The Quality of Machine Translation Assessment on Gender Markers Lingual Units

Hapni Nurliana H.D Hasibuan
Universitas Negeri Yogyakarta
Jl. Colombo Yogyakarta No.1, Karang Malang, Caturtunggal, Kec. Depok, Kabupaten Sleman, Daerah Istimewa Yogyakarta 55281
hapnihasibuan@gmail.com

ABSTRACT

Machine Translation (MT) is one of the most advanced and elaborate research fields within Translation Technology, the quality of MT output has always been a great concern, and MT evaluation is a popular research topic. This research aims to assess the quality translation on the gender markers lingual unit of the Arabic short story "ﻋَﺒْﺪُ ﷲِﻭَﺍﻟْﻌُﺼْﻔُﻮْﺭُ" which is translated to English and Indonesian using Machine Translation. Machine Translation used is Google Translation. Further, the research was qualitatively based. The subject is gender markers lingual unit that taken from the Arabic short story "ﻋَﺒْﺪُ ﷲِﻭَﺍﻟْﻌُﺼْﻔُﻮْﺭُ". The key instrument of this research is the human instrument. Additional instruments used to support this research consisted of tables of the lingual units of gender markers and a table of rating scales based on Nababans' theory (2012). The findings showed that this analysis had discovered 72 lingual units of gender markers in the short story. The most frequent type found was personal pronouns. The research was concluded that the google translate translation on the gender markers lingual units has high quality on the accuracy, acceptance, and readability level.

Keywords: translation quality assessment; machine translation; gender markers; lingual unit; short story
INTRODUCTION

With the increased globalization, translation plays a significant role in human's life. People can communicate and interact with others all around the world. Without understanding the culture or system of other languages, communication might be difficult. As the definition, translation is transfer words, phrases, clauses, or sentences from the source language to target language Catford (1965: 20). In addition, Nida and Taber (1974: 12) stated, “translating consists of reproducing in the receptor language the closest natural equivalent of the source-language message, first in terms of meaning and secondly in terms of style”. Therefore, it can be said that translation contains the reproduction into the target language as naturally as possible in terms of meaning and style.

Avoiding translation difficulties, in the sophisticated technology era, humans are no longer flipping through the dictionary or coming to the library. Instead, the process of translation cannot only be done by humans but also by machine translation machines. The first machine translation was begun with the dictionary software containing limited words and phrases. However, the rapid development of technology has improved machine translation, translating phrases, sentences, paragraphs, complete manuscripts, and books.

Several applications have been created to translate, such as google translate, bing.com, translate.com. All these machines can translate into more than 50 languages. One of them is from Arabic to English or Indonesian. However, as a machine translator, translation is not perfect because basically, machine translation provides only service with a "rough" translation still needs revision. This translation needs to be studied more widely, especially to see errors in the existing language, because many parties doubt its quality. That language error can be morphological, grammatical, syntactic, or phonological.

Many previous types of research of machine translation have been done. Koponen (2010) concluded that the error classification explained the focus on the un-equivalence of machine translation in semantic components (individual concepts and relations between them) of the source and target texts. Following research conducted by Alawneh et al. (2011), the result showed that translating the well-structure of English into Arabic through machine translation has not yet achieved the high-quality assessment. Maslihah (2018) researched how google translates accuracy from educational verses Arabic into Indonesian and its impact. The conclusion stated the Google Translate application has a deviation accuracy level, but this research also expected students to realize the weakness and deplorable impact after using google translate.

The other research conducted by Fauziah et al (2012: 15) stated that machine translation has a weakness according to the accuracy of the meaning of google translate in the book "Fathul Qorib" is not accurate. Because, google translate is only correct in translating the meaning of isim, fi'il, and letters
literally, not in a perfect sentence. In evidence, Khoiriyah (2020) stated that google translate has a good quality in vocabulary terms but no in conveying the meaning. Furthermore, in grammatical, google translate does not have grammatical arrangement.

Moreover, some of the results of the google translate translation on special meanings, religious terms, spelling, punctuation, grammar, morphology, and lexicon are inaccurate. Thus, the meaning of the resulting translation is incorrect, and the message contained in it cannot be understood. Another obstacle experienced by users of this service is that translations that are not grammatical are often found.

Considering the statement above, one of the sub-categories in grammar is gender (Muhammadun, 2016: 46). In Webster's New World Dictionary, gender is defined as the visible differences between men and women regarding values and behavior. This is reinforced by Lips (1993: 7), which defines gender as cultural expectations of men and women. Therefore, based on the close relationship between language and cultural treatment of the gender system, the lingual unit elements function as sex differentiators, both at the phonological, morphological and lexicon levels.

One language that has a complex gender marker is Arabic. All words in Arabic are masculine except the words that have a feminine gender marker. The complexities of Arabic gender have not been extensively researched. The gender markers of Arabic grammatically are nouns, personal pronouns, relative pronouns, demonstrative pronouns, adjectives, verbs, and particles (Najjar & Shahin, 2015: 256).

A word's gender can affect its form and behavior. Farghal and Shunnaq (1999: 56) that English makes very few gender distinctions in its pronominal system. However, the issue of gender markers in Indonesian rules are not discussed in detail. It can be expressed lexically.

Based on the background, assessing the quality of machine translation in Gender as a sub-category grammatical becomes an interesting idea to research. Furthermore, gender in Arabic, Indonesian, and English have different forms. Thus, the problems can be formulated in the following question: “How does the quality of Google Translate in the gender markers from Arabic to Indonesian and English, as well as its impact brought about?” The result research will give the beginner or student translator important information about the accuracy, acceptability, and readability of google translate in translating Arabic stories into Indonesian and English, especially in translation gender marker.
METHOD

This research used a qualitative approach to assess the quality of google translate results in the story. The data source was the short story of فيـيـن in Arabic as SL and in Indonesian and English as TL. To obtain accurate data in this research, the researcher applied several general techniques to collect and analyze data. In data collection, the technique starts from reading data sources related to the research topic. Then, data analysis finds the short story, translating through google translate into English and Indonesian, classifying the gender marker, asking the raters to assess the data through the instrument table of rating scale. The instruments of this research were the researcher herself, raters who are capable in the translation field, and the instrument table of rating scales based on Nababan (2012).

Table 1:
Assessment Criteria of Translation Accuracy Level
(Adapted from Nababan, et. al, 2012: 50)

<table>
<thead>
<tr>
<th>Translation Categories</th>
<th>Score</th>
<th>Qualitative parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accurate</td>
<td>3</td>
<td>The source text content that consists of word, technical term, phrase, clause or sentence is accurately transferred into the target text without distortion of meaning.</td>
</tr>
<tr>
<td>Less accurate</td>
<td>2</td>
<td>Most of the source text content, such as words, technical terms, phrase, clause or sentence is transferred accurately into the target text, but still needs improvement.</td>
</tr>
<tr>
<td>Inaccurate</td>
<td>1</td>
<td>Most of, or the whole source text content is inaccurately transferred into the target text. It includes the deletion and elimination of important information of source text.</td>
</tr>
</tbody>
</table>
Table 2:
Assessment Criteria of Translation Acceptability Level
(Adapted from Nababan, et al., 2012: 50)

<table>
<thead>
<tr>
<th>Translation Categories</th>
<th>Score</th>
<th>Qualitative parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acceptable</td>
<td>3</td>
<td>Translation feels natural; where the technical term has been commonly used and familiar to the reader; while the phrases, clauses and sentences used do not conform to Indonesian standards.</td>
</tr>
<tr>
<td>Less Acceptable</td>
<td>2</td>
<td>The translation feels natural but consists of a few uncommon technical terms or slight grammatical errors.</td>
</tr>
<tr>
<td>Unacceptable</td>
<td>1</td>
<td>Translation feels unnatural, and consists of technical terms that are uncommonly used and unfamiliar to the reader; the phrases, clauses and sentences do not conform to the Indonesian standard.</td>
</tr>
</tbody>
</table>

Table 3:
Assessment Criteria of Translation Readability Level
(Adapted from Nababan, et al. 2012: 5)

<table>
<thead>
<tr>
<th>Translation Categories</th>
<th>Score</th>
<th>Qualitative parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>High Readability</td>
<td>3</td>
<td>The content of text (words, phrases, clauses and sentences) can be understood easily.</td>
</tr>
<tr>
<td>Moderate Readability</td>
<td>2</td>
<td>Generally, the content of text can be understood, but there are certain parts that should be re-read to comprehend.</td>
</tr>
<tr>
<td>Low readability</td>
<td>1</td>
<td>The text is difficult to comprehend</td>
</tr>
</tbody>
</table>
FINDINGS AND DISCUSSION

Based on the analysis, the researcher found 72 words containing gender markers found in the narrative story " عبد الله وَ ﻪُ ﺪُ ﺒْ ﻋَ", as shown in the following table:

Table 4: The Gender Markers of 

<table>
<thead>
<tr>
<th>No</th>
<th>Gender Marker Categories</th>
<th>Utterance (Muzakkar)</th>
<th>Feminine (Muannas)</th>
<th>(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Pronoun</td>
<td>29</td>
<td>6</td>
<td>49</td>
</tr>
<tr>
<td>2</td>
<td>Demonstrative pronouns</td>
<td>3</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>Adjective</td>
<td>10</td>
<td>3</td>
<td>18</td>
</tr>
<tr>
<td>4</td>
<td>Nominal</td>
<td>18</td>
<td>2</td>
<td>28</td>
</tr>
<tr>
<td>5</td>
<td>Numeral</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>61</td>
<td>11</td>
<td>100</td>
</tr>
</tbody>
</table>

The table above shows the frequency of gender markers, there are 5 categories of reference based on Muhammadun’s theory (2016). There were 61 data markers for masculine gender from 72 data, while gender feminine funding was 11 from 72 data. This is because the story " عبد الله وَ ﻪُ ﺪُ ﺒْ ﻋَ" is a story about a little boy, a bird and his father. All characters are male nouns, meaning more masculine personal pronouns than the femininity used in the story.

The table above also concludes that the personal pronouns are most often found in stories, with 29 masculine utterances and 6 feminist utterances (48.6%). The personal pronoun is a pronoun known as اسم ضمير: Ism damir. It manifests itself in two forms, namely as an independent word, which is known as ضمير افتصل: damir munfasil. It is the pronoun that separate persons; and as a suffix (suffix) which is attached to a noun, verb, or a preposition known as ضمير مرفوع: damir muttasil. It is the pronoun that connect persons. This happens because the story " عبد الله وَ ﻪُ ﺪُ ﺒْ ﻋَ" contains many personal pronouns, either damir munfasil (standing alone) or damir muttasil (being together).
1. Assessment of Translation Quality

The story has found 72 gender marker words. In the next discussion, namely assessing the quality of the Machine Translation results in the story. This assessment is based on Nababan’s (2012) translation quality assessment theory which states that assessing the quality of the translation is divided into three aspects, namely aspects of accuracy, acceptability, and readability.

a. Aspects of Accuracy

According to Nababan's theory, accuracy is divided into accurate, less accurate, and inaccurate. For example, the story:

Table 5: The Accuracy Assessment of Google Translate

<table>
<thead>
<tr>
<th>Translation category</th>
<th>Frequency</th>
<th>Percentage (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accurate</td>
<td>55</td>
<td>76</td>
</tr>
<tr>
<td>Less-accurate</td>
<td>8</td>
<td>11</td>
</tr>
<tr>
<td>Inaccurate</td>
<td>9</td>
<td>13</td>
</tr>
<tr>
<td>Total</td>
<td>72</td>
<td>100</td>
</tr>
</tbody>
</table>

The table above showed that 55 out of 72 data (76%) were translated accurately, 8 data (11%) were translated less accurately, and 9 data (13%) were translated inaccurately. It can be concluded that the Machine Translation of the gender marker word in the story can be said to be an accurate translation. The following explicated the examples of data from the three categories of accuracy assessments based on Nababan (2012: 51):

i. Accurate Translation

Accurate translation is defined as the meaning of words, technical terms, phrases, clauses, sentences, or source language text accurately transferred to the target language; there is no distortion of meaning (Nababan, 2012). There are 55 out of 72 data that can be described as accurate translation. An example of an accurate translation can be seen below:

Data 2/II
ST: ﺑﱢ ﻧَٰ  
TT1: Rumahnya
TT2: His house

Based on observations of data number 2 / II, hi is the personal pronoun suffix in Arabic, the pronoun bi in the word baitibi refers to him as a man, namely Abdullah. Baitibi means his house in English and rumahnya in Indonesia.
The Quality of Machine Translation Assessment:...
Hapni Nurliana H.D Hasibuan
DOI: https://doi.org/10.26714/lensa.11.1.2021.54-67
inconsistent translations in every language.

iii. Inaccurate Translation

Inaccurate translation is defined as the meaning of words, technical terms, phrases, clauses, sentences or the source language text is inaccurately transferred to the language (Nababan, 2012: 50). The table showed that the inaccurate translation category on Google translate has 9 out of 72 data.

Data 1 / 1
ST: غَيْبَةُ اللَّهِ
TT1: Hamba Tuhan
TT2: Servant of God

The data above also explicated many times in this story. The translation in English and Indonesian was inaccurate. The meaning in the source language is not transferred accurately. Further, the name of the person should not be translated.

The following data was explicated as an inaccurate translation.

Data 8 /III
ST: ﻧَaliases ﻟَهَا وَحَدَّ ﺔبِ
TT1: Ia mengambil satu darinya dan memasukkannya ke dalam.
TT2: But he took one from it and took it down

Ke dalam and took it was not the accurate translation from ﺔبِ. If we can guess the meaning, the correct translation should be dengannya or with his. So it meant the translated words in both target texts were inaccurate.

b. Acceptability Aspects

Table 6: The Acceptance Assessment of Google Translate

<table>
<thead>
<tr>
<th>Translation Category</th>
<th>Frequency</th>
<th>Percentage (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acceptable</td>
<td>57</td>
<td>79</td>
</tr>
<tr>
<td>Less-Acceptable</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Unacceptable</td>
<td>12</td>
<td>17</td>
</tr>
<tr>
<td>Total</td>
<td>72</td>
<td>100</td>
</tr>
</tbody>
</table>

The table above showed that 57 out of 72 data (79%) acceptable, 3 (4%) less acceptable, and 12 (17%) unacceptable. Therefore, it can be concluded that the Google Translation results of the gender markers in this story has an acceptable translation. The following explicated the examples of data from the three categories of acceptance assessments based on Nababan (2012: 51):
i. Acceptable Translations

Translation can be said to be acceptable if the translation feels natural; technical terms are commonly used and familiar to readers; The phrases, clauses and sentences used to follow the rules of the Indonesian language (Nababan: 2012). Thus, there are 57 out of 72 data, including acceptable translations.

Data 3/I, 3/II
ST: شجرة عالية
TT1: Pohon yang tinggi
TT2: a tall tree

The data mentioned is from 2 data but compiled in 1 phrase to be analyzed. The word شجرة is a feminine nominal category in Arabic and عالية is a feminine adjective because it follows the noun in front of it. Correspondingly, all words in Arabic are masculine except those that have a feminine gender marker. The feminine form is derived from the masculine form. Because masculine is the original form, this form does not need a marker that shows its distinctiveness as masculine (Muhammadun, 2016: 50).

This data can be seen that the term شجرة is translated accurately. The term above should be translated as pohon yang tinggi/a tall tree. However, there are additional words that are usually applied in TL.

Data 4/I, 4/II, 4/III
ST: أراد أن يأخذ واحدًا منها ولقأ سمع صوتها
TT1: Dan ketika dia mendengar suaranya, dia ingin mengambil salah satu dari mereka
TT2: And when he heard its sound, he wanted to take one of it.

As the explanation before this example, this example data also were combined from three data in one sentence. All of the bold words were translations acceptable in the target language. Therefore, the readers might understand the translated words.

ii. Less-acceptable Translation

According to Nababan (2012: 51), a translation that feels natural but there is a problem with using technical terms or a slight grammatical error is called an unacceptable translation. As stated in the translation acceptability assessment table above. There are 3 out of 72 data that contain less acceptable gender markers.

Data 2/I, 13/I
ST: عبد الله
TT1: Abd Allah

The Quality of Machine Translation Assessment...
TT2: Servant of God

The data example is mentioned twice in the story. The translation has been natural but not acceptable due to a few grammatical errors in translating and writing people's names. Therefore, the ST should be translated as Abdullah in Indonesian and English.

Data 17/III
ST: ﺍِﻟْﺤُﺰْﻥ ﺃَﮐُﻮْﻥُ ﻓِﻲْ ﻏَﺎﻳَﺔِ ﻭَ ﺍﻷَْﻟَﻢِ
TT1: Aku akan sangat sedih dan sedih
TT2: Be in the end of sorrow and pain

The translated word meant the less acceptable word, Indonesian as the first target text should be translated as rasa sakit. As in English translation was acceptable. The Indonesian translation established the translation as less acceptable.

iii. Unacceptable Translation

An unacceptable translation is a translation that is not natural or feels like a translation work; technical terms used are not commonly used and are not familiar to the reader; the phrases, clauses, and sentences used are not following Indonesian language rules Nababan (2012: 51). There were 12 out of 72 gender marker data which included in the unacceptable translation by google translation.

Data 8/III
ST: ﺭَ ﻧَﻭَ ﻪِ ﺑِ
TT1: Memasukkan nya ke dalam.
TT2: He took one from it

The data is one of the unacceptable translation words. The translation is unnatural and unfamiliar to the reader. If ST translated literally, the meaning of ﺭَ ﻧَﻭَ ﻪِ ﺑِ means turun ke bawah bersamanya. In addition, hi in the word bibi is a personal pronoun suffix in the third person. The word hi refers to usfur (little bird).

Data 5/1
ST: ﻣَ ﺟَ ﺖَ ﻩَ
TT1: Mengulurkan tangannya.
TT2: Reach out his hand.

The translated sentences missed the little meaning. It meant that the data does not got translated fully by the Google Translate. Therefore, the data above is classified as the unacceptable translation.
c. Readability Aspects

Table 7: The Readability Assessment of Google Translate

<table>
<thead>
<tr>
<th>Translation Categories</th>
<th>Frequency</th>
<th>Percentage (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>High Readability Translation</td>
<td>59</td>
<td>82</td>
</tr>
<tr>
<td>Moderate Readability Translation</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Low Readability Translation</td>
<td>13</td>
<td>18</td>
</tr>
<tr>
<td>Total</td>
<td>72</td>
<td>100</td>
</tr>
</tbody>
</table>

The table above shows that 59 out of 72 data (82%) have high readability and 13 data (18%) have low readability. It concluded that the Google Translation of the gender markers in the story is legible and easy to understand. The following data were the examples from the three readability categories based on Nababan (2012: 51):

i. High Readability Translation

Qualification for high-level translation legibility is when words, technical terms, phrases, clauses, sentences or translated texts can be easily understood by the reader (Nababan, 2012: 51). There are 59 out of 72 pieces of data in this story.

Data 20 /I-IV

ST: وطلبتُ من الخادم أن يرجع ال nouveور إلى أهله

TT1: Dia meminta *pelayan* untuk mengembalikan *burung* itu ke keluarganya.

TT2: And he asked the servant to return the blurry to his family.

The example sentence above has four words containing the gender markers. Interestingly, the data above are assessed as high readability translation. *طالبُ* is a verb with he as the hidden subject. *خادم* is a masculine noun. *نوعْ* is masculine nouns from animal species and *هَيْ* in the word referred to the bird. These translations were accurate and accepted in target languages. It was caused the present data were enough to understand by reading one time.

Data 7/I-II

ST: ولكن وَلَا قَلْبُ هَيْ رَهْ لتَحَفّهَا

TT1: Tetapi *hatinya* adalah apa yang melemahkan *kondisinya*.

TT2: But his heart is what threw her condition.

The example has two gender markers. It included into high readability translation. So the translation will get easily in reading and understanding the
The Quality of Machine Translation Assessment:...
Hapni Nurliana H.D Hasibuan
DOI: https://doi.org/10.26714/lensa.11.1.2021.54-67

The Quality of Machine Translation Assessment:...
Hapni Nurliana H.D Hasibuan
DOI: https://doi.org/10.26714/lensa.11.1.2021.54-67

story.

ii. Moderate Readability Translation
The reader's moderate level of readability assessment is generally understandable; however, certain passages must be read more than once to understand the translation (Nababan, 2012: 51). By the table above, no data falls into the category of moderate readability from the results of google translate on gender words in the story of

iii. Low-Level Readability Translation
Translations are difficult for readers to understand have a low readability rating (Nababan, 2012: 51). Regarding the frequent readability table, the low-level readability translation contains 13 out of 72 data. Therefore, the following data will be presented.

Data 14/II
ST: ﴿يا عَدُّدَ اللَّه﴾
TT1: Hai hamba Tuhan
TT2: O servant of God.

Data 14/III
ST: ﴿نَ أَيْنَ أَجْنَبَ بِهِ﴾
TT1: Dari mana asal mu?
TT2: From where you came with it.”

The gender markers data above were difficult to understand. First, عَدُّدَ الله as the name of the person should not be translated literally. Second, hi as the suffix in bihi as the personal pronoun had a distortion of meaning in Indonesian translation

These examples caused the translated text to become difficult to understand. Nababan (2012: 51) stated that translation with low readability is difficult to comprehend.

CONCLUSION

Based on the research findings, the quality of 72 gender marker translation assessments is highly accurate, acceptable, and readability in Indonesian and English translation. Moreover, Arabic translation into Indonesian as TT1 and English as TT2 has basic grammatical rules.

Google translate cannot analyze the grammar and context of the sentence from the source language (Arabic) to the target languages (Indonesian and English). Furthermore, the weakness of Google Translate is unclear and inconsistent in translating the whole text.

The use of machine translation provided much convenience to all
users. Google Translate or other Machine Translation can be an economical and practical reason increasingly that creates an instant helper in solving the translation problems. However, the user of machine translation must be accompanied by an awareness of the weaknesses in machine translation. Furthermore, the translators should develop the knowledge of the language systems and cultures to avoid machine translation errors.

REFERENCES


