

A Corpus-based Comparative Study on Two English Translation Versions of “Law of People’s Republic of China on Traditional Chinese Medicine” (the First Two Chapters)

Jinfeng Pan¹

¹ Hubei University of Chinese Medicine, Wuhan, Hubei 430065, China

ABSTRACT

Traditional Chinese medicine has made indelible contributions to the reproduction and health of the Chinese nation, as well as significant contributions to global medical development and world civilization. The “Law of People’s Republic of China on Traditional Chinese Medicine” (hereinafter referred to as the “Traditional Chinese Medicine Law”) came into effect on July 1, 2017, providing legal protection for the development of traditional Chinese medicine in China. This study has selected English translations from the PKULAW Database and the official website of the World Traditional Medicine Forum, and built a bilingual corpus. AntConc, Readability Analyzer, and CUC ParaConc has been used to compare and analyze the high-frequency words, readability, and translation effects of the two translation versions, in order to provide suggestions for the research of English translation of “Traditional Chinese Medicine Law”.

Keywords: *Corpus, “Law of People’s Republic of China on Traditional Chinese Medicine”, Translation comparison, Readability analysis.*

1. INTRODUCTION: RESEARCH BACKGROUND

Traditional Chinese Medicine (TCM) is a collective term for the medicine of various ethnic groups in China, including Han and minority medicine. “It is one of the most representative categories of traditional Chinese science.” [1] In the history of human development, traditional Chinese Medicine has not only made indelible contributions to the reproduction and health of the Chinese nation, but also made significant contributions to the progress of world civilization.[2] On December 6, 2016, the White Paper “Traditional Chinese Medicine in China” was released, stating that “this is the first time the Chinese government has published a white paper on the development of traditional Chinese medicine. Traditional Chinese medicine has been elevated to a national strategy.”[3] On December 25, 2016, with the adoption of the 25th meeting of the Standing Committee of the 12th National People’s Congress,

the first national law for the revitalization of traditional Chinese medicine, the “Law of People’s Republic of China on Traditional Chinese Medicine”, was released and came into force on July 1, 2017. Traditional Chinese Medicine Law fixes the policies and guidelines of the Party and the state on the development of traditional Chinese medicine in legal form, and reflects the expectations and requirements of the people for traditional Chinese medicine in legal form. It reflects the Party and the state’s high attention to the cause of traditional Chinese medicine and has milestone significance for the development of the traditional Chinese medicine industry. After the implementation of the law, multiple translations have been released one after another. “The publication of the Chinese-English version of the ‘Traditional Chinese Medicine Law’ has contributed to the promotion of the law and facilitated the legalization and standardization of traditional Chinese medicine worldwide.”[4] It has also helped overseas people to understand and

accept traditional Chinese medicine culture and technology, and contributed to the development of the world's traditional Chinese medicine industry and the international dissemination of traditional Chinese medicine culture.

2. RESEARCH DESIGN

2.1 Theoretical Basis

Mona Baker, a British translator, pointed out that there is a corpus shift in current translation studies, that is, “the focus of translation studies needs to shift from comparing the original text with the translated text, or comparing A language with B language, to comparing the translated text with the translated text.” [5] Corpus translation studies has undergone decades of development and made remarkable progress, becoming an important paradigm in translation studies. Corpus translation studies has introduced quantitative research methods into translation studies, observing a large number of translation phenomena that cannot be detected by the naked eye based on the advantages of extensive corpus examination and data statistics.[6] Corpus has strong advantages in analyzing language features, translation features, Chinese vocabulary features, translator expression habits, specific word classes, collocations, and so on.

2.2 Research Corpus

This study has established a bilingual corpus for the first two chapters of the “Traditional Chinese Medicine Law”, including the Chinese version, the English translation from the PKULAW Database, and the English translation from the official website of the World Traditional Medicine Forum. The PKULAW Database is the first legal search platform in China to obtain registered trademark certification, jointly launched by Peking University Legal Information Center and Peking University Yinghua Technology Co., Ltd. The World Traditional Medicine Forum, founded in September 2015, is an influential international traditional medicine organization.

2.3 Research Questions

This study will analyze and investigate the high-frequency words, readability, and translation effectiveness of the Chinese and English translations of the "Traditional Chinese Medicine Law" from the perspective of corpus translation

studies. The following three questions will be examined:

- What are the overall characteristics of high-frequency words in different translations of the “Traditional Chinese Medicine Law”?
- Is there a difference in readability between the two translation versions and what are the differences?
- What are the translation effects of the two translation versions?

2.4 Research Steps

This study first uses corpus technology to establish a Chinese original corpus of the first two chapters of the “Law of People’s Republic of China on Traditional Chinese Medicine” and a one to two Chinese-English parallel translation corpus. After cleaning, segmenting, and aligning the corpus, the researchers use Corpus WordParser to segment and label Chinese corpus, and use Tmxmall platform for bilingual alignment. The retrieval software and statistical tools used in this study for analyzing Chinese and English corpora also include corpus analysis software AntConc3.5.8, parallel corpus retrieval software CUC ParaConc V0.3, and readability analysis software Readability Analyzer 1.0. The basic information of the corpus is shown in “Table 1”.

Table 1. Basic information of the corpus

Type	Name	Word Token
Original Chinese text	ch-Traditional Chinese Medicine Law	912
English translation 1	en-PKULAW	1032
English translation 2	en-WTMF	1024

The research uses AntConc 3.5.8 to extract high-frequency words from the first two chapters of “Traditional Chinese Medicine Law” and its two English translations for comparative analysis. Next, it uses BFSU_ Readability Analysis 1.0, selects readability metrics to conduct readability analysis on the two English translations. Finally, this research uses CUC ParaConc V0.3 one-to-many parallel corpus retrieval to compare and analyze the translation performance of the two versions.

3. COMPARATIVE ANALYSIS OF WORD FREQUENCY

“The core of corpus analysis methods is word frequency analysis, which is the foundation of all corpus statistical analysis.” [7] The main method adopted in this study is based on corpus based comparative analysis of word frequency. “High frequency words are a major aspect of feature

analysis in translation language and an important indicator for testing the simplification of translation language.” [8] Using the Word List function of AntConc3.5.8, the researchers sort the high-frequency words in the self-built English corpus, add a list of commonly used stop words, remove numerals, quantifiers, pronouns, function words, etc. with less reference value, and obtain the top 20 high-frequency words, as shown in “Table 2”.

Table 2. Comparison of word frequency between the original text and two translation versions

Rank	Original text of “Traditional Chinese Medicine Law”		PKULAW Database English Translation Version (Translation 1)		World Traditional Medicine Forum English Translation Version (Translation 2)	
	Freq.	Word	Freq.	Lemma	Freq.	Lemma
1	69	中医药	96	tcm	89	tcm
2	26	中医	44	shall	40	shall
3	25	医疗	29	medical	32	medical
4	22	应当	22	state	20	article
5	17	机构	20	article	18	service
6	15	服务	19	institution	17	department
7	13	技术	17	department	17	government
8	12	主管部门	17	government	17	institution
9	12	人民政府	17	service	16	people
10	12	发展	16	health	15	state
11	12	国家	15	people	14	health
12	10	事业	12	competent	12	competent
13	10	医师	12	undertaking	11	level
14	10	规定	11	practice	10	provision
15	9	县级	11	provision	9	china
16	8	举办	10	development	9	county
17	8	国务院	10	examination	9	practice
18	8	执业	10	level	8	administration
19	8	管理	10	medicine	8	chinese
20	8	范围	10	physician	8	clinic

Through the frequency comparison table of the first 20 words in “Table 2”, it is found that the two translation versions basically use the same translation “TCM” when translating the words “中医药” and “中医”. The frequency of “shall”, appearing in both translations is higher than the corresponding word “应当”, in the original text. Through CUC ParaConc bilingual search, it is found that both translations appropriately use “shall” in context to convey the meaning of the original text. In the original text, “国家” and “国务院” appear a total of 20 times. In translation 1, the use of “state” is 22 times, while in translation 2, the use of “state” is 15 times, which is less than the total number of occurrences of the two words in the original text. By comparing the high-frequency word lists of the two translation versions, it was found that there were 9 occurrences of “China” in translation 2, indicating the country as “China”. This shows the differences in word selection between the two

translation versions. This is related to the sources of two versions. Translation 1 is from the PKULAW Database in China, and translation 2 is from the World Traditional Medicine Forum website. The two have different target groups. PKULAW is an intelligent one-stop legal information retrieval platform jointly launched by Peking University Legal Information Center and Peking University Yinghua Technology Co., Ltd., mainly targeting legal professionals; The World Traditional Medicine Forum is an influential international traditional medicine organization aimed at healthcare professionals and enthusiasts worldwide. In addition, the frequency of “physician” appearing in translation 1 is exactly the same as that of the original text “医师”, while there is no corresponding translation for “医师” in the first 20 high-frequency words of translation 2. Through CUC ParaConc search, it is found that translation 2 has two translation versions, “physician” and “practitioner”.

The word “practitioner” is more appropriate according to the dictionary, but translation 1 is better than translation 2 in terms of translation consistency.

4. ANALYSIS OF TRANSLATION READABILITY

Readability is an important evaluation criterion in language acquisition, which evaluates the difficulty of reading articles through a series of indicators. Readability is now also used to evaluate the difficulty of reading translations. There are many categories included in readability indicators, such as Type/Token Ratio, Standardized Type/Token Ratio, average sentence length, average word length, etc. This study will use BFSU_Readability Analysis 1.0 to select the above

indicators for readability analysis of two English translations. The Standardized Type/Token Ratio can be used to indicate the lexical diversity of a discourse, that is, the larger the Standardized Type/Token Ratio, the richer the lexical diversity of the discourse. [9] The calculation method for the average word length of English corpus is based on the number of letters as the word length. The calculation of average sentence length is based on the number of words in a sentence, which is usually used as a parameter to measure the difficulty of a text sentence. The higher the average sentence value, the higher the difficulty. Using Readability Analyzer to analyze the Standardized Type/Token Ratio (STTR), average sentence length, and average word length of the corpus, the following data can be obtained.

Table 3. Comparison of readability analysis results of Readability_Analyzer

File	Reading Ease	Text Difficulty	Grade Level	Sentences	ASL	AWL	Word TTR	Word STTR
en-PKULAW.txt	19.80	80.20	17.59	45	35.50	5.50	0.2433	0.2252
en-WTMF.txt	18.40	81.60	15.89	50	31.40	5.60	0.2742	0.2526

From the above “Table 3”, it can be seen that the STTR value of translation 1 is 0.2252, and the STTR value of Translation 2 is 0.2526. The higher the STTR value, the more word forms are used in the text, and the more difficult it is to read. Translation 2 is greater than translation 1, which reflects that the reading difficulty of the official website translation of the World Traditional Medicine Forum is higher than that of the PKULAW Database translation.

In terms of average sentence length, translation 1 has a longer sentence length, averaging 4.1 words more than translation 2. This reflects the relatively high complexity of sentences in the translated version of the PKULAW Database.

In terms of average word length, there is not much difference between the two translation versions. The average word length of translation 1 is 5.5, and the average word length of translation 2 is 5.6, indicating that the introduction as a whole does not use overly complex words and focuses on simple and clear legal provisions. The differences in readability of translations to some extent reflect the different translation styles of different translations.

5. TRANSLATION EFFECT ANALYSIS

5.1 Accuracy of Wording

In terms of the translation of “中医医师” here, translation 1 uses “TCM physicians”, according to the Oxford Dictionary’s explanation of physician: a doctor, especially one who is specialized in general medicine and not surgery. Physicians refer more to Western medicine physicians, especially internal medicine physicians. At present, the academic community has a high acceptance of the term “practitioner”, which is used in the translation of terms such as “中医医生”, “中医医师”, and “中医从业者”. Therefore, the wording used in translation 2 is more accurate here. (“Figure 1”)

经考试取得 医师 资格的中医 医师 按照国家有关规定，经培训、考核合格后，可以在执业活动中采用与其专业相关的现代科学技术方法。
TCM physicians who have obtained the physician qualification upon examination may, according to the relevant provisions of the state, adopt modern scientific and technical methods relating to their specialties in practice activities after receiving training and passing the examination.
TCM practitioners who have obtained physician qualification upon examinations may, in accordance with relevant provisions of China, adopt modern scientific and technological methods related to their professions in their medical practice after obtaining training and passing the examination.
(一) 中医医疗机构、中医 医师 是否超出规定的范围开展医疗活动；
(1) Whether the TCM medical institution or TCM physician conducts medical activities beyond the prescribed scope.
1. Whether the TCM medical institutions or practitioners conducts medical activities beyond the registered scope;

Figure 1 Translation comparison of 中医医师.

The English translation of “**中西医并重**” here uses “giving equal emphasis to TCM and Western medicine” in translation 1 and “balance between Chinese and Western medicine” in translation 2. The Chinese phrase “equal emphasis” refers to placing equal importance on both traditional Chinese medicine and Western medicine,

emphasizing the importance of both. This means that traditional Chinese medicine treatment and Western medicine treatment should be given equal importance. Translation 1 here conveys the meaning of the original text more accurately. (“Figure 2”)

国家大力发展中医药事业，实行 中西医并重 的方针，建立符合中医药特点的管理制度，充分发挥中医药在我国医药卫生事业中的作用。
The state shall vigorously develop the TCM undertaking, carry out the principle of giving equal emphasis to TCM and western medicine, establish the management rules in compliance with TCM features, and maximize the role of TCM in the medicine and health undertaking of the state.
The state shall vigorously develop TCM undertaking, carrying out the principle of balance between Chinese and Western medicine, establishing an administrative system that conforms to TCM features and giving TCM full play in the medical and healthcare affairs in China.

Figure 2 Translation comparison of 中西医并重.

5.2 The Use of Translation Techniques

Translation conversion refers to the process of changing the role of parts of speech or words in a sentence due to differences in grammar and habitual expressions between two languages, while ensuring that the original meaning remains

unchanged, in order to make the translation conform to the expression of the target language. Both translations here convert the active form in Chinese into the passive form in English, which is more in line with the characteristic of “making good use of passive sentences” in English and helps the publicity of the law. (“Figure 3”)

发展中医药事业应当 遵循 中医药发展规律，坚持继承和创新相结合，保持和发挥中医药特色和优势，运用现代科学技术，促进中医药理论和实践的发展。
In the development of the TCM undertaking, the TCM development rule shall be observed, the integration of inheritance and innovation shall be adhered to, TCM features and advantages shall be maintained and maximized, and modern scientific technologies shall be utilized, so as to promote the development of TCM theories and practice.
In the development of the TCM undertaking, the core value and laws of TCM development shall be followed, the integration of inheritances and innovation adhered to, TCM features and advantages preserved and maximized, and modern scientific technologies utilized, so as to promote the development of TCM theories and practice.

Figure 3 Comparison of the use of translation techniques.

5.3 The Coherence of the Text

Realizing intra-textual coherence refers to the translation being able to be accepted, understood, and disseminated in the context of the target language, taking into account both the cultural background of the recipient and the cultural context of the target language. It should not only be

readable but also acceptable. Despite the division of legal provisions, compared to translation 2, translation 1 uses prepositional phrases at the beginning of the sentence to connect the preceding text, achieving intra-textual coherence and improving the acceptability of the translation. ("Figure 4")

<p>第二条 本法所称中医药, 是包括汉族和少数民族医药在内的我国各民族医药的统称, 是反映中华民族对生命、健康和疾病的认识, 具有悠久历史传统和独特理论及技术方法的医药学体系。</p>
<p>Article 2 For the purpose of this Law, "traditional Chinese medicine" (TCM) is the general term of medicine of all Chinese ethnic groups including the Han ethnic group and minority ethnic groups, reflects the understanding of the Chinese nation on life, health and diseases, and is a medicine and pharmacology system with long historical traditions and unique theories and technical methods.</p>
<p>Article 2 Traditional Chinese medicine (hereafter referred to as TCM) in this law is a general term for the medicine of all ethnic groups including the Han nationality and ethnic minorities in China. It is a medical and pharmacological science system that reflects Chinese people's understanding of life, health and disease, with a long history and unique theories and techniques.</p>

Figure 4 Comparison of text coherence realization.

6. CONCLUSION

This study has selected English translations from the PKULAW Database and the official website of the World Traditional Medicine Forum, and built a bilingual corpus. AntConc, Readability Analyzer, and CUC ParaConc are used to compare and analyze the high-frequency words, readability, and translation effects of the two versions. Research has found that there is generally little difference in the high-frequency words between the two translation versions, but there are significant differences in the translation of individual words, which is related to the sources of the two versions and their target groups. Furthermore, in the comparative analysis of high-frequency words, it is found that translation 1 outperformed translation 2 in terms of translation consistency. The readability analysis found that the reading difficulty of the official website translation of the World Traditional Medicine Forum is higher than that of the translation from the PKULAW Database. The differences in readability of translations to some extent reflect the different translation styles of different translations. Translation effect analysis found that translation 1 is generally better than translation 2 in terms of word coherence.

The corpus-based research method has practical effectiveness in revealing the readability, word accuracy, and coherence of legal translation. When translating "Traditional Chinese Medicine Law", it is necessary to fully understand the professional terminology and concepts in the fields of traditional

Chinese medicine and law, and ensure the accuracy of the wording. In addition, it is necessary to consider cultural differences, make appropriate adjustments and conversions, and further promote the international dissemination of traditional Chinese medicine.

REFERENCES

- [1] Wang Guoqiang, Promote the Revitalization and Development of Traditional Chinese Medicine with High Cultural Confidence [N]. People's Daily, 2017-02-24 (7).
- [2] Xu Jun, Song Jiayin, Research on Foreign-Related Linguists Training in Traditional Chinese Medicine from the Perspective of Building a Community with a Shared Future for Mankind: Developments, Questions and Recommendations [J]. Language Service Research, 2022(0): 22-44.
- [3] Xinhua News Agency, The State Council Information Office releases the White Paper on Traditional Chinese Medicine in China [EB/OL]. 2016-12-06. http://www.gov.cn/xinwen/2016-12/06/content_5144024.htm#2.
- [4] Cui Fang, The Chinese-English Bilingual Version of the "Traditional Chinese Medicine Law" Has Been Released for the First Time [N]. Health News, 2017-04-21 (2).

- [5] Baker Mona, Corpus-based Translation Studies: The Challenges that Lie Ahead [A]. In H. Somers (ed.) Terminology, LSP and Translation: Studies in Language Engineering, in Honour of Juan C. Sager [C]. Amsterdam and Philadelphia: John Benjamins, 1996: 175-186.
- [6] Zhang Li, Research on the Theory and Future Development Direction of Corpus Translation Studies: A Review of Corpus Translation Studies [J]. Yangtze River, 2020, 51(11): 228.
- [7] Gui Shichun, A Corpus-based Analysis of the Register of the English Linguistics [M]. Beijing: Foreign Language Teaching and Research Press, 2009.
- [8] Gong Mingyu, Research on Legal Visibility Based on Corpus [M]. Beijing: University of International Business and Economics Press, 2021.
- [9] Guan Xinchao, Corpus and Python Applications [M]. Shanghai: Shanghai JiaoTong University Press, 2018.