The Accuracy and Consistency of Machine Translation Engines: Case Study of Translating Economic & Statistical Texts from English into Arabic

Monther Battah 1 *, Suleiman Ibrahim Shelash Mohammad 2,3,4, Muslim Momani 4, Asokan Vasudevan 5, Muhammad Turki Alshurideh 1

1Department of Human Basic Sciences, Husun University College, Al-Balqa Applied University, Jordan
2Research follower, INTI International University, 71800 Negeri Sembilan, Malaysia
3Department of Business Administration, Business School, Al al-Bayt University, Jordan
4Department of Marketing, School of Business, The University of Jordan, Amman, Jordan
5Faculty of Business and Communications, INTI International University, 71800 Negeri Sembilan, Malaysia

Abstract

Objectives: This study assesses the precision and consistency of the Google Translate engine and tool in translating statistical and economic texts from English into Arabic.

Methods: The descriptive analytical approach was employed in this project, and the study sample consists of five texts extracted from statistical and economic research.

Results: The study revealed that the Google translation tool faced numerous challenges and contained lexical, grammatical, and consistency errors. Besides, a grand challenge might arise from the idioms and abbreviations, as Google Translate would use word-for-word translation and render them inaccurate meaning. In some cases, the abbreviations in the source text (ST) would be displayed as they are in the target text (TT), while in other cases, Google Translate would do a deep search to render their meanings in the TT. Further, Google Translate has a significant problem rendering technical and specialized terms into the TT, and in some cases, it doesn’t consider the contexts and has a lexical choice challenge. Finally, Google Translate renders some terms with different meanings within the same text.

Conclusions: Idioms and abbreviations present a grand challenge, as Google Translate would use word-for-word translation and render them in an inaccurate meaning. In some cases, the abbreviations in the ST are rendered as they are in the TT, and in other cases, Google Translate would do a deep search to render their meanings in the TT.

Keywords: Google Translate, Consistency, Accuracy, Coherence and Cohesion.

Keywords: Google Translate, Consistency, Accuracy, Coherence and Cohesion.

* Corresponding author:
  m.battah@bau.edu.jo


© 2024 DSR Publishers/The University of Jordan.

This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY-NC) license https://creativecommons.org/licenses/by-nc/4.0/
The accuracy and consistency of Machine …

**Introduction**

In the past, the processes of translating documents from one language into another were done by humans who were familiar with the source and the target languages. The traditional translation process encountered several constraints including the costs, time, and accuracy of the products.

Several studies have investigated the challenges, accuracy, and consistency of the traditional paper-based translated products in different languages. Some studies revealed that translators might encounter different challenges during the translation process, most of which are turn-around familiarity. Translators might have limited familiarity with the topic, vocabulary, structure, grammar, and abbreviations. Most of these studies revealed that lack of knowledge of “Familiarity”, whether with a topic or the technical terminology used in the ST field is the primary constraint that could lead to mistranslation or producing inaccurate TT (Motlaq&Mahadi, 2020).

Summing up the most famous attempts to interpret the meaning of “Translation”, it is noted that most attempts are turning around rendering and selecting the most accurate and closest equivalence in the TT that reflects the accurate meaning of the ST term and suits the context. Meeting this condition, we will move to the “Good Translation” or “Best Translation” definitions. Several scholars developed the traditional definitions of “Translation Process” and added a new criterion, which is producing a good translation that shows the accuracy and consistency of the ST terms.

Several attempts were made to provide evidence of the importance of consistency in producing a coherent and accurate TT and to develop the language level used in the TT. Some studies investigated the accuracy and consistency of the existing translation attempts of the Holy Quran into English depending on the Quran Interpretations. Other studies focused on modern Computer Assisted Translation Tools (CAT Tools). Hence, all World Wide Web sites and visitors use the machine translation tool to localize the content into the clients’ languages (Wenget al., 2023). Therefore, some studies stressed the importance of building a new Neurological Matrix that could offer accurate and consistent products.

This study aims to investigate the accuracy and consistency of machine translation to render economic & statistical texts from English into Arabic.

**Statement of the Problem**

Various organizations and individuals are using machine translation tools to translate different types of texts from one language into another within a short time, an effortless, and costless process. Google Translate Engine, provided by Google Company, is available on browsing engines, such as Google Chrome, Mozilla Firefox, Opera, and Safari.

Browsing a foreign website, visitors might need the easiest method for changing the content to their mother tongue in no time. Google Translate tool appears on the top-right corner and can deliver this service for free.

Besides, several agencies and translators tend to use Google Translate services to render different types of texts, and they believe it is a reliable tool to meet their needs. However, several studies were conducted to illustrate the gap in consistency and accuracy of the TT produced by the Google Translate engine.

This study seeks to evaluate the level of accuracy and consistency of the Google Translate engine and tool in rendering statistical and economic texts from English into Arabic.

**Methodology**

The study depends on the descriptive analytical methodology to meet the study objectives. The first process is to identify and describe the phenomenon, and then the analytical procedure will be conducted to measure the level of accuracy and consistency in the economic and statistical TT produced by Google Translate and Google Tool.

The analysis will be done by Microsoft Office Excell after computing the number of instances in which accuracy and consistency problems were identified.

**Study Sample**

The sample was 10 short statistical and economic paragraphs in English.

**Corpus**

The corpus was extracted from economic and statistical texts. The abbreviations were excluded from the study corpus. The following sentences represent the study corpus:
Questions of the study

- What is the level of accuracy in the TT statistical and economic texts produced by Google Engine and Google Tool?
- What is the level of consistency in the TT statistical and economic texts produced by Google Engine and Google Tool?
- What are the recommended development procedures and add-ins that would foster the level of accuracy and consistency of Google Engine and Google Tool?

Study Objectives

- Identify the level of accuracy in the TT statistical and economic texts produced by Google Engine and Google Tool.
- Identify the level of consistency in the TT statistical and economic texts produced by Google Engine and Google Tool.
- Propose recommended development procedures and add-ins that would foster the level of accuracy and consistency of Google Engine and Google Tool.

Study Importance

- Theoretical Importance
  This study will enrich the translation studies literature and enlighten the way for researchers interested in machine translation accuracy, consistency, and development. This paper might be a small distribution to develop the machine translation engines and facilitate the researchers’ missions to reach a sustainable and reliable machine translation tool.

- Practical Importance
  Investigating the level of accuracy and consistency of machine translation and implementing the proposed amendments will significantly affect the quality of the TT produced by translation engines used by billions of users. This paper attempts to provide practical recommendations that would increase the quality of the translation products produced by machine translation engines in terms of consistency and accuracy of the TT.

Theoretical Background and Literature Review

Theoretical Background

History of Machine Translation

The history is full of experience and trials to develop reliable Machine translation (MT) systems. Several scholars from different countries have contributed to developing the current MT engines over the last decades; however, the idea of MT took shape in the 19th century simultaneously with the emergence of computing devices. The 1950s witnessed the first attempt to develop MT through a dictionary that involved 250 words and 6 grammatical rules. However, this corpus and dataset were not a firm base of global MT; therefore, several countries started to develop their own MT systems (Melby, 2019).

In 1966, the “Automatic Language Processing Advisory Committee (ALPAC)” reported that MT-funded research had decreased significantly. In this report, ALPAC highlighted several challenges of developing a reliable MT system including the high cost, lack of accuracy, and the clumsiness compared to full human made translations. New approaches were proposed to solve these issues (ALPAC, 1966). In the late 1970s, scholars proposed the “Rule-based approach”, which is a less complicated approach compared to the “interlingua approach”, as well as the rule-based approach is suitable while working with different languages (Tuet al, 2018).

The 3rd generation of MT systems was in the 1990s. MT systems were developed to work based on example and statistical approaches, and these approaches were the basis of the MT Memory Tool (MTMT). In the 2000s, the internet was the main source of developing MT as it offered rich opportunities to attract translation memories (TM) to develop fully automated written and oral MT systems. For instance, Japan developed an MT system for a speech “Speech to Speech” system that operates on mobile phones. Furthermore, Google Company launched the Google Translate Tool in 2016 which works based on a Neural Network Approach (NNA). Besides, the modern development in MT involved neural networks.
and deep learning which sought to improve the accuracy and consistency of MT (Johnson et al., 2017).

**MT Process Advantages and Challenges**

MT systems have several advantages including low cost, multitasking, speed, and privacy. MT allows rendering different kinds of documents in one click. Besides, depending on a human translator is expensive and takes a long time, while MT provides translation services at no cost and time. Furthermore, professional translators are specialized in specific majors and specializations. However, the MT system is a multitasking system and a jack of all trades as it can render texts from different specializations. A critical issue that all individuals and organizations are focusing on is the privacy and confidentiality of their documents. For example, private emails and financial documents. MT systems ensure the privacy of the users’ data. On the other hand, MT systems encounter several challenges including lexical choice, grammatical challenges, and pronoun references. The lexical choice challenge could arise when a term has multiple meanings; therefore, the machine translation engine in this case needs to identify the closest and most suitable meaning. Besides, MT might encounter the challenge of rendering sentences that include multiple grammatical structures. In addition, pronoun references are a serious challenge for MT especially between languages that have joint pronouns references (Motlaq&Mahadi, 2020; Koehn & Knowles, 2017).

There are several MT tools available such as Google Translate, Amazon Translate, Microsoft Translate, DeepL, and Baidu Translate tool. The machine translation process passes through several stages as follows:

Figure 1: A typical MT process (Source: Okpor (2014)).

**Consistency and Accuracy of MT**

According to Thelen (2008), consistency and accuracy are the primary references and indicators of translation quality and a condition for semantic univocity and certainty regarding legal terms translated at international organizations. Besides, translation consistency is “a feature designed to provide translators with informative hints regarding how the source words were previously translated. These hints show previous translations for individual words and the number indicating how often each translation was used in the project”.

596
Discussion

Accuracy and Consistency of MT

• Example (1):

Discussion:

The abbreviation (ISIC) refers to the International Standard Industrial Classification which consists of a 4-digit classification. ISIC 1811 refers to "Printing" which means all processes of printing, copying, scanning…etc.

The first thing, the Google translate engine rendered “ISIC 1811” into “الطباعة 1811”, which means that a literal translation strategy was employed, and the translation was not accurate nor comprehensive. In general, this attempt was not successful and did not reflect the accurate scientific meaning. Besides, MT always tends to keep the abbreviation as it is in case this abbreviation is not stored in the Translation memory.

Instead, the most accurate and accurate meaning for “ISIC 1811” would be “التصنيف الدولي الصناعي المعياري لأعمال الطباعة”.

Moving to the second segment “Assume National Accounts confirm the lack of price change!”. Google tool rendered this segment into “تفترض الحسابات القومية تأكيد عدم وجود تغير في الأسعار”.

The Google tool was not successful in rendering the most accurate meaning, as it reflects the literal meaning (Word for word) in Arabic. The intended meaning is to let us assume that National Accounts does not have and confirm that price change was not captured. Therefore, the accurate translation would be “على فرض أن الحسابات القومية قد أكدّت أنها لا تمتلك بيانات حول وجود حالة من التغير / التقلب في الأسعار”.

For this example, Google Tools was not able to render the accurate meaning of the ST.

The 3rd segment “Covers 1811, 1812, and 1820 combined – review coverage – if all activities are important, they should be covered separately, if not they should be removed.”

Google Translate:

إذا كانت جميع الأنشطة مهمة فيجب تغطيتها بشكل منفصل، وإذا لم تكن كذلك فيجب إزالتها.

Google tool started its mission in rendering the ST to the TL using Word for Word translation strategy. In the beginning, the sentence’s subject which is “ISIC”, is absent. Besides, ISIC 1811 covers three commercial activities, and their codes are “1811-1812-1820”. However, Google Translate Tool rendered this segment into “1820 و 1812 و 1811 يغطي الأعوام الثلاثة”，which is entirely wrong. Moreover, “Review Coverage” indicates that the statisticians working at the national accounts department shall review the ISIC 1811 Coverage and the activities involved under this code. However, Google Translate Tool rendered it into “تغطية المراجعة”，which doesn’t reflect the accurate meaning.

The 4th segment: Includes 5 items/businesses – 4 newspapers (2-part government owned) and one printing cardboard carton.

MT:

تتضمن 5 عناصر/أعمال تجارية - 4 صحف (جزأين مملوكين للحكومة) وواحدة لطباعة علب الكرتون.

Likewise, Word for Word translation strategy was applied in this segment. The subject is absent, and the sentence is not clear in the TT. “4 newspapers (2-part government owned)” means that ISIC 1811 involves four newspapers as two of them are owned by the government. Google translate tool rendered it into “4 صحف (جزأين مملوكين للحكومة)”. The accurate translation is “وتحتوي النتيجة أربعة صحف تملك الحكومة الاثنين منها”.

• Example (2): Dairy (ISIC 1050)

Need to understand National Accounts concerns – but the index shows little change.

Index seems to cover main establishments and items.

Expect relatively little change to items – very little innovation in dairy products.

Government price limits may affect prices – need to understand better how this is implemented in practice – are there subsidies?

Also, if only raw milk prices are regulated, is there more volatility in prices for processed milk products?
The accuracy and consistency of Machine …

Monther Battah et al.

MT:

الألبان (ISIC 1050)
نحن بحاجة إلى فهم المخاوف المتعلقة بالحسابات القومية – لكن المؤشر يظهر تغييرًا طفيفًا
ببدو أن الفهرس يغطي المنشآت والعناصر الرئيسية
توقع تغييرًا طفيفًا نسبيًا في العناصر - أقليل جدًا من الأدوات في منتجات الألبان
قد تؤثر حدود الأسعار الحكومية على الأسعار - بحاجة إلى فهم أفضل لدى كيفية التقييم ذلك عالميًا - هل هناك إعانات؟
وأيضًا، إذا تم تنظيم أسعار الحليب الخام فقط، فهل سيكون هناك المزيد من التقلبات في أسعار منتجات الألبان المصنعة؟

Rendering the phrase “Need to understand” into ”نحن بحاجة لفهم” in this case, Google engine has added a subject to illustrate the sentence “نحن”. Besides, the Google engine rendered the term “Index” into “المؤشر”, which is correct, however, in the next segment google engine rendered the same term into “الفهرس”. This is a clear example of the problems that we might encounter while translating using the Google Translate engine which is consistency and accuracy. Google Translate engine can’t classify the text type to select the most appropriate and accurate meaning in the TT that suits the ST and the context. Besides, one single word might be rendered into different meanings within the same text. Therefore, rendering economic or statistical texts using the Google Translate engine would be not successful 100% as it lacks consistency and accuracy.

Rendering the sentence “Government price limits may affect prices” into Arabic was not clear, as the TT did not cover the accurate meaning of the ST. Google translate engine rendered it into “قد تؤثر حدود الأسعار الحكومية على الأسعار”, the MT did not cover the intended meaning as the ST indicates that price limits presented by the government’s mean party may affect the market prices of dairy products as the price limitations given by the prices control department would limit traders’ profit margins. The Google Translate engine used the word-for-word translation strategy, but it did not successfully render the ST into the closest and most accurate TT.

- Example (3):
When we start index development, need to think about what we will measure – mainly linked to intended use of the index.

For example: Input or output index
Activities or products covered.
Valuation basis
Primary use of PPI – deflator for national accounts. Also, economic policy (eg Central Bank) or contract escalation by businesses.

MT:

عندما نبدأ في تطوير المؤشر، نحتاج إلى التفكير فيما سنقوم به - والذي يرتبط بشكل أساسي بالاستخدام المقصود للمؤشر
على سبيل المثال:
مؤشر الإدخال أو الإخراج
الأنشطة أو المنتجات المغطاة
أساس التقييم
الاستخدام الأساسي لمؤشر أسعار المنتجين – معامل الامكانيات للحسابات القومية. أيضا السياسة الاقتصادية (مثل البنك المركزي) أو تصعيد العقد من قبل الشركات.

Google translate engine in this example used the Word for Word translation strategy, and it tried to maintain the ST style and wording. However, the Google Translate engine rendered the segment “Input or Output Index” into “مؤشر الإدخال أو الإخراج”. This reflection is not accurate nor consistent with the context. The accurate meaning is “مؤشر المدخلات أو المخرجات" which refers to statistical and economic indicators that reflect the statistical figures computed in the index. Therefore, it is noted that the Google translate engine has not rendered the accurate technical meaning of the ST. Besides, the “Valuation Basis” refers to the basis to make values and weights for the input; therefore, the MT rendered it into “أساس القيم" which is inaccurate as in this process the statistician is not making an evaluation, but rather making values and weights.
Example (4):
Know that we want:
Output PPI
Valuation basis = basic prices
= price received by producer excluding taxes and transport (where invoiced separately and including subsidies received).
Cover at least mining (ISIC section B), manufacturing (ISIC section C), electricity (ISIC section D).
Need to decide:
Coverage of activities to 4-digit ISIC (exclude unimportant activities)
User need for product-level (CPC) indices and coverage of service activities.
Need for separate domestic and export price indices.

Example (5):
The main objectives of the survey:
Calculating averages of income and expenditure and the relationship of different economic and social factors with them, such as occupation and educational level of the head of the household.
Study the distribution of households and individuals according to income and expenditure categories and analyze the factors associated with it.
Provide data on income and expenditure for the purpose of calculating poverty indicators, identifying characteristics of the poor, and preparing poverty maps.
Provide weights data (scales) that reflect the relative importance of consumer expenditures items and that are used in preparing the consumer price index.
Provide a part of the data needed for the national accounts related to final consumption and income of the household’s sector.

MT:
The accuracy and consistency of Machine …

Monther Battah et al.

Google translate rendered the term “Average” into "متوسط". However, the term average is different from the term “Mean”, as average = معدل and mean = متوسط. Besides, Google Translate doesn’t maintain the coherence and cohesion of the TT, as we can notice that punctuations are not used correctly in the TT.

Google Translate rendered the term “categories” into “فئات”, however, this meaning is not used in the statistical nor economic language as the accurate meaning would be “شرائح”.

Google Translate adopted word for word translation strategy, which was successful to a specific limit to render the accurate meaning of the ST. However, functional translation strategy would extract the most accurate and most natural meaning of the ST as using this strategy the TT would be "توفير بيانات حول الانفاق والدخل بهدف احتساب مؤشرات الفقر والتي من خلالها يمكن تحديد خصائص شريحة الطبقة الفقيرة وتحضير خرائط مكانية لأماكن تواجد الفقراء."

Conclusion and Recommendations

Summary

- Google translate engine adopts the literal translation strategy, which renders the literal meaning, but in several cases, it does not have accurate meaning. Besides, Google Translate always tends to keep the abbreviation as it is in case this abbreviation is not stored in the Translation memory. However, Google Translate would render the same abbreviation as it is in some cases and translate it in others. Therefore, Google Translate doesn’t maintain the consistency of the produced TT.
- In some cases, Google Translate renders the 4-digit figures as they are years. For example, “Covers 1811, 1812 and 1820 combined – review coverage – if all activities are important, they should be covered separately, if not they should be removed.”
- Google translate rendered the figures (1811, 1812, and 1820) as they are years. The Google translate rendered the figures into "업데이트 1811, 1812, و 1820 مجتمعة - تغطية المراجعة - إذا كانت جميع الأنشطة مهمة يجب تغطيتها بشكل منفصل، وإذا لم تكن كذلك فيجب إزالتها." which is completely wrong.
- Google tool started its mission in rendering the ST to the TL using Word for Word translation strategy. In the beginning, the subject of the sentence is absent, which is “ISIC”. Besides, ISIC 1811 covers three commercial activities, and their codes are "1811-1812-1820". However, Google Translate Tool rendered this segment into "العوام 1811 و 1812 و 1820"، which is completely wrong.
- Google engine rendered the term “Index” into “المؤشر”，which is correct, however, in the next segment google engine rendered the same term into “الفهرس". This is a clear example of the problems that we might encounter while translating using the Google Translate engine which is consistency and accuracy.
- In some cases, Google Translate did not cover the intended meaning as the ST indicates that price limits presented by the government’s middle party may affect the market prices of dairy products as the price limitations given by the prices control department would limit the traders’ profit margins.
- Google translate rendered the term “Average” into “متوسط". However, the term average is different from the term “Mean”, as average = معدل and mean = متوسط. Besides, Google Translate doesn’t maintain the coherence and cohesion of the TT, as we can notice that punctuations are not used correctly in the TT.
- Google Translate rendered the term “categories” into “فئات". however, this meaning is not used in the statistical nor economic language as the accurate meaning would be “شرائح”.

Conclusion

Google Translate is a modern machine learning-based tool that renders texts within most of the world's languages. This tool passed through several developments to reach the most accurate and closest TT that would serve the customers’ needs of translating documents and texts from one language to another. However, this translation tool encounters several challenges and the TT would have several mistakes including lexical, grammatical, and consistency mistakes.

Most of the time, idioms and abbreviations present a grand challenge, as Google Translate would use word-for-word
translation and render them in an inaccurate meaning. In some cases, the abbreviations in the ST are rendered as they are in the TT, and in other cases, Google Translate would do a deep search to render their meanings in the TT.

Regarding accuracy, Google Translate has a significant problem in rendering technical and specialized terms into the TT, and in some cases, Google Translate doesn’t consider the contexts and has lexical choice challenges.

In terms of consistency, Google Translate renders some terms in different meanings within the same text. For example, Google Translate renders the term “Index” into “المؤشر” and in other cases into “الفهرس”, which means that the Google translate tool lacks to consistency, and the produced TT needs proofreading to unify the text.

REFERENCES


