

# Development and Research of Chinese-English Slang Translation Retrieval System under Internet Environment

Nuojia Xiong

School of University of South China, Hengyang 421001, China

---

**Abstract:** Slang plays an important role in cross-cultural communication. It is frequently used in spoken language and constantly updated. Under the drive of new needs, intelligent integrated retrieval system came into being. Based on information platforms such as the Internet, big data network and cloud computing, intelligent retrieval technology classifies and integrates English slang and develops an "English slang translation retrieval system", which greatly reduces the search cost of users and provides convenience for English learners and cross-cultural communicators.

**Keywords:** Slang, Retrieval System, Internet, Translation.

---

## 1. Introduction

With the deepening of globalization, cultural exchanges between China and the UK have become increasingly close. In order to further meet the needs of cultural communication, various translation software has been born one after another. In fact, there are a large number of intelligent translation software in the domestic and foreign markets, such as Baidu, Youdao, Google, Bing and so on. In terms of non-paid translation software at this stage, conventional translation can satisfy people's communication, but for slang or more colloquial translation, most of the translation software on the market have many problems, the most obvious problems are: (1) the word order is reversed, the grammar is confused, such as: Chinese tends to put time after the subject, such as "I'm going to the airport tomorrow", while English expressions tend to put time at the end of the sentence, that is, "I'm going to the airport tomorrow". Some translation software will neglect to adjust the word order in the process of Chinese translation. Translation that does not conform to Chinese expression habits and grammar rules (2) cannot be translated in a specific context; (3) The special cultural symbol can not correspond to the conversion cost, and the local cultural symbol leads to difficulty in understanding and incomprehensibility. In view of this lack of collecting a large number of localized expressions in both Chinese and English, establishing an English-Chinese bilingual corpus has become a hot issue worth studying. Meanwhile, developing an intelligent search software or small program with the assistance of existing technical means and testing it on large-scale real corpus has also become a development direction, and scholars are gradually shifting from theoretical research to practical application. With the increasingly frequent international communication and the vigorous development of the translation industry, more and more information construction has been carried out in the translation industry, the most prominent of which is the introduction of machine translation and machine-assisted translation products, and the informatization of translation process management for replacing or assisting human activities, which can improve the efficiency of the translation industry to a certain extent. At the same time, many universities and research institutions in

China are engaged in machine translation or machine-assisted translation research, and have achieved fruitful research results, but the results are not enough to produce, especially for auxiliary translation products. Due to the need for a large number of corpus resources, its software cannot be tested under large-scale real corpus, thus limiting the quality of the entire software.

## 2. Research Status and Development at Home and Abroad

The sensitivity of Chinese scholars to slang culture is reflected in a relatively early age. In the article "The Characteristics and Application of English Slang", it is mentioned that slang, as a kind of informal language, has the important characteristics of morphology and freshness[3]. Proper use of slang can make the language novel, lively, vivid, enhance the expression, make people refreshing. A Preliminary Study on the Translation of English Slang into Chinese, published in 2003, also mentioned three principles that should be observed when translating English slang into Chinese: (1) accurate understanding of the original; (2) Convey the features of the original text; (3) Convey the personality of the slang user[7]. Chinese scholars in this stage mainly study the characteristics of slang and translation skills, and have noticed the particularity of slang culture in the translation process. In the Understanding and Application of English Slang in Language and Culture, it is discussed that people in China and the United Kingdom and the United States have great differences in culture and history. To solve this problem, it is important to have a solid English foundation, but it is more important to understand the culture and customs of English-speaking countries[5]. English Slang from the Perspective of English and its Implications for oral English Teaching mentioned that slang has a high frequency of use in oral English, which has important implications for oral English teaching[2]. During this period, Chinese scholars began to focus on the connection between slang and various aspects such as culture, education, society, history and economy, instead of studying the translation of slang in isolation. Slang is closely related to the development of society and the change of history, so slang culture has long penetrated into all aspects of social life, and the research on

slang has also begun to diversify. For a long time, foreign scholars also have great enthusiasm for the study of British slang, and the phenomenon of slang culture has aroused great repercussions among scholars, especially in recent years, the publication frequency of English slang has increased. However, judging from the previous articles, foreign scholars rarely studied the translation of English slang into Chinese expressions. The development of intelligent retrieval system for translating English slang into Chinese has become an exploratory direction. The 2023 Review of Recent Query Reconstruction Methods in Information Retrieval Systems talks about trillions of pieces of data uploaded to the Internet each year. Extracting useful information using only a few keywords has become a major challenge, and the field of query rewriting (QR) focuses on efficient retrieval of information to overcome this problem, which is widely used in information retrieval (IR) and related fields such as search engines, multimedia IR, cross-language IR, recommendation systems, etc[1].

With the continuous development of economy and science and technology, in the 1950s and 1960s, centralized computer networks began to be applied to specific industries and institutions in developed countries such as Europe and the United States. In 1969, these centralized networks began to connect, and China started late, but also entered the information age in 1984. The emergence and popularization of computers have had a significant impact on society. The amount of information, the speed of information transmission, the speed of information processing and the degree of application of information are all increasing in a geometric manner. The spread of smartphones in most regions between 2009 and 2013 also provided a strong boost to the growth of the Internet and cloud computing. The connection between various aspects of research and technological means such as intelligent intelligence has become a trend and trend. In the past two years, scholars at home and abroad have devoted themselves to exploring the connection and balance between science and technology and corpus construction, as well as the optimization and upgrading of various retrieval systems. In the paper "Research on the Construction and Application of Corpus for Translation Teaching in the Information Age" in 2021, it is proposed that the corpus data can be automatically imported into the database by the program through the import of structured text, and the supported text formats include xls, xlsx, csv, xml and other file formats[4]. In addition, the system supports manual editing and input of corpus. In addition, the previous input content can be modified and deleted. On November 30, 2022, ChatGPT (Chat Generative Pre-trained Transformer), developed by Open AI in the United States, was officially released. ChatGpt, in addition to the chat function of interacting according to context, can also write emails, video scripts, copywriting, code, and papers. At the same time, the translation function has also achieved a great breakthrough, solving some traditional problems, but there is still room for improvement, such as the lack of industry data training; There is no real-time; Unable to do large data input. The First Stage Semantic Retrieval Model Based on Multiple Expressions published in 2023 explored the obvious shortcomings of lexical item-based models: (1) Based on symbolic expression, it is easy to suffer from lexical mismatch; (2) Based on the assumption of independence and ignoring word order information, the document semantic information cannot be well captured[6].

Among the problems in the development of retrieval

system for slang translation from Chinese to Chinese, the biggest problems are that retrieval is too mechanized, which leads to lexical mismatch, and translation is too violent and ignores cultural background. When exploring the optimization scheme of retrieval system, Chinese and foreign scholars have gradually combined theory with practice, guided by theoretical research, conducted experiments on large-scale real corpus, and achieved some excellent technical results. In terms of the current development of science and technology, the development and research of intelligent system is undoubtedly a hot issue, but the research of slang culture and slang translation into Chinese is relatively decreasing. If the research of intelligent retrieval system and slang translation into Chinese is combined and put into practical experiments and tests, it will be a promising research topic. However, there are still many problems in the current retrieval system, especially in the Chinese translation intelligent retrieval system related to slang, and it is still a long way to go to develop a perfect model.

### 3. Research Significance

The development and research of the intelligent retrieval system for translating English slang into Chinese has the following significance:

#### 3.1. Provides a Convenient Tool for People to Search and Learn Slang

Many language learners, researchers and lovers of slang culture will inevitably encounter some difficulties and obstacles in the process of learning slang. The intelligent retrieval system for translating English slang into Chinese can be used as a tool and reference to help learners retrieve and learn. There have been some printed dictionaries on slang learning published in the market, but they are obviously much more inconvenient than electronic systems. Many new slang words emerging with the development of society also need to be included in dictionaries, but the frequent revision and publication of paper dictionaries will cause the loss and waste of material, financial and human resources. However, the electronic system can solve this problem well. New slang words and expressions can be easily entered into the system at any time, constantly expanding and updating the original resource library. At the same time, it can also open up a sharing knowledge platform, which is convenient for different learners to break through the limitations of time and space for academic exchange, and further optimize the resource library through various feedback. The popularity of smart phones provides an important carrier for the development of the retrieval system. The electronic system relies on smart phones, tablets, computers and other electronic devices. Learners can use the retrieval system to communicate and learn on any occasion and at any time, which makes up for the inconvenience caused by paper books.

#### 3.2. Promote People's Understanding of Slang Culture and Promote Cross-cultural Communication

With the in-depth development of economic globalization, the communication between various countries and nations is deepening day by day, and cross-cultural communication has become a hot issue. As a form of expression of different cultures, slang plays an important role in cross-cultural communication. Ancient Chinese civilization was born in the

river basin, located in the inland plain, abundant water, so agriculture developed, closed strong. Western civilization was mostly born on the coast, built cities near the sea, and many good harbours, so the navigation and trade were developed. Different geographical environments have created different languages and cultures. For example, many Chinese slang words are a summary of experience in agricultural production, such as "Beginning of autumn and end of heat, white dew and autumn equinox are cutting the fields" and "there is more rain in the Fuli and more rice in the valley". How can foreign friends understand the 24 solar terms and the dog days in traditional Chinese culture in the process of cultural exchange? This cultural mutual understanding often requires art and skill. The 24 solar terms are the knowledge system formed by ancient ancestors in accordance with the farming time, observing the movement of celestial bodies, and recognizing the climate changes in the year, climate and phenology. However, the 24 solar terms may not be suitable for climatic conditions in other regions, so the transmission of this culture must be based on the soil of the country. When translated into Chinese, this slang phrase "Any port in a storm." may be more inclined to be translated into "desperation", which is to convert this slang into a language based on Chinese cultural soil, and the main purpose is to promote understanding and communication. This idiom is used in the United Kingdom, which is geographically linked to the sea. The United Kingdom is surrounded by the North Sea, the English Channel, the Irish Sea and the Atlantic Ocean, the natural shipping conditions are excellent, and the United Kingdom has a long history of navigation. From 1480 onwards, Bristol merchants began to organize fleets to search for the legendary Atlantis, the Brazilian Islands and the Antilles Islands, and tried to find more suitable new fishing grounds, and the late British navigation and trade have been developing. During her reign, Queen Elizabeth I of the Tudor Dynasty encouraged the development of shipbuilding and maritime industry, and supported the establishment of various overseas trading companies. This historical tradition is so embedded in British slang culture that words such as "sea", "waves", "harbour", "sailor" and "adventure" are often found in oral expression. However, ancient China is far from the sea, the terrain is closed, there are many plains, hills and vast land area, irrigation industry and agriculture and animal husbandry developed, people know little about "sea", so in the historical tradition of Chinese people, "road", "field", "grain", "flowers, birds and animals" and other words will be more widely accepted. This project is aimed at the establishment of slang culture and English-Chinese parallel corpus as well as the in-depth study of artificial intelligence. It adopts the method of combining theory and practice to carry out theoretical research on the slang acquisition system of machine translation for the establishment of English-Chinese bilingual parallel corpus, and applies it into practical operation by using certain modern scientific and technological means. Compared with the traditional software, the project has carried out iterative innovation in the translation of slang and some colloquial expressions to make up for the defects of the current intelligent translation system, and also added the image processing function, so that the system can not only search through keywords, but also through text pictures. These culturally relevant things can be presented through electronic systems, and the development of relevant retrieval systems will help to promote people's understanding of slang culture to some extent, thus promoting cross-cultural

communication.

### 3.3. To Provide a New Way for the Preservation and Development of Slang Culture.

Slang refers to informal colloquial expressions that circulate among people. It is usually closely related to people's life and production, and has a strong correlation with the region. However, with the popularity of standard language, many interesting slang words are on the verge of extinction. As a special cultural form, slang should be protected and promoted. The development of electronic retrieval system undoubtedly provides a new way of thinking for the preservation and development of slang culture. First, the establishment of electronic resource database can record and preserve all kinds of slang to the greatest extent; Secondly, slang culture can be innovated and developed through people's retrieval and learning. Based on these two significant advantages, the research on Chinese-English slang translation retrieval system under the Internet + environment is of great significance.

## 4. Research Process

Wechat mini program is a lightweight application development method based on wechat ecology. Using wechat mini program as the medium to realize the development of English slang translation intelligent retrieval system is a more operational means. The whole development process consists of four stages: design, coding, testing and release.

Design phase is an important part of small program development. At this stage, developers need to formulate design specifications and interaction rules of small programs in detail according to the functions and interface characteristics of small programs, and carry out interface design, function design, data structure design, interaction design and system architecture design. The interface design should follow the principle of human-computer interaction, emphasizing user experience and ease of use; The function design should consider the business requirements and technical implementation of the small program, and determine the data storage and exchange mode based on the data structure design. System architecture design mainly determines the technical architecture, hierarchy and service interface of the small program, which has an important impact on the overall performance and stability of the small program.

Coding stage is the core work of small program development, mainly using JavaScript, CSS, HTML5, python and other technologies to achieve the functions of small programs. In the coding process, the core processing mainly builds a model based on convolutional neural network to realize the visual image processing function, divides a large number of data into test sets and training sets and puts them into the model for training, and finally puts the required images into the model as input and outputs the required elements in the pictures. In secondary processing, a series of smoothing, denoising and corrosion processing are carried out on the pictures through OpenCV library in python to enhance the clarity of the pictures. At the same time, as an important step before data processing, data enhancement is carried out on the training data, so as to further improve the training effect of the model.

Developers need to strictly abide by coding specifications to ensure code quality and maintainability. In addition, in

order to ensure the security and stability of the code, developers need to conduct unit testing, integration testing and system testing to ensure the correctness and stability of the code.

The test stage is a crucial stage in the process of small program development, which mainly includes unit test, integration test and system test. Unit testing is to test a certain module or component, integration testing is to integrate each module and component testing, and finally system testing, testing the functional completeness, usability and performance of the small program. Developers need to correct the errors in the code through the test results to ensure the quality and stability of the small program.

Finally, small programs need to carry out application review, version management and release online work. Application audit is to ensure the quality and security of the small program, version management.

## 5. Summary and Prospect

In the future, the development of Chinese-English slang translation retrieval system will have more diversified module functions. It not only has a single retrieval and collection function, but also has the characteristics of slang culture preservation and development function, and more functions related to daily life can also be integrated into it, making it a comprehensive intelligent system that combines retrieval, collection, cultural development, academic exchange and convenience of life. The future trend of the development of intelligent retrieval system for English slang is roughly reflected in the optimization of the functions of five sections: The first aspect is the application of semantic analysis technology. With the continuous development of artificial intelligence, semantic analysis technology will be more widely used, which will enable the intelligent retrieval system to understand the user's query intention more accurately and give more accurate translation results. The second aspect is the application of machine learning technology. Machine learning technology can gradually improve the translation accuracy and speed of the English slang translation intelligent retrieval system through the learning and analysis of a large number of data, and can continuously optimize the performance of the system. The third aspect is multi-language support. With the acceleration of the process of globalization, users' demand for multi-language translation is getting higher and higher. Therefore, the future development direction of the intelligent retrieval system may include more languages to meet the needs of users; The fourth aspect is the optimization of user experience, in addition to translation accuracy, user experience is also an important direction for the future development of English slang intelligent retrieval system. The system needs to be more intelligent and friendly, so that users feel the convenience and speed of use; The fifth aspect is the improvement of human-computer interaction: the intelligent retrieval system is also expected to enhance the user experience through a more humanized interaction in the future. For example, technologies such as voice interaction and image recognition can be used to make it more convenient for users to conduct queries and translation operations.

In addition to the translation function, the following other functions may be derived from the future English slang translation intelligent retrieval system. Such as learning recommendation function, the system can recommend

suitable English slang materials and related learning resources for users according to their historical query records and learning habits. Secondly, social functions may be derived. The system can provide users with a social communication platform, so that users can exchange learning experiences with other learners, and share learning resources and experiences. Then there may be intelligent memory function, the system can provide users with intelligent memory function, through intelligent memory cards, flash cards, memorizing words and other tools, to help users quickly and efficiently memorize English slang and related knowledge points. There is also a great possibility of real-time message function. The system can provide users with real-time message function, so that users can consult professional English slang translators and solve doubts anytime and anywhere. Finally, the development of intelligent assessment function, the system can carry out comprehensive and accurate assessment and feedback for users through intelligent examinations, tests, exercises and other links. The introduction of these functions can make the intelligent retrieval system better meet the needs of users, improve the user's experience, and also help improve the user's learning effect and learning outcomes.

## Acknowledgments

Time does not live, the season is like flow, unconsciously this project has been nearing the end. In this process, I would like to thank the team members for their efforts and Teacher Lu Jia for her careful guidance, so that this project can progress smoothly. At the same time, I also want to thank the University of South China for providing a broad development platform for the project, so that this project can be carried out with the support of relevant resources.

A few words, difficult to express my heart. In short, thanks to all the teachers, team members, classmates, friends, and everyone who has provided suggestions and help to improve our project.

## References

- [1] Ashutosh Dixit,Vishal Gupta. Recent Query Reformulation Approaches for Information Retrieval System - A Survey[J]. Recent Advances in Computer Science and Communications, 2023,16(1).
- [2] Handong Li. English slang and its implications for oral English teaching from a pragmatic perspective. Journal of Bengbu College, 2014, (1):117-120.
- [3] Mei Li. Characteristics of English slang and its application [J]. Modern Information,2002,(05) : 171-176.
- [4] Ruoshuang Yin. Research on the construction and application of Corpus in Translation Teaching in the Information Age [J]. Journal of Liaoning Normal University (Social Sciences Edition),2021,(02):67-69.
- [5] Xudong Wang. The understanding and use of English slang in language culture [J]. Journal of Jixi University, 2011, (11):85-86.
- [6] Yinqiong Cai, Yixing Fan, Jiafeng Guo, et al. Multi-representation based first-stage semantic retrieval model [J]. Computer Engineering and Applications, 2023, vol. 59 (4):139-146.
- [7] Zhiqin Tang. A Preliminary study on the translation of English slang into Chinese [J]. Teaching Research,2003,(03):233-235.