

A Comparison of Chinese and Western Poetry from the Perspective of Digital Humanities: Taking Li Bai and Shakespeare as Examples

Youyou Peng

*High School Affiliated to Renmin University of China, Beijing, China
13693632189@163.com*

Abstract. From the perspective of digital humanities, a comparative study of Li Bai's and Shakespeare's poetry reveals the cultural gene differences and spiritual resonances between Chinese and Western classical poetry. Through text analysis via high-frequency word statistics and language models, it is found that Li Bai's poetry conveys the poet's feeling through presenting natural imagery and geographical space, while Shakespeare's poetry focuses on human's feelings and explores the themes of time and love. Comparative studies show that the intensity of emotional fluctuations in Li Bai's poetry is significantly higher than that in Shakespeare's, that Shakespeare's poems are generally more positive, but both share the spiritual core of "transcendent value pursuit". This study confirms that digital technology can penetrate the surface of language, reveal the deep coding mechanism of cultural genes, and provide a new method for comparative poetics research.

Keywords: Digital Humanities, Poetry, Li Bai, Shakespeare

1. Introduction

Representative of their corresponding era, Li Bai and Shakespeare are two of the most prominent poets of the East and the West. Thus, the similarities or differences of many aspects of their poem can yield rich results that can reflect about the differences of the eras they lived in.

Most studies about poetry relied on close reading of a text, while they yield excellent insights when conducted by experienced researchers, this approach result in heavily subjective interpretations, and can hardly detect features in the macro level.

In this paper, we aim to explore the use digital humanity based text analysis methods such as high-frequency word statistics and language models to explore the similarities and differences of Li Bai's and Shakespeare's styles in their poems, specifically about imagery, tone, and emotions.

First, we counted the words of high frequency in the poems of the two poets, and analyzed the different common types of objects appearing their poems.

Second, we use the natural language processing library nltk and pypinyin to conduct meter analysis on the poems.

Finally, we also used the language model SnowNLP (for Chinese) to estimate the general sentiment in the works of the two poets, while using the Chinese version of Shakespeare's poems to

unify the standards.

Through this research, we not only hope to deepen our understanding of the poetic art of Li Bai and Shakespeare, but also aim to demonstrate the effectiveness of digital humanities methods in cross-cultural poetry studies, providing new perspectives for traditional literary research.

2. A comparative analysis of high-frequency words

A comparative analysis of high-frequency words in the works of Li Bai and Shakespeare can reveal the similarities and differences in the creative styles and ideological contents of the two poets. This not only helps to gain a deeper understanding of the differences and commonalities between the literary traditions of the East and the West, but also, through objective data analysis, reveals the poets' preferences for word choice, theme selection and rhetorical devices.

Table 1. High-frequency words in Li Bai's poems and their image classification

High Frequency words	Occurrence frequency	Meaning types
people	722	Characters
day	642	Nature
mountain	568	Nature
cloud	543	Nature
The wind	522	Nature
month	493	Nature
jun (you)	488	Characters
water	455	Natural
day	419	Nature
I	367	Characters
gold	333	Nature
The sea	313	Nature
heart	302	Character
flowers	297	Natural
jade	279	Natural
The autumn	276	Nature
In the spring	241	Nature

Table 2. High-frequency words in Shakespeare's poems and their classification of intentions

High Frequency words	Frequency of occurrence	Meaning types
love/love's	169 + 29	Human emotions
eyes/eye	52 + 39	Life Time
sweet	55	Human emotions
time	53	Life Time
beauty	52	Human emotions
art	50	Art Religion
heart	50	Human emotions
fair	45	Human emotions
like	35	Human emotions
live	29	Life Time
praise	28	Human emotions
world	27	Life time
best	25	Human emotions
truth	25	Life time

Li Bai's poems are known for their profound depictions of nature. By analyzing the high-frequency words, we can also see that natural imagery dominates his verses. In his poems, landscape elements and geographical space are presented very frequently and vividly, which he conveys not only their form, but also their spirit. For example, "The Heavenly Gate cuts off the opening of the Chu River, and the clear water flows eastward but takes a turn here." ("Looking at Heavenly Gate Mountain") and "The cries of the apes on both banks cannot cease, my skiff has passed over ten thousand mountains." ("Early Departure from White Emperor City") and other verses blend the beauty of the mountains and waters with the poet's emotions. Poetry often spans vast geographical spaces, from the majestic Chang'an to the remote frontier, from the hardship of the Shu Road to the vastness of the rivers and lakes, and wherever his footprints go, his pen follows. "At night, I leave Clear Stream for the Three Gorges, and I miss you upon reaching Yuzhou." (The Moon over Mountain Emei) is a prominent example.

The dominance of natural imagery in Li Bai's poetry reflects his love and reverence for nature. Through the use of natural imagery, he not only expressed his emotions but also conveyed his praise for the beauty of nature. "Nature called upon Li Bai. The Yellow River, the East Sea, and the misty spring scenery provided him with poetic materials. Li Bai did not disappoint the world's expectations and became a singer of the prosperous Tang Dynasty and a superstar of the poetic land." [1] Li Bai's landscape poems are not merely depictions of natural landscapes, but also contain profound philosophical reflections. He expressed his insights into life, the universe and nature, such as the yearning for a free and unrestrained life, the pursuit of heroic ideals, and the lamentation of the impermanence of life. The geographical space in Li Bai's poetry is not only the actual geographical location, but also the place where his emotions are expressed. Through the depiction of different geographical Spaces, Li Bai expressed his longing for his hometown, his remembrance of his friends, and his desire to explore the unknown world. "The philosophical contemplation and emotional exploration of elements in nature in Li Bai's poems have enabled his works to transcend

the limitations of time and become an important way for later readers to explore themselves and understand nature.” [2]

On the other hand, Shakespeare’s poems are more focused on human feelings instead of nature. Shakespeare constructs a torn conversational tension in his sonnets through dense second-person pronouns, such as the repeated questioning of "thou" and "thee". This deliberate strategy of breaking away from the lyrical monologue makes the poem no longer a one-way emotional outpouring, but a field of intense debate between the poet and the virtual object, the self and the mirror image. At the level of thought, the pair of high-frequency words "time" and "love" reveal the poet's hidden index of the reconfiguration of modern values. This duality precisely reflects the fundamental predicament of humanism during the Renaissance: when the eternal picture offered by religion collapses, poets attempt to re-anchor transcendent values between physical desires and artistic creation, and the sonnet is the miniature laboratory of this spiritual experiment. “Shakespeare's achievements in sonnets were so remarkable that his works are generally regarded as the pinnacle of sonnet art in England during the Renaissance.” [3] “Shakespeare's tragedies reflect the crisis of human identity during the Renaissance, that is, the issue of how to shape human identity after breaking away from the imitation of Christ, the position of human beings in the universe, and the relationship between human beings and history, as well as between human beings and nature, all need to be rethought.” [4].

3. Analysis of poetic meter

3.1. The metrical structure of Shakespeare's verses

In poetry, stress is an important element that constitutes rhythm. According to the accent markings, it can be seen that most of Shakespeare's verses follow a rhythm of "alternation of light and heavy", with five rhythmic steps (10 syllables) per line. This stable rhythm gives the recitation of poetry musicality, reflecting the Renaissance's inheritance of the classical standard cadence. “Shakespeare's sonnets are a strictly regulated lyrical poetic form.” [5] Shakespeare's sonnets typically follow the modulated five-step pattern, that is, five steps per line, each step consisting of one unstressed syllable followed by one stressed syllable. This rhythmic pattern gives the poem musicality and a sense of rhythm. The description of the core concept has an extremely high frequency of accent, which serves to reinforce the theme.

We collected the use of stress and found through screening stress: The line THEN LET NOT WINTER's RAGGED HAND DEFACE in THEE THY SUMMER ere THOU BE distilled has the most stress and reflects the poet's intense mood.

This line is the first line of Shakespeare’s Sonnet 6. The stress in "Then let not winter's ragged hand deface" falls on "winter" and "ragged hand", highlighting the cruelty and destructiveness of winter. This emphasis, through the use of stress, allows the reader to strongly sense the threat that winter's "rough hands" pose to good things.

In the poem, stress is also used to emphasize the theme of life and death. For example, the stress in "ere thou be distilled" emphasizes the urgency of prompt action, suggesting that beauty will fade if it is not acted upon in time, and is associated with death and decay. These accents not only enhance the rhythm and musicality of the poem, but also deepen the theme and emotional expression of the poem by emphasizing specific words and images. “Shakespeare's sonnets, in addition to the "pause" or "tone group", also emphasize the arrangement of light and heavy tones” [6] Through the stress, the poet successfully conveys anxiety about the passage of time, the value of beauty, and the longing for eternity, making this poem a classic among Shakespeare's sonnets.

3.2. The metrical structure of Li Bai's verses

Li Bai's poetry mainly consists of five-character ancient poems (five-quatrains, five-regulated verse) and seven-character ancient poems (seven-quatrains, seven-regulated verse), as well as miscellaneous poems (such as "Hard is the Shu Road").

According to the rules of tonal patterns, it can be analyzed that the verses generally follow the basic rule of the alternation of tonal patterns in classical poetry, and the sentence patterns start with "level"("平") or "oblique"("仄"), such as:

Five characters, starting with level: 平平平仄仄 → 仄仄仄平平

Five characters, starting with oblique: 仄仄平平仄 → 平平仄仄平

Seven characters, starting with level: 平平仄仄平平仄 → 仄仄平平仄仄平

Seven characters, starting with oblique: 仄仄平平平仄仄 → 平平仄仄仄平平

On the basis of the basic rules of the classical meter, Li Bai did not adhere to the strict tones and often broke the rules by using various techniques. "Li Bai was adept at using romantic artistic techniques to express his inner emotions through poetry. The positive and engaged Confucian scholar's mindset and the Taoist spirit of pursuing freedom contained in his poems present a powerful force of life." [7]

Also, we should take into account that some poems have their own emotional particularities and thus exhibit a unique and unrestrained style in terms of meter. This is especially true of "Hard is the Shu Road". In "Hard is the Shu Road", three, four, five and seven characters are interlaced, and the tones are adjusted flexibly according to the sentence structure. Through the combination of long and short phrases and the contrast of tones, a strong musicality is formed, such as "Oho! Behold! How steep! How high! The Shu Road is harder than to climb the blue sky!"

In general, the rhythm of Li Bai's poetry follows the traditional tonal framework and the Tang Dynasty's five-character and seven-character tonal and rhyming traditions. At the same time, on this basis, Li Bai shows the characteristics of free innovation, breaking the rigidity through inconsistent sentence lengths, twists and turns, and free rhyme changes. He controls the form with unrestrained emotions, embodying the unique character of himself, or the "poet immortal". His style is not only an inheritance of the poetry of the Tang Dynasty, but also a transcendence of the musicality and free spirit of poetry.

4. Emotional analysis

Emotion is an important subject of study in ancient poetry. "Li Bai's ontological model also encompasses multiple dimensions such as characters, time, events and emotions." [8] The degree of expression of emotion also varies in Li Bai's and Shakespeare's poems. Emotions are divided into positive and negative ones. Unlike traditional read-based sentiment analysis methods, the SnowNLP library was used to conduct sentiment analysis on the Chinese translations of all Li Bai poems and all translated Shakespeare poems and calculate the overall sentiment tendency values, resulting in a sentiment tendency value of 0.546 for Li Bai poems and 0.646 for Shakespeare poems, with the tendency value getting closer to 1. This indicates that the sentiment is more positive; The closer the propensity value is to 0, the closer the emotion is to negative. Here, the translation of all of Shakespeare's poems is intended to unify the standards and eliminate the differences in standards resulting from different ratings and standards of models for different languages, thus facilitating comparison. From this, it can be seen that Shakespeare's poems are generally more emotionally positive than Li Bai's. It can also be seen that the positive or negative distribution of the themes of love and tragedy in Shakespeare's works may differ from Li Bai's theme of seclusion.

Shakespeare's poems seem more positive than Li Bai's, which is related to both the environment they were in and the purpose of writing their poems. Shakespeare lived in Renaissance England, a time when people began to value human values, and many of his poems explored eternal themes such as love and time. At that time, the literary tradition focused on chronologic narration, but Shakespeare's works broke the conventions of time flow [9]. His most famous sonnet, Sonnet 18, compares the beauty of a lover to an eternal summer, and says that as long as someone reads it, the beauty of a lover will never fade. This way of writing turns sad emotions into positive meanings in the end. Li Bai's situation is a bit different. He lived during the period when the Tang Dynasty was on the decline, and he himself had been exiled, so two opposite emotions often appeared in his poems. The ups and downs of emotion makes it not easy to give a stable positive impression.

The influence of cultural traditions should not be overlooked in the context of emotions. Since ancient Greece, Western literature has emphasized spiritual purification through stories, and Shakespeare's poems often have this sense of elevation [10]. Chinese poetry, on the other hand, places more emphasis on the natural expression of genuine feelings. Li Bai's seemingly negative lines, such as "I raise my cup to invite the bright moon, and with my shadow we are a group of three," can actually evoke poetic beauty in solitude in the Chinese context. This cultural difference may lead emotional-analysis tools to be more inclined to identify clear positive signals in Western literature.

5. Conclusion

This study explores the similarities and differences in the characteristics of Li Bai's and Shakespeare's poems by counting high-frequency words and invoking language models. The study reveals that the works of Li Bai and Shakespeare vary greatly in style due to the influence of different eras. The study explores the application of digital methods in the analysis of literary works.

The analysis of high-frequency words shows that the images in Li Bai's poems are mostly natural landscapes such as mountains and waters, while Shakespeare's poems pay more attention to human emotions. This difference is not only due to the differences among individual poets, but also because of the differences in eras. Li Bai lived during the period when the Tang Dynasty was on the decline. Due to the country's turmoil and the lack of success in his official career, he chose to express his feelings to the mountains and waters, and thus the mountains and waters were important subjects of his emotions in his poems. Shakespeare lived during the Renaissance, when humanism was on the rise, and thus his depictions of human emotions were more profound. In terms of poetic meter, Li Bai's poems are mainly five-character and seven-character ancient poems. Although they follow the traditional framework of tonal patterns, they often break conventions, showing the poet's characteristic of free innovation. Shakespeare's sonnets, on the other hand, are stricter embodying the Renaissance's inheritance of classical standards. These differences in rhythm, on the one hand, reflect the differences in the poetic traditions of the East and the West, and on the other hand, reflect the innovations and breakthroughs of the two poets in their respective traditions. The analysis of emotions shows that Shakespeare's emotions are more positive than Li Bai's, which also confirms the difference in the eras of the two poets.

Through a comparative study of Li Bai and Shakespeare's poetry using a digital humanities approach, we not only see the differences in form and content between Chinese and Western poetry, but also discover commonalities that transcend cultural boundaries. Both poets, with their unique artistic style and profound ideological content, have left behind precious cultural legacies for future generations. This cross-cultural comparative study not only helps us gain a deeper understanding of

the literary traditions of China and the West, but also provides new perspectives and methods for contemporary literary studies.

Digital humanities, as an emerging area of research, uses objective and systematic analytical tool for literary studies, such as data mining and sentiment analysis. This study attempts to apply such methods to the comparison of Chinese and Western poetry, demonstrating the potential of combining digital technology with traditional humanities research. In the future, with the continuous development of digital technology and the deepening of humanities research, the comparison of Chinese and Western literature from the perspective of digital humanities will have broader research space and more profound academic significance.

References

- [1] Mei Yuling. Integrating Individual Life into the Flow of Greater Life: An Interpretation of the Life Theme in "Spring River Moonlit Night" [J]. *Appreciation of Masterpieces*, 2025, (17): 80-84.
- [2] Chen Xiaoli. Thoughts on Li Bai's Poem "Moon" [J]. *Extracurricular Chinese*, 2025, (01): 79-81.
- [3] Li Shaoshan. Variants and Integration of Cultural Construction: A Review of Zhang Xiangping's "Seven-Character Regulated Verse from Shakespeare's Sonnet" [J]. *Masterpieces by Renowned Artists*, 2025, (13): 1-6.
- [4] The Tragic Dimension of Shakespeare's History in the Novel Creation by Xie Mei Pasternak and Hardy [J]. *Russian literature and art*, 2025, (02) : 103-111. The DOI: 10.16238 / j.carol carroll nki rla. 2025.02.011.
- [5] Ma Qing. Tu An: The Ferryman of British and American Poetry [N]. *Social Sciences Today*, 2023-07-27(008).
- [6] Xiong Hui. Wen Yiduo's Poetic Style Concepts in Translation [C]// School of Literature, Wuhan University, Editorial Department of Literary Criticism, Chinese Wen Yiduo Research Society, Wen Yiduo Foundation, Chinese Society for Modern History. *Proceedings of the 2016 International Academic Symposium on Wen Yiduo*. Southwest University , 2016: 143-155.
- [7] Chen Ailin. Lyrical Self-Construction in Li Bai's Poetry [J]. *Peony*, 2024, (18): 65-67.
- [8] Song Ruolin. Innovative Application and Impact of Digital Humanities Technology in Ancient Chinese Studies [J]. *Chinese Character Culture*, 2025, (11): 51-53.
- [9] Chen Lili. Spatial Narrative Interpretation of Shakespeare's King Lear [J]. *Sichuan Drama*, 2025, (01): 82-86.
- [10] Wang Chunqi. Research on the Application Status and Improvement Strategies of Task-Driven Method in Poetry Teaching in the Middle and Upper Grades of Primary Schools [D] Mudanjiang Normal University, 2025.

Appendix

1. High-frequency words
2. Syllables
3. High-frequency stressed sentences

```
import re
import nltk
from nltk.corpus import stopwords
from collections import Counter

with open("shakespeares-sonnets_TXT_FolgerShakespeare-modified.txt", 'r') as f:
    sonnets = f.read()

words = re.findall(r"[a-zA-Z-']+", sonnets)

stop_words = set(stopwords.words("english"))
```

```

words_lower = [word.lower() for word in words]
words_filtered = [word for word in words_lower if word not in stop_words]

counts = Counter(words_filtered)

print(counts.most_common(50))
# Search all the words in Shakespeare's sonnets directly, remove the function
words using the nltk library, and then perform the statistics.

from collections import Counter
import re

def retain_chinese(str):
    chars = re.findall(r"[\u4e00-\u9fff]+", str)
    chinese_str = ''.join(chars)
    return chinese_str

with open("libai_poems_puretext.txt") as f:
    Youdaoplaceholder0 = f.read().split('.')

sentences = list(set(sentences))

sentences = [retain_chinese(sentence) for sentence in sentences]

allwords = ''.join(sentences)

print(allwords)

word_count = Counter(allwords)
print(word_count.most_common(50))
# Search all the words in Li Bai's poems directly, then perform data statistics
and manually filter function words.

import nltk
from nltk.tokenize import word_tokenize
from nltk import pos_tag
with open("shakespeares-sonnets_TXT_FolgerShakespeare-modified.txt", 'r') as f:

```



```

sentences = f.read().split('.')
print(sentences)
def analyze_sentence_stress(sentence):
    tokens = word_tokenize(sentence)
    tagged = pos_tag(tokens)

    stressed = ""
    stress = 0
    for word, pos in tagged:
        if pos in ['NN', 'VB', 'JJ', 'RB']:
            stressed += word.upper() + ' '
            stress += 1
        else:
            stressed += word.lower() + ' '
    if stress >= len(tokens) * 0.6:
        return stressed
    else:
        return None
    with open("sentence_stresses.txt", "w") as f:
        pass
    with open("sentence_stresses.txt", 'a') as f:
        for s in sentences:
            a = analyze_sentence_stress(s)
            if a:
                f.write(analyze_sentence_stress(s) + '\n')
# Use the nltk library to determine stress in Shakespeare's poem by word part-
of-speech.
from pypinyin import pinyin, Style
def mark_stress_in_poem(poem):
    pinyins = pinyin(poem, style=Style.TONE3, strict=False)
    result = []
    curchar = ""
    curtone = ""
    for char, py in zip(poem, pinyins):
        pinyin_with_tone = py[0]
        tone = 0
        for ch in reversed(pinyin_with_tone):
            if ch.isdigit():

```

```

tone = int(ch)
break

if char == ", " or char == "。 " or char == "！ " or char == "? ":
    result.append(curchar)
    result.append(curtone)
    curchar = ""
    curtone = ""
else:
    curchar += char
    if tone == 3 or tone == 4:
        curtone += "仄"
    elif tone == 1 or tone == 2 or tone == 0:
        curtone += "平"
    else:
        curtone += "无"
print(tone, char)
return '\n'.join(result)
with open("libai_poems_puretext.txt", 'r') as f:
    poem = f.read()
marked_poem = mark_stress_in_poem(poem)
with open("sentence_stresses_libai_pypinyin.txt", 'w') as f:
    f.write(marked_poem)
# The pypinyin library was used to analyze the stress in Li Bai's poems and
determine their tones, with many characters being judged as light tones.

```

```

from snownlp import SnowNLP
import re
with open("shakespeare_sonnet_chinese_raw.txt", 'r') as f:
    poem = f.read()
sentences = re.split(r'[。 ！ ？ ！？， ； ]', poem)
sentences = [s.strip() for s in sentences if s.strip()]
with open("shakespeare-sentiment-analysis-translated.txt", 'w') as f:
    all_sentiment = []
    for sentence in sentences:
        s = SnowNLP(sentence)
        all_sentiment.append(s.sentiments)

```

```

f.write(sentence + ' ' + str(s.sentiments) + '\n')
print(sum(all_sentiment) / len(all_sentiment))
from snownlp import SnowNLP
import re
with open("libai_poems_puretext.txt", 'r') as f:
    poem = f.read()
    sentences = re.split(r'[。!?!?, ; ]', poem)
    sentences = [s.strip() for s in sentences if s.strip()]
    with open("libai-sentiment-analysis.txt", 'w') as f:
        all_sentiment = []
        for sentence in sentences:
            s = SnowNLP(sentence)
            all_sentiment.append(s.sentiments)
            f.write(sentence + ' ' + str(s.sentiments) + '\n')
        print(sum(all_sentiment) / len(all_sentiment))

# SnowNLP was used to obtain the sentiment tendency value of each sentence in Li
Bai's poems and translated Shakespeare's poems, and the average of all the
overall sentiment tendency values was also taken.

# The overall sentiment tendency values of Shakespeare's and Li Bai's poems were
compared with the same standard.

```