Lecture Notes on Language and Literature (2024)
Clausius Scientific Press, Canada

DOI: 10.23977/langl.2024.070424
ISSN 2523-5869 Vol. 7 Num. 4

A Corpus-based Multidimensional Analysis for English Translations of Medicine: A Comparative Study of Translators' Styles

Xiaocun Wu

School of Foreign Studies, Northeastern University at Qinhuangdao, Qinhuangdao, China

Keywords: Multidimensional Analysis, Translator's Style, *Medicine*, Literary Translation

Abstract: The translator's style is a distinguishing feature that sets a translation apart from others and has substantial research worth. Having considered the methodological imperfections of previous related studies, this study applies the Multi-Dimensional Analysis framework to investigate the variations in translator styles. The study's subjects are the two English translations of Lu Xun's novel *Medicine*, which were completed by Yang Hsien-yi, Gladys Yang, and Julia Lovell. According to the results and statistics, both translations adhere to the general narrative exposition in terms of text types, and there is no statistically significant difference between them in terms of their scores on the six dimensions. In terms of linguistic features, Lovell's translation shows significantly higher scores in indicators like AWL, TTR, JJ, PIN, and [WZPAST] than Yang and Dai's. Conversely, for features like ANDC, VBD, and [WZPRES], Yang and Dai's translation significantly outscores Lovell's. Through a comparative analysis of two translations, this research seeks to further corpus-based translation studies and promote Chinese cultural exports.

1. Introduction

Medicine, a short story in Lu Xun's collection Call to Arms, one of the most influential literary works of Chinese novels in the 20th century, is renowned for its revelation of the weaknesses of old society and criticism of the ignorant national character (Hao, 2012) [1]. Its translation into several languages substantially contributed to the spread of Chinese novels in the western literary world. Among the English translations of *Medicine*, both the versions by Lovell, Yang Hsien-yi and Gladys Yang are of great research significance in guiding literary translations and facilitating the dissemination of Chinese culture abroad.

With the advent of corpus translation studies, researchers have concentrated their focus on translator style (Huang L. & Wang, 2011: 916-917) [2]. Perceived as a "thumb-print", a translator's style is expressed through a series of linguistic or non-linguistic features (Baker, 2000: 245) [3]. Baker pioneered the corpus-based study of translators' style, which correlates translators' linguistic habits with their socio-cultural positioning, to break the situation in which the "stylistic correspondence" between the original writing and the translation is the only criterion for evaluating translations. However, as for the existing studies, the research on translator's style for translations of Chinese literary works is still limited, and the notion that translation is a derivative rather than a creative

process limited the number of studies featuring translator styles as the primary study objective (Baker, 2000: 244) [3]. Considering the gap, based on Baker's approach, this study will apply Multi-dimensional Analysis (MDA) proposed by Biber (1988) [4] to explore and compare the translator's styles in two English versions of *Medicine* from a more rigorous and comprehensive perspective. The study focuses on the following three research questions: (1) What are the differences and similarities in the text types and genres that the two translations are close to? (2) What are the differences and similarities in the two translations' scores on six dimensions? (3) What are the differences and similarities in the two translations' scores on specific linguistic features?

Theoretically, the study is based on the corpus translation studies proposed by Baker (1993) [5]. Baker advocates identifying and describing the styles of translators through corpus analysis of a large number of translated texts (*ibid.*), which emphasizes the objectivity and quantitative analysis of data. Methodologically, MDA, the major methodology applied in this study, demonstrates effectiveness and applicability in describing the styles of translations. Through the methodology, a macroscopic examination of translator stylistic distinctiveness is achieved by evaluating specific translators and works in relation to all six dimensions (Biber, 1988: 203) [4].

2. Literature Review

The corpus-based methodology has been employed in previous studies on translator styles in the English translation of Lu Xun's novels. Yan and Han tried to use the corpus research method to compare two English translations of Lu Xun's novels, examining the translation styles and strategies of Yang Hsien-yi and Gladys Yang, and William Lyell from the lexical level and the sentence level respectively (2015) [6]. Li and other researchers used the same source and multiple translation analogical corpus to compare the English translations with those of Wang, Yang and Lyell, mainly to investigate the translator style of Lovell's English translation of Lu Xun's novels (Li D. et al., 2018) [7]. And a pioneering attempt was made to compare Lovell's English translation of Lu Xun's works with her translation of *I Love Dollars: And Other Stories of China* by using the analogical corpus research method of different English translations by the same translator, to further verify her style and to analyze in depth the causes (*ibid.*). Zhu combined quantitative and qualitative methods to analyze the translation styles of the two versions of *Diary of a Madman*, Lyell's and Yang Hsien-yi and Gladys Yang's versions, respectively, from the three levels of vocabulary, syntax, and parts of speech, and adopts richer metrics, such as punctuation and readability, to further investigate the causes of the translator styles in terms of the translators' contemporary background, translation standards and strategies, and translation purposes (2021) [8].

Previous studies have demonstrated their profound research value in multiple aspects, providing valuable perspectives for us to understand the dissemination and acceptance of Lu Xun's works in global cultural contexts. However, the extant research on the English translated versions of *Medicine* or other novels by Lu Xun remains limited, with the majority focusing more on describing the specific appearance of translator styles rather than generalizing them as a whole (Hu & Xie, 2017: 15) [9]. In addition, the measurement indicators are relatively homogeneous, such as STTR, sentence length, average word length, lexical density, and there is a lack of neither correlation between statistical indicators nor combination of quantitative and qualitative dimensions (Zhao, 2020: 68) [10]. The research of translator styles should prioritize the study of global style over the analysis of local style, since translator's style is not equivalent to the mere sum of many local styles (Huang L., 2018: 80) [11]. And the manner of expression that distinguishes a translator rather than isolated instances of open interference should be a primary focus for the study of translator styles (Baker, 2000: 245) [3].

3. Methods

For the simplicity of presentation, Yang Hsien-yi and Gladys Yang is referred to by Yang & Dai. Julia Lovell is denoted by Lovell.

3.1. Corpus Profile

The corpus for this study consists of two English translations of *Medicine* by Lu Xun (1919), which is part of his short story collection *Call to Arms*. The two selected English translations of the novel were respectively completed by Julia Lovell, as well as Yang Hsien-yi and Gladys Yang. Julia Lovell's translation was released in her rendition of *The Real Story of Ah-Q and Other Tales of China, the Complete Fiction of Lu Xun* by Penguin Press in 2009. The translation by Yang Hsien-yi and Gladys Yang was published by Foreign Language Press as part of their translation anthology titled *Call to Arms* in 1981. An overview of the corpus is presented in Table 1. In general, Lovell's translation has 2,857 tokens, and Yang and Dai's translation has 2,967 tokens.

Table 1: Corpus profile

Translator	Chapter	Tokens
Julia Lovell	I	783
	II	440
	III	803
	IV	831
Yang Hsien-yi and Gladys Yang	I	816
	II	464
	III	872
	IV	815

3.2. Multi-Dimensional Analysis

The study employs Multi-Dimensional Analysis (MDA), a quantitative analytical methodology for corpus research introduced by Biber (1988) [4]. Biber extracted 67 linguistic features from extensive corpora comprising 481 texts across 23 registers. Utilizing factor analysis, Biber identifies linguistic features with recurring frequencies and co-occurrences, which are then organized into 7 dimensions based on distinct patterns of co-occurrence (*ibid.*).

As for textual dimensions in MDA, Dimension 1 plays a key role in delineating the contrast between informational and involved discourse production (Biber, 1988: 115) [4]. This dimension functions as a critical parameter for discerning differences in English texts and it illuminates the juxtaposition between discourse geared towards interaction and emotion within temporal constraints, and discourse centered on the meticulous conveyance of information (*ibid.*).

Dimension 2 refers to the distinction between Narrative and Non-narrative concerns (Biber, 1988: 115) [4]. It assesses whether the text leans towards primarily narrative or non-narrative in nature. Narrative concerns typically include significant references to past time, third-person animate referents, and descriptive details. Conversely, non-narrative concerns focus on immediate time and elaboration through attributive nouns in texts like expository or descriptive writing (*ibid.*).

Dimension 3 deals with the contrast between explicit and situation-dependent reference in discourse analysis (Biber, 1988: 115) [4]. This dimension focuses on how referents are identified within a text, either through complete and direct clarification via relativization or through vague allusions to external circumstances for identification purposes (*ibid.*). It represents the distinction between discourse that is independent of context and discourse that relies on context for clarity.

Dimension 4 delves into the overt expression of persuasion in discourse, particularly linked to the speaker's articulation of their own standpoint or the use of argumentative styles to convince the audience (Biber, 1988: 115) [4]. Key features with positive weights in Dimension 4 include infinitives, prediction modals, suasive verbs, conditional subordination, necessity modals, split auxiliaries, and possibility modals, collectively serving as markers of persuasion (Biber, 1988: 111) [4].

Dimension 5 represents a specific facet of informational discourse that encompasses abstract, technical, and formal content (Biber, 1988: 115) [4]. It serves to differentiate between abstract and non-abstract information within texts (*ibid.*). Materials falling under Dimension 5 typically concentrate on conveying information in a technical, abstract, and formal fashion. Key features with positive weights on this dimension include conjuncts, agentless passives, adverbial past participial clauses, fry-passives, and past participial WHIZ deletions (Biber, 1988: 111-112) [4].

Dimension 6 serves to distinguish informational discourses created under real-time conditions from other text types (Biber, 1988: 115) [4]. This dimension is particularly focused on online information elaboration, where information is conveyed in a somewhat loose and fragmented manner due to significant constraints. Although texts categorized under Dimension 6 are inherently informational, they are generated within specific time constraints (*ibid.*).

As a prominent method in corpus linguistic discourse analysis, MDA proves to be a quantitative approach that enables a more systematic and comprehensive examination of the domain linguistic features present in English text, facilitating a more nuanced comparison of translations produced by translators with distinct styles (Zhao, 2020: 68) [10]. Consequently, this study will employ MDA due to its comprehensiveness in addressing the research objectives effectively.

3.3. Research Procedures

This study employed the Multi-Dimensional Analysis Tagger (Nini, 2019) [12] to conduct automated analysis of the translations. More specifically, it followed structured procedures as outlined below:

1) Firstly, the two English translations were saved as text files in the ".txt" format, sorted by chapter. MAT was then used to conduct multidimensional analyses on each translation.

2) Following the analysis, the data of both translations on six dimensions underwent statistical testing. A normality test was conducted, and in cases where the normality of the data distribution was confirmed (p value > 0.05), a t-test was performed to evaluate the differences in the data derived from the two translations.

3) Subsequently, statistical tests were then used to the data from the two translations concerning linguistic features. After a normality test, data following normal distribution were analyzed using t-tests, while non-normally distributed data were assessed using the Mann-Whitney U test to explore the variations in the data derived from the two translations.

4. Results and Discussion

4.1. Text Types of Two *Medicine* Translations

4.1.1. General Text Types

Based on the results of MDA, the text type of both translations are indicated to be general narrative exposition, which exhibits a preference for informational, non-narrative, situation-dependent, and implicit-persuasion textual characteristics. Specifically for each chapter, Lovell's presents general narrative exposition in each chapter, while Yang and Dai's presents general narrative exposition in chapters 1, 2, and 4, and imaginative narrative in chapter 3.

4.1.2. Genres

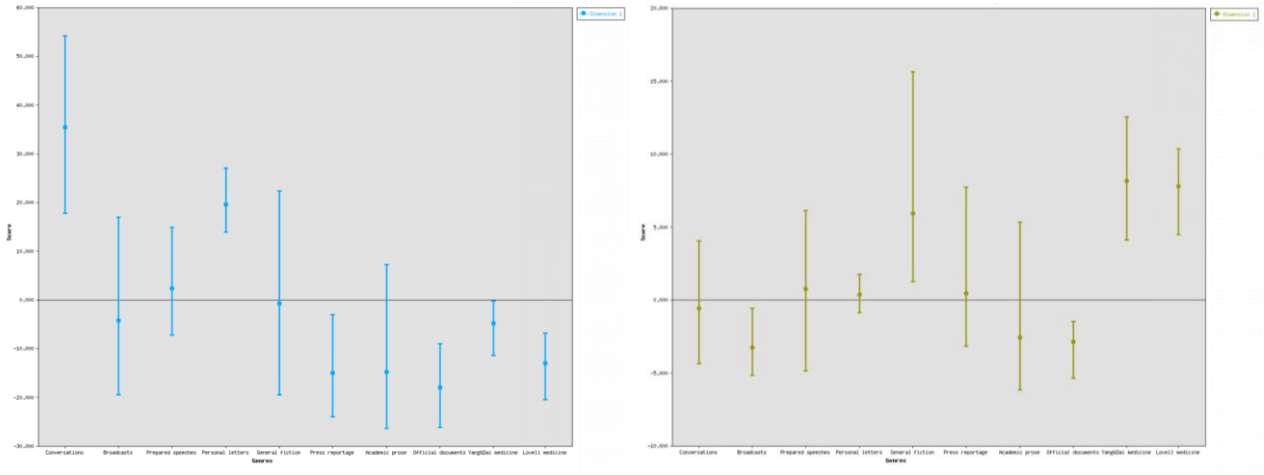

Figure 1: Comparison between Lovell's Translation and Yang & Dai's Translation on Dimension 1 and 2

As shown in Figure 1, on Dimension 1, the Lovell's translation has a lower score range compared to the Yang and Dai's. The latter is characterized by a higher level of informativeness, indicating a greater presence of nouns, long words, and adjectives. The scores of Lovell's translation on Dimension 1 lean towards the genre of academic prose, while the Yang and Dai's scores closer to broadcasts. In terms of Dimension 2, both translations exhibit higher scores, highlighting a significant emphasis on narrativeness within the text, which are consistent with general fiction.

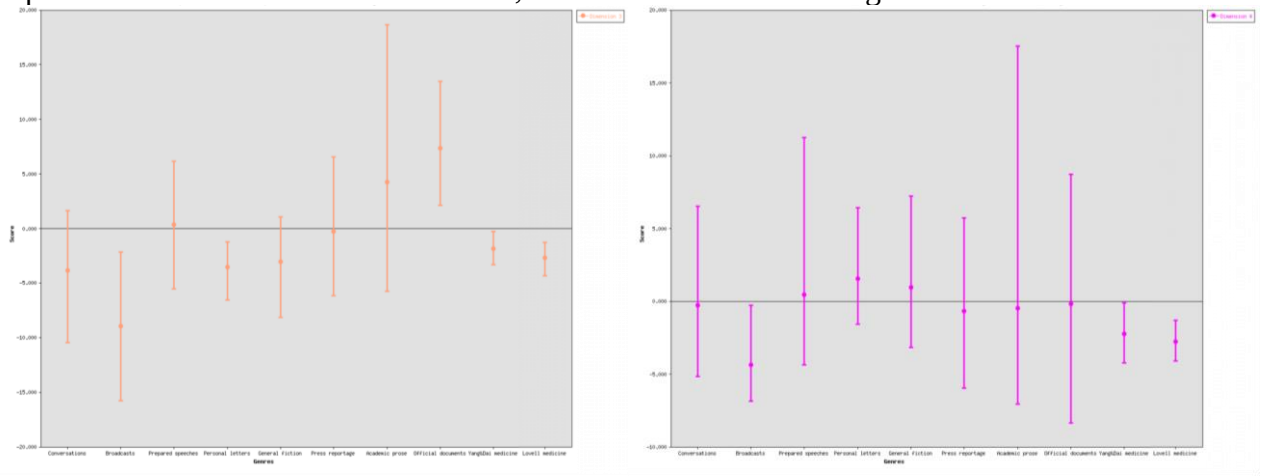

Figure 2: Comparison between Lovell's Translation and Yang & Dai's Translation on Dimension 3 and 4

Figure 2, relating to Dimension 3, indicates that both translations demonstrate similar negative scoring ranges, with Lovell's translation scoring marginally lower. Both exhibit a reliance on contextual cues within the text and the closest genres of both are general fiction. With respect to Dimension 4, it is illustrated that both translations exhibit negative scores. The Yang and Dai's translation, however, displays slightly broader scoring range compared to Lovell's. Moreover, the closest genre of Lovell's translation is broadcasts, whereas Yang and Dai's is closer to press reportage.

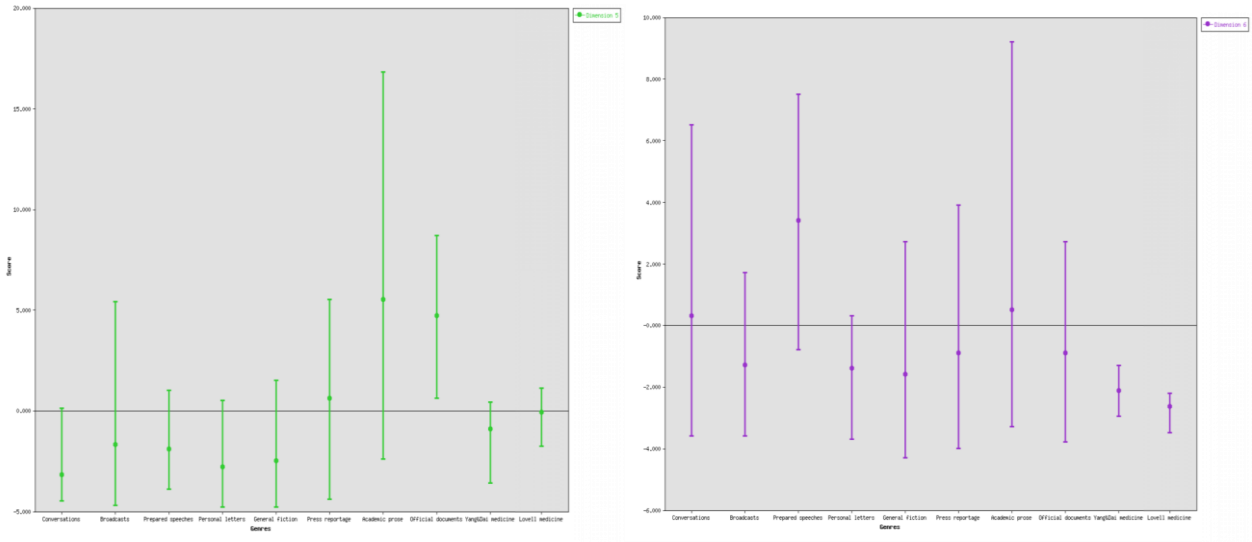

Figure 3: Comparison between Lovell's Translation and Yang & Dai's Translation on Dimension 5 and 6

As depicted in Figure 3, on Dimension 5, the scoring ranges of both translations include a mix of positive and negative scores. Notably, Lovell's translation exhibits a slightly higher scoring range compared to Yang and Dai's translation. Furthermore, the genre of Lovell's translation aligns more closely with press reportage, whereas Yang and Dai's is closer to broadcasts. Concerning Dimension 6, both translations exhibit low scores, suggesting a greater level of linguistic refinement in the organization of the text. Both translations lean towards general fiction.

4.2. Comparisons between Two *Medicine* Translations on Six Dimensions

Table 2: Results of Normality Test for Dimension Scores of Lovell's and Yang & Dai's Translations

Measures	Categorization	Shapiro-Wilk Test		
		Statistics	df	sig.
Dimension 1	Lovell	0.985	4	0.930
	Yang & Dai	0.945	4	0.684
Dimension 2	Lovell	0.960	4	0.781
	Yang & Dai	0.933	4	0.615
Dimension 3	Lovell	0.988	4	0.947
	Yang & Dai	0.963	4	0.795
Dimension 4	Lovell	0.774	4	0.063
	Yang & Dai	0.889	4	0.380
Dimension 5	Lovell	0.908	4	0.474
	Yang & Dai	0.816	4	0.134
Dimension 6	Lovell	0.836	4	0.184
	Yang & Dai	0.967	4	0.822

It is evident, according to Table 2, that the data of both translations on 6 dimensions exhibit a normal distribution (p value > 0.05), enabling further statistical analysis through the application of the t-test. Following Levene's test for equality of variance (p value> 0.05), a Student's t-test was applied to the data. The result is shown in Table 3, which reveals a non-significant mean difference between two translations on six dimensions (p value > 0.05).

Table 3: Results of T-test for Dimension Scores of Lovell's and Yang & Dai's Translations

	Levene's test for equality of variances		T-test for equality of means						
Measures	F	Sig.	t	df	sig.	Mean Difference	Std. Error Difference	95% Confidence Interval of the Differences	
								Lower	Upper
Dimension 1	0.169	0.695	-2.16	6	0.074	-8.14	3.76826	-17.36061	1.08061
Dimension 2	2.704	0.151	-0.156	6	0.881	-0.3625	2.31978	-6.03879	5.31379
Dimension 3	0.038	0.853	-0.954	6	0.377	-0.8525	0.89339	-3.03855	1.33355
Dimension 4	2.997	0.134	-0.421	6	0.689	-0.53	1.25987	-3.6128	2.5528
Dimension 5	0.643	0.453	0.76	6	0.476	0.835	1.0994	-1.85513	3.52513
Dimension 6	0.003	0.958	-1.203	6	0.274	-0.5375	0.44682	-1.63083	0.55583

4.3. Comparisons between Two *Medicine* Translations on Linguistic Features

For certain features, both translations result in zero data after applying MDA. Therefore, these features were excluded from statistical analysis and not included in the table presented.

Table 4: Results of Normality Test for Feature Scores of Lovell's and Yang & Dai's Translations

Categorization	Measures	Shapiro-Wilk Test			Measures	Shapiro-Wilk Test			Measures	Shapiro-Wilk Test		
		Statistics	df	sig.		Statistics	df	sig.		Statistics	df	sig.
Lovell	Tokens	0.73	4	0.024	NN	0.949	4	0.712	SPP2	0.917	4	0.519
Yang & Dai		0.75	4	0.038		0.905	4	0.458		0.893	4	0.396
Lovell	AWL	0.878	4	0.329	NOMZ	0.989	4	0.955	[STPR]	0.939	4	0.649
Yang & Dai		0.981	4	0.906		0.698	4	0.011		0.885	4	0.361
Lovell	TTR	0.835	4	0.182	OSUB	0.961	4	0.787	[SUAV]	0.911	4	0.486
Yang & Dai		0.897	4	0.414		0.689	4	0.009		0.848	4	0.219
Lovell	AMP	0.762	4	0.05	[PASS]	0.921	4	0.544	SYNE	0.803	4	0.107
Yang & Dai		0.768	4	0.056		0.78	4	0.071		0.67	4	0.005
Lovell	ANDC	0.993	4	0.972	[PASTP]	0.882	4	0.345	THAC	0	4	0
Yang & Dai		0.808	4	0.117		0.928	4	0.582		0.63	4	0.001
Lovell	[BEMA]	0.989	4	0.95	[PEAS]	0.933	4	0.611	[THATD]	0.862	4	0.267
Yang & Dai		0.91	4	0.484		0.973	4	0.859		0.916	4	0.514
Lovell	[BYPA]	0.762	4	0.05	PHC	0.846	4	0.212	THVC	0.63	4	0.001
Yang & Dai		0.63	4	0.001		0.821	4	0.144		0.63	4	0.001
Lovell	CONC	0.696	4	0.01	PIN	0.96	4	0.779	TIME	0.912	4	0.492
Yang & Dai		0.729	4	0.024		0.975	4	0.874		0.914	4	0.506
Lovell	COND	0.997	4	0.989	[PIRE]	0.844	4	0.206	TO	0.906	4	0.463
Yang & Dai		0.927	4	0.574		0.63	4	0.001		0.853	4	0.236
Lovell	CONJ	0.762	4	0.05	PIT	0.775	4	0.065	TOBJ	0.762	4	0.05
Yang & Dai		0.944	4	0.677		0.849	4	0.224		0	4	0
Lovell	[CONT]	0.9	4	0.429	PLACE	0.981	4	0.911	TPP3	0.803	4	0.107
Yang & Dai		0.843	4	0.206		0.776	4	0.065		0.899	4	0.425
Lovell	DEMO	0.862	4	0.269	POMD	0.832	4	0.174	TSUB	0.63	4	0.001

Yang & Dai		0.895	4	0.406		0.927	4	0.576		0	4	0
Lovell	DEMP	0.853	4	0.236	PRED	0.967	4	0.826	VBD	1	4	0.999
Yang & Dai		0.952	4	0.726		0.898	4	0.42		0.968	4	0.829
Lovell	DPAR	0	4	0	[PRESP]	0.891	4	0.389	VPRT	0.862	4	0.268
Yang & Dai		0.63	4	0.001		0.966	4	0.819		0.829	4	0.166
Lovell	DWNT	0.94	4	0.655	[PRIV]	0.834	4	0.179	[WHCL]	0.63	4	0.001
Yang & Dai		0.967	4	0.824		0.948	4	0.703		0.63	4	0.001
Lovell	EMPH	0.944	4	0.68	PRMD	0.874	4	0.314	[WHOBJ]	0.63	4	0.001
Yang & Dai		0.878	4	0.331		0.982	4	0.911		0.864	4	0.275
Lovell	EX	0.944	4	0.677	[PROD]	0.63	4	0.001	[WHQU]	0.855	4	0.244
Yang & Dai		0.905	4	0.459		0	4	0		0.949	4	0.712
Lovell	FPP1	0.797	4	0.096	[PUBV]	0.76	4	0.047	[WHSUB]	0	4	0
Yang & Dai		0.995	4	0.979		0.836	4	0.185		0.63	4	0.001
Lovell	GER	0.839	4	0.193	RB	0.91	4	0.482	[WZPAST]	0.991	4	0.963
Yang & Dai		0.915	4	0.511		0.769	4	0.057		0.893	4	0.395
Lovell	INPR	0	4	0	[SERE]	0.63	4	0.001	[WZPRES]	0.805	4	0.112
Yang & Dai		0.932	4	0.607		0.996	4	0.987		0.981	4	0.907
Lovell	JJ	0.833	4	0.177	[SMP]	0.696	4	0.01	XX0	0.862	4	0.266
Yang & Dai		0.99	4	0.959		0.992	4	0.968		0.977	4	0.884
Lovell	NEMD	0.729	4	0.024	[SPAU]	0.864	4	0.276				
Yang & Dai		0.905	4	0.459		0.82	4	0.143				

Upon conducting the Shapiro-Wilk test, it is found on 46 features that the scores of the two translations conform to a normal distribution (*p* value> 0.05). As shown in Table 4, these features include AWL, TTR, AMP, ANDC, [BEMA], COND, CONJ, [CONT], DEMO, DEMP, DWNT, EMPH, EX, FPP1, GER, INPR, JJ, NN, [PASS], [PASTP], [PEAS], PHC, PIN, PIT, PLACE, POMD, PRED, [PRESP], [PRIV], PRMD, RB, [SPAU], SPP2, [STPR], [SUAV], [THATD], TIME, TO, TOBJ, TPP3, VBD, VPRT, [WHQU], [WZPAST], [WZPRES], and XX0. Consequently, the t-test can be appropriately applied to analyze these features. Conversely, the Shapiro-Wilk test reveals a non-normal distribution of scores (*p* value< 0.05) for certain features, such as Tokens, [BYPA], CONC, DPAR, NEMD, NOMZ, OSUB, [PIRE], [PROD], [PUBV], [SERE], [SMP], SYNE, THAC, THVC, TSUB, [WHCL], [WHOBJ], and [WHSUB]. Therefore, the Mann-Whitney U test is recommended for analyzing these features.

Table 5: Results of T-test for Feature Scores of Lovell's and Yang & Dai's Translations

	Levene's test for equality of variances		T-test for equality of means						
Measures	F	Sig.	t	df	sig.	Mean Difference	Std. Error Difference	95% Confidence Interval of the Differences	
								Lower	Upper
AWL	0.071	0.798	2.643	6	0.038	0.095	0.03594	0.00706	0.18294
TTR	0.385	0.558	4.03	6	0.007	13.25	3.28824	5.20397	21.29603
AMP	12	0.013	-2.33	5.828	0.06	-0.11	0.04721	-0.22636	0.00636
ANDC	4.042	0.091	-37.725	6	<0.001	-1.27	0.03367	-1.35238	-1.18762
[BEMA]	0.323	0.59	-1.282	6	0.247	-0.5225	0.40761	-1.51989	0.47489
COND	0.112	0.75	1.037	6	0.34	0.095	0.09161	-0.12915	0.31915

CONJ	0.013	0.912	-0.928	6	0.389	-0.055	0.05927	-0.20002	0.09002
[CONT]	0.541	0.49	1.239	6	0.262	0.4675	0.37741	-0.45599	1.39099
DEMO	0.001	0.978	-1.453	6	0.196	-0.225	0.15484	-0.60388	0.15388
DEMP	0.001	0.976	-0.721	6	0.498	-0.1075	0.14913	-0.47241	0.25741
DWNT	0.318	0.593	0.321	6	0.759	0.0675	0.21055	-0.4477	0.5827
EMPH	0	0.992	-0.296	6	0.777	-0.0675	0.22772	-0.62471	0.48971
EX	0.516	0.5	-0.323	6	0.757	-0.0275	0.08504	-0.23558	0.18058
FPP1	8.986	0.024	0.572	4.758	0.593	0.1975	0.34523	-0.70369	1.09869
GER	11.362	0.015	0.297	4.159	0.781	0.05	0.16861	-0.41117	0.51117
INPR	33.607	0.001	-2.221	3	0.113	-0.2375	0.10696	-0.57788	0.10288
JJ	2.809	0.145	2.771	6	0.032	1.185	0.42772	0.13841	2.23159
NN	0	0.983	1.622	6	0.156	1.3775	0.84906	-0.70007	3.45507
[PASS]	2.045	0.203	-0.874	6	0.416	-0.14	0.1601	-0.53176	0.25176
[PASTP]	0.005	0.948	0.031	6	0.976	0.0025	0.07941	-0.19181	0.19681
[PEAS]	0.658	0.448	-1.279	6	0.248	-0.3	0.23451	-0.87383	0.27383
PHC	0.15	0.712	0.04	6	0.969	0.0075	0.18571	-0.44693	0.46193
PIN	2.996	0.134	3.038	6	0.023	2.135	0.70281	0.41529	3.85471
PIT	2.295	0.181	1.284	6	0.246	0.32	0.24919	-0.28975	0.92975
PLACE	16.93	0.006	0.823	4.205	0.455	0.2725	0.33129	-0.62991	1.17491
POMD	0.074	0.795	-1.227	6	0.266	-0.3275	0.26688	-0.98052	0.32552
PRED	1.095	0.336	-0.68	6	0.522	-0.2075	0.30509	-0.95403	0.53903
[PRESP]	0.198	0.672	1.88	6	0.109	0.44	0.23406	-0.13272	1.01272
[PRIV]	2.605	0.158	-0.64	6	0.546	-0.3575	0.55901	-1.72534	1.01034
PRMD	3.208	0.123	0.988	6	0.361	0.1375	0.13922	-0.20315	0.47815
RB	0.004	0.95	-1.339	6	0.229	-0.925	0.69075	-2.61519	0.76519
[SPAU]	0.745	0.421	-0.343	6	0.743	-0.05	0.14583	-0.40684	0.30684
SPP2	1.607	0.252	0.743	6	0.486	0.12	0.16154	-0.27528	0.51528
[STPR]	2.505	0.165	0.983	6	0.364	0.2	0.20346	-0.29785	0.69785
[SUAV]	0.22	0.656	0.337	6	0.748	0.0425	0.12609	-0.26602	0.35102
[THATD]	1.275	0.302	-1.019	6	0.348	-0.1475	0.14482	-0.50186	0.20686
TIME	1.041	0.347	0.12	6	0.908	0.0325	0.27101	-0.63064	0.69564
TO	0.786	0.409	-1.55	6	0.172	-0.46	0.29679	-1.18621	0.26621
TOBJ	937.5	0	1.729	3	0.182	0.0625	0.03614	-0.05252	0.17752
TPP3	2.057	0.201	0.422	6	0.688	0.2775	0.65823	-1.33314	1.88814
VBD	0.86	0.39	-5.894	6	0.001	-1.7825	0.30243	-2.52252	-1.04248
VPRT	1.065	0.342	0.515	6	0.625	0.175	0.33967	-0.65614	1.00614
[WHQU]	3.663	0.104	-0.132	6	0.899	-0.02	0.15122	-0.39002	0.35002
[WZPAST]	0.668	0.445	3.659	6	0.011	0.3325	0.09086	0.11016	0.55484
[WZPRES]	1.789	0.229	-2.664	6	0.037	-0.2925	0.1098	-0.56117	-0.02383
XX0	0.908	0.377	-0.238	6	0.82	-0.095	0.3998	-1.07328	0.88328

According to the result of t-test shown in Table 5, the Lovell's translation demonstrates significantly higher scores than Yang and Dai's translation on features such as AWL, TTR, JJ, PIN, and [WZPAST]. Conversely, for features like ANDC, VBD, and [WZPRES], the scores of Lovell's translation are significantly lower than those of Yang and Dai's. Moreover, Lovell's translation exhibits not significantly higher scores than Yang and Dai's on features including COND, [CONT], DWNT, FPP1, GER, NN, [PASTP], PHC, PIT, PLACE, [PRESP], PRMD, SPP2, [STPR], [SUAV], TIME, TOBJ, TPP3, and VPRT. On the other hand, for features like AMP, [BEMA], CONJ, DEMO,

DEMP, EMPH, EX, INPR, [PASS], [PEAS], POMD, PRED, [PRIV], RB, [SPAU], [THATD], TO, [WHQU], and XX0, the scores of Lovell's translation are not significantly lower than Yang and Dai's.

Table 6: Results of Mann-Whitney U Test for Feature Scores of Lovell's and Yang & Dai's Translations

Measure	Lovell vs. Yang & Dai				
	Mean rank		u	z	p
	Lovell	Yang & Dai			
Tokens	3.75	5.25	5	-0.866	0.386
[BYPA]	5.25	3.75	5	-0.992	0.321
CONC	5.25	3.75	5	-0.949	0.343
DPAR	4	5	6	-1	0.317
NEMD	3.75	5.25	5	-0.949	0.343
NOMZ	4	5	6	-0.581	0.561
OSUB	5.38	3.63	4.5	-1.042	0.297
[PIRE]	5.25	3.75	5	-0.992	0.321
[PROD]	5	4	6	-1	0.317
[PUBV]	3.13	5.88	2.5	-1.607	0.108
[SERE]	3.63	5.38	4.5	-1.076	0.282
[SMP]	4.13	4.88	6.5	-0.438	0.661
SYNE	4	5	6	-0.584	0.559
THAC	4	5	6	-1	0.317
THVC	4.5	4.5	8	0	1
TSUB	5	4	6	-1	0.317
[WHCL]	4.38	4.63	7.5	-0.189	0.85
[WHOBJ]	3.88	5.13	5.5	-0.833	0.405
[WHSUB]	4	5	6	-1	0.317

Table 6 suggests that, utilizing the Mann-Whitney U test, a non-significant difference in the distributions of the two groups is identified (p value> 0.05). In terms of some features, such as [BYPA], CONC, OSUB, [PIRE], [PROD], and TSUB, Lovell's translation scores non-significantly higher than Yang and Dai's translations. Conversely, Lovell's translation scores non-significantly lower than Yang and Dai's translations for the features [Token], [DPAR], [NEMD], [NOMZ], [PUBV], [SERE], [SMP], [SYNE], [THAC], [WHCL], [WHOBJ] and [WHSUB]. The scores of Lovell's and Yang and Dai's translations are equivalent in the feature THVC.

In particular, the linguistic features on which two translations' scores differ significantly call for special attention. Lovell's higher scores on AWL and TTR imply that her translation is more lexically complex and rich than Yang and Dai's, whereas its high score on JJ and PIN indicates that Lovell focuses more on the use of attributive adjectives and prepositional phrases than Yang and Dai, resulting in more vivid and detailed descriptions. Lovell's translation employs more past participles as determiners, as shown by a significantly higher score on [WZPAST]. In contrast, Yang and Dai's translation employs more present participles, as indicated by its significantly higher [WZPRES]. Furthermore, the Lovell translation's lower ANDC and VBD scores imply that it makes less use of independent clause coordination and past tense than the Yang and Dai's.

Lovell, a native British translator, has an authentic knowledge of English culture. She used a domestication method to convey the essence of Chinese culture to Western readers while removing possible obstacles to comprehension and avoiding repetition of vocabulary to ensure the text's

readability, which could explain why her translation scores significantly higher on AWL and TTR. Yang and Dai's translation, on the other hand, was completed through collaboration between Chinese and English translators with the goal of promoting Chinese culture abroad, conveying the ideological connotation of Chinese culture. Therefore, they prefer concise language to facilitate cross-cultural communication and adhere to a general principle of "foreignization-based and domestication-supplemented" (Ouyang, 2014: 102) [12]. The lower TTR score of their translation could be attributed to their less varied vocabulary, which preserves the Chinese language's inclination to repeat words.

5. Conclusion

In conclusion, both translations are generally categorized as general narrative expositions and exhibit an overall similarity with a non-significant discrepancy. With regards to the linguistic features, Lovell's translation significantly outperforms Yang and Dai's in features including AWL, TTR, JJ, PIN, and [WZPAST]. By comparison, for features such as ANDC, VBD, and [WZPRES], Lovell's translation scores significantly lower than Yang and Dai's. These variations might be attributed to the differences between two translators in their cultural background and translation strategies determined by specific translation purposes. This study is a practice of adopting MDA to carry out corpus-based translator style research. Future studies on translator styles can combine qualitative research and MDA to expand its depth and breadth.

References

[1] Hao T. (2012). The Theme of Medicine and a Glimpse into Lu Xun's Cultural Thought. Mangzhong Literature, 12, 35–36.

[2] Huang L., & Wang K. (2011). Corpus-based translation studies: Issues and progress. Foreign Language Teaching and Research, 43(6), 911-923+961.

[3] Baker, M. (2000). Towards a Methodology for Investigating the Style of a Literary Translator. Target. International Journal of Translation Studies, 12(2).

[4] Biber, D. (1988). Variation across Speech and Writing. Cambridge University Press.

[5] Baker, M. (1993). Corpus Linguistics and Translation Studies—Implications and Applications. In M. Baker, G. Francis, & E. Tognini-Bonelli (Eds.), Text and Technology (p. 233). John Benjamins Publishing Company.

[6] Yan Y., & Han N. (2015). A Corpus-Based Study of Translator's Style: A Case Study of Two English Translations of Lu Xun's Novels. Foreign Language Education, 36(2), 109–113.

[7] Li D., He W., & Hou L. (2018). A Corpus-Assisted Study on the Translation Style of Julia Lovell. Foreign Language Education, 39(1), 70–76.

[8] Zhu S. (2021). A Corpus-Based Study of Translation Style in Diary of a Madman. Foreign Language and Literature, 37(5), 119–128.

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

[10] Zhao C. (2020). Using corpus-based multidimensional analysis for the comparative description of translator's style. Foreign Language Learning Theory and Practice, 3, 67-73+84.

[11] Nini, A. (2019). The Multi-Dimensional Analysis Tagger. In Berber Sardinha, T. & Veirano Pinto M. (eds), Multi-Dimensional Analysis: Research Methods and Current Issues, 67-94, London; New York: Bloomsbury Academic.

[12] Ouyang Y. (2014). On Yang Xianyi's Literary Translation Thought. Jiangxi Social Sciences, 34(12), 100–104.