

Toward Crowdsourcing Translation Post-editing: A Thematic Systematic Review

Marziyeh Khalilizadeh Ganjalikhani¹, Akbar Hesabi², Saeed Ketabi³

¹Ph.D. Candidate, Department of English Language and Literature, Faculty of Foreign Languages, University of Isfahan, Isfahan, Iran, Email: marziyeh.khalilizadeh@fgn.ui.ac.ir

²Corresponding author, Associate Professor, Department of English Language and Literature, Faculty of Foreign Languages, University of Isfahan, Isfahan, Iran, Email: a.hesabi@fgn.ui.ac.ir

³Associate Professor, Department of English Language and Literature, Faculty of Foreign Languages, University of Isfahan, Isfahan, Iran, Email: ketabi@fgn.ui.ac.ir

Abstract

Crowdsourcing Translation as a Post-Editing Method (CTPE) has emerged as a rapid and inexpensive method for translation and has drawn significant attention in recent years. This qualitative study aims to analyze and synthesize the approaches and aspects underpinning CTPE research and to identify its potential that is yet to be discovered. Through a systematic literature review focused on empirical papers, we examined the limited literature thematically and identified recurring central themes. Our review reveals that the topic of CTPE requires further attention and that its potential benefits are yet to be fully discovered. We discuss the eight core concepts that emerged during our analysis, including the purpose of CTPE, CTPE areas of application, ongoing CTPE processes, platform and crowd characteristics, motivation, CTPE domains, and future perspectives. By highlighting the strengths of CTPE, we conclude that it has the potential to be a highly effective translation method in various domains.

Keywords: crowdsourcing translation, human translation, machine translation, post-editing, systematic review

Received: May 10, 2022

Revised: October 21, 2022

Accepted: December 26, 2022

Article type: Research Article

DOI: 10.22111/IJALS.2023.46015.2358

Publisher: University of Sistan and Baluchestan

© The Author(s).



How to cite: Khalilizadeh Ganjalikhani, M., Hesabi, A., Ketabi, S. (2023). Toward crowdsourcing translation post-editing: A thematic systematic review. *Iranian Journal of Applied Language Studies*, 15(2), 1-18.

<https://doi.org/10.22111/IJALS.2023.46015.2358>

1. Introduction

Translation is a crucial aspect of communication that enables people to understand content in different languages. In recent years, the volume of material that needs to be translated has increased significantly due to globalization and the growth of digital content (Azari et al., 2020). Translation service providers are facing increasing pressure to deliver high-quality translations within tight deadlines and at a reasonable cost. The first thing that might come to mind as a faster and cheaper solution for translating a massive volume of material is MT (Machine Translation). However, time and cost constraints, along with the nature of the materials to be translated, have led to the rise of another model of translation known as crowdsourcing translation (CT). There are several definitions for CT (Crowdsourcing Translation), but it generally refers to a willing volunteer community or crowd-translating online content through a platform with low or no payment (Anastasiou & Gupta, 2011). Jiménez-Crespo (2017) differentiates between crowdsourcing and similar models like online collaborative translation, also referred to as volunteer translation. However, it is similar to other translation forms, such as citizen translation, community translation, and UGT (User-generated Translation), in that they all take advantage of volunteers regardless of whether they are experts in translation or linguistics (Flanagan, 2016; Federici & Cadwell, 2018).

CT is becoming a popular translation method and replacing the traditional models today because saving money and time while providing translation services has made the service providers seek fast and inexpensive models. The cost-effectiveness, diversified operation process, speed, and flexibility of CT fulfill these requirements (Anastasiou & Gupta, 2011). However, one primary concern regarding CT is translation quality (Yan et al., 2015). Maintaining the crowdsourced translation quality has been addressed through various innovative methods indicating no standard framework defined for the task. The lack of such fixed frameworks might be due to the diverse and flexible nature of CT, allowing different employers, organizations, and service providers to set their own rules regarding translation quality. As a rising trend, CT benefits from the multilingual abilities of the internet users of a large virtual crowd (O'Hagan, 2021) and is applied in various environments according to the involved factors and intended objectives. Localization practice might be one of the most noted applications of CT, i.e., the localization of Facebook (Jiménez-Crespo, 2013). It is also used to manage the extended amount of data produced nowadays by various organizations (Sutherlin, 2013). Another area of CT application is collecting extensive annotated data for Natural language processing (Gao et al., 2015). A relatively less paid attention domain is post-editing, as seen in the work of Aikawa et al. (2012).

Since it is often necessary to ensure that translations are accurate and comprehensible, the translation industry has turned to post-editing to deliver quality translations at an affordable price. PE (post-editing) improves the quality and efficiency of the translation compared to expert translations (Guerberof, 2009; Plitt & Masselot, 2010), and it traditionally refers to human post-

editing of MT output. However, when it comes to human participation, it can also be a crowd. As stated, having a large amount of material translated using traditional methods is not practical these days. However, crowdsourcing translation post-editing contributes significantly to those seeking fast and economical translation simultaneously. Crowdsourcing has been integrated with MT since the late 1990s, used to enhance and train MT systems and to post-edit their output (Shimohata et al., 2001). As Jiménez-Crespo (2017) points out, crowd post-editing is implemented in a few different areas. As a prominent example, CTPE is found in Bing Translator and Google Translate, which enables users to post-edit the translation outputs having two purposes: helping users to achieve their desired quality level and training the machine translation system.

Different scholars have implemented crowdsourcing translation as a post-editing method through platforms. In other words, they have designed a platform that provides the participants/crowdsourcers with MT output and allows them to post-edit it. It should be noted that they might be different in terms of their details and processes (Aikawa et al., 2012; Tatsumi et al., 2012; Mitchell et al., 2014; Mitchell, 2015); another area of CTPE implementation is the evaluation of the MT output quality (Bentivogli et al., 2011; Goto et al., 2016). CTPE can also be applied to evaluating the performance of MT systems (Goto et al., 2016). As an emerging and dynamic phenomenon, CTPE's potential is yet to be discovered since there has not been a comprehensive understanding of this method (Guazzini et al., 2015). CT has a long way to go to be accepted as a standard method for post-editing and to reach its maturity level. To address this gap in the literature, this systematic thematic review aims to identify the key elements and approaches underpinning CT as a post-editing method and to explore its potential.

2. Research Question

What are the approaches and aspects underpinning Crowdsourcing Translation as a Post-Editing Method (CTPE) research and its potential?

3. Review Method

The piece is a systematic review with thematic analysis based on the Technology Acceptance Model (TAM) which is widely used for understanding users' acceptance and adoption of new technologies (Holden & Karsh, 2010). TAM posits that users' acceptance of a technology depends on their perceived usefulness and ease of use. Perceived usefulness refers to the extent to which users believe that a specific technology will improve their performance or productivity, while perceived ease of use refers to the extent to which users believe that this specific technology is easy to learn and use. The study of CT as a post-editing method is closely related to the acceptance of new technology, and the application of TAM in this review will provide insights into the factors that influence users' acceptance and adoption of this method.

As for all systematic reviews, this study also contains a research question and follows purposive sampling due to its eligibility criteria which will be explained. To select the collection of studies to answer the research question, the researchers used three academic databases: Scopus, Web of Science, and Google Scholar. The Scopus database provides comprehensive coverage of many disciplines of scholarly literature. Web of Science is also a citation database that provides access to references and citations from journals across various academic fields. Finally, Google Scholar, a free web search engine, enables users to find full text or metadata for scholarly publications across various disciplines and publishing formats. In the selection process, the aim was to include as many studies as possible; therefore, we used various keywords in our search and used different word combinations related to crowdsourcing translation: “crowdsourcing translation” “post-editing” OR “post-editing”, “collaborative translation” “post-editing” OR “post-editing”, “community translation” “post edit” OR “post-edit” “crowd translation” “post-edit” OR “post-edit,” “crowd post-edit” OR “crowd post edit.” After identifying keywords, their possible synonyms, and terms, a total of 138 papers were found. After eliminating the duplicate studies and book chapters, we reviewed, skimmed, and scanned 51 studies. A total of 14 studies were eventually selected that met the eligibility criteria. Eligibility criteria were established to ensure that the studies included in this review are relevant to the research question and meet the specific requirements of the review. In other words, Eligibility criteria refer to a set of specific conditions or requirements that a study or publication must meet to be included in this review. The inclusion and exclusion criteria of this systematic review were created based on the purpose of the research question. Among the studies addressing crowdsourcing as a post-editing method, only those meeting the criteria were selected and included in this review.

We had four criteria in mind to include the studies in this review. 1) We included the studies investigating CTPE's current state, meaning they were addressing domains that were applying this method; 2) We included the studies applying CTPE in their career regardless of their purpose because it can be applied in different areas; 3) We included the studies involving CTPE either for MT and HT (Human Translation) because in general sense post-editing is only used for MT outputs; 4) We included the studies with full online texts available to make better judgments. The exclusion criteria were, firstly, those studies only pointed to CTPE as a strategy or potential solution but did not focus on or use it; secondly, the studies that only described the CTPE process; in other words, we excluded the studies that were not empirical; and lastly, the studies that only included post-editing but not crowdsourcing and vice versa.

To be systemic and objective, we included many factors regarding the selected studies in three tables. The first table included factors such as titles, author/s, year of publication, journals, name of conference, and contexts. The second table included information such as instruments, research designs, main aims, results, and significant findings. The third table was used to classify studies with similar workflow and application fields. After extracting the data, the synthesizing process began, through which making comparisons and detecting core concepts and key issues

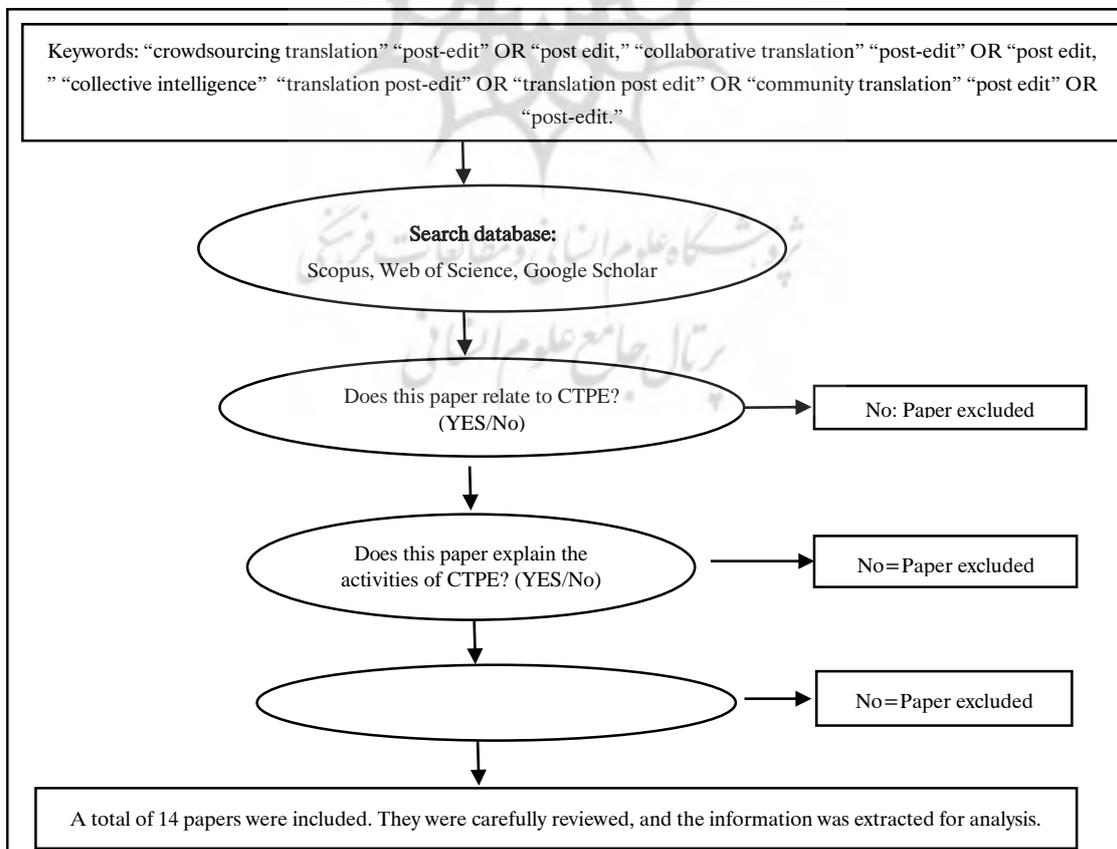
followed. In the next phase, we started organizing and writing down the findings. This paper aimed to outline the landscape of CTPE research, so there was no publication date limitation; however, our search ended on 27 September 2022. It is not surprising that this review contains a small number of studies because CTPE is still new and has a long way to go. Since notable works in this field of study were presented in workshops and conferences, the researchers included empirical conference papers.

- 1) To summarize, the following steps were taken in this systematic review:
- 2) We chose the keywords, synonyms, and terms and wrote the research question
- 3) We specified the eligibility criteria
- 4) We searched online databases: Scopus, Web of Science, and Google Scholar.
- 5) We selected the eligible studies.
- 6) We extracted the data from the selected studies.
- 7) We drew tables containing information about those studies to make the analysis easier.
- 8) We started synthesizing the needed information according to the extracted data
- 9) We organized our findings and wrote them down.

The review process is presented in Fig. 1.

Figure 1

Review Process



4. Findings and Discussion

This review includes only empirical papers, 1 of them was a journal article, and the rest were conference papers or studies presented in workshops. In 2008, the first relative study was published, but since 2012, the topic has attracted attention as the frequency of publications has increased. Based on the author's country affiliation, 9 country affiliations were identified. The United States tops the list with 5 pieces, while the UK follows with three, followed by Japan, Australia, Singapore, India, Spain, and France, each with one. The first author's affiliation was considered in those works with more than one author. The brief review presented above revealed eight fundamental concepts that emerged during the analysis of the included papers. The following concepts will be discussed:

- 1) The purpose of crowdsourcing translation as a post-editing (CTPE) method,
- 2) The areas of application of CTPE
- 3) Ongoing CTPE processes
- 4) Platform Characteristics
- 5) Crowd characteristics
- 6) Motivation,
- 7) CTPE domains
- 8) Future perspectives

It is important to note that these concepts are not present simultaneously in all papers. Depending on the research question and objectives of a study, a single paper may address several of these concepts, while another may focus on only one or two.

4.1. Purpose of CTPE

With the exponential growth of user-generated content on the internet, crowdsourcing translation as a post-editing method (CTPE) has emerged as a promising solution (Jimenez-Crespo, 2017). The three primary drivers for adopting CTPE are time, cost, and quality, although the relative importance of these factors varies depending on the specific needs and objectives of individual studies. Therefore, these factors are not always equally prioritized in all studies. This study will provide an in-depth analysis of these factors.

4.1.1. Time

The importance of time management in any project or task cannot be overstated, as failure to complete work by a set deadline can result in an unsuccessful outcome (Javadi Khasraghi & Mohammadi, 2012). This is particularly relevant for corporations, organizations, and other enterprises with multilingual users who need to provide data promptly. In response to this challenge, CTPE has been introduced as a new method that can help save time in the translation

process. Our review found that CTPE has been successful in reducing the time required for translation in more than half of the studies reviewed. For example, Aikawa et al. (2012) reported that they were able to localize most of Toyohashi University of Technology's English websites into nine languages in just two months using CTPE. Similarly, other studies have reported that CTPE has saved time in the translation process (Huynh et al., 2008; Muntés-Mulero et al., 2013; Laurenzi et al., 2013; Pavlick et al., 2014). The time savings offered by CTPE can be attributed to various factors. For instance, CTPE allows for parallel processing, where multiple translators can work on the same project simultaneously, which can significantly reduce the time required for translation (Laurenzi et al., 2013). Additionally, the use of machine translation (MT) in the CTPE process can help reduce the time required for translation, as the initial translation output can be quickly reviewed and edited by human translators (Pavlick et al., 2014). Overall, the time-saving benefits of CTPE make it an attractive option for corporations, organizations, and other enterprises with multilingual users who need to provide data in a short time.

4.1.2. Cost

One of the biggest challenges faced by organizations when it comes to translating large amounts of data is the high cost of hiring professional translators. This is particularly true for organizations that operate globally and deal with multiple languages. As a result, many organizations are constantly seeking cost-effective strategies to translate their documents. Several studies have investigated the cost-effectiveness of crowdsourcing translation for post-editing (CTPE) and have found that it has significantly reduced translation costs (Aikawa et al., 2012; Muntés-Mulero et al., 2013; Laurenzi et al., 2013; Yan et al., 2014; Dew et al., 2015). All of these studies reported that cost reduction was one of their primary objectives in applying the CTPE method, and they all found satisfactory outcomes. While several studies have investigated the cost-effectiveness of crowdsourcing translation for post-editing (CTPE), it is important to note that the scope of these studies can vary. As a result, some studies may not explicitly mention cost-effectiveness as a primary objective or may not report on it at all. For example, in the review that included Laurenzi et al. 2013 and Soares et al. 2020, the scope of the studies differed, and cost-effectiveness was not explicitly mentioned in some of them. This means that while these studies have reported on other aspects of CTPE.

4.1.3. Quality

There is considerable concern regarding the translation quality in crowdsourcing translation (Pascoal et al., 2017) and the same concern exists about CTPE because it follows the same procedure. For instance, Tatsumi et al. (2012) explicitly mention that quality control is a severe challenge in crowd post-editing and seek solutions to control the quality of the crowdsourced translations. Similarly, Yan et al. (2014) and Han et al. (2017) point out that we

must apply quality control measures if we expect the crowdsourcing output to work well. Nevertheless, it is not easy to simultaneously assess quality uniformly in crowdsourcing environments because they take advantage of innovative and dynamic methods, and there is no fixed rule or norm applied in crowdsourced translation conditions (Jiménez-Crespo, 2017), and consequently, in CTPE. The findings of this study confirm this claim since almost all studies in this review have addressed “quality” and provided different remedies to improve it or are still working on it.

By comparing the findings, we discovered that almost all studies follow a similar pattern, but there might be various procedures or lexicons referring to similar concepts. Different studies have used different contributors to elevate the quality level; for example, Aikawa et al. (2012) used moderators approved, by the web owner, to rate the post-edited content and accept them if they consider them good enough. Chen et al. (2017) would accept the post-edited results obtained from novice crowdsourcers only if three trusted contributors approved their accuracy of them. Mitchell et al. (2014) had evaluators rate the post-edited outcome to ensure its quality. What is shared by all scholars is that there should be a certain level of quality for the post-edited outputs facilitating the users’ comprehensibility. Efforts are being constantly made to optimize this quality level for various purposes; however, the translation quality level has a diverse conceptualization in different studies. For instance, in health contexts, Soares et al. (2020) mention that as far as health professionals comprehend the results of the studies that are related to their career, even if they are not fluent, the translation quality is good enough, while Dew et al. (2015) working in the same field of study, note that the post-edited output should present high quality. Due to these different quality conceptualizations, throughout our investigation, we realized that providing editing guidelines and specifying PE units would greatly help achieve the intended quality level.

• *Editing Guidelines*

The findings show that editing guidelines or instructions for post-editors vary in the studies reviewed (Aikawa et al., 2012; Chen et al., 2017; Soares et al., 2020; Iyengar., et al., 2013; Huynh et al., 2008); It is worth mentioning that lack of explicit instruction might lead to conflicting results or confusion of the crowd. Some studies have provided detailed and specific instructions (Huynh et al., 2008), while some asked the users to follow general principles (Soares et al., 2020). Based on these findings, the desired quality is only achieved if clear instructions or criteria are provided because, here, both post-editors and proofreaders are well aware of the text demands and will make the best decision. Different needs require different guidelines, so there are no general guidelines for post-editing (Hu & Cadwell, 2016). As a result, it is inferred that the intended quality of the post-edited translation is achieved with pre-defined clear guidelines.

• *PE Units*

Translation units vary based on the purpose of the study; they range from words to entire contexts; however, this is different in a crowdsourcing environment. The selection of the translation unit emerges logically as a primary consideration in crowdsourcing since tasks are grouped and distributed among different crowd participants. Therefore, the shorter the text, the more volunteer contributions are expected (Jiménez-Crespo, 2017); this is also true regarding the CTPE because, again, here, the same procedure is applied. Our findings on PE units imply that the studies attempted to provide a platform that enables the post-editors to edit or revise the translated outputs for both long and short translation units; however, we speculate that small PE units might contribute to obtaining a better post-edited output, leading to higher quality.

4.2. CTPE Areas of Application

Post-editing generally refers to a human being making some changes to make an MT output acceptable; therefore, when it comes to CTPE areas of application, one might immediately think that crowdsourcing translation is only one of the methods used for MT post-editing. However, our findings cast new light on CTPE implementation; we discovered that CTPE could be applied for both MT output and HT output, although the purpose of its implementation might vary. This section will illustrate some experimental results using CTPE for HT outputs. All studies reviewed here used CTPE for MT. However, some of them used it for revising or correcting the previously existing translations done by humans or post-editing what was already edited by other users, contributors, or colleagues. Aikawa et al. (2012) allowed the crowdsourcers to modify other crowdsourcers' translations. Their platform presents the output of MT and the translations done by the crowd members, if there are any, at the same time. The other crowdsourcers can modify either the MT output or these alternative translations. This process indicates that HT can be post-edited several times, although the first output belongs to MT. Chen et al. (2017) also stated that users might edit a segment, which in their case study is a question or an answer in the Stack overflow website, several times; again, this multiple post-editing implies HT post-editing.

Other studies also accepted that HT post-editing also improves the translation quality as well as MT post-editing. Soares et al. (2020) explicitly stated that they accept the last edition as the final translation because they believe it has the best quality. Yan et al. (2014) used a collaborative framework through which non-professional translators might provide different translations for the exact text, and different editors edit the translated sentences. Their result revealed that the outcome of this collaborative crowdsourcing process is very close to that of professional translations. Similarly, quality improvement of the already existing documents, even those post-edited by the staff, is one of the intended future purposes mentioned by Laurenzi et al. (2013). To our knowledge, these findings can open new horizons in post-editing. There are undoubtedly organizations whose already translated documents require revision or editions because of their

poor quality, and CTPE seems a practical and inexpensive solution. CTPE of HT output has excellent potential for application and cultivation.

4.3. Ongoing CTPE Trends

From the short review above, key findings of CTPE trends emerge in three significant areas: Introducing Crowdsourcing Platforms, Evaluating Current PE Processes, and Training Data

4.3.1. Introducing Crowdsourcing Platforms

Several studies have explored the use of crowdsourcing platforms for post-editing machine translation (MT) these studies introduce a platform through which crowdsourcers can post-edit the MT output. Some of these platforms also allow the crowdsourcers to edit the previously edited sections, which is a form of CT post-editing of human translation (HT) (Aikawa et al., 2012; Bernstein et al., 2010; Soares et al., 2020; Mitchell et al., 2014). The workflow is similar in these studies, with the platform providing users with MT output, and the crowdsourcers asked to edit or correct it. In cases where the translations are acceptable, bilingual members flag them or report them, and then start editing them again if necessary. One example of this type of platform is the Google Translate Community, which was used in the study by Mitchell et al. (2014). The platform provides users with MT output, and the crowdsourcers are asked to edit and improve it. The platform also allows for the editing of previously edited sections, which enables a form of CT post-editing of HT. The study found that the use of crowdsourcing for post-editing improved the quality of the MT output and reduced the time and cost of translation. Another example is the platform used by Soares et al. (2020), which was specifically designed for crowdsourcing translation. The platform provides users with MT output, and the crowdsourcers are asked to edit and correct it. The platform also allows for collaboration and communication between the crowdsourcers and project managers, which can improve the efficiency of the translation process. The study found that the use of crowdsourcing for post-editing improved the productivity of translators and reduced the time and cost of translation.

4.3.2. Evaluating Current Post-editing Processes

Another group of studies has investigated the practicality and quality of translations using existing crowdsourcing translation for post-editing (CTPE) platforms. These studies have helped to shed light on the effectiveness of CTPE as a translation solution. Laurenzi et al. (2013) investigated the impact of different editing strategies on the quality of MT output when edited by professional and non-professional translators. The study used an existing CTPE platform and found that the use of non-professional translators for post-editing resulted in a significant

improvement in translation quality. The study also found that the use of CTPE was practical for translating large amounts of data. Similarly, Iyengar et al. (2013) investigated the practicality and quality of translations using an existing CTPE platform. The study found that the use of CTPE resulted in a significant improvement in translation quality, especially for languages with low translation resources. This study also found that the use of CTPE was a practical solution for translating large amounts of data. The positive results of these studies provide evidence for the effectiveness of CTPE as a practical solution for poor-quality MT and HT translations. CTPE has the potential to significantly improve the quality of translations while increasing productivity. Additionally, the use of non-professional translators for post-editing can help to increase the diversity of translations and provide a more nuanced understanding of language use.

4.3.3. Training Data

Another group of studies has explored the use of crowdsourcing translation for post-editing (CTPE) as a tool to train machine translation (MT) and enrich translation memories (TM). Chen et al. (2017) introduced a bifunctional method that uses post-editing tools to facilitate the crowd's work in post-editing content while simultaneously training MT models and enriching TM. The study found that the use of CTPE can improve the quality and consistency of translations while reducing the time and cost of translation. The approach is based on the idea that CTPE can be used as a tool to facilitate the translation process. This approach can be seen as a dynamic system in which crowdsourcing is helping itself to improve. Similarly, Yan et al. (2014) used CTPE to provide training data for translation machines. The study found that the use of CTPE can significantly improve the accuracy of machine translations. By using CTPE to train MT models, the study demonstrated that CTPE can be used to improve the performance of MT systems. Huynh et al. (2008) also suggested that post-edited outputs can enrich the MT of existing references or constitute a new reference if there is not any. The study proposed a hybrid approach that combines MT and CTPE to improve the quality of translations. The study found that the use of CTPE can significantly improve the quality of translations, especially for languages with low translation resources. Generally speaking, the use of CTPE as a tool to train MT and enrich TM has shown promising results in improving the quality and consistency of translations while reducing the time and cost of translation. By using CTPE to train MT models and enrich TM, organizations can improve the accuracy and performance of their translation systems.

4.4. Platform Characteristics

Concerning the platforms used in different studies, simplicity, and user-friendliness were the main concerns in almost all studies. Regardless of the purpose of the study, users need clear instructions to perform the task perfectly; therefore, the more precise these instructions, the less time and energy are required. For example, Mitchell et al. (2014) provided the users with clear

instructions through a video to ensure they have a thorough knowledge of the user interface. Laurenzi et al. (2013), Dew et al. (2015), and Soares et al. (2020) explicitly state that their designed system is simple, user-friendly, and intuitive to avoid extensive user training. Other studies do not provide technical instruction because they are using or elaborating on an existing platform that is functional enough for its users. They are already familiar with its editing mechanism and need no training (Chen et al., 2017).

It is agreed on the functionality and user-friendliness of the platforms; however, they differ in some aspects. The most significant difference that might affect the workflow regarding time and cost is that when the document is outsourced to the crowd, some platforms allow crowdsources to post-edit an item (word, phrase, sentence, text) simultaneously, while others do not. As seen in the works of Laurenzi et al. (2013) and Dew et al. (2015), as soon as crowdsourcers “claim” (the term used in these studies which means “Select”) if interested to post-edit an item, other users cannot “claim” it anymore; however other users can revise the post-edited content, or work on the rest of the content if the first user stops editing or “unclaim” the item. Conversely, on platforms like CTF, introduced by Aikawa et al. (2012), several users can post-edit the same item simultaneously. Ultimately, either platform might be practical enough based on the study’s nature and purpose.

4.5. Crowd Characteristics

Models for crowdsourcing translation share several common features. One of the most fundamental ones is that they are typically performed voluntarily, by non-professional or unprofessional translators, for no or very little pay in some situations (Jiménez-Crespo, 2017). Overall, these findings regarding the crowd characteristics in CTPE are in accordance with that of crowdsourcing translation. For example, the crowd participating in the study of Soares et al. (2020) was volunteers who could read biomedical articles in both English and a second language. Similar features were obtained in others’ works when they used terms such as “foreign students with different language backgrounds” (Aikawa et al., 2012), “bilingual public health professionals” (Dew et al., 2015), or “novice contributors” (Chen et al., 2017) for the participating crowd. Except for two studies that did not specify participants, all others stated who the crowdsourcers were, and what seemed to be common to all of them was that they had a level of knowledge of what they were asked to post-edit. However, they were not trained, translators or linguistic experts.

4.6. Motivation

Motivation is another feature that plays an essential role in crowdsourcing translation. However, this review revealed that motivation had not been addressed as it should have been in the studies under investigation. Only two studies briefly pointed to the motivation of the crowd.

Aikawa et al. (2012) implied that the crowd's motivation to participate was intrinsic when they stated that their participation in the project gave them a feeling of community. Conversely, the crowd's motivation was extrinsic when they received payment for their participation (Yan et al., 2014). Although there was little evidence regarding crowd motivation in this study, one cannot deny its importance. Motivation has various shapes and forms and is essential not only for a crowd but for every single human being to perform a task because it is not logical to accept that human beings do something without any reason. We assume that solid motivation will lead to better outcomes in CTPE output; therefore, scholars should pay attention to this factor as much as other factors if they seek good outcomes for CTPE, meaning they should provide the crowd with appropriate incentives if needed.

4.7. CTPE Domains

CTPE has been used in various domains, but the findings suggest that its application is more common for environments with socio-professional content and terminology. For example, three studies focused on health content (Laurenzi et al., 2013; Soares et al., 2020; Dew et al., 2015), one study on website localization (Aikawa et al., 2012), some others on a communicative platform (Chen et al., 2017; Han et al., 2017) to be posted edited. If we refer back to section 3.5., we realize that although the crowdsourcers were not trained translators or linguists, they were knowledgeable in at least one subject area; therefore, it is evidence for the assumption that CTPE is more functional in specialized domains than general ones. Although crowdsourcing translation has no limits regarding the text type from literary works to specialized translations in the furthest corner of the world, no evidence was found regarding the post-editing of literary work or similar contexts (Jiménez-Crespo, 2017). The literary domain seems untouched in CTPE, and we presume it to be a good area for further investigation.

4.8. Future Perspective and Challenges

The findings suggest a promising area of research for CTPE since almost all studies included in this review expressed a positive attitude toward it. A shared purpose was enhancing the post-editing quality level by taking proper and effective measures. As previously mentioned, CTPE is applied in different areas, marking the future direction of the process. Studies working on creating CTPE platforms intend to improve their functionality by adding some features or removing their limitations. Others are positive about extending their workflow to other communities rather than limiting it to only one field. In a nutshell, CTPE seems to have much potential to be implemented for the task by so many organizations, communities, institutes, and those sectors that need a large and inexpensive amount of quality translated data in a short time or those who intend to improve the quality of their previously translated documents.

Although CTPE can be an effective method for training MT models, the researchers feel that there are certain possible difficulties and constraints that should be taken into account. First of all, using CTPE to train MT models takes a lot of time and effort, and crowdsourcing platforms must be properly created and set up for CTPE; otherwise, the post-editing process might take a long time. Second, gathering and processing a sizable amount of data is necessary in order to employ CTPE for training MT models. This can be difficult for languages with few resources for translation since there might not be enough data to properly train the MT models. Thirdly CTPE can be more expensive than other methods of training MT models, such as using professional translators or in-house linguists therefore convincing incentives should be set to motivate the expert crowdsourceurs to participate. Moreover, the costs associated with managing and coordinating crowdsourcing platforms, as well as compensating the crowd members for their contributions, can add up quickly. Conclusively while CTPE can be an effective tool for training MT models, it is important to consider the quality of the post-edited translations, the time and effort required the availability of data, and the associated costs when deciding whether to use this approach.

4. Conclusion

Based on the systematic review of the current literature, this study presents the findings of a crowdsourcing translation as a post-editing method. We did our best to carefully scan the studies to extract whatever was related to this topic and not miss anything. The findings of the study suggest that crowdsourcing translation as a post-editing method (CTPE) is a promising solution for improving the efficiency and effectiveness of the translation process in various fields. One of the main advantages of CTPE is its potential to save time and reduce expenses by leveraging crowdsourcing to perform post-editing. In other words, CTPE can help to reduce the workload of professional translators and enable faster turnaround times for translations, while also reducing costs compared to traditional translation methods. The study found that CTPE has the potential to improve the quality of translations by utilizing the collective knowledge and expertise of the crowdsourcing participants to identify and correct errors in machine translations or improve the quality of previously translated material. This can result in higher-quality translations that are more accurate, consistent, and appropriate for the intended audience.

Although it is generally believed that the quality of human translation is higher than that of automatic translation, except in some settings with a very narrow domain (Koehn, 2010), this review has found that CTPE is not limited to machine translation output only and is potent enough to improve the quality of existing translations. Given the potential benefits of CTPE on HT, the focus of our subsequent future work will be on an empirical study focusing on the application of CTPE on already existing translations done by humans to discover if it affects

translation quality. Such a study would contribute to a better understanding of the potential benefits of CTPE and could lead to the development of new best practices for CTPE on HT.

The current state of CTPE is characterized by three distinct fields of study that offer significant potential for future development. The first field involves introducing new CTPE platforms that can improve the efficiency and effectiveness of the post-editing process. These platforms may incorporate new features such as user feedback mechanisms and machine learning algorithms that can enhance the quality of translations easily in a short time. The second field of study involves evaluating current CTPE processes to identify areas for improvement. This may include analyzing the performance of CTPE in terms of time, cost, and translation quality, as well as identifying potential challenges and limitations of the method. By understanding these factors, researchers and practitioners can develop more effective strategies for using CTPE in different settings. The third field of study focuses on training data, which is essential for the success of CTPE. This may involve developing new methods for selecting suitable crowdsourcing participants, improving the quality of the source material that is being translated, and developing more effective training programs for crowdsourcing participants.

Furthermore, the study has revealed that CTPE platforms are typically designed to be user-friendly and easy to navigate, which may facilitate their adoption and use among a broad range of users. The simplicity and accessibility of these platforms may enable individuals who lack specialized training in translation to participate in CTPE projects, potentially enhancing the diversity and size of the crowdsourcing pool. An intriguing avenue for future research involves exploring the potential applicability of smartphones in CTPE. Smartphones have become ubiquitous and are widely used by people around the world, making them an attractive platform for crowdsourcing translation. Given their user-friendly interface and portability, smartphones may facilitate CTPE participation among a greater number of individuals and enable greater flexibility in terms of when and where translations can be performed. Further investigation into the potential advantages and limitations of using smartphones for CTPE could yield valuable insights and inform the development of new CTPE platforms that are tailored to mobile devices.

The study further uncovered that specialized domains, particularly in the medical and legal fields, are the primary users of CTPE. This could be attributed to the technical language and jargon commonly used in these fields, which may require a higher level of translation accuracy and consistency than in other domains. The finding points towards the possibility of specific use cases where CTPE may be most effective. However, further research is needed to identify these domains and to develop best practices for their application. The applicability of CTPE in other domains, such as literary translation, remains an open question. Future research in this area could provide valuable insights into the potential benefits and limitations of CTPE in literary translation and inform the development of best practices for its application in this domain.

An additional noteworthy observation from the literature is that CTPE is commonly carried out by non-professional or unprofessional translators. However, it is plausible that under

certain circumstances, a closed crowdsourcing group of experts may be advantageous. Specifically, in specialized domains where precision and domain-specific knowledge are vital, utilizing a closed crowdsourcing group of experts may enhance the quality of the translations. Despite these initial indications, further research is necessary to fully explore the potential benefits and drawbacks of using a closed crowdsourcing group of experts in CTPE. Last but not least finding is motivation. Although motivation was not an extensively addressed factor in the study, the future outlook for CTPE is optimistic. Motivation plays a crucial role in the success of crowdsourcing translation, including in the context of CTPE. Crowdsourcing relies on the voluntary participation of individuals who are not professional translators, and therefore, their motivation to contribute is essential for the success of the project. Motivation can be influenced by a variety of factors, including financial compensation, personal interest in the topic, a desire to contribute to a larger cause, and recognition for their work. In the context of CTPE, the motivation of crowdsourcing participants can be particularly important, as the quality of their work can directly impact the quality of the final translation.

6. Limitations

This work does not have an exemption from limitations. The first limitation is that this study focuses only on articles published in English. We know that using English as the standard scientific language may prevent other languages from making original contributions. Future research might remove this limitation by including non-English studies in the systematic review. The second limitation was that due to copyright laws and legal restrictions, we are currently facing in our country, Iran, in the selection phase, we only had access to the abstracts of some papers. We excluded some of them because of the exclusion criteria, which might limit the extracted knowledge, indicating that we might have missed including some relevant works.

References

- Aikawa, T., Yamamoto, K., & Isahara, H. (2012). The impact of crowdsourcing post-editing with the collaborative translation framework. *Advances in Natural Language Processing*, 1-10. https://doi.org/10.1007/978-3-642-33983-7_1
- Anastasiou, D., & Gupta, R. (2011). Comparison of crowdsourcing translation with machine translation. *Journal of Information Science*, 37(6), 637–659. <https://doi.org/10.1177/0165551511418760>
- Azari, R., Ganjalikhani, M. K., & Amirshoja'i, A. (2020). Crowdsourcing Translation: An Experience of a Cloud-based Messaging App. *Iranian Journal of Translation Studies*, 17(68), 7-22.
- Bernstein, M. S., Little, G., Miller, R. C., Hartmann, B., Ackerman, M. S., Karger, D. R., ... & Panovich, K. (2010). Soylent: A word processor with a crowd inside. *In Proceedings of the 23rd annual ACM symposium on User interface software and technology*, 313-322.
- Bentivogli, L., Federico, M., Moretti, G., & Paul, M. (2011). Getting expert quality from the crowd for machine translation evaluation. *In Proceedings of Machine Translation Summit XIII: Papers*.
- Chen, C., Xing, Z., & Liu, Y. (2017). By the community & for the community: A deep learning approach to assist collaborative editing in Q & A sites. *Proceedings of the ACM on Human-Computer Interaction*, 1(CSCW), 1-21.
- Dew, K., Turner, A. M., Desai, L., Martin, N., Laurenzi, A., & Kirchoff, K. (2015). *PHAST: A Collaborative Machine Translation and Post-Editing Tool for Public Health*. AMIA Annual Symposium proceedings. AMIA Symposium, 2015, 492–501.
- Flanagan, M. (2016). Cause for concern? Attitudes towards translation crowdsourcing in professional translators' blogs. *The Journal of Specialised Translation*, 25(1), 149–173.
- Federici, F. M., & Cadwell, P. (2018). Training citizen translators: Design and delivery of bespoke training on the fundamentals of translation for New Zealand Red Cross. *Translation Spaces*, 7(1), 20–43.
- Gao, M., Xu, W., & Callison-Burch, C. (2015). *Cost optimization in crowdsourcing translation: Low cost translations made even cheaper*. In *Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies* (pp. 705–713).
- Goto, S., Ishida, T., & Lin, D. (2016, September). Understanding crowdsourcing workflow: Modeling and optimizing iterative and parallel processes. *In Proceedings of the AAAI Conference on Human Computation and Crowdsourcing* (Vol. 4, pp. 52-58).
- Green, S., Heer, J., & Manning, C. D. (2013, April). *The efficacy of human post-editing for language translation*. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 439-448).
- Guazzini, A., Vilone, D., Donati, C., Nardi, A., & Levnajic, Z. (2015). Modeling crowdsourcing as collective problem solving. *Scientific Reports*, 5(1), 1-11.
- Guerberof, A. (2009). Productivity and quality in MT post-editing. *In MT Summit XII-Workshop: Beyond Translation Memories: New Tools for Translators MT Association for Computational Linguistics (ACL)*. <https://aclanthology.org/2009.mtsummit-btm.7>
- Han, B., Radford, W., Cadilhac, A., Harol, A., Chisholm, A., & Hachey, B. (2017). Post-edit analysis of collective biography generation. *Proceedings of the 26th International Conference on World Wide Web Companion - WWW '17 Companion*. <https://doi.org/10.1145/3041021.3054264>
- Holden, R. J., & Karsh, B. (2010). The technology acceptance model: Its past and its future in health care. *Journal of Biomedical Informatics*, 43(1), 159-172. <https://doi.org/10.1016/j.jbi.2009.07.002>
- Hu, K., & Cadwell, P. (2016). A comparative study of post-editing guidelines. *Baltic Journal of Modern Computing*, 4(2), 346–353.
- Huynh, C. P., Boitet, C., & Blanchon, H. (2008, May). SECTra_w_1: An Online Collaborative System for Evaluating, Post-editing and Presenting MT Translation Corpora. In *LREC* (Vol. 8, pp. 28-30).
- JavadiKhasraghi, H., & Mohammadi, S. (2012). An innovative Crowdsourcing approach for Amazon mechanical turk. *International Journal of Computer Applications*, 52(4), 20-13. <https://doi.org/10.5120/8190-1556>
- Jiménez-Crespo, M. A. (2013). Crowdsourcing, corpus use, and the search for translation naturalness. *Translation and Interpreting Studies*, 8(1), 23-49. <https://doi.org/10.1075/tis.8.1.02jim>

- Jiménez-Crespo, M. A. (2017). Crowdsourcing and online collaborative translations: Expanding the limits of translation studies. John Benjamins Publishing Company.
- Jiménez-Crespo, M. A. (2018). Crowdsourcing and translation quality: Novel approaches in the language industry and translation studies. In *Translation quality assessment* (pp. 69-93). Springer, Cham.
- Jiménez-Crespo, M. A. (2017). Translation crowdsourcing: Research trends and perspectives. In A. Cordingley & C. F. Manning (Eds.), *Collaborative translation from the Renaissance to the digital age* (pp. 192-211). Bloomsbury.
- Koehn, P. (2010). Enabling monolingual translators: Post-editing vs. options. In *Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, 537-545.
- Laurenzi, A., Brownstein, M., Turner, A. M., Kirchoff, K., & Kientz, J. A. (2013). A web-based collaborative translation management system for public health workers. CHI'13 Extended Abstracts on Human Factors in Computing Systems on - CHI EA'13. <https://doi.org/10.1145/2468356.2468446>
- Mitchell, L., Roturier, J., & O'Brien, S. (2013). Community-based post-editing of machine-translated content: monolingual vs. bilingual. In *Proceedings of the 2nd Workshop on Post-editing Technology and Practice*.
- Mitchell, L., O'Brien, S., & Roturier, J. (2014). Quality evaluation in community post-editing. *Machine Translation*, 28(3-4), 237-262. <https://doi.org/10.1007/s10590-014-9160-1>
- Mitchell, L. (2015). Community post-editing of machine-translated user-generated content [Doctoral dissertation, Dublin City University].
- Munková, D., Munk, M., Benko, L., & Absolon, J. (2019). From old fashioned "One size fits all" to tailor made online training. *The Challenges of the Digital Transformation in Education*, 365-376. https://doi.org/10.1007/978-3-030-11932-4_35
- Muntés-Mulero, V., Paladini, P., Manzoor, J., Gritti, A., Larriba-Pey, J., & Mijndhardt, F. (2013). Crowdsourcing for industrial problems. *Lecture Notes in Computer Science*, 6-18. https://doi.org/10.1007/978-3-642-36074-9_2
- O'Hagan, M. (2021). Community translation: Translation as a social activity and its possible consequences in the advent of Web 2.0 and beyond. *Linguistica Antverpiensia, New Series-Themes in Translation Studies*, 10. <https://doi.org/10.52034/lanstts.v10i.275>
- Pavlick, E., Yan, R., & Callison-Burch, C. (2014). Crowdsourcing for grammatical error correction. *Proceedings of the companion publication of the 17th ACM conference on Computer supported cooperative work & social computing-CSCW Companion' 14*. <https://doi.org/10.1145/2556420.2556479>
- Pascoal, S. C., Furtado, M., & Chorão, G. (2017). Crowdsourcing translation and the threats/challenges to the profession: A Portuguese translator's survey. *Current Trends in Translation Teaching and Learning E (CTTL E)*, 4, 229-263.
- Plitt, M., & Masselot, F. (2010). A productivity test of statistical machine translation post-editing in a typical localisation context. *Prague Bull. Math. Linguistics*, 93, 7-16.
- Shimohata, S., Kitamura, M., Sukehiro, T., & Murata, T. (2001). Collaborative translation environment on the Web. In *Proceedings of Machine Translation Summit VIII*.
- Soares, F., Rebecchi, R., & Stevenson, M. (2020). SciBabel: A system for crowdsourced validation of automatic translations of scientific texts. *Genomics & Informatics*, 18(2).
- Sutherlin, G. (2013). A voice in the crowd: Broader implications for crowdsourcing translation during crisis. *Journal of Information Science*, 39(3), 397-409. <https://doi.org/10.1177/0165551512471593>
- Tatsumi, M., Aikawa, T., Yamamoto, K., & Isahara, H. (2012). How good is crowd post-editing? Its potential and limitations. In *Workshop on Post-Editing Technology and Practice*.
- Yan, R., Gao, M., Pavlick, E., & Callison-Burch, C. (2014). Are two heads better than one? Crowdsourced translation via a two-step collaboration of non-professional translators and editors. *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. <https://doi.org/10.3115/v1/p14-1107>
- Yan, R., Song, Y., Li, C. T., Zhang, M., & Hu, X. (2015, June). Opportunities or risks to reduce labor in crowdsourcing translation? characterizing cost versus quality via a pagerank-hits hybrid model. In *Twenty-Fourth International Joint Conference on Artificial Intelligence*.