






Post-editing Performance of English-Major Undergraduates in China: A Case Study of C-E Translation with Pedagogical Reflections

Zhan Yu  and Ying Lu  

NingboTech University, No.1 South Qianhu Road, Ningbo, China
luying@nbt.edu.cn

Abstract. Nowadays, post-editors are in great demand in China. Therefore, this research aims to study the post-editing performance of English-major undergraduates from NingboTech University. The results indicate that: Overall, the post-editing ability of English-major undergraduates is still not sufficient for professional post-editing tasks; Their post-editing performance is related to the type of texts, and they have better performance in post-editing informative texts, while they are relatively weaker in the post-editing of expressive and vocative texts; Their post-editing quality is related to their dependence on machine translation, and the group with higher post-editing quality has a relatively lower average dependence on machine translation, but no proportional correlation is found between the two; The errors in students' post-editing versions can be mainly classified into language competence-related errors and translation competence-related errors. Based on the results, several pedagogical implications for post-editing teaching in the future are discussed.

Keywords: Machine Translation · Post-editing · Student Translators · Translation Pedagogy

1 Introduction

Nowadays, the process of internationalization is gradually deepening, and translation has become an increasingly important carrier of communication between different cultures. But the number of translators cannot meet the huge demand for translation, so machine translation emerged. However, many problems appeared while using machine systems for translation as machine-translated texts are still far from publishable quality except in some narrow domains [1]. Therefore, correction by human is necessary to make machine translation output more understandable and accurate [2]. This has led to the currently heated issue of post-editing. But the research on post-editing in China is still in its initial stage, focusing mainly on its introduction and application [3], and the teaching of machine translation and post-editing is still a weak research field [4]. Consequently, the cultivation of students for more job-oriented applications is highly necessary at the moment.

By carrying out an empirical study on the English-major undergraduates in NingboTech University, this study will focus on English-major undergraduates' performance in post-editing in the Chinese-English language pair. On this basis, this study will analyze the current deficiencies in post-editing teaching and put forward some pedagogical implications and improvement measures.

2 Literature Review

Nowadays, in order to balance productivity and quality in translation and to give full play to the advantages of human-computer interaction, the model of machine translation plus post-editing has been widely used [5]. Due to the improvement of machine translation technology, the growing market demand for translation, and the cost of human resources, post-editing will play an increasingly important role in the language service industry and translation teaching field [3]. Therefore, the combination of machine translation and human post-editing has its significance both theoretically and practically. So far, many studies have already been done in the field of post-editing, mainly focusing on its productivity, quality, and cognitive efforts. The main results will be introduced in the following paragraphs.

Many researchers have focused on productivity gain resulting from post-editing. As concluded by Plitt and Masselot, machine translation post-editing helps translators improve their throughput by 74% on average, saving 43% of their time [6]. Robert thinks that post-editing can increase translators' productivity from an average of about 2,000 words per day to about 3,500 words per day, thus contributing to an increase of about 30,000 more words per month [7]. Guerberof conducted an experiment including eight professional translators and witnessed 13%–25% of productivity gains compared with human translation [8]. However, the experiment conducted by Garcia showed that the productivity gains in the process of post-editing were marginal [9]. Therefore, at least for the moment no general conclusion can be drawn about the productivity of post-editing compared with the traditional translation process.

In contrast, studies about the quality of post-editing present rather similar results. The number of errors in machine translation is greater than in human translation in five out of eight cases in Guerberof's experiment [8]. The study carried out by Garcia in 2010 turned out that translations produced by editing machine translations were more favored in 59% of the cases [10]. Her study in 2011 further suggested that translating by post-editing was more advantageous regardless of the difficulty of the text and the capability of participants [9]. What's more, post-edited versions are of higher clarity and accuracy [11]. In a word, post-editing can improve translation quality, although the degree and aspect of its impact vary.

Apart from productivity and quality, there are also many studies about cognitive efforts in the post-editing process. O'Brien points out that pauses can indeed indicate cognitive processing in post-editing [12]. Koglin focuses on the cognitive efforts of post-editing metaphors in newspaper texts and finds that it is lower compared with manual translation [13]. Also, research has shown that post-editing can decrease the cognitive efforts in understanding source text and producing translation [14]. Although further studies are still needed in this field, it can be tentatively concluded that post-editing can relieve the cognitive burden in the translation process.

Although much research has been done about post-editing, much research gap still exists in this field. Specifically speaking, current research in this field is mainly about the translations in the same language family, especially the Indo-European language family, with little research about language pairs in different language families. What's more, the participants in these studies are mostly professional translators or postgraduates in translation major. In addition, the vast majority of them use Google Translate in the research as machine translation system, while studies concerned with Baidu Translate, which was widely used in China, are relatively few. This research will try to fill in this research gap, and the specific method will be introduced in the next section.

3 Methodology

3.1 Research Goals

The main purpose of this study is to gain some insights into the following question: What are the performances of English-major undergraduates in post-editing tasks? The sub-questions include: 1) How do English-major undergraduates perform in post-editing tasks currently? 2) What is the impact of text types on their post-editing performances? 3) Are the post-editing performances of English-major undergraduates directly related to their dependence on machine-translated text?

3.2 Participants

The participants of this study are 95 English-major undergraduates from NingboTech University in Zhejiang, China, all of whom were enrolled in 2019. They are all native Chinese speakers, with Chinese as their first language and English as their second language. They have already taken one semester of translation theories and practice lesson in which the basic principles and methods of translation are introduced, and they have taken a business English translation course. None of them have any professional training in post-editing or any work experience as post-editors. All of their personal information and performance are only used for this study and are strictly confidential.

3.3 Material

The material of this study was selected from the 2019 Social Responsibility Report of Geely Holding Group, a famous automobile manufacturer in Zhejiang, China. As Geely is a listed company on the Hong Kong Stock Exchange, it has real needs for communications in both Chinese and English, so its social responsibility report has versions in these two languages and their quality can be guaranteed. Based on Newmark's theory of text types [15], three texts (Text A, Text B, and Text C) were selected from the whole report, which were expressive, informative, and vocative respectively. The texts were abridged and modified, with brand names removed to prevent students from searching the original text directly on the Internet. Moreover, in order to reduce the effect of the difference between text difficulty on the results, they are similar in length, containing 175, 174, and 175 words respectively. After that, the three texts were pre-translated by Baidu Translate from Chinese to English (December 2021). Then they were

pasted into a Word file in a whole passage, along with the source text in Chinese. It is worth mentioning that the three paragraphs are not separate in the file and the participants are not informed of the types of each paragraph. As a matter of fact, the three texts were arranged in proper order to look like a coherent profile text of the company. This is to ensure that participants maintain the same habits and steps while post-editing the three different types of texts.

3.4 Evaluation of Translation Quality

In order to ensure the accuracy and reliability of the result, this study uses a two-dimensional evaluation method to evaluate the post-editing versions of the participants.

The first dimension is the BLEU score. It is a method of automatic translation quality evaluation proposed by IBM in 2002. The central idea of it is that the closer a machine translation is to a professional human translation, the better it is. A BLEU score ranges from 0 to 1, and it will be closer to 1 if the translation has higher quality and vice versa [16]. Due to its advantages of high speed, low cost and objectivity, BLEU has been used in translation quality assessment by many researchers [6, 17–19]. This study uses the natural language toolkit in Python to calculate the BLEU scores of each participant's post-edited version. The human translation used for reference is the official English version of the Social Responsibility Report, with some necessary modifications being made to ensure its quality. The BLEU score of every participant's translation as well as the three texts in each passage was calculated separately.

The other dimension is the marks produced by human grading. It is based on Pym's classification of errors proposed in 1992 which divides the translation errors into two basic forms: binary error and non-binary error. The former one refers to any error that is an incorrect translation, while the latter one refers to a translation which is not completely wrong, but may not be appropriate enough and should be further improved [20]. The specific grading method refers to the method adopted by Lee and Liao, which assesses students' post-editing quality in terms of sentences: 2 points will be deducted if a binary error occurs in one sentence, and 1 point will be deducted for non-binary errors. The maximum point loss would be 2 points for each sentence. As there are a total of 11 sentences in this material, the total points of the passage are 22, and every participant would get a mark ranging from 0 to 22 [21]. In order to make the evaluation more objective, the error analysis on the official website of Pigai will be referred to.

After getting the BLEU scores and the human-assessed marks, the final scores can be drawn by combining these two dimensions. To make the result easier to be analyzed, the final score adopts a 100-points system, in which the two dimensions both account for 50 points. The points of these two parts will be converted into the 50-point system and then be added together. Therefore, the participants' final scores can be got according to the following formula:

$$\text{Final score} = \text{Human - assessed mark} / \text{Full mark} * 50 + \text{BLEU} * 50 \quad (1)$$

In this way, the participants' post-editing quality can be evaluated from two aspects: the BLEU shows their similarity with the reference, indicating how "right" the translations are; while the human-scored marks take the errors they made into consideration, indicating how "wrong" they are. By using this two-dimensional evaluation method

incorporating both machine and human, the evaluation of post-editing quality would be more objective and reasonable.

3.5 Research Procedure

After determining the participants and the material of the research, the material was distributed to the participants on the official website of Pigai (www.pigai.org), a frequently used website for English writing and translation in China. Then the students were required to use the editing function in Word to post-edit the material. Considering that English majors do not have post-editing experience and need sufficient time to complete the post-editing task, this study does not record participants' time in performing this task. Students have one week to work on this task and have to submit their translations before the deadline (January 20, 2022). After completing it, they were asked to hand in their translations through the website of Pigai in the form of attachment so that they can be downloaded in the original version for analysis. Then the quality of their translations will be evaluated. After that, in each translation, the number of unchanged words borrowed from the machine-translated text will be counted. The percentage of the unchanged words in the post-editing version will be calculated in order to look into participants' dependence on the original machine output. After that, the errors in participants' translations will be categorized and analyzed by human, and some pedagogical implications can be drawn on this basis. The results will be discussed in the following section.

4 Results and Discussion

4.1 Overall Results

A total of 95 translations were received in this study, but 2 of them were not completed according to the requirements and 3 were found to have used machine translations, which have to be excluded. Altogether 90 effective samples were collected for analysis. For

Table 1. Overall results of students' post-editing

	Machine Translation	Students' Average
Unchanged words in machine output	/	236.44
Word count	342	346.49
Percentage of unchanged words	/	68.44%
Binary errors	2	2.7111
Non-binary errors	7	4.1444
Human-assessed mark/Full mark	11/22	12.43/22
BLEU	0.6677	0.6668
Final score	58.38	61.60

each of them, the percentage of unchanged words in machine output was calculated, and the post-editing quality was evaluated by the method introduced in the previous section. Table 1 shows the overall results of the participants' post-editing versions.

The average length of the edited texts are 346.49 words, with little difference compared with the machine output. The percentage of unchanged words is 68.44% on average, indicating that around one third of the machine-translated text was edited by the participants. As for their post-editing quality, they got about 60% of the full score in all of the three indicators (BLEU, human-assessed marks, and final scores), indicating that the average quality of students' post-editing work is approximately the pass level but is far from satisfying. It has also been found that most of the BLEU scores students got are similar, mostly ranging from 0.6 and 0.7, with rather small differences. This is probably due to the fact that all their post-edited versions are based on the same machine output. What's more, the non-binary errors they made are about twice as many as the binary errors. It has to be made clear that in the evaluation process, a maximum of two points would be deducted in one sentence to avoid the influence of some extreme situations. It is noteworthy that although the human-assessed mark the final score of students' translation are higher than the machine translation, their BLEU scores are slightly lower compared with that of the machine output, which may be because that the BLEU is normally used to evaluate the quality of machine translation. In addition, students made obviously less non-binary errors than machine while the frequency of their binary errors increases. The following three sections will present a detailed analysis of the results.

4.2 Impact of Text Types on Post-editing Performance

The average BLEU, human-assessed marks, and final scores of each type of text are counted to show students' post-editing performance of different text types. The results are listed in Table 2.

Table 2. Post-editing performance of different text types

	Text A	Text B	Text C	Passage
Binary errors	0.62	0.87	1.22	2.71
Non-binary errors	2.58	0.18	1.39	4.14
BLEU	0.6040	0.6170	0.5474	0.6668
Mark/Full mark	4.18/8	4.09/6	4.17/8	12.43/22
Final score	56.31	64.93	53.41	61.60

From the results we can see that judged by all of these three indicators, students have better post-editing performance in Text B, namely informative text. While their scores in the post-editing quality of expressive and vocative texts are relatively lower. The ratio between binary and non-binary errors also differs between the three text types. The occurrence of non-binary errors in post-editing expressive text is more than four times compared with binary errors, indicating that most of their errors in this text type are not totally wrong, but still have space for improvement. As for vocative text, the

frequency of these two kinds of errors is almost the same, while binary errors appear much more than non-binary errors in the post-editing of informative text. The reason for this result might be that the expressive and vocative texts focus more on the orators and readers, contributing to their diversification of translation, thus increasing the difficulty of translation and post-editing. Contrarily, informative texts focus more on facts and reality, calling for fewer translation skills.

In short, different text types do have an impact on students' post-editing performance. They have a more satisfying performance in post-editing informative text, while their performance in post-editing expressive and vocative text is relatively inferior. It is worth mentioning that as texts are diverse, not every text can be categorized into these three types and some texts may be divided into more than one types. Therefore, this conclusion may not be absolutely right in every context.

4.3 Correlation Between Students' Dependence on Machine Translation Output and Their Post-editing Quality

To study the correlation between the participants' dependence on machine-translated text and their post-editing quality, the students are divided into six groups according to three indicators of post-editing quality applied in this study: BLEU score (Group 1 and Group 2), human-assessed mark (Group 3 and Group 4) and final score (Group 5 and Group 6). The criteria for the division is the median number of each indicator. For example, the median of their marks is 12, and students whose marks are 12 and above will be divided into Group 1, and those below 12 form Group 2. For each group, their average percentage of unchanged words will be counted to show their dependence on machine output. The results are listed in Table 3.

Table 3. Groups of different post-editing quality and their dependence on machine translation

Criteria for grouping	Group	Percentage of Unchanged Words in Machine Output
BLEU	Group 1 (higher BLEU)	65.54%
	Group 2 (lower BLEU)	71.33%
Human-assessed marks	Group 3 (higher marks)	64.63%
	Group 4 (lower marks)	73.41%
Final scores	Group 5 (higher scores)	63.25%
	Group 6 (lower scores)	73.63%

As can be seen from this table, whether students are divided according to their BLEU scores, human-assessed marks, or final scores, the group with relatively better translation quality has lower dependence on machine translation. This proves that the quality of post-editing is negatively correlated with the dependence on machine translation. What's more, the average percentage of the higher quality groups (Group 1, 3, and 5) and the lower quality groups (Group 2, 4, and 6) are 64.47% and 72.79% respectively, which is

higher than the results in Lee and Liao's study (58.5% for the group from the prestigious university and 66.3% for the group from a graduate institute) [21] and lower than the results in Yamada's study (69% for the pass group and 79.9% for the fail group) [22].

On the whole, it can be concluded that participants who perform better in this task make more editions in their post-editing process while the group of students with relatively inferior post-editing performance use more words from the original machine-translated text. This result corresponds to the results of the two studies mentioned above.

In order to further probe into the relation between the dependence on machine translation and post-editing, three scatter diagrams (Fig. 1, 2, and 3) are drawn using Excel, whose horizontal axes show a participants' percentage of unchanged words in the machine-translated text and the vertical axes show his or her post-editing performance. These three figures demonstrate the relation between students' dependence on machine translation and their BLEU scores, human-assessed marks, and final scores respectively.

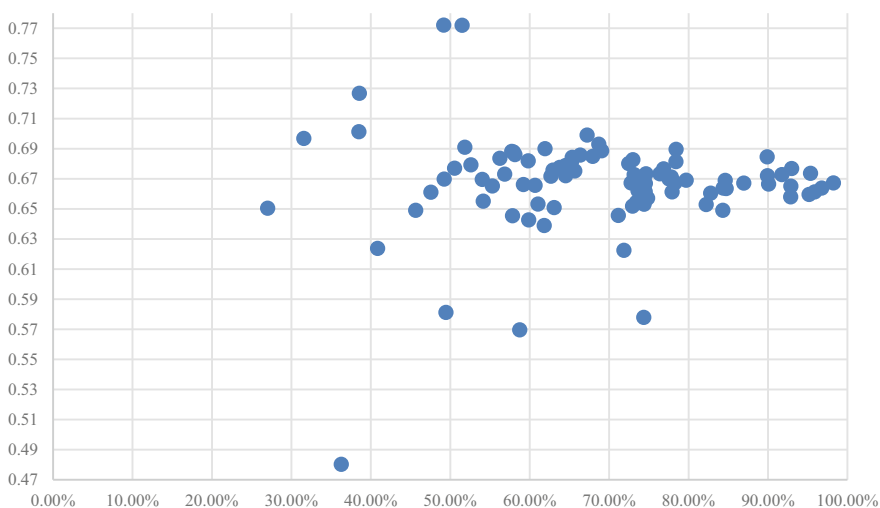


Fig. 1. Participants' percentage of unchanged words and their BLEU

From these three diagrams, we can see that participants' scores and their dependence on machine-translated texts are not proportional. A lower percentage of unchanged words in translation doesn't necessarily lead to a higher score and vice versa. Therefore, no proportional relation can be found between these two variables at least in this study. There are some possible reasons for this situation. Firstly, as we all know, translation and post-editing quality can be affected by many factors and therefore cannot be simply attributed to these two factors. In addition, in some cases in this study, the machine output is already correct without the need for modification, but the participant may mistakenly believe that the sentence has to be changed, thus making some unnecessary mistakes.

To sum up, in this study a correlation does exist between participants' dependence on the machine-translated text and their post-editing quality, but no firm conclusion can yet be drawn on this issue since no proportional relation is found between the two factors.

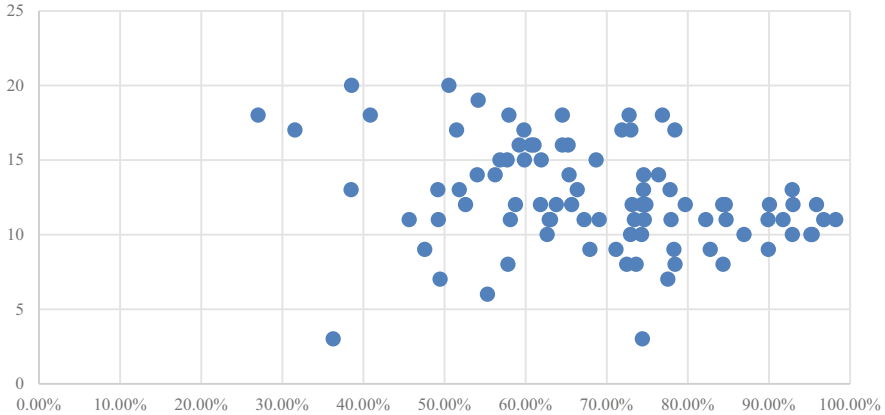


Fig. 2. Participants' percentage of unchanged words and their marks

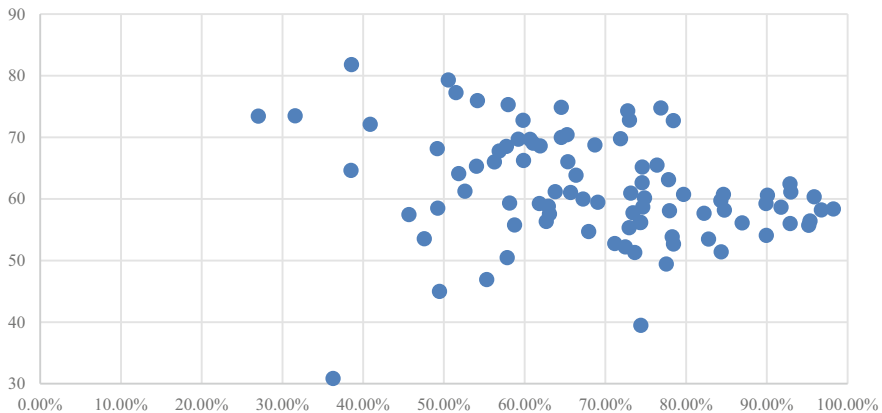


Fig. 3. Participants' percentage of unchanged words and their final score

4.4 Error Analysis

In order to get a better understanding of the errors students made in the post-editing process, the types of errors and their occurrence are analyzed. By referring to the classification of translation errors in some of the current studies [18, 23–25], this study divided the translation errors of participants into two big categories based on the root cause of the errors: language competence-related mistakes, which are caused by insufficient command of language, and translation competence-related mistakes, which are caused by inadequacy of translation capability. The former one means that the translation itself, as a piece of text, is not correct, even without considering the source text. While the latter one means that the text is not translated in the way that presents equivalent meaning and function of the original text to the readers. Based on the specific situation of the errors, this study further divided the former category into 7 small categories and the latter one into 6 small categories, with a total of 13 kinds of errors. It has to be made clear that

due to the particularity of the materials and the participants, this error analysis is only for this study and is not general. The following table lists the number and percentage of each type of errors in this study (Table 4).

Table 4. Error Analysis

	Type of Error	Number	Percentage	
Language competence-related mistakes	Logical confusion	58	9.08%	52.43%
	Error of verb form	128	20.03%	
	Error of preposition	17	2.66%	
	Miscollocation of words	111	17.37%	
	Spelling mistake	3	0.47%	
	Error of sentence structure	16	2.50%	
	Ambiguity	2	0.31%	
Translation competence-related mistakes	Redundancy	32	5.01%	47.57%
	Mistranslation of proper noun	36	5.63%	
	Word-for-word translation	217	33.96%	
	Omission	12	1.88%	
	Under-translation	4	0.63%	
	Over-translation	3	0.47%	
TOTAL		639	100%	

In general, mistakes related to students' language competence are slightly more than those caused by deficiency in translation skills. In language competence-related mistakes, error of verb form accounts for the largest part, followed by miscollocation of words and logical confusions. As for translation competence-related mistakes, word-for-word translation is the most frequent one, accounting for more than one third of the total mistakes, while mistranslation of proper noun and redundancy come second and third.

It is worth mentioning that some errors made by participants in this research can be attributed to more than one reasons and some of the mistakes made by students have as much to do with their translation skills as their language abilities. Moreover, it has also been found that frequent occurrence of language competence related mistakes doesn't always lead to a large number of mistakes related to translation competence, which disagrees with the long-held view that a good command of language abilities is the premise of translation skills. It is often the case in this study that translations with many grammatical errors also have many brilliant sentences, and the specific circumstances and reasons for this are left for further research. Due to the constraint of the space, each error type will not be explained in detail.

5 Pedagogical Implications

Based on the research results, there are several pedagogical implications for improving students' post-editing abilities and teaching of post-editing in the future.

5.1 Curriculum Provision

It is advisable that the curriculum of translation talents cultivation be divided into two parts: foreign language abilities and translation competence. As is shown from the research results, the mistakes students made in post-editing are caused by the deficiency of both language and translation competence. Therefore, these two aspects of the curriculum are needed and should be taught in a more targeted way. To be more specific, in the first and second years, courses related to basic foreign language abilities should be paid more attention to. During the third and fourth years, courses including translation theories and practice, computer-aided translation, and post-editing can be incorporated to further improve their competence in translation. Of course, these two aspects should not be separated and should reinforce each other as a whole. In this way, the building of students' translation capacities can be more comprehensive and effective.

5.2 Practice of Different Text Types

This study has proved that students have different post-editing performances when encountering different types of texts. Therefore, the practice of different types of texts should be included in the post-editing teaching. When selecting materials, teachers are suggested to take expressive, informative, and vocative texts into consideration. Moreover, the practice of these three types should follow a reasonable order, namely from easy to difficult. Teachers should let students post-edit informational texts at first and gradually teach them the post-editing skills of expressive text and vocative texts. Attention should also be paid to the diversification of theme and context, helping students become more qualified for the actual demand.

5.3 Content of the Courses

In post-editing teaching, apart from the basic principles and methods of it, the introduction of the common types of machine translation errors and their corresponding solutions should also be included. If translators can understand different error types of machine translation, it will help them to locate the errors in machine translation more quickly and accurately, thus improving the efficiency of their post-editing [4]. This can be carried out by encouraging students to analyze the mistakes in machine translation themselves instead of telling them the characteristics of machine translation directly. On this basis, they are more likely to learn more quickly during their post-editing process.

5.4 Market-Oriented

This study has shown that English-major undergraduates' post-edited versions still have a rather big gap compared with the professional version. What's more, post-editing has

not yet become an independent lesson for English-major students. As a result, students have little access to this important part in the language service industry. Therefore, if students can be provided with the opportunities to learn more about the career demand of post-editors and even do an internship in the language service industry, they are sure to gain a lot more.

In a word, with the rapid development of technologies and quick upgrade of translation software, translation teaching nowadays should also make adaptations in order to keep up with the time. The large demand for qualified post-editors and the imperfect post-editing performance of English-major undergraduates call for more attention and further improvements in this field. Hopefully these suggestions can be helpful to the translation and post-editing pedagogy in the future to some extent.

6 Conclusion

By conducting a case study on 95 students from NingboTech University, this research finds that: 1) Current performances of English-major undergraduates in post-editing are still not proficient enough to meet the requirements of competent post-editors; 2) Text types have an influence on English-major undergraduates' performance in post-editing. Students have more satisfactory performances in post-editing informative texts while their performances in post-editing expressive text and vocative text are relatively inferior; 3) Students' post-editing performances are related to their dependence on machine-translated text, but no proportional relation is found in this research; 4) The errors students made during the post-editing tasks can be mainly divided into language competence related errors and translation competence related skills. On this basis, some pedagogical implications are proposed concerning curriculum provision, material for teaching, content of courses and orientation.

References

1. Koponen, M.: Is machine translation post-editing worth the effort? A survey of research into post-editing and effort. *J. Special. Transl.* **25**, 131–148 (2016)
2. Tatsumi, M.: *Post-Editing Machine Translated Text in a Commercial Setting: Observation and Statistical Analysis*. Dublin City University, Dublin (2010)
3. Feng, Q.G., Cui, Q.L.: Research focuses and trends in post-editing of machine translation. *Shanghai J. Transl.* **06**, 67–74+89+94 (2016). <https://doi.org/10.3969/j.issn.1672-9358.2016.06.012>
4. Zhao, T.: Current situation and problems of machine translation post-editing. *Foreign Lang. Educ.* **42**(04), 100–104 (2021). <https://doi.org/10.16362/j.cnki.cn61-1023/h.2021.04.017>
5. Wang, X.L., Lai, S., Jia, Y.F.: A comparative study of HT and NMT post-editing: data from eye-tracking and key-logging on metaphor translation. *Foreign Lang. Learn. Theory Pract.* **04**, 115–126 (2021)
6. Plitt, M., Masselot, F.: A productivity test of statistical machine translation post-editing in a typical localisation context. *Prague Bull. Math. Linguist.* **93**, 7–16 (2010). <https://doi.org/10.2478/v10108-010-0010-x>
7. Robert, A.M.: Vous avez dit post-éditrice? Quelques éléments d'un parcours personnel. *J. Special. Transl.* **19**, 29–40 (2013)

8. Guerberof, A.: Productivity and quality in MT post-editing. *MT Summit XII* **39**, 137–144 (2009)
9. Garcia, I.: Translating by post-editing: is it the way forward? *Mach. Transl.* **25**, 217–237 (2011)
10. Garcia, I.: Is machine translation ready yet? *Target* **22**, 7–21 (2010)
11. Fiederer, R., O'Brien, S.: Quality and machine translation: a realistic objective? *J. Special. Transl.* **11**, 52–74 (2009)
12. O'Brien, S.: Pauses as indicators of cognitive effort in post-editing machine translation output. *Across Lang. Cult.* **7**(1), 1–21 (2006). <https://doi.org/10.1556/Acr.7.2006.1.1>
13. Koglin, A.: An empirical investigation of cognitive effort required to post-edit machine translated metaphors compared to the translation of metaphors. *Transl. Interpret.* **7**, 126–141 (2015)
14. Lu, Z., Sun, J.: An eye-tracking study of cognitive processing in human translation and post-editing. *Foreign Lang. Teach. Res.* **50**(05), 760–769+801 (2018)
15. Newmark, P.: A textbook of translation. Shanghai Foreign Language Education Press, Shanghai (2001). <https://doi.org/10.1177/026009358904000310>
16. Papineni, K., Roukos, S., Ward, T., Zhu, W.J.: BLEU: A method for automatic evaluation of machine translation. In: *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*, pp. 311–318. Association for Computational Linguistics, Philadelphia (2002). <https://doi.org/10.3115/1073083.1073135>
17. Lagarda, A., Ortiz-Martínez, D., Daniel, A., Alabau, V., Casacuberta, F.: Translating without in-domain corpus: machine translation post-editing with online learning techniques. *Comput. Speech Lang.* **32**(1), 109–134 (2015). <https://doi.org/10.1016/j.csl.2014.10.004>
18. Munkova, D., Hajek, P., Munk, M., Skalka, J.: Evaluation of machine translation quality through the metrics of error rate and accuracy. *Procedia Comput. Sci.* **171**, 1327–1336 (2020). <https://doi.org/10.1016/j.procs.2020.04.142>
19. Guo, W.H., Hu, F.M.: A study of the assessment of translations and post-editing in neural machine translation. *J. Beijing Int. Stud. Univ.* **43**(05), 66–82 (2021)
20. Pym, A.: Translation error analysis and the interface with language teaching. In: *Teaching Translation and Interpreting: Training, Talent and Experience*, pp. 279–288. John Benjamin's Publishing Company, Amsterdam and Philadelphia (1992). <https://doi.org/10.1075/z.56.42pym>
21. Lee, J., Liao, P.: A comparative study of human translation and machine translation with post-editing. *Compil. Transl. Rev.* **4**(2), 105–149 (2011). <https://doi.org/10.29912/CTR.201109.0005>
22. Yamada, M.: Can college students be post-editors? An investigation into employing language learners in machine translation plus post-editing settings. *Mach. Transl.* **29**(1), 49–67 (2014). <https://doi.org/10.1007/s10590-014-9167-7>
23. Li, M., Zhu, X.M.: Error patterns and statistical analyses of an English–Chinese machine translation corpus. *J. Univ. Shanghai Sci. Technol. (Soc. Sci. Edn.)* **35**(03), 201–207 (2013). <https://doi.org/10.13256/j.cnki.jusst.sse.2013.03.004>
24. Cui, Q.L., Li, W.: The character of error types of post-editing: perspective of machine translation based on scientific and technological materials. *Chin. Sci. Technol. Transl. J.* **28**(04), 19–22 (2015). <https://doi.org/10.16024/j.cnki.issn1002-0489.2015.04.007>
25. Wang, X.L., Wang, T.T.: A comparative study of human translation and machine translation post-editing in E–C translation: translation speed, quality and translators' attitude. *Foreign Lang. Cult.* **3**(04), 83–93 (2019). <https://doi.org/10.19967/j.cnki.flc.2019.04.009>