International Journal of Linguistics, Literature and Translation
ISSN: 2617-0299 (Online); ISSN: 2708-0099 (Print)
DOI: 10.32996/ijllt
Journal Homepage: www.al-kindipublisher.com/index.php/ijllt

IJLLT
AL-KINDI CENTER FOR RESEARCH AND DEVELOPMENT

| RESEARCH ARTICLE

A Study on Semantic Similarity of English Translations of Costume Terms in *Jīn Píng Méi* Based on Natural Language Processing

Qi Li[1] ✉ and Qiongyu Liu[2]
[12]School of Foreign Languages, Guangzhou College of Commerce, Guangzhou, Guangdong, P.R.China
Corresponding Author: Qi Li, **E-mail**: liqi2011103207@126.com

| ABSTRACT
This study aims to provide guidance for translating costume terms in Chinese classics, with a specific focus on exploring the semantic similarity of two English translations of costume terms in 《金瓶梅》(Jīn Píng Méi). To achieve this objective, the study compiled a corpus of 35 groups of costume terms (covering clothes, headwear, and footwear) and their corresponding translations from the two versions. It then used the multilingual BERT (mBERT) model, a key natural language processing tool, with relevant technical processes (including text preprocessing, semantic vector mapping, and cosine similarity calculation) to analyze semantic differences between the translations. The main results are as follows: the overall average cosine similarity of the translated terms is 0.69, indicating moderately high consistency. Footwear terms have the highest similarity (0.82) due to their strong practical attributes, while ceremonial and clothes terms show lower similarity (affected by cultural load and complex components). Additionally, about 33% of daily casual terms have high similarity, and around 11% of terms have low similarity, mainly caused by misinterpretation of cultural symbols.

| KEYWORDS
Jīn Píng Méi, costume term translations, natural language processing, semantic similarity.

| ARTICLE INFORMATION
ACCEPTED: 01 August 2025 **PUBLISHED:** 22 September 2025 **DOI:** 10.32996/ijllt.2025.8.9.19

1. Introduction

As a masterpiece of the "worldly novel" genre in the Ming Dynasty(AD1368-1644), 《金瓶梅》(Jīn Píng Méi) constructs a material and cultural map of Ming society through its delicate descriptions of costumes. The costume terms in the novel not only carry physical attributes such as material and style but also imply cultural connotations, including the hierarchical system and aesthetic concepts. With the growing demand for cultural communication, 《金瓶梅》(Jīn Píng Méi) has been translated into English in multiple versions. Among these, The Golden Lotus (1939, London) translated by the British scholar Clement Egerton and The Plum in the Golden Vase (1982–2013) completed by the American sinologist David Tod Roy after over 30 years of efforts, are widely recognized in academia as the two full English translations. These two versions have laid a solid foundation for the dissemination of this classical masterpiece in the Western world.

Owing to the novel's unique literary status and the crucial role of its translations in cultural communication, the English versions have long been a core focus of literary translation studies, with existing research covering multiple dimensions such as language, literature, and culture. Zhao Chaoyong (2020) conducted a comparative analysis using corpus tools and found significant differences in the register dimension between Egerton's and Roy's translations. Based on the Kano model, Zhao Chaoyong & Li Shi (2025) performed text mining on reader review data, constructed a "demand-satisfaction" model, and pointed out that the English translations of The Plum in the Golden Vase have achieved favorable reception in the English-speaking world. The reader demands cover five major dimensions, including text content, translation quality, and aesthetic value, with the improvement of ethical needs ranking the highest in priority. Wen Xiuying & Wang Yuchen (2025) systematically reviewed the century-long history of the English translation of The Plum in the Golden Vase and proposed that its translation strategies have evolved from

"abridged translation to full translation, from domestication to foreignization, and from simplified handling to thick translation practice". Concurrently, the representation of China in these translations has transitioned from an "exotic other" to a "subject of multicultural dialogue".

However, most existing studies focus on the overall register of translations, reader reception, or macro translation strategies. Research on the semantic comparison of culturally specific terms such as costume terms, remains insufficient, and most studies rely on qualitative analysis, lacking quantitative semantic similarity evaluation based on natural language processing (NLP) techniques. With the development of multilingual pre-trained models, multilingual BERT (mBERT) has demonstrated strong capabilities in capturing the contextual semantics of multilingual texts, providing a new technical approach for cross-translation term semantic comparison. In view of this, this study takes the costume terms in The Plum in the Golden Vase and their translations in Egerton's and Roy's versions as the research objects, and conducts a semantic similarity analysis using the mBERT model to expand the connotation of quantitative analysis at the costume term level in The Plum in the Golden Vase.

2. Methodology

2.1 Research Content

2.1.1 Corpus Selection Criteria

Taking Egerton's translation (The Golden Lotus) and Roy's translation (The Plum in the Golden Vase) of 《金瓶梅》(Jīn Píng Méi) as the corpus sources, and combining the original costume descriptions in the novel, the following selection criteria were formulated:

1. Include terms clearly referring to costume categories and their corresponding translations, excluding vague referential expressions;
2. Ensure translations are complete and free of obvious typesetting errors to guarantee the accuracy of semantic analysis;
3. Cover different chapters (Chapters 2–96) and three major costume categories (clothes, headwear, footwear) to ensure the representativeness and diversity of the corpus and avoid conclusion bias caused by single-sample limitation.

2.1.2 Basic Corpus Information

In accordance with the above criteria, a total of 35 groups of costume terms and their corresponding translations were selected. Detailed information is presented in Table 2-1.

Table 2-1 Comparison of Costume Terms in 《金瓶梅》(Jīn Píng Méi) and Their Translations in Egerton's and Roy's Versions

NO.	Costume Term	Egerton Version	Roy Version	Category	Chapter
	比甲	stomacher	vest	Clothes	2
	大红遍地金对襟罗衫儿	a gown of red silk embroidered with gold	a silk blouse of scarlet brocade that opened down the middle	Clothes	20
	翠盖拖泥妆花罗裙	a skirt with an embroidered pattern of green leaves	a long trailing skirt of kingfisher-blue figured silk	Clothes	20
	五彩洒线揉头狮子补子员领	a robe of office with long-maned lions embroidered in five colors	a round-collared robe emblazoned with a mandarin square of variegated embroidery featuring a lion with a ruffled mane	Clothes	31
	苏州绢直裰	Su-chou silk	Su-chou chiffon	Clothes	34
	沉香色妆花补子遍地锦罗袄儿	a stomacher of sandal-wood color, with embroidered flowers	an aloeswood-colored jacket of figured silk brocade	Clothes	40
	大红金枝绿叶百花拖泥裙	a scarlet skirt to match,of the hundred flower design, and all the flowers had gold branches and green leaves	a trailing scarlet skirt sprigged with gold-stemmed and green-leaved flowers	Clothes	40

	大红氅衣儿	a long red gown	a loosely cut crimson robe	Clothes	43
	青水纬罗直身	a black silk gown	a long gown of jet moiré	Clothes	61
	紫绒狮补直身	a purple gown	a long gown of purple velvet featuring only a mandarin square with an embroidered lion	Clothes	68
	白绫对衿袄儿	a white double-breasted silk coat with an embroidered hem	a jacket of white satin with purfled edging that opened down the middle	Clothes	68
	紫羊绒鹤氅	a purple woollen gown	a crane-decorated robe of purple cashmere	Clothes	69
	白绫袄子	a white jacket	a white satin tunic	Clothes	73
	青缎**五彩**飞鱼蟒衣	a dark green silken gown with a dragon in five colors	a green velvet variegated flying fish python robe	Clothes	73
	纱绿遍地金裙	a green skirt	a sand-green brocade skirt	Clothes	75
	大红通袖四兽朝麒麟袍儿	a broad-sleeved with a unicorn gown	a full-sleeved scarlet robe，decorated with a motif of the four animals representing the cardinal directions paying homage to the ch'i-lin	Clothes	96
	翠蓝十样锦百花裙	a skirt of the hundred flowers	a skirt of kingfisher-blue variegated brocade	Clothes	96
	金玲珑簪儿	golden filigree hairpins	gold openwork hairpins	Headwear	2
	缠棕大帽	a large hat of woven palm	a large palmetto hat	Headwear	7
	金九凤垫	a nine-phoenix pin	a pin for holding my chignon in place, in the shape of nine golden phoenixes	Headwear	20
	银丝云髻	a silver hair-net	an informal cloud-shaped chignon enclosed in a fret of silver filigree	Headwear	29
	乌纱	a black ceremonial hat	the black silk hat of an official	Headwear	31
	瓦楞帽儿	a tile- shaped hat	a “tile-ridge” hat	Headwear	34
	金梁缎子八吉祥帽儿	a silken cap of Good Fortune,with a gold brim	a gilt-ridged satin cap decorated with the eight auspicious symbols	Headwear	43
	忠靖冠	a white silk hat	a white satin loyal and	Headwear	69

			tranquil hat		
	貂鼠暖耳	sable ear-covers	sable ear-muffs	Headwear	69
	窝兔儿	fur cap	toque	Headwear	76
	翠蓝销金箍儿	golden band	kingfisher-blue gold headband	Headwear	77
	五梁冠	a five-arched hat	a five-ridged cap	Headwear	96
	细结底陈桥鞋儿	a pair of fine but heavily soled shoes	a pair of fine-soled Ch'en-ch'iao shoes	Footwear	2
	清水布袜儿	socks as white as the purest water	pure cotton stockings	Footwear	2
	云头巧缉山鸦	tiny shoes made like the mountain-crow, with tips embroidered to look like the claws	With a pattern of mountain peaks embroideredon the tips of their toes. Her raven-hued shoes,	Footwear	2
	绿绸子睡鞋大红提根	a pair of green silk bed-shoes, with crimson tops	a pair of sand-green pongee sleeping shoes with scarlet heel lifts	Footwear	28
	墨青素缎鞋儿	shoes were made of dark green silk	shoes of plain ink-black satin	Footwear	68
	皂靴	a pair of black shoes with white soles	white-soled black boots	Footwear	69

2.2 The mBERT Model

As a multilingual variant of BERT, the mBERT model is pre-trained on corpora in 104 languages, endowing it with strong cross-linguistic semantic representation capabilities. Currently, it is widely applied in fields such as multilingual text classification, cross-linguistic information retrieval, and translation quality assessment.

Built on the BERT architecture, the mBERT model consists of 12 Transformer encoders and achieves bidirectional semantic encoding through pre-training on multilingual corpora. Unlike monolingual BERT, mBERT adopts a unified vocabulary and encoding method for texts in different languages during pre-training, enabling it to capture semantic correlations between texts in various languages. This feature makes it particularly suitable for semantic comparison between Egerton's and Roy's English translations—even if the two versions differ in vocabulary selection and syntactic expression, the model can still extract core semantic features through contextual semantic analysis.

The mBERT model processes translated texts in three key steps:

Tokenization: The BERTTokenizer splits continuous text into subword units (e.g., "kingfisher-blue" is split into "kingfisher", "-", "blue");
Semantic Vector Generation: The Transformer encoder captures semantic correlations between subwords via the self-attention mechanism and generates a semantic vector for each subword;
Sentence-Level Semantic Representation: By extracting the [CLS] token (a sentence-level semantic representation), a 768-dimensional semantic vector of the entire translated text is obtained. This vector accurately reflects the core semantic connotation of the text, providing a numerical basis for subsequent similarity calculation.

Cosine similarity quantifies the semantic similarity of two texts by measuring the cosine value of the angle between their semantic vectors. Its value ranges from [-1, 1], where a value closer to 1 indicates higher semantic similarity, and a value closer to -1 indicates greater semantic difference. Let V_E be the semantic vector of Egerton's translation and V_R be that of Roy's translation. The formula for calculating their cosine similarity (Sim) is:

$$Sim(V_E, V_R) = (V_E \cdot V_R)/(|V_E| \times |V_R|)$$

Here, $V_E \cdot V_R$ denotes the dot product of the two vectors, and $|V_E|$ and $|V_R|$ represent the magnitudes of V_E and V_R, respectively. This method effectively eliminates the interference of vector length (text length) and focuses on the matching degree of semantic connotations, making it suitable for semantic comparison between the two translations.

2.3 Data Preprocessing

2.3.1 Text Cleaning

For the English translations of the 35 groups of costume terms in Egerton's and Roy's versions, interference information that might affect semantic extraction was targeted and processed:

Redundant punctuation (e.g., repeated commas, unnecessary parentheses), meaningless line breaks, and contextually redundant conjunctions (e.g., redundant "and" used in describing costume patterns) were removed;

Text case was standardized: Proper nouns (e.g., the place name "Su-chou", the Ming Dynasty costume-specific term "ch'i-lin") were capitalized at the first letter, while all other words were in lowercase. For example, "Kingfisher-blue" was adjusted to "kingfisher-blue" and "Sandal-wood color" to "sandal-wood color". This standardization prevents the mBERT model from misjudging the semantic relevance of cognate words (e.g., "silk" and "Silk") due to case differences.

2.3.2 Term Alignment Verification

With "consistency between costume category and description object" as the core criterion, the matching relationships between the 35 groups of costume terms (17 for Clothes, 12 for Headwear, 6 for Footwear) in Table 2-1 and their corresponding translations in the two versions were verified one by one. The verification logic was: if both translations refer to the category of the original term (e.g., "比甲" belongs to Clothes, and both translations use words for upper garments) and the description object does not deviate from the costume itself, the sample is judged as "valid". After verification, all 35 groups of samples met the requirements of "no category deviation and no object misalignment".

2.3.3 Text Tokenization and Vector Mapping

The BERTTokenizer supporting the mBERT model was used to tokenize the cleaned translations, with a focus on solving the segmentation problem of compound descriptive terms for costumes. Compound terms such as "kingfisher-blue", "sandal-wood color", and "flying fish python" were split into subword units recognizable by the model (e.g., "kingfisher", "-", "blue", "sandal", "-", "wood", "color", "flying", "fish", "python") to avoid the loss of semantic information caused by complex word structures.

Subsequently, the pre-trained "bert-base-multilingual-cased" model was invoked to input the tokenized text into the encoder. By extracting the [CLS] token (a sentence-level semantic feature), a standardized 768-dimensional contextual semantic vector was generated. Each translation corresponds to a unique vector, and the vector generation strictly follows the one-to-one correspondence of "original term → Egerton's translation → Roy's translation". Finally, the vectors were stored in NumPy array format, providing a standardized numerical basis for cosine similarity calculation.

2.3.4 Model Input Parameter Setting

Considering the sample size (35 groups) and the length characteristics of the translations (the longest translation is Roy's version of "大红金枝绿叶百花拖泥裙" in the Clothes category (No. 7): "a trailing scarlet skirt sprigged with gold-stemmed and green-leaved flowers", with a character length of approximately 60), parameters were optimized to balance efficiency and accuracy:

batch_size = 10: 10 groups of samples were processed at a time, adapting to the batch calculation of 35 samples to avoid memory overflow or computational redundancy;

max_length = 64: Covering the length of the longest translation to ensure no truncation of semantic information in long texts;
Optimizer: Adam optimizer was selected, with a learning rate of 2e-5 (adapting to the convergence speed of the mBERT model in semantic extraction of costume texts);

Framework: The model was implemented using the PyTorch 1.12.0 framework, and a random seed (seed = 42) was fixed to ensure consistent vector generation results across multiple runs and reduce random errors.

2.3.5 Similarity Calculation Process

A three-step method of "paired extraction → formula-based calculation → grade classification" was adopted for quantification. The specific steps are as follows:

Paired Vector Extraction: From the NumPy array, semantic vectors of Egerton's translations (V_E) and Roy's translations (V_R) for the 35 groups of samples were extracted in pairs according to the order of No. in Table 2-1, forming 35 vector pairs: (V_{E1}, V_{R1}) to (V_{E35}, V_{R35});

Cosine Similarity Calculation: The semantic matching degree of each vector pair was calculated using the cosine similarity formula (as shown above). The calculation was implemented using the numpy.linalg.norm and numpy.dot functions in Python;

Similarity Grade Classification: Based on the actual semantic differences of costume terms, the results were divided into three grades: high similarity (Sim ≥ 0.8), moderate similarity (0.5 ≤ Sim < 0.8), and low similarity (Sim < 0.5). No samples with semantic deviation were found. The final statistical results are presented in Table 3-2.

Table 2-2 Statistical Distribution of Semantic Similarity Grades

Semantic Similarity Grade	Sim Range	Number of Samples (Groups)	Proportion (%)
High Similarity	Sim ≥ 0.8	12	33.29
Moderate Similarity	0.5 ≤ Sim < 0.8	19	53.29
Low Similarity	Sim < 0.5	4	11.42

3. Results

3.1 Overall Semantic Similarity Calculation Results

Based on the semantic vector mapping and cosine similarity calculation of the mBERT model, the analysis of the 35 groups of costume term translations in Table 2-1 shows that the overall average semantic similarity is 0.69, presenting a characteristic of "moderately high semantic similarity". This indicates that Egerton's and Roy's versions have strong consistency in transmitting the "core attributes (category, function, key physical features)" of costume terms, but diverge in "cultural connotations (e.g., Ming Dynasty-specific patterns and craftsmanship)" and "detailed descriptions (e.g., material precision and style details)"—reflecting differences in translators' understanding of Ming Dynasty costume culture and their translation strategies.

From the perspective of similarity grade distribution, the number, proportion, and core characteristics of samples in each grade are as follows:

High Similarity Samples (Sim ≥ 0.8): A total of 12 groups, accounting for 33.29%. The core characteristic of these samples is "no deviation in core attributes and minor differences in secondary details". The two translations are completely consistent in the "category attribution (e.g., clothes, hats, shoes)", "core function (e.g., warmth retention, hair securing, wearing)", and "key physical features (basic color, core material)" of the costumes. Differences only exist in non-core information, such as "geographical identifiers, secondary materials, and decorative details", which do not affect the overall judgment of costume type.

Example 1 (Clothes category, No. 13: "白绫袄子", Sim = 0.87): Egerton's translation ("a white jacket") and Roy's translation ("a white satin tunic") both accurately convey the core attributes of "white color" and "upper garment category". The only difference is that Roy's translation supplements "satin" (a material similar to "ling", the fine silk in the original term), with no deviation in core semantics.

Example 2 (Headwear category, No. 26: "貂鼠暖耳", Sim = 0.93): Both Egerton's translation ("sable ear-covers") and Roy's translation ("sable ear-muffs") use "sable" to specify the material, and "ear-covers/ear-muffs" accurately correspond to the warmth-retention function of "ear warmers", resulting in highly consistent semantics.

Example 3 (Footwear category, No. 35: "皂靴", Sim = 0.88): Both translations convey the core features of "black color", "white soles", and "boot category". The only minor difference is the word choice between "shoes" and "boots". Considering the attribute of "zao boots" as high-top shoes in the Ming Dynasty, there is no substantial semantic difference.

Moderate Similarity Samples (0.5 ≤ Sim < 0.8): A total of 19 groups, accounting for 53.29%, which is the main body of the samples. These samples exhibit the characteristic of "consistent core semantics and divergent detailed semantics". There is no deviation in macro dimensions such as costume category and core function, but there are differences in micro dimensions such as "color precision (e.g., red vs. crimson)", "material refinement (e.g., silk vs. thin silk)", and "style details (e.g., long vs. loose)", resulting in moderate similarity.

Example 4 (Clothes category, No. 8: "**大**红氅衣儿", Sim = 0.68): Egerton's translation ("a long red gown") emphasizes "long length" and "red color", while Roy's translation ("a loosely cut crimson robe") focuses on "loose cut" and "crimson color". Both refer to "red outer garments", but differences in color precision and style description reduce the similarity.

Example 5 (Headwear category, No. 19: "**金玲**珑簪儿", Sim = 0.72): Egerton's translation ("a large hat of woven palm") describes the material as "woven palm", while Roy's translation ("a large palmetto hat") simplifies the description using "palmetto" (a type of palm plant). Both clearly convey the core attributes of "large size" and "hat category", and the detailed differences do not undermine semantic consistency.

Example 6 (Footwear category, No. 30: 细结底陈桥鞋儿", Sim = 0.75): Egerton's translation ("a pair of fine but heavily soled shoes") focuses on "fine sole" and "thick sole", while Roy's translation ("a pair of fine-soled Ch'en-ch'iao shoes") retains the geographical identifier "Ch'en-ch'iao". The core attribute of "shoe category" is consistent, and the selection of details leads to moderate similarity.

Low Similarity Samples (Sim < 0.5): A total of 4 groups, accounting for 11.42%. These samples are characterized by "substantial deviation in core semantics or key details". The two translations show significant differences in the interpretation of "Ming Dynasty-specific cultural symbols (e.g., patterns, craftsmanship)" and "key physical attributes (e.g., color, material)", resulting in low semantic matching.

Example 7 (Clothes category, No. 14: "**青**缎**五彩**飞鱼蟒衣", Sim = 0.32): Egerton's translation ("a dark green silken gown with a dragon in five colors") simplifies the Ming Dynasty-specific pattern "flying fish and python" to "dragon" and mistakenly identifies "dark blue silk (qingduan)" as "dark green silken". In contrast, Roy's translation ("a green velvet variegated flying fish python robe") accurately restores "flying fish python" and specifies "green velvet", leading to significant differences in core patterns, colors, and materials.

Example 8 (Headwear category, No. 21: "银丝云髻", Sim = 0.42): Egerton's translation ("a silver hair-net") mistakenly identifies "cloud-shaped chignon (yunji)" as "hair-net", deviating from the core category and omitting the "silver-thread craftsmanship". Roy's translation ("an informal cloud-shaped chignon enclosed in a fret of silver filigree") accurately restores the "cloud-shaped chignon" and "silver filigree craftsmanship", resulting in a large semantic difference.

Example 9 (Footwear category, No. 32: "**云**头巧缉山鸦", Sim = 0.32): Egerton's translation interprets "mountain crow (shanya)" as "mountain-crow (a bird)" and "cloud tip (yuntou)" as "claws", while Roy's translation interprets "mountain crow" as "raven-hued (the color of raven feathers)" and "cloud tip" as "mountain peaks". The core images are completely different, leading to extremely low similarity.

3.2 Semantic Similarity Analysis of Different Costume Categories

Combining the classification of the three costume categories (17 groups for Clothes, 12 for Headwear, 6 for Footwear) in Table 2-1, cross-statistics and feature analysis of their semantic similarity were conducted to interpret the laws of semantic differences from the perspective of "category attributes (practical/cultural, simple/complex)". The statistical results are presented in Table 3-1, and the specific characteristics of each category are as follows:

Table 3-1 Semantic Similarity Statistics of Different Costume Categories

Costume Category	Number of Samples (Groups)	Average Semantic Similarity	Proportion of High Similarity (%)	Proportion of Moderate Similarity (%)	Proportion of Low Similarity (%)
Clothes	17	0.65	29.41 (5/17)	58.82 (10/17)	11.76 (2/17)
Headwear	12	0.76	41.67 (5/12)	41.67 (5/12)	16.66 (2/12)
Footwear	6	0.82	66.67 (4/6)	16.67 (1/6)	16.66 (1/6)

3.2.1 Footwear Category: Strongest Practical Orientation, Highest Semantic Consistency

The Footwear category ranks first among the three categories with an average semantic similarity of 0.82, and the proportion of high similarity reaches 66.67% (4 out of 6 groups). Only 1 group (No. 34: "**墨青素**缎鞋儿", Sim = 0.48) falls into the low similarity grade. The core reason lies in its category attribute of "strong practical orientation": all Footwear terms describe daily shoes and

socks worn in the Ming Dynasty, with semantics focusing on directly observable physical attributes such as "function (shoe/sock)", "color", "basic material", and "style". There are no complex cultural connotations (e.g., hierarchical symbols or special patterns), leaving no room for ambiguity in translators' understanding of core semantics and resulting in the lowest probability of translation deviation.

The consistency of high similarity samples is reflected in "no deviation in key physical attributes":

Example 10 (No. 31: "**清水布袜儿**", Sim = 0.85): Egerton's translation ("socks as white as the purest water") conveys "white color" and "sock category" through a metaphor, while Roy's translation ("pure cotton stockings") specifies "cotton material", "white color", and "stocking category". The core semantics of "white socks" are consistent.

Example 11 (No. 33: 绿绸子睡鞋大红提根", Sim = 0.81): Both translations convey the core features of "green color", "sleeping shoes (bed-shoes/sleeping shoes)", and "scarlet heel lifts (crimson tops/scarlet heel lifts)". The only minor difference is the material description between "silk" and "pongee" (a type of thin silk), which does not affect semantic judgment.

The deviation of the only low similarity sample (No. 34: "**墨青素缎鞋儿**", Sim = 0.48) stems from the misjudgment of the key attribute "color": Egerton's translation ("shoes were made of dark green silk") mistakenly interprets "moqing (dark black)" as "dark green", while Roy's translation ("shoes of plain ink-black satin") accurately describes "moqing (ink-black)". As color is a core distinguishing feature of Footwear, this deviation directly leads to a significant reduction in semantic similarity.

3.2.2 Headwear Category: Distinct Functional Differentiation, Prominent Differences Between Daily and Ceremonial Terms

The Headwear category has an average semantic similarity of 0.76, with a high similarity proportion of 41.67%. Overall, it exhibits the characteristic of "high similarity for daily headwear and moderate-low similarity for ceremonial headwear", which is closely related to its "functional duality" attribute:

Daily headwear takes "sunshade and warmth retention" as its core functions, with semantics focusing on practical attributes such as material and shape, resulting in low understanding difficulty;

Ceremonial headwear carries the cultural connotations of the Ming Dynasty official costume system (hierarchical symbols and moral implications). Semantic transmission requires balancing both physical attributes and cultural information, making differences more likely to occur.

(1) Daily Headwear: Practical Attributes Dominant, High Semantic Consistency
High similarity samples are concentrated in daily headwear, with the core being the accurate transmission of "material + function":

Example 12 (No. 23: "**瓦楞帽儿**", Sim = 0.83): Both Egerton's translation ("a tile-shaped hat") and Roy's translation ("a 'tile-ridge' hat") use "tile-shaped/tile-ridge" to correspond to the core shape and clearly indicate the "hat category", resulting in no semantic deviation.

Example 13 (No. 27: "窝兔儿", Sim = 0.80): Both Egerton's translation ("fur cap") and Roy's translation ("toque"—a brimless fur cap) use "fur" to specify the material and refer to "warmth-retention hat category", with consistent core semantics.

(2) Ceremonial Headwear: Cultural Connotation Differences, Reduced Semantic Similarity
Moderate-low similarity samples are concentrated in ceremonial headwear, with deviations arising from "whether cultural information is transmitted":

Example 14 (No. 25: "**忠靖冠**", Sim = 0.56): The "Zhongjing Crown" is a ceremonial headwear for Ming Dynasty officials after retiring from court, implying the meaning of "loyalty and tranquility". Egerton's translation ("a white silk hat") only describes the physical attributes of "white color", "silk material", and "hat category", omitting cultural connotations. Roy's translation ("a white satin loyal and tranquil hat") supplements the implied meaning through "loyal and tranquil". Although it does not fully restore the ceremonial attribute of "crown", semantic differences have already been formed.

Example 15 (No. 29: "五梁冠", Sim = 0.61): The "ridge" in "Five-Ridge Crown" is a symbol of official rank in the Ming Dynasty (five ridges correspond to the fifth rank). Egerton's translation ("a five-arched hat") simplifies "ridge" to "arched", losing the rank information. Roy's translation ("a five-ridged cap") uses "ridge" (closer to the shape of "ridge" in the original term) but still fails to transmit the rank connotation. The different degrees of cultural information loss lead to reduced similarity.

3.2.3 Clothes Category: Highest Complexity, Most Prominent Semantic Differentiation

The Clothes category has the largest number of samples (17 groups) and the lowest average semantic similarity (0.65), with a moderate similarity proportion of 58.82% (10 out of 17 groups). It is the most typical "semantically divergent" category among the three, and the core reason lies in its "high complexity" attribute:

The Clothes category includes both "daily casual clothes" and "ceremonial costumes"; Most terms contain four layers of detailed descriptions: "style + material + craftsmanship + pattern". The room for differences in translators' "detail selection" and "cultural restoration" is much larger than that in other categories, directly leading to semantic divergence.

(1) Daily Casual Clothes: Simple Structure, Relatively High Semantic Consistency

High similarity samples are concentrated in daily casual clothes with simple structures, characterized by "fewer description dimensions and no conflicting details":

Example 15 (No. 1: "比甲", Sim = 0.80): The "比甲" is a sleeveless short jacket in the Ming Dynasty. Both translations—"stomacher (decorative chest garment)" and "vest (sleeveless jacket)"—refer to "sleeveless short upper garments", with only minor differences in functional focus (decorative vs. daily use), resulting in consistent core semantics.

Example 16 (No. 13: "白绫袄子", Sim = 0.87): As mentioned in Section 3.1, both translations clearly indicate "white upper garment", and minor differences in material description do not affect the overall judgment.

(2) Ceremonial Costumes: Complex Details, Significant Semantic Differences

Moderate-low similarity samples are concentrated in ceremonial costumes with complex details, with deviations arising from "the selection or misinterpretation of craftsmanship/pattern details":

Moderate Similarity Case (No. 4: "五彩洒线揉头狮子补子员领", Sim = 0.59): This term contains five layers of details—"color (five colors) + craftsmanship (thread-sprinkled embroidery) + pattern (maned lion) + shape (mandarin square, round collar)"—and refers to a ceremonial robe for Ming Dynasty officials. Egerton's translation ("a robe of office with long-maned lions embroidered in five colors") simplifies the "thread-sprinkled craftsmanship", "mandarin square (rank symbol)", and "round collar", retaining only the core elements of "official robe" and "lion pattern". Roy's translation ("a round-collared robe emblazoned with a mandarin square of variegated embroidery featuring a lion with a ruffled mane") fully retains the "round collar" and "mandarin square" and provides a more accurate description of the pattern. Differences in details lead to moderate similarity.

Low Similarity Case (No. 14: "青缎五彩飞鱼蟒衣", Sim = 0.32): As mentioned in Section 3.1, Egerton's translation mistakenly identifies "flying fish and python" as "dragon" and "dark blue silk (qingduan)" as "dark green silken", showing significant differences from Roy's accurate description, resulting in extremely low semantic similarity.

4. Conclusion

This study conducts quantitative semantic similarity analysis on 35 groups of costume terms (covering clothes, headwear, footwear) from《金瓶梅》(Jīn Píng Méi) and their translations in Egerton's version and Roy's version using the mBERT model, filling the gap of insufficient quantitative methods in existing research on culturally specific term translation of the novel. The results show an overall average cosine similarity of 0.69 (moderately high consistency), verifying three laws: practical attribute intensity correlates positively with similarity (highest in Footwear, average 0.82), while cultural load and element complexity correlate negatively (lower similarity in ceremonial terms and Clothes category); high similarity samples (33.29%) focus on daily casual terms with minor non-core differences, and low similarity (11.42%) stems from cultural symbol misinterpretation or key attribute errors. The study enriches 《金瓶梅》(Jīn Píng Méi) translation research methodologically (introducing NLP quantification and constructing a "practical attribute-cultural load-element complexity" evaluation framework) and provides practical guidance for costume culture translation (prioritizing core cultural symbols) and version selection (Roy's for academia, Egerton's for general readers), though it has limitations in corpus scale (35 samples) and model singularity (only mBERT); future research could expand to more samples/translations, compare cross-models (e.g., XLM-RoBERTa), deepen cultural semantic mining, and extend to other classical novels to promote more systematic quantitative research on ancient Chinese cultural specific term translation.

Funding: This research received no external funding.
Conflicts of Interest: The authors declare no conflict of interest.
Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers.

References

[1] Al-Batineh, M., & Rabadi, R. (2019). Will the machine understand literary translation? A glimpse into the future of literary machine translation through the lenses of artificial intelligence. *Studies in Translation*, 5(1), 151–169.
[2] Bhowmik, S., Lotankar, S., Dixit, M., & Munshi, A. (2025, March). Bridging classic and modern English: An NLP approach to translation and educational chatbots in English literature. In *2025 IEEE International Conference on Interdisciplinary Approaches in Technology and Management for Social Innovation (IATMSI)* (Vol. 3, pp. 1–6). IEEE. doi: 10.1109/IATMSI64286.2025.10984542.
[3] Gardazi, N. M., Daud, A., Malik, M. K., Bukhari, A., Alsahfi, T., & Alshemaimri, B. (2025). BERT applications in natural language processing: A review. *Artificial Intelligence Review*, 58(6), 1–49.
[4] Kumar, T., Mahrishi, M., & Sharma, G. (2023). Emotion recognition in Hindi text using multilingual BERT transformer. *Multimedia Tools and Applications*, 82(27), 42373–42394. https://doi.org/10.1007/s11042-023-15150-1
[5] Pota, M., Ventura, M., Fujita, H., & Esposito, M. (2021). Multilingual evaluation of pre-processing for BERT-based sentiment analysis of tweets. *Expert Systems with Applications*, 181, 115119. https://doi.org/10.1016/j.eswa.2021.115119
[6] Wen, X. Y., & Wang, Y. C. (2025). A century of English translations of *The Plum in the Golden Vase*: The evolution of translation strategies and cross-cultural representation of Chinese image. *Translation Horizons*, (02), 57–68.
[7] Zhao, C. Y. (2020). A multi-dimensional analysis of register variation in the complete English translations of *The Plum in the Golden Vase* based on corpus. *Foreign Language Teaching and Research*, 52(02), 283–295.
[8] Zhao, C. Y., & Li, S. (2025). An analysis of readers' acceptance evaluation and improvement paths based on the Kano model: Taking the English translations of *The Plum in the Golden Vase* as an example. *Foreign Languages (Journal of Shanghai International Studies University)*, 48(04), 105–116.