Vol.4 No.3 2024 Philosophy & Ideology Research
https://doi.org/10.37420/j.pir.2024.027

# A Corpus-Based Study of Translators' Style——Taking the Chinese Translations of the Three Volumes of His Dark Materials

**Chaoqun Xiong** [1*]
[1] **School of International Education, Jiangxi Normal University, Nanchang, China**
[*]**Corresponding author,E-mail:Laurencebarnes79@gmail.com**

## Abstract
*This study examines the stylistic differences demonstrated by three different translators in their respective translations of the three parts of Philip Pullman's His Dark Materials Trilogy. By constructing a corpus and utilizing Python for data processing and quantitative analysis, the study conducts a comparative analysis of the three Chinese translations at both the lexical and syntactic levels. The results reveal that each translator exhibits distinct levels of lexical richness and syntactic complexity in their respective translations, reflecting their unique translation styles.*

## Keywords
*His Dark Materials Trilogy, Translators' style, Corpus analysis*

## 1.Introduction

Since the mid-20th century, Western fantasy literature has experienced rapid development, producing numerous outstanding works, with The Lord of the Rings and Harry Potter as representative examples. Entering the 21st century, advancements in cinematography and special effects technology have made the reconstruction of the fictional worlds in fantasy novels a reality. Both The Lord of the Rings and Harry Potter were adapted into films during this period, garnering widespread acclaim. Beginning in 2011, the American television series Game of Thrones, adapted from the fantasy novel series A Song of Ice and Fire, achieved unprecedented global success, drawing increasing attention to Western fantasy literature. Consequently, many outstanding works have been translated into Chinese and introduced to China. Among them is His Dark Materials, a celebrated work by the renowned British fantasy author Philip Pullman. The trilogy has received widespread praise and won numerous awards, including the Carnegie Medal.

The Chinese translation of His Dark Materials comprises three volumes, each translated by a different translator. Different translators bring their own habits and preferences to the translation process, which shape their individual styles. While translators are generally expected to remain faithful to the original text and preserve its essence for readers, their styles can influence their translation strategies and practices, resulting in distinct effects in different translations of the same source text. Since each volume of His Dark Materials was translated by a different translator, this study uses the trilogy as research subject to analyze and compare the three

© 2024 by the author(s); licensee Mason Publish Group, this work for open access publication is under the Creative Commons Attribution International License (CC BY 4.0). (http://creativecommons.org/licenses/by/4.0/)

volumes at the lexical, syntactic, and textual levels through corpus-based methods and quantitative analysis. By examining the translators' styles, this research aims to contribute to corpus-based translation studies and provide insights into the translation of Western fantasy literature.

## 2.Research Status

### *2.1. Current Status of Translation Studies on Western Fantasy Literature*

A search for “fantasy novel translation” on China National Knowledge Infrastructure (CNKI) reveals that most relevant studies were published around a decade ago. While a considerable number of articles exist, they predominantly focus on Western fantasy classics such as A Song of Ice and Fire and The Lord of the Rings, differing mainly in perspectives and theoretical frameworks. For instance, Chen Jiakun and Liu Yunhong used the Chinese translation of The Lord of the Rings as a case study to explore the relationship between retranslations and the generative process of textual vitality, analyzing aspects such as the expansion of translational space, construction and refinement of translation strategies, and translators' value orientations[1]. Similarly, Wu Shuang (2014) analyzed the Chinese translation of A Game of Thrones, the first book in A Song of Ice and Fire, from the perspective of Skopos theory[2].

Few studies have focused on other high-quality Western fantasy literature. In terms of classification, the aforementioned works belong to the subgenre of high fantasy, characterized by expansive world-building and intricate character relationships, often requiring substantial effort from readers to engage with the texts.

### *2.2. Current Status of Corpus-Based Translator Style Studies*

Hermans argued that “a second voice always exists in translations,” sparking extensive discussions in the field about translator style[3]. Baker defined translator style as a set of linguistic or non-linguistic features that distinguish one translator from others[4]. Since 2005, corpus-based studies on translator style have gained significant attention in Chinese academia, becoming a major research focus in translation studies. With the increasing recognition of the translator’s subjectivity, researchers argue that corpus methods provide precise descriptions of translator styles across lexical, syntactic, and textual levels using scientific data.

Many corpus-based studies on translator style focus on English-to-Chinese translations. For example, Liu Zequan et al. used a parallel corpus of Dream of the Red Chamber in Chinese and English to compare and explore stylistic similarities and differences among four English translations[5]. Huang Libo and Zhu Zhiyu analyzed Howard Goldblatt’s translator style in his translations of modern Chinese novels, using Dai Naidie's translated works as a reference point[6]. These studies have significantly contributed to promoting the international dissemination of Chinese cultural works.

In contrast, there are far fewer corpus-based studies on translator styles in English-to-Chinese translations, and existing research tends to focus on classic Western literary works, such as Ulysses, 1984, and Shakespearean dramas. Studies concentrating on Western fantasy literature, which has a broader readership, remain rare.

© 2024 by the author(s); licensee Mason Publish Group, this work for open access publication is under the Creative Commons Attribution International License (CC BY 4.0). (http://creativecommons.org/licenses/by/4.0/)

### *2.3. Summary*

A review of the current status of translation studies on Western fantasy literature reveals that research predominantly focuses on works such as A Song of Ice and Fire and The Lord of the Rings, while other works receive relatively less attention. Meanwhile, although corpus-based translator style studies have made notable progress in the English-to-Chinese field, their application in the translation of Western fantasy literature remains limited.

## 3.Research Design

### *3.1. Research Corpus*

For long novels, it is common practice to divide the work into multiple volumes for publication due to the sequential nature of its composition. However, to maintain a consistent overall style in the translation, it is customary for the same translator to handle all volumes. This study uses the English original text of His Dark Materials and its three Chinese translations as the research corpus, with specific details shown in the table below. Notably, each volume of His Dark Materials was translated by a different translator, a unique aspect that lends special value to this research. For details on the corpus, see Table 1

*Table 1: Corpus Information*

| Source | Author/Translator | Language |
|---|---|---|
| His Dark Materials | Philip Pullman | English |
| The Northern Lights | Zhou Jingxing | Simplified Chinese |
| The Subtle Knife | Zhou Qian | Simplified Chinese |
| The Amber Spyglass | Chen Junqun | Simplified Chinese |

### *3.2. Research Questions*

This study aims to address the following research questions:

1. Do the three Chinese translations of His Dark Materials align stylistically with the original text?

2. What are the similarities and differences in style among the three Chinese translations of His Dark Materials?

### *3.3. Research Methodology*

This study employs Python to perform preprocessing tasks on the English and Chinese corpora, including noise reduction, cleaning, and tokenization. Subsequently, using corpus-based research methods, the study analyzes the three Chinese translations of His Dark Materials from the perspectives of type-token ratio, word length, lexical density, and average sentence/paragraph length. Through these analyses, the study explores the distinct stylistic features of the translators of the three volumes.

© 2024 by the author(s); licensee Mason Publish Group, this work for open access publication is under the Creative Commons Attribution International License (CC BY 4.0). (http://creativecommons.org/licenses/by/4.0/)

# 4.Research Results

## *4.1. Lexical Richness*

Type-token ratio (TTR) is a measure of lexical variation, where "type" refers to distinct words in a text (excluding duplicates and case variations) and "token" refers to all word occurrences[7]. Huang Libo suggested that this ratio serves as a key indicator of lexical diversity in a text[8]. Using Python and the Jieba library, this study tokenized each corpus, calculated the overall TTR, and then divided each corpus into 1000-token segments to compute the standardized TTR (STTR) as the mean of these values.

The results are shown in table 2. Among the three translations, Zhou Qian's translation has the lowest STTR at 39.19, indicating a smaller vocabulary, lower lexical richness, and less varied word usage. Zhou Jingxing's and Chen Junqun's translations exhibit similar levels of lexical richness, closer to that of the English original. However, the original English text has the highest STTR among all four corpora, suggesting a significant gap in lexical richness between the translations and the original text.

*Table 2: Lexical Richness*

| Source | Philip Pullman | Zhou Jingxing | Zhou Qian | Chen Junqun |
|---|---|---|---|---|
| Tokens | 369803 | 130538 | 102991 | 160542 |
| Types | 17977 | 10417 | 8356 | 10938 |
| TTR | 4.87 | 7.98 | 8.11 | 6.81 |
| STTR | 42.37 | 41.18 | 39.19 | 41.42 |

## *4.2. Lexical Density*

Lexical density, defined as the percentage of content words (nouns, verbs, adjectives, and adverbs) relative to the total number of words, reflects the informational content of a text (Baker, 1995)[9]. For the English text, this study used the NLTK library for part-of-speech tagging, which performs well with English corpora. For the Chinese texts, the Jieba library was used, as it is optimized for Chinese natural language processing. Content words in Chinese included nouns, verbs, adjectives, and numerals.

The results are summarized Table 3. As with lexical variation, the lexical density of the English original is the highest, while Zhou Qian's translation is the lowest. Zhou Jingxing's and Chen Junqun's translations show similar lexical densities, with Chen Junqun's slightly higher. All three translations have lower lexical densities than the original English text.

*Table 3: Lexical Density*

| Source | Philip Pullman | Zhou Jingxing | Zhou Qian | Chen Junqun |
|---|---|---|---|---|
| Noun | 78617 | 3641 | 2590 | 5269 |
| Verb | 74001 | 4037 | 3263 | 5315 |
| Adjective | 20703 | 25014 | 18894 | 29433 |
| Number & adverb | 23970 | 27049 | 21434 | 33881 |
| Percentage | 53.35% | 50.10% | 48.88% | 50.51 |

© 2024 by the author(s); licensee Mason Publish Group, this work for open access publication is under the Creative Commons Attribution International License (CC BY 4.0). (http://creativecommons.org/licenses/by/4.0/)

### *4.3. Mean Sentence Length*

Similar to TTR, mean sentence length is a marker of translator style [10]. Longer sentences with fewer total sentences indicate greater syntactic complexity and a higher degree of sentence maturity. For the English text, sentence segmentation was performed using the NLTK library. For the Chinese translations, sentences were segmented using regular expressions based on punctuation marks.

Table 4 shows that Zhou Jingxing's and Zhou Qian's translations have similar mean sentence lengths, suggesting comparable sentence complexity. Chen Junqun's translation, however, has a significantly higher mean sentence length, indicating more complex and challenging sentence structures. Compared to the original English text, Zhou Jingxing's and Zhou Qian's translations are closer in mean sentence length, reflecting efforts to preserve the syntactic style of the original.

*Table 4: Mean Sentence Length*

| Source | Philip Pullman | Zhou Jingxing | Zhou Qian | Chen Junqun |
|---|---|---|---|---|
| Sentence number | 25210 | 8801 | 6673 | 8672 |
| Mean Sentence length | 17.3 | 17 | 17.6 | 20.73 |

### *4.3. Mean Segment Length*

According to Zhu Yifan, segments are sub-sentential units separated by commas, semicolons, and other punctuation marks. The mean segment length is calculated as the total number of tokens divided by the number of punctuation marks that denote segment boundaries[11]. Qin Hongwu et al. argue that segment length better captures the characteristics of Chinese texts, which are more paratactic[12].

The results indicate that Chen Junqun's translation has the highest mean segment length, suggesting the greatest information density and the highest difficulty of comprehension. Zhou Jingxing's and Zhou Qian's translations show similar mean segment lengths, aligning with their sentence length results.

*Table 4: Mean Segment Length*

| Source | Zhou Jingxing | Zhou Qian | Chen Junqun |
|---|---|---|---|
| Sentence number | 20952 | 16568 | 23890 |
| Mean Sentence length | 5.69 | 5.7 | 6.12 |

## 5.Conclusion

This study conducted a quantitative analysis of the three Chinese translations of Philip Pullman’s His Dark Materials trilogy to explore stylistic differences among the translators. The findings reveal notable discrepancies between the translations and the English original in terms of lexical variation and lexical density. Zhou Qian's translation exhibits the lowest lexical richness, suggesting a relatively limited and repetitive vocabulary, while Zhou Jingxing’s and Chen Junqun’s translations are closer to the original in terms of lexical diversity.

© 2024 by the author(s); licensee Mason Publish Group, this work for open access publication is under the Creative Commons Attribution International License (CC BY 4.0). (http://creativecommons.org/licenses/by/4.0/)

Analysis of mean sentence length and mean segment length shows that Chen Junqun's translation features significantly longer sentences and segments compared to the other two translations, indicating a more complex syntactic style and higher information density. Zhou Jingxing's and Zhou Qian's translations, on the other hand, align more closely with the sentence lengths of the original text, reflecting a more faithful translation style.

A combined analysis of lexical and syntactic features reveals differences in translation strategies among the three translators. Zhou Qian's translation leans toward simplicity and clarity, whereas Chen Junqun's version is more detailed and intricate. Zhou Jingxing strikes a balance, with lexical richness and sentence lengths that most closely mirror the original text, presenting a more balanced translation approach.

However, this study has certain limitations. It primarily focuses on quantitative analysis of vocabulary and syntax, potentially overlooking the impact of contextual and cultural factors on the translators' styles. Translators may employ different strategies when handling culturally specific content, metaphors, and idioms, aspects that this study does not explore in depth. Additionally, translation is inherently subjective, influenced by the translators' personal backgrounds, academic training, and professional experience. While the quantitative analysis sheds light on stylistic differences, it cannot fully account for the subjective reasons behind these variations.

## References

[1] Chen, J., & Liu, Y. (2019). Re-translation and the generation of new life in texts: A case study of The Lord of the Rings Chinese translation. Foreign Language Teaching, 03, 80-85. https://doi.org/10.16362/j.cnki.cn61-1023/h.2019.03.014.

[2] Wu, S. (2014). A purposive approach to evaluating the Chinese translation of A Song of Ice and Fire: Game of Thrones (Master's thesis, Beijing Foreign Studies University).https://kns.cnki.net/kcms2/article/abstract?v=m4MC-GHErBuNSswH6KgEjUeRuigfIF5R3IGAZL4PY5coLU9voIV4fc4wyk0rCV0AnZflGBu9x6PJrYVQFzCkSyGbhMC0gW8j_lony7AKSbVV6OdoDK6mOqgdv8ZbL6zFX5WPoPYLPl1MOOaoXLix-pPJGH_eIFWx8wyjQYcbRqlbhrMhktVel6yiP5vEtSG-wVm6JCzLX8M=&uniplatform=NZKPT&language=CHS

[3] Hermans, T. (1996). The translator's voice in translated narrative. Target: International Journal of Translation Studies, 8, 23-48.

[4] Baker, M. (2000). Towards a methodology for investigating the style of a literary translator. Target: International Journal of Translation Studies, 2, 241-266.

[5] Liu, Z., Liu, C., & Zhu, H. (2011). A preliminary study of translator style in four English translations of Dream of the Red Chamber: Corpus-based statistics and analysis. Chinese Translators Journal, 01, 60-64.

[6] Huang, L., & Zhu, Z. (2012). Corpus-based study of translator style: A case study of David Hawkes' English translations of contemporary Chinese novels. Foreign Language Research, 05, 64-71. https://doi.org/10.13978/j.cnki.wyyj.2012.05.002.

[7][9] Baker, M. (1995). Corpora in translation studies: An overview and some suggestions for future research. Target: International Journal of Translation Studies, 7, 223-243.

© 2024 by the author(s); licensee Mason Publish Group, this work for open access publication is under the Creative Commons Attribution International License (CC BY 4.0). (http://creativecommons.org/licenses/by/4.0/)

[8] Huang, L. (2021). Corpus-based translation studies theory research. Foreign Language Teaching and Research Press.
[10] Olohan, M. (2004). Introducing corpora in translation studies.
[11] Hu, X. (2007). A corpus-based study of lexical features in Chinese translated fiction. Foreign Language Teaching and Research, 03, 214-220 + 241.
[12] Zhu, Y. (2016). A review of research on linguistic features in Chinese translation studies. Contemporary Foreign Language Studies, 06, 40-46 + 61 + 109.

© 2024 by the author(s); licensee Mason Publish Group, this work for open access publication is under the Creative Commons Attribution International License (CC BY 4.0). (http://creativecommons.org/licenses/by/4.0/)