AWEJ for Translation & Literary Studies
International Peer Reviewed Journal ISSN: 2550-1542

AWEJ for Translation & Literary Studies, Volume 7, Number 1. February 2023 Pp.208-219
DOI: http://dx.doi.org/10.24093/awejtls/vol7no1.16

Google Translate Errors in Legal Texts: Machine Translation Quality Assessment

Eman Rashed Alkatheery
English Language Department, College of Languages Studies
King Saud University, Riyadh, Kingdom of Saudi Arabia
ekathiri@ksu.edu.sa

Received: 11/07/2022 Accepted: 01/31/2023 Published: 02/24/2023

Abstract
Machine translation received intensive research in different language pairs; yet, the quality of specialized translation; e.g., legal translation, from Arabic into English received little attention. The present study investigates the quality of machine translation of legal texts from Arabic into English. The paper aims at examining the errors found in the machine translation of legal texts from Arabic into English. It also studies the legal discourse features in machine-translation output. The research questions tackle the accuracy of machine translation of legal discourse and error categories and frequencies in machine translation. The researcher evaluated several factors to assess the quality of Google Translate; i.e., lexical, syntactic, and register-related errors. The study data consists of five legislative texts. The researcher conducted a manual error assessment and classification. To ensure the reliability of the error analysis, an existing human translation of the documents was used as a reference to ensure the reliability of the MT quality assessment and post-editing process. Later, the errors were classified, and their percentages and frequencies were calculated. A few examples of errors in each category were discussed and analysed. The highest error category was lexical errors scoring 43.4%. The last detected error category was deletion with a percentage of 1.7%. Syntactic errors constituted one-fourth of the errors found in the data. Legal register-related errors scored 30.2%. The subcategories of legal register-related errors varied in their occurrence; e.g., one-third of the errors related to legal discourse were legal terms. The study concluded that machine translation; though it provided a comprehensible output, could not translate legal structures and terminology perfectly.
Keywords: Machine Translation, Translation Quality Assessment, Legal Translation, Translation Error Typology, Google Translate

Cite as: Alkatheery, E.R. (2023). Google Translate Errors in Legal Texts: Machine Translation Quality Assessment. *Arab World English Journal for Translation & Literary Studies 7* (1): 208-219. DOI: http://dx.doi.org/10.24093/awejtls/vol7no1.16

Introduction

Machine Translation (MT) began in the second half of the last century. It went through several phases of development, with different approaches to machine translation that led to the gradual improvement in machine translation output. The introduction of Example-Based Machine Translation (EBMT) and Statistical Machine Translation (SMT), followed by Neural Machine Translation (NMT) paved the way towards better MT output.

Several aspects of MT between Arabic and English have been investigated. For instance, some studies examined error typology, others conducted comparisons between several MT engines, and other studies analysed the progress of MT output through time (Adly & Al Ansary, 2009; Al-mahasees, 2022; Hadla, Hailat, & Al-Kabi, 2014; Hijazi, 2013; Jabak, 2019). However, Machine Translation Quality Assessment (MTQA) of legal texts from Arabic into English has not been studied before to the best knowledge of the researcher. Hence, the present study aims to explore this aspect of legal discourse and machine translation.

The significance of the study arises from the fact that the introduction of a corpus-based approach to MT is assumed to raise the percentage of accuracy of MT output since previously translated corpora are used as a reference. In addition, the study sheds light on the accuracy of MT output of legislative texts and its compliance with the features of legal discourse. To achieve the objective of the study, the research questions are as follows:

1- What are the types of errors found in the MT output of legal texts?
2- Does MT's output of legal texts maintain features of legal discourse?

The paper is divided into four parts. The first part presents a review of the literature; it gives a brief overview of features of legal discourse, machine translation, and previous related studies on MT quality assessment with an indication of to research gap. The second part outlines the study data and method of analysis. The third part presents an analysis of the study data. Finally, a discussion of the study findings in light of previous literature is addressed.

Literature Review

This section is divided into three parts. The first part discusses the features of English legal discourse such as lexical features, and syntactic features. The second part addresses the history and development of MT. It also tackles different approaches to MT quality assessment. Finally, the last part reviews relevant studies on MT quality assessment. The suggested contribution of the present study in light of previous studies is highlighted as a conclusion to this section.

Legal Discourse

Legal discourse, also known as 'legalese', is a specialized type of language that has distinctive features. Lexical features of English legal discourse include archaic adverbs such as *hereby, thereof, said, hereunder, aforesaid*, etc. Due to the influence of Latin and French on English legal discourse, several legal terms are borrowed from these two languages such as *bona fide, ipso facto*, etc. In addition, the use of the quantifier '*any'* is also found in legal texts; e.g., 'any interest and penalties' (Alfarahaty, 2014; Hijazi, 2013).

Several syntactic features are recognized in English legal discourse. Nominalization is when a noun is used to replace a verb. The purpose of using such a feature is to not indicate the tense, person, modality, agent, etc. In addition, the passive voice is very common in some legal texts such as statutes and laws. Wh-forms are usually deleted, and nouns are followed immediately by the past participle forms as in '*the notices given*'. Another syntactic feature that is highly detected in legal texts is the use of conditional clauses and restrictive connectors; for instance, *provided that, notwithstanding, subject to,* and others. The use of pronouns and demonstratives in legal discourse is avoided. Rather, words like *said, such, and the same* are used. Pronouns and demonstratives can also be avoided through the repetition of nouns. Furthermore, legal texts are characterized by long, complex sentences that contribute to clarifying the meaning and avoiding ambiguity and multiple interpretations (Alfarahaty, 2014; Hijazi, 2013).

One of the essential aspects of legal discourse is the use of models. The most frequently used models in legal discourse are *shall, may,* and *must* (Alfarahaty, 2014; Hijazi, 2013). Modals in legal discourse tend to have meanings different than general English. They are used to express obligation, prohibition and permission rather than futurity (Foley, 2002).

Machine Translation

The history of MT dates back to the 1950s. There are more than 20 approaches to machine translation that led to the fast development of MT such as Rule-Based Machine Translation (RBMT), Corpus-Based Machine Translation (CBMT), Data-Driven Machine Translation (DDMT), Example-Based Machine Translation (EBMT), Statistical Machine Translation (SMT), and Neural Machine Translation (NMT). Rule-Based Machine Translation (RBMT) is considered the first technique used in machine translation (Hutchins, 2000). RBMT is made of a collection of rules and lexical items in the form of dictionaries (Alqudsi, Omar, & Shaker; 2012). On the other hand, Corpus-Based Machine Translation (CBMT) is a technique that is based on the use of parallel corpora (Sin-wai, 2016). Statistical Machine Translation is based on statistics and probability techniques and is used by the most well-known machine translation engines globally; Google Translate, and Microsoft Bing (Sin-wai, 2016). However, Neural Machine Translation (NMT) is the most recent development in MT which proved to give better results compared to previous MT approaches. It is "an integration of neural language models into traditional statistical machine translation systems" (Almahasees, 2022, p. 23).

Arabic is considered one of the few languages that has been first researched with the introduction of MT in the USA in the 1950s (Sin-wai, 2016). Almahasees (2022) claimed that several reasons contributed to this fact such as cultural communication, and global commerce. The first Arabic machine translation company, Sakhr, was founded in 1982. Almahasees (2022) listed several issues that face machine translation of Arabic such as relatively free word order, complex morphology, subject embedding, diacritics, and punctuation.

Nevertheless, the speed of translation provided by MT engines does not guarantee the quality of the output (Adly & Al Ansary, 2009; Al-mahasees, 2022; Hadla et al, 2014; Hijazi, 2013; Jabak, 2019). Hence, researchers in translation technology proposed two approaches to assess the quality of MT output; i.e., human and automatic machine translation quality assessment (Castillho, Doherty, Gaspari, & Moorkens; 2018). Human evaluation of machine translation includes declarative evaluation which covers *fidelity* and *intelligibility*. Fidelity explores the extent to which the target text contains the information found in the source text. Intelligibility evaluates the well-formation of sentences in the target text. However, human evaluation of machine translation can be slow, expensive, and inconsistent (Way, 2018). Hence, automatic evaluation of machine translation is seen as an alternative due to its speed and objectiveness; yet, it is considered less comprehensive since it does not indicate the category of errors. Nevertheless, several automatic MTE metrics were proposed to compare the MT output with a reference translation (Castillho, et al, 2018).

Previous Studies

Several studies investigated the quality of Google Translate in translating between Arabic and English. Some researchers conducted a manual qualitative Translation Quality Assessment (TQA). Other studies followed an automatic evaluation by using certain software. The following review of the literature is divided into two parts. The first part discusses studies on the quality of MT output of legal texts in different language pairs; while the second part presents a summary of some studies that analyzed the quality of machine translation between English and Arabic.

Hijazi (2013) assessed Google Translate's performance in the translation of 14 legal texts from English into Arabic. Errors were classified into two categories: lexical and syntactic. Each category was further subdivided into subcategories. Lexical errors included polysemy, homonymy, legal doublets, and legal adverbs. On the other hand, syntactic errors included morphology, concord, and modality. The study concluded that the system did not succeed in providing a feasible and precise legal translation. Nevertheless, the researcher's overall assessment showed that Google Translate provided a gist translation that can assist readers in understanding the subject matter of the text in general. Another study that tackled legal translation with a different language pair is Killman's (2013) study. Killman (2013) examined the accuracy of Google Translate output of vocabulary items in legal texts from Spanish into English. The results of the data analysis showed that almost 70% of legal terms were translated correctly. A point to be taken is that the aforementioned studies were based on SMT architecture. The only study, as far as the researcher knows, that evaluated the MT output of legal texts after the introduction of NBMT is Wisemann (2019). The study compared the NMT output of two systems; i.e., Deepl Translator and MetaCat. The researcher followed a manual human evaluation approach and the study data consisted of Italian legal documents translated into German. Several error categories were detected such as ellipsis, mistranslation, translation of synonyms with the same word, terminology inconsistency, etc. In general, both NMT systems produced fluent translation; however, the results of the study showed that Deepl Translator performed better than MetaCat.

Al-Mahasees (2022) conducted a comparative study on three MT systems: Google Translate, Microsoft Bing, and Sakhr. Several texts were translated from English to Arabic and vice versa. The results revealed that Google Translate presented the best output of UN and WHO texts in terms of accuracy and fluency. However, in performance in translating literary texts from English into Arabic; Microsoft Bing achieved better.

Izwaini (2006) analyzed the performance of three MT systems; i.e., Google, Sakhr and Systran. He evaluated the problems faced in sentence translation. The study results showed that the output of the three MT engines was distorted and lacked cohesion and coherence. Al-Sukhni, Al-Kabi and Al-smadi (2016) compared the translation of Google Translate and Microsoft Translator of chapters from the Holy Quran into English. The output was assessed automatically by using ATEC. The study indicated that machine translation when compared to human translation, was ineffective in providing an accurate translation of Quranic texts.

Jabak (2019) assessed the quality and accuracy of Google Translate Arabic-English translations. The error analysis showed that Google Translate output contained lexical and syntactic errors. Lexical errors were the highest leading to faulty, inaccurate translations. The study concluded that the Google MT system cannot be used separately without any human post-translation interference.

Hadla, Hailat, and Al-Kabi (2014) compared Google Translate and Babylon MT systems. The study data included more than 1000 sentences distributed among the four sentence functions: declarative, interrogative, imperative, and exclamatory. BLUE was used for the automatic evaluation of the data. The results showed that Google Translate provided a more precise translation. A similar study conducted by Adly and Al Ansary (2009) explored the difference between three MT systems: Google Translate, Babylon, and Tarjim. They followed an automatic evaluation approach. The three MT systems failed to grasp the semantic cohesion and style of Arabic.

In conclusion, the previous review of the literature indicates that regardless of the accelerating advancements in MT, the MT output still contains several errors from different categories. MTQA, whether manual or automatic, highlighted a variety of error typologies; nevertheless, some MT systems provided better performance such as Google Translate and Deepl Translator. Previous work failed to address the NMT of legal texts as a very limited number of studies evaluated the quality of specialized translation. Furthermore, to the best of the researcher's knowledge, machine translation of legal texts from Arabic to English has never been studied before. Hence, the present study aims to evaluate the quality of NMT of legal texts from Arabic to English. The selected MT system is Google Translate since previous literature indicated its superiority over other MT systems.

Method

The study data is composed of five legislative documents. Five Saudi Laws were downloaded from the National Centre for Archives and Records. The researcher used Google Translate to translate all five documents. Google Translate is considered one of the most used MT systems in the world. Google Translate's approach to machine translation was statistical-based (Sin-wai, 2016). However, in 2017 Google Translate switched to Neural Machine Translation (Almahasees, 2022).

Table one lists the titles of the five Saudi Laws and the number of words per each.

Table 1. *Text titles and number of words*

Text Number	Name of the Law	Number of Words
Text 1	Saudi Post Corporation	1271
Text 2	Telecommunications Law	2457
Text 3	Anti-Cyber Crimes	1236
Text 4	Electronic Transactions	2932
Text 5	Postal Services	2493
Total		10389

The MT output was evaluated manually by the researcher. To ensure the reliability of the translation quality assessment (TQA), an existing human translation of the documents was used as a reference to achieve inter-rater reliability. Official translations of the five texts were retrieved from an official government website.

Findings

The errors found in the MT output can be classified as follows: (1) lexical errors, (2) syntactic errors, (3) omission, and (4) legal register-related errors. Lexical errors include the wrong choice of a word and mistranslation. Syntactic errors include inflectional errors and word order errors. Omission varies from the deletion of a word to the deletion of a whole clause. Errors related to legal discourse include the use of models, the use of pronouns, the use of deictic and archaic terms, and legal terminology. Table two (see appendix A) shows the percentages and frequencies of each error for each text and as a whole.

As seen in Table two (see appendix A), lexical errors came first as 43.5% of the errors found in all the five texts were lexical. The least detected error category is omission with a percentage of 1.7%. Legal register related errors in total came second with 155 errors found in the data. The highest occurring subcategory of legal register related errors is the use of pronouns and other anaphoric words constituting 11% of the MT errors, followed by legal terminology scoring 10% with 52 errors. The least occurring subcategory was the use of modals. The following part presents examples for each error category.

Lexical Errors

Lexical errors were the mainly wrong choice of a vocabulary item and mistranslations. Mistranslation errors affect the accuracy and intelligibility of the target text. The first article in all the five Laws starts with:

يُقصد **بالعِبارات** والمُصطلحات الآتية - أينما ورد في هذا النِظام ولائحتِه التنفيذية - المعاني المُبينة أمامها، ما لم يقتض السياق خِلاف ذلك:

and it was translated into:

The following terms and **expressions** - wherever mentioned in this Law and its Implementing Regulations - shall have the meanings indicated opposite them unless the context requires otherwise:

The use of the word (expression) is considered a mistranslation of the word (العبارات), and the adverb of place (أمام) was mistranslated as (opposite). In the following excerpt, the translation of the job title was inaccurate.

ولرئيس المؤسسة تفويض بعض هذه الصلاحيات إلى المسئولين في المؤسسة.

The **head** of the institution may delegate some of these powers to the officials in the institution.

The mischoice of lexical items was one of the highest recurring errors. The translation of (مجلس الادراة) was incorrect in some instances (*council*), though it was translated correctly at the beginning of the Law into (*board*).

لا يجوز لعضو **المجلس** تفويض شخص آخر للتصويت عنه عند غيابه، ولعضو المجلس المعترض تسجيل اعتراضه وأسباب الاعتراض ضمن محضر قرارات المجلس.

A **council** member may not authorize another person to vote on his behalf in his absence, and the objecting council member may record his objection and the reasons for the objection in the minutes of the council's decisions.

Finally, Google translate failed in translating near-synonyms. In the example below, the same lexical item (necessary) was used twice to translate synonyms.

للمُشغِلين دخول العقارات واستِخدامِها في الحدود **الضرورية واللازِمة**
The operators may enter the real estate and use it within the **necessary and necessary** limits

Legal Register-related Errors

30.2% of the errors were legal register-related errors. The errors related to the legal discourse were subclassified as follows: (1) the use of pronouns, demonstratives, and anaphora; (2) legal terminology, (3) archaic terms, and (4) modals. The highest occurring subcategory was the use of pronouns and other anaphoric words constituting 11% of the MT errors. Legal terminology scored 10% with 52 errors. The least occurring subcategory was the use of modals.

Anaphoric Devices

The MT output failed in replacing pronouns and demonstratives with other devices that are used in legal discourse. In the following example, the 3rd person pronoun ***their*** was used instead of other reference words that are mainly used in legal discourse such as ***said***, ***such***, ***same***, etc.

يتولى المفتشون مجتمعين أو منفردين ضبط مخالفات أحكام النظام واللائحة، والتحقيق فيها. وتحدد اللائحة قواعد وإجراءات **عملهم**.
Inspectors, jointly or severally ,shall control and investigate violations of the provisions of the Law and Regulations. The regulation defines the rules and procedures for **their** work.

The human reference translation used the word ***said***

Inspectors shall individually or collectively detect and record violations of this Law and its Regulations. The Regulations shall determine the work rules and procedures for **said** inspectors.

In the example below, the demonstrative ***these*** was used; however, the reference translation used the word ***such*** instead.

للمحكمة المختصة أن تعفي من ***هذه*** العقوبات كل من يبادر من الجناة بإبلاغ السلطة المختصة بالجريمة قبل العلم بها وقبل وقوع الضرر

The competent court may exempt from ***these*** penalties whoever among the offenders initiates to inform the competent authority of the crime before knowledge of it and before the occurrence of the damage

The competent court may exempt any offender from ***such*** punishments if he informs the competent authority of the crime prior to its discovery and prior to the infliction of damage

Legal Terminology

The data analysis showed that Google Translate could not perfectly translate all the legal terms found in the data. In some instances, a similar legal term; yet inaccurate, was used such as *System* instead of *Law* for (*نظام*), and *Regulation* instead of *Statute* for (*تنظيم*). In other cases, a plain English term was used such as ***an exception to*** instead of ***notwithstanding*** as seen in the example below.

استثناء من الفقرة (1) من هذه المادة

As an exception to Paragraph (1) of this Article

Notwithstanding paragraph (1) of this Article

In the following example, the verbs يحل and يلغي were translated into ***replaces*** and ***cancels*** respectively.

يحل النظام محل نظام البريد، الصادر بالمرسوم الملكي رقم (م / 4) وتاريخ 21 / 2 / 1406هـ، **ويلغي** كل ما يتعارض **معه** من أحكام.
The **system replaces** the postal system, issued by Royal Decree No. (M / 4) dated 2/21/1406 AH, and **cancels** all provisions that conflict **with it.**

In the human reference translation, ***supersede*** and ***repeal*** were used.

This Law shall ***supersede*** the Law of Postal Services issued pursuant to Royal Decree No. (M/4), dated 21/2/1406H, and shall ***repeal*** any provisions conflicting therewith.

An English translation of a legal expression (حسن النية) occurred twice in the data. In text (3) a literal translation was given instead of the Latin expression *bona fide*. However, in text (4), the Latin term was used.

مع عدم الإخلال بحقوق **حسني النية،** يجوز الحكم بمصادرة الأجهزة أو البرامج أو الوسائل المستخدمة في ارتكاب أي من الجرائم المنصوص عليها في هذا النظام.

Without prejudice to the rights of **good faith**, it is permissible to order the confiscation of devices, programs, or means used in committing any of the crimes stipulated in this Law

Archaic Adverbs

Archaic terms were barely used; instead, prepositions were followed by pronouns in 24 cases. In the following example, the MT output used the pronoun ***on*** followed by ***it***.

أن التعامل الإلكتروني لم يطرأ عليه تغيير منذ وضع التوقيع الإلكتروني **عليه**.

The electronic transaction has not changed since the electronic signature was placed ***on it.***

The human reference translation used instead the adverb *thereto*.

The electronic transaction has not been altered since the electronic signature was affixed ***thereto***.

Modals

The use of modals was mainly accurate. In translating articles that indicate obligations, the modal ***shall*** was successfully used with the infinitive form of the verb in translating the verb in the present tense in the majority of cases. Yet, this translation was not consistent throughout the whole data. 24 cases that required the use of modals were translated wrongly.

تُحدِّد الهيئة في التراخيص التي تُصدِرُها للمُشغلين مُتطلبات إصدار الفواتير وخدمات الاستعلامات وخدمات الطوارئ
In the licenses it issues to operators, the Authority **specifies** the requirements for billing, information services, and emergency services

Syntactic Errors

24.8% of the errors detected in the study data are syntactic constituting one-fourth of the errors. Syntactic errors were mainly related to number, definiteness, voice, part of speech, and word order. Such errors affected the intelligibility of the MT output.

وعلى المؤسسة أن تضع في مكان واضح للجمهور في المكاتب البريدية التابعة لها مقدار المقابل المالي المستحق لها على الخدمات البريدية التي تقدمها.

The institution shall put in a clear place for the public in its post offices the amount of the financial consideration due to it for the postal services it provides.

The human reference translation reads easily compared to the MT output.

Saudi Post shall display the postal service fees in a conspicuous place in its post offices

Omission

Al-Mahasees (2022) defines omission as "the deletion of an item that should appear in a well-formed and meaningful sentence" (p. 83). The omission was the least detected category of errors. Deletion includes the omission of a word, a few words, or a whole clause. Only eight cases (four words, and four clauses) were detected in the study data. The deletion of the words and clauses affected the intelligibility of the target text. However, other related error types such as addition and were not recorded in the data.

In summary, lexical errors constituted 43.5% of the errors found, followed by register-related errors with 155 errors found in the data. The highest occurring subcategory of legal register-related errors is the use of pronouns and other anaphoric words followed by legal terminology and modals. The least recorded error category is omission.

Discussion

The research's main aim is to evaluate the quality of MT output of legal texts from Arabic to English. Two research questions were proposed.

1. What are the types of errors found in the MT output of legal texts?
2. Does MT's output of legal texts maintain features of legal discourse?

The findings of the study showed that Google Translate provided a good translation; nevertheless, several error types were found.

To answer the first question, errors were classified and calculated. The data analysis revealed four error types: lexical, syntactic, register-related and omission. 512 errors out of 10389 words were found in the data constituting 5% of the data. Significantly, the data analysis indicated that lexical errors constituted almost half of the errors. Lexical errors were the highest in the study data, which is consistent with Jabak's (2019) study. the least reported error type was omission with only 8 cases.

As for the second question, the data analysis indicated that legal discourse errors came second in the study sample. The inaccurate use of modals can be attributed to the differences between Arabic and English. Verbs in Arabic can carry modality, but English verbs require other helping verbs to express modality. The same can be applied to archaic terms. The prepositional phrases in Arabic were translated instead of using archaic adverbs due to their different morphology in English.

Hence, it can be implied from the findings of the study that Google Translate failed in presenting a legal document that follows the norms of English, which comes in great agreement with Hijazi's (2013) study. In contradiction with Killman's (2013) study, legal terms in the present study were poorly translated. Al-Sukhni et al. (2016) reported that Google Translate could not translate Islamic texts efficiently; supporting the findings of the present study that Google Translate, when compared to human translation, could not provide an accurate specialized translation; i.e., legal translation.

Conclusion

The present research objective is to study the quality of Neural Machine Translation of legal texts from Arabic into English. The researcher followed a manual evaluation to assess the quality of NMT output. The study's contribution is twofold: the evaluation of the NMT architecture of Google Translate in general; and the diagnosis of legal discourse features of English NMT output in specific. The results indicated that the introduction of NMT to Google Translate has not succeeded in solving common issues related to errors in MT output. Though Google Translate provided a comprehensible output, lexical and syntactic errors were seen in the data. The paper also suggested that Google Translate failed in translating several factors of legal discourse.

About the Author

Dr. Eman Alkatheery is an assistant professor at the College of Language Studies, King Saud University. Her research interests are linguistics, translation, and quality of higher education institutions. She has a 15-year experience in teaching undergraduate and postgraduate courses in linguistics, language skills, and translation. She is a certified trainer and presented several training courses and workshops in the fields of English language, translation technology, and quality and academic accreditation.ORCID ID : 0000-0003-1920-2668

References

Adly, N., & Al Ansary, S. (2009). Evaluation of Arabic machine translation system based on the universal networking language. In *International Conference on Application of Natural Language to Information Systems* (pp. 243–257), Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-12550-8_20

Almahasees, Z. (2021). *Analysing English-Arabic Machine Translation: Google Translate, Microsoft Translator and Sakhr*. Routledge. https://doi.org/10.4324/9781003191018

Alqudsi, A., Omar, N., & Shaker, K. (2012). Arabic machine translation: A survey. *Artificial Intelligence Review*, *42*(4), 549–572.Almahasees, Z. (2021). Analysing English-Arabic Machine Translation: Google Translate, Microsoft Translator and Sakhr. Routledge. https://doi.org/10.4324/9781003191018

Al-Sukhni, E., Al-Kabi, M. N., & Alsmadi, I. M. (2016). An automatic evaluation for online machine translation: Holy Quran case study. *International Journal of Advanced Computer Science and Applications*, *7*(6), 118–123. DOI: 10.14569/IJACSA.2016.070614

Castilho, S., Doherty, S., Gaspari, F., Moorkens, J. (2018). Approaches to Human and Machine Translation Quality Assessment. In J. Moorkens, S. Castilho, F. Gaspari, & S. Doherty, (eds.), *Translation Quality Assessment. Machine Translation: Technologies and Applications*, *1*, 9-38. Springer, Cham. https://doi.org/10.1007/978-3-319-91241-7_2

El-Farahaty, H. (2014). *Arabic-English-Arabic Legal Translation*. Routledge. https://doi.org/10.4324/9781315745893
Foley, R. (2002). Legislative Language in the EU: the Crucible'. *International Journal for the Semiotics of Law*, *15*(4), 361–374. https://doi.org/10.1023/A:1021203529151
Hijazi, B. (2013). *Assessment of Google's Translation of Legal Texts,* (Unpublished Master's Thesis). University of Petra, Amman, Jordan.
Hadla, L., Hailat, T., & Al-Kabi, M. (2014). Evaluating Arabic to English Machine Translation. *International Journal of Advanced Computer Science and Applications (IJACSA)*, *5*(11), 68 – 73. DOI: 10.14569/IJACSA.2014.051112#sthash.NW2l6vl5.dpuf
Hutchins, W. J. (2000) *Early Years in Machine Translation,* Amsterdam and Philadelphia: John Benjamins Publishing Company.
Izwaini, S. (2006). *Problems of Arabic machine translation: Evaluation of three systems*. Retrieved from www.mt-archive.info/BCS-2006-Izwaini.pdf
Jabak, O. (2019). Assessment of Arabic-English translation produced by Google Translate. *International Journal of Linguistics, Literature and Translation*, *2*(4), 238-247.
Killman, J. (2014). Vocabulary accuracy of statistical machine translation in the legal context. In O,Brien, S., Simard, M., & Specia, L (Eds.), *Third Workshop on Post-Editing Technology and Practice (WPTP-3), The 11th Conference of the Association for Machine Translation in the Americas* (pp. 22-26).Vancouver: AMTA. https://www.amtaweb.org/AMTA2014Proceedings/AMTA2014Proceedings_PEWorkshop_final.pdf
Sin-wai, C. (2016). The Future of Translation Technology: Towards a World without Babel (1st ed.). Routledge. https://doi.org/10.4324/9781315731865
Way, A. (2018). Quality Expectations of Machine Translation. In J. Moorkens, S. Castilho, F. Gaspari, & S. Doherty, (eds.), *Translation Quality Assessment. Machine Translation: Technologies and Applications*, *1*,159-178. Springer, Cham. https://doi.org/10.1007/978-3-319-91241-7_8
Wisemann, E. (2019). Machine Translation in the Field of Law: a Study of the Translation of Italian Legal Texts into German. *Comparative Legilinguistics*, *37*(4), 117-153. *http://dx.doi.org/10.14746/cl.2019.37.4*

Appendix A

Table 2. *Frequencies and percentages of error categories*

Error category	Text 1		Text 2		Text 3		Text 4		Text 5		Total	%
	No.	%	No.	%	No.	%	No.	%	No.	%		
Lexical errors	23	32%	61	46%	33	44.4%	42	38.9%	64	50.4%	223	43.4%
Syntactic errors	27	37.5%	26	19.5%	24	33.3%	29	26.9%	21	16.5%	127	24.8%
Deletion	3	4.2%	5	3%	-	1.4%	-	-	-	-	8	1.7

Legal discourse errors												30.2%
Modals	2	2.8%	8	6%	-	-	6	5.6%	7	5.5%	23	4.5%
Legal terminology	9	12.5%	11	8.3%	6	8.3%	10	9.3%	16	12.6%	52	10%
Archaic terms	-	-	8	6%	2	2.8%	7	6.5%	7	5.5%	24	4.7%
Pronouns	8	11%	15	11.3%	7	9.7%	14	13%	12	9.4%	56	11%
Total	72	-	134	-	72	-	108	-	127	-	513	-