Design and Implementation of Translation System of English Based on Web

Kaixian Chen

China Tourism College, Northwest Normal University, Lanzhou 730070, China

paperiset@163.com

Keywords: English for science and technology, Translation of English, Information technology, Internet-assisted translation

Abstract. The Internet technology provides great help for the translation to a certain extent, the translator can solve many difficulties in the translation practice with the help of the network information resources, including resource sharing and network translation tools, sometimes they might even find corresponding version directly and examine the correctness of translation through the machine translation, thus making the translation more efficient and accurate. But due to the limitation of computer techniques, there are some shortcomings in machine translation, it's necessary for the translator to enhance their English ability to improve the quality of translation.

Introduction

Currently, most of technical documents are written in English. When people who do not know English would like to know such documents, they have to ask translators for help. However, as such documents are increasing in a tremendous speed, it is almost impossible to translate all of them in time, provided that translators carry out the task manually. Fortunately, the emergence of computer-aided translation software changed this awkward situation. Translation memory helps translators to store their translations, relieving them from dull routine work.

Since the computer and multimedia devices combined with the translation teaching in late 1990s, the translation instruction has been gradually shifting its focus from the traditional teacher-based teaching to the interactivity between teachers and students with the help of teachers. Meanwhile the network and information technology has enabled the improvement of translation teaching. Certainly, new material, methods and curriculum designs have been introduced by the information technology. Then, what to teach and how to present the teaching contents in the computer-aided translation courses are the major concerns in the thesis.

The traditional EST translation studies have focused on the specific analysis of grammatical features, lexical features, and stylistic features, etc, to arrive at the methods and techniques of EST translation. For example, translation of long sentences, passive voice, and conversion of part of considerable theoretical results in these areas, so this speech, etc. There have been considerable theoretical results in these areas, so this paper attempts to explore from a new perspective. Combined with the author's own experience, the thesis illustrates effects of modern translation tools in EST translation.

Now, translation practice has undergone a drastic change compared with the past. There's no exaggeration to say all translation are CAT (Computer Aided Translation). Computer-Aided Instruction (CAI) has been developed fast throughout the world since it originated in USA in late 1970s [1]. With decades of development, CAI nowadays has mainly two research directions. The first direction is the development of multimedia technology with the help of advanced audio-visual technologies. The study enables the possibility of presenting vividly the teaching contents with the help of animated films, videos, sounds and images. The other direction is CAI has reformed the educational practice, in which computer plays a significant role. New teaching ideas and teaching methods are enriching the translation course and broaden those translation educators' scopes. For example, since 1995, integrating hypermedia devices, such as Adobe Premiere and Sound Edit, for translation teaching projects, have been frequently used in the translation teaching. According to the

Computer-aided translation teaching survey, many software and tools, such as Find Reader, OmagaT, are frequently used in translation teaching to improve the quality and efficiency of translation. In China, computer-aided translation instruction has been developed gradually since the first introduction of computer-aided instruction to Chinese scholars in the early 1980 [2]. It is unanimously admitted that computer-aided instruction is an efficient way of language teaching and learning to perform several interactively joint functions of audio, visual and multi-directional stimulations and to promote revolutionarily the traditional teaching mode in Chinese classroom.

Meanwhile, to facilitate the computer-aided translation instruction, some useful software are produced, such as, Interactive Encyclopedia, Encyclopedia Britannica and other different kinds of on-line dictionaries. There are some useful tools for translation teaching, especially for network-based teaching and distance teaching system. Inevitably, the more the computer-aided translation instruction develops, the more it requires the satisfactory software, useful tools and user-friendly network facilities.

The Online searching Engines and Dictionaries

Translation is a complex activity which involves knowledge of many fields, no matter what type of translation belongs to, there are two phases: comprehension and presentation. Usually the difficulties of translation are not caused by the incomprehensible words, but due to the lack of background knowledge, especially when translating unfamiliar materials, the translator must know the basic background knowledge, but the translator's personal knowledge and information are limited, so translators need to have a large number of reference materials, such as encyclopedia, dictionaries, and so on. But it's very difficult for the translators to own these traditional references due to the objective conditions. Now, the network provides a huge shared data bank for the translator, which includes encyclopedia, the dictionaries and some special resources. All of these provide many conveniences to the resolution of translation difficulties. With the network technology, the translator can find related information to get a general understanding of the related subject through search engines, online encyclopedia, electronic dictionary, online terminology and online newspapers and magazines, thus the translation quality can be improved by reducing understanding errors. As time passes, the translator will use these knowledge in the later translation practice, thus the translator's ability to understand will be expanded and enriched.

Along with the technical and social development, the new-created things make the language renewed greatly. For example, the Ministry of Education announces 171 Chinese new glossaries in "Chinese Language Living condition Report (2006)", such as "the mortgage slave", "the flower style", "worm.nimaya", "and superman scholar" and so on. But according to The Barnhart Dictionary Companion, they collect 1500-1600 new English glossaries every year. In John Ayton's 20th Century Words he collected 5000 English new words and expressions which produced in the 20th century. "It's obvious that the renewal of traditional dictionary cannot catch the speed of the language development. Moreover the practical translation often deals with the different specialty, ranging from the popular information, computer, biology to fire prevention, marine transportation and so on. Without doubt, in order to deal with so great variety of translation materials, achieving support from the Internet through the search function is the only way to get the information promptly and accurately".

With the rapid development of Internet and the swift growth of network information, the network search engine has obtained more and more favor of scholars, researchers and translator's .The Internet is a free giant corpus to the translators. "It is very convenient for translators to use search engine to search related knowledge, the idiomatic translation of words and expressions, thus enhancing the translation efficiency and quality". It should say that the inquiry way of Internet search is a basic skill to the E-time translators.

The network inquiry is time-saving and effort saving. Take Google as an example. If you want to search the Chinese Meaning of "CPPCC", Google can find 291000 items that conformed to the inquiry result in 0.16 seconds. If you want to search the English expression of UNESCO, as long as you input the appropriate key word, Google can find 2260000 items that conformed to the inquiry

result in 0.05 seconds. Obviously the Internet search engine is very efficient. But if you want to obtain the information with the traditional inquiry way, you should read a huge collection of reference books, and sometimes your work is fruitless. Thus many problems which meet in the translation can be solved easily and rapidly through the network search. Then which one is the most reliable reference in so many search results? Actually Google "Page Rank" has already solved this problem for us. According to "Page Rank", the more front the website linked with the search result place, the more popular and credible it is. The Google search results of the above two examples are both the organization's official website links (ranked in the first place), which fully proved the authority of the translation.

Many English words have established translation, for example, the translation of proper names such as the unit name, the place name, the personal name and so on belong to this category. "A random translation of a material with many unfamiliar words will lead reader to misunderstanding, but the help of Internet will improve the accuracy of translation".

The Application of Information Technology in Translation and Problems

In the presentation phase of translation, the application of Internet-based translation tools will make the translation more efficient, especially for scientific and technical translation. Specifically, the main translation tools include a machine translation system and machine translation software. The translator can get the version of the text through machine translation and translation software when he inputs the text directly into the computer, but it is poorly readable, so the translation memory technology is always used in the process of translation to improve translation efficiency and accuracy. The user can establish memory storehouse with the existing original text and the translation by using the translation memory technology. In the translation process, the translation memory system can give the reference translation through searching the same or similar translation resources in the translation memory storehouse, thus the redundant work is avoided. There are many difficult problems of the natural language processing branches in machine translation, it's hard to get breakthrough before solving some basic problems because of the ambiguous meanings in morphology analysis, the part-of-speech tagging, the syntax analysis and the semantic analysis. The main problems in the machine translation exist in the following aspects.

Polysemy. Polysemy exists in both Chinese and Western languages. In machine translation process, sometimes at first you need to determine the specific meanings of a word, as well as the correspondence in the target language, but the relationship between any two different languages is not one-to-one correspondence, because the words can derive extended meaning in its own evolution in order to meet the need of the presentation, therefore polysemy appeared. The solution to polysemy is the word sense disambiguation, also called the semantic tagger, which means giving every word in the text its semantic encoding according to its respective context. At present, polysemy typography study is still in its beginning stage; mainly include artificial intelligence methods, dictionary-based methods and corpus-based approaches.

Word Concurrency. Theoretically it refers to the syntax distribution of one word with two or even more kinds of parts of speech, these words are called the concurrent word. Therefore, although the part of speech of a word in a sentence is definite, because it belongs to a certain context, sometimes the part-of-speech tagging cannot guarantee the correctness completely. The so-called part-of-speech tagging determines grammar of each word, thus determines its part of speech. Part-of-speech tagging automatically is the automatic elimination of ambiguous meanings of word concurrency.

Ambiguous Meanings of Phrase Structure. Different adhesive relations of nouns, prepositions and so on can result different reasonable meanings, like "She showed/her baby photos" and "She showed her/baby photos", different structure analysis has different reasonable semantics, and the ambiguous meaning is caused by the different division method. Distinguishing the structure ambiguity of the machine translation is a widespread problem, because it involves a deep analysis and comprehension of natural language. At present, we cannot do machine translation on the basis of the comprehension of natural language, what we can do now is just to study it as a kind of applied technology. Using the

strategy of ambiguity to ambiguity is only an expedient way before the system has the capacity to identify structural ambiguity.

The quality of translation cannot be guaranteed. The difficulties in distinguish of semantic ambiguity and choice of semantic meaning degrade the quality of translation, which was caused mainly by the lag behind of the theory research on the development basis of machine translation. At present, the entire translation system was established mainly on the syntactic level, the limited semantic which information system used is used for the syntactic level, the target language relies solely on simple syntactic structure conversion from the source language, but this is not enough. In addition, the language rules which the system used are still based on traditional grammar, they can not reflect the complex characteristics of natural language, the quality of the translation, has not made substantial progress, a lot of problems existed 50 years ago has not resolved now. The existing translation progress relied more on the computer-resource development, but not on the translation method. Since the 1990s, the computer and network technology developed a lot, therefore the range of available technologies are further expanded, which includes full-text search technology, voice input, network information gathering and the query engine, and so on. In short, the difficulties in machine translation mainly include two difficulties. One of the difficulties is the difference between the natural language and computer language, there are always exceptions in grammar ,which makes it impossible for the computer to store; the other difficulty is the word ambiguity. In natural language, a word can have more than one meaning, a certain meaning also can be expressed by several words, so it's quite hard for the computer to select the best one in a variety of semantic expressions, but people seldom have the same problem when using natural language because they can understand these semantics easily, so it's necessary for us to revise the version after the machine translation in order to ensure the accuracy of the translation. A sound translation needs the combination of translator's English ability and computer skills with machine translation, which can be seen in the Fig. 1.
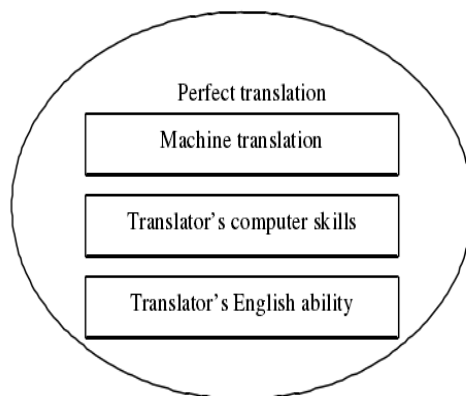

Fig.1 The constitutions of perfect translation

Conclusions

"A handy tool makes a handy man." It is not enough for us to face the challenging translation task merely by depending upon individual knowledge and information. In the Internet web-based era, the translator must be good at using Internet to assist the translation practice for the high efficiency and high quality of translation, thus" getting twice the result with half the effort". However, although the Internet technology can improve the efficiency of the translation greatly it is only a highly effective tool, the Internet tools are useless without solid basic language skills, the instruction of correct translation theories and methods, so the translator cannot ignore the training of translation basic skills when applying the Internet to enhance translation quality and efficiency.

References

[1] Zhang Jingxiang. Language polymorphism and Internet-assisted translation, China Journal of Translators for Science and Technology, February 2003, pp 32-35

[2] Wang Junli. The application of Internet sources in translation, China Journal of Translators for Science and Technology, February 2007, pp 36-40.