Applied Corpus Linguistics 5 (2025) 100125

Contents lists available at ScienceDirect

Applied Corpus Linguistics

journal homepage: www.elsevier.com/locate/acorp

Articles

# Stylistic nuances through syntactic complexity: A corpus-assisted study of narration and dialogue in two english translations of *Hongloumeng*☆

Yifeng Sun [a], Kanglong Liu [b,*]

[a] Department of English, University of Macau, Macau SAR, China
[b] Department of Chinese and Bilingual Studies, The Hong Kong Polytechnic University, Hong Kong SAR, China

ARTICLE INFO

*Keywords:*
Translation style
*Hongloumeng*
Syntactic complexity
Narration
Dialogue

ABSTRACT

*Hongloumeng* (also known as *The Story of the Stone* or *A Dream of Red Mansions*), a significant work in Chinese literature, along with its various English translations has been the subject of substantial scholarly attention. Among these, the two translated versions by David Hawkes and John Minford, and Xianyi Yang and Gladys Yang, have garnered much academic interest and sparked extensive discussions. However, there remains a significant void in the thorough analysis of syntactic complexity, a crucial aspect of their respective distinct translation styles. This study aims to address this gap by conducting a meticulous examination of the syntactic complexity in the first 80 chapters of the novel, as translated by Hawkes and the Yangs, with a specific focus on the subgenres of narration and dialogue. The analysis reveals substantial disparities, such as Hawkes employing longer linguistic units in narration and a higher frequency of subordinations in dialogue. By emphasizing the importance of syntactic complexity within the realm of translation style, this study advocates for integrating metrics that assess syntactic complexity in future explorations related to translation styles. The implications of these findings for enhancing translation research and pedagogy are also discussed.

## 1. Introduction

*Hongloumeng* (HLM), recognized as one of China's Four Great Classic Novels of Chinese literature, has earned international recognition for its intricate depiction of the social and political milieu of the waning Qing Dynasty (1644–1911). The novel provides an unparalleled glimpse into the complex cultural and ideological tapestry of feudal Chinese society, masterfully weaving themes of romance and tragedy into an enthralling saga. Crafted against a backdrop of creative historical fiction, HLM served as a clever vehicle for author Cao Xueqin (曹雪芹) to circumvent the stringent literary censorship of his time. The narrative reflects the shadow of personal adversity that clouded Cao's own life, echoing the decline of his family's fortunes and the heartrending grief of losing his son—motifs that resonate throughout the tale's exploration of hardship. The novel's initial 80 chapters are credited to Cao's literary genius, while the concluding 40 chapters were posthumously completed by Gao E three decades later, as chronicled in the 1982 edition published by People's Literature Publishing House (Cao and Gao, 1982).

Over the years, HLM and its various translations have become a staple in literary and translation studies, drawing extensive scholarly attention (Zheng and Fan, 2022). As of 2015, the novel had been translated into over 20 languages, boasting at least 20 full-length translations that capture the full expanse of the story (Chen, 2015). Among these, two English translations have achieved particular acclaim within the global literary arena. *The Story of the Stone*, translated by David Hawkes and later by John Minford from the late 1970s into the early 1980s. Hawkes (Cao, 1973, 1977, 1980a) and Minford (Cao, 1982, 1986) are credited with translating the first 80 and the final 40 chapters of HLM, respectively. Concurrently, *A Dream of Red Mansions* was translated by Xianyi Yang (also spelled as Hsien-Yi Yang) and Gladys Yang (Cao, 1978, 1980b). Xianyi primarily orally translated the text, while Gladys was responsible for transcribing and refining it (Li et al., 2011). Both Hawkes and Minford, as sinologists deeply engaged with Chinese culture, infused their translation with their scholarly insight. The Yangs, experienced translators, completed their translation under a government commission in the late 1970s. These translations have played a crucial role in comparative studies of the novel's English versions.

☆ Corpus data concerning the study are publicly available on Open Science Framework (https://osf.io/smpnc/).

* Corresponding author at: Department of Chinese and Bilingual Studies, The Hong Kong Polytechnic University, Hong Kong SAR, China.
*E-mail addresses:* sunyf@um.edu.mo (Y. Sun), kl.liu@polyu.edu.hk (K. Liu).

https://doi.org/10.1016/j.acorp.2025.100125
Received 4 April 2024; Received in revised form 20 February 2025; Accepted 27 February 2025
Available online 28 February 2025

## 2. Style in translation

Ferdinand de Saussure's concepts of *langue* and *parole* provide essential distinctions between the overall language system and its practical application. Building on this framework, Leech and Short (2007: 9) situate style within the domain of parole, defining it as "how language is utilized in a specific context, by a particular individual, and for a specific purpose". This definition highlights style as a purposeful selection from the broader linguistic arsenal. Lefevere (1992) emphasizes that translators do not simply replicate the source text; instead, they partake in an act of "rewriting." Expanding on this notion, Baker (2000: 245) uses the metaphor of "fingerprints" to describe the personal touches that translators inevitably leave on their work, which serve to distinguish their linguistic choices from those of their peers. Recent decades have seen a growing trend of using both parallel and comparable corpora to explore the style in literary translations, aiming to illuminate the translator's stylistic profile through various corpus indicators from both comparative and comparable perspectives (Sun and Li, 2020:651). Consequently, the study of translation style has risen to prominence as a key focus of corpus-based research, complementing investigations into translation universals. Methodologically, the study of translator style bifurcates into two principal approaches: the "translator style" and "translation style" paradigms (Saldanha, 2011). The former approach involves comparing multiple translations of a single source text to explore differences in specific linguistic features or indicators (Mastropierro, 2018; Winters, 2009), while the latter entails comparing the body of work of different translators to identify their distinctive language use relative to others (Bosseaux, 2007; Huang and Chu, 2014; Saldanha, 2011).

Translation involves the incorporation of the translator's unique style into the original text, creating a blend of authorial voice and translator-specific elements. This fusion constitutes a challenging stylistic domain to isolate and analyze, especially in hybrid genres combining dialogue and narration. Malmkjær (2003: 38) identifies two distinct approaches to style analysis: the "study of style," which favors form over meaning, and "stylistic analysis," which examines the reasons and processes that inform particular translation decisions. The latter approach, by exploring the impact of stylistic domains on idiosyncratic translation decisions, provides a more comprehensive understanding of a translated text. For a thorough examination of the translation style, it is essential to minimize the influence of the source text (ST) and focus specifically on the target text (TT). Consequently, Saldanha's (2014) target-text-oriented approach is widely favored for stylistic analysis of translation. By comparing multiple translations of the same ST, insights into systematic strategies and linguistic preferences of individual translators are gained, thereby shedding light on the concept of translation style.

In the realm of HLM translation research, there has been an emphasis on comparing the unique styles of different translators. Wong (2002) has conducted a qualitative study examining the techniques employed by translators in their English, French, German, Italian, and indirect English translations of HLM. The versions by Hawkes-Minford and the Yangs, in particular, have received significant scholarly attention. Hawkes-Minford's translation has been lauded internationally for its readability and elegance, while the Yangs' version is esteemed in China for its linguistic accuracy (Wang, 2016). Various studies have focused on comparing the lexical choices of the Hawkes-Minford and Yang translations. For example, Zuo (2015) identified Hawkes-Minford's approach to cultural items in HLM as involving literal translation, in contrast to the Yangs' more liberal translation style. Additionally, the translations of HLM have been compared in terms of architectural terms (Wang and Yu, 2019), titles and honorifics (Liu, 2008), and character names (Zhou, 2015). Stylistic variations have also been examined, including the use of metaphorical idioms (Su, 2021) and lexical bundles (Liu and Afzaal, 2021a).

In a similar vein, a growing number of studies have focused on comparing textual differences among translators. Ji and Oakes (2012) conducted statistical tests to analyze the frequency of conjunctions and genitives in Edward Charles Bowra's translation of HLM, noting a higher frequency of determiners in Henry Bencraft Joly's version. Li et al. (2011) employed a statistical approach to compare the translation styles of Hawkes and the Yangs, finding that Hawkes tended to paraphrase and explain cultural items, while the Yangs took a more literal approach to maintain fidelity to the ST. This emphasis on faithfulness to the ST is influenced by the translation commission by the Foreign Languages Press, a government-funded organization aiming to promote Chinese literary classics globally, where fidelity to the ST is considered the primary criterion for translation quality assessment in mainland China. Further studies, including those by Fang (2017) and Fang and Liu (2015), have supported this finding, with Li and Wu (2017) concluding that the Yangs' HLM translation demonstrated greater syntactic similarity to the ST. Stylistic differences, such as lexical choices and syntactic variations, are also frequently compared in HLM translation research. For example, Hou (2013) found that nominalization had different effects on formality in Joly's HLM translation and conciseness in the Yangs' version.

Syntactic variation serves as a distinguishing characteristic in HLM translations. Xu and Li (2021) noted that translated fiction exhibits notable differences in syntactic complexity compared to non-translated English, showcasing increased phrasal complexity, including the presence of coordinate and complex nominal phrases. Translated English prose, on the other hand, is characterized by the presence of subordinations (Xu and Li, 2021). Given that HLM encompasses hybrid genres like fictional narration and dialogue, syntactic complexity is expected to play a defining role in HLM translations. Nevertheless, it remains to be seen whether syntactic complexity serves as a distinctive feature among different HLM translations. A systematic review of syntactic complexity research conducted by Jagaiah et al. (2020) from 1970 to 2019 uncovered that key indicators of syntactic complexity include mean T-unit length, mean number of words per clause, and mean number of clauses per T-unit. These metrics are commonly used to assess writing proficiency. However, in the context of translation, syntactic complexity also addresses other aspects such as explicitation (Blum-Kulka, 1986; Baker, 1993). It has been observed that the ST tends to be simplified in the target language during the translation process (Klaudy and Károly, 2005:17), and explicitation is primarily driven by pragmatic risk-avoidance (Kruger, 2019). For instance, Wang and Li (2011) found in their analysis of two Chinese translations of *Ulysses* that dependent clauses are often positioned after the independent clause to adhere to the syntax of the English source language. This deliberate choice is made to align with the expectations of faithfulness and adherence to norms in mainland China.

The analysis of syntactic complexity in Translation Studies has primarily focused on comparing the distinction between translation and non-translation in product-oriented research (e.g., Liu and Afzaal, 2021b; Xu and Li, 2021) or understanding the cognitive demands in process-oriented research (e.g., Ma, 2021). However, there is a lack of investigation into syntactic complexity among different translated literary works. In this study, we utilized a comprehensive set of 14 syntactic complexity measures developed by Lu (2010) to compare the syntactic complexity of Hawkes's and the Yangs' HLM translations, specifically focusing on the first 80 chapters of the novel. The aim is to quantify their syntactic differences and ultimately determine the extent to which syntactic complexity influences their respective translation styles. Building on the findings of Egbert and Mahlberg (2020), which suggest that dialogue in fiction utilizes more interactive strategies to engage readers, while narration exhibits a higher prevalence of descriptive elements, our research further distinguishes between speech and narration when contrasting the syntactic complexity of Hawkes's and the Yangs' HLM translations. This approach aligns with Egbert and Mahlberg's (2020) suggestion to treat dialogue and narration as distinct registers in the analysis of translation styles.

## 3. Research questions

To explore the impact of the 14 syntactic complexity indicators on translation style, this study addresses the following research questions:

RQ1: Are there significant differences in terms of syntactic complexity between Hawkes's and the Yangs' HLM translations?
RQ2: Are there significant differences in terms of syntactic complexity between the narrative and dialogue parts of Hawkes's and the Yangs' HLM translations, respectively?

Upon pinpointing notable disparities in syntactic complexity indicators between Hawkes and the Yangs, overall and within specific genres, we conducted an in-depth examination and provided illustrations of the indicators that exhibited the largest disparities in the results section.

## 4. Methods and procedures

### 4.1. Corpus

For the purpose of this study, we compiled a parallel corpus that incorporates the translations of the first 80 chapters of HLM carried out by Hawkes and the Yang couple. The corpus was assembled through a combination of scanning hard copies and downloading digital versions from the internet, resulting in a tripartite parallel corpus: the original Chinese text, Hawkes' translation of the first 80 chapters, and the Yangs' translation of the same 80 chapters (see Table 1). Despite the greater overall word usage in the Hawkes subcorpus, the Yangs subcorpus demonstrates a higher Type Token Ratio (TTR) of 3.38, compared to Hawkes' TTR of 3.17, indicating a greater utilization of unique words. Furthermore, the Standardized Type Token Ratio (STTR), computed per 1000 words, reveals the Yangs' subcorpus exhibits a higher STTR of 44.11, in contrast to Hawkes's 42.11. This further highlights the prevalence of unique word usage across the 80 chapters.

Following Egbert and Mahlberg's (2020) approach, the translated texts were imported into a custom Python program that automatically distinguished fictional dialogues from narration by identifying quotation marks. Subsequently, the dialogue data underwent manual proofreading to ensure the inclusion of only self-conversations and conversations between characters (see Chou and Liu, 2024, for a detailed explanation of the corpus compilation). After this process, the final corpus was divided into four English sub-corpora: the Hawkes's Narration corpus (HN), the Hawkes Dialogue corpus (HD), the Yangs' Narration corpus YN), and the Yangs' Dialogue corpus (YD). The statistics illustrating the corpus breakdown are summarized in Table 1. In alignment with the general findings, both YN and YD exhibit higher TTRs and STTRs compared to HN and HD, respectively. This indicates that the Yangs utilized a more extensive range of distinct words in both the narration and dialogue segments.

### 4.2. Analytical framework

Previous research has established that translated texts differ from non-translated texts in syntactic complexity (Liu and Afzaal, 2021b; Wang et al., 2023). In our study, we utilized the 14 syntactic complexity measures outlined by Lu (2010) to quantitatively compare the differences between Hawkes's and Yangs translations. Following Lu's (2010) methodology, both English translations were initially processed using the Stanford parser (Klein and Manning, 2003), which incorporates functions for sentence segmentation, word tokenization, and Part-of-Speech (POS) tagging with Tregex (Levy and Andrew, 2006). With the POS tags obtained, syntactic structures were retrieved and calculated. For example, a clause was defined as a combination of a subject and an immediate finite verb, and the analyzer could extract and count the number of clauses by identifying the conjunction of S (subject) and VP (verb phrase). The extracted syntactic structures encompassed words (delimited by spaces and punctuation marks), sentences (segmented by punctuation marks such as periods, question marks, ellipsis, etc.), clauses (subject and finite verb phrases), dependent clauses (SBAR nodes with words like 'that' and 'when' followed by a finite clause), T-units (beginning with a ROOT node), complex T-units (beginning with a ROOT node and preceding an SBAR node), coordinate phrases (adjective, adverb, noun, and verb phrases preceding conjunction), complex nominals (combinations of noun phrases with adjectives, prepositions, participles, or relative clauses, etc.), and verb phrases (finite or non-finite verb phrases denoted by VP). Finally, the metrics were calculated, resulting in a set of 14 metrics grouped into five subconstructs: length of production unit, amount of subordination, amount of coordination, phrasal complexity, and overall sentence complexity, as presented in Table 2.

The comparison of syntactic complexity between Hawkes's and the Yangs' translations was conducted in three stages. In the first stage, we utilized a Mann-Whitney U test to compare the mean rank differences of 14 syntactic complexity metrics in HC and YC. This allowed us to pinpoint the subconstructs with the most significant differences between the two translators. In the second stage, the same test was conducted to

**Table 2**
The fourteen syntactic complexity measures based on Lu (2010).

| Syntactic complexity metrics | Definition |
|---|---|
| **Subconstruct 1: Length of production unit** | |
| Mean length of sentence (MLS) | # of words / # of sentences |
| Mean length of T-unit (MLT) | # of words / # of T-units |
| Mean length of clause (MLC) | # of words / # of clauses |
| **Subconstruct 2: Amount of Subordination** | |
| T-unit complexity ratio (C_T) | # of clauses / # of T-units |
| Complex T-unit ratio (CT_T) | # of complex T-units / # of T-units |
| Dependent clause ratio (DC_C) | # of dependent clauses / # of clauses |
| Dependent clauses per T-unit (DC_T) | # of dependent clauses / # of T-units |
| **Subconstruct 3: Amount of Coordination** | |
| Coordinate phrases per clause (CP_C) | # of coordinate phrases / # of clauses |
| Coordinate phrases per T-unit (CP_T) | # of coordinate phrases / # of T-units |
| Sentence coordination ratio (T_S) | # of T-units / # of sentences |
| **Subconstruct 4: Phrasal complexity** | |
| Complex nominals per clause (CN_C) | # of complex nominals / # of clauses |
| Complex nominals per T-unit (CN_T) | # of complex nominals / # of T-units |
| Verb phrases per T-unit (VP_T) | # of verb phrases / # of T-units |
| **Subconstruct 5: Overall sentence complexity** | |
| Sentence complexity ratio (C_S) | # of clauses / # of sentences |

**Table 1**
Statistics of the two English HLM translations.

| Genre | Translator | Tokens | Types | TTR | STTR | STTR SD | Sentences |
|---|---|---|---|---|---|---|---|
| Narration | Hawkes (HN) | 279,361 | 14,683 | 5.26 | 43.66 | 55.25 | 12,340 |
| | Yangs (YN) | 193,903 | 11,082 | 5.72 | 43.73 | 54.70 | 10,334 |
| Dialogue | Hawkes (HD) | 280,718 | 10,730 | 3.82 | 39.28 | 59.71 | 19,219 |
| | Yangs (YD) | 219,478 | 9801 | 4.47 | 42.14 | 56.56 | 17,036 |

compare the mean rank differences of the 14 metrics in HN and YN, aiming to reveal notable differences in the narrative genre. Lastly, the same procedure was repeated in the third stage to compare the mean rank differences of the 14 metrics in HD and YD. This comparison sought to identify the prevalent subconstructs in the dialogue genre. The results of these comparisons provide essential insights into the marked syntactic distinctions between the two translators and distinctive translation styles manifested in narration and dialogue.

## 5. Results

The comparison of the two English translations of HLM by Hawkes and the Yangs unveiled that Hawkes tended to employ more syntactically complex structures. Among the 14 measures used to calculate syntactic complexity, which fall into five subconstructs, four subconstructs revealed significant differences between the two translations. The findings, as illustrated by the mean rank differences in Table 3, indicate that the most significant variance was observed in the category of the amount of subordination, especially in dependent clauses per clause and dependent clauses per T-unit, closely followed by clauses per T-unit. The positive mean rank differences in these significant subconstructs suggest that Hawkes is characterized by a notably higher use of subordination and complex phrases, composes longer units, and incorporates a greater number of clauses within a sentence, in comparison to the Yangs.

The comparison between the narrative and dialogue styles of Hawkes and the Yangs reveals a substantial divergence in syntactic complexity. This contrast is more pronounced in narration, where 13 out of 14 measures demonstrate significant differences between HN and YN, in contrast to 11 out of 14 measures showing significant differences between HD and YD (See Fig. 1 and Fig. 2). Overall, the comparison (HD vs YD) follows a similar pattern, with 11 out of 14 measures indicating that Hawkes's translations exhibit higher syntactic complexity. However, some metrics, particularly those related to coordination, including coordination phrases per clause (CP_C) in narration and coordination phrases per T-unit (CP_T) and T-units per sentence (T_S) in dialogue, do not reveal significant differences. In addition, the metric of complex nominals per clause (CN_C) in dialogue also does not display significant divergence (refer to Table 4).

### 5.1. Narration

An analysis of the 14 metrics used to compare syntactic complexity between Hawkes and the Yangs in the narration subgenre reveals that sentence length (MLS), T-unit length (MLT), and the number of complex nominals per T-unit (CN_T) are the most significant indicators of the differences between the two versions (See Fig. 1). The length of production units, as demonstrated by MLS and MLT, has emerged as the most prominent syntactic difference between the two translations. The data shows that Hawkes's translation averaged 22.33 words per sentence, 18.72 words per T-unit, and 11.10 words per clause, while the Yangs' translation averaged 18.60 words per sentence, 16.09 words per T-unit, and 10.34 words per clause (See Appendix A2). This aligns with the findings of Li et al. (2011), indicating that Hawkes used more words and longer sentences in his translations. Furthermore, Hawkes was observed to infuse a personal touch into his translation, as illustrated in Excerpt (1), where he employs phrases like "drowsy state" and "drift off" to convey a dream-like state. The Yangs, on the other hand, opt for the simpler expression "in dream." This highlights that the Yangs employed a broader vocabulary range in their translations. Additionally in Excerpt (1), they succinctly use the nominal phrase "an unknown place," whereas Hawkes opted for the more elaborate relative clause "some place he could not identify" (refer to Appendix A3).

(1) ST: 夢至一處, 不辨是何地方.忽見那廂來了一僧一道, 且行且談.

"Meng zhi yi chu, bu bian shi he difang. Hu jian na xiang laile yi seng yidao, qiexing qie tan."

Hawkes: While in this drowsy state he seemed to drift off to some place he could not identify, where he became aware of a monk and a Taoist walking along and talking as they went.

Yangs: In dream he travelled to an unknown place, where he suddenly noticed a monk and a Taoist approaching, talking together.

In addition to the length of production units, Hawkes's narrative translation shows a greater utilization of complex nominals and verb phrases compared to the Yangs in terms of linguistic features. According to Lu (2010), following Cooper's (1976) definition, complex nominals can encompass combinations of nouns with adjectives, possessives, prepositions, relative clauses, or participles, as well as gerunds or infinitives occupying the subject position. Verb phrases refer to finite or non-finite verb phrases that immediately follow the subject. The results of this study indicate that, on average, Hawkes used 1.98 complex nominals per T-unit (CN_T), 1.17 complex nominals per clause (CN_C), and 2.45 verb phrases per T-unit (VP_T), whereas the Yangs used 1.59

**Table 3**
Independent-samples mann- whitney U test on syntactic complexity between Hawkes and the Yangs.

| Measure | Hawkes (HC) vs. Yangs (YC) | | | | | |
|---|---|---|---|---|---|---|
| | Mean rank | | Difference | U | z | P |
| ***Length of production unit*** | Hawkes | Yangs | | | | |
| words per sentence | 95.71 | 65.29 | 30.42 | 1983.00 | -4.15 | <0.001* |
| words per T-unit | 100.01 | 60.99 | 39.02 | 1639.00 | -5.33 | <0.001* |
| words per clause | 87.33 | 73.67 | 13.66 | 2653.50 | -1.87 | .062 |
| ***Amount of subordination*** | | | | | | |
| clauses per T-unit | 101.36 | 59.64 | 41.72 | 1531.50 | -5.69 | <0.001* |
| complex T-units per T-unit | 99.54 | 61.46 | 38.08 | 1677.00 | -5.20 | <0.001* |
| dependent clauses per clause | 103.99 | 57.01 | 46.98 | 1321.00 | -6.41 | <0.001* |
| dependent clauses per T-unit | 102.90 | 58.10 | 44.80 | 1408.00 | -6.12 | <0.001* |
| ***Amount of coordination*** | | | | | | |
| coordinate phrases per clause | 78.29 | 82.71 | -4.42 | 3023.50 | -0.60 | .547 |
| coordinate phrases per T-unit | 87.75 | 73.25 | 14.50 | 2620.00 | -1.98 | .048* |
| T-units per sentence | 77.22 | 83.78 | -6.56 | 2938.00 | -0.89 | .371 |
| ***Phrasal complexity*** | | | | | | |
| complex nominals per clause | 95.31 | 65.69 | 29.62 | 2015.00 | -4.04 | <0.001* |
| complex nominals per T-unit | 100.68 | 60.33 | 40.35 | 1586.00 | -5.51 | <0.001* |
| verb phrases per T-unit | 100.45 | 60.55 | 39.90 | 1604.00 | -5.45 | <0.001* |
| ***Overall sentence complexity*** | | | | | | |
| clauses per sentence | 95.55 | 65.45 | 30.10 | 1996.00 | -4.11 | <0.001* |

Note.
* The asterisk (*) denotes a significant difference between the mean ranks of Hawkes and the Yangs.

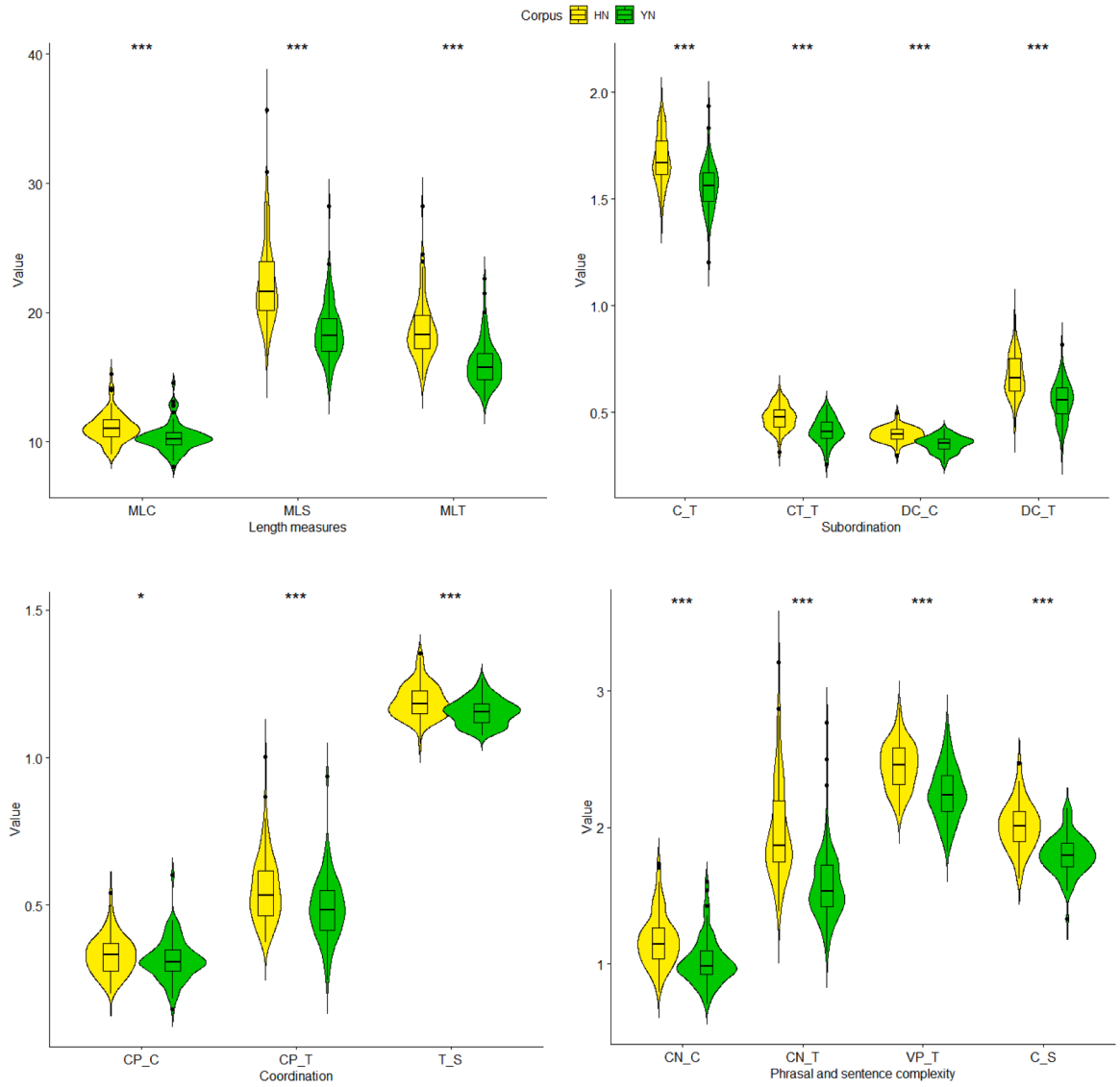


**Fig. 1.** . Comparison of syntactic complexity between HN and YN. *Note.* * denotes significant differences between the two corpora at a level of $< 0.05$, while *** indicates a difference at a level of $< 0.01$.

CN_T, 1.02 CN_C, and 2.25 VP_T (See Appendix A2). Essentially, Hawkes employed a higher frequency of complex nominals and verb phrases within single clauses or T-units in his narrative translation compared to the Yangs. An example of this can be observed in Excerpt (3), where Hawkes's description of a beautiful woman being escorted into the room by a group of female servants involves four nominal groups ("beautiful young woman", "the room behind", "the one they were sitting in", "a bevy of serving women and maids"). It also includes an additional complex nominal ("the one they were sitting in"), which was not present in the original text but was added based on Hawkes's personal interpretation. In contrast, the Yangs' translation only comprises two nominal groups ("some matrons and maids", "a young woman").

(2) ST: 心下想時, 只見一群媳婦丫鬟擁著一個麗人從後房門進來.

"Xin xia xiang shi, zhijian yiqun xifu yahuan yongzhe yige liren cong hou fang men jinlai."

Hawkes: Even as she wondered, a beautiful young woman entered from the room behind the one they were sitting in, surrounded by a bevy of serving women and maids.

Yangs: While she was still wondering, through the back door trooped some matrons and maids surrounding a young woman.

### 5.2. *Fictional dialogue*

Fig. 2 shows significant differences in syntactic complexity between Hawkes's and the Yangs' translations of HLM dialogue, particularly in terms of dependent clauses per T-unit (DC_T), dependent clauses per clause (DC_C), and verb phrases per T-unit (VP_T). The subconstruct "amount of subordination," which includes the first two metrics, emerges as the most significant factor when comparing the two translations. This suggests that the use of nested structures is a hallmark of Hawkes's translation style in HLM. Our analysis reveals that Hawkes employed significantly more subordination in his dialogue translations, averaging 0.64 DC_T and 0.38 DC_C, compared to the Yangs' 0.60 DC_T and 0.35 DC_C. As demonstrated in Excerpt (4), Hawkes's translation utilizes a more intricate sentence structure, featuring a T-unit containing a main clause ("you think") and a lengthy subordinate clause. This subordinate clause includes two independent clauses, both commencing with the conditional word "if." The subordinate clause provides additional information, rendering the sentence more complex and nuanced. The presence of two main clauses ("you will be..." and "you would be…") within the subordinate clause, along with additional layers of subordination, results in a more complex sentence structure. On the

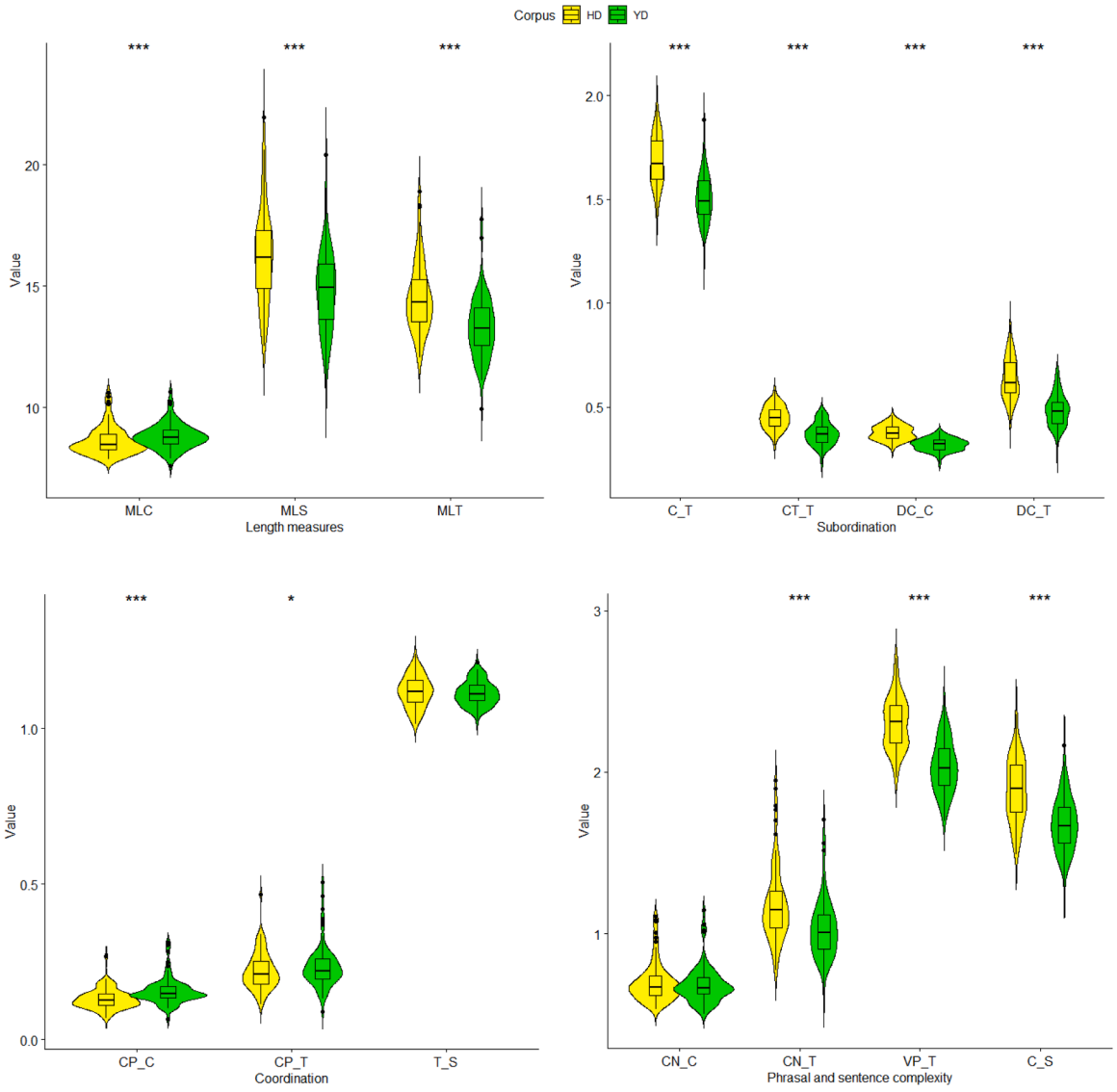


**Fig. 2.** Comparison of syntactic complexity between HD and YD. *Note*. * denotes significant differences between the two corpora at a level of $< 0.05$, while *** indicates a difference at a level of $< 0.01$.

other hand, the Yangs' translation breaks down the idea of the character's preference for independence into a separate sentence, presenting a simpler T-unit consisting of a main clause and a single subordinate clause. This yields a more explicit and simpler sentence structure.

(3) ST:「你的意思我早知道了, 守着舅舅姨母住着, 未免拘紧了, 不如各自住
着, 好任意施为.」

"Ni de yisi wo zao zhidaole, shouzhe jiujiu yimu zhuzhe, weimian ju jinle, buru gezi zhu zhe, hao renyi shi wei."

Hawkes: "But I know perfectly well what's in your mind. You think that if we are staying with your uncle or aunt you will be too restricted, and that if we were living in our own place you would be freer to do just as you liked."

Yangs: "I know what you're after. You're afraid of being under restraint if you stay with your uncle or aunt. You'd prefer to be on your own, free to do as you please."

Another syntactic aspect that underscores a distinction between the two translator pairs is the overall sentence complexity, measured by the average number of clauses per sentence (C_S). This study adopts Lu's (2010) definition of a clause as one that has a subject and is immediately followed by a finite verb. The analysis reveals that Hawkes's dialogue translation exhibits a higher level of sentence complexity, with an average of 1.89 clauses per sentence, compared to the Yangs' translation with 1.68 clauses per sentence. Illustrating this contrast, excerpt (5) exemplifies Hawkes's sentence consisting of five clauses, each structured with a subject followed immediately by a finite verb ("I haven't", "I'm, "I call", "I've", "I know"), while the Yangs' sentence only contains three such clauses ("I've", "I'm", "I've").

(4) ST: 况且我長了這麼大, 文不文武不武.雖說做買賣, 究竟戥子,算盤從沒拿過.地土風俗, 遠近道路, 又不知道.

"Kuangqie wo zhangle zhemo da, wen bu wen wu bu wu. Suishuo zuo maimai, jiujing deng zi, suanpan cong mei naguo. Di tu fengsu, yuanjin daolu, you bu zhidao."

Hawkes: And for another thing, I haven't done very much with my life to date: I'm neither a scholar nor a soldier, and though I call myself a merchant, I've never handled a pair of scales or an abacus in my life, not to mention the fact that I know nothing about the places and peoples of the empire of its roads and waterways.

Yangs: Besides, all these years I've never taken to books or soldiering, and although I'm in business I've never handled a balance or abacus and know nothing either about local customs and different parts of the country.

**Table 4**
Independent-samples mann- whitney U test on syntactic complexity (Hawkes' and Yangs' narration and dialogue).

| Measure | Hawkes Narration (HN) vs. Yangs Narration (YN) | | | | | | Hawkes Dialogue (HD) vs. Yangs Dialogue (YD) | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Mean rank | | difference | U | z | P | Mean rank | | difference | U | z | P |
| ***Length of production unit*** | Hawkes | Yangs | | | | | Hawkes | Yangs | | | | |
| words per sentence | 109.09 | 51.91 | 57.18 | 913.00 | -7.81 | <0.001* | 97.99 | 63.01 | 34.98 | 1801.00 | -4.77 | <0.001* |
| words per T-unit | 108.05 | 52.95 | 55.10 | 996.00 | -7.52 | <0.001* | 100.33 | 60.68 | 39.65 | 1614.00 | -5.41 | <0.001* |
| words per clause | 98.49 | 62.51 | 35.98 | 1761.00 | -4.91 | <0.001* | 69.26 | 91.74 | -22.48 | 2301.00 | -3.07 | <0.001* |
| ***Amount of subordination*** | | | | | | | | | | | | <0.001* |
| clauses per T-unit | 103.34 | 57.66 | 45.68 | 1373.00 | -6.24 | <0.001* | 109.30 | 51.70 | 57.60 | 896.00 | -7.86 | <0.001* |
| complex T-units per T-unit | 101.99 | 59.01 | 42.98 | 1480.50 | -5.87 | <0.001* | 108.06 | 52.94 | 55.12 | 995.00 | -7.53 | <0.001* |
| dependent clauses per clause | 106.16 | 54.84 | 51.32 | 1147.00 | -7.01 | <0.001* | 111.58 | 49.43 | 62.15 | 714.00 | -8.48 | <0.001* |
| dependent clauses per T-unit | 105.83 | 55.17 | 50.66 | 1173.50 | -6.92 | <0.001* | 112.00 | 49.00 | 63.00 | 680.00 | -8.60 | <0.001* |
| ***Amount of coordination*** | | | | | | | | | | | | <0.001* |
| coordinate phrases per clause | 86.21 | 74.79 | 11.42 | 2743.00 | -1.56 | .119 | 64.39 | 96.61 | -32.22 | 1911.50 | -4.40 | <0.001* |
| coordinate phrases per T-unit | 93.87 | 67.13 | 26.74 | 2130.50 | -3.65 | <0.001* | 96.83 | 64.18 | 32.65 | 2751.00 | -1.53 | .125 |
| T-units per sentence | 94.86 | 66.14 | 28.72 | 2051.50 | -3.92 | <0.001* | 83.87 | 77.13 | 6.74 | 2930.50 | -0.92 | .358 |
| ***Phrasal complexity*** | | | | | | | | | | | | |
| complex nominals per clause | 100.43 | 60.57 | 39.86 | 1605.50 | -5.54 | <0.001* | 82.91 | 78.09 | 4.82 | 3007.00 | -0.66 | .510 |
| complex nominals per T-unit | 108.17 | 52.83 | 55.34 | 986.50 | -7.55 | <0.001* | 96.83 | 64.18 | 32.65 | 1894.00 | -4.46 | <0.001* |
| verb phrases per T-unit | 102.43 | 58.57 | 43.86 | 1445.50 | -5.99 | <0.001* | 110.47 | 50.53 | 59.94 | 802.50 | -8.18 | <0.001* |
| ***Overall sentence complexity*** | | | | | | | | | | | | <0.001* |
| clauses per sentence | 106.46 | 54.54 | 51.92 | 1123.00 | -7.09 | <0.001* | 103.59 | 57.41 | 46.18 | 1353.00 | -6.30 | <0.001* |

Note.
* The asterisk (*) denotes a significant difference between the mean ranks of Hawkes and the Yangs.

## 6. Discussion

This study conducts a detailed analysis of two HLM translations by Hawkes and the Yangs, focusing on syntactic complexity in both narration and dialogue. Utilizing 14 syntactic complexity metrics, the study reveals notable differences between the translations in four types of syntactic complexity: length of production unit, amount of subordination, phrasal complexity, and overall sentence complexity, with Hawkes employing significantly more syntactically complex structures compared to the Yangs. Genre-based analyses indicate that the length of production unit represents the most significant divergence in the translation of HLM's narration, whereas the degree of subordination remains the most distinctive feature in the translation of dialogue. Li et al. (2011) attribute these differences to the translators' distinct motivations, contexts, and translation methods. The Yangs, commissioned by the Foreign Language Press in Beijing, aimed to present Chinese literature and culture authentically to an international audience, which resulted in the adoption of a rich lexicon. Conversely, Hawkes, as an academic and independent translator, prioritized making the text engaging and enjoyable for the reader, often elaborating on complex cultural terms with more extended sentences. Moreover, the Yangs' dual-mode translation process, involving Xianyi Yang's initial oral translation followed by Gladys Yang's written transcription (further elaborated in Li et al., 2011), could have led to a less syntactically complex version than Hawkes's exclusively written translation process.

Skopos Theory (Reiß et al., 2013) offers a valuable framework for understanding the different objectives that shaped Hawkes's and the Yangs' translation strategies. It can be said that working under the aegis of the Chinese government, the Yangs faced constraints that required adherence to precise and concise syntactic structures. In contrast, Hawkes, working with Penguin Books, a commercial publisher catering to a wide readership, had greater freedom to enhance the reading experience by weaving cultural explanations directly into the narrative, resulting in more intricately structured sentences.

The collaborative nature of the Yangs' translation effort is a unique aspect of their work on HLM. Although they worked as a team, their process involved a distinct division of labor, which is an essential factor to consider when examining their translation style and approach. As previously mentioned, Gladys Yang's role in the translation was somewhat indirect. She did not translate directly from the original Chinese text; instead, her work was based on an initial translation produced by her husband, Xianyi Yang. This initial translation served as a draft upon which Gladys would build, refining and polishing the English expression. While the Yangs' two-tiered translation process could be expected to introduce additional interpretive layers, the extent to which it impacts the readability and modifies the original text's style and nuances warrants further exploration. Xianyi Yang's initial drafts provided the syntactic groundwork, later refined by Gladys to meet English standards, which might include streamlining structures for clearer comprehension. Furthermore, this approach would require meticulous handling of literary elements to preserve the author's intent and ensure the translation's approachability for English audiences. Contrary to these expectations, however, the end result did not corroborate such assumptions; the translation retained a certain stiffness, and the anticipated ease of reading did not fully come to fruition.

Ma and Zhao (2023) conducted a detailed quantitative analysis that highlighted David Hawkes's meticulous approach to modifying syntactic structures in his translation. To a large extent, Hawkes's creative adjustments to syntax, which diverged from the original text's word order, may have enhanced the readability and stylistic appeal of the English rendition. These changes were not isolated to individual sentences; they were part of a broader strategy that considered the overall flow and connectivity between paragraphs. According to systemic functional grammar principles (Halliday and Matthiessen, 2014), Hawkes's revisions to clause relationships within larger syntactic units were designed to weave meanings more cohesively. This method is particularly prominent in his translation of dialogues, where he employed multiple subordinations to construct complex, dynamic semantic flows. These flows mimic the grammatical features often seen in spoken English, vividly and clearly bringing the characters' conversations to life. This approach not only maintains the original's nuances but also enhances the translation's readability, making the text more accessible and engaging for readers. Furthermore, reader responses to the two translations varied significantly in terms of perceived readability, influenced by the syntactic complexities. Jiang's (2014) analysis of customer reviews on Amazon revealed that Hawkes's translation received more favorable feedback than Yang's, particularly for its elegant language style. Reviewers frequently noted the high readability of Hawkes's translation. Hong (1997) argued that this readability was linked to Hawkes's strategy of text amplification. In his comparative analysis, Hong observed that Hawkes often added adjectives and clauses to culturally loaded terms, such as classical allusions, etiquettes, and social conventions. These additions, which resulted in more complex nominal phrases and longer T-units, helped readers grasp the cultural

connotations of specific expressions, further enhancing the translation's accessibility and appeal.

Biber et al. (2011) identified subordination as a key feature of conversational English, and Liu et al. (2023) demonstrated that English translations from Chinese texts often underutilize subordination in comparison to native spoken English. Given their respective linguistic backgrounds, it is reasonable to assume that Hawkes, being a native English speaker, would be more predisposed to using subordination, while Xianyi Yang's translation choices might be shaped by the syntactic patterns of the Chinese source text. The dialogues in HLM, originally penned in the vernacular *baihua* of the Qing Dynasty, show subtle differences from modern Chinese. In translating these dialogues, Hawkes introduces a higher level of syntactic complexity than the Yangs, particularly by using subordination more extensively. This technique more accurately reflects the intricacies of contemporary spoken English and adheres to the patterns commonly found in English-language fiction.

The approach and findings of this study carry significant implications for advancing the use of corpora in translation research. First, by employing 14 syntactic complexity measures, this study demonstrates the effectiveness of using a comprehensive set of metrics to analyze and compare the syntactic features of translated texts. This multi-dimensional approach enables researchers to identify and quantify the specific aspects of syntactic complexity that differentiate translations, providing a more nuanced understanding of translation (e.g., Liu and Afzaal, 2021b; Liu et al., 2023) and translator style. The 14 metrics used in this study can be readily applied to other parallel corpora, allowing researchers to investigate stylistic differences across translations of the same source text or to examine the syntactic characteristics of a single translator's work. Second, the dissection of fictional works into their narrative and dialogue components represents an innovative approach within translation-style research. This method emphasizes the importance of recognizing fiction as a genre that blends narrative and dialogue (Egbert and Mahlberg, 2020), each with their distinct stylistic features requiring specific translation techniques. By examining narrative and dialogue separately, researchers can conduct a more nuanced analysis of the stylistic elements and translation strategies used across different genres within a single literary work. This reveals the complex genre-specific translation strategies crucial for managing texts with embedded subgenres. Furthermore, this genre-based analytical approach is not limited to fiction; it can be applied to various types of texts and genres. This expansion gives researchers the means to investigate the diversity in translators' methods across different textual situations, thereby deepening our comprehension of translation by acknowledging the complexity and variety of textual forms and genres.

Corpus linguistics, which utilizes large collections of authentic texts and computer-based tools, can benefit teaching and learning by providing empirical language data, enabling data-driven learning, and facilitating the creation of tailored teaching materials and resources (Friginal and Taylor, 2021; Friginal and Roberts, 2022). The insights derived from this study present significant benefits for translator training and education, which inherently involves the study of two languages. In particular, translation curricula could be enriched by incorporating targeted training on analyzing and adjusting syntactic structures to achieve desired specific stylistic outcomes. Translation students should be encouraged to develop a heightened sensitivity to the syntactic characteristics inherent to various text types and genres within both source and target languages. Indeed, Carreres and Noriega-Sánchez (2011) emphasize a strong interconnection between translation teaching and language learning. Consequently, equipping students with the ability to effectively analyze syntactic structures of texts can greatly enhance their translation skills. For instance, when translating fiction, students should be taught to identify the unique stylistic features of narration and dialogue and to apply strategies that accurately convey the intended tone and voice of each segment. This educational approach would likely entail a detailed examination of the syntactic features in the source text's narration and dialogue, exploring how these features influence the overall style and mood. Following the analysis, practical exercises could give students the opportunity to translate excerpts, experimenting with the target language's syntactic possibilities to closely mirror the original style. In addition, the study's revelations aboutthe impact of translation purpose and methodology on syntactic complexity suggest that it is vital for students to be exposed to a variety of translation scenarios and real-world projects. This exposure is crucial for them to gain a sophisticated understanding of how these factors shape translation choices. By working across a broad spectrum of translation assignments that vary in target audiences, cultural contexts, and objectives, students can acquire a profound comprehension of how these variables affect stylistic and syntactic decision-making in translation.

## 7. Conclusion

This study represents a pioneering effort in systematically evaluating the syntactic complexity between two HLM translations, aiming to elucidate its influence on translation style. Championing the analytical distinction between dialogues and narrations to reveal stylistic nuances of HLM translations (e.g., Liu, Kwok et al., 2022; Liu, Cheung et al., 2022), our research has uncovered that Hawkes's translation of HLM displays a higher syntactic complexity compared to the Yangs. This heightened complexity is evident in aspects such as the length of the production unit, the degree of subordination, phrasal complexity, and overall sentence complexity. These findings regarding syntactic complexity are explored further in light of differing translation modes and purposes, including factors such as commission origins and target readerships, following the insights provided by Li et al. (2011).

While this study offers some insights into the syntactic complexity in HLM translations, it is important to recognize its limitations. First, our analysis did not directly address the syntactic complexity of the source text but opted for a comparable corpus design. Future research could augment our findings by comparing the syntactic complexity of the source text with that of the translations and provide a more comprehensive understanding of how the intricacies of the original text shape the translation process. Second, the scope of this study was limited to the first 80 chapters of HLM. A more comprehensive analysis could be achieved by including the second part of HLM, specifically chapters 81–120 translated by John Minford, and comparing them with the equivalent sections translated by the Yangs. Expanding the study to encompass the entire text would facilitate a more detailed examination of syntactic complexity across the entire work. Additionally, future studies might benefit from considering additional linguistic parameters, such as lexical complexity and cohesion. Exploring these aspects would provide further clarity on the stylistic features of HLM translations, thereby enhancing our understanding of the translation's overall style and quality.

## CRediT authorship contribution statement

**Yifeng Sun:** Writing – review & editing, Validation, Investigation, Conceptualization. **Kanglong Liu:** Writing – original draft, Visualization, Software, Methodology, Formal analysis, Data curation.

## Declaration of competing interest

Yifeng Sun and Kanglong Liu, the authors of the manuscript entitled "Stylistic Nuances Through Syntactic Complexity: A Corpus-Assisted Study of Narration and Dialogue in Two English Translations of Hongloumeng," explicitly declare that they have no conflicts of interest regarding this submission

## Appendix A1. Descriptive statistics for the 14 measures in the five subconstructs of syntactic complexity

| Measure and index | Hawkes | | | | Yangs | | | |
|---|---|---|---|---|---|---|---|---|
| | Min | Max | Mean | SD | Min | Max | Mean | SD |
| ***Length of production unit*** | | | | | | | | |
| words per sentence | 13.82 | 30.98 | 19.41 | 3.48 | 12.30 | 28.24 | 17.39 | 3.57 |
| words per T-unit | 12.46 | 23.05 | 16.63 | 2.47 | 11.30 | 22.55 | 14.75 | 2.43 |
| words per clause | 7.26 | 12.18 | 8.95 | 1.00 | 7.56 | 12.04 | 8.71 | 0.85 |
| ***Amount of subordination*** | | | | | | | | |
| clauses per T-unit | 1.52 | 2.25 | 1.85 | 0.14 | 1.40 | 2.07 | 1.69 | 0.17 |
| complex T-units per T-unit | 0.31 | 0.59 | 0.49 | 0.05 | 0.29 | 0.57 | 0.43 | 0.07 |
| dependent clauses per clause | 0.29 | 0.48 | 0.40 | 0.03 | 0.26 | 0.45 | 0.35 | 0.05 |
| dependent clauses per T-unit | 0.44 | 1.08 | 0.75 | 0.12 | 0.37 | 0.89 | 0.60 | 0.14 |
| ***Amount of coordination*** | | | | | | | | |
| coordinate phrases per clause | 0.13 | 0.40 | 0.20 | 0.05 | 0.12 | 0.42 | 0.21 | 0.05 |
| coordinate phrases per T-unit | 0.24 | 0.74 | 0.37 | 0.10 | 0.20 | 0.78 | 0.35 | 0.11 |
| T-units per sentence | 1.06 | 1.35 | 1.16 | 0.05 | 1.08 | 1.34 | 1.17 | 0.05 |
| ***Phrasal complexity*** | | | | | | | | |
| complex nominals per clause | 0.60 | 1.41 | 0.87 | 0.17 | 0.53 | 1.37 | 0.78 | 0.17 |
| complex nominals per T-unit | 1.10 | 3.04 | 1.61 | 0.38 | 0.84 | 2.57 | 1.34 | 0.38 |
| verb phrases per T-unit | 2.03 | 3.07 | 2.52 | 0.22 | 1.84 | 2.85 | 2.29 | 0.25 |
| ***Overall sentence complexity*** | | | | | | | | |
| clauses per sentence | 1.70 | 2.90 | 2.16 | 0.23 | 1.53 | 2.63 | 1.99 | 0.27 |

## Appendix A2. Descriptive statistics for the 14 measures in the five subconstructs of syntactic complexity in narration

| Measure and index | Hawkes | | | | Yangs | | | |
|---|---|---|---|---|---|---|---|---|
| | Min | Max | Mean | SD | Min | Max | Mean | SD |
| ***Length of production unit*** | | | | | | | | |
| words per sentence | 16.57 | 35.65 | 22.33 | 3.26 | 14.28 | 28.21 | 18.60 | 2.23 |
| words per T-unit | 14.78 | 28.25 | 18.72 | 2.34 | 13.11 | 22.62 | 16.09 | 1.76 |
| words per clause | 9.04 | 15.25 | 11.10 | 1.16 | 8.04 | 14.54 | 10.34 | 1.08 |
| ***Amount of subordination*** | | | | | | | | |
| clauses per T-unit | 1.43 | 1.94 | 1.69 | 0.12 | 1.20 | 1.93 | 1.56 | 0.12 |
| complex T-units per T-unit | 0.31 | 0.61 | 0.48 | 0.06 | 0.26 | 0.54 | 0.42 | 0.06 |
| dependent clauses per clause | 0.30 | 0.50 | 0.40 | 0.03 | 0.26 | 0.42 | 0.35 | 0.04 |
| dependent clauses per T-unit | 0.43 | 0.96 | 0.67 | 0.10 | 0.31 | 0.81 | 0.55 | 0.09 |
| ***Amount of coordination*** | | | | | | | | |
| coordinate phrases per clause | 0.20 | 0.54 | 0.33 | 0.07 | 0.15 | 0.60 | 0.32 | 0.07 |
| coordinate phrases per T-unit | 0.38 | 1.00 | 0.56 | 0.12 | 0.25 | 0.94 | 0.49 | 0.11 |
| T-units per sentence | 1.05 | 1.35 | 1.19 | 0.06 | 1.08 | 1.26 | 1.16 | 0.04 |
| ***Phrasal complexity*** | | | | | | | | |
| complex nominals per clause | 0.80 | 1.73 | 1.17 | 0.19 | 0.71 | 1.61 | 1.02 | 0.17 |
| complex nominals per T-unit | 1.38 | 3.21 | 1.98 | 0.36 | 1.08 | 2.77 | 1.59 | 0.27 |
| verb phrases per T-unit | 2.08 | 2.88 | 2.45 | 0.17 | 1.83 | 2.75 | 2.25 | 0.20 |
| ***Overall sentence complexity*** | | | | | | | | |
| clauses per sentence | 1.62 | 2.47 | 2.01 | 0.18 | 1.33 | 2.14 | 1.80 | 0.14 |

## Appendix A3. Descriptive statistics for the 14 measures in the five subconstructs of syntactic complexity in dialogue

| Measure and index | Hawkes | | | | Yangs | | | |
|---|---|---|---|---|---|---|---|---|
| | Min | Max | Mean | SD | Min | Max | Mean | SD |
| ***Length of production unit*** | | | | | | | | |
| words per sentence | 12.52 | 21.94 | 16.30 | 1.99 | 10.68 | 20.42 | 14.83 | 1.72 |
| words per T-unit | 12.07 | 18.89 | 14.55 | 1.44 | 9.95 | 17.77 | 13.31 | 1.26 |
| words per clause | 7.85 | 10.63 | 8.62 | 0.60 | 7.61 | 10.65 | 8.82 | 0.54 |
| ***Amount of subordination*** | | | | | | | | |
| clauses per T-unit | 1.42 | 1.95 | 1.69 | 0.12 | 1.20 | 1.89 | 1.51 | 0.11 |
| complex T-units per T-unit | 0.31 | 0.58 | 0.45 | 0.06 | 0.22 | 0.48 | 0.37 | 0.05 |
| dependent clauses per clause | 0.29 | 0.46 | 0.38 | 0.04 | 0.23 | 0.39 | 0.32 | 0.03 |
| dependent clauses per T-unit | 0.42 | 0.89 | 0.64 | 0.10 | 0.27 | 0.67 | 0.48 | 0.08 |
| ***Amount of coordination*** | | | | | | | | |
| coordinate phrases per clause | 0.07 | 0.27 | 0.13 | 0.03 | 0.07 | 0.31 | 0.16 | 0.04 |
| coordinate phrases per T-unit | 0.11 | 0.47 | 0.22 | 0.06 | 0.09 | 0.51 | 0.24 | 0.07 |
| T-units per sentence | 1.01 | 1.24 | 1.12 | 0.05 | 1.02 | 1.21 | 1.11 | 0.04 |
| ***Phrasal complexity*** | | | | | | | | |

(*continued on next page*)

(*continued*)

| Measure and index | Hawkes | | | | Yangs | | | |
|---|---|---|---|---|---|---|---|---|
| | Min | Max | Mean | SD | Min | Max | Mean | SD |
| complex nominals per clause | 0.54 | 1.11 | 0.70 | 0.12 | 0.50 | 1.15 | 0.69 | 0.12 |
| complex nominals per T-unit | 0.77 | 1.95 | 1.19 | 0.24 | 0.60 | 1.71 | 1.04 | 0.19 |
| verb phrases per T-unit | 1.97 | 2.70 | 2.30 | 0.16 | 1.70 | 2.47 | 2.04 | 0.16 |
| ***Overall sentence complexity*** | | | | | | | | |
| clauses per sentence | 1.49 | 2.36 | 1.89 | 0.20 | 1.28 | 2.17 | 1.68 | 0.17 |

## References


Baker, M., 1993. Corpus linguistics and translation studies: implications and applications. In: Baker, M., Francis, G., Tognini-Bonelli, E. (Eds.), Text and Technology: In Honour of John Sinclair. John Benjamins, pp. 233–250. https://doi.org/10.1075/z.64.15bak.

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

Biber, D., Gray, B., Poonpon, K., 2011. Should we use characteristics of conversation to measure grammatical complexity in L2 writing development? TESOL Quart. 45 (1), 5–35. https://doi.org/10.5054/tq.2011.244483.

Blum-Kulka, Sh., 1986. Shifts of cohesion and coherence in translation. In: House, J., Blum-Kulka, Sh. (Eds.), Interlingual and Intercultural Communication. Narr, pp. 17–35. https://doi.org/10.7575/aiac.alls.v.8n.3p.100.

Bosseaux, C., 2007. How Does It Feel? Point of View in Translation: The Case of Virginia Woolf Into French. Rodopi.

Cao, X., 1973. The Story of the Stone. Vol. 1. The Golden Days [红楼梦第一卷:枉入红尘]. (D. Hawkes, Trans.). Penguin.

Cao, X., 1977. The Story of the Stone. Vol. 2. The Crab-Flower Club [红楼梦第二卷:海棠诗社]. (D. Hawkes, Trans.). Penguin.

Cao, X., 1978. A Dream of Red Mansions. Vol. I-II [红楼梦第一卷 - 第二卷]. (X. Yang and G. Yang, Trans.). Foreign Languages Press.

Cao, X., 1980a. The Story of the Stone. Vol. 3. The Warning Voice [红楼梦第三卷:异兆悲音]. (D. Hawkes, Trans.). Penguin.

Cao, X., 1980b. A Dream of Red Mansions. Vol. III [红楼梦第三卷]. (X. Yang and G. Yang, Trans.). Foreign Languages Press.

Cao, X., 1982. The Story of the Stone, Vol. 4: The Debt of Tears [红楼梦第四卷:绛珠还泪]. (J. Minford, Trans.). Penguin.

Cao, X., 1986. The Story of the Stone, Vol. 5: The Dreamer Wakes [红楼梦第五卷:万境归空]. (J. Minford, Trans.). Penguin.

Cao, X., Gao, E., 1982. Hongloumeng [红楼梦]. People's Literature Press.

Carreres, Á., Noriega-Sánchez, M., 2011. Translation in language teaching: insights from professional translator training. Lang. Learn. J. 39 (3), 281–297. https://doi.org/10.1080/09571736.2011.567356.

Chen, K., 2015. The aims and strategies of translation: taking the translation of A Dream of Red Mansions as an example. Front. Narrat. Stud. 4, 66–70.

Chou, I., Liu, K., 2024. Style in speech and narration of two english translations of Hongloumeng: a corpus-based multidimensional study. Target 36 (1), 76–111. https://benjamins.com/online/target/articles/target.22020.cho.

Cooper, T.C., 1976. Measuring written syntactic patterns of second language learners of German. J. Educ. Res. 69 (5), 176–183. https://doi.org/10.1080/00220671.1976.10884868.

Egbert, J., Mahlberg, M., 2020. Fiction-one register or two? Speech and narration in novels. Reg. Stud. 2 (1), 72–101. https://doi.org/10.1075/rs.19006.egb.

Fang, Y., 2017. L-motif TTR for authorship identification in Hongloumeng and its translation. In: Liu, H., Liang, J. (Eds.), Motifs in Language and Text. De Gruyter Mouton, pp. 87–108. https://doi.org/10.1515/9783110476637-006.

Fang, Y., Liu, H., 2015. Comparison of vocabulary richness in two translated hongloumeng. Glottometrics 31, 54–75.

Friginal, E., Roberts, J., 2022. Materials design and corpora in the classroom. In: Jablonki, R., Csomay, E. (Eds.), The Routledge Handbook of Corpora and Language Teaching and Learning. Routledge, pp. 131–146.

Friginal, E., Taylor, J., 2021. Corpora in the classroom. In: Mohebbi, H., Coombe, C. (Eds.), Research Questions in Language Education and Applied Linguistics. Springer, pp. 189–195.

Halliday, M.A.K., Matthiessen, C., 2014. Halliday's Introduction to Functional Grammar, 4th ed. Routledge.

Hong, T., 1997. The decoder's channel and the translation method of amplification in the english translation of Hongloumeng [解码者的渠道与《红楼梦》英译本中的"扩展译法"]. Stud. A Dream Red Mans. [红楼梦学刊] 13, 284–298.

Hou, Y., 2013. A corpus-based study of nominalization as a feature of translator's style (based on the English versions of Hong Lou Meng). Meta 58 (3), 556–573. https://doi.org/10.7202/1025051ar.

Huang, L., Chu, C., 2014. Translator's style or translational style? A corpus-based study of style in translated Chinese novels. Asia Pacif. Trans. Intercult. Stud. 1 (2), 122–141. https://doi.org/10.1080/23306343.2014.883742.

Jagaiah, T., Olinghouse, N.G., Kearns, D.M., 2020. Syntactic complexity measures: variation by genre, grade-level, students' writing abilities, and writing quality. Read Writ. 33 (10), 2577–2638. https://link.springer.com/article/10.1007/s11145-020-10057-x.

Ji, M., Oakes, M.P., 2012. A corpus study of early english translations of Cao Xueqin's Hongloumeng. In: Oakes, M.P., Ji, M. (Eds.), Quantitative Methods in Corpus-based Translation Studies. John Benjamins, pp. 177–208. https://doi.org/10.1075/scl.51.07ji.

Jiang, F., 2014. Afterlife of the Stone: A Critical History of Translation and Dissemination of Hongloumeng in the English-speaking World [他乡的石头记——《红楼梦》百年英译史研究]. Nankai University Press.

Klaudy, K., Károly, K., 2005. Implicitation in translation: empirical evidence for operational asymmetry in translation. Across Lang. Cult. 6 (1), 13–28. https://doi.org/10.1556/acr.6.2005.1.2.

Klein, D., Manning, C.D., 2003. Fast exact inference with a factored model for natural language parsing. Adv. Neural Inf. Process. Syst. 15, 3–10. http://ilpubs.stanford.edu:8091/~klein/lex-parser.pdf.

Kruger, H., 2019. That again: a multivariate analysis of the factors conditioning syntactic explicitness in translated English. Across Lang. Cult. 20 (1), 1–33. https://doi.org/10.1556/084.001.

Leech, G.N., Short, M., 2007. Style in Fiction: A Linguistic Introduction to English Fictional Prose. Pearson Education.

Lefevere, A., 1992. Translation, Rewriting, and the Manipulation of Literary Fame. Routledge.

Levy, R., Andrew, G., 2006. Tregex and Tsurgeon: tools for querying and manipulating tree data structures. In: The Fifth International Conference on Language Resources and Evaluation. Genoa, Italy.

Li, D., Zhang, C., Liu, K., 2011. Translation style and ideology: a corpus-assisted analysis of two english translations of Hongloumeng. Liter. Linguis. Comput. 26 (2), 153–166. https://doi.org/10.1093/llc/fqr001.

Li, X., Wu, C., 2017. Coherence in Hong Lou Meng and its english translations: an exploratory investigation. Funct. Linguis. 4 (1), 1–14. https://doi.org/10.1186/s40554-016-0035-z.

Liu, K., Kwok, H.L., Moratto, R., 2022a. Hedges and boosters as indicators of translation style: with reference to fictional dialogues in Hongloumeng translations. In: Moratto, R., Liu, K., Chao, D.K. (Eds.), Dream of the Red Chamber: Literary and Translation Perspectives. Routledge, pp. 182–199. https://doi.org/10.4324/9781003296812-13.

Liu, K., Afzaal, M., 2021a. Translator's style through lexical bundles: a corpus-driven analysis of two english translations of Hongloumeng. Front. Psychol. 12, 633422. https://doi.org/10.3389/fpsyg.2021.633422.

Liu, K., Afzaal, M., 2021b. Syntactic complexity in translated and non-translated texts: a corpus-based study of simplification. PLoS One 16 (6), e0253454. https://doi.org/10.1371/journal.pone.0253454.

Liu, K., Cheung, J.O., Moratto, R., 2022b. Lexical bundles in the fictional dialogues of two Hongloumeng translations: a corpus-assisted approach. In: Moratto, R., Li, D. (Eds.), Advances in Corpus Applications in Literary and Translation Studies. Routledge, pp. 229–253. https://doi.org/10.4324/9781003298328-14.

Liu, Y., Cheung, K.F.A., Liu, K., 2023. Syntactic complexity of interpreted, L2 and L1 speech: a constrained language perspective. Lingua 286, 103509. https://doi.org/10.1016/j.lingua.2023.103509.

Liu, Z., 2008. Translating tenor: with reference to the english versions of Hong Lou Meng. Meta 53 (3), 528–548. https://doi.org/10.7202/019238ar.

Lu, X., 2010. Automatic analysis of syntactic complexity in second language writing. Int. J. Corpus Linguist. 15 (4), 474–496. https://doi.org/10.1075/ijcl.15.4.02lu.

Ma, H., Zhao, H., 2023. Study on the translation manuscripts of Hawkes's the story of the Stone from the perspective of genetic translation studies [翻译发生学视域下的霍译《红楼梦》手稿研究]. J. Tianjin Foreign Stud. Univ. [天津外国语大学学报] 30 (6), 67–80.

Ma, X., 2021. Coping with syntactic complexity in English-Chinese sight translation by translation and interpreting students: an eye-tracking investigation. Across Lang. Cult. 22 (2), 192–213. https://doi.org/10.1556/084.2021.00014.

Malmkjær, K., 2003. What happened to God and the Angels: an exercise in translational stylistics. Target 15 (1), 37–58. https://doi.org/10.1075/target.15.1.03mal.

Mastropierro, L., 2018. Key clusters as indicators of translator style. Target 30 (2), 240–259. https://doi.org/10.1075/target.17040.mas.

Reiß, K., Vermeer, H.J., Nord, C., Dudenhöfer, M., 2013. Towards a General Theory of Translational Action: Skopos Theory Explained. Routledge.

Saldanha, G., 2011. Translator style: methodological considerations. The Transl. 17 (1), 25–50. https://doi.org/10.1080/13556509.2011.10799478.

Saldanha, G., 2014. Style in, and of, translation. In: Bermann, S., Porter, C. (Eds.), A Companion to Translation Studies. Wiley, pp. 95–106. https://doi.org/10.1002/9781118613504.ch7.

Su, K., 2021. Translation of metaphorical idioms: a case study of two english versions of Hongloumeng. Babel 67 (3), 332–354. https://doi.org/10.1075/babel.00222.su.

Sun, Y., Li, D., 2020. Digital humanities approaches to literary translation. Comparat. Literat. Stud. 57 (4), 640–654. https://doi.org/10.5325/complitstudies.57.4.0640.
Wang, J., Yu, Y., 2019. A study on the translation of architecture in Hong Lou Meng. Theory Pract. Lang. Stud. 9 (7), 790–795. https://doi.org/10.17507/tpls.0907.06.
Wang, N., 2016. Chinese literature as world literature. Canad. Rev. Comparat. Literat. 43 (3), 380–392. https://doi.org/10.1353/crc.2016.0030.
Wang, Q., Li, D., 2011. Looking for translator's fingerprints: a corpus-based study on chinese translations of Ulysses. Liter. Linguis. Comput. 27 (1), 81–93. https://doi.org/10.1093/llc/fqr039.
Wang, Z., Liu, K., Moratto, R., 2023. A corpus-based study of syntactic complexity of translated and non-translated chairman's statements. Transl. Interpret. 15 (1), 135–151. https://doi.org/10.12807/ti.115201.2023.a07.
Winters, M., 2009. Modal particles explained: how modal particles creep into translations and reveal translators' styles. Target 21 (1), 74–97. https://doi.org/10.1075/target.21.1.04win.
Wong, L., 2002. Translating register: with reference to english, French, German, and Italian versions of the Hong Lou Meng. Babel 48 (3), 247–266. https://doi.org/10.1075/babel.48.3.04won.
Xu, J., Li, J., 2021. A syntactic complexity analysis of translational English across genres. Across Lang. Cult. 22 (2), 214–232. https://doi.org/10.1556/084.2021.00015.
Zheng, J., Fan, W., 2022. Hong Lou Meng in the english world: reception of a translated Chinese classic in digital media. The Transl. 29 (3), 312–326. https://doi.org/10.1080/13556509.2022.2119646.
Zhou, H., 2015. Translating characters' names in Hong Lou Meng during the 20th century: from seeking lexical equivalence to maintaining communicative function. Open Access Lib. J. 2 (11), 1–7. https://doi.org/10.4236/oalib.1102028.
Zuo, B., 2015. Strategies of cultural translation: a contrastive analysis of the two english versions of Hong Lou Meng. In: Ko, L. (Ed.), Translation and Cross-cultural Communication Studies in the Asia Pacific. Brill, pp. 57–72. https://doi.org/10.1163/9789004299245_005.