The Ability of English Education Department Students of University of Mataram to Conduct Post-Editing of Narrative Text for GNMT

Rani Dwi Hapsari[1*], Baharuddin[1], Lalu Ali Wardana[1], Santi Farmasari[1]

[1] English Education Department, University of Mataram, Mataram, Indonesia.

Abstract The emerging of the AI-generated Machine Translation or Neural Machine Translation has continued the role of translator to conduct adequate post-editing for the NMT. This research is aimed to analyze the ability of English Education Department Students of University of Mataram in conducting post-editing of narrative text for Google Translate, one of widely known NMTs. This research is qualitative refers to the information collected in one or more methods. The data was collected through CMC, they are the Interpreting and Translation class students' worksheet of post-edited narrative text which previously was translated by GT. The data then were analyzed using few theories Baker's textual equivalence theory, and harmonized metric of quality translation. Through the analyzation result, it could be inferred that most students are able to identify which sentences generated by GNMT that require post-editing and which that do not. Most of sentences left unedited by students are proven well-translated by the GNMT. Regardless, some sentences indeed still need to be post-edited after input to the GNMT, as it seems GNMT still struggles to translate well some phrases or terms that related to the culture of the language. Most students in the Translation and Interpreting class at the University of Mataram have proven their ability to identify and correct errors in GNMT-generated sentences through post-editing. This research is hoped to offer valuable insights for both education and translation and serves as a useful reference for readers seeking comprehension in this field.

Keywords: Google neural machine translation, Narrative text, Post-editing.

Abstrak: The emergence of AI-generated Machine Translation or Neural Machine Translation has continued the translator's role to do adequate post-editing for NMT. This study aims to analyze the ability of students of the Department of English Education at the University of Mataram in post-editing narrative texts for Google Translate, one of the widely known NMTs. This research is qualitative with a case study method. Data was collected through CMC, namely worksheets for Interpreting and Translation class students of post-edited narrative texts that had previously been translated by GT. The data were then analyzed using several Baker textual equivalence theories, and translation quality harmonization metrics. Through the results of the analysis, it can be concluded that most students are able to identify which sentences produced by GNMT require post-editing and which do not. Most of the sentences that were not edited by students proved to be well translated by GNMT. Apart from that, some sentences still need to be edited after input into GNMT, because it seems that GNMT still has difficulty properly translating some phrases or terms related to language culture. Most of the students in the Translation and Interpreting class at the University of Mataram have proven their ability to identify and correct errors in sentences produced by GNMT through post-editing. It is hoped that this research will offer

_____
Email: rani_hapsari@unram.ac.id

valuable insights for both education and translation and serve as a useful reference for readers seeking understanding in this field.

Kata Kunci: Google neural machine translation, Narrative text, Post-editing.

Introduction

Nowadays, the need for translation, stated by Ramadhan et al. (2021), from another language (Source Language) to other languages (Target Language) has been increasing widely. The translation process is crucial with the intention of the message conveyed or obtained can be comprehended properly. Along with the statement above, the mastery of translation skill also needs to be considered. Translation is one of certain issues which needs to be discussed inside the scope of foreign language learning (Sari et al., 2022). The act of translation can be done in various ways to produce translations result, one of them by using translation machine, and one of the translator machines widely used is Google Translate, this translator machine is also called as GT (Saputra et al., 2022).

Machine Translation (MT) is a technology created to help to quicken the process of translating as the demand for translation, as time goes by, keeps increasing. As for Statistical Machine Translation (SMT), The overall goal is to learn how to translate from a large corpus of equivalent source and target sentence pairs.

Henceforth, as a result, Neural Machine Translation (NMT) is proposed as a new machine translation paradigm (Cheng, 2019). Kalchbrenner and Blunsom proposed neural machine translation (NMT) in 2013, is the newest approach to machine translation (Irfan, 2017).

Unlike SMT which is only capable of translating the text based on its corpus; translating by breaking the sentences into phrases and translating the pieces and then trying to stitch the translations back together, NMT uses a neural network to consider the whole sentences when predicting the translations which make it possible to take into account the context of each word and phrase used.

However, neural machine translation, in this case, Google neural MT; the neural machine translation services offered by Google Company to provide translation services from source languages to target languages (Almira Zulaika et al., 2022), is not exceptionally perfect in doing its job as a translator. Since the words that it is able to learn the meaning of are depends on the frequency those words are being input, GNMT is ineffective in dealing with rare words and tends to fail to translate all words in the SL text (Wu et al., 2016).

Machine translation alone is not fully capable to produce a faithful translation. Argued by Wardana et al. (2022) Machine translation is not completely reliable to produce faithful translation as the act involves factor such as in difference of culture and insight. For that the results of Google Translate or machine translation results still need to be edited by human. The need of human touch in Google Translation resulted in the increasing demand in post-editing. Post-editing is, defined by Somers (cited in Tobing, 2019), tidying up the raw results, correcting mistakes, revising entire sections, or, in the worst case, retranslating entire sections.

Method

Research design is a detailed plan for collecting data in empirical research projects. Even in empirical research, it is considered a "blueprint" because its objective is to answer specific research questions or prove specific hypotheses (Bhattacherjee, 2012). It should also be noted that the purpose of research is crucial to determine the design and methodology of the study (Cohen et al., 2005). This research is a case study that examines the ability of students of the University of Mataram to conduct post-editing of narrative text for GNMT. The narrative text depicts a series of the acts or portrays a story, in other words narrative depicts an experience. The narrative's logical sequence of perception and words chronologically according to time (Yani et al., 2022). It uses a qualitative approach in which data from a translation and interpretation class were collected.

Although it is qualitative research, this research does not avoid quantifying when data is presented. Data were collected through observation and documentation. The data were in the form of the product of the post-edited narrative text of the students for GNMT, which were analyzed then. The data were students' worksheets of post-editing narrative text for GNMT given through assignment. The data were taken a form as the post-edited narrative text of students in class Translation and Interpreting for GNMT. The researcher then collected the worksheet of the students in doing post-editing of narrative text for GNMT as the data.

Qualitative method data collection refers to information collected in one or more methods (e.g., translation interviews, chaining, or capture of phenomena) and then analyzed("The Sage Encyclopedia of Qualitative Research Methods," 2008). In this

research, data were collected through observation, CMC and documentation. The data were taking the form of the post-editing narrative text of students in the translation and interpretation for the GNMT class.

It was collected through (1) observation, which was carried out in class B of the course of translation and interpretation. The researcher observed the learning activity related to the post-editing process. (2) The assignment supported by the CMC, while the student received the assignment post-edit after receiving the lesson of it. The CMC is used to help assign the task to the students in the class. (3) Documentation, while the researcher collected the paper work of the students in the post-editing of narrative text for GNMT as data.

Result and Discussion

Table 1. Finding Result

EDITED SENTENCES	ACCURACY	ADDITION	4%	28%
		OMISSION	4%	
		OVER TRANSLATION	6%	
		UNDER TRANSLATION	1%	
		MISTRANSLATION	7%	
	LOCALE CONVENTION		6%	
UNEDITED SENTENCES				72%

Stated in the table above, among total 340 sentences of the twenty students from Translation and Interpreting class, the total of the errors for the post-editing process for gnmt is 28%; where the addition 4%, omission 4%, over-translation 6%, under-translation 1%, mistranslation 7%, along with unedited sentences 72%.

a) **Edited Sentences**

In this discussion, the result of students' post-edit of narrative text for Google Translate (GT) are discussed and analyzed elaborately, the term ST is used to represent source text and ET to represent edited text.

I. Accuracy
  ▪ *Addition*
    ST   : The gods saw the tears of the farmer.
    GT   : Para dewa melihat air mata petani.
    ET   : Tanpa disadari para dewa melihat kesedihan sang petani.

In datum 1, the addition made by the editor was considered minor error —that would not lead to loss of the text purpose but brings slight change in meaning. The addition does not mislead the purpose of the text from the original which explained the minor category, it was assumed that the editor added the addition to give the text more dynamics.

The "tears" which the GNMT correctly translate into "*air mata*" in Bahasa Indonesia was replaced to "*kesedihan*" meaning "sadness" in English. It was explained as reiteration of lexical cohesion where Halliday and Hasan (1976, quoted in Baker (1992)) defined it as items of repetition of earlier item, synonym or near-synonym, a superordinate, or a general word.

  ▪ *Omission*
    ST   : Time went by they live together and got a beautiful baby boy.
    GT   : Waktu berlalu mereka hidup bersama dan dikaruniai seorang bayi laki-laki yang cantik.
    ET   : Waktu berlalu dan mereka dikaruniai seorang bayi laki-laki.

In datum 2, the editor omitted "they live together" from their version. It was assumed that the editor saw the phrase hold no significant effect to the purpose of the text which resulted in the omission. The omission then considered as neutral omission because of the insignificant change it made. The editor also decided to omit the term "beautiful" which the GNMT translate into "*cantik*" meaning beautiful/pretty (feminine) in Bahasa Indonesia. The lack of the GNMT ability to differentiate the feminine, masculine, or general words used in this translation for this proved the needs of human assistance to post-edit the machine-translated text. The omission ended up considered as neutral since the 'beauty' of the baby boy has no significant role to the story, hence, does not affect the objective purpose of the text/story.

    ST   : Because he felt tired of this happening the farmer decided to catch who should be responsible.

GT : Karena merasa lelah dengan kejadian ini, petani memutuskan untuk menangkap siapa yang harus bertanggung jawab.

ET : Petani tersebut memutuskan untuk mencari tahu siapa pelakunya.

In datum 3, the editor omitted the phrase 'because he felt tired of this happening' which resulted in minor error. The error was classified as minor since it omits the reason or the cause of why the farmer decides to find the culprit.

In addition to that, the editor replaced the phrase "who should be responsible" to "*pelakunya*" which translated as "the culprit" or "the perpetrator" in English where Baker (1992) defined the act as reiteration to create lexical cohesiveness. It was assumed to create more effective sentence as "who should be responsible" could be seen as ineffective one.

▪ *Over-translation*

ST : The farmer was alone and heart broken.

GT : Petani itu sendirian dan patah hati.

ET : Sang petani pun hidup sebatang kara dan dilema kesedihan.

In datum 4, the editor used "sebatang kara" as the translation for "alone", an Indonesian idiom meaning to live alone without any single family anymore. The editor also used "dilemma kesedihan" to translate "heartbroken" which could be translated as "dilemma of grief". Both the term "sebatang kara" and "dilemma kesedihan" are considered as subordinate terms or words thus explains the classification for the over-translation. Nevertheless, both are considered as neutral level of over-translation as they don't bring confusion to the overall story.

▪ *Under-translation*

S : A long time ago, there was a farmer. He had a small farm.

GT : Dahulu kala, ada seorang petani. Dia memiliki sebuah peternakan kecil.

ET : Pada zaman dahulu kala, hiduplah seorang petani yang memiliki sebuah ladang yang kecil.

In datum 5, the sentence is considered as minor under-translation since the editor translate the term "small farm" into "ladang yang kecil" which if it was to be back translated to English would be "small land" or "small field". The term used by the editor is seen to be superordinate and could lead to more questions that demand details.

▪ *Mistranslation*

S : Because he felt tired of this happening the farmer decided to catch who should be responsible.

GT : Karena merasa lelah dengan kejadian ini, petani memutuskan untuk menangkap siapa yang harus bertanggung jawab.

ET : Hal ini membuat petani itu jengkel sehingga ia memutuskan untuk menangkap pelaku yang bertanggung jawab atas kejadian itu.

In datum 6, the sentence is considered as over-translation in minor level and addition in neutral level. There are several factors as to why it is considered minor translation; the editor decided to replace "karena merasa lelah" (because he felt tired in English) with "hal ini membuat petani itu jengkel" (this thing made the farmer annoyed in English) which literally change the whole meaning of the phrase, but still considered doesn't affect the general purpose of the text as the phrase still deliver a general meaning of "the farmer disagree with the action occurred".

The next phrase of "to catch who should be responsible" also was slightly changed by the editor into "untuk menangkap pelaku yang bertanggung jawab atas kejadian itu", the terms "pelaku" and "atas kejadian itu" were added by the editor with assumption to make the text more cohesive and dynamic, thus made the sentence only minor over-translation and neutral addition as the changes don't affect the general purpose of the text.

▪ *Locale Convention*

ST : They lived happily for a long time, until one day, the star maiden found her magical wings near their house

GT : Mereka hidup bahagia untuk waktu yang lama, sampai suatu hari, gadis bintang menemukan sayap ajaibnya di dekat rumah mereka.

ET : Mereka hidup bahagia untuk waktu yang lama sampai suatu ketika,

bidadari tersebut menemukan selendang ajaibnya didekat rumah mereka.

In datum 7, the term "magical wings" translated by gnmt as "sayap ajaib" is replaced by "selendang ajaib", meaning magical scarf in English, in the post-edited version. It is believed that the term "selendang ajaib" is more suitable to use in the target language as it is a term that culturally widely known in the target language as a tool of magical creature to fly; so, the target reader could relate more with the stories as the local is culturally more familiar with the term. The term itself also become widely known for there is similar story or narrative in the target language to the source language called "Jaka Tarub dan Tujuh Bidadari" where in that story, the protagonist steal and hide one of the "selendang ajaib" which also was used to fly by the bidadari, the beautiful magical creature who usually has a form similar to female body, then ended up marrying her. The term "selendang ajaib" is seen as neutral level of locale convention as it still delivers the original objective of the term "magical wings".

b) **Unedited Sentences**

ST   : It was surprising him because there were three of star maidens there.

GT   : Itu mengejutkannya karena ada tiga gadis bintang di sana.

In datum 8, it is found that some students decide to use the term or the translation that the GNMT generated. Above is the example sentence that left unedited or untouched. Should we analyze the sentence in the source language and the one GNMT generated, it is found that the latter is well-translated. The GNMT is able to identify the reference for "him" in the source language and translated it to "nya" correctly. Though the term "star maiden" was translated literally and considered culturally less relatable with the local and preferrable to be edited, it is inferred that the one GNMT generated is still considerably accurate to the source language.

ST   : When they saw the farmer was approaching, they tried to escape.

GT   : Ketika mereka melihat petani itu mendekat, mereka mencoba melarikan diri.

In datum 9, also showed a sentence that left unedited by some students. Through the analyzation, it is found that the sentence generated by GNMT or Google Translation is profound already. Baker (1992) explained that reiteration for lexical cohesion is necessary to maintain dynamic in a text, hence the sentence generated by the GNMT is considerably dynamic as it is able to translate the term "the farmer" into "petani itu" and maintain the specification of which or what farmer that the source language is discussed. Thus, inferred the students' judgment to left the sentence unedited is justifiable.

SL   : A long time ago, there was a farmer.

GNMT   : Dahulu kala, ada seorang petani.

In datum 10, is shown the sentence that some of the students left unedited. This sentence generated by GNMT is well specified the term "a farmer" into "seorang petani". The term considered well-specified as the article "a" in English is generally used for both animate and inanimate object. In the other hand, in Indonesian the articles used for animate and inanimate object are different; e.g., "seekor, seorang" for animate object and "sebuah, sebongkah" for inanimate one.

The GNMT also well-translated the phrase "a long time ago" to "dahulu kala" instead of literally as in "sebuah panjang waktu lalu" proven the GNMT ability to translate sentence correctly unlike machine translation with SMT.

Conclusion

Based on the analyzation, it could be inferred that most students are able to identify which sentences generated by GNMT that require post-editing and which that do not. Most of sentences left unedited by students are proven well-translated by the GNMT. Regardless, some sentences indeed still need to be post-edited after input to the GNMT, as it seems GNMT still struggles to translate well some phrases or terms that related to the culture of the language.

The most of the students in the Translation and Interpreting class of English Education Department of University of Mataram, through their product of post-editing, are proven to have adequate ability to comprehend and identify errors in sentences generated by GNMT and later conduct post-editing for GNMT and produce sentences more suitable to the target language. It is anticipated that this research could bring more insight and use in both education and translation field. It is also expected that the research could become a valuable reference and help for the readers who seek comprehension or insight of this field.

Referensi

Almira Zulaika, B., Ali Wardana, L., & Farmasari, S. (2022). Students' Ability to Conduct Pre-Editing of Text Procedure for Google Neural Machine Translation. *Journal of Language*, 4(2), 173–183.

Baker, M. (1992). *In Other Words: A Coursebook on Translation* (3rd ed.). Rouletdge.

Bhattacherjee, A. (2012). Social Science Research: Principles, Methods, and Practices. In *Social Science Research: Principles, Methods, and Practices* (2nd ed., Vol. 2). Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License. http://scholarcommons.usf.edu/oa_textbooks http://scholarcommons.usf.edu/oa_textbooks/3

Cheng, Y. (2019). *Joint Training for Neural Machine Translation* [Tsinghua University]. http://www.springer.com/series/8790

Cohen, L., Manion, L., & Morisson, K. (2005). *Research Methods in Education 5th Edition* (5th ed.). RoutLedgeFalmer.

Irfan, M. (2017). *Machine Translation*. https://www.taus.net/academy/timelines/translation-

Rizki Ramadhan, N., Baharuddin, & Ali Wardana, L. (2021). An Analysis of Translation Method Used in the Novel Earth Translated by Gill Westaway. *LISDAYA: Jurnal Linguistik (Terapan), Sastra, Dan Budaya*, 17(2), 30–37.

Saputra, A., Sumiati, & Baharuddin. (2022). The Analysis of Google Translate Accuracy in Translating Procedural and Narrative Text. *JEEF (Journal of English Education Forum)*, 2(1), 7–11.

Sari, N., Arifuddin, & Baharuddin. (2022). The Equivalence in The Translation of English Idiomatic Expression into Indonesian by Students of English Education Department University of Mataram. *Culturalistics: Journal of Cultural, Literary, and Linguistic Studies*, 6(1), 48–58. http://ejournal.undip.ac.id/index.php/culturalistics

The Sage Encyclopedia of Qualitative Research Methods. (2008). In L. M. Given (Ed.), *Qualitative Research Methods* (Vols. 1–2). Sage Publication Inc.

Tobing, M. (2019). *Post-Editing Process of Machine Translation: A Case Study of Google Translate in Local Government Website*.

Wardana, L. A., Baharuddin, B., & Nurtaat, L. (2022). Kemampuan Mahasiswa melakukan post-editing terhadap Hasil Terjemahan Machine Translation. *Jurnal Ilmiah Profesi Pendidikan*, 7(1), 53–61. https://doi.org/10.29303/jipp.v7i1.392

Wu, Y., Schuster, M., Chen, Z., Le, Q. V., Norouzi, M., Macherey, W., Krikun, M., Cao, Y., Gao, Q., Macherey, K., Klingner, J., Shah, A., Johnson, M., Liu, X., Kaiser, Ł., Gouws, S., Kato, Y., Kudo, T., Kazawa, H., … Dean, J. (2016). *Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation*. http://arxiv.org/abs/1609.08144

Yani, S., Ezir, E., Daulay, I. K., & Manugeren, M. (2022). Teaching Narrative Text Through MindMapping Technique. *Journal of Language*, 4(1), 73–82.