# Detecting "Parasitic Poems": Quantifying Poetic Style in Late Imperial Chinese Fiction

Jiayu Liu[1] ⬛, Rongqian Ma[2] ⬛, and Keli Du[3] ⬛

[1] Department of Statistics, College of Liberal Arts & Sciences, University of Illinois Urbana-Champaign, Champaign, IL, USA

[2] Department of Information and Library Science, Luddy School of Informatics, Computing, and Engineering, Indiana University Bloomington, Bloomington, IN, USA

[3] Trier Center for Digital Humanities, University of Trier, Trier, Germany

## Abstract

Embedded poetry is a defining feature of late imperial Chinese fiction, yet its narrative function remains contested. While some critics regard these poems as "parasitic"—reiterating surrounding prose with minimal contribution—others argue for their integral aesthetic and rhetorical roles. This study aims to explore if parasitic poems exist in late imperial Chinese fiction and how they can be systematically identified. We develop a computational framework to detect such poems across a corpus of Qing-dynasty novels, combining proxy-based measures (cosine similarity and mutual information) with prompt-based large language models (LLMs). Using a manually annotated dataset of 300 poem-context pairs, we evaluate each method's alignment with human judgments. Our preliminary findings show that proxy models achieve higher accuracy but exhibit limited sensitivity to nonparasitic cases. A multilingual prompt-based approach yields a more balanced performance, suggesting LLMs can approximate literary interpretation when effectively prompted. Our work offers tools for analyzing Chinese poetry and demonstrates the potential of LLMs in modeling literary analysis.

**Keywords:** Late imperial Chinese fiction, parasitic poems, computational literary analysis

## 1 Introduction

In late imperial Chinese fiction (1368-1912), the interplay between prose and poetry is a defining stylistic feature, with poems frequently embedded within the narrative fabric of novels. For example, in the Ming-dynasty (1368-1644) classical Chinese short story collection *New Tales Under the Lamplight* (Jiandeng Xinhua 剪燈新話) by Qu You 瞿佑 (1347-1433), more than half of the stories contain multiple poems, with some works devoting over 30% of their content to poetry [9]. Late Ming classic novels such as the *The Journey to the West* (Xiyou Ji 西遊記) and *The Plum in the Golden Vase* (Jinping Mei 金瓶梅) also contain over 600 poems, which drastically increased since the emergence of the "interweaving prose and verse" (sanyun jiehe 散韻結合) literary phenomenon in the Tang Dynasty (618-907).

The earliest critique of the poems embedded in popular storytelling narratives can be traced to Zhao Yanwei's 趙彥衛 *Miscellaneous Notes from the Studio of Yunlu* (Yunlu Manchao 雲麓漫鈔) in the Southern Song Dynasty (南宋, 1127–1279) [15]. In his assessment, the primary motivation for including poetry in prose narratives was to showcase the writer's literary skills (蓋此等文備眾體，可以見史才，詩筆，議論)[15]. This view has been echoed by contemporary critics, who emphasize the social function of such verses as a means of self-display, signaling the writer's

creativity and aligning the work with elite literary traditions [5]. Another school of thought explains this unique literary phenomenon from the need for performance. They demonstrated that since the Song-Yuan (960-1368) period, professional storytellers used the poem as a teaser to engage the audience across contexts (e.g., social gatherings, public events) [8].

Due to the complex motivations underlying their inclusion, the rhetorical function of embedded poems in fiction has long been a subject of debate among literary scholars, both historically and in contemporary scholarship, within China and beyond. A particular view of these poems sees them as "parasitic" (or "decorative"). Parasitic poems, according to the Chinese literary scholar Zhao Yishan 趙義山, are poems that merely repeat or rephrase content already conveyed in the surrounding prose, offering little to no advancement of plot or character, and often impeding narrative flow [16]. Western scholars such as John Bishop view poetic insertions as once integral but later reduced to commentary or aesthetic embellishment, treating the parasitic poem phenomenon as a "limitation" of the late imperial Chinese fiction [1].

This perspective, however, has been contested by other literary scholars. Rao Longsun 饒龍隼, for example, explicitly opposes the label of "parasitic," arguing instead that verse and prose in late imperial fiction form a "holistic organism," and should not be analytically separated [10]. According to Rao, the poetic elements are integral to the aesthetic, rhetorical, and structural coherence of the narratives. A third group of scholars takes a non-polemical stance, focusing on the function and origins of embedded verse rather than labeling it. For instance, Zhang Zhejun 張哲駿 sees the use of poetry as a means to confer textual authority, drawing upon classical precedents such as the *Book of Songs* (Shijing 詩經) [14]. Similarly, Guo Jie 郭傑 traces the tradition of combining prose and verse with historiographical practices in the pre-Qin literature, suggesting that this hybrid verse-within-prose form has deep cultural roots [4]. Other literary scholars used a traditional close reading approach to analyze the narrative functions of specific poems incorporated in classic Ming-Qing novels, such as the *Dream of the Red Chamber* (Honglou meng 紅樓夢), showing the nonparasitic status as well as the storytelling value of these poems [13].

However, whether the poems embedded in late imperial Chinese fiction are truly "parasitic," and how they are distributed across both canonical and lesser-known novels, remains an open question. Traditional literary criticism, grounded in the close reading approach and interpretive judgment, tends to focus on a small set of canonical works, and struggles to consistently distinguish parasitic poems from those that perform meaningful narrative functions. Given this methodological limitation, it is valuable to develop a quantifiable approach to evaluating the parasitic nature of poems embedded in late imperial Chinese vernacular fiction.

In this paper, we aim to address this problem by asking one simple research question: ***How to measure and evaluate the parasiticism of poetry in late imperial Chinese fiction?*** We define parasitic poems as those that merely reiterate information already conveyed in the surrounding prose. These poems typically contribute little to no advancement of plot, character development, or thematic depth, and may even disrupt narrative flow. We aim to detect subtle stylistic patterns that differentiate parasitic from nonparasitic poems across a broad corpus of Ming-Qing fiction and test the efficacy of two different measures in capturing this stylistic distinction. Our method enables a more scalable and empirically grounded analysis of this literary phenomenon, moving beyond the anecdotal and interpretive limitations of prior scholarship. In doing so, we offer a more comprehensive view of the traits of so-called "parasitic poems," highlighting how popular fiction writers engaged with elite literary forms in both aspirational and formulaic ways.

## 2 Data and Methods

### 2.1 Dataset

Combining *Daizhige Classics* 殆知閣古籍 [1] and *Chinese Text Project*,[2] we identified 614 titles of late imperial Chinese fiction in plain text. From the corpus, we randomly sampled 18 Qing dynasty (1644-1912) titles as our pilot dataset for this study. To extract poems from these novels, we used a two-step process. Chinese poetic lines typically follow fixed-length patterns, usually five or seven characters per line. In the first step, we automatically extracted all lines up to 20 characters long, along with the two preceding and following sentences. While this captured many embedded poems, it also included unrelated content such as couplets, chapter titles, or prose misidentified due to formatting inconsistencies. To refine the results, we used ChatGPT (GPT-3.5) with the prompt: "Analyze the text below to see if it contains poems, and if so, retell all the poems you find." Despite being primarily trained on modern Chinese, ChatGPT effectively identified poems in both vernacular and classical Chinese. The final dataset included 339 poems. For each, we extracted 500 words of surrounding text both before and after the poem to form a unit of analysis comprising the poem and its narrative context. We then excluded all the "opening poems" (ruchang shi 入場詩) and "concluding poems" (xiachang shi 下場詩), as they are accompanied by context on only one side—either before or after the poem—rather than both. This yielded 300 complete poem-context pairs.[3] Figure 1 presents an example from the final dataset.



**Figure 1:** Sample unit of analysis from the final dataset

### 2.2 Methods

To distinguish "parasitic" poems from other nonparasitic poems within fictional narratives, one author with academic training backgrounds in Chinese literature manually annotated the 300 poems through close reading. This annotation dataset served as the ground truth for subsequent computational evaluation. We then explored two different approaches to model this task: (1) Operationalizing the concept of parasitic poems using two proxies: cosine similarity and mutual information.

---

[1] https://github.com/garychowcmu/daizhigev20

[2] https://ctext.org/

[3] Complete dataset is available at `https://github.com/dkltimon/parasitic_poems`

Cosine similarity between vectorized representations of poems and their adjacent prose passages was used as a proxy for content redundancy [6]. Mutual information was used as a measure of contextual informativeness, quantifying the reduction in uncertainty about the poem when conditioned on its prose context [7]. (2) We also experimented with prompting-based methods using large language models to assess whether a given poem introduces new narrative or thematic elements beyond what is already expressed in the surrounding text. For each method, we evaluated performance against our annotated dataset using accuracy and F1 score, and we analyzed cases of agreement and divergence between computational predictions and human annotations. Details of model selection, embedding strategies, mutual information estimation, and LLM prompts and specifics are provided in the following sections.

### 2.2.1 Operationalizing "parasitic poem" using two proxies

To identify parasitic poems, we employed cosine similarity as a measure of semantic redundancy between each poem and its surrounding context. Cosine similarity is commonly used in natural language processing to assess the degree of similarity between two text segments based on the angle between their vector representations [6]. It is well-suited for this task because it captures content-level overlap while being insensitive to text length. To generate vector representations for poems and their contexts, we used the pretrained BERT-base Chinese model provided by Hugging Face [3].[4] We applied mean pooling over the token embeddings to obtain a fixed-length sentence-level representation for each text segment. For each poem in our dataset, we then computed the cosine similarity between its vectorized representation and that of its immediate narrative context, defined as the 500 words preceding and following the poem. The calculated **cosine similarity is our first feature** of representing the poems.

In addition, we replaced each original poem with 10 randomly selected poems from the dataset and calculated the average cosine similarity between each of these replacements and the same context. Our hypothesis is that for parasitic poems, the cosine similarity of the original poem should exceed the average similarity of the 10 randomly inserted alternatives, suggesting that its content was more redundant with the surrounding prose than what would be expected from a randomly chosen poem. **The difference between the two values is our second feature**.

In addition to cosine similarities, we used **mutual information as our third feature**. It quantifies the informational relationship between a poem and its surrounding context. If a poem and its context have similar information, then it is a parasitic poem. In contrast, if a poem contributes more information to its surrounding context, then it is not a parasitic poem. To estimate this relationship, we used the GPT-2 model [5] to approximate mutual information through negative log-likelihoods (NLL). NLL measures how surprising a text is under the predictions of a model: higher NLL values indicate lower predictability, and lower values indicate greater predictability. Calculating mutual information involves four steps:

- First, we prepare three textual inputs: the context alone, the poem alone, and the full combined text with the poem inserted into its context. In our experiment, the context is defined as 500 Chinese characters both before and after the poem.

- Second, we use the language model to compute the NLL of the poem without considering its context, denoted as NLL(poem).

- Third, we compute the NLL of the poem considering its context, which is the difference between the NLL of the poem plus the context and the NLL of the context alone:

$$NLL(poem|context) = NLL(poem + context) - NLL(context). \qquad (1)$$

---

[4] https://huggingface.co/google-bert/bert-base-chinese
[5] `https://huggingface.co/uer/gpt2-distil-chinese-cluecorpussmall`

- Finally, we calculate the approximated mutual information:

$$I(poem; context) = NLL(poem) - NLL(poem|context). \qquad (2)$$

Higher mutual information suggests greater redundancy between the poem and its context, signaling a parasitic relationship between them, whereas lower values indicate the poem contributes novel content and is therefore a nonparasitic poem.

After calculating the mutual information of all the poems, we also performed the random replacement approach. For each poem, we replaced it with 10 random poems and calculated the average mutual information based on each of these replacements and the same context. We hypothesize that for parasitic poems, the mutual information score should be higher than the average mutual information score of the 10 randomly inserted alternatives. **The difference between these two scores is then our fourth and last feature**.

After the feature engineering process, we performed classification on the poems to evaluate how good these four features can represent the difference between parasitic and nonparasitic poems. The classification was done as a 5-fold cross-validation with a linear SVM classifier. To deal with the class imbalance in our data, the SVM classifier was trained in the "balanced" mode to adjust the weights of the two classes.

### 2.2.2 Prompting

In addition to the proxy-based approach described above, we also adopted a more direct method, i.e., asking LLMs to read the poems and distinguish between "parasitic" and "nonparasitic" works. To this end, we experimented with three prompts using the GPT-4o model: In the first prompt (i.e., Prompt 1), we provided the model a short definition of a parasitic poem along with the poem itself and 500 Chinese characters before and after it as context. In the second prompt (i.e., Prompt 2), we simply described the ideas and characteristics of "parasitic poems" without explicitly defining them. Our motivation for experimenting with two prompting strategies stems from NLP research showing that reverse-dictionary tasks and descriptive prompts improve the performance of LLMs [12]. The third prompt is a multilingual one mixing Chinese and English instructions. We chose to try this because research literature suggests that multilingual prompts improve the performance of LLMs [11]. We then used the three prompts to ask the model to identify the poem type and evaluated its performance by comparing its responses with our annotated ground truth.

- **Prompt 1:** Please determine whether the following poem is used in a "parasitic" manner relative to its surrounding context. "Parasitic" means the poem does not add much meaningful information to its immediate contexts (e.g., repeating what has already been said in the contexts).
  Please answer only "Yes" or "No."
  Previous context: {before}
  Poem: {poem}
  Following context: {after}
  If it is parasitic, please respond with 1; if it is not, please respond with 0. No explanation is needed.

- **Prompt 2:** Please determine whether the following poem does not add much meaningful information to its immediate contexts (e.g., repeating what has already been said in the contexts). Please answer only "Yes" or "No."
  Previous context: {before}
  Poem: {poem}
  Following context: {after}
  If yes, please respond with 1; if no, please respond with 0. No explanation is needed.

- **Prompt 3:** 请判断下面诗句相对于其上下文是否是"装饰性"使用。"装饰性"指：Current definition for "decorative poems" is that the poem does not add much meaningful information to its immediate contexts (e.g., repeating what has already been said in the contexts). 只回答"是"或"否"。
  上文: {before}
  诗句: {poem}
  下文: {after}
  如果有请直接回答1，如果没有请直接回答0，不需要解释。

## 3 Results

Table 1 presents the classification results for distinguishing parasitic from nonparasitic poems using two approaches: proxy-based modeling and prompt-based LLM methods.[6] Accuracy and F1 score are reported to evaluate alignment with human annotations. The proxy-based method, which relies on cosine similarity and mutual information to quantify semantic redundancy and contextual informativeness, achieved the highest accuracy of 0.70. This suggests that proxy measures capture broad trends in the annotated data, particularly those associated with the majority class. However, its F1 score is substantially lower (0.41), reflecting its limited ability to correctly identify nonparasitic poems.

Among the prompt-based approaches, the bilingual Prompt 3 performs the best, with an accuracy of 0.67 and an F1 score of 0.59. The confusion matrix (Figure 2) provides further insight: 170 of the 211 parasitic poems were correctly classified, resulting in a recall of 0.81 and a precision of 0.75, for an F1 score of approximately 0.78. This indicates that Prompt 3 aligns well with human judgments for the parasitic class. Performance was weaker for nonparasitic poems. Only 32 out of 89 were correctly identified, while 57 were misclassified as parasitic. This resulted in a precision and recall of 0.36, and an F1 score of 0.45 for the nonparasitic class. The macro-averaged F1 score for Prompt 3 was approximately 0.61, reflecting moderate overall performance but highlighting the ongoing difficulty with minority-class detection.

Prompt 2 shows a moderate improvement over Prompt 1 (accuracy and F1 score of 0.52), while Prompt 1, the least effective prompt, results in the lowest accuracy and F1 score (0.42), indicating that its formulation probably failed to guide the model toward relevant interpretive cues.

In general, all the approaches exhibited a tendency to over-classify poems as parasitic. This likely stems from both the class imbalance and the interpretive challenge of identifying subtle or indirect narrative contributions. Models appear to default to the more frequent class when uncertain, reflecting a broader limitation in computational approaches to literary analysis where interpretive nuance matters.

|  | Proxies-based | Prompt1 | Prompt 2 | Prompt3 |
| --- | --- | --- | --- | --- |
| accuracy | 0.53 | 0.42 | 0.52 | 0.67 |
| F1-score | 0.50 | 0.42 | 0.52 | 0.59 |

**Table 1:** Results of classification using two approaches

---

[6] Full results are available at `https://github.com/dkltimon/parasitic_poems`
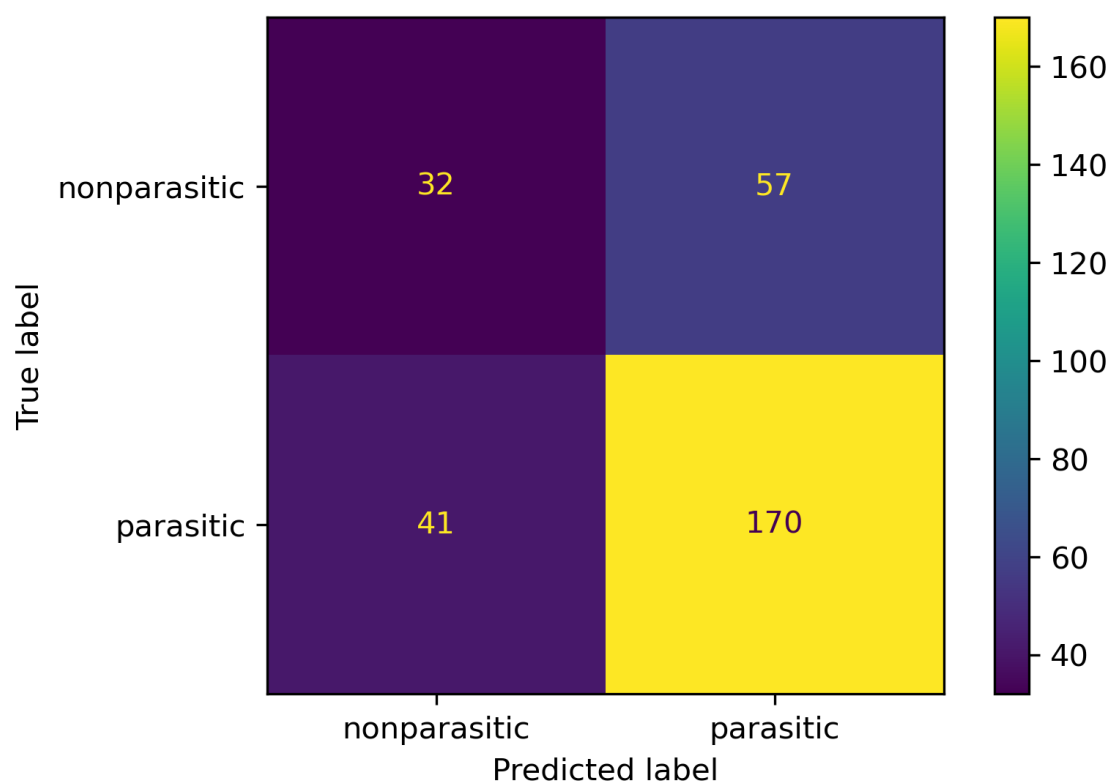
**Figure 2:** Confusion matrix of the classification result using Prompt 3.

## 4 Discussion and Future Work

The results suggest that parasitic poems are prevalent in late imperial Chinese fiction, although their boundary with nonparasitic poems remains ambiguous. While both cosine similarity and mutual information capture some features of parasitic poems, the complexity of the category probably explains the superior performance of prompt-based LLMs. These findings also shed light on the relationship between human annotation and computational classification, pointing to new directions for modeling literary analysis.

### 4.1 Rethinking "ground truth" in literary analysis

The use of human annotations as ground truth in this study raises fundamental questions about the epistemological status of "truth" in literary interpretation. Unlike tasks with objective outcomes (e.g., authorship attribution, named entity recognition), the identification of "parasitic" versus "nonparasitic" poems depends on contextual judgment and interpretation, which are shaped by disciplinary norms and individual perspectives. While our annotated dataset provides a necessary reference point for evaluating computational models, it should be seen less as a fixed standard and more as a situated, interpretive construct. Although introducing more annotators may help reduce this subjectivity, we argue that this type of uncertainty is inherent to literary analysis, which makes the goal and approach of developing a definitive "ground truth" for computational methods both difficult and, to some extent, unrealistic. Literary meaning is often multivalent, context-sensitive, and shaped by divergent interpretive frameworks. What one reader considers parasitic, another may find meaningful or enriching.

Therefore, we call for a more reflexive approach to evaluation and propose two possible directions to address this issue. Instead of treating human annotation as a singular and definitive "ground

truth," future work could adopt more flexible and interpretively sensitive approaches, such as collecting annotations from multiple readers to capture diverse perspectives (multi-perspective annotation), asking annotators to explain their reasoning (annotator rationales), or using probabilistic labeling (e.g., 70% parasitic and 30% nonparasitic) to reflect uncertainty or disagreement in how a poem is classified. More importantly, given this inherent subjectivity of the task, we think that computational disagreement with human annotations in literary analysis should not automatically be framed as model error. Instead, such divergence can offer a meaningful approach to discussing literary ambiguity, assumptions, and diversity. In this light, annotations may be better conceived not as fixed labels, but as reference points to be cross-analyzed with model outputs.

### 4.2 Prompt engineering as an alternative to metric-based models in computational literary analysis

Our experiments also show that prompt-based LLMs (especially Prompt 3) matches or even surpasses the traditional proxy-based approaches in both accuracy and interpretive alignment. Unlike proxy models that rely on predefined metrics such as cosine similarity or mutual information, prompt-based methods leverage the LLMs' latent knowledge and contextual reasoning. The strong performance of Prompt 3, which outperformed the proxy-based method in F1 while offering competitive accuracy, suggests that LLMs can internalize and apply interpretive heuristics when prompted effectively, even in nuanced domains like literary analysis. In future work, we will also test the prompting approach across various models, such as Gemini, Claude, Mistral, as well as those trained with Chinese resources, to provide guidance for applying LLMs for premodern Chinese literary analysis. Furthermore, given that existing models are likely trained primarily on modern Chinese texts, we will also explore the possibility of fine-tuning our own models using written vernacular Chinese texts in premodern periods [2]. At the same time, we observed that prompts written in different languages can influence the results. Therefore, we plan to continue testing prompts written entirely in Chinese, entirely in English, and in mixed Chinese-English formats.

This opens the door to a new paradigm of literary modeling that uses prompting not just for classification, but potentially for explanation, comparison, and interpretive commentary as well. The challenge, then, shifts from designing numerical proxies to crafting cognitively meaningful prompts. This has important methodological implications: literary NLP tasks may benefit more from prompt engineering and zero-/few-shot LLM setups than from traditional supervised pipelines, especially when human interpretive judgments are the gold standard. While our experiments highlight the varying effectiveness of prompting strategies, including definition-based, description-based, and multilingual approaches, a deeper understanding of their impact on literary modeling and how this understanding can inform the design of more effective prompts remains an important direction for future research.

## References

[1] Bishop, John L. "Some limitations of Chinese fiction." In: *The Far Eastern Quarterly 15*, no. 2 (1965), pp. 239–247. DOI: `https://doi.org/10.2307/2941769`.

[2] Blouin, Baptiste, Huang, Hen-Hsen, Henriot, Christian, and Armand, Cécile. "Unlocking Transitional Chinese: Word Segmentation in Modern Historical Texts". In: *Proceedings of the Joint 3rd International Conference on Natural Language Processing for Digital Humanities and 8th International Workshop on Computational Linguistics for Uralic Languages*, ed. by Mika Hämäläinen, Emily Öhman, Flammie Pirinen, Khalid Alnajjar, So Miyagawa, Yuri Bizzoni, Niko Partanen, and Jack Rueter. Tokyo, Japan: Association for Computational

Linguistics, Dec. 2023, pp. 92–101. URL: `https://aclanthology.org/2023.nlp4dh-1.11/`.

[3] Devlin, Jacob, Chang, Ming-Wei, Lee, Kenton, and Toutanova, Kristina. "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding". In: *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, ed. by Jill Burstein, Christy Doran, and Thamar Solorio. Minneapolis, Minnesota: Association for Computational Linguistics, June 2019, pp. 4171–4186. DOI: `10.18653/v1/N19-1423`.

[4] Guo, Jie 郭傑. "Zhongguo gudian xiaoshuo zhong shiwen ronghe chuantong de yuanyuan yu fazhan 中國古典小說中詩文融合傳統的淵源與發展". In: *Zhongguo wenxue yanjiu* 中國文學研究 *2* (1995). CNKI:SUN:ZWXY.0.1995-02-001, pp. 11–17.

[5] Hegel, Robert E. "Distinguishing Levels of Audiences for Ming-Ch'ing Vernacular Literature: A Case Study". In: *Popular Culture in Late Imperial China*, ed. by David Johnson, Andrew J. Nathan, and Evelyn S. Rawski. Berkeley and Los Angeles: University of California Press, 1985, pp. 112–142.

[6] Hendrickx, Iris, Daelemans, Walter, Marsi, Erwin, and Krahmer, Emiel. "Reducing Redundancy in Multi-document Summarization Using Lexical Semantic Similarity". In: *Proceedings of the 2009 Workshop on Language Generation and Summarisation (UCNLG+Sum 2009)*, ed. by Anja Belz, Roger Evans, and Sebastian Varges. Suntec, Singapore: Association for Computational Linguistics, Aug. 2009, pp. 63–66. URL: `https://aclanthology.org/W09-2812/`.

[7] Kreer, J. G. "A question of terminology". In: *IRE Transactions on Information Theory 3*, no. 3 (1957), p. 208. DOI: `10.1109/TIT.1957.1057418`.

[8] Luo, Ye 羅燁. *Zuiweng tanlu* 醉翁談錄. Taipei 臺北: Shijie Shuju 世界書局, 1965.

[9] Qu, You 瞿佑. *Jiandeng xinhua* 剪燈新話. Chinese. Beijing 1st edition. Beijing 北京: Zhongguo wenshi chubanshe 中國文史出版社, 2001.

[10] Rao, Longsun 饒龍隼. "Mingdai xiaoshuo zhong de shiciqu jisheng shuo bianyi—yu Zhao Yishan xiansheng shangque 明代小說中的詩詞曲寄生說辨異——與趙義山先生商権". chinese. In: *Zhongshan daxue xuebao (shehui kexue ban)* 中山大學學報（社會科學版）*63*, no. 01 (2023), pp. 29–37. DOI: `10.13471/j.cnki.jsysusse.2023.01.003`.

[11] Wang, Qihan, Pan, Shidong, Linzen, Tal, and Black, Emily. "Multilingual Prompting for Improving LLM Generation Diversity". In: *arXiv preprint arXiv:2505.15229* (2024). URL: `https://arxiv.org/abs/2505.15229`.

[12] Xu, Ningyu, Zhang, Qi, Zhang, Menghan, Qian, Peng, and Huang, Xuanjing. "On the Tip of the Tongue: Analyzing Conceptual Representation in Large Language Models with Reverse-Dictionary Probe". In: *arXiv preprint arXiv:2402.14404* (2024). arXiv:2402.14404v1 [cs.CL], 22 Feb 2024. URL: `https://arxiv.org/abs/2402.14404`.

[13] Yeh, Chia-ying 葉嘉瑩. "Mantan Hongloumeng zhong de shici 漫談紅樓夢中的詩詞". chinese. In: *Shaanxi shifan daxue xuebao (zhexue shehui kexue ban)* 陝西師範大學學報（哲學社會科學版）*33*, no. 3 (2004), pp. 58–64. DOI: `CNKI:SUN:SXSS.0.2004-03-010`.

[14] Zhang, Zhejun 張哲駿. "Proven with a Poem: Confucian Classic Origins and Functions of the Phrase in Novels and Traditional Operas (Youshi weizheng: Xiaoshuo, Xiqu taoyu de jingxue yuanyuan yu gongneng 有詩為證：小說、戲曲套語的經學淵源與功能)". In: *Journal of Macao Polytechnic University 3* (2015), pp. 154–161.

[15]  Zhao, Yanwei 趙彥衛. *Yunlu Manchao* 雲麓漫鈔. Online version (Chinese Text Project). Accessed 2025-10-26. Page cited: 35. Beijing 北京: Zhonghua Shuju 中華書局, 1966. URL: `https://ctext.org/library.pl?if=gb&file=82214&page=35`.

[16]  Zhao, Yishan 趙義山. *Mingdai jisheng ciqu yanjiu* 明代小說寄生詞曲研究. Beijing 北京: Shangwu Yinshuguan 商務印書館, 2014.