

Evaluation of Facebook Translation Service (FTS) in Translating Facebook Posts from English into Arabic in Terms of TAUS Adequacy and Fluency during Covid-19

Almahasees Zakaryia^{*1}, Al-Taher Mohammad¹, Helene Jaccopard²

¹Applied Science Private University, Department of English Language and Translation, Faculty of Arts and Science, Amman, 00962, Jordan

²University of Western Australia, Department of Translation, School of Humanities, Perth, 0061, Australia

ARTICLE INFO

Article history:

Received: 08 December, 2020

Accepted: 30 January, 2021

Online: 25 February, 2021

Keywords

Facebook Translation service (FTS)

Facebook and COVID-19

Facebook adequacy and Fluency

Translation during COVID-19

English and Arabic

ABSTRACT

The study aims to verify the capacity of Facebook Translation Service in translating English Facebook posts into Arabic in terms of two criteria: adequacy and fluency in line with the Translation Automation User Society (TAUS) scales. To ensure consistency and objectivity as recommended by TAUS, six evaluators, native speakers of Arabic and near-native speakers of English, rated the same data on each scale. The evaluators were acquainted with fluency and adequacy scales along with MT limitations and potentials. Once the corpus was uploaded and sent to the evaluators using TAUS tools, they had to assign scores online on 1-4 rating scales. Then, each report was displayed online on the TAUS reports tool. Evaluators' responses were combined in thematic categories and were calculated to obtain frequencies and percentages. The study found that FTS provided fluent output with highest percentage of the scale good equal to 3 on a scale from 1 to 4, where the output is assessed as flowing smoothly with minor linguistic errors. Moreover, FTS succeeded in generating an adequate output with the highest percentage of responses as 'most' equal to 3 on a scale from 1 to 4, where almost the full meaning of the source is deemed to be transferred in the target language. This study is useful since it highlights the role of Facebook Translation service in translating, educating the public and fighting COVID-19. Consequently, such research would encourage the use and research on the potentiality of MT and FTS in dealing with abrupt crises, such as COVID-19.

1. Introduction

Translation is a medium of human communications. It bridges the gap between human communities and ensures the best way of communicating among global communities[1]. Nowadays, the world is facing the most frightening Coronavirus (COVID-19) outbreak in the last decade. The health authorities, international health organizations, political leaders, healthcare providers and the public have social media sites at their disposal. They post information regarding COVID-19 daily. The flow of social media posts about COVID-19 is unprecedented, and would be beyond the capacity of human translation, let alone an economically viable proposition to have them translated by humans. During lockdowns around the globe, Machine Translation (MT) helped to translate such posts into Arabic. The author in [2] states, "The challenge is how to communicate rapidly changing data across language

borders so that essential information is not lost in translation". In this regard, he goes as far as stating that machine translation is a strong ally in the fight against COVID-19. Social media sites have become popular among young Arabs, who tend to use them daily. The [3] shows that Facebook is the most popular social media site in the Middle East with 200 million daily users, representing almost 70% of Middle East population. Facebook has a translation service, which offers translation for more than 80 languages.

Facebook wide use makes it necessary to evaluate the content and verify its effectiveness and limitations to the end users. English ranks as the top language of social media with 25% of internet.

This means that many Arab Facebookers will be exposed to posts written in English with some able to read and understand it, and a large portion of those who are not competent in English. Facebook has been offering their own translation service since 2011, FTS.

*Corresponding Author: Zakaryia Almahasees zmhases@yahoo.com

www.astesj.com

<https://dx.doi.org/10.25046/aj0601141>

Users can easily activate the translation service on their profiles in two ways. They can click on top right hand of their page and select setting and privacy, then choose language and region. After that, users have to click on the language into which they would like to have posts translated into. The current study scrutinizes FTS in terms of adequacy and fluency in rendering English posts into Arabic during the COVID-19 lockdown.

1.1. Machine Translation

Machine Translation (MT) is the study of the computer systems or online applications in transferring the Natural Languages from one language into another. Author in [1] shows that MT systems are “applications or online services that use machine-learning technologies to translate large amounts of text from and to any of their supported languages. The service translates a “source” text from one language to a different “target” language”. The availability of such online systems for free or at low costs makes it necessary to verify the effectiveness of MT systems in dealing with natural languages especially with the languages which belong to different families such as English and Arabic.

Translation has been integrated into technology thanks to the giant technological progress of the last 70 years, starting with the first successful MT project at the University of Georgetown in 1954 as a result of collaboration between Georgetown University and International Business Machines (IBM). The [2] indicates that the success of the first experiment “attracted a great deal of media attention in the United States. Although the system had little scientific value, its output was sufficiently impressive to stimulate the large scale funding of MT research in the USA and to inspire the initiation of MT projects elsewhere in the world and notably in the USSR”.

Research in [3] shows that 1980 is considered a flagship moment of MT with new developments emerging whilst “more dramatic development took place in MT in the 1990 since computers became more powerful with much higher storage capabilities”. This crowned with the invention of the internet, a source of translation for the general public [4]. Automatic translation tools available on the Internet translate billions of documents daily, which would take human translators months. Whereas MT service was not free in the early history of the Internet, nowadays, we have a set of MT platforms that provide translation service for free such as Google Translate, Microsoft Translator, FTS and others. The author in [5] states, “Machine Translation (MT) is being deployed for a range of use-cases by millions of people on a daily basis. Google Translate and Facebook provide billions of translations daily across many languages”.

1.2. Facebook Translation

The author in [5] indicates that Google translate and Facebook Translation users reach more than 1 billion monthly. The research in [6] shows that Facebook, firstly, used Google Translate service to translate comments into 50 languages. Facebook launched its first Translation Service in 2011, called In-line Translation Facebook Tool [7]. Where [8] states, “After Google integrated the “translate” feature into its social network, which allows users to translate posts and comments into 50 different languages, Facebook, as usual has followed the

footprints of Google+, and quietly announced the launch of “translate” button – powered by Microsoft’s Bing”, setting itself apart from Google. The author in [9] describes the first Facebook Translation tool, “Facebook has quietly introduced a new tool that makes instant inline language translations appear with a single click”. For instance, if a non-Spanish user on a Facebook public page gets across a post in Spanish, then he/she will click on the translation button and then he/she will be able to see the translation of the post in his/her language. FTS works on individual posts on Facebook and not on users’ profiles yet, including comments. Currently, FTS offers translation service for 89 languages and they will continue to add more languages.

FTS adopted phrase to phrase MT approach, thus translating whole sequences of words of differing lengths. The author in [10], the founder of Facebook, explained that translation is the best mean to connect human globally. “Understanding someone’s language brings you closer to them, and I’m looking forward to making universal translation a reality. To help us get there faster, we’re sharing our work publicly so that all researchers can use it to build better translation tools”. In the same year, Facebook research team developed a new MT approach using Convolutional Neural Networks (CNNs), allowing to translate languages more accurately (read: increase quality on a BLEU scale) and up to nine times faster than the traditional Recurrent Neural Networks (RNNs). The author in [11] describes the new MT approach as “Today we’re publishing research on how AI can deliver better language translations. With a new neural network, our AI research team was able to translate more accurately between languages, while also being nine times faster than current methods”, a superior speed confirmed by [12]. The author in [13] has shown that “CNNs have been very successful in several machine learning fields, such as image processing. However, Recurrent Neural Networks (RNNs) are the incumbent technology for text applications and have been the top choice for language translation because of their high accuracy”.

In the late of 2019, Facebook research has declared new advances in NLP, which boost the accuracy of Facebook Translation. They have also introduced a new self-supervised pretraining approach, RoBERTa, that surpassed all existing Natural Language Understanding (NLU) systems on several language comprehension tasks. They have also collaborated with New York University (NYU), DeepMind Technologies, and the University of Washington (UW) to promote their future research [14].

In a first study on Facebook Translation evaluation, it evaluated Facebook translation service in handling low-resources languages: Lao, Kazakh, Haitian, Oromo, and Burmese. They evaluated the translation of English posts into these languages. They implemented different strategies: LASER, back-translations, self-training, multilingual modeling, to improve the translation from English to low source languages. “For instance, Sinhala to English and Nepali to English translations on Facebook have improved from “useful,” which are just accurate enough to understand the meaning, to “good,” which generate full meaning but may have typos or grammatical errors”[15].

The study is mainly concerned with Facebook translation of English posts related to COVID-19 released by international

organizations such as WHO, political leaders, medical specialists and the general public. The study aims to verify the efficiency of FTS in translating posts from English into Arabic and provides constructive feedback about the degree of fluency and adequacy of FTS and whether FTS is considered a reliable source of information or not.

1.3. Machine Translation Evaluation

MT systems Evaluation (MTE) is crucial to the development of MT systems. The author in [4] shows that the evaluation of MT is central to determine the effectiveness and performance of MT systems. Machine Translation Evaluation is essential to all end-users: researchers, designers and users to select the best system to use [16]. In similar way, the development of MT systems depends on the evaluation of MT systems, limitations and strengths of the systems. MT evaluation sheds light on the capacity of the system: what the system can or cannot do. The author in [17] emphasizes that the evaluation of MT system shows accuracy and fluency sought for the audience and purpose and complies with the other features negotiated between the requester and supplier, taking into consideration end-user needs”.

Many evaluation methods have been used over the history of MTE. The first MT evaluation dated back to 1954 to assess the ability to translate 250 words from Russian into English, and they succeeded in translating this small number of words. The success of the experiment attracted a lot of funding. Consequently, in 1962 the first committee, the Automatic Language Processing Advisory Committee (ALPAC), was formed for evaluation. They found that it was not worthy to spend more money on useful MT system. The report ended with nine recommendations to evaluate MT, three of them encouraged to do further research on MT systems [18].

MT systems have achieved high performance close to human accuracy. The authors in [19], [20] shown that MT systems ‘performance have achieved near human level performance. Moreover, the widespread of MT reached millions and therefore it has become a source of information for millions of people. In [3] Google Translate has been used by 500 million daily. However, many studies have highlighted the shortcomings of MT in dealing with Natural language processing (NLP). The have indicated that the main challenge of MT is having large parallel corpora. However, these corpora are not available in all NL [21]-[23]. There are very limited parallel corpora for the majority of language pairs. On the other hand, most of the previous studies depend on monolingual corpora in each language [24], [25]. According to [23] the success of the previsions of MT research is about using inferred bilingual dictionary. Moreover, it is indicated that the success of the previous studies is about using model training in sequence to sequence systems [26]-[28]. Previous studies have adopted back translation strategy in supervising i.e. generating inputs to train the target models and vice versa. Despite the fact that there are many methods designed for evaluation, there is still no generally accepted methodology to evaluate MT systems. Yet, MT needs evaluation, which can be done manually and automatically [22].

To assess MT system, there are three stages to evaluate their performance and efficiency: firstly, the design of the system; secondly, the development of the system; and thirdly, the evaluation of the system by potential customers [27]. With regards

to the third point, MT evaluation is divided into three categories. The first is adequacy, which is used to assess the end user’s needs, such as readability and costs. The second is diagnostic evaluation, where the designers and developers examine the output of MT and its relevance to the input. The third is the performance evaluation in order to assess the systems’ performance in specific areas to assist the developers and the designers of the system [28].

MT could be evaluated manually and automatically [29], [30]. The author in [31] states manual evaluation investigates the systems’ usability via human participants by means of Error Analysis [...], whereas automatic evaluation examines MT outputs through the text’s similarity to a referenced translation. Similarly, MT evaluation has two aspects: intrinsic and extrinsic. The former highlights the language quality, while the former highlights the capacity and efficiency of the system. He also adds that there are two ways to evaluate MT output: automatic and manual evaluation. Automatic evaluation relies on the usage of metrics that could approximate the similarity between the output and the human referenced translation without human interventions, while manual evaluation depends on human evaluators to assess and rank MT output [32]. The author in [15] sums up the advantages of automatic evaluation as fast and cheap: there is no need for bilingual speakers, it requires minimal human labor and can be used as an ongoing evaluation process during the design of the system.

MT could be evaluated automatically in terms of Edit-Distance metrics, Precision and Recall, F-Measure and Word Order. The author in [33] indicates that Edit Distance evaluates MT in terms of additions, omissions, substitutions, which are the requirements for adequate output. Precision and Recall evaluates the degree of n-gram matching between MT and human translation based on referenced translation, such as BLEU and NIST. The F measure is used to measure the overall quality performance. The word order counts the word order sequence between the source sentence and the output of MT. The author in [34] reminds us why BLEU is the most popular metrics among researchers:

one of the reasons why the metric is popular in the community seems to be for its simplicity for MT developers at least. Another reason why BLEU is widely used is that it has the best correlations with human judgments of translation quality. It estimates the similarity between the translated sentence and the referenced sentence.

In fact, automatic evaluation has some advantages, but its shortcomings outweigh its advantages. Automatic evaluation assesses the text similarity between the output and the referenced human translation. It does not look for the meaning transference from the ST to the TT, which is a fundamental requirement for accuracy in translation. Bilingual Evaluation Understudy (BLUE) looks for n-gram similarity not meaning [35]. The author in [36] further argues that BLEU evaluates text similarity rather than meaning. Moreover, the Translation Automation User Society [37] agrees that automatic evaluation could provide one side of quality, which could not reflect the genuine quality of MT output. For these reasons and others, the current study adopts manual methods in evaluating MT output holistically in terms of adequacy and fluency scales provided by TAUS, rather than automated evaluations using BLEU for instance.

In fact, manual evaluation of MT is considered as golden standard to evaluate MT and cannot be overwhelmed by automatic evaluation. Manual evaluation of MT could be done in terms of quality assessment, translation ranking, error analysis, information extraction, comprehension test and post editing. Translation ranking results provide the end users with the degree of the output intelligibility.

In the same vein, [36] developed several tools for MT quality evaluation. They offered these tools on their website to conduct MT evaluation in terms of adequacy and fluency. The adequacy tool enables the evaluators to assess the output based on 1-4 rating scales to verify how much of the meaning is contained in the TT as shown in Table 1 & 2.

Table 1: Adequacy Rating Scale [34]

Adequacy Rating Scales		
4	Everything	All of the meaning in the source is contained, no more, no less
3	Most	Almost all the meaning in the source is contained in the translation.
2	Little	Fragments of the meaning in the source are contained in the translation.
1	None	None of the meaning in the source is contained in the translation.

On the other hand, [36] also provides a fluency tool to evaluate MT output in terms of structural rules as accepted by native speakers of the TL. Fluency evaluates MT output in terms of 1-4 rating scales.

Table 2: Fluency Rating Scale [34]

Fluency Rating scales		
4	Flawless	Refers to a perfectly flowing text with no errors.
3	Good	Refers to a smoothly flowing text even when a number of minor errors are present.
2	Disfluent	Refers to a text that is poorly written and difficult to understand.
1	Incomprehensible	Refers to a very poorly written text that is impossible to understand.

The present study aims to verify the best results of MT output in terms of adequacy and fluency scales provided by TAUS. The study has adopted TAUS since it provides a constructive feedback of the Facebook Machine Translation. Such feedback helps the system’s developers to improve the efficiency of the systems and provide the users with insights about how much is the FTS is adequate and fluent.

2. Literature Review

As discussed above, MTE is essential in evaluating MT output to provide the end users with feedback about the strength and limitations of the systems. Facebook is the most prominent

social media sites globally. It is also classed as the most popular social media site in the Arab region. Therefore, Facebook has become a major source of information for the majority of Facebookers during the COVID-19 outbreak. No research to date has been conducted to evaluate the efficiency of Facebook Translation in rendering English posts into Arabic in general and COVID-19 in particular.

The author in [37] conducted a study to trace the percentage of FTS usage among Jordanian during COVID-19 lockdown. They found that 94.3% use Facebook daily; 87.1% of the participants activated Facebook Translation Service (FTS). Moreover, 62.2% of the participants considered Facebook as a primary source of information regarding COVID-19 and 27.8% as secondary source. The authors in [38] have conducted a descriptive study on medical students in Jordan to assess knowledge, attitude, perceptions and precautionary measures toward COVID-19 among a sample of students in Jordan. They have found that 83.4% used social media sites as their preferred source of information regarding COVID-19. Moreover, [39] have conducted another study in Jordan to assess knowledge, practice and attitude of university students regardless of their majors. They have found that the sources of information for the University students are social media, internet and television. No significant difference was noticed between medical and non-medical college students on the sources of their information.

On the other hand, [40] has shown the social media sites have become a useful tool in confronting the crisis and connecting people together during the crisis. Moreover, public health experts have used social media sites to educate the public about the effects of COVID-19 and to discuss COVID-19 with specialists. The acceleration in the digital life changes the way of approaching the health information. She shows public health experts will use social media sites to spread correct information regarding health problems [41]. The author in [42] shows that Facebook has become a source of discussion and information exchange for 100,000 health care providers regarding the COVID-19 outbreak.

Social media sites have become platforms for survival against social isolation during the COVID-19 to “help people avoid the detrimental effects of social isolation during this pandemic” [42], with Facebook playing a major role as well. Social networks usage has rocketed covering different aspects of our life, including medicine, geography and business growth. Facebook has been used widely during COVID-19 as a platform for business, charity, and community service [43]. The [44] goes as far as suggesting that social media sites during the crisis preached religion to the followers and provided a form of religious instruction and support. On the other hand, social media have become the platform for sharing hurtful messages for people of China. They blamed China as a source of the crisis [45]. However, Facebook and other social media clamped down the spread of fake news concerning COVID-19 and fake news

This research brings the importance of MT systems in handling translation across languages during world crises. The next part highlights the methodology used in assessing MT role during COVID-19 Crisis.

3. Methodology

The study evaluates the effectiveness FTS has in providing adequate and fluent output for English posts related to Coronavirus into Arabic. It provides an answer to the following question: To what extent is FTS capable in providing adequate and fluent translation for English COVID-19 posts into Arabic? To answer this question, six evaluators evaluated the output of FTS in terms of the above TAUS adequacy and fluency scales.

The corpus of the study consists of 300 English posts related to COVID-19, collected over March and April 2020, the peak months of the pandemic. The posts were selected from various Facebook pages that belonged to political leaders, healthcare providers, medical specialists and the public. To ensure consistency and objectivity as recommended by TAUS creators, six evaluators, native speakers of Arabic and near-native speakers of English rated the same data on each scale. The evaluators were acquainted with fluency and adequacy scales along with MT limitations and potentials. Once the corpus was uploaded and sent to the evaluators using TAUS tools, they had to assign scores online on 1-4 rating scales. Then, each report was displayed online on the TAUS reports tool. The data were analyzed using SPSS Statistics to get Pearson interrater agreement among evaluators in ranking the output at two different times. The results showed that there is a high inter-rater reliability among the raters.

Table 3. Interrater correlation

N	Fluency	Adequacy
1	.947**	.969**
2	.973**	.936**
3	.931**	.967**
4	.839**	.782**
5	.976**	.867**
6	.944**	.985**

The above table shows Interrater correlation using Person Correlation between first and second evaluation for each rater, resulting in a statistically significant strong correlation between first and second evaluations for each rater.

4. Results and Discussion

4.1. Fluency

The following example illustrates how FTS translated a World Health Organization (WHO) COVID-19 instructional post into Arabic.

Example 1

Source Text: Many people are making great sacrifices to stay home and protect their health and that of others from COVID-19.

FTS output: الكثير من الناس يقدمون تضحيات كبيرة للبقاء في المنزل وحماية صحتهم وصحة الآخرين من كوفيد-19

alkthyr mn alnas yqdmwn tdhyat kbyrh llbqa' fy almnzl whmayh shthm wshh alakhryn mn kwfyd-19

Back translation: many people are offering great sacrifices to stay home and protect their health and that of other from COVID-19.

This example shows how FTS rendered the English WHO post into Arabic. The analysis shows that FTS rendered the posts fluently. Moreover, four of the raters gave the above example a rate of 4, flawless, while two of them ranked it as 3, Good. The analysis shows that FTS provided intelligible and fluent output.

Example 2

Source Text: Now, anyone returning from overseas is being forced into quarantine for 14 days when they arrive back in Australia. This is mostly occurring in hotels, and at Government expense.

FTS output: الان، اي شخص عائد من الخارج يجبر الى الحجر الصحي لمدة 14 ايام عندما يعود الى استراليا. يحدث هذا في الغالب في الفنادق، وعلى نفقة الحكومة.

alan, ay shkhs 'ea'ed mn alkharj yjbr ala alhjr alshy lmdh 14 ayam 'endma y'ewd ala astralya. yhdth hda fy alghalb fy alfnadq, w'ela nfqh alhkwmh.

Back Translation: Now, anyone returning from abroad is being forced into quarantine for 14 days when they return to Australia. This is mostly occurring in hotels, and at Government expense.

The above example shows how FTS rendered into Arabic a Facebook post of Mark McGowan, Premier of Western Australia, regarding the quarantine for arrivals to Australia. The analysis shows that the raters rated the post in a balanced way. Three of raters rated the output as flawless, where the post is perfectly translated. Moreover, three of them ranked FTS output as Good, where the sentence is fluently translated despite minor errors. In fact, these minor errors did not inhibit the intelligibility of the text.

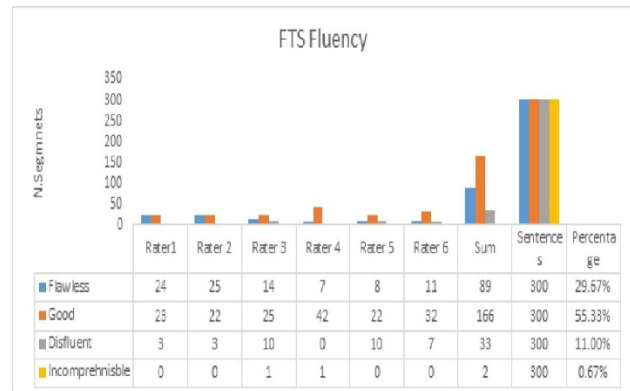


Figure 1: FTS Fluency

The above chart illustrates the degree of fluency obtained by FTS in providing fluent translation for English posts related to COVID-19 into Arabic. The highest fluency scale goes for **good** among the six evaluators with a percentage of 55.33%, where FTS output is easily understood without difficulty even when a number of minor errors are present. The lowest scale goes for **incomprehensible** scale, where the output is poor and impossible to understand which is equal to 0.67%. The first and second raters had similar evaluation. They indicated that the FTS output is held between **good** where the output has minor errors with 55.33% and **flawless** with 29.67%, a perfect output with no errors. The other four raters ranked FTS as **good**, followed by **Flawless** and then **disfluent**, where the text is poorly written and difficult to understand. According to the raters, FTS is mostly adequate in

conveying Source the ST to the TT where the target side translation is grammatically well informed, without spelling errors and experienced as using natural/intuitive language by a native speaker.

Adequacy

Example 3

Source Text: The conclusion I reached is that the Government should end the lockdown after Easter and return to a mitigation strategy, with self-quarantining limited to those most at risk" writes Toby Young.

FTS Output: الاستنتاج الذي توصلت اليه هو انه ينبغي للحكومة انهاء الاعلاق بعد عيد الفصح والعودة الى استراتيجية التخفيف، مع الحجر الصحي الذاتي. يقتصر على اولئك الاكثر عرضة للخطر. يكتب توبي يونغ.

alastntaj aldy twslt alyh hw anh ynbghy llhkwmh anha' alaghlag b'ed 'eyd alfsh wal'ewdh ala astratyjyh altkhfyf, m'e alhjr alshy aldaty, yqtsr 'ela awl'ek alakthr 'erdh llkht, yktb twby ywng.

Back Translation: The conclusion I reached is that the government should end the lockdown after Easter and returned to mitigation strategy, with self-quarantining limited to those at most risks, writes Toby Young.

This example shows how FTS rendered the Telegraph Facebook Posts into Arabic. The results show that the source text meaning is contained in the output translation. The four raters rated the FTS output as Most, where almost all the meaning in the source text is contained in the translation. However, two raters gave a scale of 4 Everything. The current study concurs with the last rating, Everything, that all the information contained in the source text is found in the target output.

Example 4

Source Text: When the world’s doctors come together, amazing things can happen. We’re all hopeful that plasma therapy can help treat COVID-19.

I’m glad Bloomberg Philanthropies can play our part to support Arturo Casadevall's research.

FTS Output:

عندما يجتمع اطباء العالم معا، يمكن ان تحدث اشياء مذهلة. نحن جميعا نامل ان العلاج بالبلازما يمكن ان يساعد في علاج كوفيد-19.

Bloomberg Philanthropies يمكن ان يلعب دورنا لدعم ابحاث ارتورو كاسادفال انا سعيد ان

'endma yjtm'e atba' al'ealm m'ea, ymkn an thdth ashya' mdhlh. nhn jmy'ea naml an al'elaj balblazma ymkn an ysa'ed fy 'elaj kwfyd-19.

ymkn an yl'eb dwrna ld'em abhath artwrw kasadfal Bloomberg Philanthropies ana s'eyd an

Back Translation: When the world’s doctors come together, amazing things can happen. We are hopeful that Plasma thereby can help cure COVID-19. I happy that Bloomberg Philanthropies can take part in supporting Arturo Casadevall's research.

The example shows how FTS transferred the English post into Arabic. The six assessors ranked the output differently. Four of them rated it as Most, while the other two raters ranked it as Little. The study correlates with the four raters that almost all the meaning in the source text is contained in the translation. However, the study indicates that keeping the noun ‘Bloomberg Philanthropies’ untranslated does not inhibit the intelligibility of the text

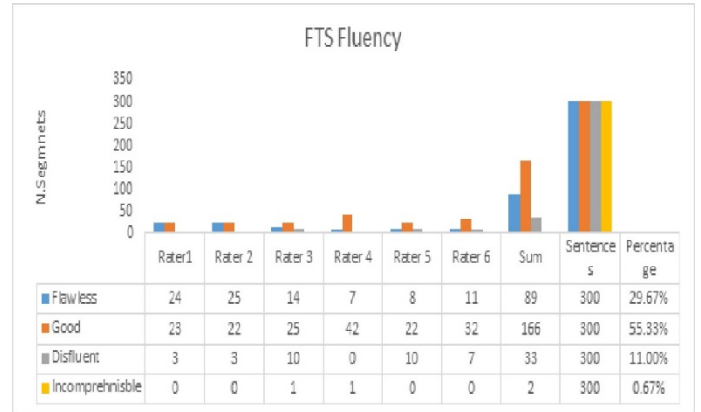


Figure 2: Adequacy Scales

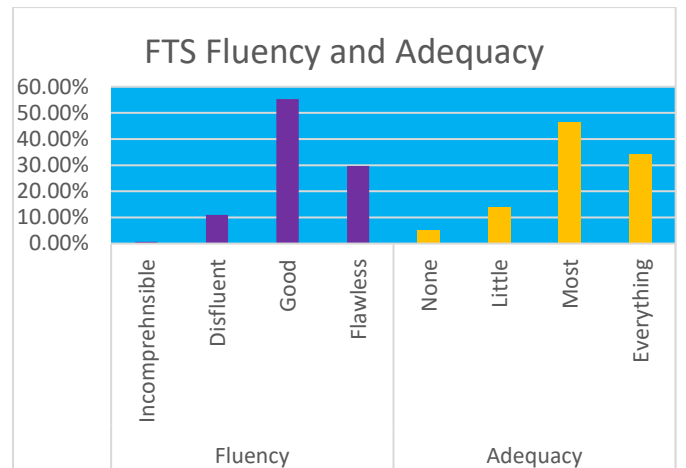


Figure 3: FTS Fluency and Adequacy Performance

The above chart shows the degree of adequacy obtained by FTS in providing adequate translation for English Facebook Posts related to COVID-19 into Arabic. The analysis shows that the highest adequacy scale goes for **most** scale with a percentage of 47%, where almost all of source text meaning is contained in the target text, while the lowest scale goes for **none**, none of the meaning is contained in the translation which is equal to 5.33%. The six raters had similar views that most of the outputs are scaled as **most**, followed by **everything**, all the meaning of the source text is contained in the translation, and then little in the third rank and None with lowest rate. The chart below illustrates FTS performance based on TAUS adequacy and Fluency.

The above chart illustrates the ratings of fluency and adequacy obtained by FTS in rendering English Facebook posts related to COVID-19 into Arabic. The analysis shows that the highest degree of adequacy achieved by FTS is Most, with a percentage of 46.33, followed by Everything and then Little and None, indicating that almost all of the ST meaning is contained in

the TL. Similarly, the highest degree of fluency achieved by FTS is Good, where FTS succeeded in providing Fluent output with minor grammatical categories, followed by Flawless, Disfluent and then Incomprehensible. These two degrees indicate the FTS could provide fluent and adequate translation to the end users.

The results of the study have indicated that Facebook have conducted rigorous research to ensure the best translation quality for the end users. The study has shown that FTS has achieved a high degree of fluency starting with *good* scale, followed by *flawless, disfluent and incomprehensible*. The study has shown that FTS achieved a higher degree of adequacy starting with *most*, followed by *everything, little* and *none*. These results agree with [11] that Facebook introduced Neural Machine Translation (NMT) to provide an accurate translation for the end users. Moreover, the study confirms the findings by [37], that the NMT introduction enhances the accuracy and the speed of MT across languages. The study correlates with [13] that back-translations, self-training, multilingual modeling could improve Facebook translation accuracy across human languages.

5. Conclusion

This research is the first research to fill a gap in the literature about the FTS degree of adequacy in contributing to translate the English Facebook posts related to COVID-19 into Arabic. It provides useful data on FTS adequacy and fluency criteria.

Overall, Facebook translation service obtained a high degree of adequacy and fluency. This is because FTS is trained well to deal with Arabic content. More importantly, we could benefit from FTS with Human Assisted Machine Translation (HAMT), in which a computer system does most of the translation, appealing in case of difficulty to a human translator for post editing services. Therefore, FTS helps in fighting COVID-19 since it facilitates the process of information exchange. Moreover, the study recommends further studies to conduct diachronic evaluation for FTS output over a period of time, since the current study is limited to assess FTS output during the COVID-19 outbreak in Jordan. Despite the fact that FTS achieved a higher degree of adequacy and fluency scales, the current analysis does not cover all genres and therefore there is a lack of integrative evaluation studies that pave the research for the holistic analysis and the error analysis studies. The FTS is still far from reaching fully adequate and fluent translation of a quality obtained by human translators.

6. Limitation and Study Forward

This study has some limitations as it investigates the adequacy and fluency of Facebook Translation Service during COVID-19 on a particular topic, namely COVID-19, from the instructors' perspective, within a defined timeframe (post COVID-19 era), and geographical context (Jordan) to answer three research questions. Consequently, further research will be necessary to examine the adequacy and fluency of FTS in other fields and different timeframes. That the current study focuses on the period of the outbreak of the COVID-19 pandemic, it is recommended that other researchers examine the long-term adequacy and fluency of FTS in all fields. It is also recommended that researchers conduct similar research on the adequacy and fluency of FTS in translating from English into their languages at different situations.

References

- [1] Microsoft.com, Machine Translation - Microsoft Translator for Business, 2020.
- [2] J. Hutchins, Compendium of Translation Software Compendium of Translation Software Directory of commercial machine translation systems and computer-aided translation support tools behalf of The International Association for Machine Translation by The European Association, 2010.
- [3] Z. Almahasees, Diachronic Evaluation of Google Translate, Microsoft Translator and Sakhr in English-Arabic Translation — the UWA Profiles and Research Repository, 2020.
- [4] F. Gaspari, J. Hutchins, "Online and Free! Ten Years of Online Machine Translation: Origins, Developments, Current Use and Future Prospects," 2007.
- [5] Bernard Scott, Machine Translation: Technologies and Applications | SpringerLink, 2018.
- [6] C. Amendola, How to Use Facebook Translate Button – Multilizer Translation Blog, 2020.
- [7] Singh, How to enable new Facebook In-Line Translation Tool on Facebook Pages | Tech 24 Hours, 2020.
- [8] T. Martín-Blas, A. Serrano-Fernández, "The role of new technologies in the learning process: Moodle as a teaching tool in Physics," Computers and Education, **52**(1), 35–44, 2009, doi:10.1016/j.compedu.2008.06.005.
- [9] White, Facebook Launches New In-Line Translation Tool, 2011.
- [10] M. Zuckerberg, Mark Zuckerberg - Today we're publishing research on how... | Facebook, 2017.
- [11] J. Gehring, M. Auli, D. Grangier, D. Yarats, Y.N. Dauphin, "Convolutional sequence to sequence learning," in 34th International Conference on Machine Learning, ICML 2017, International Machine Learning Society (IMLS): 2029–2042, 2017.
- [12] Facebook Engineering, A novel approach to neural machine translation - Facebook Engineering, 2017.
- [13] Facebook AI, New research awards in NLP and machine translation, 2019.
- [14] EuroMatrix, [PDF] 1.3: Survey of Machine Translation Evaluation | Semantic Scholar, 2009.
- [15] G.S. Koby, P. Fields, D.R. Hague, A. Lommel, A. Melby, "Defining Translation Quality," Tradumàtica: Tecnologies de La Traducció, (12), 413, 2014, doi:10.5565/rev/tradumatica.76.
- [16] N.R. Council, Language and Machines, National Academies Press, 1966, doi:10.17226/9547.
- [17] H. Hassan, A. Aue, C. Chen, V. Chowdhary, J. Clark, C. Federmann, X. Huang, M. Junczys-Dowmunt, W. Lewis, M. Li, S. Liu, T.-Y. Liu, R. Luo, A. Menezes, T. Qin, F. Seide, X. Tan, F. Tian, L. Wu, S. Wu, Y. Xia, D. Zhang, Z. Zhang, M. Zhou, "Achieving Human Parity on Automatic Chinese to English News Translation," arXiv preprint arXiv, 2018.
- [18] Y. Wu, M. Schuster, Z. Chen, Q. V. Le, M. Norouzi, W. Macherey, M. Krikun, Y. Cao, Q. Gao, K. Macherey, J. Klingner, A. Shah, M. Johnson, X. Liu, L. Kaiser, S. Gouws, Y. Kato, T. Kudo, H. Kazawa, K. Stevens, G. Curian, N. Patil, W. Wang, C. Young, J. Smith, J. Riesa, A. Rudnick, O. Vinyals, G. Corrado, et al., "Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation," arXiv preprint arXiv 2016.
- [19] P. Koehn, "Neural Machine Translation," Tradumàtica: Tecnologies de La Traducció, (15), 66, 2017.
- [20] P. Isabelle, C. Cherry, G. Foster, "A Challenge Set Approach to Evaluating Machine Translation," EMNLP 2017 - Conference on Empirical Methods in Natural Language Processing, Proceedings, 2486–2496, 2017.
- [21] R. Sennrich, B. Haddow, A. Birch, "Improving Neural Machine Translation Models with Monolingual Data," 54th Annual Meeting of the Association for Computational Linguistics, ACL 2016 - Long Papers, **1**, 86–96, 2015.
- [22] G. Lample, A. Conneau, L. Denoyer, M. Ranzato, "Unsupervised Machine Translation Using Monolingual Corpora Only," ArXiv, 2017.
- [23] M. Artetxe, G. Labaka, E. Agirre, K. Cho, "Unsupervised Neural Machine Translation," ArXiv, 2017.
- [24] D. Bahdanau, K.H. Cho, Y. Bengio, "Neural machine translation by jointly learning to align and translate," in 3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings, International Conference on Learning Representations, ICLR, 2015.
- [25] I. Sutskever, O. Vinyals, Q. V. Le, "Sequence to Sequence Learning with Neural Networks," Advances in Neural Information Processing Systems, **4**(January), 3104–3112, 2014.
- [26] P. Vincent, H. Larochelle, Y. Bengio, P.A. Manzagol, "Extracting and composing robust features with denoising autoencoders," in Proceedings of the 25th International Conference on Machine Learning, Association for

- Computing Machinery (ACM), New York, New York, USA: 1096–1103, 2008, doi:10.1145/1390156.1390294.
- [27] W. Hutchins, H. Somers, “An introduction to machine translation,” Undefined, 1992.
- [28] L. Hirschman, L. Hirschman, H.S. Thompson, “Overview of Evaluation in Speech and Natural Language Processing,” 1997.
- [29] S. Chan, Routledge Encyclopedia of Translation Technology - 1st Edition - Sin-W, 2014.
- [30] Z.M. Almahasees, “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, 2017, doi:10.18178/ijll.2017.3.1.100.
- [31] M. Przybocki, G. Sanders, A. Le, *Edit Distance: A Metric for Machine Translation Evaluation*, 2006.
- [32] E. Stephens, “Human Evaluation on Statistical Machine Translation,” Undefined, 2014.
- [33] E. Reiter, A structured review of the validity of BLEU, *Computational Linguistics*, **44**(3), 393–401, 2018, doi:10.1162/COLI_a_00322.
- [34] TAUS, Adequacy/Fluency Guidelines - TAUS - The Language Data Network, 2019.
- [35] R. Tatman, Evaluating Text Output in NLP: BLEU at your own risk | by Rachael Tatman | Towards Data Science, 2019.
- [36] TAUS, TAUS Quality Dashboard. An Industry-Shared Platform for Quality Evaluation and Business Intelligence September, PDF Free Download, 2015.
- [37] Z. Mohammad, Adam Benvie, Improving the accuracy & speed of translations with Neural Machine Translation - Watson Blog, 2018.
- [38] A.I. Khasawneh, A.A. Humeidan, J.W. Alsulaiman, S. Bloukh, M. Ramadan, T.N. Al-Shatanawi, H.H. Awad, W.Y. Hijazi, K.R. Al-Kammash, N. Obeidat, T. Saleh, K.A. Kheirallah, “Medical Students and COVID-19: Knowledge, Attitudes, and Precautionary Measures. A Descriptive Study From Jordan,” *Frontiers in Public Health*, 2020, doi:10.3389/fpubh.2020.00253.
- [39] H. Alzoubi, N. Alnawaiseh, A. Al-Mnayyis, M. Abu-Lubad, A. Aqel, H. Al-Shagahin, “Covid-19 - Knowledge, attitude and practice among medical and non-medical university students in Jordan,” *Journal of Pure and Applied Microbiology*, 2020, doi:10.22207/JPAM.14.1.04.
- [40] Zakrzweski, The Technology 202: Coronavirus could change how social networks approach public health - The Washington Post, 2020.
- [41] H. Ouyang, “At the Front Lines of Coronavirus, Turning to Social Media,” *The New York Times*, 2020.
- [42] A. Ertzoni, The Sociology of Surviving the Coronavirus | The National Interest, 2020.
- [43] Goldstein, Facebook Unveils New Feature to Help Those in Need amid Coronavirus | PEOPLE.com, 2020.
- [44] Heilweil, Coronavirus social distancing leads to empty churches and a rise in apps - Vox, 2020.
- [45] Wong, Coronavirus sparks a different kind of problem for social networks - CNET, 2020.