

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

hapniasibuan@gmail.com

Article History: Submitted on 5th April 2021; Accepted on 3rd June 2021;
Published on 30th June 2021

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 "وَالْعَصْفُورُ لِلَّهِ عِنْدُ" 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 "وَالْعَصْفُورُ لِلَّهِ عِنْدُ". 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 *عَبْدُ اللَّهِ وَالْعَصْفُورُ* story. The data source was the short story of *عَبْدُ اللَّهِ وَالْعَصْفُورُ* 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)

Translation Categories	Score	Qualitative parameters
Accurate	3	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.
Less accurate	2	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.
Inaccurate	1	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.

Table 2:
 Assessment Criteria of Translation Acceptability Level
 (Adapted from Nababan, et. al, 2012: 50)

Translation Categories	Score	Qualitative parameters
Acceptable	3	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.
Less Acceptable	2	The translation feels natural but consists of a few uncommon technical terms or slight grammatical errors.
Unacceptable	1	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.

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

Translation Categories	Score	Qualitative parameters
High Readability	3	The content of text (words, phrases, clauses and sentences) can be understood easily.
Moderate Readability	2	Generally, the content of text can be understood, but there are certain parts that should be re-read to comprehend.
Low readability	1	The text is difficult to comprehend

FINDINGS AND DISCUSSION

Based on the analysis, the researcher found 72 words containing gender markers found in the narrative story "عَبْدُ اللَّهِ وَالْعَصْفُورُ", as shown in the following table:

Table 4:
 The Gender Markers of *عَبْدُ اللَّهِ وَالْعَصْفُورُ*

No	Gender Marker Categories	Utterance		(%)
		Masculine (Muzakkar)	Feminine (Muannas)	
1	Pronoun	29	6	49
2	Demonstrative pronouns	3	0	4
3	Adjective	10	3	18
4	Nominal	18	2	28
5	Numeral	1	0	1
Total		61	11	100

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 "عَبْدُ اللَّهِ وَالْعَصْفُورُ" is a story about a little boy, a bird and his father. All characters are male nouns, meaning more masculine personal pronouns than the feminism 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 منفصل *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 متصل *damir muttasil*. It is the pronoun that connect persons. This happens because the story *عَبْدُ اللَّهِ وَالْعَصْفُورُ* contains many personal pronouns, either *dhamir munfasil* (standing alone) or *dhamir muttasil* (being together).

1. Assessment of Translation Quality

The story *عَبْدُ اللَّهِ وَالْعَصْفُورُ* has found 72 gender marker words. In the next discussion, namely assessing the quality of the Machine Translation results in the *عَبْدُ اللَّهِ وَالْعَصْفُورُ* 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 *عَبْدُ اللَّهِ وَالْعَصْفُورُ* story:

Table 5:
The Accuracy Assessment of Google Translate

Translation category	Frequency	Percentage (%)
Accurate	55	76
Less-accurate	8	11
Inaccurate	9	13
Total	72	100

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 *عَبْدُ اللَّهِ وَالْعَصْفُورُ* 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: *بَيْتِهِ*

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.

Data 4:

ST: أَنْ أَرَى أَنْ يَأْخُذَ وَاحِدًا مِنْهَا

T^T1: *Dia ingin mengambil salah satu dari mereka*

T^T2: He takes **one** from it

From the sentences above, it can be seen that the pronoun as the subject was translated as *He* and *dia*. Therefore, these translated words were accurate.

The researcher concluded that the translation is a high-accuracy translation or an accurate translation. The meanings of words, technical terms, phrases, clauses, and source language sentences (ST) are accurately transferred to the target languages (TT); there was no distortion of meaning and nothing was deleted. In Indonesian, there is no categorization of masculine or feminine in the third person. However, English is masculine in the third person and Google translated the word correctly.

ii. Less Accurate Translation

Less accurate is defined as most of the meanings of words, technical terms, phrases, clauses, sentences or the source language text have been transferred accurately into the target language. However, there are still distortions in the meaning or translation of multiple meanings or the meaning is omitted, which disturbs the integrity of the message (Nababan, 2012: 50). Eight out of 72 data were described as inaccurate translations. Example of translation data:

Data 1 / II

ST: الْعَصْفُورُ

T^T1: *Burung gereja*

T^T2: **Sparrow**

The data above was repeated many times in the translation cases, it caused the stories told about the little bird. The noun الْعَصْفُورُ should mean little bird. However, the translation is *burung gereja* in Indonesian and/or sparrow in English. This translation is inaccurate. It disturbs the integrity of the message and there is a small distortion of meaning. At the same time, the عَبْدُ اللَّهِ word will be explained in an inaccurate translation.

The next example is in data 6/III

ST: فَصَاحَتْ لِعَصَافِيرُ خَوْفًا وَفَرَعًا

T^T1: *Burung-burung berteriak, dengan ketakutan dan teror.*

T^T2: The birds cried in fear and **panic**.

The data has distortion meaning in T^T1, it should be translated as *panik* as like in English as T^T2. It concluded that *google translation* still has

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: عَبْدُ اللَّهِ

T^T1: *Hamba Tuhan*

T^T2: **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: بَلْ أَخَذَ مِنْهَا وَجِدًا وَنَزَّ بِهِ

T^T1: *Ia mengambil satu darinya dan memasukkannya ke dalam.*

T^T2: But he took one from it and **took it** down

Ke dalam and **took it** was not the accurate translation from بَلْ. 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

Translation Category	Frequency	Percentage (%)
Acceptable	57	79
Less-Acceptable	3	4
Unacceptable	12	17
Total	72	100

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: شَجَرَةٌ عَالِيَةٌ

T^T1: *Pohon yang tinggi*

T^T2: a tall tree

The data mentioned is from 2 data but compiled in 1 phrase to be analyzed. The word شَجَرَةٌ is a feminine nominal category in Arabic and عَالِيَةٌ 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 شَجَرَةٌ عَالِيَةٌ 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: أَرَادَ أَنْ يَأْخُذَ وَاحِدًا مِنْهَا وَلَمَّا سَمِعَ صَوْتَهَا

T^T1: *Dan ketika dia mendengar **suaranya**, dia ingin mengambil **salah satu dari mereka***

T^T2: 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: عَبْدُ اللَّهِ

T^T1: *Abd Allah*

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: أَكُونُ فِي غَايَةِ الْحُزْنِ وَالْأَلَمِ

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: وَنَزَّ بِهِ

TT1: *Memasukkannya 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 وَنَزَّ بِهِ 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: إِلَيْهَا وَمَدَّ يَدَهُ

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

Translation Categories	Frequency	Percentage (%)
High Readability Translation	59	82
Moderate Readability Translation	0	0
Low Readability Translation	13	18
Total	72	100

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 Goggle Translation of the gender markers in the story *عَبْدُ اللَّهِ وَالْغُصْفُورُ* 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: *وَطَلَّبَ نَ الْخَادِمَ أَنْ يَرْجِعَ الْغُصْفُورَ إِلَى أَهْلِهِ*

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. *طَلَّبَ* is a verb with he as the hidden subject. *الْخَادِمَ* is a masculine noun. *الْغُصْفُورَ* is masculine nouns from animal species and hi in the word *أَهْلِهِ*, 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: *وَلَكِنْ وَلَكِنَّ قَلْبَهُ مَارَقَ بِحَالِهَا*

TT1: *Tetapi hatinya adalah apa yang melemparkan 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

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 عَبْدُ اللَّهِ وَالْغُصْفُورُ.

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: يَا عَبْدَ اللَّهِ

TT1: *Hai hamba Tuhan*

TT2: O servant of God.

Data 14/III

ST: مِنْ أَيْنَ جِئْتَ بِهِ

TT1: *Dari mana asalmu?*

TT2: From where **you** came **with it.**"

The gender markers data above were difficult to understand. First, عَبْدُ اللَّهِ as the name of the person should not be translated literally. Second, *bi* as the suffix in *bibi* 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

- Alawneh, M., Omar, N., & Sembok, T. M. (2011). Machine Translation from English to Arabic. *Engineering and Technology*, 11, 95–99.
- Catford, J. C. (1965). *A linguistic theory of translation*. London: Oxford University Press.
- Creswell, W. J. (2013). *Research design pendekatan kualitatif, kuantitatif, dan mixed (edisi ketiga)*. (Terjemahan Achmad Fawaidi). Yogyakarta: Pustaka Pelajar.
- Farghal, M. and Shunnaq. (1999) *A. Translation with reference to Arabic and English: A practical guide*. Irbid, Jordan: Dar Al-Hillal for Translation.
- Khoiriyah, H. (2020). Kualitas Hasil Terjemahan Google Translate Dari Bahasa Arab Ke Indonesian. *Al Mi'yar: Jurnal Ilmiah Pembelajaran Bahasa Arab Dan Kebahasaan*, 3(1), 127. <https://doi.org/10.35931/am.v3i1.205>
- Koponen, M. (2010). Assessing Machine Translation Quality with Error Analysis. *Electronic Proceedings of the VIII Kätu Symposium on Translation and Interpreting Studies*, 4, 1–12.
- Lips, M. Hillary. (1993). *Sex and Gender; An Introduction*. California, London, Toronto: Mayfield Publishing Company.
- Maslihah, R. E. (2018). Akurasi Penggunaan Translation Machine pada Penulisan Skripsi Mahasiswa. *Cendekia: Jurnal Kependidikan Dan Kemasyarakatan*, 16(2), 245. <https://doi.org/10.21154/cendekia.v16i2.1295>
- Muhammadun, Muzdalifah, (2016). Penanda Gender dalam Perspektif Bahasa Arab dan Indonesian. *Jurnal Al-Maiyyah, Volume 9 No. 1 Januari-Juni*
- Nababan, M., Nuraeni, A., & Sumardiono. (2012). Pengembangan model penilaian kualitas terjemahan. *Kajian Linguistik dan Sastra*, 24(1), 39–57.
- Najjar, O. and Shahin, S. (2015). The Translational Impact of Gender Sensitization on the Palestinian Society. *Arab World English Journal* 4, 253-268
- Nida, E. A., & Taber, C. B. (1982). *The theory and practice of translation (second photomechanical reprint)* (8th Edition). Netherlands: E.J. Brill