Classical Texts, Real or Fake?

Masakatsu Murakami[+]

Quantitive analyses for determining in authorship of two Japanese classical texts are introduced. One treats the question of the authorship of an important religious text called "Sandai Hiho Bonshoji (Teaching on Tripartite Fundamental Secret) 三大秘法禀承事" and the other the authorship of "Genji Monogatari 源氏物語", which is considered the greatest accomplishment in Japanese classical literature.

1. Authorship Problem of "Sandai Hiho Bonshoji"

1.1 The question about authorship

"Sandai Hiho Bonshoji" is a work attributed to Nichiren 日蓮 (1222 〜 1282), the founder of the Nichiren Sect, which is one of the biggest sects of Japanese Buddhism.

The work is relatively short, containing no more than 1280 words.

In spite of its brevity, however, the work is of unusual importance in the study of Nichiren's religious thought. Nichiren's original text is lost and all that remain there are transcriptions. This has often given rise to doubts as to the authorship of the document.

The purpose of our present study is to determine authorship of "Sandai Hiho Bonshoji" by means of quantitive analysis of Nichire's writing, in as much as the question now seems to be beyond the reach of bibliographical research. In addition, we intend to use the same approach to evaluate authorship of four other alleged writings of Nichiren on which doubt exists today, M51, M52, M53 and M54.

[+] Institute of Statistical Mathematics & Graduate University for Advanced Studies

1.2 Procedure for determination of authorship

Authenticity of "Sandai Hiho Bonshoji" (M00 in the subsequent discussion) is studied according to the following procedure. First, a quantitative comparison is made between Nichiren's writing and those writings known to be done by others, in order to identify textual peculiarities (characteristics) of Nichiren's writing. Once Nichiren's textual characteristics have been identified, M00 is subjected to analysis in order to see if it contains such characteristics. If M00 does not show any such characteristics, then it can be considered a fake. If, on the other hand, M00 does show characteristic traits unique to Nichiren's writing, the text is then compared with writings of two disciples of Nichiren, namely Nikko 日興 and Nichijun 日順. If the text shows clear difference from the disciples' writings, then the text, M00, is determined as authentic to Nichiren.

1.3 Analysis

There are two types of data we used for this analysis. The first type is comprised of the to length of sentences, length of words, frequency of parts of speech, size of vocabulary and other data which mainly characterize the structure

of sentences. The second type concerns individual words used to see if there are words Nichiren liked and used often.

a. Analysis using data on sentence structure

Figure 1 shows an example of the result obtained from cluster analysis of data on the structure of sentences. It can be seen from the dendrogram that 50 texts chosen for the purpose of analysis are divided into two clusters, depending on whether they are authentic writing of Nichiren or not, with the exception of six, S01, S04, S07, S11, S12 and S35. These have long been considered authentic writings of Nichiren, but they are not included in that cluster. Thus, the cluster does not comprise the complete authentic writings of Nichiren, if these six texts are genuine. With regard to five texts whose authorship is questionable, M00 and M54 are included in the cluster of genuine texts, while three others, M51, M52 and M54, appear in the other cluster.

We then altered the variables for sentence structure and repeated the analysis. The most important text, M00, was classified in the cluster of authentic texts 35 times out of 35 analyses. As for M54, the text was in the cluster of Nichiren's writing in all cases except one. With regard to the other three texts, M51 and M52 were never included in the cluster of authentic texts, while M53 came in that cluster only three times out of 35.

b. Analysis using data on vocabulary

As we did in our analysis on sentence structure, analysis was repeated using different words. It shows that M00 is in the cluster of Nichiren's work in every analysis.

1.4 Authorship of "Sandai Hiho Bonshoji"

We found that the text M00 always falls in the cluster of Nichiren's own writings according to the cluster analyses done on sentence structure data and vocabulary data.

As to similarity of M00 to texts of Nichiren's disciples (Nikko and Nichijun), Nichijun's D62 was classified as a text of his master in 5 case out of 16 according to the analysis using vocabulary data. On the other hand, all other texts of these disciples were excluded from Nichiren's cluster in the
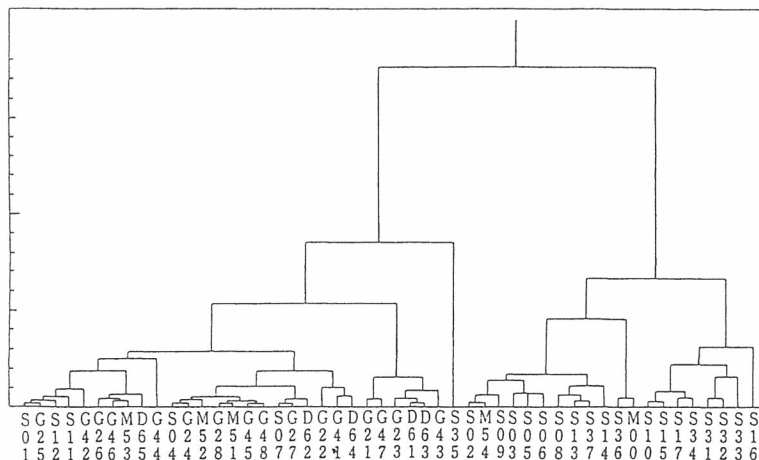

Fig.1 Classificaton according to flexible methods(S···is a work of Nichiren. G···is apocrypha,M···is a work whose authorship is questioned. D···is a work of disciples.)

analyses using sentence structure data as well as vocabulary data. These result show that there is no similarity between M00 and the text written by Nikko and Nichijun.

On the basis of these findings, and within the scope of the quantitative analyses discussed above, we can conclude that M00 is an authentic writing of Nichiren. As for M54, the text shows characteristics similar to those of Nichiren, and therefore it is quite likely to be Nichiren's. On the other hand, chances are quite high that the three other texts, M51, M52 and M53, are apocryphal.

2. Authorship Problem of Genji Monogatari

2.1 The question regarding the last ten volumes of Genji Monogatari

Genji Monogatari 源氏物語, which is rightly considered the greatest accomplishment in Japanese classical literature, is generally thought to have been written toward A.D. 1004, and because of its prestige, the work has been the subject of intensive study for almost ten centuries. Yet, in spite of this effort, there are a great number of problems which remain to be answered. One of the important questions regarding Genji Monogatari is whether Murasaki Shikibu 紫式部 (Lady Murasaki) is the sole author of all its 54 volumes.

As is well known, there have been some influential scholars who hold that the last ten chapters from "Hashihime 橋姫" (the 45th) to "Yumeno-Ukihashi 夢浮橋" (the 54th) were actually written by Dainino-Sanmi 大弐の三位, the daughter of Lady Murasaki. In fact, this view is typical of many others who consider Genji Monogatari to be the product of plural authors.

The author compared the frequency of the main part of speech appearing in the former part of Genji Monogatari (from the 1st to 44th chapters) with those in "Uji-Jujyo" (from 45th to 54th chapters). The results are given in Table 1, (We may think that difference in the frequency is greater as the absolute value of t increase). These show that there is no significant difference in frequency so far as verbs, adjective verbs and adverbs are concerned, but for adjectives and particles there are some differences. As to nouns and auxiliary verbs, the difference becomes clearly meaningful.

Table 2 shows the results of

Table 1. The frequency of main parts of speech

Part of speech	1st to 44th chapter	45th to 54th chapter	t-value
noun	0.18189	0.16423	5.1624
verb	0.16190	0.16505	-1.3581
adjective verb	0.02388	0.02426	-0.2963
particle	0.31517	0.31872	-1.6130
adjective	0.05907	0.05566	1.8597
auxiliary verb	0.11309	0.12391	-5.0477
adverb	0.04000	0.04106	-0.8770

Table 2. The frequency of words

word	1st to 44th chapter	45th to 54th chapter	t-value
hito 人 (noun)	0.01437	0.02000	-2.9273
mono 物 (noun)	0.00429	0.00581	-1.9426
omoi おもい (verb)	0.00251	0.00146	3.0332

comparison of the frequency of certain words often appearing throughout Genji Monogatari. We can see that "hito 人" (man or woman; noun), "mono 物" (matter, thing; noun) and "omoi おもひ" (thought, feeling; declinable conjunctive form of the verb "omou おもふ"), three words which appear often in the work, differ in their frequency in the first 44 chapters and the later Uji-Jujyo.

Figure 2,3,4 and 5 are graphical presentations of the usage of nouns, auxiliary verbs,"hito 人" and"mono 物".

These figures seem to support the view that the author of the last ten chapters of Genji Monogatari is not the same person who wrote the 44 chapters preceding them, but we would be wrong to assert so, inasmuch as it is quite possible that Lady Murasaki's style did change gradually over the long time she took to accomplish the work.
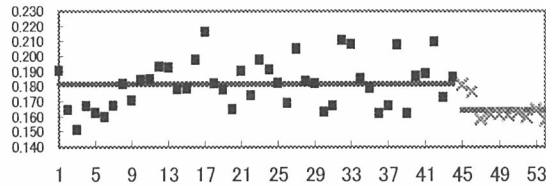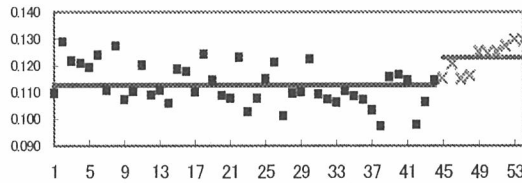
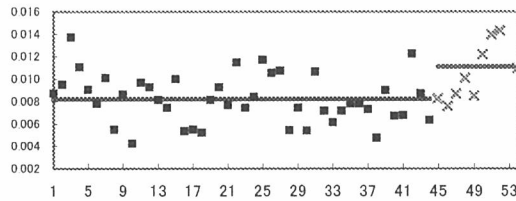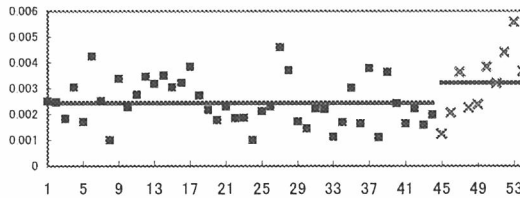Fig.2 The frequency of noons

Fig. 3 The frequency of auxiliary verbs

Fig.4 The frequency of "hito 人"

Fig.5 The frequency of "mono 物"