| RESEARCH ARTICLE

A Corpus-Based Study on China English in the English Translation of *Tao Te Ching*

Jiao Jiaqi

Postgraduate student, School of Foreign Languages, Ocean University of China, Qingdao, China; Postgraduate student, Faculty of Arts and Humanities, University of Macau, Macau, China

Corresponding Author: Jiao Jiaqi, **E-mail**: jiaojiaqi1005@163.com

| ABSTRACT

This study aims to examine the features of China English in the translation of Chinese classics by comparing two versions of *Tao Te Ching* based on corpus data. Of the two English versions, one was translated by a well-known Chinese translator—Xu Yuanchong, and the other was translated by an American sinologist—Arthur Waley. This study found that Xu's translation indicates more features of China English compared with Waley's translation according to three major aspects. First, Xu's translation is more concise, employing fewer words to translate *Tao Te Ching*. Second, Xu's version features fewer clauses and more clear sentences. Third, the paratactic nature of China English is reflected in Xu's translation, which has more content words and less cohesiveness. This study reveals the characteristics of China English in translation texts and partly fills the research gaps regarding the quantitative research in this field.

| KEYWORDS

China English, translation of *Tao Te Ching*, corpus-based translation study

| ARTICLE DOI: 10.32996/ijllt.2022.5.1.8

1. Introduction

In the context of English as a Lingua Franca, China English represents a linguistic identity that allows Chinese people to maintain their cultural identity in an English-speaking society. Because ancient Chinese philosophical texts can reflect China's language and culture, their English translations are highly representative and valuable for studying China English. *Tao Te Ching* is a well-known Chinese classic worldwide that has been translated into English around 200 times. Most translations of *Tao Te Ching* were completed by native English speakers who were concerned about English readers and attempted to make their works more understandable, which means their translations were more idiomatic, and the prosody of the original Chinese version may have been lost. In comparison, Chinese translators would pay close attention to the distinctive Chinese culture to reproduce the prosody when translating the Chinese classics. Consequently, this study aims to examine the features of China English in the translation of Chinese classics by comparing two versions of *Tao Te Ching* based on corpus data. Of the two English versions, one was translated by Xu Yuanchong (1921-2021), a well-known Chinese translator, and the other was translated by Arthur Waley, an American sinologist (1888-1966).

2. Literature Review
2.1 China English
English in China has been given a variety of terms that reflect different perspectives on the language. For instance, terms like "Chinglish," "Sinicized English," and "Chinese English" are usually deemed negative from the perspective of semantic prosody. The term "China English" is considered a more objective way of referring to English in China. Many experts have been trying to define China English since the 1990s. According to Xie (1995), China English is an "interference variety used by Chinese in cross-cultural communication, which manifests itself at various levels, including language itself, schema, and culture." Jia and Xiang (1997) define China English as "a variety of English used by Chinese speakers, based on standard English, but with inevitable Chinese characteristics or those that help disseminate the Chinese culture." In this paper, China English refers to various English used by Chinese people that incorporate cultural elements.

2.2 *Tao Te Ching* and its Translation

As a classic Chinese work, *Tao Te Ching* was completed during the Chunqiu Period by Lao Tzu. *Tao Te Ching* has been translated into more than 40 languages, including English, French, Russian, German, and Japanese. It is one of the most widely translated Chinese classics. Arthur Waley, a well-known American sinologist and translator, translated *Tao Te Ching* in 1936. His English version of *Tao Te Ching* was highly received and has been reprinted numerous times, establishing a positive influence in the English-speaking world. Many scholars in China have conducted studies based on Waley's version of *Tao Te Ching* (e.g., Li, 2010; Zhou, 2020), but few studies of Tao Te Ching are based on Xu's English version. Xu has devoted his life to translating classic Chinese classics. His translation of *Chu Ci*, *One Hundred Tang Poems*, and *Analects* has also earned positive reviews in the English-speaking world. He is known for his translation theory of "Three Beauty," which states that while translating Chinese into English, three kinds of "beauty" should be transmitted: the beauty of Chinese form, the beauty of Chinese image, and the beauty of Chinese prosody. His translation of *Tao Te Ching* was published in 2011 under the guidance of the "Three Beauty" theory.

2.3 Corpus-Based Translation Study

Corpus-based translation study appeared in 1993, following *Corpus Linguistics and Translation Studies: Implications and Applications* written by Mona Baker. Since then, the corpus linguistics method has been involved in translation studies. Based on the corpus, many studies focused on topics like universal aspects of translation language, translator styles, and translation education (e.g., Baker, 2000; Mundy, 2001; Zanettin, 1998). For example, Baker (2000) examined Peter Clark and Peter Bush's different translation styles from type/token ratio (TTR), average sentence length, and reporting structures based on the self-built corpus data. It was found that Clark's translations had a lower TTR, shorter sentences, and higher use of reporting verbs. The analysis method in Baker's (2000) study will be adopted in this study. In addition, this study will introduce more data to examine two translation works, one from an English native speaker and the other from a Chinese speaker.

3. Research Design

3.1 Research Questions

This study uses Xu Yuanchong's translation of *Tao Te Ching* as a representative work of China English and compares it to the work of English native speakers to determine the features of China English in translated texts. This study aims to address the research question: What are the differences between Xu Yuanchong's and Arthur Waley's version of *Tao Te Ching*?

3.2 Research Methods

The two translated texts of *Tao Te Ching* are compared on three levels: the lexical level, sentence level, and textual level utilizing corpus tools like WordSmith and TreeTagger. At the lexical level, word information such as the number of types and tokens, the standardized type/token ratio (STTR), high-frequency words, word length, and keywords will be measured. This study will mainly measure the number of sentences, the average length of sentences, the number of clauses, and the number of passive voice sentences at the sentence level. The degree of cohesiveness will be measured at the textual level by comparing cohesion strategies such as reference, substitution, and conjunction.

4. Research Results and Data Analysis

4.1 At the Lexical Level

4.1.1 Lexical Richness

Lexical richness indicates the number of different words used in a text, which can be measured by STTR, proportional to lexical richness. The statistic of the lexical richness of the two texts is shown in Table 1.

Table 1: *Statistics of Lexical Richness*

Measures	Waley	Xu
Types	1767	1386
Tokens	9300	6786
STTR	37.97	36.35

Table 1 shows that Waley's text has a higher STTR than Xu's text. Moreover, Xu's text contains far fewer tokens, indicating that Xu's translation is more concise and contains fewer descriptive terms.

4.1.2 Lexical Density

Data on lexical density can be calculated using TreeTagger's Part-of-Speech tagging, which represents the degree of formality in a text. A higher lexical density in a text indicates that it is more formal and difficult to comprehend. The statistics of the lexical density of the two versions are shown in Table 2.

Table 2: *Statistics of Lexical Density*

Measures	Waley	Xu
Nouns	1721	1323
Verbs	1176	852
Adjectives	556	705
Adverbs	519	241
Total number	3971	3121
Lexical density	0.4270	0.4599

It can be demonstrated that Xu's text has a higher lexical density, indicating that it contains more information while having fewer words. It also reflects that China English may utilize fewer words to convey the same notion in comparison to standard English.

4.1.3 High-Frequency Words
WordList function in WordSmith is used to generate a list of high-frequency words, as is shown in Table 3.

Table 3: *Statistics of High-Frequency Words*

Number	Waley		Xu	
	Word	**Freq.**	**Word**	**Freq.**
1	The	603	The	498
2	Is	296	Is	223
3	To	266	To	184
4	Of	233	And	172
5	And	227	Be	151
6	It	183	Not	138
7	*That*	164	A	133
8	Not	143	It	126
9	He	135	Of	102
10	Be	132	*Will*	100
11	In	115	In	89
12	A	107	*Law*	66
13	But	83	*You*	66
14	*Them*	82	*So*	63
15	*Are*	74	*Divine*	62
16	*All*	72	*May*	57
17	*Who*	68	Can	54
18	*As*	67	But	52
19	Can	63	He	52
20	*What*	62	*World*	50

It can be shown that the top ten words in both texts are nearly identical, indicating that some function words are inextricably linked in the translation of *Tao Te Ching*. Between the eleventh and twentieth words, over half of them are owned by only one of the two texts. Moreover, Xu's translation has more content words, while Waley's translation has more function words.

4.1.4 Word Length
The word length can reflect a translator's preferred word choice. Long words will increase the complexity of a text. Table 4 demonstrates the statistics of word lengths in the two texts.

Table 4: *Statistics of World Length*

Word Length	Waley		Xu	
1 letter words	172	1.84%	181	2.67%
2 letter words	1827	19.65%	1299	19.14%
3 letter words	2181	23.45%	1617	23.83%
4 letter words	1852	19.91%	1237	18.23%
5 letter words	1098	11.80%	814	12.00%
6 letter words	756	8.13%	668	9.84%
7 letter words	575	6.18%	395	5.82%
8 letter words	345	3.70%	206	3.04%
9 letter words	255	2.74%	166	2.45%
10 letter words	155	1.67%	118	1.74%
Above 10 letter words	74	0.80%	85	1.25%

The word length of the two texts is nearly identical, indicating that the two texts have few differences in terms of word difficulty.

4.2 At the Sentence Level
4.2.1 Sentence Length
The length of a sentence reflects the complexity of the sentence to some extent. The longer the average length of a sentence is, the more complicated the syntactic structure of the text is, and the more difficult the sentence is to understand. The standardized deviation of sentences infers the consistency of sentence length throughout the text. The statistics of the sentence length of the texts are shown in Table 5.

Table 5: *Statistics of Sentence Length*

Measure	Waley	Xu
Total number of sentences	527	534
The average length of sentences	17.65	12.71
Standardized deviation of sentences	9.88	6.58

While Xu's text comprises five more sentences than Waley's, the average length of each sentence is significantly shorter and more constant, showing that Xu's text is more concise and less variable in length.

4.2.2 Number of Clauses
The number of clauses in a sentence can also reflect the sentence's complexity. As is shown in Table 6, the number of compound sentences in Xu's text is much lower than in Waley's text, indicating that Xu's text has more superficial sentence structures.

Table 6: *Number of Clauses*

Clauses	Waley	Xu
"that" clauses	136	12
"when" clauses	40	35
"if" clauses	37	8
"which" clauses	18	49
Total number	231	104
Proportion	43.83%	19.48%

4.2.3 Number of Passive Voice Sentences (PVS)
Compared to English, Chinese contains fewer passive voice sentences, which is reflected in Xu's translation, as shown in Table 7.

Table 7: *Number of PVS*

Measure	Waley	Xu
Number of PVS	167	150
Total Number of Sentences	527	534
Proportion of PVS	31.69	28.09%

Xu's text includes less PVS than Waley's text, while the proportion of PVS of the two texts is similar. It can be concluded that Xu's text follows the pattern used in English writing.

4.3 At the Textual Level
According to Halliday and Hasan (1976), reference, substitution, and conjunction are fundamental aspects of text cohesiveness. The degree of cohesiveness will be assessed at the textual level. The marked words can be used to test the cohesion-forming strategies of the two texts.

4.3.1 Reference
Reference can be divided into personal, demonstrative, and comparative. Regarding the three categories of references in the two texts, the total number of references used in Xu's text is less than in Waley's text, as is shown in Table 8.

Table 8: *Statistics of Reference*

Personal Reference	Waley	Xu
I	32	33
Me	3	3
You	31	66
He	135	52
Him	12	4
She	0	0
Her	0	2
It	183	126
They	58	35
Them	82	18
We	20	9
Us	2	0
Total	558	348
Demonstrative Reference		
This	48	10
These	11	3
Those	35	25
Here	1	0
Now	7	0
Then	16	7
Total	118	46
Comparative Reference		
Same	4	1
Different	3	2
Similar	0	0
Otherwise	1	1
The Comparative Degree	19	25
The Superlative Degree	41	7
Total	68	36

4.3.2 Substitution

As is shown in Table 9, the most common use of substitution word in both texts is "one," indicating that nouns are substituted the most in both texts, although the overall number of substitutions in Xu's text is lower than Waley's text, like the number of references.

Table 9: *Statistics of Substitution Words*

Substitution Words	Waley	Xu
One	59	36
Ones	1	0
Do	3	4
Doing	1	4
So	15	5
Not so	0	0
Total	79	49

4.2.3 Conjunction

Table 10 reveals that Xu's text has lower conjunction than Waley's text, and Xu's text also has a lower overall number of cohesions.

Table 10: *Statistics of Conjunction and Cohesion Words*

Conjunction Words	Waley	Xu
And	227	171
But	83	52
Or	16	26
Because	22	2
Though	13	3
Total	361	254
Cohesion		
Reference	744	430
Substitution	79	49
Conjunction	361	254
Total	1184	733

5. Conclusion

Previous research on China English has centered on influencing factors from the political, economic, and cultural perspectives. This study takes a more micro approach, attempting to describe the characteristics of China English through translation and providing data from authentic texts for the study of China English.

The data analysis reveals that Xu's translation differs significantly from Waley's translation. Xu's translation exhibits the following features as a typical work of China English: (1) When compared to a native speaker's translation, Xu's translation is more concise, employing fewer words to translate *Tao Te Ching*; (2) Xu's version features fewer clauses and more clear sentences; (3) The paratactic nature of China English is reflected in Xu's translation, which has more content words and less cohesiveness. This study also has its limitations as a brief corpus-based investigation. Qualitative analysis may be used in the subsequent investigation. Moreover, feedback from the readers on the rating platform can be analyzed from the Appraisal Theory (Martin & White, 2005) for further study.

Conflicts of Interest: The author declares no conflict of interest.

References

[1] Baker, M. (2000). Towards a methodology for investigating the style of a literary translator. *Target*, *12*(2), 241–266. https://doi.org/10.1075/target.12.2.04bak

[2] Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. Longman.

[3] Jia, G., & Xiang, M. (1997). In defence of China English. *Foreign languages and their teaching*, *101*(5), 11–12.

[4] Li, Y. (2010). *A comparative study on three English versions of Tao Te Ching—From the perspective of post-colonial feminism* (Master's dissertation). Soochow University.

[5] Martin, J. R., & White, P. R. R. (2005). *The language of evaluation: Appraisal in English*. Palgrave Macmillan. https://doi.org/10.1057/9780230511910

[6] Munday, J. (2001). *Introducing translation studies: Theories and applications*. Routledge.

[7] Xie, Z. (1995). China English: Interference variety in cross-culture communication. *Modern Foreign Languages*, *70*(4), 7–11.

[8] Zanettin, F. (2002). Bilingual comparable corpora and the training of translators. *Meta*, *43*(4), 616–630. https://doi.org/10.7202/004638ar

[9] Zhou, L. (2020). *Comparing Arthur Waley and D. C. Lau's English translations of Tao Te Ching: Centered on culture-loaded words* (Master's dissertation). Harbin Institute of Technology.