

# A Corpus-based Study on Translation Norms of Prose: Taking Zhu Ziqing's *The View of My Father's Back* as an Example

Ya Zhao

Xi'an Polytechnic University, Xi'an, China

---

**Abstract:** This paper builds a small-scale corpus consisting of Zhu Ziqing's original work *The View of My Father's Back* and its English translations by Zhang Peiji, Yang Xianyi and Gladys Yang. Through the corpus, the lexical and syntactic features of prose are studied to explore the translation norms of prose. The quantitative and qualitative analysis of Zhang's and Yang's translations shows a tendency towards conventionalized expectations at the levels of words and sentences. This study hopes to provide translators with helpful insights on translating Chinese prose and promote overseas dissemination of Chinese prose works and cultural communication.

**Keywords:** Corpus, Translation norms, Prose, *The View of Father's Back*.

---

## 1. Introduction

Corpus-based translation studies refer to the systematic analysis of the essence and process of translation by observing and statistically analyzing a large number of translation facts or phenomena using corpus methods. As the quintessence of Chinese literature, Chinese prose has a history of thousands of years, containing profound Chinese cultural implications. The profound connotations, refined language, and lasting appeal of Chinese prose make it not only translatable, but also worth translating. In this paper, a small-scale corpus is built to use AntConc 4.0.11 to analyze statistics of Zhu Ziqing's prose *The View of My Father's Back* and its two English translations by Zhang Peiji and by Yang Xianyi and Gladys Yang. By comparing type/token ratios, high frequency words, and average sentence lengths between the two translations, the syntactic features of the translations are analyzed to summarize the translation norms of prose.

A CNKI search on "Corpus" and *The View of My Father's Back* returns 3 papers applying corpus methods in translation studies, but all focus on the translator's style instead of translation norms. In *The View of My Father's Back*, Zhu Ziqing uses plain language to profoundly yet delicately express a father's love for his children through ordinary events, sincerely touching readers. The corpus-based study of translation norms will help translators explore norms in prose translation, promoting overseas dissemination of Chinese prose works and cultural communication.

## 2. Translation Norms

A norm is a sociological concept referring to social constraints, including moral and ideological requirements and standards in a society and culture, as well as universally accepted behaviors. Since the 1960s, translation scholars like Toury have studied translation norms from a sociological perspective, arguing that translation is constrained by norms, influenced not only by the source text and culture but also by the target language and culture. Afterwards, Chesterman et al. also began to study the concept of translation norms, which has become an important topic in contemporary translation studies.

Based on this, Toury (1995) divides translation norms into three categories: initial norms, which determine the overall selection of the translated text, requiring translators to decide the text type based on considerations from the source and target cultures and societies, including political, cultural, economic factors, to analyze which works are worth translating; preliminary norms that constrain the overall orientation of the translation, requiring translators to make overall choices and determine translation orientations, i.e. choose between source language/culture norms versus target language/culture norms. Adhering to source norms and representing source language/culture patterns is adequacy translation; conforming to target language/culture norms and patterns is acceptability translation; operational norms guide translators in making a series of decisions during the actual translation process, affecting the composition of the text, involving specific translation strategies from lexis to syntax and textual levels. Chesterman (1997), Nord (1991) and others have also proposed expectancy norms, professional norms, text-linguistic norms, etc. Based on these scholars' definitions, Xu Xianyao summarizes translation norms as two types: external norms and internal norms. External norms include expectancy norms and selection norms, referring to translation policies, general expectations imposed on translations, and how social groups "select" translated texts and source languages. These two norms do not involve concrete translation operations. Internal norms include operational norms that directly relate to specific translation strategies and process.

Expectancy norms include "foreignization expectations" and "domestication expectations". Foreignization expectations refer to readers' expectation for foreign elements and exotic flavors in the translation to bring novelty to the target language. Domestication expectations refer to conforming to target language conventions and cultural traditions, at least linguistically, to be easily understood, since "comprehensibility" is a prerequisite for translation. The coexistence of these two expectations in contemporary society indicates that operational norms must also follow certain patterns. According to Baker et al.'s "universals of translation" and contrastive linguists like Lian Shuneng

(2010), four operational norms regarding contemporary prose translation are hypothesized: (1) sense-for-sense or word-for-word; (2) simplification or elaboration; (3) explicitness or implicitness; (4) conventionality or creativity. If the translated text shows sense-for-sense, simplification, explicitness or conventionality, it reflects domestication expectations; otherwise, foreignization expectations.

Operational norms include lexical norms and syntactic norms. Corpus methods will be used to examine these norms at different levels and verify the hypothesized operational norms.

### 3. Building a Parallel Corpus

Corpus-based translation studies refer to the systematic analysis of the essence, process and phenomena of translation based on corpora as materials, using statistical and theoretical analysis, and drawing on theories in linguistics, literature, culture and translation studies. In this paper, a small parallel corpus is built with AntConc 4.0.11 to analyze statistics of Zhu Ziqing's *The View of My Father's Back*, Zhang Peiji's translation, and Yang & Yang's translation. By comparing type/token ratios, high frequency words, and average sentence lengths between the two translations, syntactic features are analyzed to summarize prose translation norms.

#### 3.1. Research Content

*The View of My Father's Back* is a 1925 autobiographical prose by Zhu Ziqing. It consists of several small stories: the father and son returning home for a funeral, the father worrying about his son traveling alone and seeing him off at the train station, the father using unflattering language to hire someone to carry the son's excessive luggage, the father carefully selecting a train seat for his son, and the father climbing over platforms to buy oranges for his son. The depiction of the father's back view climbing up and down the platform to buy oranges for his son is most profound. Zhu expresses the father's caring and thoughtfulness towards his son, and the son's respect and reluctance to part with his father through the subtle, commonplace event of seeing each other off at the train station. Therefore, translators should focus on grasping the theme and background when translating. This study compares Zhang Peiji's English translation of *The View of My Father's Back* versus Yang & Yang's, analyzing lexical and syntactic differences to study translation norms. Zhang's version is referred to as Zhang's translation, and Yang & Yang's version as Yang's translation.

#### 3.2. Research Tools and Methods

This study uses Wordsmith Tools 8.0, online alignment tool Tmxmall, and AntConc 4.0.11 for quantitative analysis at the lexical and syntactic levels to obtain statistical data for the two translations. By analyzing and comparing the data, the translation norms followed in each translation are identified, providing insights on compiling a bilingual corpus of Chinese prose and improving existing English translations of modern Chinese prose.

### 3.3. Research Process

This study mainly includes four aspects: corpus collection, noise reduction, alignment, and statistical analysis and application.

#### 3.3.1. Corpus Collection

For corpus collection, Zhu Ziqing's original text, Zhang's translation and Yang's translation were purchased and downloaded from online sources. To obtain editable DOCX and TXT files for the next noise reduction step, the PDF texts were converted into DOCX and TXT formats using ABBYY software.

#### 3.3.2. Noise Reduction

Since the original text was short and clean, not much noise reduction was needed. To fully clean the texts, mainly empty lines and indentation containing spaces or tabs, tags in brackets, and annotation information were removed. Any Chinese content in the English corpus was also deleted. The Text Toolkit's clean whitespace and paragraphs functions cleaned spacing and paragraphs. Punctuation was consolidated using the clean punctuation function to fix any fullwidth/halfwidth issues. After these steps, the texts were further examined for any remaining issues before obtaining the final cleaned Zhu Ziqing original, Zhang's translation, and Yang's translation for research. Meticulous corpus cleaning ensures high-quality data for later statistical analysis and reduces errors.

#### 3.3.3. Alignment

Corpus alignment matches the correspondence between the source text and target text to build a parallel corpus. By aligning corpora, translated texts can be imported into translation memories to improve efficiency.

This study uses the online alignment tool Tmxmall for sentence alignment. Tmxmall supports importing 36 formats like commonly used docx, xlsx, pdf, txt, and can export to tmx, xlsx, docx, txt. After importing files, Tmxmall automatically aligns texts with consistent paragraphing, requiring little manual work since source and target sentences are mostly aligned. Minor adjustments were made to achieve 100% accuracy before generating the final TMX file for research.

## 4. Results and Discussion

### 4.1. Lexical Level

#### 4.1.1. Type/Token Ratio

Type/token ratio (TTR) measures lexical variety in a text. "Tokens" refer to total words in a text, while "types" are unique words. The same recurring word is counted as one "type". TTR indicates the proportion of unique words in a text, objectively reflecting lexical diversity. Higher TTR means more varied vocabulary usage. Using AntConc 4.0.11, the numbers of types and tokens in both translations are counted. Since the texts are short, the standard TTR is not used. Instead, type/token ratio percentage is used, with results in Table 1.

**Table 1.** Type/Token Ratio

	Zhu Ziqing's Original	Zhang's Translation	Yang's Translation
Tokens	1317	1137	969
Types	512	463	405
Type/Token	42.71	40.72	41.80

Table 1 shows Zhang's translation has 1137 tokens versus Yang's 969 tokens. In types, Zhang's translation has 463 types and Yang's 405 types. "Tokens" represents vocabulary size, indicating Yang's translation uses simpler vocabulary. Moreover, Zhang's and Yang's TTR are 40.72% and 41.80% respectively, with little difference between them. This is related to the small corpus size, but shows Yang's translation has slightly more varied vocabulary, which may make it more difficult to read than Zhang's.

Overall, both translations have significantly lower TTR than the Chinese original prose, indicating translated language has simpler vocabulary compared to original writing.

This result supports the hypothesis that translated language uses simpler vocabulary versus original texts, consistent with Baker et al.'s "simplification" universal of translation.

#### 4.1.2. High Frequency Words

Word frequency refers to how often each word form appears in a text. Frequency lists are a basic statistical measure in corpus analysis. Analyzing high frequency words helps compare lexical diversity between original and translated texts. Using AntConc 4.0.11, the top 15 most frequent words in Zhu Ziqing's original work and both translations are generated, as shown in Table 2.

**Table 2.** High Frequency Words

Rank	Words	Zhu Ziqing's Original	Words	Zhang's Translation	Words	Yang's Translation
1	我	47	the	57	to	47
2	他	37	to	42	the	42
3	的	30	I	36	he	32
4	了	27	and	32	I	32
5	到	11	of	26	and	30
6	去	11	a	24	a	19
7	父亲	11	he	20	his	16
8	着	10	was	20	me	16
9	走	10	me	19	my	16
10	不	8	in	15	was	16
11	又	7	father	14	for	10
12	已	7	for	13	him	10
13	是	7	him	11	of	10
14	看	7	his	11	father	9
15	见	6	on	10	had	9
	要					

Table 2 shows both translations repeatedly use certain words with higher frequencies. This tentatively proves the hypothesis that prose translations reuse limited vocabulary with smaller ranges. Lexically, contemporary translators mainly adopt simplification as an operational norm, reflecting domestication expectations.

#### 4.2. Syntactic Level

Average sentence length refers to the average number of tokens or words per sentence, indicating sentence complexity in a text. Generally, higher average sentence length correlates to more complex syntax and lower readability, and vice versa.

For Chinese, statistics are based on characters, and English on words. The total characters/words in each text excluding extraneous information like author and title are divided by the total number of sentences to obtain the average sentence length. The same method calculates overall average sentence length for each corpus. To clearly demonstrate translators' syntactic operational norms, corpus tools WordSmith Tools 8.0 are used to analyze sentence data in the translations versus original English texts, as shown in Table 3.

Table 3 shows 66, 65, and 59 sentences in Zhu Ziqing's original work, Zhang's translation, and Yang's translation respectively, with average lengths of 19.95, 17.49, and 16.42. The number of sentences is similar between the source text and translations, without excessive splitting or combining, demonstrating that contemporary translators adopt operational norms that conform to target language conventions. The translations have lower average sentence lengths than the original work. But due to the limited length, both translations largely retain the original syntax. Overall, syntactically, the translators mainly adopt simplification and conventionality as operational norms, reflecting domestication expectations.

### 5. Conclusion

This study analyzes type/token ratio, high frequency words, and average sentence length to summarize prose translation norms in Zhu Ziqing's *The View of My Father's Back*. Based on the data, contemporary translators adopt systematic strategies reflecting norms. Overall, lexical and syntactic choices mainly reflect domestication expectations. However, further research is needed on discourse-level prose translation norms, and the small corpus size limits generalizability of the conclusions. Still, the corpus-based study on *The View of My Father's Back* facilitates understanding of its translations and promoting overseas dissemination of Chinese prose and cultural communication.

Translation is an important means of cultural communication. Chinese prose has a history of thousands of years, with profound cultural implications and refined

**Table 3.** Average Sentence Length

	Zhu Ziqing's Original	Zhang's Translation	Yang's Translation
Total chars	1317	1137	969
Sentences	66	65	59
Avg. length	19.95	17.49	16.42

language reflecting Chinese cultural heritage. With growing cultural confidence in China, scholars should help disseminate Chinese prose abroad, seizing opportunities to showcase Chinese prose culture under the “going global” initiative, and promote international cultural exchange through translating more literary works.

## References

- [1] Fu, Jingmin. (2013). Translation norms from a global perspective. *Shanghai Journal of Translators*, (4), 11-15.
- [2] Toury, G. (1995). *Descriptive translation studies and beyond*. Amsterdam: John Benjamins Pub. Co.
- [3] Xu, Xianyao. (2007). A corpus-based study on lexical features in Chinese translation of fiction. *Foreign Language Teaching and Research*, (03), 214-220+241.
- [4] Zhang, Jiguang, & Zhang, Zheng. (2014). A corpus-based study on translation norms of contemporary English essays translated into Chinese. *Foreign Language Teaching Theory and Practice*, (04), 83-91+95.
- [5] Hu, Kaibao, & Mao, Pengfei. (2012). A review of corpus-based translation studies abroad. *Contemporary Linguistics*, 14(04), 380-395+437-438.
- [6] Hu, Kaibao. (2012). Corpus-based translation studies: Connotation and significance. *Journal of Shanghai International Studies University*, 35(05), 59-70.
- [7] Zhang, Jiguang, & Zhang, Lanhui. (2010). A corpus-based study on syntactic operational norms in contemporary prose translation (English-Chinese). *Journal of Xihua Normal University (Philosophy and Social Sciences)*, (04), 60-65.
- [8] Pan, Zhidan. (2009). Plain and ornate, each has its appeal: English translations of Ming and Qing dynasty chuanqi plays (Master's thesis). Soochow University.