Exploration and Research on Computer Aided Translation Technology and English Translation Practice

To cite this article: Guoping Chen 2019 J. Phys.: Conf. Ser. 1213 042011

View the article online for updates and enhancements.
Exploration and Research on Computer Aided Translation Technology and English Translation Practice

Guoping Chen
Guangxi Lantian Aviation Technical College
Laibin, Guangxi, China

ABSTRACT: In recent years, the rapid development of science and technology and economic globalization have made the demand for translation more and more in all walks of life. Therefore, computer-aided translation technology has received more and more attention. Combining modern information technology with traditional translation theory, this paper expounds the principle of computer-aided translation technology, and uses the combination of translation memory technology and mixed strategy of statistics and examples to realize computer-assisted translation system. According to the specific English translation practice research, it is proposed to use translation memory technology to optimize translation process, improve translation quality and translation efficiency, and hope to bring work help to the majority of people engaged in English translation, so that their work can be carried out more smoothly.

1. INTRODUCTION
In recent years, globalization of economic development and science and technology are also constantly developing and the traditional field of translation is also being hit by unprecedented impact. Economic globalization has increased the exchanges between languages. This change has brought more opportunities to the translation industry and broadened the market for translation. The continuous development of science and technology has also made the translation technology mature. However, opportunities and challenges are inseparable, and opportunities have certain challenges. The rapid development of the global economy puts higher demands on the people engaged in translation work, not only to ensure the quality of translation, but also to put forward higher requirements for the speed of translation. Translators are required to meet growing market demands in terms of quality, quantity and speed. It is always an urgent desire for people to use advanced information technology to achieve efficient translation between different languages and improve the efficiency of information acquisition. At the beginning, people didn't know how to realize the combination of information technology and translation activities. After continuous research, they finally found the combination of information technology and translation activities and continued to study in depth. In this process, they deeply understood the understanding of translation activities. The summary of translation theory has played a very important role.

2. THE BASIC PRINCIPLE OF COMPUTER-AIDED TRANSLATION
Computer-aided translation is a translation solution based on translation memory (TM) technology, referred to as CAT (Computer Aided Translation). Computer-assisted translation is a translation strategy used by translators to participate in the translation process using computer programs. The basis of computer-assisted translation is the corpus and the thesaurus of system storage.

In general, the translation process consists of four steps: analysis, transfer, reorganization and
inspection. Analysis refers to dividing a complete sentence into independent units and understanding its basic meanings as much as possible; translating refers to converting one language symbol into another, and recombining refers to using another language. The process of expressing the source information by the symbol; the inspection refers to checking whether the information content of the two language symbols is the same, and whether the content of the original text is accurately translated.

The working principle of computer-aided translation technology is: CAT system divides the complete sentence into independent sentence units, such as words and phrases, according to the judgment of the code program and according to the system's own thesaurus, the system memory of the thesaurus, users The established thesaurus, etc., gives advice on how to translate; at the same time, the system automatically searches for similar or identical existing translation resources (such as sentences, paragraphs, etc.) in the translation memory, if a search is given and a reference translation is given. In this process, the translator can adjust the division of the sentence unit by the CAT system, and then reorganize the sentence unit according to the translation suggestions given by the system, complete the information reorganization, reassemble into sentences, or according to the reference translation displayed by the system or Adopt or modify the reference translation.

The biggest advantage of computer-assisted translation is the retrieval and storage of information. Using computer-assisted translation, the system can store a large number of the professional vocabulary, thus reducing the labor of the translator in the translation process; the computer-assisted translation technology can automatically store the content translated by the user according to its automatic memory function. When the user translates a sentence, the system automatically searches for the stored content. If the currently translated sentence is the same as the sentence that the user has translated, the previous translation result can be automatically given; for similar sentences, can give translation suggestions and references. In the integrated network's translation memory system, system data is stored on the LAN server, so translators can share the database regardless of the geographical distance between team members.

3. KEY TECHNOLOGY ANALYSIS OF CAT SYSTEM
At present, the internationally famous CAT software has TRANDOS and there is Yaxin in China. According to Internet survey data in 2008, TRANDOS market share accounts for 80% of the entire industry, and in 2014, this data dropped to 70%. The main reason is that the emergence of a large number of CAT software in recent years has seized the market share of TRADOS. In general, CAT software generally has a series of functions such as term base management, translation memory, translation project management, corpus processing and application. The architecture diagram of the CAT system is shown in Figure 1.
3.1 Preprocessing Module
The auxiliary translation preprocessing means that for the translated text uploaded by the user, the system first recognizes the language and performs segmentation, clause and the like. According to the mark of the XML file format, the related data such as the number of pages, the number of segments and the line number of the translated text are stored in the database.

3.2 Term Management
One of the main functions of CAT software is terminology management. Each article contains many words. Checking the consistency of words becomes the key content of editorial proofreading. If there is an omission in this work, it will seriously affect the quality of the translation, greatly hinder the reader's understanding and even misunderstand the reader. To solve this problem, the CAT system has developed a terminology management module for standardizing terminology. When translating, the system can automatically identify the corresponding terminology translation in the terminology library established by the translator. At present, the CAT system has a term management function. The content in the terminology library can be imported through CAT tools, dictionaries, manuals, etc., or directly into an Excel table, and the imported content is stored in the term management module.

3.3 Translation Memory
Translation memory technology (TM technology) is the core technology of CAT. The system memory is a file stored locally. At the beginning, the system memory is initialized to an empty file and the user continues to use it through continuous correction. The translation is adjusted and the original text and the translation are separately saved to the translation memory, so that the scale of the translation memory is continuously expanded, thereby improving the accuracy of translation and thus establishing a more perfect system memory. The CAT system creates a memory library in terms of terms or sentences. When translating, it manually interprets the incorrect translation of terms and sentences during project translation, translates into correct translations without errors and sets the source language and target. Language, which stores terms, sentences and corresponding translations into the system memory.
3.4 CAT System Process
The system flow of CAT imports the source language documents in doc, txt and other formats into the CAT system. The CAT system first classifies the text, uses the word segmentation tool to segment each sentence in turn and performs the similarity of the sentences according to the TM technology. Calculate to check if it can match completely, directly output this sentence. If it can't match, the translator needs to manually proofread the translation of the bureau, then store the translation of the sentence into the translation memory and output the sentence translation until the sentence translation of the entire document is completely translated, the target language text is output, and finally the text document is exported in doc, txt and the like. The CAT system flow chart is shown in Figure 2.

![Figure 2 CAT system process](image)

The basic flow of the preprocessing module is shown in Figure 3 below:

![Figure 3 Preprocessing module workflow](image)

4. ENGLISH TRANSLATION RESEARCH COMBINED WITH CAT
Computer-aided translation software provides a new perspective and method for studying English translation. This paper combines CAT to conduct some research on English translation, including the following aspects. Collecting CAT for English translation can not only improve the speed of translation,
but also improve the efficiency of English translation.

4.1 Extraction of Terms, Proverbs, Idioms
For the translation of English majors, the English term extraction function is quite important, which can reduce the workload of translators. Translators use computer translation software to extract terms from the beginning of the translation work, import and save to the term base, and the unified work of terminology can play a multi-functional role. However, due to limitations in translation technology, computer-aided translation software can only perform terminology management functions for terminology applications of individual words. However, in higher quality translations, the translation of colloquialism, multi-word terms and idioms should be consistent in the whole text. However, current software is not able to achieve this effect and still needs to be verified by the translator. Translators must pay attention to this aspect when using the CAT system for English translation.

4.2 Translation Memory Retrieval and Matching
Before the translation staff can translate the translation memory through CAT, find similar or identical texts, bilingual translation cases of sentences, and translate English according to the case. Translation memory retrieval is performed in the same way as general search engine retrieval, including fuzzy matching and full matching. In the translator's actual work, the probability of being able to retrieve in the memory is generally low, such as financial translation, literary translation, etc.; but for technical documents with high repetition rate, the probability of searching for matching cases is relatively high. Therefore, for English translators, technical documents are easier to match, saving translation time and improving efficiency. However, the fuzzy matching of computer-aided translation technology is not ideal at present. For fuzzy matching, it is necessary for the translator to perform manual intervention to ensure the correctness of the translation.

4.3 Computer Aided English Translation Quality Control
For translators, quality is important. At present, most CAT systems have a simple review function, which can give hints for some obvious simple errors, avoid obvious problems such as missed translation and grammatical errors, and have certain effects on quality control. However, this review function is not powerful enough and it is not possible to review whether the more complicated terminology is appropriate and whether the language expression conforms to the language habits of the target language. In the future development, there may be cooperation between computer-aided software and review software to further improve the review function and improve the quality control effect.

4.4 Exploration of English Translation Practice
In order to ensure that the results of the practice are more objective and comprehensive, this paper uses TRADOS to conduct multi-domain scientific English translation.

First of all, this paper selects a scientific paper on concrete. Then the whole paper is divided into two parts: the part that is not input to TRADOS (about 2,000 words) and the part that is input into TRADOS (about 4,000 words). The purpose of this paper is: the part that is not input to TRADOS is to detect the degree of recognition of the same or similar translation content by TRADOS; the part input to TRADOS is to store certain term bases and sentence libraries for TRADOS. And use this to carry out a case study summary: the degree of assistance played by the degree of matching and the cause of the degree of matching.

Secondly, for the 4000 words input to TRADOS, this article basically follows four steps: First, input to TRADOS 1000 words, that is, system word inventory into 1000 words; secondly, input to TARADOS 1000 words, that is, system word inventory into 2000 words; again, input to TDRADOS 1500 words, the system word inventory into 3500 words; finally, input to TRADOS 500 words, the system word inventory into 4000 words. Based on the deposit of the font, this article will randomly select some of the content that has not been saved to TRADOS for testing, and the degree of recognition of the new sentence when TRADOS accumulates a certain number of translation memories.
Finally, through the input of the above four steps and the corresponding test result data, it is analyzed and summarized: from the perspective of translation memory capacity formed by different vocabulary, the influence of translation memory size on matching degree and auxiliary effect has changed. From the structure of the article, when the number of translated words input into TRADOS reaches 4000, the system saves the content including the verification experiment and the auxiliary experiment in the experimental report, as well as the matching degree and auxiliary effect of TRADOS. For the part that is not input to TRADOS, the test shows that the matching degree is low, so the auxiliary effect is poor. On the contrary, the matching degree is high and the auxiliary effect is better.

Practice has proved that the use of CAT for English translation can improve the quality and efficiency of English translation.

5. CONCLUSION
Combined with the development of modern information technology and traditional translation theory, this paper expounds the basic principles of computer-aided translation and uses the combination of translation memory technology and mixed strategy of statistics and examples to realize computer-assisted translation system. According to the specific English translation practice research, it is proposed to use translation memory technology to optimize translation process, improve translation quality and translation efficiency and hope to bring work help to the majority of people engaged in English translation, so that their work can be carried out more smoothly.

REFERENCES
[1] FU Yanfu. Translation memory theory and a few computer aided translation software review[J]. Journal of Social Science of Human Medical University, 2010, 12 (2): 107-108. (in Chinese)
[2] XU Bin. Computer aided translation: concept and tools[EB/OL]. [2010-01-31]. http://www.docin.com/p-290885839.html. (in Chinese)
[3] JIN Guangsa. Present situation and development trend of the technology of computer aided translation[J]. Journal of Shenyang Institute of Engineering(Natural Science), 2010, 6 (3): 264-267. (in Chinese)
[4] Zhu Yubin. Technology is based on the road, the Taoist device is the same – thinking about the computer-assisted translation course of MTI in local colleges and universities [J]. Chinese Translation, 2012, (3): 63-65.
[5] Zhong Weihe, Zhao Junfeng. Interpretation of the National Standards for Teaching Quality of Undergraduate Majors[J]. Foreign Language Teaching and Research, 2015 (2) : 289-296.
[6] Xu Yunfeng. Computer-assisted translation theory and practice based on information technology [J]. Aviation Science and Technology, 2014, (6): 54-57.
[7] Li Tongliang. Higher vocational application English major CAT course construction [J]. Education and Teaching Forum, 2015, (1): 202-203.
[8] Zhang Yijun and so on. 2013. Computer-assisted translation: theory and practice [M]. Xi’an: Shaanxi Normal University Press
[9] Hou Xiaofang. Computer Aided Translation Software Usability Study - Memo Q and Trados Case Study [D]. Jinhua: Zhejiang Normal University, 2013.
[10] Zhu Yubin, Chen Xiaojian. Comparative study of four common computer-aided translation softwares at home and abroad [J]. Foreign Language Electrification Teaching, 2013(1): 71.