A Corpus-based Study: A Stylistic Analysis of English Translation of the Articles of Civil Code of the People's Republic of China Involve Intellectual Property

Can Yang, Ziyi Zhao, Ran Zhuo, Chuhan Zhang, Qianru Zhang, Xiaowen Ji

University of Shanghai for Science and Technology, Shanghai, 200000, China

Abstract

On May 28, 2020, the Third Session of the 13th National People's Congress (NPC) adopted the Civil Code of the People's Republic of China. But the relevant parts of intellectual property rights still remain to be analyzed in the English version of the Civil Code issued by the General Office of the NPC Standing Committee. Therefore, our study aims at analyzing the frequency of words to get the main content of the code, and analyzing the average sentence length, sentence structure and difficulty of the text to present the acceptability of the text. The study analyses vocabulary, sentence and discourse by means of CROWN corpus created by Xu Jiajin and Liang Maocheng, and tools such as antconc 3.5.7, range3.2.3, Readability Analyzer. The purpose of this study is to provide stylistic reference and translation suggestions for translators and legal practitioners in their legal translation practice, and to show China's presence in cross-cultural communication in the field of law. More further researches are expected.

Keywords

Civil Code; Intellectual property; Corpus; Stylistic characteristics.

1. Introduction

The Civil Code of the People's Republic of China (hereinafter referred to as the Civil Code) occupies a fundamental position in the legal system, and its promulgation and implementation will greatly promote the rule of law in China and the construction of socialism with Chinese characteristics. As the "encyclopedia of social life", the Civil Code also covers intellectual property protection, for example: An important change in intellectual property protection in the Code is to clarify the punitive provisions of intellectual property. The early study of stylistics in linguistics mainly depended on the intuition of researchers. With the development of science and technology, corpus method and experimental method are applied to stylistic research, making it an empirical science combining quantitative and qualitative research [2]. Corpus linguistics and stylistics both focus on the characteristics of language use, and both focus on the form and meaning of language, which determines the rationality and necessity of the combination of the two. Corpus of research methods to study work style characteristics, for different research purposes, corpus research in stylistics can be done from different aspects, such as parts of speech distribution features, syntactic features, word frequency, subject headings, collocations, etc., from different aspects to study can make readers understand the contents of the text, character, theme expression, and so on.

This study takes the relevant parts of the English version of the Civil Code issued by the General Office of the Standing Committee of the National People's Congress as the texts to be analyzed. 218-222) the 32 indexes (such as nominalization, subordination, proper nouns, etc.) listed in the "Stylistic Component Table" are labeled and counted to further explore its stylistic characteristics. Editplus3 text collator is used to clean the text and remove redundant Spaces, garble and other influencing factors to improve the credibility of the research results. At the

same time, annotated statistics of the original text are also carried out, and the comparison between the original and the translated text and the attribution analysis of similarities and differences are carried out. Using the research method of corpus to make a quantitative analysis of the text, it changes the defect of the previous stylistic research only from the author's intuition. Therefore, based on the Corpus of contemporary American English, this study studies the Civil Code from the perspective of stylistics, in order to provide some enlightenment for the study of translation commonality and legal translation practice in the field of intellectual property.

2. A Study of Stylistic Feature at the Lexical Level

2.1. Type Token Ratio

Token is a language unit, which refers to the total number of words in the text. Type is a statistic, which refers to the number of words that are not repeated in the corpus [3]. The use of words in the corpus can be judged by type token ratio (TTR). Baker [4]believes that the TTR value is directly proportional to the richness and diversity of the author's or translator's vocabulary use. Standard type token ratio (STTR) refers to calculating the class token ratio of text in batches within a certain text length, and then calculating the average value of each class token. Compared with TTR, STTR is not affected by the capacity of corpus, so it is more reliable [5]. The author uses antconc 3.5.7 to count the number of type and token, divides the text into groups of every 1200 tokens, and obtains the civil code of the people's Republic of China (hereinafter referred to as the civil code) according to the calculation formula of type token ratio (TTR), that is, TTR = type / token * 100% The STTR of the articles related to intellectual property rights is 24.33%, and the STTR of the reference corpus crown is 46.53%.

Range3.2.3 (http://www.vuw.ac.nz/lals/staff/Paul_Nation.) The tools are divided into level 1, level 2 and level 3 vocabularies. Level 1 and level 2 vocabularies are the most commonly used and commonly used vocabularies respectively, which are from the general English vocabulary [6] compiled by Michael West. Level 3 vocabularies are academic English vocabulary. Liang Maocheng [7] have pointed out that the vocabulary outside the three-level vocabulary of class symbols is a reliable criterion for judging the difficulty. Therefore, it is concluded that the civil code In the relevant provisions of intellectual property rights, the proportion of the third level vocabulary and words outside the vocabulary of the category symbol is 24.48%, which shows that the text is not difficult, easy to understand and easy for people to understand, so that people can better use legal tools to protect their legitimate rights and interests.

2.2. Statistics on Word Frequency of English Translation of Some Provisions of Intellectual Property in Civil Dode

Table 1. The statistics on word frequency

Number	Word List	Frequency	Number	Word List	Frequency
1	the	389	11	technological	56
2	a	150	12	shall	53
3	of	128	13	by	46
4	and	124	14	patent	43
5	to	87	15	for	39
6	or	83	16	development	39
7	contract	69	17	is	36
8	technology	68	18	parties	35
9	in	63	19	work	35
10	article	58	20	where	33

The author uses antconc 3.5.7 to count the word frequency in the text. The function words in the relevant provisions of intellectual property in the English version of the civil code rank first. Because notional words can reflect the main content of the text. The author retrieved the top ten high-frequency notional words (see Table 2), from which we can see that the core contents of the text include contract, technology, article, technical patent, development, parties, right, person and association. It can be seen that the laws related to intellectual property in the civil code mainly protect the intellectual property of civil legal subjects from the perspective of contract law. At the same time, the intellectual property part of the civil code focuses more on From this, we can know that the civil code effectively guides people how to protect their intellectual property rights through patents, contracts, etc.

Table 2. Word list of notional words

Number	Word List
1	contract
2	technology
3	article
4	technological
5	patent
6	development
7	parties
8	right
9	person
10	accordance

3. A Study of Stylistic Features of Sentence

3.1. Statistics of Average Sentence Lenghth of English Translation of Some Povision of Intellectual Property in Civil Code

In order to explore the sentences related to intellectual property in the English translation of the Civil Code of the People's Republic of China, we need to take into account the particularity of the legal text. The legal text not only involves the seriousness of the text itself, but also takes into account the needs of the reading object, even in a clear understanding. The calculation of average sentence length is based on the number of words contained in the sentence. This value is usually used as a parameter to measure the difficulty of the text. The greater the average sentence length, the higher the difficulty. However, there are different results for specific texts. The average sentence length is calculated by dividing the total number of words by the number of sentences, and through search, statistics and calculation in Microsoft office word. In the Civil Code of the People's Republic of China, the average sentence length of the intellectual property part in the English translation is 31.28. From the data, it belongs to a complex text, but if you look carefully at the structure of the text, you will find that although there are only 117 punctuation marks at the end of the sentence, there are 219 commas. [8] Because the texts related to scientific and technological property rights will involve many parallel forms and rigorous considerations. In the Civil Code of the People's Republic of China, the comma in the intellectual property part of the English translation not only does not increase the complexity of the text, but also makes the entries intuitive and easy to understand. [9]

In the English translation of the Civil Code of the People's Republic of China conforms to the logic and way of English expression, and the terms are listed clearly for foreigners to understand and read. The method of payment for the price, compensation, or royalty shall be

agreed by the parties in a technology contract, and the payment may be made in a lump sum or by installation based on one-time calculation, or based on the method of commission payment or such payment plus an upfront fee. (article 846) This sentence introduces the payment methods of price, remuneration and royalties, clarifies the subject with specified power, and uses or to connect four feasible payment methods. The structure is clear and the meaning is clear, which meets the special needs of strict and accurate legal texts.[10]

3.2. A Study on the Passive Voice in the English Translation of Some Provisions of Intellectual Property in the Civil Code of the People'S Republic of China

Generally speaking, English passive voice is used to highlight the receiver of action. In legal texts, it is usually used to emphasize the legal subject or behavior. Many passive voice sentences play a general role in an item. In this sentence, the trustee is connected after by, that is, the subject of expenses, from which the reader can clearly understand the responsible party. In the Civil Code of the People's Republic of China, there is a sentence——Where is no agreement in a technology consultation contract or a technology service contract on the bearing of the necessary expenses for the trusted person to carry out the normal work, or the relevant agreement is unclear, the said expenses shall be borne by the entrusted person. (Article 886) Passive voice can also make sentences coherent, emphasize more, and more close to the expression habit of English. For example, A work for hire is a technical achievement that is compiled as a result of performing the tasks assigned by a legal person or unincorporated organization or that is incorporated mainly by using the materials and technical resources of the said legal person or unincorporated organization. (article 847) The use of the past participle of be + verb in this sentence makes the sentence structure clear and concise.

4. An Analysis of the Readability of the English Version of the Civil Code

Using the corpus analysis tool: readability analyzer to measure the difficulty of the text can show the acceptance of it. The text readability detection tool analyzes the readability of articles through Flesch reading ease. Flesch reading ease is calculated according to the statistical method of Dr. Rudolf Flesch of the United States. [11] The calculation is based on the number of words and syllables contained in the sentence. The value is between 0 and 100. The larger the number, the easier the article is to read. The calculation formula is as follows:

$$206.835 - 1.015 \left(\frac{\text{total words}}{\text{total sentences}} \right) - 84.6 \left(\frac{\text{total syllables}}{\text{total words}} \right)$$

The Flesch reading ease of the content of science and technology projects in the civil code is 12.07 (see Table 1), which shows that the text difficulty of the content of science and technology projects in the English translation of the civil code is very difficult. Readers need to have a certain vocabulary reserve. [12]

Table 3. Comparison of difficulty

0 – 30	Very difficult
30 – 50	Difficult
50 – 60	Fairly difficult
60 – 70	Standard
70 – 80	Fairly easy
80 – 90	Easy
90 – 100	Very easy

5. Conclusions

Using the research method of corpus linguistics, this paper studies the English translation of the civil code of the people's Republic of China from the perspective of stylistics. The results show that the text has low vocabulary density, low difficulty and rigorous wording; Because the text adopts the form of general long sentences and clause listing, the discourse is more difficult. However, its structure is clear and has distinct text characteristics, so it has a certain readability.

References

[1] General Office of the Standing Committee of the National People's Congress (2020). Civil code of the people's Republic of China. Beijing: China democracy and legal system press.

[2] Lei Xi, Zhang Delu. New development of research methods of modern stylistics [J]. Modern foreign languages, 2016 (02).

[3] Baker, Mona.Corpus linguistics and translation studies:implications and applications [A]. In M.Baker.et al(eds.).Textand Technology[C].Amsterdam:Benjamins.1993:233-252.

[4] Baker,Mona.Corpora in translation studies: an overview and some suggestions for future research [J].Target,1995(2): 223-243.

[5] Baker M.Towards a Methodology for investigating the Style of a Literary Translator [J]. Target, 2000 (02): 241-266.

[6] Michael West A General service List of English Words ,1953.

[7] Liang Maocheng, Li Wenzhong, Xu Jiajin. Corpus application course [M]. Beijing: foreign language teaching and Research Press, 2011.

[8] Si Junlong, Stylistic features of the English translation of the white paper progress in China's human rights cause based on corpus, Journal of Shangqiu Vocational and technical college, 2018, (05): 48-49.

[9] Liu Zequan, Liu chaopeng, Zhu Hong, On the translator's style of four English versions of a dream of Red Mansions——corpus based statistics and analysis, China translation, 2011, (01): 61-63.

[10] Huang Libo, Zhu Zhiyu, A corpus study of translator's style —— Taking Ge Haowen's English translation of modern and contemporary Chinese novels as an example, Foreign language research, 2012, (05): 65-67.

[11] Wang Shuangshuang, corpus based analysis of lexical features in the English translation of the dust settled [J]. English Square, 2015 (06):36-37.

[12] Li Yao, a corpus based study of English translation manifestation -- Taking the 2020 government work report as an example [D], overseas English, 2021 (16):35-36+54.