Machine Translation and Computer Aided English Translation

Chuanhua Xu* and Qianqian Li
Shandong Management University, Shandong, China

*Corresponding author e-mail: 14438220120014@sdmu.edu.cn

Abstract. With the development of globalization, the contacts and exchanges between countries in the world are getting closer and more frequent. As the most widely used language in the world, English is favored and valued by many non-English speaking countries. In order to facilitate communication, people use computers technology has developed a computer translation tool. The purpose of this article is to explore the difference between machine translation and computer-assisted translation in computer translation, compare the advantages and characteristics of the two in English translation, and point out the direction for future English translation trends. This article uses Google Translator, Youdao Translator and Baidu Translator in machine translation software and Wordfast, WordFisher and iCAT Huoyun Translator in computer-aided translation software as the experimental research objects to compare and analyze the different features and advantages between machine translation and computer-aided translation, and compare the error rate and match rate of the two in English translation. Finally, the experimental results show that the total number of errors in computer-aided translation software translation is less than that of machine translation software. Among them, WordFisher has the least number of errors, 25 errors in vocabulary, 31 errors in sentences, and 3 errors in punctuation, a total of 59 errors; with the improvement of translation matching rate, the number of translation errors of English professional terminology of machine translation software is increasing, and the error rate is getting higher and higher, when the matching rate is 90-100, Google Translate becomes the software with the most translation errors, and the number of errors reached 78. Studies have shown that computer-assisted English is better and more accurate than machine translation.

Keywords: Machine Translation; Computer-Aided Translation; English Translation; Translation Memory; Comparative Analysis

1. Introduction
Under the wave of globalization, along with the advancement of the information age, the application of computer technology in English translation has become more and more extensive. As the universal language of the world, English has become the official language of communication and exchanges between countries. However, for non-English speaking countries, it is still difficult to learn English. Therefore, in order to facilitate economic, cultural and even political exchanges between countries,
computer translation has appeared. Computer translation includes machine translation and computer-aided translation. They are both language translation tools based on computer technology. However, there are certain differences between the two in essence. The purpose of this article is to study and analyze the different characteristics of machine translation and computer-aided translation. Compare the advantages of the two, provide help for better English translation, and also point out the direction for future English translation trends.

In fact, as early as the 1950s and 1960s, there were people who had studied computer translation. However, computer technology was not mature at that time, and people only initially explored the role of computers in language translation. Later, with the development and maturity of computer science, computer translation really entered the public's field of vision. Huang G, Zhang J, and Zhou Y pointed out that computer-assisted translation (CAT) systems are the most popular tool used to help human translators perform language translation effectively. They said that in order to further improve translation efficiency, people are more and more interested in applying machine translation (MT) technology to upgrade CAT [1]. Al-Furaih RF introduces a new resource to support researchers from different disciplines, such as computational linguistics, descriptive translation research, computer-aided translation technology, Arabic machine translation applications, cognitive science and translation pedagogy, etc. [2]. Torregrosa D, Juan Antonio Pérez-Ortiz, and Forcada M L believe that Interactive Translation Prediction (ITP) is a form of computer-assisted translation that assists professional translators by providing continuous, context-based computer-generated suggestions. They stated that although most of the most advanced ITP systems use the glass box approach, which means they are tightly coupled with improved machine translation systems, the black box approach does not require access to the internal working principles of the bilingual resources used [3]. In addition, Arcan M, Turchi M, Tonelli S and others extract and integrate automatically aligned bilingual terms into a statistical machine translation (SMT) system in a computer-aided translation scheme. They evaluated the proposed framework, took a small set of parallel documents as input, collected bilingual terms in specific fields and injected them into the SMT system to improve translation quality [4]. El-Rahman S A, El-Shishtawy T A and El-Kammar R A introduced the real-world technology of a machine-assisted translation system. In this technology, they divide the system dictionary into a multi-module structure, which is used to quickly retrieve the Arabic features of English words. Each module is accessed through an interface containing the necessary morphological rules, which leads the search to the correct sub-dictionary [5].

Looking at previous achievements, we can see that the application of computers in language translation is relatively mature. From the initial simple machine translation to the later computer-aided translation, the quality and efficiency of English translation are improving. However, among these research results, it is not difficult to find that they are all researches on a certain aspect of computer translation, that is to say, they are purely studying machine translation or computer-aided translation. Few people have explored the difference and comparison of advantages between the two. Based on this, this article aims to study and analyze the differences between the two in English translation, and compare their advantages and characteristics.

2. Basic Theories of Machine Translation and Computer-Aided Translation and Their Application in English Translation

2.1. Machine Translation

2.1.1. Principles of machine translation
Machine translation is the process of using computer technology to completely transform the natural source language of human beings into another natural target language. The realization of this process does not involve any human involvement, and the computer system analyzes the grammar and structure features of the language and then automatically generated and converted into the target language. The recognition of the source language by machine translation is an automatic recognition
based on computer vision. However, because this recognition is based on the automatic recognition of the machine, it is inevitable that there will be deviations and errors in the conversion of complex sentences and grammar, so the accuracy rate of machine translation is low [6].

2.1.2. Defects of machine translation

Human natural language is a very flexible and changeable symbol. The expression and conversion of sentence meaning between different languages have a certain degree of artistry, which can not be expressed clearly by simple literal translation. In the past, in manual translation, in order to reflect the artistry and beauty of the original sentence, the translator would be cautious and flexible in choosing words and constructing sentences. While expressing the meaning of the sentence completely, it would also skillfully add sentences that conform to local language habits. This is something that machine translation cannot do. Machine translation is to analyze the words, grammar and structure of the source language, and then call the system language database to rearrange and combine the source language to form a target language similar in structure to the source language [7]. It is essentially a process of literal translation. This literal translation lacks an understanding of the target language habits. Therefore, the translated sentences are relatively rigid and sometimes grammatical deviations may occur.

2.2. Computer Aided Translation

Unlike machine translation, which takes the computer as the protagonist, computer-aided translation takes the translator or translator as the protagonist. When the translator is translating, the computer assists the translator through various means. The principle is to build a translation memory in the background with the help of translation software or tools, and use the data storage function of the computer system to convert the previously translated content and the text is stored, and when it is translated next time, the computer can automatically search for and generate similar sentences, grammar, and sentence patterns based on these translation memories, and the translator can then decide whether to adopt and use them among these sentences [8]. In essence, this kind of translation memory of computer-assisted translation is essentially a translation suggestion, and the final decision object is still the person itself, which improves translation quality while also improving translation efficiency.

2.3. Comparison of Machine Translation and Computer-Aided Translation in English Translation Applications

Machine translation and computer-aided translation are collectively referred to as computer translation. The two are language translation tools based on computer science and technology, but they are essentially different to a certain extent. Compared with machine translation, the application effect of computer-aided translation is better and more significant, and its conversion accuracy of grammar, sentence pattern and sentence structure will be higher [9]. The differences between the two are specifically manifested in the following aspects:

2.3.1. The time of appearance is different

Machine translation was born in the initial stage of the development of computer technology, and computer-aided translation emerged on the basis of the development of machine translation, and is the product of the mature development of computer technology.

2.3.2. Different design concepts

The original intention of machine translation design is to allow machines to replace humans in translation, thereby reducing the workload and difficulty of human work. It is more about pursuing speed and efficiency, while ignoring the quality of translation. The audience of machine translation is more people who lack English and other foreign language foundation or have low foreign language proficiency, to help them solve basic translation problems. Computer-aided translation is to help
senior translators with English and other foreign language translation skills, and use computers to assist in improving their translation level and work efficiency. The original intention of computer-aided translation is to serve more senior translators.

2.3.3. The scope of translation is different
Machine translation can only translate some simple sentences, such as daily conversations, simple business or travel vocabulary, and some other non-literary text materials. For literary works, due to its high requirements for the translator's translation level, not only simple sentence conversion, but also need to convey the author's thoughts and feelings, so this is far from being achieved by machine translation.

2.3.4. Different translation engines
Since machine translation has no human participation in the whole process, the translation result is the final result. Therefore, in order to improve the accuracy of the translation, machine translation software often has a strong vocabulary and grammar library, and the translation engine is mainly derived from the system to built a language database. The computer-assisted translation relies on the memory of translation, and forms a translation memory through the continuous accumulation of historical translation records. In subsequent translations, the memory is automatically searched and matched, and the translation with higher conformity is selected as a reference [10].

3. Comparative Experiment of Machine Translation and Computer-Aided English Translation
In order to further compare the different characteristics and advantages between machine translation and computer-assisted translation, to explore when the quality of computer-assisted translation is better than machine translation, and whether computer-assisted translation can guarantee the consistency of English translation and the original sentence. Taking a translation software as an example, detailed research and discussion on machine translation and computer-assisted translation are carried out. For machine translation, we selected Google Translate, Youdao Translator, and Baidu Translator as experimental subjects; for computer-assisted translation, we selected Wordfast, WordFisher, and iCAT Huoyun Translator as experimental subjects, and compared the terminology translation errors of these translation software in English translation. Rate and translation matching rate to analyze the advantages of computer-aided translation in English translation.

3.1. Source Language Data Collection
In this experiment, the English translation task of a translation company is selected as the language source text. This English source text mainly involves information about computers, mathematics, and artificial intelligence. There are a large number of professional terms in the text. The whole article has 5 pages, 5,012 words, 213 paragraphs, 97 grammatical sentence patterns and 254 professional nouns.

3.2. Translation of English Source Text
According to different translation software, the English source text is translated separately. Among them, Google, Youdao and Baidu in machine translation software use general full-text translation; Wordfast, WordFisher and iCAT in the computer-assisted translation software use a combination of man and machine for translation. The professional translators of the translation company perform translation with the help of computer-assisted translation software. The three different computer-aided translation software are equipped with three different professional translators, and the three translators translation experience is about 3 years. Finally, statistics and comparison of translation results.

3.3. Common Error Types in English Translation
In English translation, our common translation errors mainly include:
Vocabulary errors, including mistranslation of professional nouns, mistranslation of terms, mistranslation of word abbreviations, mistranslation of vocabulary or substitution, and untranslated vocabulary, etc.;

(2) Sentence errors, such as grammatical errors, word order errors, phrase errors, sentence voice and tense errors, etc.;

(3) Punctuation marks are wrong.

3.4. Experimental Detection Indicators
The purpose of this experiment is to compare the different characteristics and advantages of machine translation and computer-aided translation, and explore the consistency and translation errors of the two in English translation. Therefore, the experiment will compare the translation error rate of machine translation and computer-aided translation. Compare with the match rate of professional terms.

3.5. Data Processing and Analysis
The experimental data will be analyzed and processed using SPSS22.0 mathematical statistical analysis software to calculate the number and types of errors between machine translation and computer-aided translation, and compare the variance and standard deviation of the two. The calculation formula is as (1) and (2), and discuss the results using a combination of quantitative and qualitative methods.

\[
s^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}
\]

(1)

\[
\sigma = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}}
\]

(2)

4. Comparative Experiment Results of Machine Translation and Computer-Aided English Translation

4.1. Comparison of Translation Error Rate Between the Two
According to the comparative experiment in the third part, we divide them into 6 groups according to different translation software, and calculate the error rate of the translation results of different translation software. Common vocabulary errors, sentence errors and punctuation errors Divide, calculate the number of each group at these error points. The results are shown in Table 1 and Figure 1.

|                             | Google Translate | Youdao Translation | Baidu translator | Wordfast | WordFisher | iCAT Fire Cloud Translator |
|-----------------------------|------------------|--------------------|------------------|----------|------------|-----------------------------|
| Vocabulary error            | 56               | 52                 | 49               | 32       | 31         | 39                          |
| Sentence error              | 31               | 45                 | 42               | 36       | 25         | 35                          |
| Punctuation errors          | 5                | 6                  | 6                | 4        | 3          | 4                           |
| Total errors                | 92               | 103                | 97               | 72       | 59         | 78                          |
It can be seen from Table 1 and Figure 1 that the overall amount of errors in computer-aided translation software translation is less than that of machine translation software. WordFisher has the least number of errors in computer-assisted translation, with 31 vocabulary errors, 25 sentence errors, and 3 punctuation errors. The total number of errors is 59; iCAT Fire Cloud Translator has the most errors with 78 errors. Google Translate, which has the fewest errors in machine translation, also made 92 errors, including 56 vocabulary errors, 31 sentence errors, and 5 punctuation errors; Youdao translation had the most errors, with 103 errors in total. It can be seen that the translation error rate of computer-assisted translation is lower than that of machine translation, and the accuracy rate is higher.

4.2. Comparison of the Matching Rate of the Two English Professional Terms Translation

According to the content of the English source text, the text content is grouped according to the matching rate of 60-70, 70-80, 80-90, and 90-100. Each group needs to include the same amount of fragments. Since the full text has 213 paragraphs, we take 50 fragments and place them evenly in each group, and calculate and analyze the consistency of the translation of English professional terms by these 6 translation software. The results are shown in Table 2 and Figure 2.

| Matching rate | Google Translate | Youdao Translation | Baidu translator | Wordfast | WordFisher | iCAT Fire Cloud Translator |
|---------------|------------------|--------------------|------------------|----------|------------|-----------------------------|
| 60-70         | 56               | 52                 | 49               | 32       | 31         | 39                          |
| 70-80         | 59               | 58                 | 53               | 29       | 25         | 35                          |
| 80-90         | 65               | 63                 | 60               | 25       | 24         | 31                          |
| 90-100        | 78               | 70                 | 72               | 12       | 11         | 17                          |

Figure 1. Translation error rate of each translation software
It can be seen from Table 2 and Figure 2 that with the improvement of translation matching rate, the number of translation errors of English professional terminology of machine translation software is increasing, and the error rate is increasing. When the matching rate is 90-100, Google translate changes has become the software with the most translation errors, with 78 errors, followed by Baidu Translator with 72. On the contrary, the number of errors of computer-assisted translation software is decreasing with the improvement of the matching rate, and the error rate is significantly reduced. When the matching rate is 90-100, WordFisher with the least number of errors has only 11 professional term errors, the most iCAT Huoyun Translator only had 17 errors. This shows that with the increase in the amount of translation tasks, the advantages of computer-assisted translation have become more and more obvious, and its consistency with the English source text is relatively high. When faced with a small number of translation tasks, the advantages of machine translation and computer-aided translation are not obvious. Both have certain translation errors, but the error rate of computer-assisted translation is lower.

5. Conclusion
Although machine translation is far inferior to computer-assisted translation in terms of performance and effect, it’s convenient and efficient translation methods and channels have also brought people a lot of convenience. The development and application of software such as Google and Youdao solved the dilemma of ordinary people's foreign language communication difficulties, and to a large extent promoted economic and cultural exchanges between the country and the people. Computer-aided translation greatly improves the work efficiency of translators, saves translation time and energy, and improves the translation quality of English and other foreign languages. We believe that in the near future, computer translation will be further widely and deeply applied, and we will usher in a new era of computer translation.

Acknowledgments
Shandong Management University
Qihang Project Number: QH2020R13 Data-driven Teaching Research and Practice of Business English Based on Online Learning.
References

[1] Huang G, Zhang J, Zhou Y, et al. Input Method for Human Translators: A Novel Approach to Integrate Machine Translation Effectively and Imperceptibly[J]. ACM transactions on Asian language information processing, 2019, 18(1):4.1-4.22.

[2] Alfuraih R F. The undergraduate learner translator corpus: a new resource for translation studies and computational linguistics[J]. Language Resources and Evaluation, 2020, 54(3):801-830.

[3] Torregrosa D, Juan Antonio Pérez-Ortiz, Forcada M L. Comparative Human and Automatic Evaluation of Glass-Box and Black-Box Approaches to Interactive Translation Prediction[J]. Prague Bulletin of Mathematical Lingus, 2017, 108(1):97-108.

[4] Arcan M, Turchi M, Tonelli S, et al. Leveraging bilingual terminology to improve machine translation in a CAT environment[J]. Natural Language Engineering, 2017, 23(pt.5):1-26.

[5] El-Rahman S A, El-Shishtawy T A, El-Kammar R A. A Knowledge-Based Machine Translation Using AI Technique[J]. International journal of software innovation, 2018, 6(3):79-92.

[6] Sen S, Raghunathan A. Approximate Computing for Long Short Term Memory (LSTM) Neural Networks[J]. IEEE Transactions on Computer Aided Design of Integrated Circuits & Systems, 2018, 37(11):2266-2276.

[7] Duan H, Wang L, Zhang C, et al. Retrosynthesis with attention-based NMT model and chemical analysis of "wrong" predictions[J]. RSC Adv. 2020, 10(3):1371-1378.

[8] Thawabtah M A. Complementarity between Translation Memories and Computer-Aided Translation Tools: Implications for Translator Training[J]. SKASE Journal of Translation and Interpretation, 2018, 11(2):2-15.

[9] Shadiev R, Huang Y M. Facilitating cross-cultural understanding with learning activities supported by speech-to-text recognition and computer-aided translation[J]. Computers & Education, 2016, 98(C):130-141.

[10] Chantelle, Gray, van, et al. Ethico-aesthetics and the machinic repetition of literature in translation [J]. Translation and Interpreting Studies. The Journal of the American Translation and Interpreting Studies Association, 2017, 12(1):71-90.