Research on Computer-Aided Translation in TCM Texts Translation

A Case Study of *Introduction to Chinese Medicine*

Jin Liu[1] Hanya Yang[2]

[1,2] *Hubei University of Chinese Medicine, Wuhan, Hubei 430065, China*

ABSTRACT

The purpose is to explore the limitations of Computer-Assisted Translation (CAT) in English translation of Traditional Chinese Medicine (TCM). Through literature research and case analysis, combining the "CAT(Computer-Assisted Translation) plus MT(Machine Translation) plus PE(Post-Editing)" translation model with the study of Chinese medicine translation, this paper takes the chapter "Five-zang organs and Six-fu organs" in the first chapter of *Introduction to Chinese Medicine* by Qin Bowei as an example for translation, putting forward that computer-assisted translation still has some inherent limitations in Chinese medicine translation, though it has certain advantages in some aspects.

Keywords: *Computer-aided translation, Chinese medicine translation, Introduction to Chinese Medicine, CAT.*

1. INTRODUCTION

Computer-assisted Translation (CAT) refers to a method of using computer technology to assist in translation. CAT system can provide translation memory, terminology base, machine translation and other functions to effectively improve the efficiency of translation. The emergence and wide application of computer-aided translation technology have promoted the development and direction of translation studies. At present, computer- aided translation is applied in various researches of TCM translation. However, due to the uniqueness and complexity of TCM language, the result of computer-aided translation is often not expressive by not handling complicated Chinese sentences and ambiguous sentences correctly, so it has some shortcomings in the application of TCM translation. Therefore, with the help of SDL Trados, a computer-aided translation tool, CAT(Computer-Assisted Translation)+MT(Machine Translation)+PE(Post-Editing) translation mode is applied in the translation practice of *Introduction to Chinese Medicine*[1]. The consequences have shown that there are various limitations of computer-aided translation in TCM translation ranging from vocabulary selection, syntactic structure and discourse expression, which are mainly due to the unique cultural background and language characteristics of TCM. Consequently, the accuracy and fluency of computer-assisted translation in TCM English translation still need to be further improved with the support of manual intervention and professional judgment.

2. OVERVIEW OF COMPUTER-AIDED TRANSLATION

Computer-aided translation refers to the process of using computers and software tools to assist professional translators in their translation work. It combines human translation and computer technology, aiming to improve the efficiency and quality of translation[2]. Featuring automation, normalization and standardization, computer-aided translation software can help translators quickly choose applicable words and avoid problems such as repeated translation and wrong translation. At present, there are various computer-aided translation software, such as SDL Trados, MemoQ, etc., which can effectively automate the translation process. In computer-aided translation, the translator inputs the source language text into the

translation tool, which provides translation suggestions based on a preset dictionary, terminology base, context and other information, automatically completes part of the translation, and memorizes the translated part. The translator can accept or modify these automatic translation results according to his or her own judgment, thus improving the efficiency of translation. Simultaneously, computer-aided translation also provides some other auxiliary functions, such as term management, consistency check, format conversion, etc., to help translators more conveniently carry out translation work.

3. CHARACTERISTICS OF TCM TRANSLATION

TCM is a traditional medical system with Chinese characteristics, which has a long history and rich cultural connotations. The translation of TCM has played an important role in the international communication of TCM, but there are many challenges due to the particularity and complexity of TCM language.

First of all, most TCM terms come from ancient Chinese and involve a large number of professional terms such as Yin and Yang, Zang-xiang theory, and five elements, which often have no direct equivalences in English language[3]. Therefore, it is needed to express TCM concepts by explaining, describing or borrowing similar English words in TCM translation, so that readers can understand the principles and treatment methods of TCM.

Secondly, there are differences between the theoretical system of TCM and Western medicine. The traditional and ancient beauty of TCM language requires reflecting in TCM translation. In the process of translation, it is necessary to pay attention to the use of sentence patterns, words and rhetorical devices in line with the characteristics of TCM, so as to create a translation style with long ancient rhyme, solemn and elegant.

Finally, with the development of modern medicine and the increase of international communication, TCM is also integrating with modern medicine, forming a model of integration of traditional Chinese medicine and Western medicine. Therefore, it is necessary to associate TCM theories and treatment methods with modern medicine through the rational application of modern medical terms and expressions, making the translation better meet the requirements of modern medicine and providing a bridge for international readers to understand TCM.

To sum up, the characteristics of TCM translation are mainly reflected in the translation of terms, translation styles and the combination of tradition and modernity. By effectively applying these characteristics, TCM concepts and thoughts can be better disseminated to international readers.

4. RESEARCH STATUS OF THE APPLICATION OF COMPUTER-AIDED TRANSLATION TO ENGLISH TRANSLATION OF TCM

At present, the research on the application of computer-aided translation in TCM translation is mainly focused on the English teaching of TCM and the application of various TCM books, including TCM classics and instructions of proprietary Chinese medicine. In terms of TCM translation English teaching, Zhang Jing [4] (2023) has proposed the combination of computer-aided translation and TCM English translation teaching, pointing out that the use of computer-aided translation tools can fully mobilize learners' autonomy, creativity and interaction in the "learner-centered" TCM English translation teaching, so as to improve the teaching effect. In the aspect of TCM books, it mainly involves the active application of computer-aided translation in the English translation of TCM classics and instructions of proprietary Chinese medicine. Li Binbin and Yao Xia [5] (2020) have studied the feasibility and limitations of computer-aided translation in the translation of TCM classics. They have explored the feasibility of applying computer-aided translation in the translation of TCM classics from three aspects: unifying related terms, using retrieval tools, and building SDL Trados translation memory bank of TCM classics. Wang Xiaomin [6] (2015) has studied the role of computer-aided translation in promoting the translation of TCM classics, improving the efficiency of TCM classics translation and ensuring the quality of translation by using the function of computer-aided translation, and hoped to better promote the culture of TCM to the world stage through this method. Zhang Lei [7] (2023) has translated the instructions for proprietary Chinese medicines by using the mode of "Machine translation (MT) + Computer-Aided Translation (CAT) + Post-editing (PE)", improving the translation mode, standardizing the current English translation of the instructions for

proprietary Chinese medicines, improving its translation quality and readability.

To sum up, some achievements have been made in the application of computer-aided translation in TCM translation, but the scope of the research is relatively narrow and there are still some limitations. Future studies can further explore the application of computer-aided translation in other fields, such as the translation of integrated Chinese and Western medicine. With the continuous development of computer technology, computer-aided translation will play a more important role in TCM translation and make greater contributions to the dissemination and promotion of TCM culture.

5. APPLICATION RESEARCH OF "CAT + MT + PE"

The "CAT + MT + PE" translation mode refers to Computer-Assisted Translation (CAT) combined with Machine Translation (MT) and Post-Editing (PE) [8]. In this mode, the translator uses CAT tools such as terminology case and translation memory bank to perform the translation work, while the results of the machine translation are first translated as drafts or references, and finally proofread and edited by humans. The CAT tool used in this paper is SDL Trados, which has a bilingual interface and provides a good environment for the translation process. In addition, its powerful translation memory bank and terminology base functions can not only improve the quality and efficiency of translation, but also ensure the consistency of translation in the project team[9]. Furthermore, the results of the translation can accumulate translation assets for future projects.

5.1 Building a Terminology Base

When applying computer-aided translation software for TCM translation, it is necessary to establish a terminology base in advance. As a computerized collection of terms, the advantage of the terminology base is that it can be quickly queried and updated in time. Therefore, it can ensure the use of uniform terms before and after the translation, prevent the confusion of terms, and ensure the professionalism and consistency of the translation.

First of all, in this translation practice, this paper selects the *WHO International Standard for Chinese Medicine Terminology* issued by the World Health Organization in 2022 as the standard term, and extracts the high-frequency glossary of "Five-zang organs and Six-fu organs" in the first chapter of *Introduction to Chinese Medicine* with the help of online corpus analysis tools. SDL MultiTerm 2019 (a powerful terminology base management and maintenance software) is used to produce the terminology base, so that the base could be imported into SDL Trados for later translation.

5.2 Building a Memory Bank

Translation Memory (TM) is a bank that stores sentence pairs (matching source and target sentence segments) and is the core of translation technology. This is a bank of sentences and phrases that have been translated. It improves translation efficiency, reduces duplication of effort, and ensures consistency in terminology and style. Building and using a translation memory bank helps to maintain the consistency of translation, improve the efficiency of the translator's localization, and improve the quality of translation.

Secondly, this paper chooses 17 sentences that appear more frequently in the book and uses ABBYY Aligner (a bilingual alignment tool that aligns the translation in sequence) to make them in TMX format and imports them to a new memory bank created in SDL Trados.

5.3 Setting up a Translation Project

The role of the SDL Trados translation project includes translation memory bank management, document preprocessing, terminology management, collaboration and project management, automatic translation, and machine learning. SDL Trados can be integrated with a machine translation engine to provide automatic translation capabilities[10]. Users can choose to use machine translation for pre-translation followed by manual revision. The role of the SDL Trados translation project is to provide a comprehensive toolset to support all aspects of a translation project, improving translation efficiency, and maintaining translation quality and consistency.

Then, this paper establishes a new project named "Introduction to Chinese Medicine" on SDL Trados, imports the original text of "Five-zang organs and Six-fu organs" in the first chapter of *Introduction to Chinese Medicine* into it, loads Youdao translation engine into the software for automatic translation, imports the previously made terminology base and memory bank, and checks and proofread the machine-translated text with the terminology base and memory bank, analyses the

translation results output by Youdao Translation engine.

5.4 Case Analysis

Introduction to Chinese Medicine, written by Qin Bowei, systematically introduces the basic theories, diagnosis methods, treatment principles, and TCM treatment methods for common diseases. The book explains the concepts and theories of TCM in clear and concise language, avoiding overly technical terms and complicated sentence structures and long-winded descriptions, making it easy for readers to understand.

In this paper, the "Five-zang organs and Six-fu organs" in the first chapter of *Introduction to Chinese Medicine* is selected as the translation practice for a case analysis on the results of computer-assisted translation through post-editing (PE). It can be concluded that there are lexical, syntactic and textual limitations of computer-assisted translation in the translation of Chinese medicine.

5.4.1 An Analysis of the Limitations of Computer-aided Translation in the Translation of TCM

5.4.1.1 The Limitations of Vocabulary Selection

The vocabulary selection in a computer-assisted translation system is limited by the vocabulary of its database, and the terms of TCM are very varied and complicated, involving many aspects of TCM theory, diagnosis method, medical therapy and so on. However, the vocabulary base of most computer-aided translation systems does not cover all TCM terms, which leads to the failure of the system to accurately select corresponding terms during translation.

Example 1

Source Text:

根据内脏的性质和作用分为五个脏、六个腑，又把另外的一部分称为奇恒之府和传化之府。

Computer-aided Translation:

According to the nature and function of internal organs, it is divided into five internal organs and six internal organs, and the other part is called Qiheng House and Chuanhua House.

Post-edited translation:

According to the nature and function of internal organs, it is divided into five zang organs and six fu organs, and the other part is called extraordinary fu organs and transmissive-transformative fu organs.

Analysis:

The machine translation of "五个脏" and "六个府" in this sentence adopts a literal translation. According to the *WHO International Standard of Chinese Medicine Terminology*, the term standard should be translated into "five zang organs" and "six Fu organs". The translation of "奇恒之府" and "传化之府" are transliterated, and they should also be translated into "extraordinary fu organs" and "transmissive-transformative fu organs" according to the terminology standard. The error of the computer-aided translation here mainly lies in the translation of terms, because SDL Trados cannot automatically identify terms and replace the original translation under the loading of the machine translation, so post-editing is needed.

Example 2

Source Text:

脏和腑俱为内脏，其区别是:五脏藏精气而不泻，六腑传化物而不藏。

Computer-aided Translation:

Both viscera and viscera are internal organs, and the difference is that the five viscera store essence without diarrhea, while the six viscera transmit chemicals without hiding them.

Post-edited translation:

Zang organs and fu organs are both internal organs, and the difference is that the five zang organs only store up essence qi and will not discharge it, and the six fu organs only transport and transform food and will not store it up.

Analysis:

The words "脏" and "腑" do not keep to the terminology standard. In addition, the translation of some words is not very accurate, such as the translation of the former "藏" is inconsistent with the latter "藏"; "泻" not only literally has the same meaning as "diarrhea" in the English language, but its actual meaning is wider, it should be translated as "discharge"; "传" "化" are two parallel verbs "transport and transform", but they are mistranslated into verb-object combination "transmit chemicals".

5.4.1.2 The Limitation of Syntactic Structure

When dealing with TCM translation, it is often difficult for computer-aided translation systems to accurately capture the grammatical structure of Chinese sentences. Chinese syntactic structure is more flexible, focusing on subjective expression and rhetorical devices, while English pays more attention to objective expression and rigor of sentence structure. Computer-aided translation systems often fail to accurately identify such differences, leading to a decline in the accuracy of the translation results.

Example 1

Source Text:

三焦为决渎之官，主行水。

Computer-aided Translation:

The san jiao is known as "the dredging and draining organ", governing circulating water.

Post-edited translation:

The san jiao is an organ that controls water circulation.

Analysis:

The original meaning of "决渎" in this sentence is "dredging waterway", which is synonymous with "主行水" in the later text, so it can be translated into a sentence with an attributive clause, which is simple and easy to understand. Most of the translation here is literal translation. For the unique language expression form in TCM, the relationship between the sentences is not recognized by computer-aided translation, resulting in lengthy and obscure translation results.

Example 2

Source Text:

心脏本身不健全，或受情志的刺激，或因病邪的侵犯，就会出现心悸、惊惕、失眠,或善忘、喜笑失常，或谵语、神识昏迷等症。

Computer-aided Translation:

If the heart itself is not perfect, or stimulated by emotions, or invaded by pathogenic factors, palpitations, convulsions, insomnia, forgetfulness, abnormal laughter, delirium, coma and other diseases will occur.

Post-edited translation:

When the heart itself is not sound, or is stimulated by emotions, or is invaded by pathogenic factors, it will feel palpitations, convulsions, insomnia, forgetfulness, abnormal laughter, or delirium, coma and other diseases.

Analysis:

The computer-aided translation of this sentence is not very readable. Some of the words are not translated accurately or properly. For example, the word "健全" translated to "sound" is more appropriate to describe the heart. At the same time, the language in the machine translation is a little blunt and does not highlight the subject of the sentence. In the later section "就会出现" here lacks of subject, which corresponds to the heart in the front, while the machine translation directly translates into "will occur", which is not in line with the natural expression.

5.4.1.3 The Limitations of Discourse Expression

The limitations of text expression in TCM are mainly reflected in the accurate understanding and correct use of TCM terms. TCM terms have rich connotations and specific cultural backgrounds, and their meanings often cannot be directly corresponded to English words. This requires translators to have sufficient professional knowledge and contextual understanding when using computer-aided translation systems in order to correctly express TCM concepts.

Example 1

Source Text:

心生血，主藏神。肝藏血，主谋虑。

Computer-aided Translation:

The heart produces blood, and the Lord hides God. The liver stores blood, and the mastermind cares.

Post-edited translation:

The heart generates blood vessels and stores the spirit. The liver stores blood and controls planning.

Analysis:

The original sentence is concise and comprehensive, and the beauty of this language should be reflected as much as possible in translation. The first sentence and the second sentence have only one subject. The word "主" here is not the meaning of "Lord", what it wants to express is "the heart and liver are mainly responsible for...", only the following verb "藏" can be translated, the machine translation mistranslates

into “Lord”. The following “神” is the meaning of “spirit”, the computer-aided translation also wrongly translates into “God”.

Example 2

Source Text:

肝开窍于目，其充在筋，其华在爪……其华在发。

Computer-aided Translation:

The liver begins to resuscitate in the eye, its filling in the tendons, and its splendor in the claws……its splendor in the hair.

Post-edited translation:

The liver opens into the eyes, and the tissue of the liver is sinew, the lustre of the liver shows in the nails…the lustre of the liver shows in the hair.

Analysis:

The original sentence reflects the vagueness and abstractness of the Chinese medicine language, the original meaning is “visual function depends on the nourishing of liver blood, the tissue in the liver is tendon, the color of the nail can reflect the waxing and waning of liver blood and liver qi”. “开窍” refers to where it connects with it; and “其华” refers to the external expression. However, the computer-aided translation is completely literal and fails to accurately translate the meaning expressed in the original text.

5.4.2 Discovery

Although computer-aided translation software has achieved remarkable results in improving the efficiency and quality of translation, it still has some limitations, mainly including the following:

- Terminology standards are not uniform. When the terminology base is imported into SDL Trados in advance, the recognized terms cannot be automatically replaced with the original machine-translated terms when the translation engine is loaded, and post-editing is required later.
- The sentence matching of the translation memory bank is low. In the context of loading the translation engine, the translation memory bank fails to match the sentence preferentially. Pre-translation is required before machine translation. Searching and matching corpus in the translation memory bank has shown a low matching degree, and some corpus fails to be matched successfully. If the matching degree exceeds the set minimum value but is not high, the system will directly apply the matching corpus without automatic correction.
- The translation of the target text is too literal. The translation language used in machine translation is too literal, failing to reflect the sentence structure of the original text, and some sentences are too straightforward.
- The wording of the target text is not accurate. Machine translation has a big problem in the use of words. It often literally translates the words in the original text according to its own understanding, resulting in inaccurate wording in the target text.
- The logical relationship between the sentences is not clear. Machine translation does not properly deal with the logical relations between sentences, which leads to mistranslations and omissions in the translation.

To sum up, the application of computer-aided translation software in TCM translation is feasible. The efficiency and accuracy of translation can be improved through the establishment of a terminology base and memory database, but the translation quality still needs to be modified manually. Therefore, it needs to be flexible in practical application, combined with professional background and experience, in order to achieve better translation results.

6. CONCLUSION

By establishing a terminology base and unified management of the vocabulary of TCM professional terms, computer-assisted translation technology can automatically identify terms in the text and provide translation suggestions to ensure the standardization and accuracy of TCM terms. At the same time, by establishing a memory bank, it can automatically save the translation results of translators and provide automatic translation suggestions when the same or similar sentences are encountered next time. This is very useful for common sentence patterns and forms commonly seen in TCM translation, which can improve translation efficiency and consistency. However, because the corpus of TCM translation is relatively scarce compared with other fields, it brings some difficulties to the training and evaluation of computer-aided translation systems. In practical applications, researchers often need to rely on a

manual glossary and professional vocabulary database, and there is no complete and reliable corpus for professional fields. Secondly, the transfer of professional knowledge in the field of TCM requires certain background knowledge and experience, which is often difficult for the existing machine-assisted translation system to handle such professional translation tasks. For example, some special concepts and theories in TCM culture often cannot be directly translated into English through machine translation, which requires translators to have relevant professional knowledge and translation experience. At the same time, there are some differences in vocabulary and grammatical structure between Chinese and English, and these differences will also exist in the translation of TCM. For example, Chinese uses more four-character words and allusions, while English uses fewer of these expressions. Therefore, computer-aided translation systems need to be able to recognize and process these language features to generate translation results that conform to English grammar and expression habits.

At present, the systematic research and application of computer-aided translation systems in TCM translation into English are still limited in academia and industry. It is necessary to further strengthen basic research, build a corpus of professional fields, and further improve the construction of TCM vocabulary, so as to provide more comprehensive and accurate translation resources. In addition, the application of more effective computer-aided translation systems in TCM English translation can be explored by combining machine learning and artificial intelligence technologies.

ACKNOWLEDGMENTS

Funds Project: 1. General Project of Philosophy and Social Sciences of Hubei Provincial Department of Education in 2022 (No. : 22Y083). Project Name: Research on the Translation and Communication of Traditional Chinese Medicine Classics based on the English Translation of Compendium of Materia Medica.

2. “Traditional Chinese Medicine Inheritance and Innovation Plan” of Hubei University of Chinese Medicine in 2023. Project Name: Research on the Current Situation and Strategies of Traditional Chinese Medicine Culture Communication in Jingchu.

REFERENCES

[1] Qin Bowei. Introduction to Traditional Chinese Medicine [M]. *Beijing: People’s Medical Publishing*, 2018: 23-28.

[2] Wen Jun, Ren Yan. A Review of Domestic Research on Computer-aided Translation [J]. *Foreign Language Audio-visual Teaching*, 2011(03): 58-62.

[3] Sun Junfang. Cultural Characteristics of TCM Language and Translation Strategies [J]. *China Journal of Basic Medicine of TCM*, 2012, 18(10): 1151-1153.

[4] Zhang Jing. Application of Computer-aided Translation in Learner-centered English Translation Teaching of Traditional Chinese Medicine [J]. *Western Chinese Medicine*, 2023, 36(01):45-47.

[5] Li Yunyun, Yao Xia. Feasibility and Limitation of Computer-aided Translation in TCM Classics translation [J]. *English Teachers*, 2019, 20(11): 90-92.

[6] Wang Xiaomin. The Promotion Effect of Computer-aided Translation on the Translation of Traditional Chinese Medicine Classics - A Case Study of Siku Quanshu [J]. *Journal of Guiyang University* (*Social Sciences Edition*), 2015, 10(03):85-87+95.

[7] Zhang Lei, Tan Xuejiao, Gu Xingmei. Application of Computer-aided Translation Software Trados in the Translation of Proprietary Chinese Medicine Instructions [J]. *Chinese Pharmacy*, 2023, 34(03): 366-369+379.

[8] Xiao Zhiqing, Wei Guangfeng. A Study on the Collaborative Translation Model of CAT + MT + PE Teachers and Students under the Concept of Human-Computer Interaction [J]. *Language Education*, 2019, 9(02):69-74.

[9] Wu Lihua. An Empirical Study on the Quality and Efficiency of Machine-assisted Translation of Medical Texts [J]. *Chinese Journal of Science and Technology Translation*, 2018, 31(01):33-36+52.

[10] Feng Zhi-Wei, Zhang Deng-Ke. Complement of Machine Translation and Human Translation [J]. *Journal of Foreign Languages (Shanghai International Studies University)*, 202, 45(06):77-87.

[11] Hu Caixia. A Study on TCM Vocabulary Translation Based on Jinghua Yinyun [J]. *English Square*, 2021(14):26-29.

[12] Sun Chao, Zhang Hongying. Analysis on the Metaphorical Cognitive Features of TCM Language [J]. *China Journal of Library and Information of Traditional Chinese Medicine*, 2017, 41(01):20-24.

[13] Zhang Zhe, Xu Li, Min Ling. Research on the Translation Standards of TCM Terms Based on the English Corpus of TCM [J]. *China Medical Review*, 2022, 19(20):189-193.

[14] Wang H S, Wang Y. A Review of Domestic Research on Computer-aided Translation (1980-2021) [J]. *Journal of Foreign Languages*, 2019, 39(02):130-138.

[15] Xu Bin, Guo Hongmei. Practice of Non-technical Text Translation Based on Computer-aided Translation Technology [J]. *Chinese Journal of Translation*, 2015, 36(01):71-76.

[16] Sun Jiayi. Application of Computer-aided Translation in Subtitle Translation [J]. *Electronic Technology*, 2023, 52(02): 374-375.

[17] Man Jing, Liu Minnan. Production Instruction Translation under Computer-Aided Translation [J]. *Computer Knowledge and Technology*, 2019, 16(09):284-286.