



# A Corpus-Based Study on Liu Yukun's Translation Style as Reflected in *Folding Beijing*

Yuting Chen<sup>1</sup>, Wei Gong \*

<sup>1</sup>*School of Foreign Languages, Wuhan Business University, Wuhan, P.R. China*  
1959577228@qq.com

\*Corresponding author: 851100874@qq.com

## Abstract:

*Folding Beijing*, written by Hao Jingfang and later translated by the famed science fiction writer Liu Yukun (also known as Ken Liu), has been a hit overseas, promoting the spread of Chinese culture and winning the Hugo Award for Best Novelette in 2016. This paper will take Liu Yukun's translation as an example, using quantitative and qualitative methods to analyze Liu Yukun's translation style. The corpus technology, the software Antconc, is used to construct the self-built corpus of *Folding Beijing* translated by Liu Yukun. And then the data analysis is expounded from the aspects of standard type-token ration (STTR), lexical density and word frequency, average sentence length and passive voice sentences, and narrative tenses. Based on the big data and corpus technology, it is found that the translation of *Folding Beijing* is basically faithful to the original text, and at the same time has its own unique style, which is close to the habits of the target language.

**Keywords:** *Corpus Technology, Quantitative Analysis, Sci-Fi Translation, Folding Beijing*

## 1 INTRODUCTION

In the wake of the popularization and application of computers, the corpus has exerted more and more influence on the field of language research, and it has also gradually become the main tool for people to study translation style, which approaches to analyzing translation style shift from qualitative to quantitative. Through self-built corpus and then analysis based on the data, the research conclusions of translation style are more objective, authentic and reliable. Based on the advancement of science and technology, science fiction which was originated from the western world and carried by novels, has been gradually well-known by people at home and abroad. Since the 21st century, there has been a boom in translating Chinese science fiction into English. However, before this phase, it was mostly the translation of foreign works. Two key events spurred this boom: Liu Cixin's *The Three-Body Problem* received the Hugo Award for Best Novel in 2015, and Hao Jingfang's *Folding Beijing* received the Hugo Award for Best Novelette in 2016. What is worth mentioning is that Liu Yukun played an indispensable and important role as the translator of these two award-winning novels, with unique translation styles. [3]

To date, translation style does not arouse enough scholars' attention. Scholars pay greater heed to the author's style than the translator's. Translation was used to be considered as a derivative rather than creative activity [1]. So in the past, most people thought that the translator could not have his or her own style. Nevertheless, this kind of thinking has been changed gently. At present, the researches on translator's styles conducted by domestic scholars are not a minority, but most of them pay attention to the comparative study of two or more translators' styles based on the corpus. Besides, most foreign research on this topic adopts a qualitative approach, without taking corpus as the main source of analysis. In other words, these scholars still use a qualitative approach, mainly based on previous translation theories to analyze the translator's style. All in all, there is a large practice blank in relevant research based on corpus both at home and abroad.

This paper will take *Folding Beijing* which is written by Hao Jingfang and translated by Liu Yukun as the main research subject. Through the use of the software Antconc to build a corpus, accurate data will be concluded. And then by combining quantitative and qualitative methods to analyze the translation version, characteristics of the word and sentence selection, as well as text during the translation of Liu Yukun, will be found.

Finally, Liu's translation style could be concluded with the aid of big data and corpus technology.

## 2 LITERATURE REVIEW

Advances in computer technology since the turn of the millennium have encouraged an unprecedented surge of corpus building around the world. In China, many people began to use the self-built corpus for analyzing the translator's style. Meanwhile, Chinese science fiction *The Three-Body Problem* and *Folding Beijing* have won Hugo Awards successively, which also makes some scholars' corpus-based research focus shift to sci-fi translator's style.

### 2.1 Previous studies on translation of *Folding Beijing*

Written by Hao Jingfang, *Folding Beijing* tells the story of the protagonist, Lao Dao, who works as a cleaner in the third space. In order to earn money for his adopted daughter to go to school, he illegally ventured across the folding space and helped the people in the second Space to deliver messages to the first Space. However, he ran into problems in the first Space and was eventually able to return to the third Space with the assistance of a resident who were from the third Space but struggled in the first Space. *Folding Beijing* was then translated by the Chinese American translator and writer Liu Yukun (also known as Ken Liu) and won the 74th Hugo Award for Best Novelette. Since then, this book has aroused great notice globally. While most of the scholars in this field focus on the author and Chinese Science fiction novels, only a small number of scholars pay heed to the translation of *Folding Beijing* and the translator.

Liu Jinyu and Hu Jiajia (2018) claimed that Liu Yukun's translation is faithful to the original text, and retains the cultural words with Chinese characteristics as much as possible, making Chinese culture better known to the world. [7]

Besides, some papers also pay attention to the subjectivity of the translator, such as *On the Translator's Subjectivity in Ken Liu's Translation of Folding Beijing* by Chen Yinlong (2017) [2]. He summarized the subjective factors which are conducive to Liu Yukun's perfect translation.

To date, this topic, though very important to the research of Chinese Sci-fi translation, has drawn very few scholars' attention due to the shorter pages of this novel and there has been much less discussion from western countries. As a result, the studies of translation style are more difficult to find.

### 2.2 Previous studies on corpus-based translation style

Corpus technology has exerted more and more influence on the field of language research, gradually having become the main tool to study translator's style, which makes the study of translator's style shift from qualitative to quantitative. A number of academics conducted important study in this area.

Hu Kaibao, and Kyung Hye Kim (2020), concluded that the translator's style can be deduced from lexical idiosyncrasy and syntactical sequence based on the bilingual corpus. [5]

Meanwhile, another one is *Style in Translation: A Corpus-Based Perspective* by Huang Libo. "Baker proposed the notion of translator style which refers to a translator's particular way of translation" (quoted in Huang 2015) [4]. According to this, he also summarized some indicators of the translator's style, including discourse presentation translation, readability, direct and inverse translation and so on.

Liu Kanglong and Muhammad Afzaal conducted a case study. Having adopted a corpus-driven approach, they used lexical bundles as a feature that can reflect the difference between two translator's styles, and came to a conclusion that the richness of lexical bundles can be influenced by many factors, especially the translator's language background, the translation skopos. [8]

Among previous studies on corpus-based translation style, it is obvious that the main research abroad focuses on the theoretical level and few on case studies. In China, indeed, a host of scholars use this new tool and theory to analyze the translator's styles. What is worth mentioning is that the majority of them pay attention to the comparative study of two different translators of the same original work. So there are few studies about the translator's style of a single translation.

Based on the shortcomings of the above literature, this paper is aimed at analyzing Liu Yukun's translation style as reflected in *Folding Beijing* based on the corpus. By creatively combining both quantitative and qualitative methods, this research will take statistical data as the first source and analyze the translation style of Liu Yukun.

## 3 CORPUS CONSTRUCTION

For the further development of the study, constructing a small self-built corpus is required. And target corpora of this study are the original text of *Folding Beijing* and its translation version by Liu Yukun. The number of types and tokens of the translation are 3052 and 16100 respectively. As for the reference corpus, Ame06\_FicSic in Antconc corpus software is selected as the reference corpus for this study. The number of types and tokens of this corpus is 3039 and 12311 respectively, which is close

to the number of Liu Yukun's translation. In this study's tables, the corpus of the original text and the translation will be abbreviated as Hao and Liu respectively, while the Antconc corpus will be abbreviated as Ant. These abbreviations are uniform throughout the text. The process of corpus construction is quiet cumbersome and energy-consuming, including the three main steps shown in the following chart:

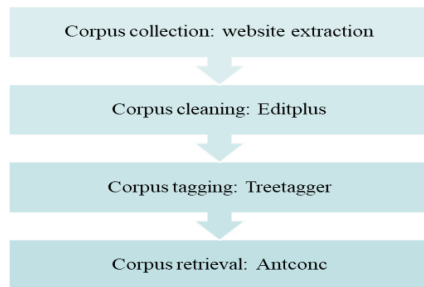


Figure1: Process of corpus construction

### 3.1 Corpus collection and cleaning

At the beginning of the research, the target corpora can be downloaded from the official website of Uncannymagazine and Baidu, which could not be downloaded directly due to copyright problems, so they can only be copied and pasted and finally be saved as a txt version. During this period, there will be some problems such as garbled characters, non-standard punctuation marks and space conversion, which will affect the accuracy of data analysis of corpus software and the rigor of this study. Then Editplus, a text-cleaning software, is needed to clean up the texts, remove redundant Spaces and garbled characters, and replace the non-standard punctuation marks in batches, and finally export the edited document.

### 3.2 Corpus tagging

This study attempts to examine Liu Yukun's translation style in terms of words, sentences, and discourse. At the first level, namely the word level, the number of words of different classes needs to be counted respectively, which needs to assign codes to this text. The TreeTagger is a tool for tagging text with part-of-speech and lemma information. First of all, import the translation which has been txt version into Treetagger software, and then choose UTF-8 as encoding form and run. Finally, the tagged text can be stored. Again, this text is imported into Antconc, and the number of words of various classes can be obtained by using regular expressions, which is significant for subsequent research.

### 3.3 Corpus retrieval

In Antconc (Version 4.0.5), mainly functions used in this study of the menu bar are KWIC and Word. Word function can be used to retrieve the frequency of each

Word, and KWIC (Key Words in Text) can see the context of each word, which is convenient to analyze the characteristics of words and sentences in combination with the context. By using and observing the calculated data, what can be obtained is the data of type-token ratio, lexical density, word frequency and average sentence length of the translated version, which is essential for the subsequent analysis of Liu Yukun's translation style from three aspects: lexical, sentence and discourse levels.

## 4 CORPUS DATA ANALYSIS

### 4.1 Standard type-token ratio

Type is similar to words we commonly speak, while the token is the number of types that are not counted repeatedly in a text. By importing the translation version into the corpus software, the type-token ratio (TTR) can be concluded, which is about 18.96. Baker (2000) pointed out that the ratio is related to the richness and diversity of the words used by the translator. The more abundant the words used by the author, the bigger the ratio, and vice versa. However, this correlation will be restricted by the length of the article because of the frequent use and occurrence of functional words. Therefore, for accuracy, the algorithm of standardized type-token ratio (STTR) is adopted, which is showed in the following table:

Table 1: STTR

	Liu	Ant
STTR	45.04	39.26

From the table, it is clear that the standardized type-token ratio of *Folding Beijing* translated by Liu Yukun is about 45.04, while the ratio in the corpus software Antconc is about 39.26. The ratio in Liu Yukun's translation is obviously larger than the reference corpus. So it can be concluded that Liu Yukun's vocabulary is more abundant and varied, which is also inseparable from Liu Yukun's personal experience. His hometown is China. After eight years in China, he began his journey to the United States. Influenced by American culture, he used English as his first language, which enabled him to have a good command of English and have a rich English vocabulary reserve.

### 4.2 Lexical density

English words can be roughly divided into two categories, content words and functional words (Zhang 2020) [11]. "According to Olohan, lexical density refers to the proportion of vocabulary" (Peng Ming, 2018). He also contends the ratio of lexical density helps to reveal the difficulty and information capacity of a text, that is, the higher the ratio is, the more information contained in the corpus. [10]

After being tagged, the file was loaded into Antconc, so as to obtain the number of content words and lexical density of the text based on regular expressions:

Table 2: Lexical density

	Liu	Ant
noun	3932	2787
verb	2437	1850
adjective	988	728
adverbial	1256	889
total Content words	8613	6254
total tokens	16100	12311
lexical density	0.535	0.508

According to Table 2, it is not difficult to find that the density of words in Liu Yukun's translation is over 50%, so the number of content words accounts for the majority, which indicates that Liu's translation contains high information content and rich meaning.

Liu Yukun's life experience enables him not only to extract accurate information from Hao Jingfang's original text but also to convey it in English. His translation has a high lexical density, rich and accurate meaning, and clearly describes the main story of the science fiction *Folding Beijing*.

### 4.3 Word frequency

Li Fan (2020) said that the word with high frequency indicates its importance in the text and the translator's familiarity with the word. It also reflects a translator's style for the fact that the translator will unintentionally use the word with high frequency. The data of the word frequency in Liu Yukun's translation go as follows: [6]

Table 3: Word frequency

Rank	Liu	Ant	Rank	Liu	Ant
1	the	the	11	had	his
2	and	a	12	that	that
3	he	and	13	as	you
4	to	of	14	at	i
5	a	to	15	t	said
6	of	he	16	i	s
7	was	was	17	you	she
8	in	in	18	with	at
9	his	it	19	for	had
10	Lao Dao	her	20	him	with

The word with high frequency in this table can be roughly classified into 6 groups: prepositions, articles, conjunctions, nouns, pronouns, verbs, which can be further discussed from the following three perspectives:

The first type is function words. According to the table above, the top five of the two corpora are prepositions, articles, conjunctions and so on, which are important components of sentences. Meanwhile, the words commonly used by Liu Yukun in this table are basically consistent with the reference corpus. This demonstrates that Liu Yukun's translation is attempting to approach the target language as closely as possible, which is also consistent with English expressing practices.

So take the most frequently occurred conjunction "and" for example. Through using Antconc for retrieval, this conjunction is mostly used in the description of the folding scenes of the three spaces in Beijing and movements of characters which are also heavily portrayed in the source text because of its theme and which are aimed at promoting the plot of the stories. The appearance of conjunctions also reflects the characteristics of hypotaxis in English, which is also a feature of Liu Yukun's translation. One of the most significant distinctions between Chinese and English is hypotaxis and parataxis. Chinese tends to express meanings rather than emphasize the form consistency. Nevertheless, English is inclined to pursue the unity of content and form.

What's more, in this table, nouns with high frequency are Lao Dao and space. The reason is that the protagonist of this book is Lao Dao, and the full text develops as Lao Dao shuttles in three spaces. The repeated use of the main character's name reinforces the reader's memory and promotes the development of the story. As for the word "space", through retrieval, it is found that most of the time the word is written as Space, and the capitalization of the first letter shows the uniqueness of the word, which is referring to the three folding Spaces. In his translation, Liu Yukun achieves consistency between the context, which also makes the whole translation more coherent and readable.

In addition, in the word frequency table, the words that are used less often also need to be analyzed. In Liu Yukun's translation, "t", also known as not, appears frequently, but rarely in the table of the corpus. Through retrieval, it shows that in Liu Yukun's translation, "not" is often used together with "did" and "can" to indicate negation. There are three Spaces in the science fiction novel *Folding Beijing*, which are metaphors for the high, medium, and lower classes. Class solidification is serious, and the bottom people can only constantly compromise with the status quo of life. Liu Yukun often uses "not" in his translation, which represents people's helplessness and conforms to the main meaning of the original text,

which reflects his translation innovation based on faithfulness to the original text.

#### 4.4 Average sentence length

According to Li Fan (2020), the average sentence length is counted through the total token divided by the total sentences in the file[6]. In Huang Libo's opinion (2015), it can be indicative of the difficulty of the text, that is, the longer the average sentence length is, the more difficult the text is. It is well-known that a short sentence is easier to analyze its sentence structure and to grasp its idea. However, the long sentence has complex structures but it shows the internal logical relationship and carries more information. [4]

Table 4: Average sentence length

	Liu	Ant
token	16100	12311
sentence	1156	933
average sentence length	13.9	11.9

From Table 4 above, what can be easily found is that the average sentence length of Liu Yukun's translation is 13.9, only slightly longer than the reference corpus in Antconc, which means Liu Yukun's translation style is similar to American writing habits. Furthermore, "Butler contends that the sentence of 1-9 words is a short sentence while 10-24 is a medium sentence and a sentence with more than 25 words is called a long sentence" (quoted in Li 2020) [6]. So the average sentence length of Liu Yukun's translation belongs to the medium sentence, which is not difficult for the readers to understand and will not increase their reading burden. Thereby, it is natural to draw a conclusion that Liu Yukun uses medium sentences to ensure that his translation is both readable and informative.

#### 4.5 Passive voice sentences

In English, there are two voices, passive voice and active voice (Zhang 2015). [12]

Table 5: Passive voice sentences

	Hao	Liu	Ant
passive voice sentence	25	85	46
sentence	910	1156	1012
Percentage (%)	2.74	7.3	4.5

According to Table 5, Liu Yukun's translation has 85 passive voice sentences of 1156 sentences, and the passive voice sentences account for 7.3%. Generally

speaking, English writing is inclined to use passive voice and emphasizes the impersonal, while Chinese tends to use active voice and pay attention to the personal (Jia, 2020). So Liu Yukun's choice of voice closely conforms to English habits.

#### 4.6 Narrative tenses

The tense that Liu Yukun's translation chooses in his translation can be concluded from Table 3, Liu Yukun mainly uses the past tense in his translation, which can be seen from the frequency of "was, had, were" in the word frequency table, which is consistent with the narrative habit of most English novels. Through the retrieval, what can be concluded is that there's not too much emphasis on tenses in the original text. However, the translation version's choice of the tense can add more heaviness to the specific place, the folding Beijing. Readers can also feel the heaviness of being here, the rigidity of class and people's powerlessness about the status quo. The translator has his own unique views on the choice of tense, rather than just sticking to the original text.

As for the tense, Chinese verbs themselves do not change, while in English the tense is manifested by the verb transformation (Liu, 2009). During the translation, unlike the original text, the translator Liu does some inflectional changes of verbs to highlight the time, which is in accord with English grammar. [9]

## 5 CONCLUSION

Liu Yukun's translation style is collected and evaluated from three different perspectives: words, sentences, and discourse. From the perspective of words, by comparing the STTR, lexical density and word frequency of his translation with those of the corpus, it can be concluded that words which convey rich information in Liu's translation are rich and varied. In addition, the translation connects sentences with conjunctions, which is similar to the target language. From the perspective of sentence level, this paper chooses average sentence length and passive voice sentences to analyze and concludes that the average sentence length of Liu's translation is moderate, which makes the translation both informative and readable. And the use of plenty of passive voice sentences also shows that his translation is closer to the target language. At the level of discourse, this paper finds that in the aspect of narrative tense, Liu has his own ideas and chooses the past tense as the main tense, which is more suitable for the theme of the original text. In general, based on the data of corpus, it can be found that Liu Yukun's translation, on the basis of being faithful to the original text, has made appropriate innovations to make it close to the target culture.

Based on the self-built corpus, this paper adopts classic translation analysis theories to analyze a single

translator's style, but there are still many shortcomings, such as the limitations of the analysis angle, and the analysis results are also affected by subjective factors. At the same time, due to the development of corpus software, some perspectives can not provide accurate data. Although there are some problems due to technical and subjective factors in this study, it also gives a reference to analyze translation style.

## ACKNOWLEDGEMENTS

This research is funded by the Scientific Research Project of Wuhan Business University entitled "A Diachronic Study of Chinese Cultural Image Based on COCA", Grant No. 2020KY002.

## REFERENCES

- [1] Baker, Mona. Towards a Methodology for Investigating the Style of a Literary Translator[J]. *Target. International Journal of Translation Studies*, 2000, 12(2): 241-266.
- [2] Chen Yinlong. On the translator's subjectivity in Ken Liu's translation of *Folding Beijing*[D]. China Foreign Affair University, 2017.
- [3] Hao Jingfang. *Folding Beijing* [M]. 2012. Uncannymagazine. Mar. 2022. <<https://uncannymagazine.com/article/folding-beijing-2/>>
- [4] Huang Libo. *Style in Translation: A Corpus-Based Perspective*[M]. Xi'an: Shanghai Jiao Tong University Press, 2015.
- [5] Hu Kaibao, and Kyung Hye Kim. *Corpus-Based Translation and Interpreting Studies in Chinese Contexts. Present and Future*[M]. Cham: Springer International Publishing, 2020.
- [6] Li Fan. A Corpus-Based Study on The Translator's Styles as Reference in the Two English Versions of *Luotuo Xiangzi* [D]. Shanghai International Studies University, 2020.
- [7] Liu Jinyu, and Hu Jijia. Going Out and Going In! A Study of Translation Norms and Strategies of Translating Chinese Science Fiction by Ken Liu into English-Spoken Countries[J]. *Proceedings of the Seventh Northwest Asia International Symposium on Language, Literature and Translation*, 2018, 6: 197-204.
- [8] Liu Kanglong, and Muhammad Afzaal. Translator's Style Through Lexical Bundles: A Corpus-Driven Analysis of Two English Translations of *Honglouloumeng* [J]. *Frontiers in Psychology*, 2012, 12: 633-422.
- [9] Liu Yuqin. A Comparison of Tense Expression in Chinese, English and French[J]. *Success (education)*2009, (01): 199-200.
- [10] Peng Ming. A Corpus-Based Research on Translator's Style[D]. Dalian International Studies University, 2018.
- [11] Zhang Weiyu. *Lexicology: A Course Book (Third Edition)* [M]. Wuhan: Central China Normal University Press, 2020.
- [12] Zhang Zhenbang. *A New Concise English Grammar* [M]. Shanghai: Shanghai Foreign Language Education Press, 2015.

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

