# Translator Stylistics: A Corpus Analysis of Two English Translations of Traditional Chinese Medicine Classic *Da Sheng Bian*

Yang Zhang[†], Yue'e Zhu*

Central South University of Forestry and Technology, Changsha 410004, China
[†]*First author:* Yang Zhang, Zhangyang98@163.com

*Corresponding author:* Yue'e Zhu, yuezhu@csuft.edu.cn

**Abstract:** *Da Sheng Bian*, a significant work on obstetrics and gynecology that emerged in the early Qing Dynasty, was initially published as "A Treatise on Midwifery" in 1842 by William Lockhart, a British missionary to China. In 1894, John G. Kerr, an American missionary, translated the text as "The Tat Shang Pin". This paper conducts a comparative study of the two English translations using a self-constructed English-Chinese parallel corpus of *Da Sheng Bian*. The study explores the translation styles of the two translators by examining the token-types ratio and frequency at the lexical level, mean sentence length at the syntactic level, and the use of conjunctions at the discourse level. The observed differences in translation styles between the two translations are analyzed in relation to the translators' backgrounds and translation strategies.

**Keywords:** Corpus; Da sheng bian; Translator style

## 1. Introduction

*Da Sheng Bian* ( 达生编 ), also known as *Da Sheng Pian*, stands as a renowned monograph on obstetrics from the Kangxi era of the Qing Dynasty. The author, known as Ji Zhai Jushi, originally hailing from Xuning County in Anhui Province but spent a significant period in Huoshan, was revealed by Li (2009) to be Feng Ye, elegantly styled as Wei Feng, after conducting thorough research and consulting the Huoshan County Records. Ji Zhai continued to be his preferred pseudonym. He was proficient in medicine and served in the Nanchang County government during his middle years. Upon retirement, he dedicated his life to aiding the common people with his medical knowledge [1]. The book, approximately 10,000 words long, is a distinctive work on obstetrics from the early Qing Dynasty, published in the 54th year of the Kangxi era (1715).

The English translation of this medical work began with early Protestant missionaries to China at the end of the Qing Dynasty. The earliest English translation, "A Treatise on Midwifery" by William Lockhart, was published in *The Dublin Journal of Medical Science* in January 1842 [2]. Another English translation, "Tat Shang Pin," or "Midwifery Made Easy," was done by John Glasgow Kerr, a missionary from the United

States [3,4]. This version was published in 1894 in *the Annals of Gynecology and Pediatric* and *The American Gynecological & Obstetrical Journal*. Both translations found their place in national medical journals, indicating the translators' intention to target medical professionals and individuals interested in medicine in their respective countries, emphasizing the professional guidance embedded in this Chinese text.

Currently, research results on the English translation of *Da Sheng Bian* are relatively scarce, focusing mainly on three aspects: the dissemination of the two translations [5], the English translation methods and techniques of Lockhart under the alienation strategy [6], and research on the historical context, purpose, motivation, strategy, method, and quality of Kerr's translation [7]. Comparative studies of the two translations under other translation theories are still in their infancy. This paper aims to bridge this gap by comparing the two translations at both the linguistic and non-linguistic levels, analyzing the translator's style in both instances.

## 2. Translator Style

Hu and Xie (2017) have summarized that the primary focus of translator-style research encompasses two key aspects: the linguistic feature level and the non-linguistic feature level. Examination of translator style at the linguistic feature level involves the study of vocabulary, syntax, collocation, semantic rhyme, and part of speech, respectively [8]. Given that Chinese medical texts fall within the realm of scientific and technological texts, the integration of corpus technology with research on the English translation of scientific and technological texts is predominantly observed in the construction of corpus ontology and the application of corpus technology [9].

The translator's style at the non-linguistic feature level primarily pertains to translation strategies and methods. This paper delves into the analysis of the translation styles employed in the two translations, focusing on vocabulary, syntax, and discourse. Additionally, it seeks to further explore the origins of the translator's style through an examination of translation strategy and translation background.

## 3. Linguistic features of English versions of *Da Sheng Bian*

Currently, two full translations of *Da Sheng Bian* are available worldwide: William Lockhart's translation (1842) and John Glasgow Kerr's translation (1894), both published in the nineteenth century.

Initially, a parallel corpus of the original and multiple translations of *Da Sheng Bian* was constructed through key steps, including corpus electronicization, corpus cleaning, manual error correction, and corpus alignment. Subsequently, AntConc 4.2.0, ParaConc, WordSmith 4.0, and Online Word Counting Tool were utilized to examine the translation styles of the two versions at the lexical, sentence, and discourse levels.

### 3.1. Lexical level

The vocabulary level of the two translations was primarily explored from three aspects: type-token ratio, average word length, and vocabulary frequency.

#### 3.1.1. Type-token ratio and mean word length

Since Mona Baker (2000), many researchers have employed the standardized type-token ratio (STTR) to analyze translators' linguistic features in translation [10]. Tokens refer to the total number of words in a text, while types are the number of non-repeated tokens. The types/tokens ratio reflects the richness of vocabulary, with a higher STTR indicating greater vocabulary variation and text richness. STTR refers to the ratio of types to tokens per 1,000 words, which is relatively more objective and accurate.

**Table 1.** Characteristics in Lockhart's translation and Kerr's translations at the lexical level

|  | Tokens | Types | STTR | Mean word length |
|---|---|---|---|---|
| Lockhart's translations | 12,505 | 2,338 | 42.77% | 4.32 |
| Kerr's translations | 8,570 | 1,675 | 40.26% | 4.28 |

The analyzed data, as shown in **Table 1**, indicates that Lockhart's translation has more tokens, suggesting a more comprehensive interpretation of the original test and a tendency toward explicit translation to convey implicit meanings. In terms of vocabulary richness, Lockhart's translation slightly surpasses Kerr's translation. In the final sections of "Maxim" ( 格言 , *geyan*), "Prescription" ( 方药 , *fangyao*), and "Appendix" ( 附录 , *fulu*), both translators omitted to translate the "Appendix" section, while Kerr also omitted to translate the "Maxims" and "Prescriptions" chapters. The fact that Kerr's translation leaves out relevant chapters from the original work and lacks numerous details is another factor contributing to its significantly shorter length than Lockhart's version. For instance, the subheading " 胞 衣 不 下 " (*baoyibuxia*) was interpreted as "Regulating Cases in which the Membranes Do Not Descend" by Lockhart, but Kerr left this title untranslated. Likewise, the title "乳少" (*rushao*) was translated as "Deficiency of Milk" by Lockhart, yet Kerr opted not to translate it. This approach was also applied to the chapters "Maxim" and "Prescription," which Kerr opted to leave untranslated.

Mean word length is defined as the average length of types in a translation, where greater mean word length indicates greater long and difficult words used in the translation, thereby making the translation difficult to understand [11]. Ordinary texts consist of more words with 2–5 letters and have an average word length of about 4 letters. Both versions exhibit similar mean word lengths (Lockhart: 4.32, Kerr: 4.28). Both translations predominantly use medium-length words, resulting in texts of moderate difficulty to comprehend.

### 3.1.2. Word frequency

Word frequency, reflecting the frequency of vocabulary use in a corpus, is often employed in translation studies to examine authors' or translators' preferences and text characteristics in terms of vocabulary use. **Table 2** illustrates the statistical results of the 36 most frequently used words in the self-constructed corpus of Lockhart's and Kerr's translations by ParaConc Demo software.

**Table 2.** High-frequency words (HFWs) in both translations

|  | HFWs in Lockhart's | % | HFWs in Kerr's | % |
|---|---|---|---|---|
| High-frequency pronouns and nouns | it | 1.43 | it | 2.29 |
|  | she | 0.92 | they | 0.64 |
|  | they | 0.68 | child | 0.63 |
|  | her | 0.61 | her | 0.51 |
|  | woman | 0.42 | she | 0.48 |
|  | child | 0.39 | patient | 0.39 |
| High-frequency conjunctions | and | 4.52 | and | 3.93 |
|  | of | 2.76 | of | 2.42 |
|  | or | 1.16 | when | 0.77 |
|  | if | 0.75 | if | 0.76 |
|  | but | 0.74 | but | 0.74 |
|  | then | 0.45 | or | 0.73 |
|  | so | 0.40 |  |  |

The highest frequency word in both translations is "it," with Kerr's version using it more frequently despite having fewer tokens. This suggests Kerr's preference for vague reference. Regarding personal pronouns and nouns, both translations share high-frequency words such as "they," "child," and "her," with differences in frequency. The words "woman," "her," and "child" are found in both translations, showing the main objects of "obstetrics" in this text are "maternity" and "infant." Lockhart's translation emphasizes "woman" more, while Kerr's translation opts for "patient," which shows that Lockhart is more inclined to express the expectant mother's gender characteristics when describing puerpera, while Kerr's translation aligns with its focus on Western medical practitioners, differing from Ji Zhai Jushi's original intent to present *Da Sheng Bian* as a universal and popular folk obstetrics book.

Regarding the use of high-frequency conjunctions, Lockhart's translation utilizes a more varied range of conjunctures compared to Kerr's translation, enhancing the coherence of the text. The specific high-frequency conjunctions used differ slightly between the two translations, indicating stylistic distinctions.

## 3.2. Syntactic level

Linguistic features at the syntactic level encompass mean sentence length and sentence standard deviation. Mean sentence length denotes the average number of words in each sentence, while sentence standard deviation measures the variation in sentence length from the mean. The syntactic level primarily gauges the translator's style through average sentence length, a factor directly correlated with syntactic maturity and complexity [12]. Wang (2003) noted a common trend of increased word count in translations [13]. In this study, the online statistical tool Online Word Counting Tool and online software WordSmith 4.0 were utilized to calculate the average sentence length and other data for the two translations.

**Table 3.** Sentence-level data features

|  | Lockhart's translation | Kerr's translation |
| --- | --- | --- |
| Total number of sentences | 380 | 423 |
| Mean sentence length | 32.92 | 10.12 |
| Sentence standard deviation | 30.11 | 11.76 |

Analysis of the data reveals that the total number of sentences in Lockhart's translation is fewer than in Kerr's translation, suggesting that Kerr's translation provides a more detailed elaboration of the original text. The average sentence lengths for Lockhart and Kerr are 32.92 and 20.12, respectively, with standard deviations of 30.11 and 11.76, respectively. Butler (1985) classified sentences into three categories based on length: short sentences (1–9 words), medium-length sentences (10–25 words), and long sentences (more than 25 words) [14]. Both translations predominantly utilize medium and longer-length sentences. Lockhart's translation, with its longer average sentence length, favors the use of extended sentences, aligning with its aim to convey science-based information in an informative manner. A notable difference lies in the standard deviation of sentence lengths between the two translations, with Lockhart's translation exhibiting a high standard deviation of 30.11. This can be attributed to the excessive use of punctuation, particularly in listing various types of medicinal herbs in the "Prescription" chapter.

## 3.3. Discourse level

"Cohesion" and "coherence" stand as the fundamental properties of English discourse [15]. The linguistic components that convey semantic connections form the articulation of discourse. Conjunctions serve as a

primary means of expressing the articulation of a text. Through a comparison of conjunction usage, this study aims to discern the linguistic characteristics of the two translators at the discourse level.

The original text of *Da Sheng Bian*, originating from the Qing Dynasty, belongs to the ancient Chinese language and is crafted in a classical writing style. Commonly employed conjunctions in ancient Chinese, such as "与" (*yu*), "及" (*ji*), "而" (*er*), "则" (*ze*), "且" (*qie*), "况" (*kuang*), "但" (*dan*), "抑" (*yi*), "虽" (*sui*), "然" (*ran*), "故" (*gu*), "苟" (*gou*), "倘" (*tang*), "令" (*ling*), "第令" (*diling*) "籍使" (*jishi*), "之" (*zhi*), and more, dominate the entire text of *Da Sheng Bian*. Given its reliance on literary phrases, scholars investigating the discourse logic of Chinese medical texts have compared the distinctions between Chinese medical texts and English discourse. They have proposed that English translations should visualize methods of articulation. Any coherent English discourse is composed of semantically linked sentences, and this semantic relationship must be conveyed through specific means of utterance articulation [16].

### 3.3.1. Classification of conjunctions

Connectives play a crucial role as formalized markers that deepen the logical relationships between sentences, contributing to the articulation and coherence of discourse [17]. Scholars differ in their classifications of connectives; Halliday and Hasan (1976) and Quick (1985) abroad [15,18]; Huang (2012) and Ju (2018) in China have proposed different classifications [19,20]. Ju (2018) synthesized various scholars' classifications and, considering the characteristics of English-translated discourse of canonical books, formulated seven classifications for connectives in English translation discourse of classics books: listing, additional, contrastive, appositional, resulting, inferential, and summative [20].

Given that the traditional Chinese medicine (TCM) obstetrics text *Da Sheng Bian* and Ju's research topic are both Chinese canonical texts, and their translations exhibit characteristics of classical English-translated discourse, this paper, relying on Ju's categorization of conjunctions, conducts a statistical analysis of the distribution of each type of conjunctions.

### 3.3.2. The overall distribution of conjunctions in the two translations

In analyzing the distribution of conjunctions, conjunctions in both translations were initially screened and classified based on semantics, followed by the compilation of frequency data. The statistically examined conjunctions include "first," "next," and "then" under "listing"; "and," "and also," "also," and "too" under "additional"; "but," "yet," "however," "though," "although," "instead," and "on the contrary" under "contrastive"; "that is" and "for instance" under "appositional"; "at the same time," "hence," "so," "therefore," "thus," and "in consequence" under "resulting"; "therefore" and "otherwise" under "inferential". **Table 4** shows the most frequent conjunctions of each type.

**Table 4.** The frequency of conjunctions in two translations

| Types | Conjunctions | Lockhart's translation | Kerr's translation |
|---|---|---|---|
| Listing | Then | 61 | 25 |
| Additional | And | 555 | 337 |
| Contrastive | But | 86 | 63 |
| Appositional | That is | 5 | 0 |
| Resulting | So | 52 | 28 |
| Inferential | Therefore | 5 | 6 |

The table reveals that in both translations, "then" in "listing," "and" in "additional," "but" in "contrastive," "that is" in "appositional," "so" in "resulting," and "therefore" in "inferential" are the most frequently used conjunctions. The frequencies in both translations are 61/25, 555/337, 86/63, 5/0, 52/28, and 5/6, respectively. Overall, the frequency of conjunctions in Lockhart's translation is higher than in Kerr's translation, emphasizing its prominence in text articulation and coherence.
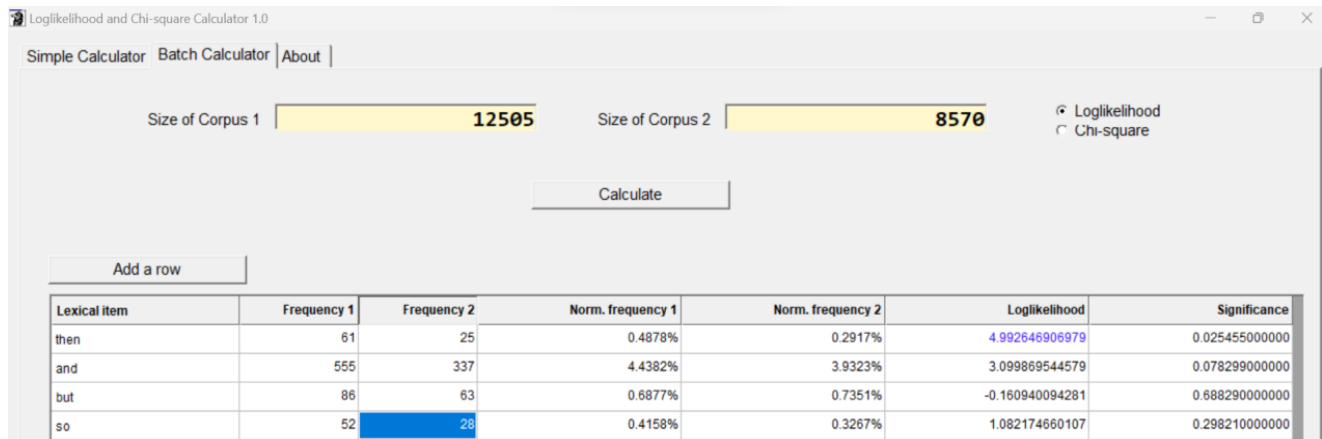


**Figure 1.** Log-likelihood of high-frequency conjunctions in both translations

Next, the log-likelihood of high-frequency conjunctions in both translations was examined, and their frequencies were compared using the chi-squared and log-likelihood calculator (**Figure 1**). A log-likelihood greater than 3.8 indicates a significance below 0.05. The standard frequency of "then" in Lockhart's translation is 0.48, compared to 0.29 in Kerr's translation, with a significance of 0.0254, indicating a significant difference. The log-likelihood of "and," "but," and "so" in both translations are 3.09, 0.16, and 1.08, respectively, all below 3.8. The significance values are 0.078, 0.688, and 0.298, respectively, all exceeding 0.05, indicating no significant difference. This suggests general consistency in the use of the three types of conjunctions between the two translators in both translations.

In summary, the frequency of conjunctions is generally higher in Lockhart's translation, mainly attributed to the prominent use of "then". This alignment in discourse indicates that both translators adhere to the linguistic features of their respective languages, employing an explicit translation strategy to enhance sentence cohesion. Throughout the translation process, the emphasis on English sentence rigor, inclusion of appropriate conjunctions, and clarification of logical relationships within the discourse contribute to translations that are objective, accurate, comprehensive, and easily understandable for readers.

## 4. The non-linguistic factor of translator style in English versions of *Da Sheng Bian*

This section delves into the non-linguistic aspects influencing the formation of the translator's style, examining the translation context and strategies employed in the two translations.

### 4.1. Translation background

Lockhart completed the translation in 1842, following his arrival in Guangzhou at the end of 1839, where he joined the Chinese Medical Missionary Association. During intermittent periods working at a hospital in Macau while studying Chinese in Guangzhou, he finalized the translation of *Da sheng bian* and published it in *The Dublin Journal of Medical Science*, edited by Thomas Ledwich, a prominent Irish anatomist and

surgeon. Lockhart's translation occurred during his missionary work in China, serving the missionary cause of his country. It marked the initial significant effort to introduce *Da Sheng Bian* to the English-speaking world. Fleetwood Churchill, Lockhart's friend, concluded that this translation was Lockhart's first attempt at translating Chinese medical works. Churchill hoped the journal would continue to spotlight Lockhart's medical practice in China, making the missionary community's work known to medical readers in English-speaking countries [2].

Kerr encountered *Da Sheng Bian* during an obstetrical procedure and developed an interest in it. His purpose in translating the work was to present the various popular solutions to obstructed labor in the book, along with highlighting the sophistry and inconsistencies in the original author's views [3,4]. Kerr's translation does not cover all the contents of the original work, and the number of tokens is significantly fewer than in Lockhart's translation.

In the 19th century, Western missionaries in China, besides spreading Christianity, aimed to explore TCM to complement Western medicine. Fan (2017) noted that missionaries studied and translated Chinese medicine to address the shortage of Western medicine supplies for practicing medicine in China [21]. Lockhart's extensive time in China led him to actively translate numerous Chinese medical works. Kerr's translation, on the other hand, reflects the perspective of some missionaries who believed that Chinese medicine lacked systematic anatomical knowledge, opting to impart medical knowledge based on their experiential understanding. Consequently, negative translations of Chinese medicines and therapeutic approaches were prevalent in their work.

## 4.2. Translation strategy

Lockhart adopts a foreignizing translation strategy, aiming to propagate Chinese TCM culture [8]. His translation approach leans toward faithfulness, advocating for the enrichment of Western medicine through the promotion of Chinese medical culture. For instance, consider Lockhart's suggestions for the puerpera in the original text:

The original text: "只宜闭目静养，勿令熟睡。恐倦极熟睡，血气上壅，因而眩晕。"

Lockhart's translation: She ought to shut her eyes, and be kept quiet; but she must not sleep soundly, for should she be very weary and sleep soundly, the blood and heart will flow upwards and occasion dizziness of the head.

The language is simple and easily comprehensible, effectively conveying the principles of Chinese medicine. The translation incorporates appropriate conjunctions into the original clauses, enhancing sentence coherence and facilitating acceptance by target language readers. Lockhart remains faithful to the original text while considering the needs of the translated language readers. Given its status as a widely circulated book on obstetrics, the original text is easily understood, featuring many short sentences. Lockhart's high STTR better reflects the characteristics of the original short sentences, making the translation more readable.

Kerr's translation of *Da Sheng Bian* stems from his interest in Chinese obstetrics and his critique of the outdated obstetrics concepts within the work. Consequently, Kerr's translation is notably shorter than the original text, omitting certain content, such as the two chapters on "Maxim" and "Prescription." Despite this, its linguistic features align with the reading habits of target language readers overall. For instance, in translating the four-word idiom "揠苗助长" (*yamiaozhuzhang*) in the sentence "观'揠苗助长'四字，即知将试痛认作正生之弊矣", Kerr rendered it as "pulling up grain to make it grow," enhancing readability in the target language. In the translation of medicinal prescription names, Kerr combined literal translation and transliterations, such as translating "加味芎归汤" (*jiaweixiongguitang*) as "the anodyne medicine" and "回生丹" (*huishengdan*) as "the restoring life powders." This approach also contributes to the promotion of TCM and Chinese culture.

# 5. Conclusion

In conclusion, this paper has undertaken a comprehensive comparison of the translation styles employed by Lockhart and Kerr in their respective translations of *Da Sheng Bian*, examining lexical, syntactic, and discourse levels. The analysis reveals that Lockhart's translation exhibits meticulous attention to detail, aligning with a foreignizing translation strategy. In this approach, nuances from the original text are explicitly conveyed to the readers in the translated language, employing explicit strategies. On the other hand, Kerr's translation is more succinct, adopting a freer translation approach.

Despite both translators operating in the same era, their distinct translation purposes result in divergent presentations. In the context of the big data era, this paper utilizes corpus analysis to compare the English translations of *Da Sheng Bian*. This objective is to capture the interest of experts and scholars in the field, fostering advancements in the translation of Chinese medical texts and promoting the dissemination of Chinese medicine worldwide.

# Funding

# Disclosure statement

The authors declare no conflict of interest.

# References

[1] Li Y, 2009, Examination of the Authorship of the Famous Obstetrics Book *Da Sheng Bian*. Traditional Chinese Medicine Culture, 4(4): 31–32.

[2] Lockhart W, 1842, A Treatise on Midwifery. The Dublin Journal of Medical Science, 20: 333–369.

[3] Kerr JG, 1894, Tat Shang Pin. Annals of Gynaecology and Paediatry, 7(6): 326–333.

[4] Kerr JG, 1894, Midwifery Made Easy. The American Gynecological & Obstetrical Journal, 5(8): 461–468.

[5] Li J, Guo Q, 2016, The Translation and Dissemination of *Da Sheng Bian* and the Evaluation of Chinese Obstetrics. Journal of Guangzhou University of Traditional Chinese Medicine, 33(6): 896–898.

[6] Yong W, Zhu Y, The First English Translation of Traditional Chinese medicine *Da Sheng Bian*. Chinese Science and Technology Translation, 35(1): 52–55.

[7] Cheng H, 2022, Distorted Evidence: Research on J. G. Kerr's Translation of *Da Sheng Bian*, thesis, Guangxi University.

[8] Hu K, Xie L, 2017, Towards a Corpus-Based Study of Translator's Style. Chinese Translators Journal, 38(2): 12–18 + 128.

[9] Xu M, Pei W, 2022, A Review of Corpus-Based Studies on the English Translation of Chinese Sci-Tech Classics: The State, Significance, and Prospects. Foreign Languages and Their Teaching, 326(5): 116–124 + 149.

[10] Baker M, 2001, Towards a Methodology for Investigating the Style of a Literary Translator. Target International Journal of Translation Studies, 12(2): 241–266. https://doi.org/10.1075/target.12.2.04bak

[11] Pan H, 2022, Research on Construction of English Corpus of Suwen: Great Theory on Yin-Yang Corresponding to Nature. Journal of Basic Chinese Medicine, 28(4): 630–634 + 666.

[12] Hu K, 2017, Corpus Translation. Shanghai Jiao Tong University Press, Shanghai, 230.

[13] Wang K, 2003, Sentence Parallelism in English-Chinese/Chinese-English: A Corpus-Based Investigation. Foreign Language Teaching and Research, 35(6): 410–416.

[14] Butler C, 1985, Studies in Linguistics. Longman Group Limited, London, 121.

[15] Halliday MAK, Hasan R, 1976, Cohesion in English. Longman Group Limited, London.

[16] Sun X, 1996, "Cohesion" and "Coherence" in English Discourse Analysis. Foreign Language Research, 19(1): 25.

[17] Ma Y, Sun H, Jia Q, 2020, Corpus-Based Contrastive Study of Conjunctions in Chinese Classic Translation. Shanghai Journal of Translators, 155(6): 40–45.

[18] Quirk R, Greenbaum S, Leech G, et al, 1985, A Comprehensive Grammar of the English Language. Longman Group Limited, London.

[19] Huang G, 2012, "Paraphrase" as a Strategy in Translating *Lun Yu* (*The Analects*) into English: A Functional Linguistics Perspective. Journal of University of Science and Technology Beijing (Social Sciences Edition), 3: 16–21.

[20] Ju Y, 2018, A Corpus-Based Comparative Analysis of Sentence Connectors in Two English Translations of *The Analects*. Foreign Language and Literature Research, 4(1): 59–72.

[21] Fan Y, 2017, Study on the Activities of TCM Translation and Its Influence. Suzhou University Press, Suzhou.