

## Sound Japanese corpus to help Jiangmen's cultural and tourism industry develop in high-quality

Xilin Tan <sup>a</sup>, Jieying Zhou <sup>b</sup>, Huijun Wu <sup>c</sup>

School of Foreign Languages, Wuyi University, Jiangmen 529020, China;

<sup>a</sup>tan23839131@163.com, <sup>b</sup>3458208643@qq.com, <sup>c</sup>1617724312@qq.com

### Abstract

Nowadays, epidemic prevention and control are fully liberalized, the domestic economic situation is good, and the tourism fever climbs. The hit of the TV series "Rampage" made Jiangmen tourism become a popular tourist city for a while. For this reason, it is urgent to improve and supplement the translated foreign propaganda materials of Jiangmen attractions promptly, and at the same time increase its language variety, to help Chinese and foreign cultural exchanges, and then enhance the international influence of Jiangmen attractions. By analyzing the current situation of Japanese corpus translation of domestic attractions, combining with the existing Japanese corpus translation of foreign attractions that can be learned from, using the knowledge learned to try to supplement the Japanese translation materials of Jiangmen attractions, and providing reference opinions for the key issues raised.

### Keywords

Japanese Corpus, attractions in Japanese, cultural and tourism industry, cultural exchange.

### 1. Status of the Japanese Corpus of Domestic Attractions

According to research, the lack of small language outreach materials will have a certain negative impact on the dissemination of culture, so in recent years, the construction of the corpus of each small language has been paid more and more attention.

Taking Changsha's red tourism resources outreach website as an example, the proposal of the "Belt and Road" initiative has brought more development opportunities for Changsha's tourism industry. It is understood that Changsha has successfully organized red tourism cooperation and exchange activities, and is also actively improving the construction of independent foreign language websites and webpages in Japanese and other languages for Changsha's attractions, which has made a significant contribution to improving the international recognition of Changsha's red cultural resources.[1] There is also a Japanese corpus constructed based on Anhui culture, which is reconstructed based on the contents of professional teaching materials, students' interests, tourism hotspots and professional needs in Anhui, making the content of Japanese language teaching in tourism richer and more systematic, which not only provides a source of corpus for the study of applied Japanese, but also provides a reference for some of the attractions that have been affected by the lack of foreign propaganda materials and the impact on the cultural popularity and acceptance.[2]

However, at the same time, there are also several kou to be solved, specifically reflected in the following aspects.

#### 1.1. Lack of Japanese translation materials for domestic attractions

In the notice boards and website promotion of many domestic attractions, there are few Japanese translations of attraction entries and one of the important reasons is that there is a

lack of Japanese translations of Jiangmen's attractions in domestic academic circles. The lack of translated information prevents international tourists from gaining a deeper understanding of the attractions' natural landscapes, folk cultures, etc. Meanwhile, some of the attractions that lack publicity materials also suffer from the lack of translated information, which affects the efficiency and accuracy of the translation of attractions' entries, and is not conducive to the dissemination of attractions' cultures to the outside world.

### **1.2. Lack of accuracy and rigor in translating Japanese language materials for attractions**

Due to the lack of Japanese translation materials for domestic attractions, many attractions choose to use machine translation to make up for the lack of publicity materials. insufficient. Machine translation can complete a large amount of work in a short period of time at a small cost, but there are also problems of improper use of words and hard direct translation, which may cause misunderstanding among foreign tourists and is not conducive to the export of China's excellent culture.

### **1.3. The Problem of Relatively Lagging Research on the Construction of Japanese Language Corpus in Sightseeing Spots**

Domestic research on the Japanese corpus of attractions started relatively late compared to Europe, America, Japan, and so on, and compared to the flourishing development of the English corpus, the research and construction of the Japanese corpus has not been paid attention to by universities and relevant research institutions, resulting in the construction of corpus has a relative lag in the research, and has not formed a systematic and normative text and other problems.

## **2. Foreign data collection on corpus**

Foreign research on corpora started earlier and more maturely than that in China. The application of Japanese corpus in a broad sense can be traced back to the surveys sponsored by the National Diet Research Institute of Japan in the 1960s, but the corpus of these surveys was basically only used for specific research and did not have generality. After entering the 1990s, some corpora serving language research appeared, and the common ones are as follows: EDR corpus, Kyoto University corpus, Spoken Japanese corpus, Solar corpus, Japanese Classical Literature Article database, RWC text database, IPA corpus, ATR dialogues database, all of which are closed corpora, although they are more generalized and of a certain scale than the survey corpora of the 1960s. Although the beginning of Japanese corpus research is not very late compared to Europe and the United States, the existing corpora and similar textual materials lack scale, balance, and representativeness, and cannot meet the needs of modern language research. It was in recognition of this problem that the National Institute of Languages launched a huge corpus construction project called KOTONOHA, which was improved in terms of planning, corpus collection, and post-processing compared with the previous corpus construction. [3]In addition, it is necessary to have an in-depth understanding of European and American corpora. Because modern corpus research originated in Europe and the United States, European and American corpora, especially English corpora, have the longest history, the richest variety, and the most mature design. For example, in terms of corpus classification, European and American corpora have a large number of classifications and a complete range of categories, which not only generally have clear design concepts, but also ensure the balance and representativeness of the collected corpus through the method of systematic sampling.

### 3. Supplementary Japanese Corpus Translation for Attractions

In the early stages of the research, the project team encountered several difficulties, such as unclear categorization of attractions, which led to confusion in conveying the information and hindered the project to a certain extent. In this regard, the project team plans to initially classify the collected attraction entries by using the App for making mind maps, then use an Excel table to separate the Chinese, Japanese kana, and kanji of the entries for easy searching, then sort out six very representative entries to make public tweets, and finally apply the knowledge spectrum graph to embed the internal association of the words to show the results of the project, and at the same time, sort out the rules and translation methods of the attractions into Japanese, and propose improvements accordingly. At the same time, the current rules and methods of Japanese translation are sorted out, and corresponding improvement proposals are put forward.

#### 3.1. Crawling existing Jiangmen local corpus on foreign websites

Since the number of existing Japanese translations of attractions and their translation materials in domestic academic circles is relatively small, it is necessary to collect the local corpus of Jiangmen with the help of Japanese websites, such as NHK, TBS, KOTONOHA Modern Japanese Balanced Written Corpus, Chunagon, Syonagon and the National Diet Library, and so on. Collecting Jiangmen's local language corpus through Japanese websites can effectively understand the latest research results, means and methods of foreign countries in this field, so as to provide a more scientific and reliable material base for the translation of Japanese in Jiangmen's attractions.

#### 3.2. Crawling the existing Japanese corpus of attractions on the Japanese version of the People's Daily Online website

The Japanese version of People's Daily covers the latest news reports in politics, economy, society, science, and other fields, which is not only a tool for Japanese learners to advance their studies, but also an important channel for the Japanese people to learn about China's current affairs. In the Japanese version of the People's Daily, there are many news reports about places of interest, which can effectively reduce the possibility of choosing inappropriate words in the process of translation by using it to gain a deeper understanding of the translation method of places of interest.

For example, there are two ways of translating the word "former residence" in the Japanese translation of People's Daily, one is "kyuukyo", which is translated as "yuumeijin no kyuuko" when referring to the former residence of a famous person, and the other is "kokyo", which is translated as "Rojin Kokyo" when translating the former residence of Lu Xun. In the Chinese context, "gu" has the meaning of death, and "Guju" is more likely to be used to describe the house where a dead person lived; for the house where a living person lived, some people use "Jiuju" to describe it. for those who are still alive, some people use the term "old residence" more often. [4]Therefore, when translating the former residences of some celebrities who have passed away as attractions, the term "former residence" is preferred, for example, "Hyoushi Kokyou".

#### 3.3. Attempted translation of Japanese words for sightseeing spots in conjunction with professional Japanese translation experience

After two years of Japanese language study, we have accumulated a certain amount of Japanese grammar, and have a preliminary knowledge and understanding of the principles and methods of translating attractions. For example, "Kaiping 3A Internet-Renowned Toilet" should be written as "Kaihei 3A Toire" in Japanese translation, and try to use authentic Japanese expressions in the process of translation.

In the process of project implementation, the project team organized the words and phrases of the attractions collected from the network and fieldwork into an Excel table and tried to translate them. The project team focused on the collection of words related to domestic tourism outreach, and in order to find the relevant words more efficiently and accurately, the project team carried out a preliminary classification, which was subdivided into six categories: food, sightseeing spots, natural landscapes, folklore and culture, recreation and leisure, as well as accommodation and transport. At the same time, taking into account the existence of two ways of pronunciation in Japanese, in order to be more in line with the Japanese people's language habits, the project team in the process of translation borrowed some of the Japanese local vocabulary.[5] For example, Japan's "Misato" is from the merger of the village of Donghe, Yancheng and Zaodaotian Village, and the name of SanXiang Ancient Town in China is similar to this, so "SanXiang Ancient Town" was translated as "Misato Kochin". In addition, when the project team encountered words like Caishen Temple, we phonetically transliterated Caishen as Zaishin, and added a note "Cyuugoku no Ebesu", so that it would not lead to misunderstandings while retaining the Chinese characteristics. The Chinese characters were chosen by the project team. In terms of the choice of Chinese characters, the project team referred to Japanese word usage and translated "Long" as "Tatsu" instead of "Ryu" in the translation of local names. However, since Japanese people tend to use "Ryu" in their names, the character "Ryu" was used in the translation of Longxizi commemorative museum. It is hoped that this will help the translators who translate Japanese for the corpus of Jiangmen attractions in the future to reduce repetitive translations, improve translation efficiency, and reduce translation errors to provide certain assistance, so as to ensure that the Japanese translation of Jiangmen attractions is accurate and rigorous.

### **3.4. Production of public tweets, step-by-step addition of terminology entries, definitions, citations, etc.**

In today's social media age, creating public tweets is one of the effective ways to engage users to view and follow expertise. Terminology entries can help readers better understand what the tweets are conveying and improve the professionalism of the tweets. The explanation section can explain complex concepts clearly so that readers can understand them easily. At the same time, by citing actual cases or examples to illustrate the specific application scenarios of a term, the credibility and persuasiveness of the tweets can be enhanced to attract more readers' attention.

The public number is based on the classification of terminology collection form, planning the tweet theme and creating related content, and tweeting in the form of graphic correspondence, with interesting styles and processing elements such as SVG, so that readers can understand the content of the tweet at a glance without losing the aesthetics.

The gradual addition of terminology entries, explanations, citations, etc. not only improves the quality of tweets but also provides readers with more comprehensive information and knowledge. Clearly categorized entries can also help readers to search for content quickly and provide convenient services.

When conditions permit, the public number will organize the knowledge involved in the content of this project into a series of compilations in the future, which will be convenient for Japanese learners to understand and read, and may broaden the relevant knowledge accordingly as the creator's Japanese literacy improves, bringing readers more knowledge points.

### **3.5. Applying knowledge graphs to embed internal associations for words**

In the research process, the project team gained an in-depth understanding of the knowledge graph, systematically sorted out the translated attractions, and drew the collected Japanese

corpus of Jiangmen attractions into a visual graph for mining and analysis. Based on the classification of the Japanese translated corpus of Jiangmen attractions in the Excel table made by the project team above, the knowledge graph is used to show the internal structure and relationship of each classification in the form of nodes and the logical relationship between them is established through the connection of lines, which is convenient for organizing and analyzing the collected Japanese words of attractions and at the same time making them more systematic and standardized, which not only helps to memorize the vocabulary in a deeper way but also facilitates the convenient and efficient searching and management of the collected Japanese words of attractions. It not only helps to memorize the vocabulary more deeply but also makes it easy and efficient to search and manage the collected Japanese words of attractions.

### **3.6. Collection of current Japanese translation principles and methods for attractions**

In order to translate the attractions in Jiangmen City into Japanese, the project team need to understand the rules of translating attractions, and then classify the attractions according to the collected information and data. For nouns with Jiangmen characteristics, such as "Chenpi" and "Pukui", they can be translated as "chinpi" and "syuro no ki" according to the translation guidelines and the characteristics of the products. Common nouns are directly translated into Japanese. For example, "mountain" and "lake" can be directly translated into Japanese as "yama" and "mizuumi". The second is proper nouns, such as the names of attractions, roads, and restaurants. By using Japanese kanji directly for proper nouns such as names of attractions, roads, and restaurants, the meaning of the original noun can be conveyed to Japanese tourists concisely and accurately. For example "gushan", is the Chinese character "gu", the Japanese kanji also has the character "ko", and has the same meaning as the Chinese character, it can be directly translated as "Kozan". Thirdly, the translation of the names of facilities within a region should be unified. When there are multiple methods of translating the names of public facilities such as stations, attractions, and bicycle rental points within the same area, attention should be paid to uniformity. For example, a rental bicycle can be translated as "rentaru jitensya", "renta saikuru" and so on.

## **4. Attractions Outreach Development Strategy**

### **4.1. Cultivating talents for translating outreach materials for attractions**

At present, there are a lot of problems to be solved in the translation of Japanese materials for attractions in China, not only because of the lack of Japanese translation materials for Jiangmen attractions in the academic session but also because the research and construction of Japanese corpus of attractions have not been paid attention by universities and relevant research institutions and there is a relative lack of professionals engaging in the translation of attractions publicity materials or the construction of Japanese corpus of attractions.

In order to improve the quality of attraction publicity materials and the professionalism of translation practitioners, the management department can set up a mechanism for monitoring and auditing the translation of attraction publicity materials and conduct regular spot checks to ensure the accuracy and rigor of the translated materials of attractions publicity materials at the initial stage. [6] At the same time, lectures and activities on attraction publicity translation should be carried out to encourage more talents with relevant professional qualities to devote themselves to attraction publicity work; while professionals engaged in attraction publicity translation work should also show their professional ability and serious attitude, not only need to understand the history and culture behind the attractions, but also need to have an in-depth understanding of the country of the translated language to avoid misunderstandings due to the

use of inappropriate words. Misunderstandings. At the same time, foreign language majors in universities can update their teaching materials according to local tourism hotspots, students' interests, and professional levels, and encourage teachers and students with professional abilities and interests in attraction publicity translation to try to participate in attraction publicity work, so as to provide talents for attraction publicity translation work.

#### **4.2. Diversified approach to attraction outreach**

According to the survey, at present, most of the domestic attractions outreach has the problem of a single way of publicity. [7]With the rapid development of the Internet, many attractions have chosen to launch English websites to attract foreign tourists. However, with the development of information technology and the introduction of major social apps, the attraction outreach mainly on English websites seems to be a bit monotonous.

Attraction outreach is not only for English-speaking countries but also to make China's excellent traditional culture to the international community. To this end, online need to set up more small language attractions outreach websites to attract more attention of tourists from small language countries, so that they have a preliminary understanding of China's excellent traditional culture; at the same time, can be in Weibo, Red, Tiktok and other foreign user base of the APP attractions outreach, to provide more channels for international tourists to understand China's attractions and culture. Offline, we can increase the frequency of attraction outreach lectures, exhibitions, and activities related to attraction outreach translation. Combined with the online and offline promotion and dissemination at the same time, but also need to strengthen cultural exchanges with overseas countries, and increase international marketing efforts, for China's excellent traditional culture export to provide effective help.

### **5. Conclusion**

To sum up, to better realize the export of China's excellent traditional culture to foreign countries, firstly, we should further improve the domestic tourism outreach materials, such as including more attractions corpus, improving the quality of the translation of tourism outreach materials into Chinese and foreign languages, which is also helpful for improving the construction of domestic tourism outreach websites; secondly, the construction of the Japanese corpus of domestic literature and traveling, and the construction of sound Japanese corpus of domestic literature and traveling not only helps academic research, but also facilitates foreign tourists traveling to China to access attraction-related materials, and reduces misunderstandings caused by improper language expressions. helps academic research, but also facilitates foreign tourists traveling to China to check the relevant information of the attractions, reduces the misunderstanding brought by improper language expression, and allows more foreign tourists to have further knowledge and understanding of our country's excellent traditional culture in the process of traveling, to realize the cross-cultural dissemination of our country's excellent traditional culture; lastly, the translation of Japanese and other foreign languages on the publicity websites of some attractions in China appears to be improperly used or culturally deficient. [8]Finally, some domestic attractions in the foreign translation of Japanese and other foreign languages on the publicity websites of the phenomenon of improper use or cultural loss, essentially because of the lack of senior Japanese translators in the tourism industry, therefore, the teaching of Japanese language in colleges and universities can be introduced into the Japanese corpus, students through the retrieval of Japanese vocabulary, exposure to a large number of accurate, authentic Japanese expression, can greatly enhance the students' Japanese language level, for the Japanese language students to be able to better cross-cultural communication to pave the way for the future.

## Acknowledgements

Natural Science Foundation.

## References

- [1] ZHANG Yi, ZHANG Shuang. "Study on the Status Quo and Countermeasures of Japanese Translation of Outreach Websites under the Belt and Road Initiative--Taking Changsha Red Tourism Resources Outreach Website as an Example[J]. Journal of Hubei Open Vocational College,2020,33(14):184-185.
- [2] Wang Xiaozhen. A study on the construction of Japanese corpus based on the translation of Anhui tourism outreach materials[J]. Contemporary Tourism,2020,18(33):68-69.
- [3] Mao Wenwei. An overview of the current status of Japanese corpus construction[J]. Japanese Language Learning and Research,2009,(06):42-47.
- [4] Gan Zhengqi. Are "former residence" and "old residence" the same thing[J]? Talent resource development,2019,(09):54.
- [5] Wang Hui. A study on the expressive features of Japanese translation of tourism texts--an analysis based on an analog corpus[J]. Foreign Language e-Learning,2020(05):31-36+5.
- [6] Liu Juanjuan. Research on Translation Problems and Enhancement Strategies of Tourism and Cultural Outreach[J]. China Nationalities Expo,2022,(18): 92-95.
- [7] Chu Tianxia. Exploration of the Status Quo and Optimisation Path of Red Tourism Outreach Translation in Zhuhai[J]. Tourism Overview,2022,(13): 121-123+127.
- [8] Wang Hui. Study on Expression Characteristics of Japanese Translation of Tourism Texts - Analysis Based on Analogue Corpus[J]. Foreign Language e-Learning,2020(05):31-36+5.