which deals with the computational modelling of natural languages. Machine Translation was firstly anticipated in 1930 to translate natural languages by George Artsuni and Trojanski. George Artsuni, a French engineer, proposed Mechanical Brain, which aimed to translate languages. He got a patent for his device, Mechanical Brain. However, it did not see the light of the day due to the inadequacy of its device to modern computers (Henisz-Dostert et al., 1979). In 1936, Trojanski suggested the first detailed process of translating across natural languages with the aid of machines. However, his project was not successfully applied to MT (Z. Almahasees, 2020) (Henisz-Dostert et al., 1979). Weaver 1949 is considered the father of MT (Z. Almahasees, 2020) since he mapped out the science of MT in his 'memorandum of Translation'.

At the rise of the cold war between the USA and the Soviet Union (now Russia) in 1954, Leon Dostert and Peter Sheridan conducted the first experiment on translating 250 words, and they did succeed. The success of the first experiment attracted a significant scale of funding to develop MT and its potential to translate across human languages. A committee followed the first success formed by the US government, ALPAC, in 1962 to evaluate MT. It issued its report in 1966 with a conclusion that "there is no predictable prospect of useful machine translation" (ALPAC, 1966, p.5). This report was described as catastrophic since it shut the door to further research on MT. Therefore, MT research halted in the USA and other countries except for Japan and France. They continued their research to use MT in weather forecast translation. 1980 was the year of MT revival research due to the new developments of technology, and it became dominant in the 1990s with the emergence of the Internet. However, in the first years of the Internet, MT service was paid due to the high costs of running MT systems and the Internet.

Currently, several MT platforms offer free MT service for all end-users, such as Google Translate and Microsoft Translator. The current study has chosen Google Translate since it is widely used system. Google Translate offers a free automatic translation service into 109 languages, including Arabic. MT service provided by Google is powered by Neural Machine Translation (NMT) approach. Moreover, it is widely used by more than 500 million users daily, with an estimated 100 billion words translated daily (Google Translate, 2021). To understand how Google Translate works, we should first understand the MT approaches that run the systems.

1.2. MT approaches

Historically, MT systems use machine-learning technologies to translate natural languages from one language into another. The first MT approach is Rule-Based MT (RBMT), which relies on linguistic information about the source and target texts retrieved from dictionaries and grammars. Then, Statistical-Based MT (SBMT) generates translation across languages based on statistical models from bilingual text corpora. Then, Neural Machine Translation (NMT) is designed to imitate the human brain in translation. It is an approach that uses neural networks to learn linguistic rules, which results in faster and accurate translation. The study adopts the MT system with NMT, Google Translate since it aims to mimic human brains in translation. Google Translate adopted NMT in 2017 due to its potential to mimic human translation. (ASIA Digital, 2021) describes NMT as, "universally accepted as the most accurate, fluent, and versatile approach to automatic translation." For this reason and others, it is of great importance to assess Google Translate understudy's capacity to translate English COVID-19 texts into Arabic.

1.3. Machine Translation Evaluation (MTE)

Since the primary function of MT is to provide an instant translation to the end-users, the notion of MT has significantly improved. Even though MT is still far from reaching human translation, it provides to some degree acceptable translation due to the system adaptability and training. In other words, each MT output should show its quality in terms of fluency and adequacy. Therefore, the evaluation of MT systems is considered an essential step in designing and accepting the system by the end-users (Z. Almahasees, 2020). In most cases, translation quality looks for output clarity, adequacy, and fluency as prerequisites to determine output acceptability (Z. M. Almahasees, 2017). Translation quality requires comprehension to determine various kinds of translation equivalence and identify translation errors (Chan, 2014). MT users should bear in their minds that MT performance is considerably improving during the first months of the system installation. However, MT evaluation is central to highlight the system's capacity since MT could not render linguistic issues in translation such as emotional impact and style (Hutchins & Somers, 1992).

There are two ways to evaluate MT systems: manual and automatic. Manual evaluation has counted on human evaluators. It is considered subjective, costly, and inconsistent since humans have different perspectives on each issue. On the other hand, automatic evaluation involves using automatic metrics to assess translation quality without human interventions. It is considered objective, cheap, and cost-effective since it provides instant evaluation.

1.3.1. Automatic Evaluation

Automatic evaluation relies on verifying translation quality in terms of text similarity through comparing MT output to human referenced translation, i.e., how much MT output is close to human translation. The prominent automatic metric to assess MT output is Bilingual Evaluation Understudy (BLEU). BLEU is, the first MT metric, designed by (Papineni et al., 2001) and the most prominent for evaluating MT output. Even though automatic evaluation is objective and cost-effective, there are some limitations. Automatic metric tells little about the translation quality (Pan, 2016). Additionally, they provide "only one side of the story about quality, which is not always useful in a production environment" (Panic, 2019). They are also considered an imperfect alternative for human translation quality evaluation (Kral & Václav, 2013) (Callison-Burch et al., 2006). For these reasons and others, the current paper adopts manual evaluation to ensure best practices and provide an overall assessment for the chosen systems under study.

1.3.2. Manual evaluation

Although manual evaluation has been described as subjective and inconsistent, it is regarded by the researcher and (Chan, 2014) as the best method to evaluate MT outputs, and automatic metrics cannot replace it. The manual evaluation focuses on the quality of MT output and the usefulness of MT in dealing with the specific task that MT is expected to do. MT can be evaluated manually in terms of intelligibility, accuracy, and error analysis (EA). Intelligibility evaluates MT to identify grammatical errors, mistranslations, and untranslated errors. Accuracy checks whether MT output preserves the ST meaning. Error analysis is the criterion for identifying errors found in MT output. (Costa et al., 2015a) show that error analysis is essential to all MT systems. Therefore, the current paper adopts error analysis to evaluate the output of Google Translate.

1.3.2.1. Error Analysis

Error Analysis works for identifying and classifying individual errors in the MT system's output. Such an evaluation highlights the strengths and limitations of an MT system. EA aims at identifying the error and the cause of unsuccessful language (Yang, 2010). It has been an essential part of MT assessment to highlight the limitations and improvements (Llitjós et al., 2005b). It is vital to find MT errors to compare MT output with referenced human translation. It scrutinizes MT output to provide information about ways to shed light on the improvements needed to provide an acceptable translation (Vilar et al., 2007) (Condon et al., 2010). Such evaluation provides the end-users with feedback concerning system designation, development, purchase, or use (Hutchins & Somers, 1992). Therefore, the present study adopts error analysis to classify errors, provide clues about the causes of errors and give a solution for the two chosen systems under study in rendering English into Arabic.

Several taxonomies have been proposed for MT error analysis (Flanagan, 1994) (Vilar et al., 2007)**;** (Frederking et al., 2004) and (Farrús et al., 2010). Several ones have been conducted on error analysis, but the most used referred taxonomy is (Vilar et al., 2007). Vilar et al.'s classification are hierarchical. They held (Llitjós et al., 2005a) classification out and divided errors into five categories: missing words, word order, incorrect words, unknown words, and punctuation errors. Missing words show that some words are missing from the translation output. Incorrect words show that some words are wrongly chosen. Word order resorts to the word order of the MT output. Unknown words are translated simply by changing the letters using the romanization strategy. Punctuation errors represent errors in punctuation marks such as addition and omission of such marks.

Similarly, (Vilar et al., 2007) introduced five categories to classify MT errors: missing words, word order, incorrect words, unknown words, and punctuation. Likewise, (Condon et al., 2010) developed a similar classification based on (Vilar et al., 2007) to scrutinize MT errors from English into Arabic. They investigated MT errors in a corpus of 100 translations. They categorized MT errors into deletions, insertions, word orders, and substitutions. They found out that MT errors occurred at the level of pronouns in Arabic and English translations. (Bojar, 2011) utilized Villar et al.'s classification to classify English-Czech errors of four MT systems: Google, PC translator, TectoMT, and CU-Bojar. He classified errors into punctuation marks, missing words, word order, and incorrect words. He found that systems with a statistical MT approach can achieve better results than systems with previous MT approaches. Previous studies classified MT errors in terms of punctuation, lexis, and structure. However, none of the previous studies have analyzed MT in terms of semantics and texts' discourse.

In 2015, all-inclusive taxonomy was introduced by (Costa et al., 2015b). They extended the previous MT error taxonomies to cope with the analysis of Romance languages. They conducted a thorough analysis to scrutinize MT errors in four MT systems: Google Translate, Systran, and two in-house MT systems in orthography, lexis, grammar, semantics, and discourse. They found out that there are several challenges of English into European Portuguese translation. They also concluded that the recurrent errors are due to wrong choice problems and the inability to find the proper choices. Therefore, (Costa et al., 2015a) are considered the best taxonomy to assess translation quality, as shown in figure 1.

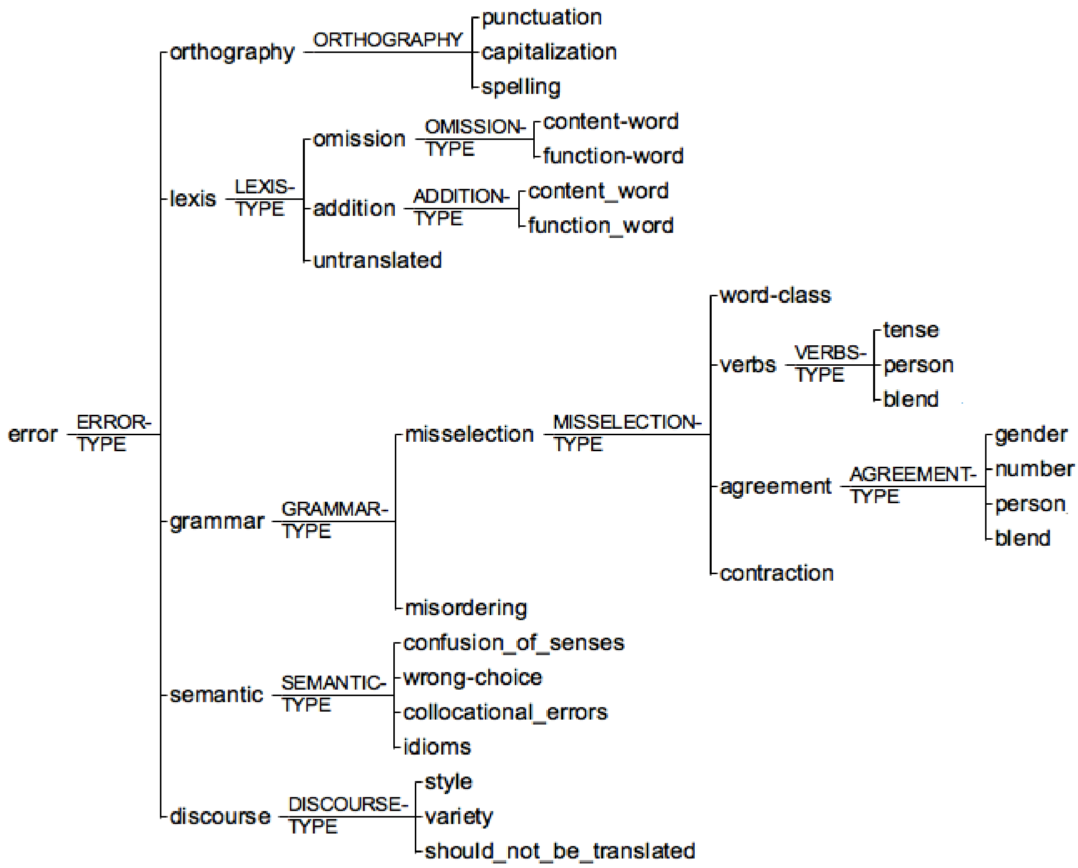


Figure 1. Taxonomy of Errors

In summary, (Costa et al., 2015a) framework would lead to constructive feedback about the capacity of the two systems understudy in the light of error analysis. This feedback would help the systems' developers improve1+3 efficiency in translating health information during catalyst times, such as COVID-19.

2. Literature Review

Several studiesn (Al-Salman & Haider, 2021a) have been conducted to verify the strengths and limitations of MT. However, a small scale of studies was conducted on MT during Covid-19.

(Way et al., 2020) indicates the number of infected people with COVID-19, and the fertility rate was high in European countries. The health professionals and the general public were keen to update their information on COVID-19. They developed an MT system, which helps to translate COVID-19 information published in German, French, Italian, and Spanish into English. (Z. Almahasees & Jaccomard, 2020) conducts a paper on Facebook Translation Service (FTS). They distributed a survey to know the percentage of FTS in Jordan and its usage during COVID-19. They found out that FTS helped to disseminate information about COVID-19 and worked as an aid to Jordanians.

In another study, (Z. Almahasees et al., 2021) scrutinizes the adequacy and fluency of FTS from English into Arabic. They found out that FTS provided an adequate and fluent translation output for general information about COVID-19. However, it could not translate medical information correctly; human translators should post-edit and review FTS output to ensure its quality. The above studies contributed to the field, but they do not detail the errors committed to translating English into Arabic COVID-19 content. (Dalzell, 2020) indicated that the Australian Federal Government used Google

Translate to send critical health information about COVID-19 to multicultural communities. Mohammad Al-Khafaji, the chief executive of the peak multicultural body, the Federation of Ethnic Community Councils Australia (FECCA), indicated that Google Translate was unacceptable and risky to translate critical health information to multicultural communities in Australia. (Moreno, 2021) shows that the Department of Health at Virginia state uses Google Translate to translate critical COVID-19 and vaccine information. He indicates that Google Translation provides wrong information due to the inability to provide accurate translation for vital information. He explains that it is not acceptable to use Google Translate to translate vital information to immigrant communities if a professional translator has not posted and reviewed the translation first. (Goodman, 2021) shows that Google Translate can help translate general information, but it could not translate vital information about COVID-19.

3. Methodology

The present research has chosen (Costa et al., 2015b) framework to assess the MT output of Google Translate in this language pair. The corpus of the study has been selected from credible health organizations: WHO, UNICEF, U.S. FDA, FCDO, and ECDC. The rationale behind this choice is the credibility of these sources in providing reliable information to the end-users about the global pandemic, COVID-19. The adopted research method for this study aims to provide the end-users with a solid feedback about the quality of the translation in terms of error analysis, which is regularly employed for evaluating the quality of machine translation. Costa et al.'s error analysis framework is the most prominent one that relates mutually between human error analysis frameworks and all previous MT error taxonomies, as shown in Figure 1. The errors were identified, tabulated, and counted in both systems. Such errors were shown through examples accompanied by explanations and the systems' rankings in dealing with COVID-19 content. Moreover, back translation was used when relevant to provide an accurate translation of the given examples.

4. Results

The analysis of the errors includes orthographic, lexical, grammatical, and semantic errors.

4.1. Orthographic Errors

An orthography is a group of rules that govern the writing of a language (Merriam-Webster, 2020). It includes spelling, capitalization, and punctuation marks. Orthographic rules are different among languages. For example, unlike Arabic, there are capitalization rules in English. Orthographical errors occur when translating any text into another language due to the differences among languages. The analysis of Google Translate output shows that the system has achieved significant progress in rendering English texts into Arabic free of spelling errors. It also shows that Google commits punctuation errors.

4.1.1. Punctuation Errors

Punctuation marks facilitate reading since they guide readers to deduce the meaning of the text. In translation, they have an essential role in reflecting the fluency of the translation. The improper usage of punctuation may inhibit the fluency of the output. The following examples illustrate punctuation errors committed by Google Translate:

Example 1:

Source Text: "Their already-dire situation has been compounded by the pandemic, which forced the government to introduce a lockdown that left many in the country out of work and with no income (UNCIEF, 2021)".

Google Translate: وقد تفاقم وضعهم المزري بالفعل بسبب الوباء، الذي أجبر الحكومة على فرض حظر ترك الكثيرين في البلاد عاطلين عن العمل وبدون دخل.

The above example illustrates how the improper usage of punctuation impacts the fluency of the translated text. The output has an Arabic relative pronoun الذي, which refers to the masculine singular noun الوباء pandemic. The output pinpoints the inability of Google Translate to render texts without considering the relative pronouns between two different languages. In Arabic, relative pronouns are not preceded by commas, unlike in English. In the given an example above, the relative pronoun الذي is preceded by a comma, which does not conform with Arabic syntax. Google Translate here translates the text by imitating the punctuation marks of the source text since Arabic, unlike English, does not have systematic punctuation rules. Therefore, the system copies the ST rules of punctuation.

Example 2:

Source Text: If COVID-19 is spreading in your community, stay safe by taking some simple precautions, such as physical distancing, wearing a mask, keeping rooms well ventilated, avoiding crowds, cleaning your hands, and coughing into a bent elbow or tissue. Check local advice where you live and work. Do it all!" (WHO, 2021b).

Google Translate:

إذا كان COVID-19 ينتشر في مجتمعك ، فابق آمنًا من خلال اتخاذ بعض الاحتياطات البسيطة ، مثل التباعد الجسدي ، وارتداء قناع ، والحفاظ على تهوية الغرف جيدًا ، وتجنب الازدحام ، وتنظيف يديك ، والسعال في كوع مثني أو منديل. تحقق من النصائح المحلية في المكان الذي تعيش فيه وتعمل. قم بها كلها !

The error in the above example is the usage of the exclamation mark. Exclamation marks are used to express exclaim, protest, command, surprise, or astonishment. They are different in English and Arabic. Google translate imitates the usage of punctuation in English in handling English texts into Arabic. In Arabic, it starts with the exclamation particle *ma* ما *and* on the comparative form of *afal* *أفعل* of the appropriate adjective while in English, it can be used with imperative sentences like 'Do it all!'. Google Translate applied ST exclamation marks while dealing with the Arabic text. The use of the exclamation article in (قم بها كلها!) is incorrect because an exclamation article does not follow imperative sentences in Arabic; it should end with a full stop (قم بها جميعها.).

4.2. Lexical Errors

The lexical level relates to the usage of words and vocabulary in a language. A lexical error concerns the influence of using a wrong word in the wrong context throughout the translation process and how such a type of error affects the whole meaning. Lexis errors include omitting, adding, and untranslated errors. Omission errors indicate the deletion of words that should appear in the translated text. Addition errors represent the addition of new words to the translated text which do not exist in the source text. However, not all additions and omissions are considered errors unless they affect the comprehensibility of the text. In translation, the lexical combination has an essential meaning on the text; the inappropriate translation impacts the intelligibility of the translated text.

4.2.1. Omission Errors

Example 3:

Source Text: “However, several NPIs can have a negative impact on the general well-being of people, the functioning of society, and the economy” (ECDC, 2021).

Google Translate: ومع ذلك ، يمكن أن يكون للعديد من المؤسسات التي لا تستهدف الربح تأثيرًا سلبيًا على الرفاه العام للأفراد ، وعمل المجتمع ، والاقتصاد.

The source text contains an acronym. The acronym is a shortened form of a written word or phrase. The acronym NPIs stands for Non-pharmaceutical interventions, which is mistranslated into Arabic as المؤسسات غير الربحية. MT systems usually tend to keep the abbreviations untranslated if the system is not familiar with the given abbreviation. However, Google Translate in this context provides an incorrect translation for NPIs as Non-profit organizations المؤسسات الغير ربحية. Such a translation affects the meaning of the translated text. The system here provides the translation of the acronym, which does not relate to the source text. The correct translation for the acronym should be التدخلات الغير صيدلانية. The mistranslation error in this example indicates that the chosen system could not recognize the connection between the abbreviation and its context. Therefore, the study recommends creating specialized lists for different domains and trains Google Translate on dealing with abbreviation translation based on its domain reference.

Example 4:

Source Text: “Easy access to testing and timeliness of testing is critical for the effectiveness of measures such as contact tracing and isolation of cases (ECDC, 2021)”

Google Translate: يعد الوصول السهل للاختبار وحسن توقيته أمرًا بالغ الأهمية لفعالية التدابير

The above example shows an incomplete translation for the given sentence. The chosen system here omits the translation of the underlined phrase. Such an omission is critical and inhibits the intelligibility of the text. The translation of the above example 5 should be as follows:

Back Translation: يعد الوصول السهل للاختبار وحسن توقيته أمراً بالغ الأهمية لفعالية التدابير مثل تعقب وعزل الحالات.

Example 5:

Source Text: Maintain at least a 1-metre distance between yourself and others to reduce your risk of infection when they cough, sneeze or speak.

Google Translate: احتفظ بمسافة لا تقل عن متر واحد بينك وبين الآخرين لتقليل خطر الإصابة بالعدوى عند السعال أو العطس أو التحدث

The above example shows that the possessive adjective ‘your’ and subject pronoun ‘they’ have been omitted. The advice in the ST is for the public. Therefore, the omission does not affect the meaning of the text.

Back Translation: ابتعد مسافة متر واحد على الأقل عن الآخرين للحد من مخاطر الإصابة بالعدوى عندما يسعلون أو يعطسون أو يتكلمون(WHO, 2021a) .

4.2.2. Untranslated Errors

Example 6:

Source Text: • Avoid the 3Cs: spaces that are closed, crowded or involve close contact(WHO, 2021a).

Google Translate: تجنب 3 :CSالمساحات التي هي جlosed ، ج rowdedأو تنطوي ج الاتصال تفقد

The above example illustrates how the chosen system translated the COVID-19 preventive measures advice from English into Arabic. The text asks the public to avoid three words that start with the letter C, ‘closed, crowded or involve close contact.’ The chosen system here dealt with 3Cs as an abbreviation. It translated the letter ‘C’ letter into three mentioned words start with the letter C. Keeping the words untranslated inhibits the comprehension of the output. ج with

Back Translation: .تجنب الميمات الثلاثة: الأماكن المغلقة أو المكتظة أو التي تنطوي على مخالطة قريبة (لصيقة)

4.2.3. Addition Errors

Example 7:

Source Text: WHO has published Q&As on ventilation and air conditioning for both the general public and people who manage public spaces and buildings(WHO, 2021a).

Google Translate:

وقد نشرت منظمة الصحة العالمية Q & باسم التهوية وتكييف الهواء لكل من الجمهور و الناس الذين يديرون الأماكن العامة والمباني.

The ST has a Q&A expression used in sessions to give the audience time to ask specific issues and topics. The expression has been kept untranslated. On the other hand, the chosen system adds the proper name, باسم, which is a proper Arabic name. This addition is wrong since it impacts the comprehension and intelligibility of the text.

4.3. Grammatical Errors

Grammar is a set of rules that govern the structure of a language. Grammar errors cover subject-verb agreement, conjugations, and word order. The misuse of derivations in language's morphological and syntactic aspects causes grammatical errors that affect the target text's structure and meaning. In the present analysis, we identified and highlighted two types of errors: Misselection errors and Misordering errors. Misselection errors represent the morphological problems that occur due to word class level (when a noun is needed but the translation engine translates it as a noun), verbal level (in terms of tense and person), and agreement error (includes gender, person, and number).

4.3.1. Misselection error (word class)

Example 8:

Source Text: “Today, the U.S. Food and Drug Administration approved the antiviral drug Veklury (remdesivir) for use in adult and pediatric patients” (FDA, 2020).

Google Translate:

وافقت إدارة الغذاء والدواء الأمريكية اليوم على العقار المضاد للفيروسات (remdesivir) Veklury للاستخدام في المرضى البالغين والأطفال

Misselection error occurs due to the misuse of the word class for (use in) when translating into Arabic. The ST text contains the preposition that comes with the noun ‘use.’ Google Translate translates the preposition "in" literally as في to indicate the place. In contrast, ST text indicates "in" as medicine. The ST preposition is equivalent to the Arabic preposition ل.

4.3.2. Misselection error (verb level: tense)

Example 9:

Source Text: “The COVID-19 pandemic has taken a devastating toll on hundreds of millions of people across the globe” (UNCIEF, 2021).

Google Translate:

تسبب جائحة COVID-19 في خسائر فادحة لمئات الملايين من الأشخاص في جميع أنحاء العالم.

Nevertheless, the verb “تسبب “can be considered as a present tense or past tense based on the context. The present perfect tense “has taken” means that COVID-19 caused a devastating toll, and it should be translated into Arabic by using the past simple tense تسبب. A Misselection error happens in the Arabic translation due to the mismatch between the feminine word “pandemic=جائحة “and the verb tense تسبب. To avoid such a type of error, the pronoun reference “ت “to the female subject should be added to the end of the verb “تسبب “for two reasons; to indicate the past tense of the verb and to match the feminine word following it "" .تسببت جائحة

4.3.3. Misordering Error

Example10:

Source Text: There were no statistically significant differences in recovery rates or mortality rates between the two groups.

Google Translate:

ولم تكن هناك ذات دلالة إحصائية الاختلافات في معدلات الشفاء أو معدلات الوفيات بين المجموعتين.

The translation of “statically significant differences” has resulted in a misordering error. In Arabic, the noun precedes the adjective, while in English, the adjective precedes the noun. The chosen system translates “statistically significant differences as ذات دلالة إحصائية الاختلافات which places the adjective before the noun. The correct translation of “statistically significant differences” should be “ فروق ذات دلالة إحصائية.”

4.4. Semantic Errors

Semantic errors are issues that relate to the meaning of the words. The semantic errors are of three types: confusion of senses, wrong choice, collocation, and idiomatic errors. The confusion of senses is when the translated word constitutes one of its possible meanings, but the chosen translation is not accurate. The wrong choice is when the translation does not relate to the source text word. Collocation and idiomatic errors occur when the system is not capable of rendering these two-word combinations correctly. In most cases, the equivalent for these combinations is very different from the system translation.

4.4.1. Confusion of Senses

Example 11:

Source Text: “authorize the drug’s use for treatment of suspected or laboratory confirmed COVID-19 in hospitalized pediatric patients” (FDA, 2020).

Google Translate:

للسماح باستخدام الدواء لعلاج COVID-19 المشتبه به أو المختبر في مرضى الأطفال في المستشفى

The above example illustrates confusion of senses error at the level of contextual translation. The Arabic translation of laboratory has one of its possible meaning as المختبر as a place to conduct experiments. The translation is incorrect; the context deals with confirmed cases of COVID-19, not the place of conducting experiments. Moreover, the translation of "hospitalized paediatric patients" also has confusion of senses errors since the translation does not indicate whether the patients are regular visitors to the hospital or hospitalized there. Based on the English sentence, the drug can treat hospitalized paediatric patients whose cases are suspected or laboratory-confirmed of having COVID-19.

Back Translation:

للسماح باستخدام الدواء لعلاج المرضى والأطفال الذين أدخلوا المستشفى ممن يشتبه إصابتهم أو أثبتت إصابتهم مخبريا بكوفيد19.

4.4.2. Wrong choice errors

Example 12:

Source Text: “stay safe by taking some simple precautions, such as physical distancing, wearing a mask”(WHO, 2021b).

Google Translate: فابق آمنًا من خلال اتخاذ بعض الاحتياطات البسيطة ، مثل التباعد الجسدي ، وارتداء قناع

The above example contains the wrong choice error for the colocation ‘wearing a mask.’ Google Translate here mistranslates the noun mask as قناع. The word قناع indicates the cover for the whole face for the sake of disguise and entertainment. In contrast, the context indicates wearing the mask to protect your mouth and nose for medical purposes, a surgical mask, or a medical mask to prevent airborne infections. The translation for the ‘mask’ should be كمامة طبية or الأقنعة الطبية.

4.4.3. Idioms Errors

Example 13:

Source Text: As business evaporated, so too did their savings.

Google Translate: مع تبخر الأعمال ، تبخرت مدخراتهم أيضًا

The above example illustrates the impact of the pandemic on humans that their businesses evaporated (lost). The chosen system translates the phrase literally as تبخر الاعمال, which does not convey the exact meaning of the translated text. The context indicates that people lost their jobs and their saving is in danger too. The correct translation of this idiom is ‘مع تلاشي فرص العمل تلاشت مدخراتهم’.

Example 14:

Source Text: A government-led emergency cash transfer program for informal workers in urban areas has provided a lifeline for parents struggling to put food on the table

Google Translate: وفر برنامج التحويلات النقدية الطارئ الذي تقوده الحكومة للعمال غير الرسميين في المناطق الحضرية شريان حياة للآباء والأمهات الذين يكافحون من أجل وضع الطعام على المائدة

The above example shows how Google Translate rendered the idiom ‘put food on the table.' The chosen system translates the idiom literally as وضع الطعام على المائدة ‘put food on the table.’ However, the idiom means "to earn enough money to cover all the necessities for oneself and his/her family." Therefore, the correct translation of the idiom should be توفير قوت يومهم.

5. Discussion

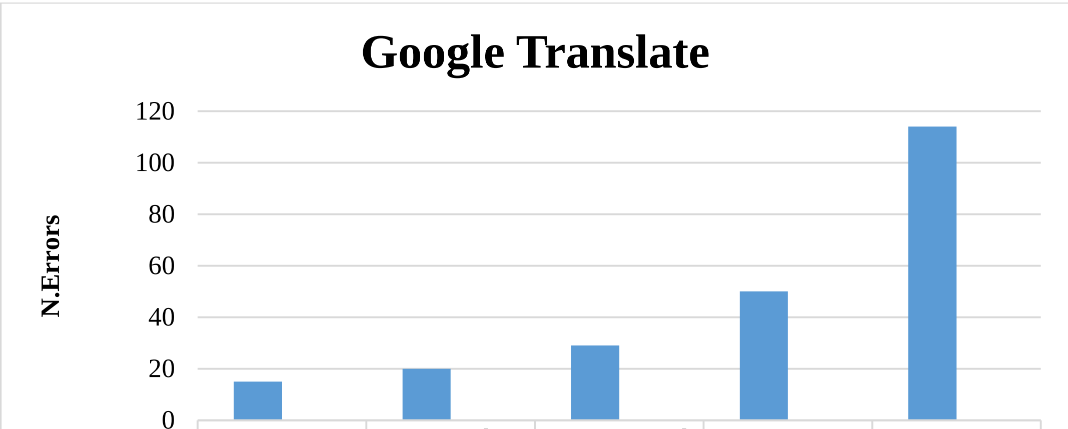
Several MT taxonomies and methods have been proposed for the assessment of MT systems. The study indicates that the most suitable method for evaluating the output of MT systems is the manual method since humans can judge the quality of MT systems in terms of adequacy, fluency, and intelligibility of the output. This conclusion agrees with (Chan, 2014) (Vilar et al., 2007), (Costa et al., 2015b) and (Z. Almahasees, 2020) that manual evaluation is the golden standard method for assessing MT outputs.

The development of MT systems has improved significantly. Different approaches have been proposed for improvement along with evaluative studies. The prominent approach is NMT, initiated by Google Translate in 2017 for English into Arabic and vice versa. The study aligns with (Z. Almahasees, 2020) and (Alkhawaja et al., 2020) that NMT proved significant progress in translating English into Arabic.

The study shows that Google Translate performs well in rendering COVID-19 content into Arabic. However, there are still mismatches in translation. Google Translate can help translate general safety instructions about Covid-19, but it is risky and not trusted in translating critical information about COVID-19. The study’s analysis shows that Google Translate rendered some abbreviations incorrectly in the corpus. These errors are due to the linguistic difference between English; unlike English, Arabic does not have a systematic system for punctuation. The system commits errors in rendering abbreviations due to the unfamiliarity of the systems with the newly invented abbreviation, as in example 4. Moreover, the system imitates ST punctuation marks, which impacts the fluency and intelligibility of the output since punctuation marks are different in English and Arabic.

Similarly, there are also significant errors in terms of lexis that inhibit the output's intelligibility, as shown in Figure 2. The meaning of lexis words emanates from the context. The context contains surface and underlying meaning. The analysis of context is attainable only by humans. This error emphasizes that Machine Translation in general and Google Translate, in particular, is still incapable of dealing with the contexts like a human.

On the other hand, Goggle commits grammatical errors since the word order, and the grammatical structure of this language pair is different. This finding aligns with what (Z. M. Almahasees, 2017) and(Z. Almahasees, 2020) found in their studies. The following chart illustrates the distribution of errors over (Costa et al., 2015a) taxonomy of errors.



	Punctuation Errors	Lexical Errors	Grammatical Errors	Semantic Errors	Sum
N.Errors	15	20	29	50	114
Percentage	1.46%	3.90%	4.26%	8.53%	18.15%

Figure 2. Google Translate Performance in dealing with COVID-19 translation texts into Arabic

Google Translate committed a set of errors while translating English COVID-19 texts into Arabic. The paper reveals that Google Translate committed the highest number of semantic errors with 8.53% of the whole texts. In comparison, the lowest number of errors goes for punctuation errors with a percentage of less than 2% of the whole texts, equal to 1.46%. On the other hand, grammatical errors come after semantic errors with 4.26% and lexical errors with 3.90%.

6. Conclusion

The present study has illustrated the importance of integrating technology and translation in coping with the demand for translation. The study verified the capacity of the most common MT system, Google Translate, with daily users of 500 million in translating a wide range of selected COVID-19 texts from international organizations: WHO, UNICEF, ECDC, FDA from English into Arabic. The paper shows that Google achieved a significant improvement in translating English COVID-19 texts into Arabic. However, it committed punctuation, lexis, grammatical and semantic errors. In this regard, the highest number of errors committed by Google is related to semantic errors, which inhibited the intelligibility of the texts, followed by grammatical and then lexical errors. The study recommends that a review by a trained translator should post edit the output of MT systems to ensure the quality of the output. Even though MT helps provide the gist of texts, it will never replace humans.

7. Ethics Committee Approval

The authors confirm that the study does not need ethics committee approval according to the research integrity rules in their country.

8. Conflict of Interest

The Authors declare that there is no conflict of interest.

References

Al-Salman, S., & Haider, A. S. (2021a). Covid-19 trending neologisms and word formation processes in english. *Russian Journal of Linguistics*, *25*(1), 24–42. https://doi.org/10.22363/2687-0088-2021-25-1-24-42.

Al-Salman, S., & Haider, A. S. (2021b). Jordanian University Students' Views on Emergency Online Learning during COVID-19. *Jordanian University Students' Views on Emergency Online Learning during COVID-19 Online Learning Journal*, *25*. https://doi.org/10.24059/olj.v25i1.2470

Alkhawaja, L., Ibrahim, H., Ghnaim, F., & Awwad, S. (2020). Neural Machine Translation: Fine-Grained Evaluation of Google Translate Output for English-to-Arabic Translation. *International Journal of English Linguistics*, *10*(4), 43. https://doi.org/10.5539/IJEL.V10N4P43

Almahasees, Z. (2020). *Diachronic Evaluation of Google Translate, Microsoft Translator and Sakhr in English-Arabic Translation — the UWA Profiles and Research Repository*. https://research-repository.uwa.edu.au/en/publications/diachronic-evaluation-of-google-translate-microsoft-translator-an

Almahasees, Z., & Jaccomard, H. (2020). Facebook translation service (FTS) usage among jordanians during COVID-19 lockdown. *Advances in Science, Technology and Engineering Systems*, *5*(6), 514–519. https://doi.org/10.25046/aj050661

Almahasees, Z. M. (2017). Assessing the Translation of Google and Microsoft Bing in Translating Political Texts from Arabic into English. *International Journal of Languages, Literature and Linguistics*, *3*(1), 1–4. https://doi.org/10.18178/ijlll.2017.3.1.100

Almahasees, Z., Mohammad, A.-T., & Jaccomard, H. (2021). Evaluation of Facebook Translation Service (FTS) in Translating Facebook Posts from English into Arabic in Terms of TAUS Adequacy and Fluency during Covid-19. *Advances in Science, Technology and Engineering Systems Journal*, *6*(1), 1241–1248. https://doi.org/10.25046/AJ0601141

ALPAC. (1966). *Language And Machines Computers In Translation And Linguistics*.

ASIA Digital. (2021). *Neural Machine Translation - AISA Digital*. https://www.aisa.digital/neural-machine-translation/

Bojar, O. (2011). Analyzing Error Types in English-Czech Machine Translation. *The Prague Bulletin of Mathematical Linguistics NUMBER*, 63–76. https://doi.org/10.2478/v10108-011-0005-2

Callison-Burch, C., Osborne, M., & Koehn, P. (2006). *Re-evaluating the Role of BLEU in Machine Translation Research*. https://www.aclweb.org/anthology/E06-1032

Chan, S. (2014). *Routledge Encyclopedia of Translation Technology - 1st Edition - Sin-W*. https://www.routledge.com/Routledge-Encyclopedia-of-Translation-Technology/Chan/p/book/9780367570439

Condon, S., Parvaz, D., Aberdeen, J., Doran, C., Freeman, A., & Awad, M. (2010). *Evaluation of Machine Translation Errors in English and Iraqi Arabic*. http://www.lrec-conf.org/proceedings/lrec2010/pdf/106_Paper.pdf

Costa, Â., Ling, W., Luís, T., Correia, R., & Coheur, L. (2015a). A linguistically motivated taxonomy for Machine Translation error analysis. *Machine Translation*, *29*(2), 127–161. https://doi.org/10.1007/s10590-015-9169-0

Costa, Â., Ling, W., Luís, T., Correia, R., & Coheur, L. (2015b). A linguistically motivated taxonomy for Machine Translation error analysis. *Machine Translation*, *29*(2), 127–161. https://doi.org/10.1007/s10590-015-9169-0

Dalzell, S. (2020). *Federal Government used Google Translate for COVID-19 messaging aimed at multicultural communities - ABC News*. https://www.abc.net.au/news/2020-11-19/government-used-google-translate-for-nonsensical-covid-19-tweet/12897200

ECDC. (2021). *Guidelines for the implementation of non-pharmaceutical interventions against COVID-19*. https://www.ecdc.europa.eu/en/publications-data/covid-19-guidelines-non-pharmaceutical-interventions

Farrús, M., Farrús, F., Costa-Jussà, M. R., Mariño, J. B., Mariño, M., & Fonollosa, J. A. R. (2010). *Linguistic-based Evaluation Criteria to identify Statistical Machine Translation Errors*. http://translate.google.com

FDA. (2020). *FDA Approves First Treatment for COVID-19 | FDA*. https://www.fda.gov/news-events/press-announcements/fda-approves-first-treatment-covid-19

Flanagan, M. A. (1994). Error Classification for MT Evaluation. *Amta*, 65–72.

Frederking, R. E., Taylor, K. B., Elliott, D., Hartley, A., & Atwell, E. (2004). A Fluency Error Categorization Scheme to Guide Automated Machine Translation Evaluation. In *LNAI* (Vol. 3265). http://eprints.whiterose.ac.uk/82298/

Goodman, B. (2021). *Lost in Translation: Language Barriers Hinder Vaccine Access*. https://www.webmd.com/vaccines/covid-19-vaccine/news/20210426/lost-in-translation-language-barriers-hinder-vaccine-access

Google Translate. (2021). *Google - About Google, Our Culture & Company News*. https://about.google/?hl=en

Haider, A. S., & Al-Salman, S. (2020). Dataset of Jordanian university students' psychological health impacted by using e-learning tools during COVID-19. *Data in Brief*, *32*, 106104. https://doi.org/10.1016/j.dib.2020.106104

Henisz-Dostert, B., Macdonald, R., & Zarechnak, M. (1979). *Machine Translation - Bozena Henisz-Dostert, R. R. Macdonald, Michael Zarechnak - Google Books*. https://books.google.mu/books?id=St4iXxXoIIAC&printsec=frontcover#v=onepage&q&f=false

Hutchins, W., & Somers, H. (1992). An introduction to machine translation. *Undefined*.

Kral, P., & Václav, M. (2013). *Text, Speech, and Dialogue: 18th International Conference, TSD 2015, Pilsen ... - Google Books*. https://books.google.jo/books?id=SQuVCgAAQBAJ&pg=PA218&lpg=PA218&dq=imperfect+su

bstitute+for+human+assessment+of+translation+quality&source=bl&ots=3zNabTwKIG&sig=ACfU3U2FkEIcGZMi9i50MBc0VL8MaT3o1A&hl=en&sa=X&ved=2ahUKEwjU3tv26YfvAhUJGewKHZNYBysQ6AEwBHoECA

Llitjós, A. F., Carbonell, J. G., & Lavie, A. (2005a). *A framework for interactive and automatic refinement of transfer-based machine translation*. https://www.aclweb.org/anthology/2005.eamt-1.13

Llitjós, A. F., Carbonell, J. G., & Lavie, A. (2005b). *A Framework for Interactive and Automatic Refinement of Transfer-based Machine Translation* (Vol. 87).

Merriam-Webster. (2020). *Orthography | Definition of Orthography by Merriam-Webster*. Https://Www.Merriam-Webster.Com/Dictionary/. https://www.merriam-webster.com/dictionary/orthography

Moreno, S. (2021). *Virginia uses Google Translate for COVID vaccine information. Here's how that magnifies language barriers, misinformation | Richmond Local News | richmond.com*. https://richmond.com/news/local/virginia-uses-google-translate-for-covid-vaccine-information-heres-how-that-magnifies-language-barriers-misinformation/article_715cb81a-d880-5c98-aac5-6b30b378bbd3.html

Pan. (2016). *How BLEU Measures Translation and Why It Matters | Slator*. https://slator.com/technology/how-bleu-measures-translation-and-why-it-matters/

Panic, M. (2019). *DQF-MQM: Beyond Automatic MT Quality Metrics - TAUS*. https://blog.taus.net/dqf-mqm-beyond-automatic-mt-quality-metrics?utm_campaign=Quality Dashboard&utm_content=82798944&utm_medium=social&utm_source=twitter&hss_channel=tw-112671713

Papineni, K., Roukos, S., Ward, T., & Zhu, W.-J. (2001). BLEU: a method for automatic evaluation of machine translation. *ACL*, 311–318. https://doi.org/10.3115/1073083.1073135

S Haider, A., & Al-Salman, S. (2020). COVID-19'S IMPACT ON THE HIGHER EDUCATION SYSTEM IN JORDAN: ADVANTAGES, CHALLENGES, AND SUGGESTIONS. *Humanities & Social Sciences Reviews*, *8*(4), 1418–1428. https://doi.org/10.18510/hssr.2020.84131

UNCIEF. (2021). *Protecting families from the economic impact of COVID-19 | UNICEF*. https://www.unicef.org/coronavirus/protecting-families-economic-impact-COVID-19

Vilar, D., Xu, J., Fernando D'haro, L., & Ney, H. (2007). *Error Analysis of Statistical Machine Translation Output*. http://www.tc-star.org/

Way, A., Haque, R., Xie, G., Gaspari, F., Popovic, M., & Poncelas, A. (2020). Facilitating Access to Multilingual COVID-19 Information via Neural Machine Translation. *ArXiv*. http://arxiv.org/abs/2005.00283

WHO. (2021a). *Advice for the public*. https://www.who.int/emergencies/diseases/novel-coronavirus-2019/advice-for-public

WHO. (2021b). *When and how to use masks*. https://www.who.int/emergencies/diseases/novel-coronavirus-2019/advice-for-public/when-and-how-to-use-masks

Yang, W. (2010). A Tentative Analysis of Errors in Language Learning and Use. *Journal of Language Teaching and Research*, *1*(3), 266–268. https://doi.org/10.4304/jltr.1.3.266-268

AUTHOR BIODATA

Dr. Almahasees is an Assistant Professor of Translation in the Dept. of English Language and Translation at Applied Science Private University, Jordan. He earned his PhD in Translation Studies, from The University of Western Australia, Australia (2020). His research interests include Translation Theories, Translation Evaluation, Comparative Translation, Computer Assisted Translation (CAT), Computational Linguistics, and Machine Translation

Lect. Meqdadi is a lecturer of Linguistics. She held a Master's degree in English Language-Literature from Hashemite University in 2015. She taught as a part-time lecturer at the University of Jordan and the Hashemite University. She also worked as a free-lance translator and she taught Arabic as a foreign language. She worked in the Department of English Language and Translation at Applied Science Private University from 2017 to July, 2021.

Albudairi is a lecturer at the College of Languages and Translation, Imam Mohammad Ibn Saud Islamic University, Riyadh. He has an MA in English-Arabic translation and currently he is a PhD student in translation. He has an experience as a translator and as a freelancer. Among his research interests are translation strategies, translation evaluation and machine translation.