

# Chapter 10

## Quantitative Research into Narrative: Statistical Analysis of “The Tale of Genji”

Gen Tsuchiyama  
*Doshisha University, Japan*

### ABSTRACT

*Stylometry is the application of quantitative analysis, primarily to written language, to identify variations in style. Statistical analysis of linguistic characteristics is applied to identify authorship, creation period, and creation order. This study involves stylometric research into “The Tale of Genji,” which is a Japanese classical literary work. “The Tale of Genji” is a long story consisting of 54 volumes. However, in the last 13 volumes, the content is different from that in the other volumes, and the writing quality differs from that of the previous 41 volumes; thus, it has long been theorized that the 13 volumes was written by a different author. The result of an analysis using the word frequency for auxiliary verbs found no evidence that positively support the theory of separate authorship. Therefore, the results indicate that the possibility that there are multiple authors of “The Tale of Genji” is low.*

### INTRODUCTION

This study involves quantitative research into the formation process of “The Tale of Genji,” which is a Japanese classical literary work, using a stylometric method. The stylometry is the application of quantitative analysis, primarily to written language, to identify variations in style. Statistical analysis of linguistic characteristics is applied to identify authorship, creation period, and creation order. Using different approaches, many studies have discussed authorship or the creation periods of historical manuscripts, such as classical literature, by focusing on descriptive content or investigating historical facts. In contrast, stylometry collects quantitative data obtained from sentences, analyzes the data to understand the literary styles, and draws conclusions.

In stylometry, it is assumed that an author’s personality is reflected in his or her writing style and that this relationship can be determined empirically. This assumption is echoed in the French maxim

DOI: 10.4018/978-1-5225-0432-0.ch010

## Quantitative Research into Narrative

“Le style c’est l’homme.” The maxim is said to be derived from “Le style c’est l’homme même,” a line from a speech delivered by Georges-Louis Leclerc, Comte de Buffon (1707-1788), on his admission to the French Academy of Sciences. Leclerc was a naturalist and scientist famous for “Buffon’s needle problem.” In P. Guiraud (1954), Leclerc’s maxim was construed to mean “A style reveals a human’s true self,” and it is understood that readers form impressions of author based on their styles. However, there are various and diverse concepts of style, depending on the field of study.

In stylometry, a style is a countable and descriptive form that considers quantitative elements in sentences, such as the frequencies of letters and words and the average lengths of words and sentences. The personality of an author or a literary work is considered to be reflected in these formal elements. Word meaning and descriptive content are not considered. In this respect, stylometry differs from humanistic literature studies. Stylometry uses statistically analyzed quantitative data to explain differences between the styles of different authors, a specific author’s stylistic changes over time, and other matters.

## BACKGROUND

This study involves quantitative research into the formation process of “The Tale of Genji” using a stylometric method, which involves multivariate analysis. We clarified the quantitative characteristics and trends in the main text using a statistical method. In this study, we present the analysis results with a level of transparency capable of withstanding quantitative debate. Based on these results, the first objective was to quantitatively investigate the different author theory and then to investigate the formation process.

In addition, although, as mentioned previously, there have been quantitative investigations into the research of multivariate analysis, which has not commonly been used to study “The Tale of Genji.” In particular, to the best of our knowledge, no study has investigated author identification using multivariate analysis. Therefore, the originality of this study is that it is based on the quantitative method.

“The Tale of Genji” is a classic work from the Heian period, which has been widely read throughout the ages. The work describes the customs and culture of the aristocratic society of the era. It is one of the oldest known long stories thought to be penned by Murasaki Shikibu—a prominent female author in the Heian period. At that time, the ability to write using Chinese logograms was considered an important cultural accomplishment. However, the tale was probably written entirely in Japanese *kana* (phonograms) because it was written for female readers. In this respect, it is a significant work in the history of Japanese literature.

The main theme in the first part of the story is the romantic life of the main character, Hikaru Genji, who is the son of an ancient Japanese emperor. Hikaru’s mother dies when he is an infant, and he is brought up in the imperial court by his young stepmother. When he grows older, Hikaru and his stepmother fall in love. However, Hikaru recognizes that their love is forbidden and he tries to forget this love affair.

When he visits a rural town, he encounters his “ideal” girl. Although he has already married for political reasons, the girl (Murasaki no Ue), eventually becomes his beloved partner.

His legitimate wife gives birth to a son; unfortunately, the child dies in infancy. Then, because of his scandalous relationship with Murasaki no Ue, the Emperor forces Hikaru to move to a distant land.

In the capital, the Emperor is troubled by bad dreams, and he believes that he is having these bad dreams because he has banished Hikaru Genji. Therefore, the Emperor orders Hikaru Genji to return to the capital and transfers the purple, the color traditionally associated with the Emperor, to the new Emperor—the son born to Hikaru Genji and his stepmother. The new Emperor knows about their love

25 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:  
[www.igi-global.com/chapter/quantitative-research-into-narrative/159629](http://www.igi-global.com/chapter/quantitative-research-into-narrative/159629)

## Related Content

---

### **AmbiLearn: Multimodal Assisted Learning**

Jennifer Hyndman, Tom Lunney and Paul Mc Kevitt (2011). *International Journal of Ambient Computing and Intelligence* (pp. 53-59).

[www.irma-international.org/article/ambilearn-multimodal-assisted-learning/52041](http://www.irma-international.org/article/ambilearn-multimodal-assisted-learning/52041)

### **The Virtual Twin: A Socialization Agent for Peer-to-Peer Networks**

Alexandre Gachet and Pius Haettenschwiler (2005). *International Journal of Intelligent Information Technologies* (pp. 56-67).

[www.irma-international.org/article/virtual-twin-socialization-agent-peer/2384](http://www.irma-international.org/article/virtual-twin-socialization-agent-peer/2384)

### **Agentic AI for Autonomous Vehicle Safety and Reliability**

Qurat Ul Ain Zam Zam, Uswa Ihsan, Javairia Malik and Rimsha Kanwa (2026). *Agentic AI for Autonomous Vehicles: Safety, Reliability, Law, and Ethics* (pp. 119-212).

[www.irma-international.org/chapter/agentic-ai-for-autonomous-vehicle-safety-and-reliability/409663](http://www.irma-international.org/chapter/agentic-ai-for-autonomous-vehicle-safety-and-reliability/409663)

### **The Role of Artificial Intelligence in Business Model Innovation**

Cemerala Sai Charan Reddy and Sanjeev Sharma (2025). *AI-Driven Business Model Innovation* (pp. 163-190).

[www.irma-international.org/chapter/the-role-of-artificial-intelligence-in-business-model-innovation/371361](http://www.irma-international.org/chapter/the-role-of-artificial-intelligence-in-business-model-innovation/371361)

### **Enhanced Security Using Modified Logical Keyless Encryption Algorithm (MLKA)**

P. Abinaya, Senthil Kumar J., Selva Kumar B. and Selva Yogiraam C. (2025). *Sustainable Information Security in the Age of AI and Green Computing* (pp. 369-384).

[www.irma-international.org/chapter/enhanced-security-using-modified-logical-keyless-encryption-algorithm-mlka/380055](http://www.irma-international.org/chapter/enhanced-security-using-modified-logical-keyless-encryption-algorithm-mlka/380055)