Computer and Translation

Yanjun Yang and Yixuan Wang
Nanchang Institute of Technology, Nanchang 330099, China
Shanghai Jiao Tong University, 800 Dongchuan Road, Shanghai, China
Email: 1286351169@qq.com

Abstract. In the big data era, due to the fast pace of society and development of information technology, translation and computer technology have become increasingly related with each other. Computer translation, with its efficiency and accuracy, has aroused more and more concern from all over the world. This paper starts from the history of computer translation. Then it illustrates five forms of computer translation, including translation automation software, computer aided translation, localization tools, translation tools in internet, and other computer tools; Afterwards the paper discusses the benefits and drawbacks of computer translation compared with human translation. Although there are some shortcomings, artificial intelligence is a trend, computer translation technology will be greatly improved and bring about more convenience for human beings.

1. Introduction

In this big data era, with the development of information technology, translation and computer technology have become more and more related with each other. Translation is one of the communicative activities involving comprehension of a message in a source language and producing an equivalent message in a target language. Computers allow for fast, efficient communication among human beings in any part of the world, and help us to cut the time we spend doing research and consulting dictionaries, and even translating, in an effort to adapt our work to our new clients' needs. Therefore, computers help humans do translation more efficient and conveniently to some extent.

2. The History of Computer Translation

Using computational tools as facilitators for translation and interpreting experience a comparatively long period. The invention of the first computer brought about the idea of using computers to do translation. In 1949, Warren Weaver, an official of the Rockefeller Foundation, suggested using cryptographic techniques, statistics and universals of language to mechanize the translation process. Punch cards had already been used to perform word-for-word translation of abstracts of scientific articles; in 1951 the first research group dedicated to machine translation was set up in the USA, and in 1954 the first public demo of a machine translation system was given. The system translated 49 pre-selected sentences from Russian into English, using a vocabulary of 250 words and only six grammatical rules (Hutchins, 1992: 6). At that time, the goal of machine translation (MT) was "the creation of fully automatic high-quality translation systems producing results indistinguishable from those of human translators" (Hutchins, 1992). The basic idea for research was that translation by machines was viable as long as it was conceived as a sort of message decodification, which meant translating word for word, a strategy we now know seldom leads to success.

Such negative view also brought a bad result: In 1966 the ALPAC (Automatic Language Processing Advisory Committee) published its now famous report, commenting that the speed of machine translation was slower, and its cost was much higher than human translation, so they made a
conclusion “there were no immediate or predictable prospects of making machine translation useful.” As a result, funding for this type of project was cut, and it was recommended that useful tools to assist human translation be developed.

However, the scientists never gave up the dream to use computers in translation. They developed the system of Systran, which is currently used by the European Commission to translate hundreds of thousands of pages a year; and the Eurotra project, which during the 1980s consolidated the efforts of many research teams throughout Europe for the purpose of creating a machine translation system that could translate all languages of the European Community into each other.

By then, machine translation was removed from research laboratory and began to be applied for actual commercial use. A more effective and sophisticated translation technology has already been in use in many companies and international organizations such as Caterpillar, CompuServe, Ericsson Language Services, Rank Xerox or the European Commission translation service (Lockwood et al, 1995: 191-240)

3. Forms of Computer Translation

With the fast development of social economy, a great number of documents and materials in industrial and commercial world need to be translated. However, an individual translator cannot carry out the task of managing an entire project alone in a reasonable amount of time unless he or she works in a team; second, that this team needs to automate as many parts as possible of the process if it is to provide a quick response to the client. Therefore, there are two key factors to measure the usefulness of computer translation, namely: the growth of the market and the need for a quick response. (Nogueira, 1998)

In order to meet these two practical requirements, the researchers have developed and invented different forms of computer translation tools. The following are the main kinds of translation tools with computer technology.

3.1. Translation Automation Software

Nowadays, there are a lot of translation automation systems in the open market. However the key technology of these systems is machine translation.

Machine translation sometimes referred to by the abbreviation MT, is a sub-field of computational linguistics that investigates the use of computer software to translate text or speech from one natural language to another. At its basic level, MT performs simple substitution of words in one natural language for words in another. Using corpus techniques, more complex translations may be attempted, allowing for better handling of differences in linguistic typology, phrase recognition, and translation of idioms, as well as the isolation of anomalies.

It is estimated that more than 1,000 kinds of translation automation systems have been sold in the market. Because more and more people use them, these kinds of software have become popular and prevalent, so the prices of them have decreased rapidly. Sometimes we can even enjoy them for totally free. In Chinese market, there are nearly 20 kinds of famous translation automation programs, such as Oriental Express, Yixing, Jin Shan Kuai Yi, Tong Yi, etc. However, there is much debate over these programs. Some people claim that human translators will soon be a thing of the past, while others say that MT systems will never be able to replace humans due to their lack of imagination and intuition. The deficiencies in machine translation systems is that many translation automation software always does translation word by word, or phrase by phrase translation, and is baffled when faced with a choice of synonyms. So sometimes the translated text would be awkward and seems unnatural to target language. For example: the simple sentence “How are you?” is translated by some translation automation programs into “怎么是你?” in Chinese. That is really a big joke and will become an unforgiving mistake. Therefore, the day machines will pose a serious threat to professional translators is still a long way off.

3.2. Computer Aided Translation

Computer-aided translation, or CAT is a form of translation wherein a human translator translates texts using computer software designed to support and facilitate the translation process. It is a broad
and imprecise term covering a range of tools, from the fairly simple to the more complicated, including spell checker, grammar checkers, terminology database, bitexts, translation memory, etc. Translation memory is the key technology in computer aided translation tools. It is to store translated sentences in a special database for re-use or shared use on a network, a tool that looks for matches (words, expressions or even sentences) between the text to be translated and those stored in the database and, whenever a match is found, (Esselink: 134). A translator can store their translations in a specially designed database so that they don’t need to translate the similar content again. A translator can make full use of the information in their own translation memory, extracting the translation which matches the segments to be translated. The newly generated text can be regarded as the draft translation and then be modified by translation.

The core of Translation memory is the memory, a complex database where source text sentences are aligned side by side with the corresponding target text sentences in advance. Translation memory in particular is of significant use for improving efficiency in repetitive tasks, maintaining consistency, efficiently updating previous translations, ensuring easy terminology management and fast, efficient project management, speeding up translation and protecting codes in tables, graphs, figures, notes and images. In short, translation memories are good at improving productivity and reducing translation costs.

Currently, there are some popular computer aided translation software in the translation market, such as Trandos, Memoq, Transit, Omegat, etc, which are mainly used overseas. In China, we always use Yaxin, Snowball, etc. Take the case of Trandos for example, translation results would be automatically memorized during translation work. When doing translation next time, computer search and compare the memory database, the similar translated words would be marked as reference translation work; Besides, when translator sets up the glossary at first, all the terminologies will be identified and translated correspondently during the work, so that it can realize the unification and improve the quality and efficiency of the translation. However, there are also some disadvantages in using this software. Firstly, Winalign is not intelligent or sensitive to language. If one of the segments are not matching with the relevant one, people have to disconnect the following all segments and set up the alignments or matchings by themselves. That's the most complicated work during whole process. Besides, Workbench usually requires the format of target text should be strictly identical to the format of source text in database, so before translating the text, people have to do a lot of edition work.

3.3. Localization Tools

Localization has appeared as a new, revolutionary market for translation. Language localization is the process of translating a product into different languages or adapting a language for a specific country or region. Many languages, especially world languages with a large number of native speakers, have spread geographically and are nowadays used in many countries and regions. Take Chinese for example, there are two different character systems: simplified and traditional Chinese. Simplified characters are primarily used in Mainland China, traditional characters are used in Taiwan, Hong Kong and Macau. Thus it is important for companies to take such differences into account, for example, when creating websites for specific regional markets.

Localization tools are certain software to address the problem of localization for translators. It will adapt the translated text to its final market, for making sure legal standards are met and quality conforms to local customers' requirements, for evaluating the potential success of the product in a particular market, or for compiling terminological glossaries to adapt the product to local consumers. For example, the software of “Catalyst” can help translators to pick up the information from source text directly and bring convenience for translator to do menu translation. In addition, it also provides statistic information for translators to accurate cost budget. The similar localization software includes: Language Studio, Multilizer, Passolo Software Localizer, RC-WinTrans, VB Language Manager Pro, Visual Localize, etc.
3.4. Translation Tools in Internet
We have already talked about relevant topic in the beginning of the semester, but actually it is an essential part of computational technology in translation. There are various types of Internet translation tools, the following are the most frequent ones:

3.4.1. Search engine and encyclopedia
Search Engine and Encyclopedia are the searching tools in Internet. Generally, they store and keep a plenty of different information, such as common sense, scientific findings and data, latest news and reports, etc. Translation is a complex activity which requires accuracy in both the lexical level and semantic level, so it is important for translator to know the background information of the source text. Searching the relevant cultural background and context situation through search engine and encyclopedias can help translate more smoothly and accurately. For example: Google, Yahoo, Wikipedia, etc.

3.4.2. Online dictionaries / thesauruses
Dictionaries are basic tools for translation, and it is hard for translators to get all kinds of traditional dictionaries because most paper dictionaries are expensive and inconvenient to look up. However, the advent of online dictionaries helps translators overcome these inconveniences and make translation more efficient. Online dictionaries always provide more detailed explanation than traditional dictionaries, for example, they list all potential explanations in different scientific fields, such as biology, geography, chemistry, medical, etc..

3.4.3. Translation forum and blog
Forum and blog are the new things in Internet. They provide an interactive platform for different people to exchange their opinions and share resources. Translators can look up some translation information, such as translation skills, specific terms, and latest translation research in the forums. In addition, you can also visit some personal blogs to learn some other translators’ experiences and feelings.

3.5. Other Computer Tools
For a large translation company and a professional translator or interpreter, it is necessary for them to learn how to operate text process tools so that the efficiency of translation will be greatly improved. Such kinds of tools are optical character recognition, speech recognition, etc.

Optical character recognition, usually abbreviated to OCR, is the mechanical or electronic translation of images of handwritten, typewritten or printed text (usually captured by a scanner) into machine-editable text.

OCR is a field of research in pattern recognition, artificial intelligence and machine vision. Though academic research in the field continues, the focus on OCR has shifted to implementation of proven techniques. Optical character recognition (using optical techniques such as mirrors and lenses) and digital character recognition (using scanners and computer algorithms) were originally considered separate fields. Because very few applications survive that use true optical techniques, the OCR term has now been broadened to include digital image processing as well.

4. The Usefulness and Drawback of Computational Translational Technology

4.1. Usefulness of Computational Translation
The progress achieved in the field of information technology (IT) has combined with modern communication requirements to foster translation automation. With the help of huge and abundant database, text that contains much technical and specialized matter is actually easier for a computer to translate. From this perspective, computer translation saves both energy and time for human beings.
4.2. Drawback of Computational Translation
The current technology in computational translation is not very mature and perfect. There still exist some drawbacks.

4.2.1. Lack of openness of meaning
Firstly, the computer translation cannot have as sufficient knowledge as human beings. The understanding of the source text includes not only the understanding of lexical words or sentences, but also the context meaning or non-linguistic meaning. Machine translation always fails to work whenever it is required to translate cultural background information, because computers do not have any memory about complicated culture. In this sense, machine translation is more “decoding” rather than “understanding” the original text. Machine translation processes the words in a mechanical and static way while human beings interpret the words in a more complicated and extensive perspective.

4.2.2. Lack of creativity of expression
Secondly, although most computational translation programs have got large quantities of synonyms, they will usually select the first one in line. Often this is not the best one to use. English is especially wealthy in synonyms, and this exacerbates the problem, since the exact meanings of the synonyms may vary substantially. The computer has no way of “knowing” which shade of meaning was intended by the original writer. The frequent problem for machine translation is to translate word by word according to source text, but human translation is a complicated work combined with understanding, analyzing, judging, selecting and creating, so human translator will analyze and output according to their own proficient knowledge, cultural background and language context. However, computer has no intelligence and emotion, it is hard to do the translation accurately and completely. The lack of aesthetic sense in machine translation is particularly reflected in literary translation.

4.2.3. Lack of political nature of translation
Translation work can reflect translator’s subjectivity, that is to say, translation work can show translator’s ideology and power tendency of translator. Obviously, machine translation can do the work strictly according to the original text one by one, it has no any other emotion and bias. However, human translator can express their own attitudes towards people, country, statue, class, race and sexual distinction, etc.

5. Conclusion
The development computer and Internet technology and its universal access to information and instant communication between users have created a physical and geographical freedom for translators that was incredible in the past. Comparatively speaking, computational translation is a newly cultivated technology, we have sufficient evidence and time to wait for its bright future. At that time, translation will become a new profession. Therefore, Translators need to accept the new technologies and learn how to use them to their maximum potential as a means to increased productivity and quality improvement. The openness of meaning, creativity of expression, and the political nature of translation seem to be the stopover and problems to be solved for machine translation. Computer and translation should not be opposite against each other, but should be more complement. With the development of artificial intelligence and neural network, the future translation must be a positive interactivity between computer and human beings.

6. Acknowledgement
This paper is supported by Humanities and Social Sciences Foundation for Universities of Jiangxi Province of 2018 Annual Project—Research on Ecological Compensation Strategy of Information Loss in Conference Interpreting (Project Number: YY18205)

7. References
[1] Celia Rico: From Novelty to Ubiquity: Computers and Translation at the Close of the Industrial Age
[2] Esselink, B. (1998): A Practical Guide to Software Localization. Amsterdam/Philadelphia: John Benjamins.
[3] Hutchins, J. y H. Somers (1992): An Introduction to Machine Translation. London: Academic Press.
[4] Huashu Wang, A Probe into Searching Ability of Translators in Big Data Era, Chinese Science & Technology Translators Journal[J], 2018(11).
[5] Lockwood, R., J. Leston y L. Lachal (1995): Globalisation. Creating New Markets with Translation Technology. Ovum Reports. Ovum Limited, London.
[6] Melby, A.K. and C. T. Warner (1995): The Possibility of Language. A Discussion of the Nature of Language, with Implications for Human and Machine Translation.
[7] Nogueira, D. (1998): "The Business of Translation" in Translation Journal, vol. 2, No. 4, October 1998.
[8] Olivia Craciunescu: Machine Translation and Computer-Assisted Translation: a New Way of Translating?
[9] Wei Chen, the Effect of Machine Translation on Translator’s Subjectivity – Future Foothold of Human Translation, Foreign Languages Research[J], 2020(2).
[10] William N. Tavolga: Comments on Machine Language Translation.