

# Research on the Application of English Translation Accuracy Based on Big Data Search Technology

Wenmiao Hu

GuangDong NanFang Institute of Technology, Jiangmen 529000, China

**Keywords:** Big data technology, English translation accuracy, Grammar, Word order

**Abstract:** In the era of big data, the surge in data has increased the demand for communication, which in turn has increased the demand for language-based services. The important feature of big data is that the technical factors are very active, and technological development and technological innovation occupy a leading position, which plays a vital role in the development of the translation industry. In the English translation industry, in the past, many difficult-to-quantify information has been transformed into data for storage and processing. Many complex translation projects urgently need to use big data methods for value mining. Based on this, the paper will apply the big data technology application and English translation, and hope to improve the accuracy of English translation.

## 1. Introduction

It is generally believed that the language service industry includes four major business areas: translation and localization services, language technology tool development, language teaching and training, and multilingual information consultation. The development of the language service industry is inseparable from the high-speed processing of massive information. However, in the context of economic globalization, information has grown exponentially. The amount of data generated in the last two years is almost the sum of the amount of data produced by humans in all previous eras. Knowledge growth and differentiation have far exceeded the maximum human tolerance. Scope, so in the information age, the rapid development of society must rely on information processing technology. Big data computing technology came into being, solved the technical problem that the data scale is too large, the traditional computing method cannot complete the analysis and processing in a reasonable time, big data technology and statistical language-based natural language processing technology in speech recognition, machine translation, semantics Major technological breakthroughs have been made in the technical field. In recent years, the language service and technology market has continued to grow and develop, and the annual growth of the market has increased year by year, from \$25 billion in 2009 to \$40.27 billion in 2016. This change has in turn brought the language service industry to a new era of information: the content of language services is no longer limited to interpretation and translation, but has become increasingly diversified. Big data has spawned many new types of business, and the structure of the language service market has changed a lot: although in 2016, the most important business of the language service industry is traditional translation and on-site interpretation. The share increased from 57% in 2013 to nearly 73%. However, compared with 2013, the conference services interpreter (3.32%), localization (0.51%), and game localization (0.54%), search engine optimization (0.35%) and subtitle translation (1.08%), these emerging industries have small market share, but are relatively stable [1]. At the same time, the diversification of the market has led to changes in the service model. For example, in today's cross-border e-commerce, real-time multi-language communication and translation are usually required, so the instant, dynamic and fragmented micro-language service model is also Beginning to enter the historical arena, the diversified language strategy should keep up with the pace of the market and change.

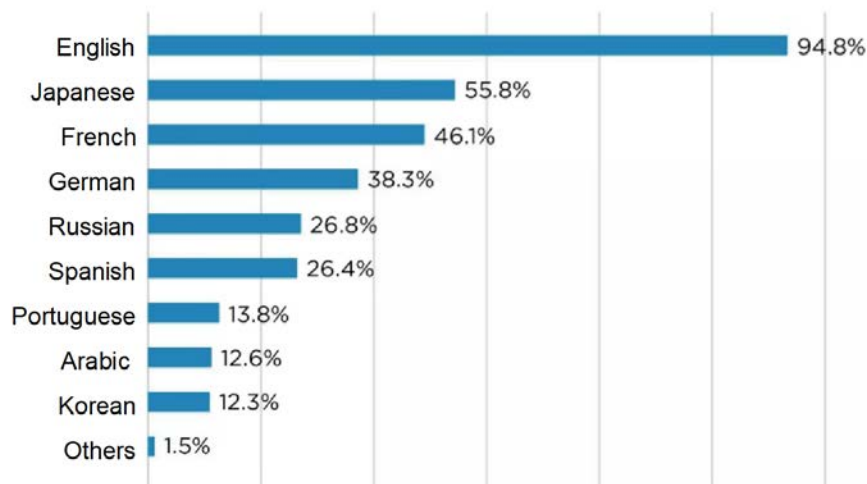


Figure 1. The proportion of different languages translated into Chinese

## 2. Translation and cultural construction in the era of big data

Is translation only a language conversion or a combination of language and culture? History has proven that translation plays an important and unique role in the process of world civilization. In China, the translation career has experienced five climaxes since the translation of Buddhist scriptures in the Eastern Han Dynasty. They are the translation of Buddhist scriptures from the Han Dynasty to the Tang and Song Dynasties, the scientific and technological translations in the late Ming and early Qing dynasties, the Western Translation of the Opium War to the May Fourth Movement, and the early days of the founding of the People's Republic of China. The translation of Eastern and Western literature before the "Cultural Revolution" and the translation from the 1970s to the present. From the spread of religion, the spread of Marxism-Leninism and world literature, and the two-way exchange of Chinese and Western cultures, it has been extended to exchanges in various fields. The whole process of translation development is the process of cultural exchange. The essence of translation is cultural communication. Translation is based on language and higher than language. As Ge Zhaoguang said, "Language is not only a language, but also a cultural communication. There is no translation of cultural common sense, as if all kinds of computer software are now making people laugh and cry." Xu Wei also proposed to define translations through cross-cultural communication activities, to understand translation and understand translation from the height of culture. Mr. Ji Xianlin believes that the desire of Chinese culture can always be youthful, and the mystery lies in translation.

The role of translation in cultural communication is indispensable. It is a constructive force, as Lefevere said: "Translation has always been the main influence force that determines the development of world culture." Language is a communicative act in a sociocultural context. As a kind of social activity, translation will also be influenced and restricted by various factors such as politics, ideology, economy and culture. When the translation is counteracted on the original text, it becomes a force to shape and influence the target language culture. At this time, translation has become a strategy for cultural development. Ninanjana also pointed out: "The role of translation is two-way, affecting the source language culture and the target language culture, and shaking the previous ideas." Venuti said that translation is essentially a cultural and political act. Through the shaping of cultural identity, translation has a lasting and far-reaching impact on society, and "will strengthen alliances, confrontation and inter-state relations in geopolitical relations. Hegemony." Therefore, translation is not only closely related to culture, but also a profound manifestation of translation politics. Especially in the context of the great strategy of the Chinese government to improve the "cultural soft power" and the export of Chinese culture, the development of translation also provides a broad platform and effective way for the construction of Chinese cultural soft power.

Therefore, the translation activity itself is rich, complex and creative. It is not only the conversion of language, but also the function of language service. It should be recognized and understood in the position and vision of multicultural communication.

### **3. Accuracy analysis of English translation in the era of big data**

Despite the difficulties of English translation, accuracy is still regarded as the top priority of English translation. Some scholars even think that accuracy is the soul of language. Accuracy should be certain, clear, and precise; requirements should be strict, objective, and normative, and no extension, reasoning, or speculation should be accompanied by a slight emotional color. Such as: "at a rate of 10 billion bits per second", "orbiting the Earth every 90 minutes", etc., these words are very clear and clear [2].

In this sentence, an orbiting solar power plant carries a distance clause that would collect ... or focused collect-tors, and the clause contains another three attributive clauses that would boil ...would generate electricity. These four attributive clauses are all led by that, and the loops are stacked and stacked. Although the sentence structure is complex, the relationship is clear and logical. On the surface, it seems that the translation seems to be lengthy and dragged down against the principle of simplicity. However, the proper addition of the translation is very accurate. In addition, the scientific use of terminology makes an accurate and standardized summary and expression of the working mode of the solar power station in the universe.

### **4. Taking the advertising industry as an example to analyze the accuracy of big data technology in English translation**

#### **4.1 Impact Analysis**

In the traditional commercial advertising operation, whether it is the production of advertisements or the placement of advertisements, there is a certain ambiguity. Often, the judgment of the media and the allocation of advertising costs are based on the audience's rough estimate. However, due to reasons such as the consumer's consumer behavior and psychological understanding, the estimation tends to be biased, so that some resources are not used properly, and the advertisement is difficult to obtain the effect expected by the advertiser. However, with the development of the Internet and the massive overlap between the Internet and the consumer's life trajectory, the life and consumption trajectory left by consumers on the Internet can be monitored and recorded. Advertising companies can use massive data to obtain massive consumer data and comprehensively analyze and predict the behavior of consumers, to achieve accurate positioning of target consumers.

Similarly, big data can also be applied to ad translation. Translation is not just a simple code-switching process, but a specific form of human behavior dominated by purpose [3]. In advertising translation, it is obviously insufficient to regard translation as a way of finding the equivalent of the source language in the target language. Multinational advertising must achieve good sales results in foreign countries. The translation of the language also needs to translate or even reproduce the source language text for the target language readers, which will cause consumers to resonate with emotions and evoke the consumer's desire to purchase the product. With the rise of big data technology, translators can get detailed information of target audiences and related analysis results through databases and consulting companies. Based on a comprehensive understanding of target audiences, they can find more acceptable audiences and promote consumers. The language pattern of the corresponding consumer behavior is generated for translation, and the purpose of advertisement translation is better realized.

#### **4.2 Accuracy and accuracy analysis**

Although the principle of targeting readers' preferences should be recognized in advertising translation, how to translate based on the preferences of target consumers is inconclusive. There are many factors that determine whether advertising translation can achieve its marketing purposes, such

as source text, consumer preferences and cultural background, advertisers' willingness, and the product itself. How can translators use these factors to form consumers? Buying a translation of desires? First, in order to enable consumers to generate the desire to purchase goods through advertising, translators need to "precisely locate" their consumers. On this basis, effectively combine the characteristics of target consumers, product features and source texts. Targeted and targeted "precise translation".

#### **4.2.1 Target positioning accuracy**

To achieve "precise translation", we must first accurately target the aesthetic taste of the target audience and the shopping tendency. If an advertisement is translated, it will be "faithful, daring, and elegant". If it does not meet the aesthetics of the target audience, it cannot be mobilized by consumers. It can only be considered as a failure case. Different types of goods often have their own specific consumer groups. For example, the consumer groups targeted by make-up products are 18 to 45-year-old fashion women, and the target groups of milk powder products are mostly infants and young children. The marketing case of selling the comb to the monk is classic but not universal. In the translation of advertisements, the translator needs to research and locate the main characteristics of the target group, and effectively combine the information to be expressed in the source text with the psychological preference of the target group, to achieve the maximum validity of the advertisement translation.

The slogan "Don't be shy, shine" of Dior's enchanting lipstick is "playing color, and dare to be outstanding". As one of the luxury brands, Dior has its own fixed consumer group – high-end people under the age of 40. These people can be divided into two main categories, social elites and young new sharps. A detailed analysis of its consumption habits, lifestyle, education level, etc. can be found that these two types of consumers generally value their own tastes, and generally agree with independent living concepts. Therefore, in the translation of such advertisements, the translator can appropriately integrate the elements of self-expression based on highlighting the elegant taste of the products. Therefore, the translator's use of the words "excellent color" and "excellent" in the translation process can not only reflect the product characteristics of this lip gloss, but also reflect the higher taste and social status of such people from the side. Adding the words "play" and "dare" before the words "beautiful" and "excellent" can also fully cater to the psychological characteristics of modern women who want to be confident and courageous.

#### **4.2.2 Quality Positioning Accuracy**

According to the definition of the quality of goods in the Introduction to Modern Commodity, the quality of goods refers to the comprehensive characteristics of the products that are suitable for their use under certain conditions of use [4]. Regardless of the form in which commercial advertisements are disseminated, what kind of content is disseminated is, in the final analysis, the marketing of the quality of goods. In order to achieve the purpose of advertising translation, it is inevitable that the translator will "recreate" the original text in the translation process [5]. But no matter how the translator "re-creates", it should be based on the precise positioning of the quality of the goods. Highlighting the distinctive features of merchandise in ad translation can help consumers fully understand the merchandise itself and quickly build a merchandise image in mind. Especially in the process of transnational marketing of goods, some products are not familiar to foreign consumers themselves. If the characteristics of goods cannot be highlighted in the marketing process, it is easy for foreign consumers to be confused about the products themselves.

When the iPhone 6 and iPhone 6s were released in 2014, Apple customized the "bigger than bigger" slogan for the two phones. The mainland translation was "greater than big". Once this translation was published, it immediately caused an uproar. Because in the English slogan, the author cleverly uses two "bigger" to highlight the two new features of the iPhone 6 and iPhone 6s, that is, compared to the previous iPhone, these two phones not only increase the size of the phone, but also The performance of the mobile phone has also become stronger. However, the translation of the mainland official website "larger than the big" is not only inconsistent with the Chinese expression

habits, but also does not reflect the characteristics of the goods contained in the original text. In contrast, the Hong Kong-Taiwan version of the translation "stops at the big" not only conforms to the Chinese expression habits, but also reflects the information that the appearance of the product becomes larger and the performance becomes stronger.

## 5. Analysis of New English Translation Techniques in the Age of Big Data

### 5.1 Crowdsourcing translation

In the Web 2.0 model, an open, self-service and popular online platform enables users to actively participate in the creation, sharing and consumption of information. With the rapid growth of people's demand for cross-cultural communication, ordinary netizens have many translation needs for foreign film, tourism, culture, news and other information, and this demand is often highly time-sensitive [6]. Therefore, some young people acted as "cultural porters". Based on the principle of voluntary participation, they used the form of "crowdsourcing and translation" to complete many translation tasks in a short period of time, forming a translation under the network technology platform design framework. Organizational model. A group of netizens with better foreign languages are keen to search and browse foreign film and television resources of interest on the Internet. At the same time, they are also willing to translate these video resources to friends or other enthusiasts. Subtitle translations initiated by volunteer translators in the early days were created in this context. "The whole activity is not for the purpose of profit, but more because of personal interest, the need for information as a carrier, and a utopia character that is ideal. Civilian volunteers at the online market took on the translation tasks that were originally performed by professional translators under the cathedral model. This new type of translation mode, which is supported by Internet technology and translated by many grassroots translators, is called the crowdsourcing translation model. Figure 2 shows the crowdsourcing translation mode.

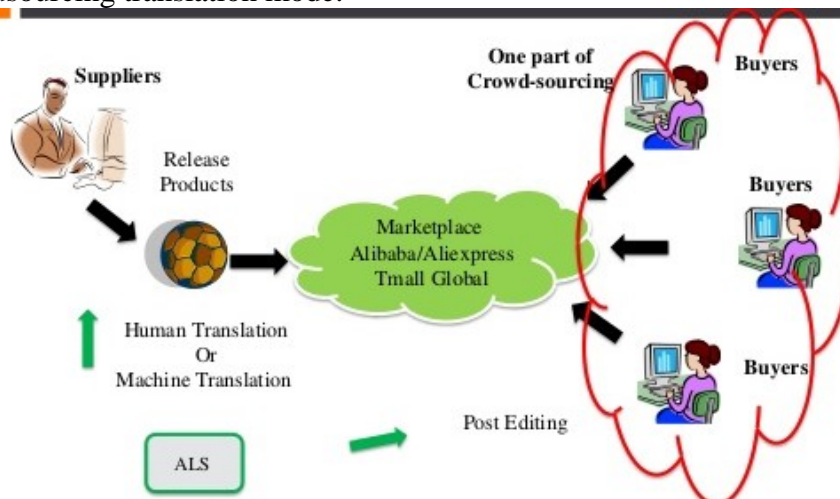


Figure 2. Crowdsourcing translation mode

### 5.2 Cloud Translation

Cloud computing is based on big data operations. "With the huge computing power of the Internet, the processing program is automatically split into countless smaller subroutines and then handed over to a large system composed of multiple servers. After searching and calculating. And analysis returns the processing results to the user." Cloud computing has the characteristics of large scale, security, scalability and on-demand services. The data processing capability has been rapidly improved, and the overall management cost has been greatly reduced. According to the SPI three-tier architecture service concept of cloud computing, some language service providers have introduced cloud computing technology into the translation industry by building a cloud translation platform. Specifically, the language service provider "stores data and programs in the cloud, and uses a number of open APIs to turn the network into a large multi-functional operating system. Information

can be exchanged between websites, and users can register users on the Internet. Forms have their own data and can be used on different websites." Figure 3 shows the K2K cloud computing translation platform structure [7].

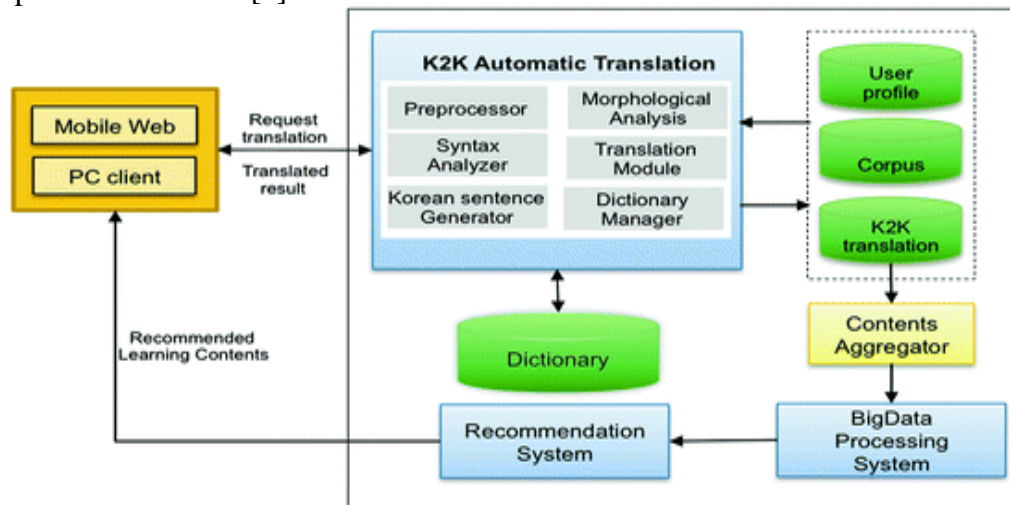


Figure 3. K2K cloud translation platform

The cloud translation platform leverages powerful network technologies and language processing technologies to integrate translation resources scattered throughout the region. Through the optimal configuration and recombination of resources, technologies and services, the efficiency of the entire translation industry has been greatly improved. This process is actually "crowdsourcing service providers turn the value created by network users into a part of their paid commercial resources." The translation platform and product customization services built through cloud technology improve management efficiency and satisfy customers' translation. The demand for "low price, high efficiency and high quality" products.

## 6. Conclusion

In the era of big data, translation technology is essentially the fusion of humanities and technology under the digital humanism. The two influence each other, interact and symbiosis. Translation technology is the construction and promotion of translation activities and translation society. In the new era, translation technology has become an inescapable fate of translation practitioners. We should embrace the development of new technologies with an open mind, fully recognize the relationship between humanity and technology of technology, and give full play to modern translation technology. The advantages of the market, according to the development of the market, advance with the times, adjust the talent training strategy and syllabus, and cultivate modern language service talents with comprehensive literacy that adapt to the development of the times and market demand.

## References

- [1] Kim, T., Hwang, M., Hwang, M., Song, S., Jeong, D., & Jung, H. Translation of technical terminologies between English and Korean based on textual big data. *Software Practice & Experience*, Vol. 8 (2015) No.45, p. 1115-1126.
- [2] Saz, O., Lin, Y., & Eskenazi, M. Measuring the impact of translation on the accuracy and fluency of vocabulary acquisition of English. *Computer Speech & Language*, Vol. 1 (2015) No.31, p. 49-64.
- [3] Pizer, S. D. Falsification testing of instrumental variables methods for comparative effectiveness research. *Health Services Research*, Vol. 2 (2016) No.51, p. 790-811.

- [4] Nielson, J. L., Guandique, C. F., Liu, A. W., Burke, D. A., Lash, A. T., & Moseanko, R., et al. Development of a database for translational spinal cord injury research. *Journal of Neurotrauma*, Vol. 21 (2014) No.31, p. 1789-1799.
- [5] Westra, B. L., Latimer, G. E., Matney, S. A., Park, J. I., Sensmeier, J., & Simpson, R. L., et al. A national action plan for sharable and comparable nursing data to support practice and translational research for transforming health care. *J Am Med Inform Assoc*, Vol. 3 (2015) No.22, p. 600-607.
- [6] Hemingway, H., Asselbergs, F. W., Danesh, J., Dobson, R., Maniadakis, N., & Maggioni, A., et al. Big data from electronic health records for early and late translational cardiovascular research: challenges and potential. *European Heart Journal*, Vol. 16 (2018) No.39, p. 1481-1495.
- [7] Lecroq, T., & Soualmia, L. F. Managing large-scale genomic datasets and translation into clinical practice. *Yearbook of Medical Informatics*, Vol. 1 (2014) No.23, p. 212-214.