

Application of Translation Technologies in the Translation of IMTFE Transcripts

Danhua Huang¹

¹ English College, Zhejiang Yuexiu University, Shaoxing, Zhejiang, China

Correspondence: Danhua Huang, English College, Zhejiang Yuexiu University, Shaoxing, Zhejiang, China.
E-mail: david8301@qq.com

Received: January 20, 2021 Accepted: February 21, 2021 Online Published: February 26, 2021

doi:10.5539/ells.v11n1p51 URL: <https://doi.org/10.5539/ells.v11n1p51>

Abstract

Machine translation has grown very rapidly in recent times due to the developments in big data, artificial intelligence, and cloud computing software and techniques. The first generation of Rule-Based Machine Translation has been replaced by fourth generation of Neural Machine Translation based on complex deep learning and networking models. Translation models have also undergone tremendous changes. The traditional translation model fails to meet the needs of the modern language service industry. There still exists doubt whether Translation technologies could enhance the quality in the translation of non-technical texts. So far no one has discussed the application of the technologies in the translation of IMTFE Transcripts. This paper aims to prove that translation technologies can be applied to the translation of IMTFE Transcripts which the author works on. By analyzing the text features of IMTFE Transcripts, and applying different translation technologies in the translation, the author finds that the quality and efficiency of translation have been greatly enhanced. It is concluded that translation technologies can be used to facilitate translation of texts of different kinds. In spite of their drawbacks, translators can still benefit a lot from the adoption of translation technologies.

Keywords: machine translation, translation technologies, International Military Tribunal for the Far East transcripts, deep learning

1. Introduction

1.1 Background of the Study

After Japan was defeated in the Second World War in 1945, The Allied Power established the International Military Tribunal for the Far East (hereinafter referred to as IMTFE) in Tokyo, to try Japanese war criminals in accordance with the Cairo Declaration 1943, the Potsdam Proclamation, and the International Military Tribunal for the Far East Charter. China, the United States, the Soviet Union, the United Kingdom, France, the Netherlands, Australia, New Zealand, Canada, India and the Philippines sent a total of 11 judges to participate in the trial. After the establishment of IMTFE from May 1946 to November 1948, which lasted about two and half years with a total of 817 Trial sessions (in this paper a session refers to one day). Of those 7 convicted Class A criminals were sentenced to death penalty, 16 convicted Class A war criminals were sentenced to life imprisonment, and 2 convicted Class A war criminals were sentenced to imprisonment for 20 years and 7 years respectively. Of the other three Class A war criminals, they were not tried because one suffered from mental illness, and the other two died during the trial (Cheng, 2014, pp. 153–159).

IMTFE Transcripts are recorded in Japanese and English, detailing the history of Japan's war of aggression. But due to historical reasons, China does not have the complete transcripts. Over 60 years following the trial, National Library of China and the Shanghai Jiaotong University Press not only managed to obtain nearly 50,000 pages of photocopied transcripts from US National Archives, but painstakingly pieced together the missing 7,000 pages of the transcripts in 2012, and published 80 volumes of the transcripts in English.

The author discovered while searching the CNKI with the academic paper's title that 3 retrieved papers had all been published in 2014. Similarly, a search corresponding to IMTFE (in Chinese) with the paper's title revealed that a total of 48 papers had been published between Sept 1979 and Sept 2017. However, as IMTFE is more commonly and informally known as the Tokyo Trial. Hence the author repeated the search using the Chinese expression of Tokyo Trial in the search bar and found that 540 papers had been retrieved and subsequently

published between March 1947 to Sept 2020 of which 385 had been published from August 2001 to Sept 2020. The author further classifies the search criteria in accordance with the institutional affiliations of those authors; Shanghai Jiaotong University ranks 1st with 28 papers, which is partially explained by the existence of the Tokyo Trial Research Center at the university which it had established on its 65th anniversary in 2011. Nanjing University ranks 2nd with 21 related papers, followed by the National Library of China with 11 papers. Further search shows that Professor Song Zhiyong of Nankai University published 16 papers, and Professor Cheng Zhaoqi of Shanghai Jiaotong University published 12 papers. The search reveals that a total of 9 papers with the corresponding Chinese translation of Tokyo Trial were published in 2020. (The above retrieval information is as of December 25, 2020). It has been over 70 years since the completion of the trials, and the total number of relevant research papers cannot be considered large. Interestingly 93% of the papers were published after 2000 reflecting the growing interest of scholars in research on IMTFE. Most of the Chinese scholars tend to use the corresponding Chinese expression of Tokyo Trial rather than the corresponding Chinese expression of IMTFE, and the main research focus is mostly on the historical significance of the trial, the fairness of the trial, the comparison between IMTFE and its counterpart Nuremberg Trial in the same era.

A search using the Chinese expression of Tokyo Trial reveals only 6 publications from 2015–2019 disclosing a lack of research interest in translation and interpretation of IMTFE amongst Chinese scholars. In contrast two scholars from Japan did an in-depth study of the interpretation mechanism of IMTFE. Namely; Kayoko Takeda from Rikkyo University in Japan. She explored the interconnections among the interpreters, monitors and the language arbiters in IMTFE from the sociopolitical perspective. And Tomie Watanabe who elaborated the interpretation and analyzed the interpretation based on the monitors' intervention in the Class A criminal Tojo's cross examination case.

The translation of IMTFE Transcripts in China did not start until early 21st century, but great progress has been made since. Shanghai Jiaotong University Press and the National Library of China Press jointly published International Military Tribunal for the Far East Index and Appendix in 2013, which includes the index of the names, the index of evidence, the index of people present in court, the index of important events and the appendix. Evidence index is the first multilingual index compiled in three languages i.e., Chinese, English and Japanese. The appendix includes the relevant system, historical events, documents, and research results.

In March 2019, under the leadership of Professor Ye Xingguo of Zhejiang Yuexiu University, a translating team was founded in collaboration with Shanghai Jiaotong University Press for English to Chinese translation project of IMTFE Transcripts (the Second Volume). As a member of the team, the author has the honor to complete the translation of about 300,000 words which is one book covering the proceedings from October 1, 1947 to October 14, 1947. As one of the classic publishing projects of Shanghai Jiaotong University Press, the second volume of IMTFE was awarded the National Publishing Fund in 2020. The translation of proceeding transcripts is of great academic value, historical and practical significance, since it would be a rich resource base for the Chinese researchers and an important historical document for unraveling the historically important aspects of World War II.

1.2 Literature Review

Cui (2016) states that the texts can be divided into three levels. The first level is reference text, which is used to obtain basic information, and it is used for reference. It is acceptable to have some parts untranslated. The second level is the ordinary text, which is used to give readers' guidance. The renderings should be faithful to the original text, with accurate terminology and complete meaning. The third level is publishing text, which needs to be faithful, fluent and even elegant as it is used for printing or publishing. He argues that different strategies are used for the texts of different levels. According to him, human translation is mainly adopted in the translation of publishing text.

Zhang (2020) states that the connotation and function of legal translation reflect the differences between legal translation and translation of other texts. The preciseness and professionalism of legal translation fundamentally means that application of machine translation to the field of legal translation is still facing multiple difficulties.

Xu (2015) argues that it is unreasonable to think that CAT is not suitable for literary translation. The fundamental problem is those users are not proficient in CAT tools and fail to formulate and follow reasonable translation and proofreading procedures. These are not the problems of CAT tools, but the users' instead.

Fu (2015) states that translation in the modern times is a translation which integrates translation technologies and is supported by translation technologies. Translation technologies have become an important part in the field of translation that cannot be ignored.

It can be concluded that some scholars argue that translation technologies especially CAT tools are not always reliable. Others think they can be widely used in translation. However, few studies have been done on the feasibility and necessity of application of translation technologies to the translation of IMTFE Transcripts. It is the author's interest to carry out an empirical study to prove the efficiency and effectiveness of translation technologies in the translation of IMTFE Transcripts.

1.3 Methodology

It may be possible to apply the translation technologies to the translation of IMTFE Transcripts. The current research seeks whether the translation technologies could greatly enhance the translation productivity and accuracy. This is a qualitative and descriptive study based on the application of the different translation technologies. Although only several examples are analyzed in this paper, this choice is made after observing and analyzing the text. The examples have been selected to draw the conclusions, which will be complemented with further studies. The empirical study consists of 4 stages. The first stage analyses textual characteristics of IMTFE Transcripts. The second stage focuses on the pre-process preparation especially the establishment of term base and translation memory. The third stage focuses translating or editing the renderings with the aid of translating technologies. The final stage is to get the conclusion and evaluate the relationship between translators and translation technologies.

2. The Textual Characteristics of IMTFE Transcripts

2.1 Rich Terminology

One important feature of the IMTFE Transcripts is that it abounds with repeated terminologies.

2.1.1 Terms, Names and Titles

Over 100 criminal suspects including 28 Class A war criminal were arrested by the US soldiers after their landing in Japan. Subsequent criminal proceedings that followed were attended by judges, prosecutors, defendants, counsels and witnesses. The transcripts contain the names such as the Class A criminal Hideki Tojo who was later hanged, Sir William Flood Webb the chief justice of the court, Joseph Berry Keenan the chief prosecutor, and many other Chinese people, who attended the trials.

The trails involve military and government officials of different ranks. Such as War Minister, Foreign Minister, Commander-in-Chief of the Kwantung Army, the Chief of the General Staff, the Chief of the Naval General Staff, the President of the Privy Council, Minister of Finance etc.

2.1.2 Toponymic Terms

IMTFE Transcripts also contains the names of places where many incidents occurred, such as the Changkufeng, Tangku, Fushian, kuling, Ryojun, Canton, Marco Polo Bridge. When the trial is about China on a specific day, such kind of names of places could appear in one Trail transcript repeatedly although they do not appear in every Trail transcript.

2.1.3 Terms of Historical Events

Major incidents are mentioned in the trail such as the China Incident, the Nanking Incident, the Mukden Incident, the Manchurian Incident Chang Ku Feng Incident, Shimpeitai Incident, The Nakamura Incident, to name just a few.

2.1.4 Terms of Different Organizations

Various Chinese and Japanese government departments, companies, and organizations find a mention in the transcripts, and the use of institutional acronyms is also common. Such as Diet, Foreign Ministry, Asahi Shinbun Sha, the Concordia Society, Manchukuo, the Special Service Section of Mukden, Kokusui-Kai, IPS etc.

2.1.5 Legal Terms

During the Trial transcript, the proceedings are conducted in Japanese and English and they contain a lot of legal terms. Such as indictment, exhibit, count, affidavit, deposition, cross examination, testimony etc.

2.2 Interrogative Sentences

Direct examination and cross-examination are the main interrogation methods adopted in the court. During the proceedings, prosecutors ask defendants and witnesses a lot of carefully framed questions. Sometimes in order to clarify misunderstandings and ambiguity, the questions are repeated. Sometimes it is repeated by prosecutors as a part of well-crafted strategy of interrogating the defendants. For example, in the Trail transcript of October 8th 1947 altogether there are 181 interrogative sentences.

2.3 Alternation Between Simple Short Sentences and Difficult Long Sentences

In the trial transcript, some expressions are very short & simple, thus it is very easy to understand, such as the following expression:

“I don’t recall that particular speech for sure, but I may have made such a speech.” (Trial transcript of October 14th 1947)

Some sentences are difficult to understand because of the terms they contain.

For example, in connection with the last two answers of the witness, prosecutors wish to refer the Tribunal to exhibit 1103, at record page 10,062.

“Before we go into that, I would like to get it straight on the record that this affidavit is being presented as part of the defense of the accused Minami.” (Trial transcript of November 19th 1947)

The second expression is not so easy to understand as the first one, since it contains the words that are not always used in our daily life or its meaning is different from the ordinary one. Such as the “record” refers to the document which is written down in the court, and the “affidavit”, “defense” and “accused” are legal terms.

Although short sentences are very common in the proceedings. There are also plenty of long sentences which mainly are subject clauses, predicative clauses, object clauses, attributive clauses, adverbial clauses, sometimes one paragraph consist of one sentence. For example,

“Next with regard to my activities, I have performed the duties of Chief of the Naval Affairs Bureau in accordance with the established regulations governing the functions of the Navy Ministry and my department, and pursuant to the orders and instructions and at the supervision of my chief, the Minister of the Navy, performing my duties faithfully in accordance with the regulations and orders.” (Trial transcript of November 19th 1947)

3. Definition of Translation Technologies and Their Benefits

3.1 Introduction of Translation Technologies

Machine translation (MT) refers to translating text from one natural language into another natural language with the use of the machine (usually a computer) (Hu, 2013, p. 25). Artificial intelligence, big data, and cloud computing have brought unprecedented opportunities and challenges to human being, affecting all walks of life, including translation industry. Machine translation is the application of artificial intelligence in the field of translation and which is known as the jewel in the crown of artificial intelligence. Since the 1970s, machine translation has experienced four stages from its birth, development to prosperity. The first stage is Rule-Based Machine Translation, the second is Example-Based Machine Translation, the third is Statistical Machine Translation, the fourth is Neural Machine Translation based on neural networks. Sutskever et al. (2014) proposed a sequence-to-sequence model. This is a general framework that uses neural networks to map one sequence to another, and uses neural networks for machine translation. But none except Google, paid any attention to the subject initially. Two years later, Google launched a neural machine translation engine with deep learning capabilities, bringing machine translation to a new level. Subsequently, China’s Internet giants such as Baidu, Youdao, Sougou also launched their own neural machine translation engines, each with its own special and outstanding features.

Computer-Aided Translation technologies(CAT) refer to all computer software that can assist translators in translation, including optical character recognition software, speech recognition systems, corpus analysis tools, terminology management, translation memory systems, and localization tools. Some Chinese scholars only mentioned the use of CAT tools such as Trados, MemoQ, Déjà vu or tools with similar function when studying computer-aided translation technologies. This is a narrowly defined CAT technology, which is, a translation technology that uses translation memory matching technology to improve translation efficiency (Wang, 2014, pp. 92–97).

Translation technologies can be divided into broad sense and narrow sense. In the broad sense, they refer to all forms of information technology that can be used to help language translation and other cross-language communication activities, while in the narrow sense, they mainly refer to CAT tools (Lu, 2014, p. 63). Translation technologies involve application of various technologies in human translation, machine translation and computer-aided translation, including the application of basic computer tools and special translation tools (Lynne Bowker, 2002, p. 6). Different scholars have different definitions about CAT technology and translation technology. Some use the two different terms to refer to the same notion. The author uses the term “translation technologies” which should consist of CAT tools instead of using CAT to refer to general tools and specific tools

as applied in translation.

In modern times, translation research and practice require effective and efficient solutions to process large volumes of texts. Translation Technologies, especially machine translation (MT) and computer-assisted translation (CAT) tools, provide such a solution and have become the core of the translation process, as well as an integral part of interlingual communication and the global language service industry.

3.2 Advantages of Translation Technologies

In order to adapt to the development of the modern language service industry, more and more language service providers (LSP) want to employ qualified translators who are adept at using translating tools. Practice has proved that the application of translation technologies can effectively improve effectiveness and efficiency in the translation of various texts. Some Chinese scholars have explored the possibility of applying translation technologies in the translation of the areas such as medicine, tourism, science and technology, legislation, manuals, film and television subtitles, but so far no one has studied the application of translation technologies in the translation of IMTFE transcript. Traditional translation, which means translating without relying on the translation technologies, often has the following two problems. One is that the terms are not consistent, which means the same terms could lead to the different meanings or interpretations. Even if the term is made into an Excel glossary, the translator still needs to search for it on its own. For a voluminous translation project of millions of words, it is still quite time-consuming. Shanghai Jiaotong University Press has published *International Military Tribunal for the Far East Index and Appendix* in 2013, which greatly facilitates the retrieval of terms in the translation. However, the aforementioned problems still exist in traditional retrieval methods. The second problem is that the translator cannot make full use of the translated resources, which means the translator has to translate every single sentence even it has been translated before. As a result, it leads to the unnecessary repetitive work. During the proceedings, there are a lot of same or similar expressions in the interrogation among prosecutors, defendants, witnesses, and judges. Every Trail transcript of the court contains formulaic expressions. Such as “May it please the Tribunal.” issued by the defendant’s counsel very frequently as soon as he is going to speak in the court. When an evidence is submitted and admitted by judge, the following expression is always used by the judge. “Admitted on the usual terms”. Without relying on translation technologies, translators can only rely on one’s own memory for translation. It still takes time to identify and is likely to cause problems in the translation such as the same content in the source language with totally different renderings in the target language, as a result the consistency and style of the translation cannot be guaranteed.

The analysis of the IMTFE Transcripts, by the author has revealed that the repeated content accounts for not only a considerable proportion, but that the same content also overlaps among Trail transcripts of different dates, the same terms and formulaic expressions also abound in the Trail transcripts. This implies that it is feasible to apply translation technologies to make use of the repeated parts of the transcripts. Taking the transcripts of October 19, 1947 as an example, the author counted the repeated segments with Trados, one widely used CAT tool. With the calculation of Trados the total number of words were 14,695, the repeating segments in the transcript were 1,506. The repetition rate is 10.25%, the same segments are mainly names and simple short answers in the interrogation. When the CAT tool is used, a name just needs to be translated once, when it appears again, there is no need for the translator to retranslate it. When a large-scale high-quality translation memory is applied, the matching rate between the original text and the translation memory will be much higher, thereby reducing unnecessary repetitive work and enhancing translator’s productivity, allowing the translator to focus on the quality of the untranslated text. During the translation process, the author integrates a variety of translation technologies, using Trados as the main translation platform, utilizing term bases and translating memories which effectively improves the efficiency and the quality of translation which in turn demonstrates the feasibility and necessity of applying translation technologies in this translation project.

4. Application Strategies

The development of artificial intelligence has brought unprecedented changes to the translation model. Translation is divided into fully automated high-quality translation (FAHQT), machine-aided human translation (MAHT), human-aided machine translation (HAMT), and human translation (HT). At present, machine translation can provide fully automatic translation, but it is not yet possible to achieve high-quality translation of the texts of various subjects. In machine-aided human translation, humans are mainly responsible for the entire translation process, but the machine provides us with different types of support and assistance; in human-aided machine translation, the translation is essentially carried out by the tools and humans help the process by pre-editing or post-editing the text (Zhou, 2013, pp. 91–95).

The translation of IMTFE Transcripts will be published by the publishing press which poses high requirements

for the quality and efficiency of translation. Therefore, it is not a desirable choice to solely rely on machine translation or human translation to achieve both high quality and efficiency in the meantime. In the translation project, the author has adopted MAHT and HAMT which effectively applies translation technologies to the whole translation workflow. The author has divided it into three stages: pre-translation, during translation and post-translation.

4.1 Pre-Process for Translation

4.1.1 Conversion and Processing of Text Format

The original document of IMTFE Transcripts is a noneditable PDF format which in fact is converted from the photos. Such kind of PDF cannot be translated with Trados. Therefore, OCR software has to be used to convert PDF into editable MS Word documents. ABBYY Finereader, PDFelement, AdobeAcrobat could be used to complete this task, on the whole the overall quality of text converted by ABBYY Finereader is the highest, but there are still some flaws due to the printing-issues in the original text. So, spelling mistakes have to be checked first. When the author translated the converted word document, for the first time, an annoying problem arose. i.e., with Trados the name instead of being in the same segment, appeared separated in two different segments such as TAKAZUMI OKA is separated and in two different segments. Batch merging cannot be done, merging separated segments consumes a lot of time and energy since such problems are quite common in the whole text. It is annoying distraction in the translation work. In order to solve the problem, the author used the search and replace function in MS Word, enter the original English name such as TAKAZUMI OKA and replace it with Takazumi Oka with one batch. It turns out that such problems caused by the upper case are easily solved.

4.1.2 Preparation and Management of Term Base

There are different ways to get term base: searching or purchasing on line, sharing or making by the translator. It is difficult to guarantee the quality of terminology with first two ways. It is a good way to share their term base among the project members. Although it takes a long time for the author to establish the term base by himself, it ensures quality since it is doublechecked by the author. In the translation project of IMTFE Transcripts, Shanghai Jiaotong University Press provided some terminology. Project members also shared the Excel glossary online through Tencent documents. project members can edit and update the terminology. In the translation process, the terminology is continuously enriched and can be searched in real time. Members can share resources, promoting translation efficiency and quality, but the disadvantage is that translators still need to frequently search Tencent documents, which requires internet support. The strategy adopted by the author is to use the terminology conversion tools SDL MultiTerm Desktop and MultiTerm Convert to convert the terminology provided by the publisher into a term base. The completion of the conversion requires the use of two software. MultiTerm only recognize and accept files in XML format, so MultiTerm Convert is used to convert the Excel glossary into XML format. Another term conversion tool used by the author is Glossary Converter. This process is relatively hassle-free. After setting the export format, the author just needs to drag the Excel glossary into the software, and term base of Sdlb format will be automatically exported to the designated folder. As for the terms on the online Tencent documents, the author first downloads and saves them in the format of Xlsx for checking the accuracy of the terms. As all the project members can use and edit Tencent documents, duplication, there may be erroneous or repetitious terms due to carelessness. MS excel is used to delete duplicate terms, and after checking, the author can convert it into term base of Sdlb format in the aforementioned ways.

4.1.3 Preparation and Management of Translation Memory

The working principle of translation memory (TM) is that the user uses the existing original text and its corresponding renderings to create one or more translation memories. During the translation process, the system will automatically search for the same or similar expression in translation memories. Translation resources (such as segments) provide reference translations so that the user can avoid unnecessary repetitive work and focus on the translation of new content (Christensen & Schjoldager, 2016, p. 90). The ways to establish the translation memory are the same as those of the term base. It is unlikely to search and purchase high-quality parallel bilingual corpora related to IMTFE Transcripts since they are only published by Shanghai Jiaotong University Press in China, and there is currently no ready-made translation memory. The most reliable way is to accumulate through translation practice. When translators use CAT tools in translation, they will update the final renderings into a translation memory. With the accumulating process, the translation memory will become increasingly large and can be used for future translating projects. Although the translation memory obtained through translating practice takes time, the quality is assured. Another way is to utilize tools. Firstly, the original text is translated into Chinese with the help of a machine translation engine, and then the bilingual documents (the original text and its renderings) are made into a translation memory pursuant to the segment alignment. As for the online

machine translation, whether it is Google, Baidu, Youdao, or Sogou, there is a limit of 5000 characters, user will not be able to get a complete machine translation for free if the words are over 5000. This problem can be solved by splitting the document into multiple ones. However, if the text is large, it is wasteful of time and effort, so the best way is to obtain the complete translation with machine translation. The author can obtain a complete translation with the help of machine translation in the following three ways.

The first is the adoption of Google Chrome browser. The steps are listed as follows: open the original document, save it as an htm format document, and then open the document with the Google Chrome browser, Set the source language and target language for translation, and choose “Always translate English”, then it will perform translation automatically. If it does not, click the right mouse button on the web page and select “Translate to 中文” to start translation, drag the scroll bar of the web page down until the translation is completed, and finally copy the renderings to the new MS Word document.

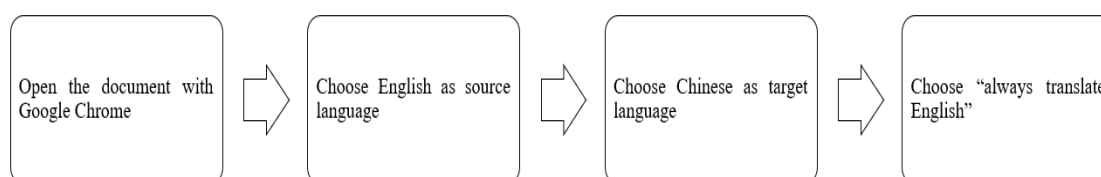


Figure1. The flowchart of Google Chrome translation

The second is to use Twinslator software to obtain the renderings completed by machine translation. The steps are listed as follows: register through the official website (<https://pe-x.iol8.com/>), launch the Twinslator, import the document, select machine translation, choose specific machine translation engine, then export the renderings when the translating is done by the machine.

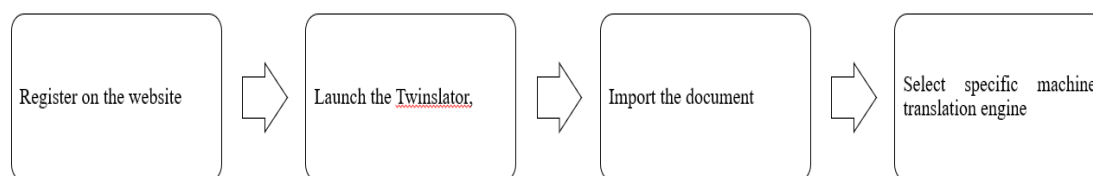


Figure 2. The flowchart of Twinslator translation

The third is to use the Transmate software. The steps are listed as follows: register the Baidu Translation API on Baidu's official website (<http://api.fanyi.baidu.com/>) to obtain the account and key, launch Transmate, enter the API account and key in the API settings, create a new project, import the original text, choose Baidu translation engine to obtain the renderings, export the renderings after completion, or directly export it as a TMX format translation memory.

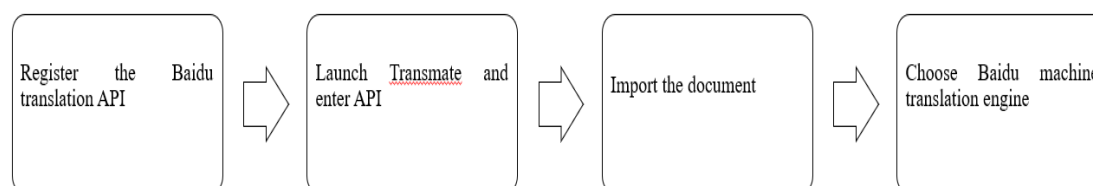


Figure 3. The flowchart of Transmate translation

In the translating practice, judging by the accuracy and expressiveness of the language the author found that, when translating from English to Chinese, the renderings of Baidu machine translation and Sogou machine

translation is far more accurate and hence more desirable. Thus, the author tends to adopt the renderings obtained from these two machine translations. The outstanding advantage of Transmate is that it can align the original text with the renderings obtained through machine translation, and export it as a TMX translation memory.

Translation memory is a parallel corpus and an important language asset for translation research and practice. However, the original text and its corresponding translation cannot automatically become translation memory. Alignment tools have to be used to process the original text and machine-translate text into a translation memory. The commonly-used tools by the author are Abbyy Aligner and LF aligner. The way to use these two types of software are basically the same. The steps are listed as follows: set the language pair, import the file of source language and the file of target language obtained through the aforementioned ways, then perform the automatic alignment. It is found that alignment accuracy could be as high as 99%. Those unaligned segments are manually edited to achieve full alignment. In the aligning process, the following two principles should be followed: basically, sentence alignment is based on full stop, in long sentences alignment based on semicolon. When the alignment is completed, export and save it as a translation memory of TMX format, which is a universal format that can be used in all mainstream CAT tools such as Trados or MemoQ. On the whole, the user interface of ABBYY Aligner is more user-friendly. The interface of LF aligner is relatively simple.

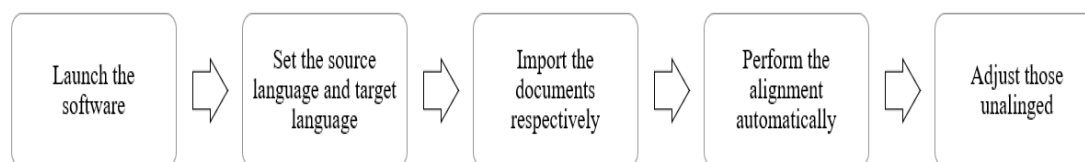


Figure 4. Flowchart of alignment

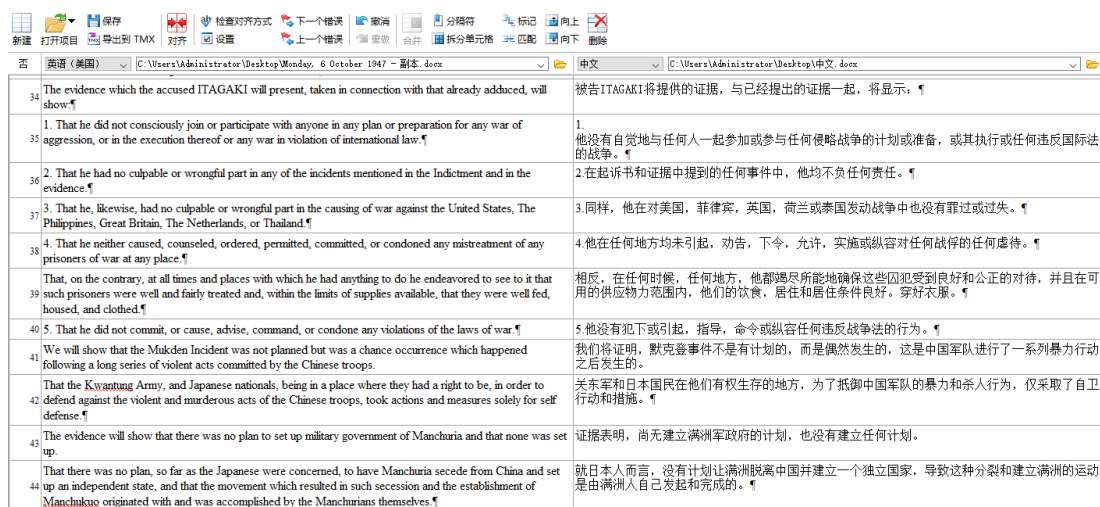


Figure 5. Screenshot of ABBYY Aligner user interface

The other two alignment tools are online tools which are also user-friendly, namely Tmxmall and Yeekit. The operating process is similar to the above-mentioned alignment tools. Users import the target file, perform alignment, edit and adjust the segments, and export and save it as translation Memory of TMX format. Different from the above alignment tools, users need to get a count from the official website and such intelligent alignment is completed online. One irritant the user may not like is that Tmxmall is not free now as it used to be.

4.2 Editing and Translating

In the stage of during-translation, the author uses Trados, a widely used CAT tool, to create a new translation project, set translation language pairs, add files, load the prepared term base, the translation memory to assist the author in translation. From the figure below, it shows that with application of term base, when the terms in the original text matches those in the term base, the term base set window will be automatically displayed for the author.

The author only needs to insert the correct term into the specified position with the shortcut keys. The whole process is fast and efficient, and repeated input of the same translation is avoided. For those terms with mistakes, the author can edit and modify at any time during the translation process, and those modified terms will be automatically updated to the term base, so the term base can be enriched by continuous addition.

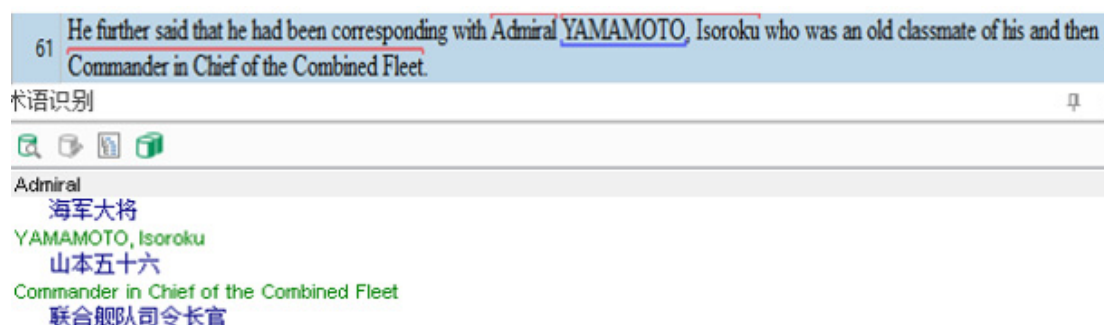


Figure 6. Screenshot of the original text segment and term base window

The core asset of the CAT tool is the translation memory. As mentioned above, there are various ways to make the translation memory. The one that can be used in translation is the memory accumulated through the previous translation with a CAT tool, and the other translation memory is obtained by machine translation and alignment tools. The two different translation memories could be used simultaneously or separately. The strategy adopted by the author is to use these two translation memories separately. First import the translation memory accumulated by the translator during the translating practice with Trados, and set the matching rate to 70% which is the minimum matching rate. Trados will automatically search the translation memory to match the original text, and the matched translation will be automatically filled in the translation column. After the matching is done the author can change the state of the matched segments and lock them. The purpose is to prevent the matched content from being polluted by the newly-imported translation memory, because the memory made by machine translation has a higher matching rate, but the quality is not as good as that of the previous one. If matched segments are not locked, the previously matched translation will be overwritten by the new translation memory.

During the translation process, the author can check, choose or edict the matched contents. For the fully-matched correct translation, the author just uses it. For the segment with a high match, the author simply has to translate the missing part. if there is a minor problem, it needs to be modified by the author himself. Those segments with more serious problems need active intervention of the author which would be demonstrated in Example 3. The author will apply different translating skills to correct those mistakes or even retranslate those segments with problems. The following examples will show the strategies.

Table 1. Examples of machine translation and post-editing

Example 1	At the close of court yesterday the last question had not been answered.
Baidu Translation	昨天庭审结束时，最后一个问题还没有回答。
Example 2	Did the Navy Ministry discuss with officials of the Foreign Office the plan called “Tentative Plan for Policy towards Southern Regions,” which is the exhibit which I just referred to?
Baidu Translation	海军部是否与外交部官员讨论了我刚才提到的“对南部地区政策的初步计划”展览？
Post-editing	海军省与外务省官员是否讨论了“对南部区域的初步计划”也就是我刚才提到的物证？
Example 3	Will you examine this document and state whether it is not a secret order issued by the Chief of the Naval Affairs Bureau at the time you held that office to the Chief of the Army Military Affairs Bureau concerning the transfer of 38 prisoners of war from Ofuna prisoner of war camp?
Baidu Translation	请您检查一下这份文件，并说明这是否是您担任海军事务局局长时，海军事务局局长向陆军军事事务局局长发出的关于从大船渡战俘营转移 38 名战俘的秘密命令？
Post-editing	仔细检查这份文件，说明你担任海军省军务局局长时是否发密令给陆军省省军务局局长，要求将 38 名战俘从大船战俘营转移出去？

It can be seen from Example 1 that the rendering of Baidu Translation for simple sentences is quite satisfactory. In this case, the author only needs to confirm the translation instead of typing each word. In Example 2 because

of the terminology, Baidu Translation does give the correct translation of “the Japanese Navy Ministry, the Foreign Office, and exhibit”. However, because the corresponding Chinese expressions for the three English expressions are already stored in the term base, the author only needs to select the correct translation in term base window. In Example three, the rendering is very heavy such as the unnecessary repetition of “the Chief of the Naval Affairs Bureau” in Chinese. In the English “secret order” is followed by two attributes which are “issued by the Chief of the Naval Affairs Bureau” and “concerning the transfer of 38 prisoners of war from Ofuna prisoner of war camp”. In Baidu’s translation the Chinese translation of the two attributes proceed the “secret order” which is tedious and unidiomatic. Therefore, the author deletes the unnecessary repetition, and rearranges the order of the sentence structure by stating whether the Chief of the Naval Affairs Bureau issued the order first and then the content of the order in Chinese which is more idiomatic in Chinese. Since this is the question from the prosecutor to the Japanese witness, the author uses imperative expression in the translation of “Will you examine this document” instead of the polite way from Baidu Translation. With human’s active intervention, the translation quality is guaranteed when the problem arises from machine translation. It can be seen that although machine translation is featured by deep learning capabilities, it cannot analyze the text according to the context, so it may cause inaccuracies or ambiguities in translation. In this case, the translator needs to actively improve and optimize the rendering. Practice has proven that loading the term base and translation memory in the translation project can greatly reduce the burden on human-memory, and unnecessary repetitive typing.

4.3 Quality Assurance After Translation

When the translation is completed, post-translation review is indispensable. The author needs to check spellings, missing translations, term consistencies, rendering consistency and other issues. During the review phase, the author had used Xbench software, which can detect missing untranslated segments; segments with the same original text but different renderings, segments with the same translation but different original texts, numerical errors, segments containing double blanks. Before using the tools, the author still needs to check the renderings carefully, because Xbench cannot replace humans to detect all the errors, such as typos in Chinese characters which need to be proofread by humans.

Another common problem is that since the translating interface of Trados is tabular, there are basically two columns, one is the source text and the other is the target text to be translated. The whole text is mainly divided into sentence segments pursuant to the punctuation such as period or semicolon. The translator completes the translation sentence by sentence. Because English and Chinese belong to different language families, there can be many differences in the translations. One of them is that English is hypotactic, while Chinese is paratactic. In the proceedings, the majority of the English sentences contain personal pronouns. There is no problem when reading the renderings alone, but there will be problems when the renderings are viewed holistically. Personal pronouns, which are indispensable in English are redundant in many cases in Chinese. Therefore, the author must check the translation in Trados to ensure quality with Xbench, it is still necessary to export the translation and proofread the final translation in a holistic way.

Segment	Source Text (English)	Target Text (Chinese)
1	No. 118 Kakinokizaka, Meguro-ku, Tokyo, Japan.	日本市目黒区麁木11号。
2	Friday, 5 December 1947	1947年12月5日星期五
3	Court House of the Tribunal War Ministry Building Tokyo, Japan	日本东京都旧陆军省大厦内远东国际军事法庭
4	Marshal of the Court:	法庭书记官:
5	The International Military Tribunal for the Far East is now in session.	远东国际军事法庭现在开庭。
6	Acting President:	代理庭长:
7	All the accused are present except MATSUI, who is represented by counsel.	除了由律师代理的松井外, 所有被告都出庭。
8	We have a certificate from the prison surgeon at Sugamo certifying that he is ill and unable to attend the trial today.	我们有巢鸭监狱外科医生的证明, 证明他病了, 不能参加今天的审判。
9	The certificate will be recorded and filed.	医生证明将记录存档。
10	Mr. Takahashi:	高桥先生:
11	Mr. Takahashi:	高桥先生:
12	We next call former Admiral SAWAMOTO, Yorio	接下来我们传唤前海军大将泽本鹤雄。
13	The witness, Admiral SAWAMOTO, was a former Vice-Minister of the Navy under both the accused Admiral SHIMADA and his predecessor, Admiral OKAWA, Kotoiro.	证人海军大将泽本是前海军次官, 隶属于被告海军大将岛田繁太郎和他的前任海军大将及川古志郎。
14	He will testify concerning the actions of the accused SHIMADA before and during his tenure as Navy Minister.	他将就被告岛田繁太郎在担任海军大臣之前和期间的行为作证。
15	Marshal of the Court:	法庭书记官:
16	Mr. President, the witness SAWAMOTO is in court.	庭长阁下, 证人泽本已经到庭。
17	He has previously been sworn before this Tribunal.	他曾在本法庭宣誓。
18	Acting President:	代理庭长:
19	You are warned that you are still under your previous oath.	我们警告你, 你之前的宣誓证明词仍然有效。
20	Q/O Mr. Takahashi:	高桥先生问:

Figure 7. Screenshot of the Trados tabular interface

4.4 Language Asset Management

When the author has completed all the above process, the author will update the project memory and the main memory. In the process of translation, the author constantly accumulates, expands and optimizes the translation memory. The final step for the author is to update the project translation memory and the main translation memory in Trados, which can be edited, updated, backed up, and shared and used at any time for future translation research and translation practices (Wang, 2015, pp. 82–85) .

5. Conclusion

A lot of researches have been done on the possibility and feasibility of combining machine translation and human translation. However, empirical research on how to apply the various translation technologies to the translation of trial transcripts is rarely found. The main contribution of the paper is proposing an approach of utilizing different translation technologies combined with human's active intervention in terms of translating quality and efficiency. First the author analyzes the features of the text which shows that there are many identical or similar expressions throughout the whole texts. Then, the author demonstrates what specific translation technologies are used and how they are used in different steps of the translation. Finally, by analyzing examples, it is found that there are merits of machine translation as well as setbacks. The study results show the application of translation technologies can enhance the translation efficiency, while the collaboration between translators and translation technologies can ensure both the translation efficiency and quality. Only when the translators are critically aware of the available translating tools can they be well prepared for the future changes.

Acknowledgments

Fund program: 2019 Zhejiang Province's Enterprise-University Collaborative Education Project for the 13th Five-Year Plan: Application of Translation Technologies in the Context of Artificial Intelligence

References

- Cheng, Z. Q., & Zhao, Y. H. (2014). Transcripts of International Military Tribunal for the Far East, Retrieval Tools and Name Issues. *Modern Chinese History Studies*, 5, 153–159.
- Cui, Q. L., & Lei, X. F. (2016). Human-Computer Interaction Translation Strategy Based on Text Classification, *Contemporary Foreign Languages Studies*, 3, 46–50.
- Fu, J. M., & Xie, S. (2015). Development of Translation Technologies and Translation Teaching. *Technology Enhanced Foreign Language Education*, 6, 37–41.
- Hu, Z. L. (2013). *Linguistics* (p. 25). Beijing: Peking University Press.
- Ilya, S., & Oriol, V. (2014). Sequence to Sequence Learning with Neural Networks. *NIPS*, 12.
- Kristine, B., & Tina, P. (2016). Translator-Computer Interaction in Action—An Observational Process Study of Computer-Aided Translation. *The Journal of Specialised Translation*, 90.
- Lu, W. Z., & Chen, H. (2014). Translation Technology and Cultivation of Professional Translators. *Technology Enhanced Foreign Language Education*, 3, 62–67.
- Lynne, B. (2002). *Computer-Aided translation technologies: A Practical Introduction* (p. 6). Ottawa: University of Ottawa Press.
- Wang, H. S. (2014). Research on Computer Aided translation technologies in The Information Age. *Foreign Studies*, 92–97.
- Wang, H. S., & Yang R. L. (2015). The Application of Computer Aided Translation (CAT) Technology in Modern Translation. *Language Education*, 3(01), 82–85.
- Xu, B., & Guo, H. M. (2015). Practice of Non-Technical Text Translation Based on Computer Translation Technology. *Chinese Translators Journal*, 36(01), 71–76.
- Zhang, F. L. (2020). On Machine Translation Technology in Legal Translation. *Technology Enhanced Foreign Language Education*, 1, 54.
- Zhou, X. H. (2013). Computer-Aided Translation Teaching: Methods and Resources. *Chinese Translators Journal*, 34, 91–95.

Copyrights

Copyright for this article is retained by the author, with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).