

A Corpus-Based Study on the Translation of Jia Baoyu's Character Image in *Hong Lou Meng*

Ying Li

Tongji University, Shanghai, China

Abstract: Based on a self-built corpus, this paper adopts both quantitative and qualitative methods to conduct a systematic comparative analysis of the translation of Jia Baoyu's character image. The research focuses on the original text of *Hong Lou Meng* as well as its two English versions translated by Yang Xianyi & Gladys Yang and by David Hawkes & John Minford respectively, examining multiple dimensions including lexis, syntax, punctuation, and descriptions of characters' behaviors and appearances. The study reveals that the Yang version generally demonstrates a reproducing translation orientation. It tends to be close to the original expression at the lexical and syntactic levels with concise language, retains many structural features of Chinese in the translation of address forms, and presents a relatively implicit style in characters' dialogues. It also shows considerable restraint in the use of punctuation and the depiction of actions, thus preserving the narrative style of the original text to a certain extent. In contrast, the Hawkes version is characterized by an enhancing translation approach. On the basis of faithfulness to the original content, it employs more flexible and diverse translation strategies: the characters' utterances are more colloquial, the address forms and syntactic structures are more in line with English usage conventions, and the frequent use of exclamation marks strengthens emotional expression. In the descriptions of actions and appearances, Hawkes enriches details through amplification and domestication, enhancing the perceptibility and expressiveness of the character image. Overall, the differences in strategic choices across multiple dimensions between the two versions have shaped distinct pathways of character construction, reflecting the multi-dimensional mechanism of character image formation in literary translation.

Keywords: *Hong Lou Meng*, Corpus-based translation studies, Character image, Jia Baoyu.

1. INTRODUCTION

Hong Lou Meng is one of the pinnacles of classical Chinese fiction and also holds an important position in world literature. The work is renowned for its sophisticated and elaborate character system as well as its exquisitely detailed descriptions. Through superb narrative art, Cao Xueqin endows the characters with a strong sense of authenticity and three-dimensionality in their speech, behavior and emotions.

Research on *Hong Lou Meng* and its English translations has a long history with abundant academic achievements. Existing studies have mostly been conducted from the perspectives of translator's style, linguistic features and cultural connotations. Especially in the past two decades, the introduction of corpus-based methods has driven the shift of research from experiential judgments to data-driven empirical analyses. However, compared with macro-level translation comparisons or investigations of linguistic features, systematic research on the translation of character images remains insufficient. Most current studies focus on partial analyses at the level of character dialogue or vocabulary, and rarely conduct a comprehensive investigation into the overall construction mechanism of character images from multi-dimensional linguistic resources.

In view of this, this paper takes Jia Baoyu, the central character of the novel, as the research object. Relying on a self-built corpus and combining quantitative and qualitative methods, it compares and analyzes the version by Yang Xianyi & Gladys Yang (hereafter referred to as "the Yang version") and the version by David Hawkes & John Minford (hereafter referred to as "the Hawkes version") from multiple aspects, including characters' discourse, syntactic features, and descriptions of actions and appearances. This paper aims to reveal how different translation strategies participate in the reconstruction of character images through multi-dimensional coordination, so as to deepen the understanding of character construction mechanisms in literary translation and provide new analytical approaches and empirical support for literary translation research from a corpus perspective.

2. LITERATURE REVIEW

Since the advent of English translations of *Hong Lou Meng*, research on its translation has been developing

continuously, with diverse research orientations formed at different stages. Generally speaking, existing studies can be roughly divided into three main directions: macro-level translation comparison, character translation research, and the application of corpus-based approaches.

First, in macro-level studies on the translated versions, scholars including Wang Hongyin (2002) and Feng Qinghua (2006, 2008) have carried out systematic investigations into the English translation of *Hong Lou Meng* from the perspectives of translation style, linguistic features and cultural connotations. Most of such studies adopt a holistic textual perspective to conduct comprehensive reviews of translators' strategies and version characteristics, laying a fundamental framework for research on the English translation of *Hong Lou Meng*. Nevertheless, these studies tend to emphasize macro-level description, paying insufficient attention to the micro-construction mechanisms of specific character images, and in particular lack an integrated analysis of characters' multi-dimensional expressive resources such as discourse, behavior and appearance.

Second, although relevant achievements have been accumulated in the field of character translation research, such studies are still characterized by scattered research objects and single-dimensional analysis on the whole. For instance, Liu Zequan (2010, 2012) built a self-designed corpus and took Grandmother Liu as the research subject. From the perspective of interpersonal metafunction in systemic functional linguistics, he made a detailed analysis of differences in the representation of character speech in different translations, demonstrating the research potential of combining corpus methods with functional linguistics. However, this study mainly focused on a single character and the dimension of dialogue, without further extending to other representational dimensions such as behavior and appearance. Peng Wenqing (2016) explored the translation differences of commendatory and derogatory terms in *Hong Lou Meng* and *Romance of the Three Kingdoms*, revealing how translators' lexical choices influence the evaluative meaning of character images. Yet her research centered on value orientations at the lexical level, lacking a systematic perspective on the overall construction of characters. Liao Chunxia (2019) analyzed character dialogue at the lexical and syntactic levels and explored discourse differences in combination with gender, personality and social status, expanding the depth of research on character speech. Nevertheless, her analysis was still largely confined to the linguistic dimension and failed to cover non-verbal elements such as actions and physical appearances.

In addition, Dang Zhengsheng (2012) devoted a special chapter in his monograph to the translation of character language, and Feng Qinghua (2006) also selected multiple characters to analyze the linguistic individuality of figures in the novel. Although these studies have recognized the importance of character differentiation, they still take "linguistic representation" as the core and lack a systematic integration of the multi-modal construction mechanisms of character images.

Third, at the methodological level, the introduction of corpus linguistics has provided a new path for literary translation research. Scholars such as Yang Huizhong (2002) and Hu Kaibao (2017, 2021) have demonstrated the effectiveness of corpus-based methods in translation studies, gradually shifting research from intuitive judgments to data-driven empirical analysis. However, most existing corpus-based studies focus on the statistics of linguistic features or explorations of translation universals, and lack systematic and in-depth application to how core character images in classical Chinese novels are reconstructed in the process of translation.

As for overseas research, relevant discussions mainly focus on the evaluation of translated versions and translation strategies. For example, McDougall's evaluation of the Hawkes translation and Stephen Owen's exploration of translation strategies for classical Chinese fiction have both provided important references from the perspectives of reception and cultural transformation. Nevertheless, such studies mostly remain at the macro level and pay limited attention to the translation mechanisms of specific character images.

To sum up, although existing studies have carried out beneficial explorations into the English translation of *Hong Lou Meng* from different perspectives, they still have the following deficiencies. First, most research targets individual characters or partial linguistic phenomena, lacking systematic and in-depth investigations into core characters such as Jia Baoyu. Second, in terms of analytical dimensions, most studies concentrate on the level of character speech and rarely integrate elements such as actions and appearances into a unified analytical framework. Third, methodologically, qualitative and quantitative studies are often separated from each other without organic integration.

In view of this, this paper takes Jia Baoyu, the "rebellious core character", as the entry point. Relying on a self-built corpus and combining quantitative and qualitative methods, it conducts a comprehensive investigation into

the character construction mechanisms in different translations from multiple dimensions including discourse, syntax, actions and appearances, so as to make up for the deficiencies of existing research in research object selection, analytical dimensions and methodological integration.

3. RESEARCH DESIGN

3.1 Research Data

The data for the small-scale self-built corpus in this study are drawn from the original Chinese text of *Hong Lou Meng* and its two English versions, namely the Yangs' version and the Hawkes-Minford version. The Yangs' version is *A Dream of Red Mansions*, published by Foreign Languages Press in 2003, the Hawkes-Minford version is *The Story of the Stone*, published by Shanghai Foreign Language Education Press in 2012. As this study aims to explore the representation and construction of Jia Baoyu's character image in translation, the data are selected in accordance with the principle of "character-generating elements", which specifically include: character utterances, descriptions of character behaviors (covering both active and passive behaviors), character portrait descriptions, and character psychological descriptions. These elements together constitute the core dimensions of character formation in literary narration, thus providing an interpretable data foundation for subsequent analyses.

During the collection of utterance data, instances were found where the translations have adjusted the original quotation patterns, such as converting direct speech into indirect speech or even omitting parts of the utterances. In response to this issue, this study takes the original text as the criterion for identification, and conducts unified extraction and alignment of contents identified as Jia Baoyu's utterances. This ensures the comparability of different translated versions in terms of "utterance attribution", laying a foundation for the analysis of character discourse construction.

3.2 Research Content and Research Questions

One of the defining features of literary works is their figurativeness, and in narrative texts, this feature is mainly embodied through the construction of character images (Hu, 2021: 163). As one of the central figures in the narrative structure of *Hong Lou Meng*, Jia Baoyu is also a vital character embodying emotional values and anti-Confucian ethical consciousness. His personality combines naivety and obstinacy with delicacy and sensitivity, reflecting both his alienation from and rebellion against feudal ethical orders and his highly sensitive emotional experiences. Therefore, how the translated versions represent this complex character will directly affect the target-language readers' understanding of the ideological connotations and aesthetic values of the work.

From the perspective of narrative construction, character images are not formed in a single dimension, but are co-constructed through multiple layers such as discourse, behavior, portrait and psychological descriptions (Hu, 2021: 169). On this basis, this study operationalizes the investigation into the translation of Jia Baoyu's character image into three interrelated dimensions: discourse construction, behavior construction, and descriptive construction. By combining corpus methods with close textual reading, it analyzes the differences between the two versions in the above dimensions and further explores how such differences shape the representation of the character image.

Specifically, this study attempts to answer the following research questions:

(1) In the translation of character behaviors and portrait descriptions, are there systematic differences in the translation strategies (e.g., amplification, omission, domestication, etc.) adopted by the two translators? To what extent do these differences alter the presentation of the character's behavioral patterns and external image?

(2) Based on the integrated analysis of discourse, behavior and description, what types of Jia Baoyu images are constructed in each of the two translations? Is there a characterization shift between the two versions, and if so, what are its specific manifestations?

3.3 Research Tools and Methods

This study adopts an integrated approach of quantitative and qualitative analysis. First, a small-scale parallel corpus is manually constructed, involving data collection, text cleaning and manual alignment of the original Chinese text of *Hong Lou Meng* and its English translations, so as to form a comparable data foundation.

On this basis, multi-dimensional statistical analyses of the corpus are conducted with the help of WordSmith 8.0 and AntConc 3.5.8, focusing on such indicators as Standardized Type-Token Ratio (STTR), average sentence length, high-frequency word distribution, and punctuation use. It should be noted that this study does not merely describe differences at the level of linguistic form, but attempts to establish an interpretive connection between “corpus data – linguistic features – character image construction”. For instance, lexical diversity can reflect the expressiveness and mental complexity of a character’s speech, syntactic structure and sentence length variation can reveal the rhythm and intensity of emotional expression, and the use of punctuation helps uncover the emotional orientation and tonal features of character utterances.

On the basis of quantitative analysis, this study further carries out qualitative analysis with typical examples, focusing on the translators’ handling of character portrait and behavior descriptions and the strategic choices behind such treatments. Finally, through a comprehensive investigation of the three dimensions of discourse, behavior and description, an overall evaluation is made of the ways in which Jia Baoyu’s character image is represented in the two translated versions.

4. COMPARATIVE ANALYSIS OF LINGUISTIC FEATURES IN CHARACTER DISCOURSE

The discourse system of *Hong Lou Meng* is exemplary in ancient Chinese fiction. In terms of length, the total volume of character dialogue exceeds 300,000 Chinese characters, accounting for nearly half of the entire novel, which demonstrates its central position in the narrative structure. Functionally, character discourse not only portrays personality traits through lexical choice and syntactic organization, but also reveals the relational structure, social identities and psychological states among characters via tone, rhetoric and modes of expression, while fulfilling the narrative function of advancing the plot. Therefore, the translation of character discourse is not merely a matter of linguistic transfer, but also a crucial part in the representation of character images.

Against this background, this paper examines the translation of Jia Baoyu’s discourse from the lexical and syntactic levels, and attempts to establish an interpretive pathway of “linguistic features – expressive modes – character images”, so as to reveal the differences between the two translations in the construction of character discourse and their implications for character portrayal.

4.1 Lexical Comparison

4.1.1 Standardized Type-Token Ratio (STTR)

A type refers to a distinct lexical item in a text, while a token refers to the total number of occurrences of lexical items (Baker, 1995). The Type-Token Ratio (TTR) is commonly used to measure lexical richness and diversity in a text (Yang, 2002: 153). Given that TTR is susceptible to text length, this study adopts the Standardized Type-Token Ratio (STTR) to improve comparability across different texts.

Statistical analysis of the corpus using WordSmith 8.0 yields the following results: Jia Baoyu’s discourse in the original text amounts to 37,329 Chinese characters, corresponding to 26,579 words in the Yang version and 35,586 words in the Hawkes version. In terms of total word count, the Hawkes version is significantly longer than the Yang version and closer to the length of the original text. To eliminate interference from individual cases, this study further selects the discourse of Shi Xiangyun and Grandmother Liu as reference data. The statistical results also show a consistent trend that the word count in the Hawkes version is generally higher than that in the Yang version (Shi Xiangyun: original 5,539 characters, Yang version 3,990 words, Hawkes version 5,086 words, Grandmother Liu: original 4,388 characters, Yang version 3,267 words, Hawkes version 4,221 words). This phenomenon indicates that in the translation of character discourse, the Yang version overall tends to be concise, whereas the Hawkes version prefers to expand expressions for full semantic articulation.

Further investigation into the STTR index shows that the Yang version registers 42.67 and the Hawkes version 41.23. Under the same text length condition, the Yang version displays slightly higher lexical diversity than the Hawkes version. On the surface, this result seems to contradict the observation that the Hawkes version contains more words, yet in fact it reflects two distinct translation orientations: the Yang version achieves relatively diverse lexical distribution within a more limited length, while the Hawkes version increases lexical quantity to enhance expressive concreteness and semantic explication.

From the perspective of character discourse construction, this difference carries significant implications for portrayal. Higher lexical diversity usually signifies conciseness and generality of expression, helping to preserve the implicit features of the original discourse, the increase in lexical items and expanded expression, by contrast, better reinforce semantic clarity and emotional explicitness. Accordingly, in constructing the image of Jia Baoyu, who is characterized by both sensitivity and delicacy as well as rebelliousness and uninhibitedness, the Yang version tends to maintain the implicit expression of the original through relatively concise yet varied lexical choices, lending a degree of restraint to his speech. The Hawkes version, through lexical expansion and elaboration, renders the character's discourse more concrete and explicit, thereby strengthening the perceptibility of emotional expression to a certain extent.

In other words, the disparities in STTR and word count not only reflect the translators' different strategies in language use, but also shape the expressive modes of character discourse, ultimately contributing to the construction of Jia Baoyu's image: the former leans toward an "introverted – implicit" discourse style, while the latter approximates an "explicit – concrete" mode of expression. This difference provides important quantitative evidence for the subsequent analysis of the character's emotional expression and personality representation.

4.1.2 High-Frequency Words

High-frequency words refer to lexical items whose occurrence frequency in a given corpus is markedly higher than that of other words. Within a corpus-based framework, high-frequency words not only reflect the basic lexical makeup of a text but also reveal tendencies in language use. In general corpus analysis, high-frequency words are mostly function words, such as articles, prepositions and pronouns in English. In the study of character discourse, however, the analytical significance of high-frequency words extends beyond the grammatical level, lying more in the discourse orientation and semantic focus they embody.

Accordingly, in this study, on the basis of high-frequency word statistics, emphasis is placed on semantically loaded items, especially those that reflect character relationships, addressees and emotional orientation. The high-frequency words obtained through statistical analysis of the corpus using AntConc 3.5.8 are presented in Table 1 (not fully listed due to space constraints).

From an overall distributional perspective, the high-frequency words in both translations contain numerous personal pronouns (e.g., *I*, *you*, *we*, *me*) and anaphoric expressions (e.g., *her*, *it*), indicating that Jia Baoyu's discourse is highly interactive and oriented toward interpersonal communication. This is consistent with the character trait of Baoyu in the original novel, who prioritizes qing (human affection) and values emotional bonds with others.

Further observation reveals that both versions feature prominent female-referring terms among high-frequency words, such as *her*, *she* and address terms related to female identity. This phenomenon suggests that Jia Baoyu's utterances largely revolve around female characters, reflecting the gender orientation of his emotional focus. Such a female-centered discourse structure constitutes a linguistic manifestation of his defining personality trait: affinity toward women and alienation from patriarchal male norms.

At the specific lexical level, the two translations display interpretively significant differences. The high-frequency words in the Yang version include *sister* (134 occurrences), whereas the Hawkes version features *cousin* (316 occurrences) as a representative address term. Contextual review of the original text shows that both terms mostly correspond to the Chinese term *jiejie* (elder sister). The two translators, however, adopt divergent strategies in rendering this address: the Yang version prefers *sister*, a relatively generalized term carrying intimate connotations, while the Hawkes version tends to use *cousin* to specify the kinship ties between characters.

This divergence is more than a mere lexical choice, it reflects two distinct models of relational construction. Specifically, although *sister* in English can refer to close female companions without blood ties, its semantic core lies in emotional closeness, thereby blurring rigid boundaries of family structure and rendering interpersonal relations more affective and less institutionalized. By contrast, *cousin* clearly marks intra-clan kinship, strengthening the characters' familial identity positioning and lending relations a more structural and social quality.

From the perspective of character construction, such differences in address terms directly shape the representation of interpersonal relations in Jia Baoyu's discourse. The use of *sister* in the Yang version to some extent reinforces Baoyu's emotional identification and affectionate attitude toward women, rendering his speech mild and intimate,

and thus highlighting his disposition of valuing affection over ritual. The use of *cousin* in the Hawkes version, while maintaining clarity in the network of character relations, embeds Baoyu's discourse more firmly within the clan structure, somewhat weakening the unregulated and rebellious quality of his emotional expression.

In addition, the concentrated appearance of female-referring terms and kinship address terms among high-frequency words collectively forms a stable pattern of discourse orientation: Jia Baoyu's utterances consistently center on the female group. This linguistic feature is reflected not only in his choice of addressees but also reinforces the value orientation at the heart of his character, which takes women as the emotional center. In other words, high-frequency words reveal not merely a lexical distribution but a structural discourse-focusing mechanism. This mechanism is retained or reconstructed to varying degrees in the translation process, thereby influencing target-language readers' overall perception of Jia Baoyu's image.

In summary, the analysis of high-frequency words demonstrates that the differences between the Yang version and the Hawkes version in character address and reference systems are not confined to the lexical level but fundamentally reflect two distinct strategies of discourse construction. By shaping the representation of interpersonal relations and emotional orientation, these differences further contribute to the reconstruction of Jia Baoyu's character image, providing important linguistic evidence for the subsequent analysis at the syntactic and tonal levels.

Table 1

	Yang's version		Hawkes' version	
	word	frequency	word	frequency
1	I	762	THE	1096
2	YOU	629	TO	924
3	AND	613	I	837
4	THE	594	A	757
5	A	528	AND	648
6	HAVE	429	YOU	629
7	OF	343	IT	571
8	THAT	307	HER	536
9	WE	258	WAS	487
10	ME	249	SAID	405
11	IT	214	COUSIN	316
12	YOUR	178	MY	302
13	BUT	173	OF	278
14	SISTER	134	YOUR	273

4.1.3 Address Terms

As an important component of character discourse, address terms serve as a highly condensed expression of interpersonal relationships and emotional attitudes. Unlike ordinary vocabulary, address terms not only perform a referential function but also carry rich sociocultural information and emotional orientation. Their choice often directly reflects the speaker's perception of the interlocutor's identity and emotional stance. Therefore, in literary translation, the handling of address terms involves not only linguistic transfer but also the accuracy of representing interpersonal relationships and constructing character images.

From a cross-cultural perspective, the Chinese address system is distinctly relation-oriented, emphasizing generational hierarchy, intimacy, and social status. In contrast, English address terms tend to be individual-oriented, weakening kinship structures and prioritizing communicative functions in many contexts (Feng, 2006). Accordingly, the translation of address terms inevitably entails cultural adaptation, and different approaches reflect the translator's trade-off between relationship retention and naturalness of expression.

A comparison between the Yang version and the Hawkes version reveals consistent divergent tendencies in their treatment of address terms. On the whole, the Yang version leans toward preserving the original address structure, reproducing the network of interpersonal relationships through relatively literal translation. The Hawkes version, by contrast, prioritizes naturalness in the target language and, without compromising comprehension, applies a certain degree of domestication to address terms to bring them in line with English pragmatic conventions.

This difference is particularly evident in the following examples:

Example (1)

Original: 宝玉便问道：“周姐姐，你作什么到那边去了。”

Yang: What were you doing over there, Sister Zhou?

Hawkes: What were you doing at my aunt's place, Zhou?

Example (2)

Original: 宝玉笑道：“好哥哥，你别委曲，我明儿请你。”李贵道：“小祖宗，谁敢望你请，只求听一句半句话就有了。”

Yang: “Cheer up, good brothers,” replied Baoyu with a smile. “I’ll give you a treat tomorrow.”

Hawkes: ‘Don’t be upset, old chap!’ said Bao-yu. Tomorrow, I’ll treat you all.’

In Example (1), the original address *Zhou Jiejie* (Sister Zhou) not only identifies the character but also conveys tones of intimacy and respect. By using *Sister Zhou*, the Yang version directly reproduces the formal structure of the original address, retaining the Chinese practice of extending kinship terms to mark social relations. In comparison, the Hawkes version simplifies the address to *Zhou* and supplements contextual information (*my aunt's place*) to maintain referential clarity. This approach somewhat dilutes the emotional content inherent in the address itself, shifting interpersonal relations from “affectively charged address” to “contextually grounded reference.”

The difference in Example (2) is even more representative. The original *hao gege* (good brother) is a typical affective address: it corresponds neither precisely to actual kinship nor serves as a mere generic term, but functions in specific contexts to express closeness and reassurance. The Yang version achieves formal correspondence with *good brothers*, preserving the original address connotation as much as possible. The Hawkes version, however, uses *old chap*, a common colloquial expression in English, converting it into a conventional utterance with communicative function. While this transformation enhances the naturalness of the translation, it also replaces the original kinship-metaphorical address with a generalized interpersonal term.

From the perspective of discourse construction, such differences exist not only at the level of translation strategy but also directly shape the representation of relationships in Jia Baoyu's speech. By retaining address structures such as *jiejie* (sister) and *gege* (brother), the Yang version reconstructs a relationship network bonded by emotion, blurring boundaries between kinship and social acquaintance. This renders Baoyu's discourse more intimate and warmer, reinforcing his affectionate disposition. Through domestication, the Hawkes version adapts address terms to English pragmatic norms, making interpersonal relations clearer and more standardized, yet at the same time somewhat reducing the emotional density established by address terms in the original.

Furthermore, variations in address terms also bear on the representation of the character's anti-normative traits. In the original novel, Jia Baoyu frequently uses kinship addresses for non-immediate family members or even non-blood relations, a usage that itself loosens and reconstructs traditional hierarchical order. The Yang version retains this linguistic feature through formal preservation. The domestication in the Hawkes version, however, normalizes such transgressive expressions to a certain extent, potentially weakening the anti-Confucian implications embedded in Baoyu's discourse.

In summary, the translational differences in address terms reflect two distinct strategies of relational construction: one oriented toward structural preservation, maintaining the cultural characteristics of interpersonal relations by reproducing the original address system, the other oriented toward pragmatic adaptation, adjusting address forms to enhance the naturalness of the target text. These two strategies respectively shape the emotional expression and relational presentation of character discourse, ultimately contributing to the reconstruction of Jia Baoyu's character image.

4.2 Comparison of Syntactic Categories

Compared with the lexical level, syntactic structure has a more direct impact on the organization and expressive rhythm of discourse, serving as a key indicator of a character's speech style. In literary translation, syntactic choices not only affect the fluency and readability of the target text but also largely shape the emotional intensity and mode of expression of character utterances. Therefore, differences at the syntactic level can be regarded as

one of the important means of character discourse construction.

Against this background, this paper conducts a comparative analysis of the syntactic features in the Yang version and the Hawkes version in rendering Jia Baoyu's discourse from three aspects: average sentence length, quotation patterns, and punctuation use, and further explores their influence on the representation of the character image.

4.2.1 Average Sentence Length

Average sentence length is an important index for measuring syntactic complexity and information-carrying capacity. Generally speaking, longer sentences correlate with more complex syntactic structures and more hierarchical information organization, whereas shorter sentences tend to feature compact rhythm and direct expression (Hu, 2017: 230). Accordingly, average sentence length can reflect not only the translator's linguistic style but also their strategic choices in information organization and discourse presentation.

According to the statistical results from WordSmith 8.0, the average sentence length of Jia Baoyu's discourse in the Yang version is 11.76, compared with 14.21 in the Hawkes version, the latter being significantly longer. This difference indicates that the Hawkes version prefers longer syntactic structures to accommodate more information units, while the Yang version tends to use relatively concise sentences, resulting in strong rhythmicity and compactness of expression.

Further textual analysis reveals that this disparity is closely related to the two translators' approaches to handling the information in the original text. In the translation process, the Hawkes version often explicates implicit information in the original through syntactic expansion, such as supplementing culturally loaded items (e.g., address terms, idioms) or integrating multiple semantic units via clause structures, thus endowing the target text with high information density and explanatory power. In contrast, the Yang version tends to preserve the original information distribution with minimal syntactic expansion, retaining to some extent the paratactic features of Chinese and presenting a succinct and condensed style.

From the perspective of character discourse construction, differences in sentence length are not merely formal variations but directly influence the rhythm and emotional expression of the utterances. Shorter sentences carry greater immediacy and colloquialism, making the character's speech more natural and direct, which helps convey the sensitivity and candor in Jia Baoyu's emotional expressions. Longer sentences, by integrating and elaborating information, achieve more complete and explicit expression, yet may somewhat reduce the intensity of fleeting emotions in the original.

It can therefore be argued that the relatively concise sentences in the Yang version preserve to a certain extent the characteristic of Jia Baoyu's speech as "emotion stirred within and uttered at once", lending his discourse a sense of spontaneity and emotional fluctuation. The Hawkes version, through syntactic expansion and structural integration, renders the character's utterances more organized and narrative, enhancing comprehensibility while also imparting a degree of mitigation and rationalization to emotional expression.

4.2.2 Quotation Patterns

Quotation patterns constitute a key part of the presentation of character discourse and directly bear on the boundary between narrative voice and character voice. In fictional narration, direct speech usually preserves the original tone and emotional tension of a character's utterances, whereas indirect speech embeds the discourse within the narrative context via the narrator's rendition, thereby weakening its immediacy and individuality to a certain extent.

Manual statistics show that the original text contains 469 instances of direct speech. The Yang version retains 449 of these, with only a small number (18) of speech-type shifts, demonstrating a high degree of formal correspondence overall. In contrast, the Hawkes version shows 42 instances of speech-type conversion, revealing a more flexible handling tendency.

Further analysis reveals that in certain contexts, the Hawkes version tends to convert short, loosely structured direct speech from the original into indirect speech to achieve coherence and naturalness in textual flow. For example:

Table 2

Treatment of Speech	Yang's Version	Hawkes' version
Retain original speech pattern	449	417
Shift speech pattern	18	42
Omit speech	2	10

Example (3)

Original: 宝玉听说，千姐姐万姐姐的央求：“只住一日，明儿必回去的。”

Hawkes: he promised faithfully to return on the morrow without demur.

Example (4)

Original: 宝玉因让“林妹妹吃茶。”

Hawkes: Bao-yu invited Dai-yu to have some tea.

In the above examples, the original uses direct speech to present the character's immediate verbal behavior, marked by distinct colloquialism and situationality. By contrast, the Hawkes version incorporates the character's utterances into the narrative frame through reported speech, resulting in smoother expression yet somewhat reducing the independence of the character's voice.

From the perspective of character construction, the choice of quotation pattern directly affects the reader's perception of the immediacy of character discourse. By retaining a large number of direct speeches, the Yang version presents Jia Baoyu's utterances to the reader with minimal mediation, strengthening the directness of his voice and the immediacy of his emotions, which helps portray his candid and sentimental side. By appropriately increasing indirect speech, the Hawkes version partially transforms character utterances into narrative content, enhancing narrative fluency while somewhat weakening the expressive force of the character's speech.

It can thus be argued that the differences between the two quotation patterns reflect distinct narrative orientations: the former emphasizes the direct presentation of character voice, while the latter leans toward discourse reconstruction under narrative integration. This divergence further shapes the representation of Jia Baoyu's image, endowing him with more pronounced traits of an emotional subject in one version and of a narrative role in the other.

4.2.3 Punctuation

In literary translation, punctuation is not merely a reflection of writing conventions but also an important device for discourse organization and emotional expression. As a supra-syntactic resource distinct from lexicon and syntax, punctuation directly acts on tone and discourse rhythm, thereby influencing the reader's perception of the character's emotional state. Consequently, differences in punctuation usage carry considerable significance for character construction in the translation of character discourse.

Previous studies have noted that the selection and variation of punctuation often reflect the translator's individual style and aesthetic orientation (Wan, 1991). In this study, a comparative analysis of the original text and the two translations reveals significant differences between the Yang version and the Hawkes version in the use of exclamation marks.

Statistical results show that Jia Baoyu's discourse in the original contains 69 exclamation marks, the Yang version uses 58, slightly fewer than the original, while the Hawkes version uses as many as 127, notably higher than the original and nearly twice the number in the Yang version. This disparity indicates that the two translators adopt different strategies regarding the explicitness of emotional expression in character discourse.

Specifically, as a punctuation mark that directly signals emotional intensity, a higher frequency of exclamation marks usually corresponds to more explicit and intensified emotional expression. The abundant use of exclamation marks in the Hawkes version lends more distinct emotional fluctuations to the tone of the character's speech, enhancing its expressive appeal and dramatic effect. This approach amplifies the emotional coloring in Jia Baoyu's utterances to a certain extent, presenting him as more extroverted and unrestrained in expression.

In comparison, the Yang version uses fewer exclamation marks than the Hawkes version, even slightly fewer than

the original, indicating a relatively restrained strategy at the punctuation level. By reducing explicit emotional markers, the Yang version maintains the comparatively implicit mode of expression in the original, allowing the character's emotions to be conveyed more through context and lexicon itself, resulting in a more introverted and restrained discourse style.

Combining this discrepancy with the earlier syntactic analysis reveals internal consistency: the Hawkes version tends toward syntactic expansion and information integration at the syntactic level, and strengthens emotional expression by increasing exclamation marks at the punctuation level, together forming an explicit and intensified discourse construction model. The Yang version, by contrast, adopts relatively concise sentences and restrained punctuation, creating a condensed and implicit mode of expression. These two models respectively shape the emotional presentation of character discourse, bringing different facets of Jia Baoyu's personality to the fore in each translation.

From the perspective of character construction, differences in punctuation further influence the reader's understanding of Jia Baoyu's emotional traits. The intensified use of exclamation marks in the Hawkes version lends greater emotional tension and expressiveness to Baoyu's speech, highlighting his overtly emotional side. The relatively restrained punctuation strategy in the Yang version renders his emotional expression more reserved and implicit, closer to the original feature of profound affection without full disclosure.

This study therefore finds that punctuation is not a mere formal element but an important resource participating in the construction of character discourse. The divergence between the two translators in the use of exclamation marks essentially reflects different orientations toward the explicitness of emotional expression, ultimately shaping the representation of Jia Baoyu's character image and giving him differentiated traits of explicit intensification on the one hand and reserved implicitness on the other.

5. COMPARATIVE ANALYSIS OF THE TRANSLATION OF CHARACTER ACTIONS

The depiction of character actions constitutes one of the important means of constructing character images in literary narration. Compared with direct expression at the discourse level, actions and behaviors often present personality through showing rather than telling, conveying emotional attitudes and psychological states without explicit statement. In *Hong Lou Meng*, Cao Xueqin employs delicate and layered action descriptions to endow characters with strong dynamism and complexity. Cao Xueqin devoted meticulous effort to character portrayal, which is particularly evident in the detailed depiction of each character's actions. "Only by examining the text of *Hong Lou Meng* from a literary perspective can we achieve a proper interpretation of the ideological and artistic value of this great literary work" (Zhang, 2007: 259). Investigation into the text reveals that Jia Baoyu's utterances are frequently introduced by specific behavioral markers, among which "笑道" (*said with a smile*) is especially frequent and representative. This expression not only serves as an introductory marker of speech but also acts as a crucial medium connecting the character's emotional state and verbal behavior. Therefore, the translation of "笑道" involves not only lexical transfer but also the representation of the character's emotional expression and personality traits.

Meanwhile, comparative analysis also indicates that the two translations display systematic differences in the presentation of action details, which further contribute to the construction of Jia Baoyu's character image.

5.1 Translation of "笑道"

"笑道" is a typical speech-reporting structure in *Hong Lou Meng*, occurring 2,201 times in the first eighty chapters (Liu & Yan, 2010). Functionally, "笑道" not only identifies the speaker but also adds emotional coloring to the utterance through the component "smile," lending specific tone and attitude to the character's speech. It can thus be regarded as an affective discourse-introducing mechanism, performing multiple pragmatic functions in different contexts, such as greeting, bantering, comforting, and irony.

"笑道" appears 196 times in Jia Baoyu's lines. Statistical results show that the Yang version fully retains the dual structure of "smile and speak" in 127 instances, whereas the Hawkes version does so in 164 instances. At the same time, approximately one-third of instances of "笑道" in the Yang version are omitted or simplified, using zero translation or retaining only the verb of saying. In contrast, the Hawkes version makes similar adjustments in only

32 cases, demonstrating a higher overall retention rate.

On the surface, this difference might suggest that the Hawkes version is formally more “complete.” Further analysis, however, reveals that the two translators adopt distinct narrative strategies in rendering “笑道”.

First, in terms of structural choice, the Hawkes version tends to follow standardized English narrative patterns, using structures such as *said Bao-yu and adverb* (e.g., *said Bao-yu entreatingly*, *said Bao-yu encouragingly*). This approach conforms to the narrative conventions of English fiction, supplementing the tone of the utterance with adverbs to explicate emotional information while maintaining syntactic fluency. Nevertheless, such adverbial treatment to some extent converts emotions originally conveyed through actions into narrative labels, shifting emotional expression from behavioral showing to semantic stating.

In comparison, the Yang version exhibits greater diversity in rendering “笑道”. In addition to the conventional *said and adverb* structure, it extensively employs semantically specific verbs such as *grin*, *chuckle*, *tease*, *cry* to directly represent the action of smiling. By differentiating types of smiles through the inherent semantics of these verbs, this method preserves and even strengthens the expressiveness of action description in the target text.

From the perspective of character construction, this difference is highly significant. As a pivotal link connecting action and speech, the translation of “笑道” directly shapes the pathway of emotional expression. The Hawkes version, through standardized structures and adverbial modification, renders emotional expression clearer and more interpretable, yet at the cost of weakening the expressive function of the action itself, leaving character emotions largely dependent on narrative labeling. By contrast, the Yang version uses varied verbs to represent different forms of smiling, allowing emotions to emerge more directly from behavior, thus enhancing the vividness and dynamism of the character’s portrayal.

Furthermore, this divergence relates to the representation of different facets of Jia Baoyu’s personality. In the original novel, Baoyu frequently uses smiling as a key mode of emotional expression, and this mechanism of “expressing emotion through smiling” is inherently implicit and polysemous. By retaining and refining this action, the Yang version lends layered emotional tension to Baoyu’s speech, reinforcing his sensitive, delicate, and volatile character. The standardized treatment in the Hawkes version simplifies emotional expression into more unitary and explicit forms, improving readability while potentially diluting the original technique of conveying emotion through actions.

It can therefore be concluded that the two translators’ divergent treatments of “笑道” reflect not merely formal linguistic choices but two distinct narrative orientations: one oriented toward behavioral presentation, constructing character emotion through detailed actions, the other toward semantic explication, clarifying character attitude through narrative labeling. These two approaches shape the expressive mode of character discourse and ultimately contribute to the reconstruction of Jia Baoyu’s image.

5.2 Translation of Actional Micro-details

Compared with stylized behavioral markers such as “笑道”, the translation of concrete actional micro-details more directly reflects the translator’s understanding of the intensity of a character’s behavior and emotional tension. Differences at this level are therefore manifested not only in linguistic form but also profoundly affect the dynamic presentation of the character image.

Comparative analysis shows that in the translation of Jia Baoyu’s actional micro-details, the Yang version and the Hawkes version display distinctly different orientations: the former generally adheres to the expressive path of the original, focusing on representation, the latter frequently employs amplification strategies to elaborate the action process, thereby enhancing the visuality and expressiveness of the actions.

Example (5)

Original: 登时发作起痴狂病来，摘下那玉，就狠命摔去，骂道……

Yang: Tearing off the jade he flung it on the ground.

Hawkes: This sent Bao-yu off instantly into one of his mad fits. Snatching the jade from his neck he hurled it violently on the floor as if to smash it and began abusing it passionately.

Take Example (5) as an illustration. The original constructs a scene of sudden and violent behavior through the continuous action chain of “摘下——摔去” (tear off – fling away) coupled with the degree adverb “狠命” (desperately). The Yang version largely follows the original structure, using “Tearing off... he flung it...” to reproduce the sequence of actions, maintaining the compact rhythm and directness of the original. The Hawkes version, by contrast, provides multiple layers of expansion: it supplements the starting point of the action (from his neck), explicates the behavioral intention through “as if to smash it”, and further extends emotional expression by adding “began abusing it passionately”. This treatment shifts the action from event presentation to process intensification, significantly heightening the dramatic tension of the behavior.

Example (6)

Original: 宝玉禁不住大叫：“了不得！”一脚踹进门去。

Yang: “Heavens above!” As Baoyu charged into the room

Hawkes: ‘Good lord!’ He cried out involuntarily, and kicking open the door, strode into the study.

The difference is equally evident in Example (6). The original combines “禁不住大叫” (could not help shouting) and “一脚踹进门去” (kicked the door open and entered) to form a tight linkage between emotional outburst and action execution. The Yang version adopts a relatively concise expression for compressed representation of the action, presenting a rhythm similar to the original yet somewhat weakening the suddenness of the emotional eruption. In comparison, the Hawkes version clearly marks the involuntary nature of the emotion by adding “cried out involuntarily”, and decomposes and extends the action through “kicking open the door, strode into...”, lending the behavioral process greater layeredness and visibility.

From the perspective of narrative function, this difference can be summarized as two distinct mechanisms of action representation: one is compressed representation oriented toward structural equivalence, emphasizing the integrity of the action chain and consistency of rhythm, the other is expanded representation oriented toward perceptual intensification, providing readers with more concrete sensory experience by supplementing action details and psychological motivations.

Furthermore, such divergent treatments directly shape the presentation of the character’s personality. Through relatively restrained action descriptions, the Yang version brings Jia Baoyu’s behavior closer to the narrative rhythm of the original, with intense yet still somewhat implicit emotional expression. By supplementing details and extending actions, the Hawkes version lends the character’s behavior greater impact and visual effect, thus highlighting his emotionally expressive and uninhibited side.

6. COMPARATIVE ANALYSIS OF THE TRANSLATION OF CHARACTER APPEARANCE

The depiction of character appearance constitutes one of the important dimensions of character portrayal in literary works. Through descriptions of physical features, attire and overall demeanor, the text not only provides readers with an intuitive visual impression but also implies multiple layers of information including the character’s identity, social status and cultural-aesthetic orientation. Appearance depiction thus serves both as a foundational layer of image construction and a vital carrier of aesthetic consciousness.

In *Hong Lou Meng*, Cao Xueqin offers an elaborate portrayal of Jia Baoyu’s appearance in highly rhetorical language upon his first appearance. This depiction not only presents his physical features but also constructs an aesthetic image of grace, agility and faint femininity through a series of metaphorical images such as “the moon of the Mid-Autumn festival”, “spring dawn flowers” and “autumn waves”, embodying the aesthetic tradition of softness and grace in classical Chinese literature.

A comparison of the translations reveals that the Hawkes version intervenes more noticeably in the translation of character appearance than the Yang version, chiefly through amplification and domestication.

Example (7)

Original: 面若中秋之月，色如春晓之花，鬓若刀裁，眉如墨画，面如桃瓣，目若秋波。

Hawkes: As to his person, he had: a face like the moon of Mid-Autumn, a complexion like flowers at dawn, a hairline straight as a knife-cut, eyebrows that might have been painted by an artist's brush, a shapely nose, and eyes clear as limpid pools.

In Example (7): The original constructs a unified image through a sequence of parallel metaphors, marked by high density of imagery and rhythmic cadence. While largely retaining the original metaphorical structure, the Hawkes version adds the description "a shapely nose", which is absent from the original. This supplement disrupts the original imagery system and introduces a new visual focus to the character's portrayal in the target text.

From an aesthetic perspective, this treatment carries distinct cultural orientation. The original emphasizes overall demeanor and associative imagery, creating a subtle and fluid beauty through natural motifs. The specific facial feature added in the Hawkes version, by contrast, aligns more closely with the Western literary aesthetic centered on facial structure. This shift reflects a transition from imagery-based aesthetics to figurative realism.

Furthermore, in rendering "眉如墨画" (eyebrows as if painted in ink), the Hawkes version replaces the culturally specific "ink" with "painted by an artist's brush", converting the distinctly Chinese calligraphic-painting context into a more universal artistic reference. This process of domestication lowers the barrier of cultural comprehension yet also dilutes the calligraphic tradition and cultural connotations inherent in the original.

From the perspective of character construction, these translational strategies collectively reshape the image of Jia Baoyu. The original's layered imagery conveys an aura of grace, subtlety and gentle femininity, which is partially reconstructed in the Hawkes version into a concrete, clear and structurally defined appearance more in line with Western aesthetic logic. While enhancing visualizability, this transformation may to some extent alter the reader's overall perception of the character's temperament.

In comparison, the Yang version tends to preserve the rhetorical structure and imagery system of the original in its handling of appearance descriptions, thereby sustaining the aesthetic style of the source text to a certain degree. Although this approach may increase interpretive difficulty, it helps maintain the unique aesthetic atmosphere and character demeanor conveyed in the original.

7. CONCLUSION

Taking the character image of Jia Baoyu as the research object, this paper conducts a systematic comparison of the two English translations from multiple dimensions, including lexicon, syntax, punctuation, actions and appearance, by combining quantitative analysis of character discourse translation with qualitative analysis of action and appearance descriptions. The study finds that different translators' choices at various linguistic and narrative levels jointly contribute to the reconstruction of the character image, revealing differences in translation orientation with internal consistency.

Overall, the Yang version demonstrates a strong source-text-oriented feature across all dimensions. At the lexical and syntactic levels, its translation strategies mainly adhere closely to the original expression, tending to preserve the structure and rhythm of the source text. In the use of punctuation and the depiction of actions, it adopts a relatively restrained approach. This reproduction-oriented translation path helps sustain the narrative style and aesthetic characteristics of the original, maintaining an implicit and restrained mode of expression in the character's discourse and behavior as a whole. However, due to limited intervention in emotional explicitation and action intensification, the weakening of certain details reduces the perceptibility of the character's emotions to a certain extent, resulting in a more subdued portrayal of Jia Baoyu's personality and diminished expressive tension of the character image.

In contrast, the Hawkes version exhibits a reader reception-oriented, intensifying translation orientation. At the lexical level, it enriches expressions such as address terms and modal particles, and enhances the sense of immediacy of discourse through relatively colloquial language. At the syntactic and punctuation levels, it strengthens tone and emotional expression via sentence expansion and frequent use of exclamation marks. In the depiction of actions and appearance, it further supplements and reconstructs behavioral details and physical features through amplification and domestication strategies. Such multi-layered explicitation renders the character's emotions and behavioral motivations clearer and more identifiable, thereby boosting the expressiveness and appeal of the translation and lending a sharper, more three-dimensional quality to Jia Baoyu's image.

Taken together, the two translations present systematic differences in character construction mechanisms: the Yang version tends to achieve overall reproduction of the original character image through formal linguistic correspondence and restrained expression, while the Hawkes version reshapes and reconstructs the character image to a certain degree via multi-dimensional detail intensification and emotional explicitation. The former represents a reproductive approach to character construction, emphasizing continuity with the original style, the latter embodies an intensifying approach, focusing more on readers' perceptual experience and narrative effect.

This study indicates that character images in literary translation are not merely the product of transformation at a single linguistic level, but emerge from the combined effects of multiple factors, including lexical choice, syntactic organization, punctuation use, and descriptions of actions and appearance. Translators' strategic choices at different levels ultimately converge into an overall narrative orientation, producing differentiated character images in the target text. Therefore, examining character translation from an integrated multi-dimensional perspective helps to achieve a more comprehensive understanding of translator style and the underlying mechanisms by which literary images are reconstructed in translation.

REFERENCES

- [1] Baker, M. (1995). Corpora in translation studies: An overview and some suggestions for future research. *Target*, 7(2), 223–243.
- [2] Baker, M. (2000). Towards a methodology for investigating the style of a literary translator. *Target*, 12(2), 241–266.
- [3] Dang, Z. S. (2012). *A comparative study of the English translations of Hong Lou Meng: Based on the translations by Hawkes and Yang Xianyi*. Peking University Press.
- [4] Feng, Q. H. (2006). *The art of translating Hong Lou Meng: A study on translation artistry*. Shanghai Foreign Language Education Press.
- [5] Hu, K. B. (2017). *Corpus-based translation studies*. Shanghai Jiao Tong University Press.
- [6] Hu, K. B., & Li, Y. (2021). *Corpus-based approaches to literary translation studies*. Foreign Language Teaching and Research Press.
- [7] Liao, C. X. (2019). *A quantitative and qualitative study on the English translations of dialogues in Hong Lou Meng* (Unpublished master's thesis). Shanghai International Studies University.
- [8] Liu, Z. Q. (2010). *The creation and application of a Chinese-English parallel corpus of Hong Lou Meng*. Guangming Daily Press.
- [9] Liu, Z. Q., Liu, C. P., & Zhu, H. (2011). A corpus-based analysis of the translator styles in four English versions of *Hong Lou Meng*. *Chinese Translators Journal*, 32(1), 60–64.
- [10] Liu, Z. Q., & Yan, J. M. (2010). A corpus-based study of translator style and translation strategies: A case study of reporting verbs in *Hong Lou Meng* and their English translations. *Journal of PLA University of Foreign Languages*, 33(4), 87–92.
- [11] Peng, W. Q., & Feng, Q. H. (2016). A study on the reconstruction of literary characters through the translation of commendatory and derogatory terms. *Technology Enhanced Foreign Language Education*, (2), 28–34.
- [12] Wan, C. S. (1991). *A handbook of English punctuation*. Sichuan People's Publishing House.
- [13] Wang, H. Y. (2002). A comparative study of two English translations of the chapter titles in *Hong Lou Meng* from the perspective of rhetorical appeal. *Foreign Languages and Their Teaching*, (1), 54–57.
- [14] Yang, H. Z. (2002). *An introduction to corpus linguistics*. Shanghai Foreign Language Education Press.
- [15] Zhang, Q. S. (2007). Interpreting *Hong Lou Meng*: The importance of delving into the text. *Studies on A Dream of Red Mansions*, (1), 255–259.