

Translating Social Items From the Tibetan Epic “King Gesar” Into English: What a Corpus-Based Analysis Reveals and How It Can Help

Yuqiao Zhou

Faculty of Modern Languages and Communication, University Putra Malaysia, Selangor, Malaysia;
School of Humanities, Creative Industries and Social Science, The University of Newcastle, Newcastle, Australia

Hazlina Abdul Halim

Faculty of Modern Languages and Communication, University Putra Malaysia, Selangor, Malaysia

Diana Abu Ujum

Faculty of Modern Languages and Communication, University Putra Malaysia, Selangor, Malaysia

Ling Yann Wong

Faculty of Modern Languages and Communication, University Putra Malaysia, Selangor, Malaysia

Abstract—The *Ge Sa-er* (also known as *Gesar*) saga has been handed down for above 1000 years by Tibetan Chinese, Mongolian Chinese, the Tu and the Naxi, describing a heroic story in present-day Kangba Tibetan area. This article analyse the English translation of social discourse in the *Ge Sa-er* version retold by Alai (2009). Due to different knowledge shared by source and target readers, it is difficult to represent the source social meaning in the English translation. Aiming to resolve the translation problems revealed by social cognition, this article uses statistical methods together with corpus to conduct stylistic analyses on *Ge Sa-er Wang* (Alai, 2009). An unexpected finding of this research is that the data from statistical methods can better predict the translation difficulties of *Ge Sa-er Wang* than the data from corpus. Based on the stylistic features of *Ge Sa-er Wang*, we propose a social discourse translation model for Chinese ethnic minority literature.

Index Terms—social item translation, corpus stylistics, ethnic minority literature, Tibetan, *Ge Sa-er Wang*

I. INTRODUCTION

With about one thousand years of history, the *Ge Sa-er* saga plays a critical role in understanding the culture of the Chinese Tibetan areas¹. As one of the ethnic minorities in China, Tibetans are centered in Tibet Autonomous Provincial Region, Sichuan Province, Qinghai Province, Gansu Province and Yunnan Province. The *Ge Sa-er* saga shows indigenous culture of the Tibetan. However, the structures of indigenous cultures are unavoidable to be influenced by modern civilization. It is therefore crucial to preserve ancient epics and folklores (Gui, 2012), as they contain cultural and social elements about the ancient ideational framework.

The translation history of the *Ge Sa-er* legend can be traced back to around 300 years ago, when people first translated several chapters in *The Epic of Ge Sa-er Wang* into Russian, German, French, English and other languages after the sage was recorded in the form of printed book. The latest English version of the *Ge Sa-er* saga was accomplished in 2013 when Goldblatt² & Lin co-translated the Chinese version of 《格萨尔王》 *Ge Sa-er Wang* (Alai, 2009). *Ge Sa-er Wang* is a celebrated work by Alai - a prize-awarded Tibetan Chinese writer. To translate *Ge Sa-er Wang*, an unavoidable obstacle is how to transfer social items which demonstrate the socially shared knowledge, ideology and attitude of people in Kangba-Tibetan area. Social items, in the first chapter of *Ge Sa-er Wang*, consist of 6,140 tokens, which is to say, 9.11% of the total number of tokens (67,403). This research takes an initial step to figure out a translation model for social items in *Ge Sa-er Wang* by conducting a corpus-based stylistic analysis.

However, corpus fails to give a clue about the features of the frequencies and the correlation between the wording and plot, although corpus can be used to describe word frequency and sentence structure. Statistical methods can explicate the correlation between the items used by the author and the episode designed in each part, and statistical methods are also conducive to exploring the characteristics of these frequencies. As a result, this research will use the statistical software, Statistical Package for Social Sciences (SPSS), together with corpus to analyse the stylistic features

¹ There are mainly five Tibetan areas in China, including Kangba, Weizang, Anduo, Jiarong and Gongbu.

² Goldblatt has translated more than 80 works by over 30 Chinese writers, including many novels by Yan Mo who was a Nobel Prize winner in 2012.

of *Ge Sa-er Wang*. Based on the stylistic features and translation analyses, this research suggests a translation model for social epithets in Chinese ethnic minority myths or other Chinese myths.

II. TAXONOMY OF SOCIAL DISCOURSE AND STYLISTIC ANALYSIS

Corpus stylistics has become increasingly popular in the past two decades since the “corpus turn” (Short & Leech, 2007, p. 286) in stylistics. Onanuga (2018, p. 136) mentions, “stylistic analyses enable the immediate understanding of utterances and texts”. Corpus stylistics provides descriptive concepts and approaches to research on literary stylistics, “thus situating the field within the wider context of the digital humanities” (Mahlberg, 2017, p. 388). Meanwhile, corpus stylistics is retaken by many scholars to explore the effect of linguistic choice on interpreting literature (e.g., Youdale, 2020; McIntyre & Walker, 2019; Mahlberg & Wiegand, 2020; Mahlberg, 2013, 2017; Stockwell & Mahlberg, 2015; Flowerdew, 2012), and to examine how linguistic description associates with literary appreciation (Mahlberg, 2013).

Compared with corpus stylistics, statistical or computational stylistics fails to receive high attention. Only few scholars conduct statistical or computational stylistics. For instance, Youdale (2017, 2020) takes computer-aided approaches to outline the corpus-based stylistic features of literary works, exploring the stylistic effects on translation. Youdale’s research proves the significance to use computational/statistical methods and corpus stylistics together to conduct translation studies. Biber (2011) also advocates the integrated use of statistical methods and corpus stylistics. This article will combine SPSS and corpus to analyse the distribution features of social items, and explore how the distribution of social items is associated to plot.

To collect social items, we created a corpus of the first chapter of *Ge Sa-er Wang*, consisting of 67,403 Chinese characters, among which 6,140 words were of social discourse. The corpus of the first chapter is sufficient to analyse the features of this ethnic minority literature according to Biber’s (1990) corpus research which proves that a 1,000-word sample can show linguistic features in a stable manner. In the first chapter, 184 social items are collected based on van Dijk’s (2014, 2016) socio-cognitive approach. Socio-cognitive approach is suggested by Shao and Zhou (2022a, 2022b) as an important model to analyse underlying social meaning in literary works when Shao and Zhou explore the English translation of Alai’s works. Socio-cognitive approach explores the associations among discourse, cognition and society, viewing social cognition as the interface between discourse and society. It points out that social structure shapes people’s understanding of discourse, while discourse is not processed directly but, instead, via a set of cognitive connections. Socio-cognitive approach groups social discourse mainly into three types: socially shared knowledge (SSK), ideology and attitude. SSK points to the interpretation of the world “shared with other members of the same epistemic community” (van Dijk, 2016, p. 69). Socially shared attitude decides public ideational systems regarding different phenomena, while ideology, as a fundamental cognition, influences people’s attitude and behavior. Accordingly, the taxonomy for the social items within *Ge Sa-er Wang* is based on van Dijk’s (2016) socio-cognitive approach: SSK-items, ideological and attitudinal items.

III. RESEARCH DESIGN AND DATA COLLECTION

A. Text Analysis

According to the legend, Ge Sa-er is a son of the 天王 “Tian King” and 天母 “Tian mother”³. Upon samsara, he is reincarnated in human form, named 崔巴噶瓦 “Cuiba Ga-wa”. Being a warrior throughout his life, Ge Sa-er unifies more than 150 tribes and consequently, he creates the country, named Ling. He is the person who brings civilization to the people of Ling.

In the ancient time, the way people preserved the folklore of Ge Sa-er was to recount it as sagas or Thangka paintings before it was textualized (Honko, 2000). The oral epic of Ge Sa-er consists of more than one million lines and approximately 20 million words; it is narrated in “chronological order” (Karmay, 1995), listed as Intangible Cultural Heritage by UNESCO. The West also embraces Ge Sa-er as a model for “engaged”, “secular”, and “enlightened-warrior” Buddhism (FitzHerbert, 2016). As a pivotal tool to comprehend the social life and culture of Tibetan Chinese, the Ge Sa-er legend has attracted increasing attention of scholars, which builds on the development of the story.

Although it is difficult to translate such a long history and abundance of socio-cultural resonances into a foreign-language system where the story has no cultural presence, many translators (see Schmidt, 1839; Francke, 2000; Alai, 2013) have undertaken to translate this story, spreading the Ge Sa-er saga universally. This article takes *Ge Sa-er Wang* by Alai (2009) as a case to conduct a stylistic analysis and translation study on social items. Social items serve as an important means that people use to comprehend the socially shared knowledge, ideology and attitude of Tibetan Chinese in Kangba area.

Ge Sa-er Wang comprises three chapters, including 66 short stories. The first chapter, 神子降临 *Shenzi Jianglin* “Divine Descent”, consists of 21 short stories, covering 98 pages and 67,403 Chinese characters. In this chapter, the spatiotemporal background, location, characters’ relations and other background information are elaborated. This article

³ The Chinese character 天 tian refers to sky, as well as the place where deities live.

respectively, which are not social items either. *Ga* is especially included in the items of names and places, with a total frequency of 234, while *mo* appears 153 times, involved in 22 lexes.

Based on the corpus, we find that, out of 184 social items, 85 items are stylistic-featured characters (see Appendix), covering 4,231 epithets, accounting for 85.63% of the total frequency of 4,941 social items. If readers can understand the meaning of 85 stylistic-featured characters, they can presumably interpret most social discourse in the first chapter of *Ge Sa-er Wang*. Similarly, if the translators are able to resolve the problems generated by those 85 stylistic-featured characters, they can almost transfer above 80% of social expressions. In addition, the type/token ratio (TTR) of social items is 3%, while the TTR of featured characters is 2.01%; both TTRs exhibit a low quantity of types compared with tokens.

B. Statistical Approach: The Phenomenon of 0

Among all social items counts, the highest percentage of distribution is the frequency of 0. 0 of an expression means that this item is not used in every part; that is to say, this item may appear in one part, two parts and even three parts, but it does not be used in at least one part. 0 in part 1 points to the items used in other parts rather than part 1, for instance. The higher the frequency of 0 is, the more difficult it is for target readers to interpret the text because the high frequency of 0 predicts lots of different words in each part. Readers may still need to interpret many new words after they finish reading one part, due to the high number of different social items in the following part.

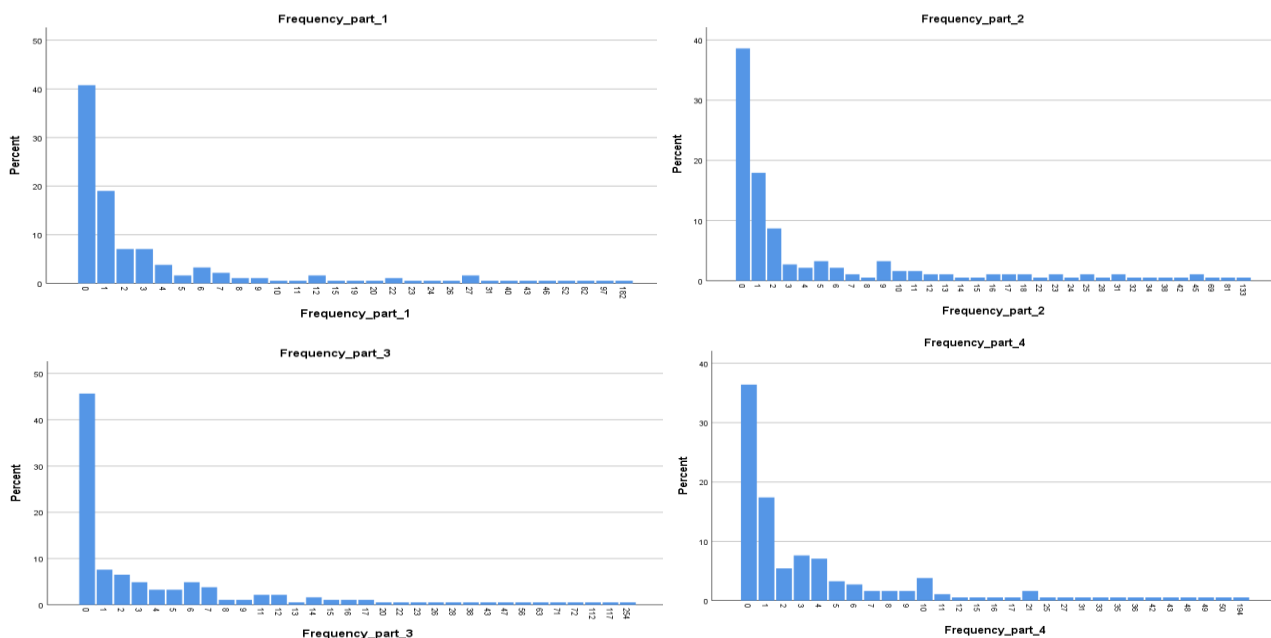


Figure 2. The Distribution of 0

Figure 2 shows above 40% zero items in Part 1, 3 and 4, and around 40% in Part 2. This result reveals a high possibility of having social items that exist exclusively in one special part along with different plots. After exploring the frequencies of each item, we find that there are 73 epithets with three 0s, that is, that are used exclusively in one part, accounting for 39.63% of 184 social items, which may complicate interpretation and translation. SSK-, ideological and attitudinal items all exhibit a similar distribution to that in Figure 2, where 0 accounts for the highest percentage across the four parts.

C. Statistical Approach: Correlation Between Language Choice and Plot Development

Based on the frequencies collected by the corpus, the correlation between the wording of social items and plot is analysed by using SPSS. Since the socially shared knowledge, ideology and attitude of the author may be comparatively stable, one may question whether the social discourse structures will remain consistent as the plot develops. Without applying quantitative method, we could just infer the association according to our personal feelings by reading the text, while SPSS helps answer this question in an objective way.

The selection of a special test depends mainly on the distribution of data and the source and size of samples. This article consists of four groups of independent samples without being chosen randomly and the data are not distributed normally, so Spearman's test is chosen to explore the correlation between social items' wording and plot. Spearman's test on 184 social items reveals that the correlation between them is significant, at the level of 0.01 (see Table 1), which means that the application of social items is strongly related to the arrangement of plots: the story's different plots stimulate the writer's social cognition and his use of the corresponding social items, and this however consistent the writer is in terms of his or her socially cognitive structure of SSK, ideology and attitude.

TABLE 1
CORRELATIONS BETWEEN WORDING AND PLOT DEVELOPMENT

		Part_1	Part_2	Part_3	Part_4
Spearman's rhoPart_1	Correlation Coefficient	1.000	.505**	.518**	.403**
	Sig. (2-tailed)	.	.000	.000	.000
Part_2	Correlation Coefficient	.505**	1.000	.551**	.514**
	Sig. (2-tailed)	.000	.	.000	.000
Part_3	Correlation Coefficient	.518**	.551**	1.000	.551**
	Sig. (2-tailed)	.000	.000	.	.000
Part_4	Correlation Coefficient	.403**	.514**	.551**	1.000
	Sig. (2-tailed)	.000	.000	.000	.

** . Correlation is significant at the 0.01 level (2-tailed).

Whereas the Spearman results of SSK-items and ideological items remain the same as the results of 184 social items to be significant different at 0.01, the results of attitudinal items exhibit no significant difference between Part 1 and 2, Part 1 and 4, as shown in Table 2. All the four parts depict people’s attitude towards the relationship across humans, evil beings and deities, and people’s ideas about Ge Sa-er storytellers and dreams. Deities with big power save people out of their misery brought about by evil beings who can hide inside humans’ hearts. People’s attitude, at that time, towards professional Ge Sa-er bard was that only the person chosen by deities could learn the story episode from dreams. It is said that deities make the story shown in a chosen bard’s dreams.

TABLE 2
CORRELATIONS BETWEEN ATTITUDE-SI WORDING AND PLOT DEVELOPMENT

		Part_1	Part_2	Part_3	Part_4
Attitude_part_1	Correlation Coefficient	1.000	.218	.486**	.002
	Sig. (2-tailed)	.	.224	.004	.991
Attitude_part_2	Correlation Coefficient	.218	1.000	.495**	.396*
	Sig. (2-tailed)	.224	.	.003	.022
Attitude_part_3	Correlation Coefficient	.486**	.495**	1.000	.384*
	Sig. (2-tailed)	.004	.003	.	.028
Attitude_part_4	Correlation Coefficient	.002	.396*	.384*	1.000
	Sig. (2-tailed)	.991	.022	.028	.

** . Correlation is significant at the 0.01 level (2-tailed).

* . Correlation is significant at the 0.05 level (2-tailed).

The correlation results predict that the story’s development may influence the choice of social items, while attitudinal items tend to be less influenced by plots. That is to say, SSK- and ideological items may cause more difficulties to translation due to its alternation along with plots.

D. Statistical Approach: The Distribution of Social Items

The data on the features of frequency and correlation facilitate stylistic analyses of social discourse, while the differences shown by data cannot totally reflect readers’ feelings. For literary works, the judgement of social items’ wording differences in the four parts may vary from one reader to another. Some readers may hold the idea that there is no wording difference across the four parts in that they see many social items repeating in each part. Although 39.63% of 184 social items are used in only one part, the sum of these items is 249 which is 5.04% of the total social items; that means almost 95% social epithets exist in at least two parts. The repetition of the 95% social items would make readers perceive no wording difference across the four parts. However, for the readers who take plots into consideration, they may propose a different suggestion that wording may distinguish each other as the social items are used according to different plot.

When Independent-samples Kruskal-Wallis test - a nonparametric test - is used to analyse the differences across the four parts, the result illustrates that the repetition of social items in the four parts of the first chapter is not significantly different. It also shows that the quantity of ideological items among the four parts does not vary from each other significantly. However, the pairwise results reveal that in spite of no significant difference across the four parts, the distributions of SSK-items between Part 3 and 4, Part 1 and 4 show significant difference at the level of 0.05, as depicted in Table 3. The choice of social items in Part 1, describing the human realm before divine descent, is different from the use of social items to depict the deity’s re-exile in Part 3. Furthermore, the events narrated in Part 2, after divine descent, do not remain the same as that which recounts the deity’s exile in Part 3.

TABLE 3
PAIRWISE COMPARISONS OF SSK-ITEMS IN EACH PART

Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.
3-1	2.263	15.675	.144	.885
3-2	29.304	15.675	1.869	.062
3-4	-36.784	15.675	-2.347	.019
1-2	-27.041	15.675	-1.725	.085
1-4	-34.521	15.675	-2.202	.028
2-4	-7.479	15.675	-.477	.633

In respect to attitudinal items, the difference between Part 1 and 3 is significant at the level of 0.01 (see Table 4). Compared with Part 1 which explains the current situation among humans, evil beings and deities, Part 3 mainly shows deities' great power but not the complicated relationship between deities and other living beings. The very different plot may distinguish the attitudinal wording of Part 1 from that of Part 3.

TABLE 4
PAIRWISE COMPARISONS OF ATTITUDINAL ITEMS IN EACH PART

Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.
3-2	7.318	8.340	.878	.380
3-4	-14.894	8.340	-1.786	.074
3-1	22.939	8.340	2.751	.006
2-4	-7.576	8.340	-.908	.364
2-1	15.621	8.340	1.873	.061
4-1	8.045	8.340	.965	.335

E. Translation Difficulties Analysis

Based on the statistical data, SSK-items show a strong correlation between the wording and plot, and have a significant difference between two pairs; ideological items show no significant difference of the distribution of social items across the four parts, but a strong correlation between the wording and plot; attitudinal items show difference only between one pair and less influenced by the development of plots. From the SPSS results, can we infer that it is difficult to translate SSK-items due to the difference in distribution and the correlation with plots; ideological items may rank the second type of SI to bring translation obstacles; and it may be easy to transfer attitudinal items?

However, based on the TTRs calculated by the corpus, the answer is totally opposite to SPSS. TTR could be used to predict the glossary difficulties within a text (Baker, 1995), while the TTR of SSK-, ideological and attitudinal items equals 2.39%, 4.03% and 4.42% respectively. According the TTRs, the attitudinal items may be the most difficult ones, and SSK-items are the easiest one. Then which result could better predict the translation difficulties?

V. ENGLISH TRANSLATION OF SOCIAL ITEMS

A. SSK-Items Translation

The frequency of SSK-items amounts to 3,274, with 4,053 tokens, accounting for 66.26% of the total frequency of 4,941 social items. The following table illustrates the top 10 most frequently used SSK-items, including eight socially stylistic-featured characters.

TABLE 5
TOP 10 SSK-ITEMS

SSK-SI	Part 1	Part 2	Part 3	Part 4	Total
人 <i>ren</i> "humans"	182	133	254	194	763
神 <i>shen</i> "deities"	97	81	56	25	259
噶 <i>ga</i>	19	69	112	36	236
觉如 <i>Jueru</i>	0	38	117	43	198
天 <i>Tian</i> "the place deities live"	43	45	71	27	186
魔 <i>mo</i> "evil beings"	82	18	22	31	153
师 <i>shi</i> "a model/a person with specialised skills"	27	45	16	4	92
妖 <i>yao</i> "attractive evil beings"	40	12	15	21	88
菩萨 <i>Pusa</i> "Bodhisattva"	22	1	14	50	87

Table 5 demonstrates that 人 *ren* remains the highest ranking of the 184 social items in Chapter One; 神 *shen*, 魔 *mo* and 噶 *ga* appear as top 10 social items, as well as in the type of SSK. *Ren* refers to higher-order animals that are able to manufacture tools and use them for work (Wei, 2020, p. 812). In terms of socio-cultural level, *ren* means people living in the human realm; it is used to show the distinctions between the realms of humans, deities and evil beings. *Shen* points to the creator or the ruler of the universe (Wei, 2020, p. 856), and, socio-culturally, it refers to deities living in Tian with strong power over nature, representing a particular quality. *Mo* refers to an evil ghost that can puzzle people and cause them to die (Wei, 2020, p. 658), while, socio-culturally, it points to anything preventing people from behaving positively in a way leading to a good future. In the English version, *ren* is translated as (*the*) *humans/the people* and *shen* as (*the*) *deities*. By applying the translation procedure of cultural substitution, *mo* is translated as (*the*) *demons* (shown in Extract 1).

(1) The source text

说来也怪，神不下界，魔也就消失了。也许魔折腾人，只是为了向神挑衅，如果只是欺负软弱的人，自己都觉得没什么劲头。(Alai, 2009, pp. 1-2)

(1) Gloss

It is strange. When deities did not go down to the earth world, the evil beings were disappeared either. Maybe

the reason for the evil beings to afflict humans was that they wanted to provoke deities, and they felt boring to only bully wicked humans.

(1) The target text

Yet once the deities stopped meddling, the demons seemed also to disappear. Perhaps they had plagued humans simply to taunt the deities, and lost interest when only the feckless humans remained. (Alai, 2013, p. 3)

Cultural substitution refers to a translation procedure of using special-existing epithets in target cultural system but containing the same annotations and language functions to replace the items exclusively used in source culture (Olk, 2013). *Demon* refers to evil being/spirit in Christianity (Du, 1994). Although *demon* shares the same meaning with *mo* to point to evil beings and show the same function to be a special item in target language system, it will mix the source story with target mythological system. The mix of source and target mythologies together will impede target readers' understanding of *mo*, and thus results in a breakdown of the construction of a cognitive schema of *mo*.

The item 噶 *ga* implies a person with power or denotes a special constitution to give orders. In this myth, *ga* is used together with 岭 “Ling” as 岭噶 “Ling-ga”, the place into which Ge Sa-er descends to save the inhabitants. In addition, this character is also included in such names as 崔巴噶瓦 “Cuiba Gawa” and 嘉察协噶 “Jiacha Xiega”. Cuiba Gawa is the deity who descends into the human realm to be named Ge Sa-er, and Jiacha Xiega is his elder brother. The two of them have the same father, an official in the clan of 幼 “youth”, but have different mothers. Both have a chance to lead the Darong tribe. The words *kar* and *Ga* are used to translate 噶 *ga*, which is also sometimes omitted in the English translation. *Glingkar*, *Thosba Gawa* and *Gyatsa Zhakar* are applied to translate the names Ling-ga, Cuiba Gawa and Jiacha Xiega. However similar the pronunciations of *kar* and *Ga* are to 噶 *ga*, they fail to transfer the social meaning implied by the Chinese character *ga*, without connotating power.

B. The Translation of Socially Shared Ideological Items

Socially shared ideology points to the ideas shared by a specific group of people, guiding them how to make decisions (van Dijk, 2016). Chapter One uses ideological discourse to emphasise human inability to conquer evil beings, and the great power held by deities to relieve humans of their suffering (Zhou et al., 2021). The ideological discourse makes it reasonable that one deity should be chosen to save the human realm. Based on the corpus, 54 ideological items have been collected, totaling 1197 epithets, with 1340 tokens, accounting for 24.22% of the total frequency of 4941 social items. In Chapter One, the top 5 most frequently applied items are shown in Table 6.

TABLE 6
THE TOP 5 IDEOLOGICAL ITEMS

Ideology-SI	Part 1	Part 2	Part 3	Part 4	Total
王	9	11	72	35	127
上	24	34	47	15	120
大	52	32	20	9	113
国	27	11	43	6	87
下	26	25	16	11	78

The socially stylistic-featured characters of 上 *shang* and 下 *xia* rank top of the 5 most frequent expressions in Chapter One. In this myth, in addition to the linguistic meaning of *shang*, which refers to a higher place or position than something (Wei, 2020, p. 848), and *xia*, which means a lower place or position (Wei, 2020, p. 1027), their ideological meanings are mainly used to describe the hierarchy between deities who live in Tian above and mankind who live in the human realm below. Deities are believed by people to have extremely powerful abilities to enable humans to overcome adversity in any circumstance. Furthermore, *shang* and *xia* demonstrate the hierarchy between the people with power and those without. The former usually guides and controls the behavior of the latter in the human realm. In this text, Alai sometimes uses *shang* to exhibit people's praying for support from deities or Bodhisattvas by looking up Tian. Extracts are listed below to illustrate whether the translation represents the social structure of the item *shang*.

(2) The source text

...他说：难道上天也要如此戏弄我们吗？ (Alai, 2009, p. 45)

(2) Gloss

...He said, “Does it mean that the above Tian plays such tricks on us too?”

(2) The target text

... “Is it possible that Heaven is merely toying with us?” they wondered. (Alai, 2013, p. 54)

The Chinese item 上天 *shangtian* “above Tian” in Extract (2) refers to the place where various deities live. The character 上 *shang* metaphorically shows the hierarchical structure between deities and ordinary people: deities are in a higher social status than the mortals. *Shangtian* is usually translated as *sky* or *Heaven* by Goldblatt and Lin. *Sky* is used to express the linguistic meaning of 天 *tian*, while *Heaven* is applied to translate *Tian* when it refers to the place where deities live. However, *Heaven* is typically used to express the place believed to be the home of God, where good people go after they die, as well as to refer to God, when the first letter is capitalized. God points to the being or spirit that is worshipped and is believed to have created the universe in Christianity, Judaism, and Islam (Hornby, 2004, p. 752).

Heaven conveys similar meaning with *shangtian*. The item *shangtian* is an exclusive concept in source language system, and *Heaven* has the same function to be a special item in target language system. Nevertheless, the substitution of *Heaven* for *shangtian* will impede target readers' construction of the meaning of *shangtian*, as well as of the other items closely related to *shangtian*, such as Buddhas, Bodhisattvas and the Tian King. Although like *Tian*, *Heaven* connotes that it is a place higher than the human realm, it fails to emphasize the ideological meaning of the different level between deities and humans due to the omission of *shang*.

Next, *xia* is applied to describe the things happening in the human realm from the point of view of deities in Tian; and sometimes *xia* points to the lower position of the people who are possessed and controlled by the powerful.

(3) The source text

...他挥动手臂，是让下人们更大声地歌唱。(Alai, 2009, p. 46)

(3) Gloss

...He waved his arm to ask the lower people to sing louder.

(3) The target text

... With a wave of his arm, he ordered his servants to sing louder. (Alai, 2013, p. 55)

In Extract (3), 下人们 *xiarenmen* refers to those people who are legally owned by big wealthy or powerful families who force them to work. This sentence is applied to narrate the action of arm-waving by Sen Lun, the father of Ge Sa-er, in a noble family like royal one; Sen Lun is a candidate to rule the tribe. 下 *xia* in this expression refers to people who are opposed to each other in the master-slave hierarchy. However, a servant points to "a person who works in another person's house, and cooks, cleans, etc. for, them" (Hornby, 2004, p. 1586). Servant is a kind of job, which may not refer to a lower-graded social class, failing not only to show the hierarchy between masters and slaves, but also to explain the linguistic significance.

C. The Translation of Socially Shared Attitudinal Items

Socially shared ideology affects the way people make choices and thus plays a pivotal role in people's attitudes, which are in relation to people's emotion to judge whether a choice is positive or negative, right or wrong (van Dijk, 2016). In this myth, in addition to the items *would like*, *cannot* and *still want*, many other items without modal verbs also show socially shared attitudes. Altogether, 33 attitudinal items were collected, totalling 470 epithets with 747 tokens, which is 9.51% out of the total frequency of the social items.

TABLE 7
THE TOP 4 ATTITUDINAL ITEMS

Attitude-SI	Part 1	Part 2	Part 3	Part 4	Total
让	27	31	63	33	154
梦	46	42	26	21	135
唱	23	31	6	21	81
艺人	4	13	5	2	24

Table 7 shows that the stylistic-featured words 让 "let/allow", 梦 "dream", 唱 "sing", and the phase of 艺人 "artist" rank in the top 4 of the most frequently used attitudinal items. 让 "let/allow" underlies the great power of deities to control everything; deities control human behaviors and determine human rights. In addition, 让 "let/allow" also shows that some people's fate is decided by those in power.

As mentioned above, people believe that deities usually hand down the detailed Ge Sa-er story in the dreams of chosen singers or bards. This myth illustrates that people inside the fictional world may share the same attitude about the designation of the singer with the people of that time in the real world. Table 7 shows that the items relevant to the attitude towards singers ranks in the top 4 attitudinal items: 梦 "dream", 唱 "sing" and 艺人 "artist". Out of the 21 stories in Chapter One, 8 stories describe how a shepherd becomes a professional singer of Ge Sa-er songs. The title of each story specifies whether the narration is about Ge Sa-er or the shepherd (a bard). Extract (4) retrieved from the corpus explains the translation of attitudinal items in the title.

(4) The source text

[说唱人：牧羊人的梦] (Alai, 2009, p. 12)

(4) Gloss

[The bard to tell and sing: the dream of the shepherd]

(4) The target text

The Storyteller

The Shepherd's Dream (Alai, 2013, p. 14)

Goldblatt and Lin use *storyteller* to translate 说唱人 *shuochangren*. Actually, *shuochangren* means the person who sings the saga to the accompaniment of a *qin* rather than just tells stories. A *qin* is a classic instrument similar to a guitar. However, although the word *storyteller* fails to show the linguistic meaning of *shuochangren*, Goldblatt and Lin narrate the way a bard sings Ge Sa-er stories in another sentence, as shown below.

(5) The source text

“...听老夫抚琴歌唱。” (Alai, 2009, p. 41)

(5) Gloss

“...listen to me, this old fellow singing with a lute.”

(5) The target text

“...listen to this old man sing with his lute.” (Alai, 2013, p. 49)

Extract (5) explains clearly that the story is sung by a singer with an instrument, rather than simply being spoken or passed on in written format. Specifically, the version of the legends of *Ge Sa-er Wang* goes through three approximate stages: songs, xylograph and printed books. As a result, a specific explanation helps target readers to better understand the distinctions between three types of versions. Moreover, Extracts (4) and (5) show that although the title fails to show the specific way a bard uses to tell the saga, explanations in the text demonstrate the traditional way to perform the story. As a result, the English translation conveys source attitudinal meaning in a highly equivalent way.

D. A Translation Model for Social Items

The translation analyses reveal that compared with attitudinal items, the translators face problems to represent the connotated social meaning of SSK- and ideological items, so the inference based on statistical results rather than corpus data can better predict the translation difficulties. The target text can reproduce the underlying attitude of the original text to a high degree, in that attitude is more relevant to the judgement of being positive versus negative, agreement versus disagreement, which is less complicated to express in another language.

The translation analyses point out two translation problems of social discourse: (1) the absence of social meaning and (2) a mixed mythological system in the translation. Although the social meaning of some social items fails to be represented in the TT, target readers can still interpret the meaning according to the context. However, if it is transferred to a special item existing exclusively in target language system, target readers will confuse the source myth with target mythological system. It is therefore essential to address the second problem to facilitate the comprehension of Tibetan culture within the TT.

Since people are more likely to interpret new concepts based on their own existing knowledge system, the translators may apply the target special items occasionally and unconsciously during translating. The mix of source and target mythological systems is a problem of social cognition, which cannot be resolved by applying one specific translation procedure. This problem should be addressed by constructing a systematic cognitive system in the translation. As a result, we propose a translation model to form a consistent mythological system of social discourse for the translation of Chinese ethnic minority/Chinese myths, as shown in Figure 3.

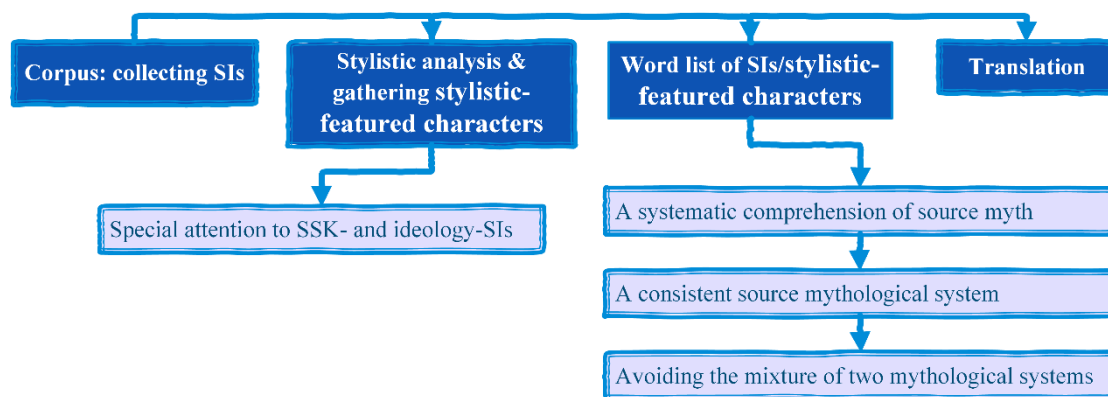


Figure 3. A Translation Model for Social Items

Figure 3 illustrates that all social items are used to form a socio-cultural system of the original text, and the core task is to translate this system. To achieve this goal, the translation model mainly consists of four steps. Firstly, it is suggested to build a corpus which guarantees a relatively more accurate collection of social discourse. In addition, socially stylistic-featured characters can also be gathered in the step. The second step is to integrate statistical approach into corpus to facilitate the stylistic analysis of social discourse. Statistical results tell translators the difficult points to transfer social discourse. Thirdly, a word list of social items or socially stylistic-featured characters can be formulated to explore the way to translate them. Putting social items together is conducive to constructing a systematically mythological framework of the source myth. Translating social items by viewing them as a mythological system may be a good method to prevent the mixing of the source and target mythologies. After all these analyses, the last step goes to translate the whole story.

VI. CONCLUSION

This research denotes that the integration of statistical approaches into corpus can be a significant tool for translation studies. An unexpected finding is that statistical analyses can be used to predict the translation difficulties of social

discourse because it can explain word distribution features and correlations between plots and wording.

Based on the stylistic and translation analyses, we suggest a corpus-based translation model for social discourse in Chinese ethnic minority/Chinese myths, emphasizing the importance to have a systematic understanding of the social discourse in the text. It is important to translate social expressions as an integrated socio-cultural system rather than as individual and separate words. An integrated socio-cultural system will facilitate target readers' interpretation of the translation. In building the translation model, stylistic-featured characters play a significant role, since they are involved in numerous social epithets. 85 stylistic-featured characters (Appendix) of social discourse, which help understand 85.63% of the total social items in the corpus, were gathered for a reference point for the translation of other Chinese myths. It has been shown how these stylistic-featured characters can be referenced to build a social discourse corpus, and how interpreting the underlying social meaning within the items is conducive to translating social discourse within *Ge Sa-er Wang* and other Chinese ethnic minority/Chinese myths.

Although this research seeks to prove that combining corpus and statistical approaches facilitates the stylistic analysis of social discourses in an objective way, and suggests a translation model for Chinese ethnic minority/Chinese myths, other unaddressed questions need to be posed in order to take this research further; these include whether this translation model can be applied to other items in addition to social items, whether such quantitative approaches are suitable for exploring style beyond the specific remit of the social, and what contributions the taxonomy and stylistic features of social discourses may make to building a more powerful corpus. It is hoped that this article will lay a platform for research that can address these and other related questions.

APPENDIX

TABLE 8
THE FREQUENCY OF STYLISTIC-FEATURED CHARACTERS IN EACH PART OF CHAPTER ONE

SI	Part 1	Part 2	Part 3	Part 4	Total
菩萨	22	1	14	50	87
人	182	133	254	194	763
魔	82	18	22	31	153
神	97	81	56	25	259
妖	40	12	15	21	88
天	43	45	71	27	186
师	27	45	16	4	92
咒	8	6	6	1	21
法	20	17	15	16	68
邪	7	4	1	3	15
缘	5	14	2	7	28
噶	19	69	112	36	236
僧	8	0	5	42	55
福	2	5	6	1	14
世	11	6	8	10	35
云	4	9	0	10	23
凡	7	7	0	4	18
庙	0	0	0	49	49
幻	6	1	12	8	27
宝	0	7	0	0	7
佛	0	18	4	48	70
光	7	24	3	9	43
示	9	5	12	7	33
生	12	9	11	11	43
修	5	10	1	2	18
界	22	23	6	7	58
金	1	13	0	5	19
经	1	0	0	9	10
教	3	5	9	8	25
术	6	12	9	1	28
祭	1	0	11	3	15
殿	0	1	0	10	11
祥	6	1	3	3	13
善	0	0	6	3	9
孽	2	1	3	5	11
悟	1	1	0	5	7
命	7	3	5	1	16
施	1	1	4	1	7
赐	2	1	3	0	6

祷	0	1	4	5	10
显	3	1	7	6	17
寺	0	2	0	10	12
德	3	4	0	0	7
魂	2	2	1	0	5
圣	3	0	0	0	3
仙	0	3	0	1	4
供	2	2	0	5	9
灵	12	5	6	3	26
祝	3	1	0	0	4
俗	0	2	0	6	8
授	3	1	1	2	7
卦	0	0	3	0	3
莲	2	1	0	3	6
氏	0	25	0	0	25
王	9	11	72	35	127
下	26	25	16	11	78
上	24	34	47	15	120
国	27	11	43	6	87
领	4	16	28	4	52
苦	15	9	6	4	34
妃	1	9	12	0	22
杀	3	3	23	17	46
族	1	28	6	1	36
龙	6	16	8	1	31
民	3	9	11	3	26
税	0	0	5	4	9
罪	1	0	7	0	8
恶	4	4	5	12	25
降	6	22	12	6	46
争	5	1	13	5	24
德	0	2	0	0	2
拜	0	1	2	4	7
令	0	1	3	2	6
大	52	32	20	9	113
主	0	0	2	4	6
兵	3	10	7	8	28
将	0	2	7	3	12
部	1	23	38	4	66
高	12	8	3	2	25
低	1	0	0	0	1
唱	23	31	6	21	81
盐	4	0	2	0	6
娶	0	6	0	0	6
让	27	31	63	33	154
梦	46	42	26	21	135
Total	1013	1045	1220	953	4231

ACKNOWLEDGEMENTS

The authors would like to appreciate the discussions with Dr Marie-Laure Vuaille-Barcan and Associate Professor Dr Alistair Rolls about social items. The authors are very grateful to Associate Professor Dr Alistair Rolls for his great help to proofread the English. The authors, however, bear full responsibility for the Paper.

REFERENCES

- [1] Alai. (2009). *Ge Sa Er Wang* [King Ge Sa-er]. Chongqin Publisher.
- [2] Alai. (2013). *Ge Sa Er Wang* [King Ge Sa-er] (H. Goldblatt & S. L. Lin, Trans.). Chongqin Publisher. (Original work published 2009)
- [3] Baker, M. (1995). Corpora in translation studies: An overview and some suggestions for future research. *Target*, 7(2), 223-243.
- [4] Biber, D. (1990). Methodological issues regarding corpus-based analyses of linguistic variation. *Literary and linguistic computing*, 5(4), 257-269.
- [5] Biber, D. (2011). Corpus linguistics and the scientific study of literature: Back to the future? *Scientific Study of Literature*, 1(1), 15-23.

- [6] Du, Y. (1994). *Jian Ming Ying Han Han Ying Shi Jie Zong Jiao Ci Dian* [A Concise English-Chinese & Chinese-English Dictionary of Religious Terms Around the World]. China Translation & Publishing House.
- [7] Fitzherbert, G. (2016). Constitutional Mythologies and Entangled Cultures in the Tibeto-Mongolian Gesar Epic: The Motif of Gesar's Celestial Descent. *Journal of American Folklore*, 129(513), 297-326.
- [8] Flowerdew, L. (2012). *Corpora and language education*. Palgrave Macmillan.
- [9] Francke, A. H. (2000). *A Lower Ladakhi Version of the Kesar Saga*. Asian Educational Services.
- [10] Gui, Z. (2012). Shen Hua Chong Shu De Li Shi Zhui Wen Yu Wen Hua Xun Gen – Ping A Lai De *Ge Sa Er Wang* [Historical Inquiry and Cultural Root of Myth Retelling - Comments on *Ge Sa-er Wang* by Alai]. *Xiao Shuo Ping Lun* [Fiction review], 28(1), 167-171.
- [11] Honko, L. (2000). Text as Process and Practice: The Textualization of Oral Epics. In L. Honko (Ed.), *Textualization of Oral Epics 2* (pp. 3-56). Mouton de Gruyter.
- [12] Hornby, A. S. (2004). *Oxford Advanced Learner's English-Chinese Dictionary* (6th ed.). Oxford University Press.
- [13] Hunston, S. (2010). How Can a Corpus Be Used to Explore Patterns? In A. O'Keeffe & M. J. McCarthy (Eds.), *The Routledge Handbook of Corpus Linguistics* (pp. 152-166). Routledge.
- [14] Karmay, S. G. (1995). The Social Organization of Ling and the Term "phu-nu" in the Gesar Epic. *Bulletin of the School of Oriental and African Studies, University of London*, 58(2), 303-313.
- [15] Mahlberg, M. (2013). *Corpus stylistics and Dickens's fiction*. Routledge.
- [16] Mahlberg, M. (2017). Corpus stylistics. In M. Burke (Ed.), *The Routledge handbook of stylistics* (pp. 378-392). Routledge.
- [17] Mahlberg, M., & Wiegand, V. (2020). Literary stylistics. In S. Adolphs & D. Knight (Eds.), *The Routledge Handbook of English Language and Digital Humanities* (pp. 306-327). Routledge.
- [18] McIntyre, D., & Walker, B. (2019). *Corpus Stylistics*. Edinburgh University Press.
- [19] Olk, H. M. (2013). Cultural references in translation: a framework for quantitative translation analysis. *Perspectives*, 21(3), 344-357.
- [20] Onanuga, P. (2018). Exploring Post-Colonial Leadership Crisocial items in Africa: A Linguistic Stylistic Analysis of Osundare's, The State Visit and Two Plays. *Journal of Literary Studies*, 34(1), 133-149.
- [21] Schmidt, I. J. (1839). *Die Thaten Bogda Gesser Chan's, des Vertilgers der Wurzel der zehn Übel in den zehn Gegenden* [The Deeds of Bogda Gesar Khan, the Destroyer of the Root of the Ten Evils in the Ten Regions]. DerKaiserl Akademie der Wissensehaften.
- [22] Shao, L., & Zhou, Y. (2022a). Wen Ben Shi Jie Yu Fan Yi De She Hui Ren Zhi Yan Jiu – Yi A Lai Duan Pian Xiao Shuo *A Gu Dun Ba Ying Yi Wei Li* [A Sociocognitive Approach to Text World and Translation: Illustrated by the English Rendering of Alai's Short Story "Akhü Tenpa"]. *Wai Yu Yan Jiu* [Foreign Language Research], 39(5), 91-96, 103.
- [23] Shao, L., & Zhou, Y. (2022b). Di Yu Xing Yu Yan De Fan Yi Dui Wen Ben Shi Jie Gou Jian De Ying Xiang – Yi A Lai Xiao Shuo Ying Yi Bi Jiao Wei Li [The Influence of the Translation of Regional Characteristic Language on Text-World Construction: The Case of the English Translation of Alai's Novels]. *Wai Yu Yu Wai Yu Jiao Xue* [Foreign Languages and Their Teaching], 327(06), 133-143, 149.
- [24] Short, M. H., & Leech, G. N. (2007). *Style in fiction: A linguistic introduction to English fictional prose*. Pearson Education.
- [25] Stockwell, P., & Mahlberg, M. (2015). Mind-modelling with corpus stylistics in David Copperfield. *Language and Literature*, 24(2), 129-147.
- [26] van Dijk, T. A. (2014). Discourse, Cognition, Society. In J. Angermüller & D. Maingueneau (Eds.), *The Discourse Studies Reader Main Currents in Theory and Analysis* (pp. 121-146). John Benjamins.
- [27] van Dijk, T. A. (2016). Critical Discourse Studies: A Sociocognitive Approach. In R. Wodak & M. Meyer (Eds.), *Methods of Critical Discourse Studies* (2nd ed., pp. 60-87). Sage.
- [28] Wei, L. (2020). *Zhong Hua Da Zi Dian* [Chinese Dictionary]. The Commercial Press.
- [29] Youdale, R. (2017). *Translating Literary Style: Close and Distant Reading in the Translation of Gracias por el Fuego by Mario Benedetti* [Unpublished doctoral dissertation]. University of Bristol, UK.
- [30] Youdale, R. (2020). *Using Computers in the Translation of Literary Style: Challenges and Opportunities*. Routledge.
- [31] Zhou, Y., & Halim, H. A., Ujum, D. A., & Wong, L. Y. (2021). Ideological square in stylistic analysis and its translation of 《格萨尔王》 *Ge Sa-er Wang*: Negative ingroups and positive outgroups. In H. A. Halim & L. H. Ang (Eds.), *Proceeding from MICFL2021: the 7th Malaysia International Conference on Foreign Languages*. University Putra Malaysia.

Yuqiao Zhou is a jointly awarded PhD candidate in Translation and Interpretation at University Putra Malaysia and in Modern Languages at the University of Newcastle, Australia. She is Tutor, teaching Translation, at Guangdong Innovational and Technical College, China. Her research interests are in literary translation studies and corpus stylistics. Email: GS59613@student.upm.edu.my/c3388090@uon.edu.au. ORCID: 0000-0001-9027-6964

Hazlina Abdul Halim (corresponding author) is Associate Professor in the Faculty of Modern Language and Communication, University Putra Malaysia. She has written chapters for 15 books, and has published over 60 articles in international journals. Her research interests are in literary translation and French applied linguistics. Email: hazlina_ah@upm.edu.my. ORCID: 0000-0003-3599-9195

Diana Abu Ujum is Senior Lecturer in the Faculty of Modern Language and Communication, University Putra Malaysia. Her research interests are in literary translation and contemporary fiction. Email: a_diana@upm.edu.my. ORCID: 0000-0002-6080-7565

Ling Yann Wong is Senior Lecturer in the Faculty of Modern Language and Communication, University Putra Malaysia. Her research interests are in literary translation and Chinese historical linguistics. Email: wonglynn@upm.edu.my