



ENGLISH DEPARTMENT STUDENTS' ABILITY IN CONDUCTING POST-EDITING OF IDIOMATIC EXPRESSIONS IN LITERARY TEXT OF YANDEX NMT

Hani Noviani¹, Baharuddin², Lalu Ali Wardana³

^{1 2 3} English Education Department, Faculty of Teacher Training and Education,
University of Mataram, Indonesia

*Corresponding Author: Haninoviani1599@gmail.com

Abstract: Many translators have used machine translation to get rapid translation output and achieve excellent outcomes. One of the current challenges for Machine Translation is handling significant texts, such as literature, that include sarcasm, metaphor, irony, idioms, and other ambiguous elements of language. Many previous studies have shown that Neural Machine Translation can impact translation productivity when followed by post-editing. To achieve high-quality output, the students use post-editing. In this study, the researcher purposed to find out students' ability in conducting post-editing of idiomatic expression in literary text input to the Yandex NMT. The participants of this research consisted of students of the English Education Program, Faculty of Teacher Training and Education, Mataram University, who were taking "Translation and Interpreting" in the fifth semester in Academic Year 2021/2022. This data was collected from assignments conducted by 30 students in 5TI-1. The supporting data were gathered through observation by entering and sitting in on the class. The data were analyzed using a short story with the title "The Theft of the Monalisa." The results of this study showed that the quality of the text produced by the students who conducted post-editing was acceptable. The results of post-editing can still be understood by the readers, even though there are small translation errors with irrelevant information. Thus, the more time and effort put into post-editing the text, the more likely it is that Yandex NMT will produce higher quality translations. The abstract is written in a single paragraph and single-spaced and consists of brief introduction, objective(s), methods, significant findings, and conclusion. It is no more than 250 words.

Keywords: Yandex Neural Machine Translation; Post-editing; Literary Text; Idiomatic Expression

INTRODUCTION

Machine translation is the translation performed by a machine using a formula embedded in a program to aid translation results. The machine fully translates after receiving input in the form of text from the source language into the target language (Wardana et al., 2022). However, machine translation has a weakness. From the start, there is skepticism about the capacity of machine translation to translate things that smell, taste, and live up to humans' knowledge of the terms, different cultures, and ways of life for those who speak two or more languages (Baharuddin et al., 2022). One of the current challenges for machine translation (MT) is handling significant texts, such as literature, that include sarcasm, metaphor, irony, idioms, and other ambiguous elements of language that are likely to result in a word-by-word translation. In other circumstances, it might be challenging to translate a message due to the language style, cultural differences, and the purpose of the speaker (Sari et al., 2022).

Many researchers in the field of translation have discovered a possibility of translating more texts efficiently with Neural Machine Translation systems appearance. Neural Machine Translation (NMT) is an approach for machine translation. There are several types of Neural Machine

Translation. One of them is Yandex Translate. Yandex (formerly Yandex.Translation) is a Yandex-provided web service designed to translate text or web pages into other languages. Yandex uses a statistical method for its machine translation. Statistically based machine translation has recently received much attention in the scientific community as it uses translated text sources. Many previous studies have shown that NMT can significantly impact translation productivity when followed by post-editing. Post-editing NMT output for literature will increase productivity for literary translators (Guerberof-Arenas & Toral, 2020). On the other hand, professional translators may identify the incorrect meanings and structures generated by NMT and take appropriate post-editing procedures to eliminate such mistakes. Based on the descriptions of MT issues and post-editing provided above, research on post-editing is still a relatively new subject in translation (Cetiner & Isisag, 2019).

Moreover, in the Indonesian context, little empirical evidence has been revealed regarding undergraduate students' classroom practice in activities in both English-Indonesian and Indonesian-English language pairs. (Latifah et al., 2022). Literary translators must deal with the creativity of authors to achieve the same level of quality as the original text. Idiomatic expressions are one of the most interesting and innovative aspects to study. Idioms have their characteristics and patterns as a language form, and they are frequently used in both written and spoken communication since they can convey a richness of linguistics (Wang & Wang, 2013). Not all idioms can be translated into idioms in the source language since idioms are set words or expressions with specific meanings and forms. Therefore, a translator must understand the true meaning of idioms in the source language while translating them. Moreover, to understand the meanings of idioms, people must understand the background cultures of both the source and target languages.

Based on the primary issues mentioned above, the researcher needs a more in-depth study to examine students in conducting post-editing in translating text inputs using YandexTranslation. This is the reason the researcher needs to investigate it to conduct this study with the name "English Department Students' Ability in Conducting Post Editing of Idiomatic Expressions in Literary Text of Yandex Neural Machine Translation."

In this study, an attempt was made to find an answer to the following question: "How is it possible for students to change idiomatic expressions in literary text to resemble the original literary text?" The purpose of this investigation is to determine whether or not students are capable of engaging in post-editing of material prior to its use in Yandex Translate. This study is also to see the progress of MT and how it may be utilized to help students learn and communicate ideas in English with the help of that machine. Specifically, this is something that will be looked at in this study.

RESEARCH METHODS

In this study, a qualitative descriptive method will be used. This descriptive research method is applied in this study to examine, review, and draw conclusions on the impact of post-editing on translating activities. This study fits qualitative research as Creswell (2012) claims that qualitative research identifies a problem, develops a detailed understanding of a central phenomenon, and gathers information based on words from a small number of individuals. The descriptive research is appropriate to be conducted since the aim of this research is to find out the English Department students' ability in conducting post-editing of idiomatic expressions in literary text output. The data is presented descriptively based on observation and documents.

Data for this study will be gathered through observation and assignment. Firstly, this study uses observation as the data collection. This study applies participatory observation to gather study data through observation and sensing while engaging in the activities of students. Then, students in the class are given direct research assignments, which provide the data for the needs of this study. The task is not specifically designed for research reasons. The student's work will be checked and graded for the lecture's requirement and evaluated for research purposes.

According to Miles et al. (2014), the steps of analysis are data condensation, data display, and conclusion drawing/verification. In the condensation stage, several sections of the acquired data are created based on the condition of the data that suit the translation theories described in the framework. The data is carefully analyzed and applied to help in answering the research question. The display of data is second step is data analysis. Looking at the displays will help the researcher to comprehend what is happening and will do something for further analysis for understanding. Sentences from text that has previously completed editing from the original source material are further discussed in the study. The framework consisted of a collection of translation theories created for this research was used to understand and explain the categorize research data in the discussion. Following the presentation of the data then, the conclusion is drawn.

FINDINGS AND DISCUSSION

The process of gathering data is observed through classroom observation. The students are motivated to learn and active in trying to conduct post-editing. The student is keen in learning and active in attempting to conduct post-editing of a lecturer-provided text. The researcher identifies the challenges that students face when studying the "Translation and Interpreting" course with the post-editing material that has been used. When learning is clear, there is positive and active interaction between the students and the lecturer.

In student learning activities, it is necessary for students to understand the text before they conduct post-editing by correcting components like grammar, word choice, punctuation, subject/object, and identifying the vocabulary to produce acceptable and appropriate sentences on the text. In order to create an appropriate target text in Yandex Neural Machine Translation. It is crucial to practice comparing the vocabulary between the source text and the target text that is appropriate to the content of the text. Trying to match the vocabulary between the source text and the target text therefore requires practice. YandexTranslate needs the use of a dictionary to help determine the vocabulary or an appropriate phrase to produce an appropriate target language.

The data was gathered through the observation of the classroom. In the class, the process of identifying the sentences were observed. The students put their new translation of sentence one by one by writing on the telegram group that made for this purpose of the study. From previous data collected through condensation step, the suitable data will be taken for discussion. There is one short story with title "The Thief of the Mona Lisa" consists of 406 words and 12 idioms expressions. The modification between the source text and the raw text that occurred after post-editing is already categorized in the data used for discussion. It produces text targets that can be analyzed using suitable translation theories.

After the lecturer gave the text during learning in the classroom. The participants conducted post-editing through light post-editing. Since the participants did the tasks including grammatical and usage errors. Not only that, but the participants also produced outcomes with the poor quality of idiomatic expressions. The process of gathering data is observed through classroom observation. The students are motivated to learn and active in trying to conduct post-editing. The student is keen

in learning and active in attempting to conduct post-editing of a lecturer-provided text. The researcher identifies the challenges that students face when studying the "Translation and Interpreting" course with the post-editing material that has been used. When learning is clear, there is positive and active interaction between the students and the lecturer.

Table 4.1 An example of text that produces the quality of the with good criteria in conducting post-editing

Source Text	Yandex Translate	Edited Translation
The police even interviewed Peruggia twice, but he must have mixed them up because they said he couldn't possibly be the man behind the burglary. The sixty detectives that were on the case had failed to find the thief.	Polisi bahkan mewawancarai Peruggia dua kali, tetapi dia pasti mencampuradukkan mereka karena mereka mengatakan dia tidak mungkin menjadi orang di balik perampokan itu. Enam puluh detektif yang ada dalam kasus ini gagal menemukan pencuri itu.	Polisi bahkan mewawancarai Peruggia dua kali, tetapi Peruggia selalu membuat polisi kebingungan dan mengatakan dia tidak mungkin terlibat dalam pencurian tersebut. Enam puluh detektif yang ada dalam kasus ini gagal menemukan pencuri lukisan itu.

In sentence “*mixed them up*” translated into *mencampuradukkan mereka*. This idiom has meaning make them confuse. It means Yandex Translate did not produce the result well. However, the participant put the correct and exact meaning in the edited Translation which meaning “*membuat polisi kebingunan*” and when it translates to English become “*make them confuse.*” Here, we can identify that participant using idioms of similar meaning and form translating this English idiom. The replacement of the word “*bulgary*” from “*perampokan*” into “*pencurian*” is related to the naturalness and acceptability of translation in the target language. Besides that, the article “*the*” has been changed from “*itu*” to “*tersebut.*” This case does not reduce the exact meaning and message from the text. The editor also added “*pencuri lukisan itu*” instead of “*menemukan pencuri itu*” thus, it makes the target results seem more effective and clearer. Based on the sentence in the edited result, the changes obtained in post-editing text by making improvements such as word choice in were successful in producing the appropriate outputs.

Table 4.2 An example of text that produces the quality of the with good criteria in conducting post-editing

Source Text	Yandex Translate	Edited Translation
He couldn't believe it that the police would twist his words, and call him crazy and sentence him to seven months jail.	Dia tidak bisa percaya bahwa polisi akan memutar kata-katanya, dan memanggilnya crazy and hukuman dia tujuh bulan penjara	Dia tidak bisa percaya bahwa polisi akan memutar kata-katanya, dan memanggilnya crazy dan kalimat hukuman dia tujuh bulan penjara.

In the sentence “*He couldn't believe it that the police would twist his words.*” The idiom is “*twist the word*” is translated to *memutar kata-katanya*. However, it is easier to be comprehend if the editor put “*memutar-balikan kata-kata.*” But both will be understood because it is familiar to each other. Here, we can identify that the participant used more neutral word because it is the closest to equivalent for the word *memutar kata-katanya*. In the next sentence “sentence him to seven months jail”, the participant produced post-editing results poorly and sounds awkward. Meanwhile YandexTranslate produced more natural with the sentence “hukuman dia tujuh bulan penjara.” In addition, some other words in the target language are not good choice to put in since it makes the readers confused of the message of the text.

Table 4.3 An example of text that produces the quality of the with acceptable criteria in conducting post-editing

Source Text	Yandex Translate	Edited Translation
It was signed "Leonardo." The man was of course Peruggia. Peruggia, who had a lot of Italians in his family tree, claimed to have stolen the artwork to return her to her native Italy.	Pria itu tentu saja Peruggia. Peruggia, yang memiliki banyak orang Italia di pohon keluarganya, mengaku telah mencuri karya seni untuk mengembalikannya ke Italia asalnya	Pria itu tentu saja Peruggia. Peruggia, yang memiliki banyak orang Italia di pohon keluarganya, mengaku telah mencuri karya seni untuk mengembalikannya ke Italia daerah asalnya.

In the English idiomatic expression “*family tree*” translated “*pohon keluarga.*” This idiom has meaning *silsilah keluarga*. The result sounds unnatural. Moreover, the participant did not edit and still put the result of YandexTranslate into edited translation. Here, the participant and YandexTranslate did not find the accurate translation. Thus, the result translated has poor quality. In the sentence “native Italy” the results post-editing in edited translation is clear and appropriate translated from the results of YandexTranslate. It makes easier to be comprehended by reader with adding the word “Italia daerah asalnya” instead of “Italia asalnya.” It means that Italia is a place.

Table 4.4 An example of text that produces the quality of the with acceptable criteria in conducting post-editing

Source Text	Yandex Translate	Edited Translation
For the first time there were lines outside the Louvre, just to see the empty space where the painting had hung. X marks the spot where the painting used to hang. The theft launched it into becoming a household name for people who had no interest in art.	Untuk pertama kalinya ada garis di luar Louvre, hanya untuk melihat ruang kosong tempat lukisan itu digantung. X menandai tempat di mana lukisan itu digunakan untuk menggantung. Pencurian itu meluncurkannya menjadi nama rumah tangga bagi	Berita hangat dan menggegerkan ini membuat orang-orang antri untuk pertama kali di Louvre hanya karena ingin melihat tanda X di dinding semula Mona Lisa tergantung. Dengan ini, Vincenzo Peruggia dengan tidak sengaja berhasil menumbuhkan rasa penasaran dan ketertarikan

	orang-orang yang tidak tertarik pada seni.	orang awam terhadap seni khususnya untuk lukisan.
--	--	---

The idiom has been changed and combined with another sentence into one. In the first sentence, the word and the structure have been changed. “*Untuk pertama kalinya ada garis di luar Louvre, hanya untuk melihat ruang kosong tempat lukisan itu digantung,*” has changed to “*Berita hangat dan menggegerkan ini membuat orang-orang antri untuk pertama kali di Louvre hanya karena ingin melihat tanda X di dinding semula Mona Lisa tergantung.*” However, the meaning is acceptable because it does not reduce the messages of the sentence. By changing and replacing with another words make the sentence is easier to be understood by the reader. The participant arranged the structure differently and combining two sentences into one sentence. Thus, it makes better in well-formedness. Also, the last sentence has been changed by adding “*dengan*” or conjunction.

Table 4.5 An example of text that produces the quality of the with nonsense criteria in conducting post-editing

Source Text	Yandex Translate	Edited Translation
He thought that he would <u>get up on a soap box</u> and be a national hero.	Dia berpikir bahwa dia akan bangun di kotak sabun dan menjadi pahlawan nasional..	Dia berpikir bahwa dia akan bangun di kotak sabun dan menjadi pahlawan nasional.

This part contains one idiom. The idiom reads “*get up on a soap box.*” Both Yandex Translation and Edited Translation are poor translation. The meaning of this idiom is stop telling what to think but the editor did not edit from machine translation correctly. The category of this idiom is pure idiom which means that it is impossible to decode the meaning by combining the meaning of their parts.

Table 4.6 An example of text that produces the quality of the with nonsense criteria in conducting post-editing

Source Text	Yandex Translate	Edited Translation
It was 24 hours before anyone even noticed the Mona Lisa was missing. The French newspapers <u>made war</u> with the government which ran the Louvre. The image of the Mona Lisa was splashed across the world.	Itu 24 jam sebelum ada yang menyadari Mona Lisa hilang. Surat kabar Prancis berperang dengan pemerintah yang menjalankan Louvre. Gambar Mona Lisa terciprat ke seluruh dunia.	Mona Lisa telah hilang, 24 jam setelah kejadian pihak museum baru menyadari lukisan itu sudah lenyap dari dinding museum. Surat kabar Perancis melakukan pemberitaan besar-besaran terhadap tidak kompetennya pihak manajemen Museum Louvre.

“*War*” is translated as “*berperang*” in the sentence. The editor did not include the outcome of the Yandex Translate translation. Moreover, the editor changed the idioms' meaning in the edited translation. Here, we can see the participant has changed and created a different meaning. It can be

said that machine translation produced better results in this case. The edited translation from Yandex translation results differed significantly. For example, "It was 24 hours before anyone even noticed the Mona Lisa was missing." The French newspapers made war with the government that ran the Louvre. "*Mona Lisa telah hilang, 24 jam setelah kejadian pihak museum baru menyadari lukisan itu sudah lenyap dari dinding museum,*" the participant said. The participant also edited the sentence by adding comma punctuation marks in the text.

According to the findings above, the quality of the text produced by the students who conducted post-editing can result in acceptable text. Since the original text's message can probably still be understood, even though there are a few small translation errors with irrelevant information, the basic idea of the original text can still be understood despite significant parts of the original text being missing or poorly translated. The results of the analysis and interpretation above show the abilities of each participant when post-editing a literary text, yet some of them still require practice. Thus, they will be able to produce results that are correct. Some of the texts that had previously undergone post-editing were excellent. However, they were still unable to obtain the correct meanings from the source text in the target text, which may be the reason for Yandex Translation's weaknesses. Due to this, post-editing abilities should be improved in order to produce better and more accurate results while using YandexTranslate.

CONCLUSION

The findings indicate that most students have the necessary skills to perform post-editing on the source material after it has been entered into YandexTranslate. In addition, for them to be able to fulfill the demand to produce better text, they will need to improve their capacity for post-editing. When measured with Baker's idea of equivalence above the word level, the ability is still at a lesser level than it was before. Most post-editing results are regarded as satisfactory. When it comes to interpreting idioms, the students should have a better awareness of the ways in which Indonesian and English differ from one another. Students should be familiar with and have a good understanding of those categories of information since idioms play a key role in the post-editing process.

REFERENCES

- Allen, J. (2003). Post-editing. *Benjamins Translation Library*, 35, 297-318.
- Angelone, Erik, Maureen Ehrensberger-Dow, and Gary Massey. 2019. *The Bloomsbury Companion to Language Industry Studies*. London - New York - Oxford - New Delhi - Sydney: Bloomsbury Academic.
- Anggrina, B., Pramudita, K. E., & Suparmi, S. (2017). EFL LEARNERS' POST-EDITING ON GOOGLE ENGLISH-INDONESIAN TRANSLATION OUTPUT. *Proceedings of ISELT FBS Universitas Negeri Padang*, 5, 130-137.
- Baharuddin, B., Amin, M., Thohir, L., & Wardana, L. A. (2021). Penerapan Teori Terjemahan pada Editing Hasil Terjemahan Google Translate pada Teks Akademik oleh Mahasiswa Universitas Mataram. *Jurnal Ilmiah Profesi Pendidikan*, 6(4), 816-824.
- Baker, M. (2018). *In other words: A coursebook on translation*. Routledge.
- Çetiner, C., & İşısağ, K. U. (2019). Undergraduate level translation students' attitudes towards machine translation post-editing training. *International Journal of Languages' Education and Teaching*, 7(1), 110-120.

- Creswell, J. W. (2012). Educational research: planning. *Conducting, and Evaluating*, 260, 375-382.
- Guerberof-Arenas, A., & Toral, A. (2020). The impact of post-editing and machine translation on creativity and reading experience. *Translation Spaces*, 9(2), 255-282.
- Latifah, N. W., Baharuddin, B., & Udin, U. (2022). An Analysis of Translation Shift in Novel Shine by Jessica Jung and Its Translation. *Culturalistics: Journal of Cultural, Literary, and Linguistic Studies*, 6(2), 11-17.
- Lommel, A., Görög, A., Melby, A., Uszkoreit, H., Burchardt, A., & Popović, M. (2015). Harmonised metric. Project Report, QT21 project (funded by the European Union's Horizon 2020 program for ICT) Retrieved from <https://www.taus.net/evaluate/dqf-tools# error-typology>.
- O'Brien, S., Balling, L. W., Carl, M., Simard, M., & Specia, L. (2014). Post-editing of Machine Translation. *Processes and Applications*. Newcastle-upon-Tyne: Cambridge Scholars Publishing.
- Sari, N., Arifuddin, A., & Baharuddin, B. (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.
- Wang, L., & Wang, S. (2013). A study of idiom translation strategies between English and Chinese. *Theory and practice in language studies*, 3(9), 1691.
- 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.

