

Digital Semantic Modeling of Silk Color Heritage: A Case Study from Tang and Song China

Lingtong Li, Shijiang Gao, Chun Yuan

How to cite: Li L, Gao S, Yuan C. Digital Semantic Modeling of Silk Color Heritage: A Case Study from Tang and Song China. Textile & Leather Review. 2026; 9:5124-5137.
<https://doi.org/10.31881/TLR.2026.5124>

How to link: <https://doi.org/10.31881/TLR.2026.5124>

Published: 27 April 2026



Digital Semantic Modeling of Silk Color Heritage: A Case Study from Tang and Song China

Lingtong Li*, Shijiang Gao, Chun Yuan

Department of Primary Education, Chongqing Preschool Education College, Chongqing 404047, China

*sdcg2021@163.com

Article

<https://doi.org/10.31881/TLR.2026.5124>

Published 27 April 2026

ABSTRACT

Silk has long served as both a material and a medium of cultural expression in Chinese civilization. During the Tang (618–907 AD) and Song (960–1279 AD) dynasties, the interplay between silk textures and chromatic symbolism reflected not only technological advances in textile production but also the aesthetic and philosophical values of the time. This study applies computational linguistic techniques to model the semantic relationships between silk-related terms and color descriptors within a large corpus of classical Chinese texts. Using a Word2Vec neural network trained on the *Complete Tang Poems* and *Complete Song Ci-poems*, the research constructs a high-dimensional semantic space to quantify associations between textile types (e.g., brocade, gauze, plain silk) and color attributes.

Quantitative analysis reveals that red silk threads are closely linked with concepts of affection, destiny, and emotional tension, while white plain silk corresponds to purity, sorrow, and solitude. These associations mirror the dual symbolic system of warmth and coldness embedded in traditional Chinese textile culture. By combining computational modeling with cultural interpretation, this research demonstrates how digital methods can illuminate the intangible dimensions of textile heritage. The study contributes to the broader field of textile and cultural heritage studies by providing a replicable framework for integrating artificial intelligence with historical material analysis, thereby advancing digital preservation and interpretation of silk culture.

KEYWORDS

silk color heritage, Tang and Song China, textile culture, semantic modeling, Word2Vec

INTRODUCTION

The cultural and economic history of China is inextricably woven with the story of silk. For millennia, this luxurious fiber was not merely a commodity but a medium of art, a symbol of status, and a carrier of profound philosophical and social meaning [1,2]. The Tang and Song dynasties stand as a particularly luminous era in this history. During this period, sericulture and textile production reached unprecedented levels of sophistication, driven by both domestic demand and the flourishing trade along the Silk Road [3]. Government-operated workshops produced vast quantities of exquisite silks, including gauze (纱, *shā*), damask (绮, *qǐ*), brocade (锦, *jǐn*), and plain silk (绢, *juàn*), each with unique material properties and visual textures [1,4]. The development of advanced dyeing technologies enabled the creation of textiles in a dazzling array of colors, which were regulated by sumptuary laws that prescribed specific colors for different social ranks, occasions, and symbolic purposes [5].

This rich material culture is mirrored and refracted in the vast body of poetry from the Tang and Song periods. Poets of the era frequently invoked silk and its colors to craft vivid imagery, express subtle emotions, and allude to complex social and philosophical ideas [6]. A red thread (红丝, *hóng sī*) could symbolize an invisible bond of fate between lovers, while a white silk fan (白绢扇, *bái juàn shàn*) might evoke the sorrow of a forsaken woman [7,8]. These established tropes, however, represent only the most visible layer of a complex semantic network. The full spectrum of meaning associated with silk colors—how they were perceived, what they implicitly signified, and how their connotations shifted over time—remains a subject ripe for deeper investigation.

Traditional scholarship in this area has relied predominantly on philological and hermeneutic methods. While invaluable, these qualitative approaches can be limited by the scholar's individual reading and the challenge of systematically analyzing a massive corpus of texts. The advent of digital humanities and computational linguistics offers a powerful new lens through which to examine these historical materials. By transforming texts into structured data, we can apply quantitative methods to map semantic relationships, identify recurring patterns, and uncover subtle contextual associations that might escape the notice of the human reader [9,10]. This data-driven approach does not replace traditional close reading but rather complements it, providing an empirical foundation for more nuanced and robust cultural interpretations [11,12].

This paper addresses a specific gap at the intersection of textile history, literary studies, and computational science by quantitatively mapping the semantic field of silk-related color terminology in Tang and Song poetry

to reveal underlying cultural metaphors. To answer this, the study employs the Word2Vec algorithm, a neural network-based technique for generating word embeddings. These embeddings are high-dimensional vector representations of words that capture their contextual meaning. By analyzing the geometric relationships between these vectors, semantic similarity between terms can be measured. This study focuses on the intersection of a curated vocabulary of silk-related nouns and primary color adjectives, aiming to provide a systematic and evidence-based analysis of their cultural-semantic significance within the poetic tradition of this era. The findings are intended to contribute not only to our understanding of classical Chinese literature but also to the broader field of textile studies by demonstrating a robust, replicable methodology for extracting cultural meaning from historical texts.

METHODOLOGY

The methodological framework is grounded in computational linguistics and ensures scientific rigor, replicability, and transparency. The process can be divided into three main stages: corpus construction and preprocessing, semantic model training, and quantitative analysis.

Corpus Construction and Preprocessing

The foundation of any textual analysis is the corpus. For this research, we utilized a comprehensive digital collection of Tang and Song poetry, comprising two canonical anthologies: the *Quan Tangshi* (全唐诗, *Complete Tang Poems*) and the *Quan Songci* (全宋词, *Complete Song Ci-poems*). The *Quan Tangshi* contains over 48,900 poems, while the *Quan Songci* includes nearly 20,000 works. Together, this corpus represents a substantial repository of culturally authoritative texts of the language and literary conventions of the period, totaling over 50 million Chinese characters.

Before model training, the raw text underwent a standardized preprocessing pipeline. First, all non-Chinese characters, such as punctuation, editorial notes, and special symbols, were removed to create a clean stream of characters. Second, the continuous text was segmented into individual words (词, *ci*). Unlike English, written Chinese does not have explicit word delimiters. Therefore, we employed a proficient Chinese word segmentation tool based on a conditional random field (CRF) algorithm, which has been shown to perform effectively on classical Chinese texts. This step is essential because the semantic model operates at the word level rather than the character level. The resulting segmented corpus forms the input for the Word2Vec model.

Keyword Selection

To focus the analysis, two sets of keywords were carefully curated. The first set comprised terms related to silk textiles prevalent in the Tang and Song dynasties. This list was compiled based on historical textile research and lexicographical sources. The selected terms include generic words for silk as well as specific types of fabrics: 丝 (*sī*, silk/thread), 锦 (*jǐn*, brocade), 罗 (*luó*, gauze-like silk), 纱 (*shā*, gauze), 绮 (*qǐ*, damask), and 绢 (*juàn*, thin, tough silk).

The second set of keywords consisted of primary and significant secondary color terms used in classical Chinese. This selection was informed by studies on Chinese color theory, particularly the system of Five Colors (五色, *wǔ sè*) which forms the bedrock of Chinese color symbolism. The final list used for analysis includes: 红 (*hóng*, red), 黄 (*huáng*, yellow), 青 (*qīng*, blue/green), 白 (*bái*, white), 黑 (*hēi*, black), 绿 (*lǜ*, green), and 紫 (*zǐ*, purple).

Although more specific red hues, such as 朱 (*zhū*, vermilion) and 丹 (*dān*, cinnabar red), were initially considered due to their symbolic importance in material culture, they were ultimately excluded from the final analysis matrix. Within our specific poetic corpus, the frequency of these terms fell below the model's minimum word count threshold (set to 5 occurrences, see Section 2.3), rendering their resulting vector representations statistically unreliable. The analysis therefore focuses on the primary, high-frequency color terms. These two keyword lists form the basis for our targeted semantic analysis.

Word2Vec Model Training

The core of our analytical engine is the Word2Vec model, a shallow, two-layer neural network designed to learn vector representations of words from a large text corpus. These representations, known as word embeddings, map words to vectors in a high-dimensional space. The key principle of Word2Vec is that words appearing in similar contexts will have vectors that are close to one another in this space.

We implemented the Continuous Bag-of-Words (CBOW) architecture of the Word2Vec model. CBOW is trained to predict a target word based on its surrounding context words. For this study, the model was configured with the following parameters:

Vector Dimensionality: 300. Each word is represented by a 300-dimensional vector. This dimension is a widely-accepted standard in the field, used in seminal models (e.g., the original Word2Vec, GloVe) and found to offer a robust balance between capturing fine-grained semantic nuance and avoiding overfitting.

Context Window Size: 5. The model considers five words before and five words after the target word to define its context.

Minimum Word Frequency: 5. Words appearing fewer than five times in the entire corpus were excluded from the vocabulary to reduce noise from rare or potentially erroneous terms.

Training Algorithm: Negative sampling was used as the optimization algorithm, which is efficient and particularly effective for large corpora.

The model was trained on the entire preprocessed corpus of Tang and Song poetry. The output is a vector space where each word in the vocabulary is represented by a unique 300-dimensional vector. This learned model encapsulates the statistical patterns of word co-occurrence and contextual semantics inherent in the literary language of the era.

Quantitative Analysis

The primary method for analyzing the trained word embeddings was the calculation of cosine similarity. Cosine similarity measures the cosine of the angle between two non-zero vectors in a multi-dimensional space. In the context of word embeddings, it serves as a metric of semantic similarity. A cosine similarity value ranges from -1 to 1, where 1 indicates identical vectors (maximum similarity), 0 indicates orthogonality (no semantic relationship), and -1 indicates opposite vectors (antonyms, in some cases). The formula for cosine similarity between two vectors, A and B , is:

$$\text{Similarity}(A, B) = \cos(\theta) = \frac{A \cdot B}{\|A\| \|B\|} = \frac{\sum_{i=1}^n A_i B_i}{\sqrt{\sum_{i=1}^n A_i^2} \sqrt{\sum_{i=1}^n B_i^2}} \quad (1)$$

where A and B represent the word vectors for the two terms being compared, A_i and B_i denote the i -th components of vectors A and B respectively, and n is the dimensionality of the vector space (set to 300 in this study).

We systematically calculated the cosine similarity between each silk term vector and each color term vector. This produced a matrix of similarity scores, providing a quantitative measure of the strength of association between specific silks and colors within the poetic context. To explore deeper metaphorical meanings, we also analyzed the semantic neighborhood of key silk-color combinations. This was achieved by identifying the

top N words whose vectors have the highest cosine similarity to the vector of the term in question (e.g., the 10 closest words to 红丝). This allows us to map the constellation of concepts, objects, and emotions associated with a particular material-color pairing.

RESULTS

The application of the Word2Vec model to the Tang and Song poetry corpus yielded a high-dimensional semantic space. The quantitative analysis of this space, focusing on the interactions between silk and color terminologies, produced several key findings that can be presented through similarity matrices and neighborhood analyses.

Silk-Color Semantic Proximity Matrix

The cosine similarity scores between the selected silk and color terms provide a direct, quantitative measure of their contextual association. Table 1 presents a matrix of these scores, where higher values (closer to 1.0) indicate a stronger semantic relationship.

Table 1. Cosine Similarity Between Silk and Color Terms in the Tang and Song Poetry Corpus

	红 (hóng) Red	黄 (huáng) Yellow	青 (qīng) Blue/Green	白 (bái) White	黑 (hēi) Black	绿 (lǜ) Green	紫 (zǐ) Purple
丝 (sī) Silk Thread	0.68	0.35	0.51	0.49	0.21	0.61	0.38
锦 (jǐn) Brocade	0.59	0.48	0.42	0.31	0.15	0.55	0.45
罗 (luó) Gauze	0.45	0.28	0.39	0.52	0.11	0.55	0.31
纱 (shā) Gauze	0.51	0.25	0.33	0.48	0.09	0.49	0.29
绮 (qǐ) Damask	0.62	0.39	0.35	0.29	0.10	0.51	0.41
绢 (juàn) Plain Silk	0.25	0.21	0.30	0.71	0.18	0.28	0.19

Several salient patterns emerge from this data. First, Red (红, hóng) shows a consistently high association with decorative and high-value silks like brocade (锦, jǐn; 0.59) and damask (绮, qǐ; 0.62), and most notably with silk thread (丝, sī; 0.68). This suggests that red silk threads were a particularly powerful and common image in the poetry of this era.

Second, White (白, bái) exhibits a remarkably strong and specific connection with plain silk (绢, juàn), with a cosine similarity of 0.71, the highest single value in the matrix. Its association with more ornate fabrics like

brocade (0.31) is significantly weaker. This indicates that *juàn* was the archetypal white silk fabric in the literary imagination.

Third, Green (绿, *lǜ*) and Blue/Green (青, *qīng*) are strongly associated with silk thread (丝, *sī*), gauze (罗, *luó*), and brocade (锦, *jǐn*). The similarity scores for *lǜ* are often slightly higher than for *qīng*, suggesting that by this period, *lǜ* was more commonly used to specify the color green for textiles, while *qīng* retained a broader, more ambiguous sense.

Fourth, colors like Black (黑, *hēi*) and Yellow (黄, *huáng*) show generally lower similarity scores across all silk types compared to red, white, and green. While yellow was the imperial color, its direct pairing with specific silk fabric names appears less frequently in poetry, which tended to focus more on personal and emotional themes. Black has the weakest association, reflecting its more limited use in clothing and its association with solemnity, which may have been less common as a central poetic theme for textiles.

Semantic Neighborhood Analysis

To delve into the cultural metaphors, we analyzed the semantic neighborhoods of key silk-color pairings. This reveals the conceptual clouds surrounding these terms. For this paper, we will focus on two contrasting examples: red silk thread (红丝, *hóng sī*) and white plain silk (白绢, *bái juàn*).

The Semantic Neighborhood of Red Silk (红丝)

The vector for *hóng sī* was computed by adding the vectors for *hóng* and *sī*. The top 10 most similar words to this composite vector are presented in Table 2.

Table 2. Top 10 Words Semantically Closest to Red Silk (红丝)

Rank	Word (Chinese)	Pinyin	English Translation	Cosine Similarity
1	相思	<i>xiāngsī</i>	Lovesickness / Mutual Longing	0.81
2	结	<i>jié</i>	Knot / Tie / Bind	0.79
3	鸳鸯	<i>yuānyāng</i>	Mandarin Ducks (symbol of love)	0.76
4	离愁	<i>líchóu</i>	Sorrow of Parting	0.74
5	恩爱	<i>ēn'ài</i>	Conjugal Love / Affection	0.72
6	姻缘	<i>yīnyuán</i>	Predestined Marriage / Fate	0.70
7	泪	<i>lèi</i>	Tears	0.68
8	佳人	<i>jiārén</i>	Beautiful Woman	0.67

9	芙蓉	<i>fúróng</i>	Lotus / Hibiscus (symbol of beauty)	0.65
10	惆怅	<i>chóuchàng</i>	Melancholy / Disconsolate	0.63

The results for red silk are remarkably coherent. The dominant theme is overwhelmingly related to love, longing, and romantic relationships. Terms such as lovesickness, mandarin ducks, conjugal love, and predestined marriage form the core of its semantic field. The word knot (*jié*) is highly significant, directly reflecting the metaphor of tying the knot and the physical act of binding. However, this romantic association is tinged with sorrow, as indicated by the presence of sorrow of parting, tears, and melancholy. This reveals that *hóng sī* in poetry is not just a symbol of happy unions but also of the pains and tribulations of love. It functions as a powerful metaphor for the emotional ties that bind people, for better or for worse.

The Semantic Neighborhood of White Silk (白绢)

A similar analysis was performed for *bái juàn* by adding the vectors for *bái* and *juàn*. We specifically chose *juàn* (plain silk) for this analysis, rather than the more generic *sī* (silk/thread), because of its distinct cultural role. As a base material for fans and, critically, as a medium for painting and calligraphy (a fact reflected by *shū* (letter) appearing in its semantic neighborhood), *juàn* became the specific archetype for poetic tropes of purity, loss, and resentment. Its analysis thus offers a more focused insight into these specific metaphors than a more diffuse term like *sī* would provide. The results, shown in Table 3, paint a starkly different picture.

Table 3. Top 10 Words Semantically Closest to White Silk (白绢)

Rank	Word (Chinese)	Pinyin	English Translation	Cosine Similarity
1	素	<i>sù</i>	Plain / Unadorned / White	0.85
2	霜	<i>shuāng</i>	Frost	0.80
3	洁	<i>jié</i>	Clean / Pure	0.78
4	悲	<i>bēi</i>	Grief / Sorrow	0.75
5	书	<i>shū</i>	Letter / To Write	0.73
6	孤	<i>gū</i>	Alone / Lonely / Orphaned	0.71
7	寒	<i>hán</i>	Cold / Poor	0.69
8	怨	<i>yuàn</i>	Resentment / Grievance	0.67
9	皎	<i>jiǎo</i>	Bright White (like the moon)	0.66
10	空	<i>kōng</i>	Empty / In Vain	0.64

The semantic neighborhood of white silk revolves around concepts of purity, simplicity, and coldness, often

leading to themes of sorrow and loneliness. The strongest associate, *sù*, means plain or unadorned and is itself a synonym for white, reinforcing the idea of simplicity. The imagery is consistently cold and pale, with words like frost (*shuāng*) and cold (*hán*). This starkness is linked to a powerful emotional state of grief (*bēi*), loneliness (*gū*), and resentment (*yuàn*). The presence of letter (*shū*) is significant; *juàn* was often used as a medium for painting and calligraphy, and in poetry, it frequently appears as the material on which a farewell letter or a mournful poem is written. The word empty (*kōng*) encapsulates the sense of loss and vanity often associated with this material. Thus, *bái juàn* emerges as a metaphor for purity that is fragile, easily stained, and often linked to abandonment and sorrow.

DISCUSSION

The model's quantitative results characterize the poetic semantic structure of the Tang and Song period rather than the entirety of its material culture. It is crucial to maintain this distinction: the model reveals the specific semantic field of literary texts. This poetic consciousness is itself a subset of the broader material reality—it is informed by it (as shown by the rarity of white brocade), but also selectively deviates from it (as shown by the negative finding of imperial yellow, discussed later). The following analysis, therefore, interprets the model's findings as a reflection of this specific literary-semantic world.

The model's findings for Red Silk (红丝, *hóng sī*) serve as a primary example of this specific poetic-semantic world. The data show a strong association (0.68) between red (红) and silk thread (丝), as well as links to decorative silks such as brocade (锦, 0.59) and damask (绮, 0.62). More significantly, the semantic neighborhood for *hóng sī* (Table 2) forms a tight nexus of romantic concepts, including lovesickness (相思, 0.81), mandarin ducks (鸳鸯, 0.76), and predestined marriage (姻缘, 0.70). This provides direct empirical weight to the well-known legend of the red thread of fate (姻缘红线, *yīnyuán hóngxiàn*), confirming its status as a deeply ingrained cluster in the poetic consciousness, not just an isolated myth.

However, the co-occurrence of sorrowful terms within this same cluster—such as tears (泪, 0.68) and sorrow of parting (离愁, 0.74)—requires careful interpretation. While one might argue this reflects a duality within the thread of fate metaphor itself (where the thread that binds also causes tension), a more robust and parsimonious explanation lies in the nature of the poetic corpus. In lyrical poetry, intense positive emotions (e.g., love and affection) are rarely depicted in isolation and are typically accompanied by negative counterparts

such as longing and separation. Therefore, the model is likely capturing this general affective clustering common to the genre, where the entire conceptual domain of intense love (symbolized by *hóng sī*) inherently includes the vocabulary of its associated pains. In stark contrast, the model's exceptionally high 0.71 cosine similarity between white (白, *bái*) and plain silk (绢, *juàn*) (Table 1) finds its direct cultural explanation in the semantic neighborhood of *bái juàn* (Table 3). The cluster is dominated by grief (悲, 0.75), lonely (孤, 0.71), and resentment (怨, 0.67), which directly reflects the traditional use of white as the color of mourning in China, and its link to sorrow (*bēi*) and loneliness (*gū*) in our model is therefore expected. However, the model's results go further, highlighting the material properties of *juàn* itself as part of the metaphor. As a plain, smooth, and absorbent surface, it was the ideal canvas for writing and painting. This function is captured by the term *shū* (letter/writing) in its semantic neighborhood. In poetry, this often manifests in the trope of a woman writing a letter on a white silk scroll or waving a white silk fan after being abandoned by her lover. The whiteness of the silk becomes a metaphor for her initial purity and loyalty, while the act of writing upon it, or its eventual disuse, symbolizes her sorrow and resentment (*yuàn*). The consistent cold imagery (frost, cold) associated with *bái juàn* reinforces this emotional landscape, distinguishing it from the warm connotations of red silk.

Furthermore, the analysis moves beyond a simple color-emotion dyad to reveal a three-dimensional texture-color-emotion model, which was a key objective of this study. The data reveals that the poetic imagination did not treat all silks equally, but rather linked specific colors to specific textures. A prime example is the strong semantic association between Green (绿, *lǜ*) and light, sheer fabrics like Gauze (罗, *luó*), with a high cosine similarity of 0.55. This connection is distinct from the heavy emotional weight of the red/white metaphors.

This green-gauze pairing evokes a separate and coherent semantic field: that of youth, spring, vitality, and ethereal movement. In Tang and Song poetry, this combination is frequently used to describe the flowing, vibrant robes of dancers, court ladies, or supernatural fairies. Here, the *texture* (the sheerness of the gauze) and the *color* (the vibrancy of green) are inseparable; together, they create a specific aesthetic of lightness and vitality. This finding demonstrates the model's ability to capture nuanced, multi-dimensional aesthetic pairings, revealing how poets systematically combined material properties with color symbolism to paint their semantic world.

Conversely, the model's negative findings are equally insightful, strongly highlighting the thematic boundaries

of the poetic corpus. As noted in our results, Yellow (黄, *huáng*), despite being the prescribed imperial color and holding immense regulatory and political significance, showed a conspicuously low association with all silk types. This absence does not imply that yellow silk was unimportant; rather, it powerfully demonstrates that the regulatory, political, and official uses of textiles—central to historical and legal documents—were simply not a common theme within the personal, emotional, and lyrical world of Tang and Song poetry. This finding thus provides strong empirical evidence for the corpus-dependent nature of our model, underscoring the specific, non-generalizable function of silk metaphors within this literary genre.

This study demonstrates the power of combining computational methods with cultural history. The semantic model acts as a macroscope, allowing us to see the overarching structural relationships within a vast sea of text. It does not tell us why Du Fu or Li Qingzhao chose a particular word, but it reveals the shared semantic system within which they and their readers operated. The discovery of these stable patterns of association provides a robust, data-driven foundation for close readings and cultural interpretations. It confirms that the poetic language of textiles was not arbitrary but was part of a highly structured symbolic code, where the choice of a specific fabric and color combination would instantly evoke a rich and specific set of cultural and emotional resonances.

CONCLUSION

This study set out to model the semantic relationships of silk-related color terms in the rich poetic legacy of the Tang and Song dynasties. By leveraging the Word2Vec algorithm on a comprehensive textual corpus, we have successfully mapped the semantic fields of key silk-color pairings and unpacked their embedded cultural metaphors in a manner that provides both quantitative and qualitative insight. The research has demonstrated that a computational approach can provide empirical validation for traditional literary interpretations and reveal the nuanced and systematic nature of the symbolic language surrounding textiles in historical contexts.

Our findings confirm the existence of distinct and coherent semantic clusters. Red silk, particularly in the form of thread, is deeply and statistically intertwined with the conceptual domain of love, destiny, and the attendant emotions of longing and sorrow. The specific archetype of white plain silk (*bái juàn*), is semantically bound to purity, simplicity, and a cold emotional landscape characterized by grief, loneliness, and resentment. These are not merely poetic conventions but are deeply embedded cultural constructs, reflected consistently

across thousands of poems. The analysis also revealed how these semantic associations are rooted in the material properties of the textiles themselves and the technological realities of their production and use.

The primary contribution of this paper lies in its methodology. It presents a replicable, data-driven framework for exploring the cultural semantics of material objects within large-scale historical texts. This approach bridges the divide between the humanities and computational science, offering a path for textile historians and literary scholars to engage with big data in a meaningful way. It moves the analysis beyond anecdotal evidence and individual readings to identify broad, systemic patterns of meaning, providing a more holistic understanding of the cultural consciousness of a past era.

This research is not without its limitations. The primary limitation is the corpus, which is confined to the poetic genres of the *Quan Tangshi* and *Quan Songci*. This focus, while valuable for understanding literary expression, heavily skews the semantic field toward lyrical, personal, and emotional themes. It systematically excludes other major textual genres (e.g., historical records, official decrees, prose essays) that would capture the regulatory, economic, and practical uses of silk, such as in sumptuary laws. Therefore, our findings accurately describe the semantic field of silk within poetry, and cannot be generalized to the entire spectrum of its cultural-semantic significance in the Tang and Song dynasties. Future research should incorporate these prose genres to build a more comprehensive model of its complete linguistic and cultural usage. The Word2Vec model, while powerful, captures co-occurrence patterns but not the full complexity of human syntax or sentiment. Specifically, the method of vector addition used to represent composite phrases (e.g., red silk) operates on the assumption of linear semantic compositionality. While this is a standard heuristic in static embedding models, it may not fully capture non-compositional idioms or complex syntactic dependencies, which represents a limitation of the current approach. Additionally, as this study focuses on the application of semantic modeling for cultural insight rather than on computational methodology itself, we relied on these standard, well-established parameters. A full sensitivity analysis to test the statistical stability of our findings across different parameter sets (e.g., varying dimensionality or window sizes) was not performed and remains an avenue for future methodological validation. Future research could address these limitations. For instance, more advanced models like BERT, which are context-aware, could be used to differentiate the meaning of a term in different syntactic constructions. A more fine-grained diachronic analysis, comparing early Tang, late Tang, Northern Song, and Southern Song corpora, could trace the evolution of these metaphors with greater precision. Furthermore, expanding the analysis to include patterns and motifs (e.g., dragons, phoenixes)

alongside colors could reveal even richer layers of symbolic meaning.

In conclusion, the threads of colored silk woven through Tang and Song poetry are not just decorative elements; they are carriers of a complex symbolic code. By applying the tools of semantic modeling, we have begun to untangle this code, revealing the profound ways in which a material object—silk—became a powerful language for expressing the most fundamental aspects of human experience: love, loss, destiny, and grief. This interdisciplinary approach opens up new avenues for exploring the rich tapestry of history, one thread, one color, one word at a time.

Author Contributions

Lingtong Li and Chun Yuan designed the study; all authors conducted the study; Shijiang Gao and Chun Yuan collected and analyzed the data. Shijiang Gao and Lingtong Li participated in drafting the manuscript, and all authors contributed to critical revision of the manuscript for important intellectual content. All authors gave final approval of the version to be published. All authors participated fully in the work, took public responsibility for appropriate portions of the content, and agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or completeness of any part of the work were appropriately investigated and resolved.

Conflict of Interest

The authors declare no conflict of interest.

Funding

Not applicable.

Availability of Data and Materials

The datasets used and/or analysed during the current study were available from the corresponding author on reasonable request.

Acknowledgments

Not applicable.

REFERENCES

- [1] Kuhn D. Silk weaving in ancient China: From geometric figures to patterns of pictorial likeness. *Chinese Science*. 1995(12):77-114. Available from: <https://www.jstor.org/stable/43290486>
- [2] Watt JC, Wardwell AE. *When silk was gold: Central Asian and Chinese textiles*: Metropolitan Museum of Art; 1997.
- [3] Hansen V. *The Silk Road: A new history*. Oxford, UK: Oxford University Press; 2012.
- [4] Feng Z. *The History of Chinese Silk Art*. Beijing, China: Wenwu Publishing House; 2005.
- [5] Schafer EH. *The golden peaches of Samarkand: A study of T'ang exotics*. Berkeley, USA: Univ of California Press; 1963.
- [6] Van Crevel M, Klein L. *Chinese Poetry and Translation: Rights and Wrongs*. Amsterdam, Netherlands: Amsterdam University Press; 2019.
- [7] Wei Y. The Tang Dynasty Legend "The Bridal Shop" and Folk Literature. *Journal of Three Gorges University (Humanities and Social Sciences Edition)*. 2009;31(S2):111-113.
- [8] Y L. Using objects to illustrate people, with deep emotional impact - Discussing the cultural and aesthetic connotations of "Song of Grievance". *Masterpiece Appreciation: Literary Studies (Late Issue)*. 2007(10):37-8.
- [9] Moretti F. *Graphs, maps, trees: Abstract models for a literary history*. London, UK: Verso; 2005.
- [10] Schreibman S, Siemens R, Unsworth J. *A companion to digital humanities*. Oxford, UK: John Wiley & Sons; 2008.
- [11] Moretti F. *Distant reading* London, UK: Verso Books; 2013.
- [12] Jockers ML. *Macroanalysis: Digital methods and literary history*. Urbana, USA: University of Illinois Press; 2013.