manuscript cultures

DOI: 10.15460/mc eISSN 2749-1021

Quantitative Orthographic Analysis
Paul S. Atkins and Michael R. Zeng | Seattle

Quantitative Orthographic Analysis of the Ogura *Shikishi*

Paul S. Atkins, and Michael R. Zeng
University of Washington
patkins@uw.edu
zengrf@uw.edu
DOI: 10.15460/s8e93666

Peer-reviewed article
Submitted: 14 August 2025 | Accepted: 8 April 2026 | Published: 2 June 2026

Recommended citation:
Paul S. Atkins and Michael R. Zeng (2026),
'Quantitative Orthographic Analysis of the Ogura *Shikishi*',
manuscript cultures, 26(1): 1–60.

CENTRE FOR THE STUDY OF MANUSCRIPT CULTURES
OPEN ACCESS
Sponsored by: DFG
Hosted by: Staats- und Universitätsbibliothek Hamburg Carl von Ossietzky
Hamburg University Press
Verlag der Staats- und Universitätsbibliothek Hamburg Carl von Ossietzky

Article

Quantitative Orthographic Analysis of the Ogura *Shikishi*

Paul S. Atkins and Michael R. Zeng | Seattle

Abstract

The Ogura *shikishi* 小倉色紙 are a group of squarish cards inscribed with classical Japanese poems in the distinctive calligraphy associated with the later years of the illustrious poet Fujiwara no Teika 藤原定家 (1162–1241). About fifty of these *shikishi* have been recorded and some are extant in prestigious collections. All of the inscribed poems appear in the famous anthology *Hyakunin isshu* 百人一首 (*One Hundred Poems by One Hundred Poets*), whose compilation has been traditionally attributed to Teika. The Ogura *shikishi* were said to have been brushed by Teika himself, an assertion which is generally doubted, but there exists no clear evidence on whether it is true.

We approach the question of authenticity by analyzing patterns in the use of *jibo* 字母, the Chinese characters that are highly cursivized in order to create hiragana, a phonetic syllabary. In premodern Japanese, the same sound could be represented by multiple hiragana, each based on a different *jibo*. As the choice of *jibo* seems largely a matter of personal preference, it holds demonstrated potential for authenticating premodern Japanese handwriting that includes hiragana.

After comparing *jibo* data derived from a group of manuscripts attributed to the hand of Teika with high confidence and comparing it with the data of *jibo* used in the Ogura *shikishi*, we conclude that the probability that Teika brushed any of the Ogura *shikishi* ranges from low to extremely low. That is, all of the Ogura *shikishi* were probably inscribed by persons other than Teika. Finally, we consider the inscriptions that appear on the versos of two of the Ogura *shikishi* and, by the same methods used to evaluate the rectos, conclude that the verso inscriptions were indeed probably brushed by Teika and reused to make the Ogura *shikishi*, perhaps in an attempt to endow them with a veneer of credibility.

Keywords

compositional data analysis, classical Japanese calligraphy, *kuzushiji*, *hentaigana*, *jibo*, scribeship

1. Introduction

The authentication of Medieval Japanese literary manuscripts is an essential and difficult problem. Forgery was not uncommon, and an early modern tradition of authentication by professional calligraphy appraisers (such as the Kohitsu 古筆 family, who were bestowed the surname meaning ‘antique brushes’ in recognition of their services) created an independent body of authentications with which modern scholars must reckon. Authenticity is a problem often encountered in the

history of art, especially when art intersects with commerce in the buying and selling of objects: a work known to be a forgery can fetch but a fragment of the price of the real thing. It is also important in writing literary history, despite the ongoing trend to downplay authorship and, by extension, authenticity. Without knowing who inscribed a particular manuscript, we are often unable to answer other, perhaps even more important, questions, such as when that manuscript was created and why. Since Teika loomed so large over the history of classical Japanese literature during the Late Medieval Era, he was a natural choice for forgers looking to invest their works with the imprimatur of authority, and it is no exaggeration to claim that there are more poetic treatises falsely attributed to Teika than there are authentic ones.[1]

In the case of the Ogura *shikishi*, however, we are not interested in authorship – all of the poems involved were drawn from imperial anthologies, in which the authors' names were clearly indicated. We are interested in scribeship: the question of whether a manuscript was inscribed by a given person, regardless of who composed the text. In light of their tight association with *Hyakunin isshu*, whose compilation by Teika has been contested over the centuries, the Ogura *shikishi* may in turn shed some light on the genesis of *Hyakunin isshu* (*One Hundred Poems by One Hundred Poets*), arguably the most famous anthology of Japanese poetry, and an extraordinarily influential force in the development and reception of waka.

The Ogura *shikishi* are among the most famous constellations of calligraphy in the history of Japanese literature and art (see Figs 1 and 2). *Shikishi* are backed sheets of paper, about twenty centimeters square, sometimes colored or decorated, typically used for inscribing poems. Each of the Ogura *shikishi* bears a single waka from the anthology *Hyakunin isshu*, which is said to have been compiled by the court poet Fujiwara no Teika (1162–1241) in 1235. Every poem is brushed in four lines in Teika's distinctive handwriting, with little use of ligatures between the characters; the names of the poets are omitted. The Ogura *shikishi* are said to have been inscribed by Teika himself and that Teika displayed them on the walls of his villa on Mt. Ogura 小倉, west of Kyoto, hence the name. Therefore, if authentic, the Ogura *shikishi* are the autograph version of *Hyakunin isshu* insofar as they are in the hand of the compiler.

Since their emergence in the late fifteenth century, two centuries after Teika's death, the Ogura *shikishi* have been coveted by connoisseurs of waka poetry and the tea ceremony. They were perhaps the first works of calligraphy to be displayed in tea rooms that were not written in classical Chinese.[2] During the Early Modern Period, wealthy merchants and powerful *daimyō* 大名 added the Ogura *shikishi* to their personal collections, and they traded at very high prices. Today, about forty of the Ogura *shikishi* are extant, some in private collections, others in the collections of some of the most prestigious public and private museums in Japan.

Despite their prominence and prestige, none of the Ogura *shikishi* has been designated by the government of Japan as an Important Cultural Property or National Treasure. Indeed, their authenticity has been questioned for centuries. They lack provenance, the styles of paper vary and even the calligraphy, although undeniably in Teika's style, varies in its similarity to known samples

[1] For a survey in English of Teika's writings on poetics, see Atkins 2017, 88–123.

[2] Komatsu 1970, vol. 1, 570–571.

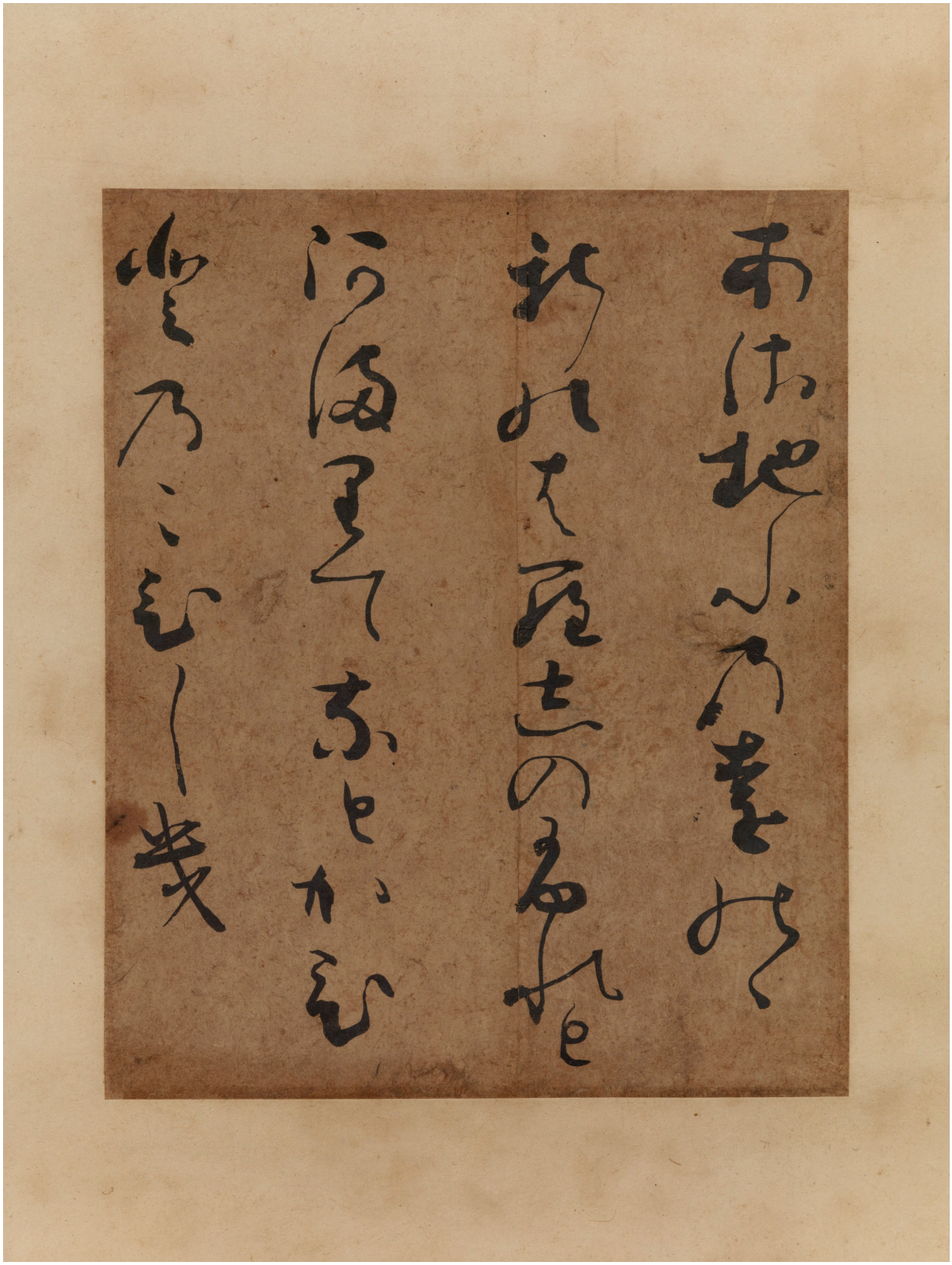

Fig. 1: Ogura *shikishi* 'Asajiu no', attributed to the hand of Fujiwara no Teika; Dazaifu, Kyushu National Museum, item no. B46.

DOI: 10.15460/s8e93666

of his handwriting. While doubts have been voiced by many on the basis of connoisseurship, there has been no systematic effort to date to authenticate or discredit the Ogura *shikishi* using objective methods.

In this paper we offer a statistical analysis of the orthography used in the Ogura *shikishi* in order to estimate the probability that they were indeed inscribed by Teika. Because the Ogura *shikishi* are clearly not a matched set, we evaluate each *shikishi* individually. Specifically, we analyze only the hiragana phonetic characters that appear in the Ogura *shikishi*, ignoring *kanji* 漢字 ('Chinese characters'). Hiragana are highly cursivized versions of Chinese characters used in Japan. The cursive form of hiragana signals to the reader by convention that they are to be treated solely as representing sounds, not meanings; they are stripped of semantic value. In contemporary Japan, each sound is represented by one and only one hiragana, but this was not the case in premodern Japan. Scribes could and did use multiple hiragana, each derived from a different Chinese character, to represent a single sound. For example, today the sound A is represented by the hiragana あ, which is derived from the *kanji* 安, but in premodern Japan, scribes also used other hiragana, derived from the *kanji* 阿, 愛, or 悪. These now-obsolete *kana* are referred today by the unfortunate term *hentaigana* 変体仮名 ('variant kana') but in the premodern era there was no such concept. The characters that are simplified to form hiragana are called *jibo* 字母 ('matrix characters').

1.1 Provenance of the Ogura shikishi

Where did the Ogura *shikishi* come from? If they are authentically in the hand of Teika, then they probably originated in 1235 when, as Teika records in his diary *Meigetsuki* 明月記, the retired general Utsunomiya Yoritsuna 宇都宮頼綱 (1172–1259) asked him to inscribe some famous poems that Yoritsuna would use to decorate his villa, which was located next to Teika's in the Saga 嵯峨 area west of Kyoto. Although Teika was ashamed of his handwriting, which he found unsightly, he found it difficult to refuse Yoritsuna's request – not only was his neighbor sincerely devoted to the way of poetry, he was wealthy, powerful, and happened to be the father of the wife of Tameie 為家, Teika's son. Here is the entry from *Meigetsuki*:

> 予本自不知書文字事、嵯峨中院障子色紙形、故予可書由、彼入道懇切、雖極見苦事、慙染筆送之、古来人歌各一首、自天智天皇以来、及家隆・雅經公卿。
>
> I have never known how to write characters. The lay monk has taken the trouble to press me for square slips of colored paper [*shikishi-gata*] for the sliding room dividers at Chūin in Saga. Although they are extremely unsightly, against my better judgment I wet my brush, and sent them. One poem by each person, from olden times until now, from Emperor Tenji all the way to Lords Ietaka and Masatsune.[3]

[3] Entry of Bunryaku 2.05.27 (1235). Reizei-ke Shiguretei Bunko 2018, 528. The original is not extant, which is not atypical for the later years of *Meigetsuki*, and the transcription is based on a copy made by Reizei Tamehisa 冷泉為久 (1686–1741). This and all other translations are by Paul S. Atkins unless indicated otherwise.

Proponents of the idea that Teika compiled the text known as *Hyakunin isshu* today cite this passage as evidence in support of their thesis.[4] Indeed, *Hyakunin isshu* does begin with a poem attributed to Emperor Tenji 天智 (626–671) and includes toward the end poems by Fujiwara no Ietaka 藤原家隆 (1158–1237) and Asukai Masatsune 飛鳥井雅経 (1170–1221). However, this passage does not say how many poems Teika inscribed and the oldest known copy of *Hyakunin isshu* dates from 1445, more than 200 years after Teika's death in 1241.[5] This is a yawning gap which many have attempted to fill, to little avail. We should note in passing that Teika did not inscribe *shikishi*, but *shikishi-gata* 色紙形. *Shikishi* are stiff cards that can be handled individually; *shikishi-gata* are slips of paper of the same shape and size as a *shikishi*, intended to be mounted on a standing screen, a sliding door, or some other surface to provide suitable backing.[6]

A generation or so after Gyōkō 尭孝 made the earliest extant copy of *Hyakunin isshu* in 1445, there seems to have occurred some sort of *Hyakunin isshu* 'boom'. In 1479, the renga poet Sōgi 宗祇 mentioned *Hyakunin isshu* in a list of essential reading for poets that he included in a treatise.[7] A similar anthology of one hundred Chinese poems by Japanese authors bears the same title and was compiled sometime between 1480 and 1493.[8] Moreover, *Shin Hyakunin isshu* 新百人一首, a collection with the same structure as *Hyakunin isshu* but with different poets and poems, was compiled by the shogun Ashikaga Yoshihisa 足利義尚 (1465–1489) in 1483.[9] Perhaps most importantly, Sōgi produced in 1478 the earliest extant commentary on *Hyakunin isshu*. He says that it is based on lectures given to him in 1471 by the warrior and poet Tō Tsuneyori 東常縁 (1401–1494). Sōgi's origin story for *Hyakunin isshu* deviates somewhat from the information gleaned from Teika's brief diary entry. He writes:

> 右百首は京極黄門小倉山庄障子色紙和歌也。其を世に百人一首と号する也。これをえらびかきおかるる事は新古今集の撰定家卿の心にかなはず。
>
> The preceding sequence of one hundred waka are the waka of the *shikishi* on the sliding screens of the villa on Mount Ogura of the Kyōgoku Middle Counselor. It is popularly called *Hyakunin isshu*. With regard to it being compiled and written down, the selection of poems included in the *Shin Kokinshū* was not to Lord Teika's liking.[10]

---

[4] An alternate view proposed by Ariyoshi Tamotsu 有吉保 holds that the collection Teika selected and brushed for Yoritsuna in 1235 was that known today as *Hyakunin shūka* 百人秀歌, comprising 101 poems, of which all but a handful are included in *Hyakunin isshu*, and that the text now known as *Hyakunin isshu* was selected by Teika after the death of Cloistered Emperor Go-Toba 後鳥羽 in 1239 and before Teika's own death in 1241; it is this later text that survives today as the Ogura *shikishi*. Cited in Mostow 1996, 25. Joshua Mostow summarizes the current thinking on the origins of *Hyakunin isshu* (which necessarily intersect with the origins of the Ogura *shikishi*) in Mostow 2024, 2–4.

[5] The manuscript is in the hand of Gyōkō 尭孝 (1391–1455). See Eiga no taigai 詠歌大概, Tokyo, Kunaichō Shoryōbu Toshoryō Bunko, 503: 236, <https://shoryobu.kunaicho.go.jp/Toshoryo/Viewer/1000653280000/> (accessed on 13 August 2025). *Hyakunin isshu* is bound alongside other texts in the same volume, including Teika's treatise *Eiga no taigai*, and begins with image 19 of 44.

[6] See Sugiura 2002.

[7] *Oi no susami* 老のすさみ. Cited in Akisada 2002, 117.

[8] Ōsen Keisan 黄川景三, compiler, *Hyakunin isshu* 百人一首. See Hibino 1986.

[9] Takei 2013.

[10] Tokyo, Kunaichō Shoryōbu Toshoryō Bunko, Takatsukasa 138, image 4 of the digitised version available at <https://shoryobu.kunaicho.go.jp/Toshoryo/Viewer/1000419380000/> (accessed on 13 August 2025). This manuscript, titled

Teika had served as a co-compiler of the *Shin Kokinshū*, the eighth imperial anthology, which was closely supervised by its commissioner, Cloistered Emperor Go-Toba. Although Teika and Go-Toba famously clashed over the compilation, among other matters, there is no other evidence to suggest that Teika ever repudiated the *Shin Kokinshū* or its style. Sōgi then explains that Teika believed the poems selected for the *Shin Kokinshū* sacrificed substance for showy effect, but could not express that view (presumably for fear of offending Go-Toba). He continues:

> されば黄門の心あらはれがたき事を口惜く思ひ給ふ故に、古今百人の哥をえらび我山庄に書おかるる者也。
>
> Therefore, the Middle Counselor [Teika] felt that it would be deeply regrettable if he could not reveal his feelings, and so he chose the poems of one hundred persons of ancient and recent times and wrote them out at his mountain villa.[11]

This harkens back to the opening passage, which refers to *Hyakunin isshu* as the 'Ogura mountain villa *shikishi* waka'. That is, in this telling, Teika did not select and inscribe the poems to decorate Yoritsuna's villa at the latter's request, but rather his own villa, to vent his frustration over the *Shin Kokinshū* and express his true ideas about poetic style. The improbability of this scenario is apparent, given that Teika noted in his diary how much he disliked his own calligraphy; it is unlikely he would have decorated his own villa with it. Sōgi states that the style of *Hyakunin isshu* may also be seen in the ninth imperial anthology, *Shin chokusen wakashū* 新勅撰和歌集, compiled by Teika alone in 1235 (just before the diary entry cited above). Moreover, this style shared by *Hyakunin isshu* and the *Shin chokusen shū* is identical to that professed by the Nijō 二条 school, the group of poets associated with a branch of Teika's descendants that had died out, which Sōgi eventually led. The Nijō are regarded as the most artistically conservative of Teika's descendants, preferring a relatively plain, forthright, and declamatory style, sometimes called the *shōfūtei* 正風体 ('the correct style'), *ushintei* 有心体 ('the sincere style'), or *uruwashiki tei* 麗しき体 ('the splendid style'). *Hyakunin isshu* was mobilized by Sōgi to justify and legitimate the Nijō school and its style; the Ogura *shikishi*, in turn, help to authenticate *Hyakunin isshu*.

A few years after Sōgi's commentary we encounter the oldest extant records of the Ogura *shikishi* themselves. The courtier Sanjōnishi Sanetaka 三条西実隆 (1455–1537) was an accomplished calligrapher and avid poet who, improbably, took a man from nowhere, Sōgi, as his teacher.[12] Like Teika, Sanetaka also kept a diary, which has proven to be a treasure trove of information about his life and times.

Hyakunin isshu shō 百人一首抄, was copied in the year An'ei 5 (1776) and putatively reproduces a manuscript dated to the year Bunmei 10 (1478).

[11] Tokyo, Kunaichō Shoryōbu Toshoryō Bunko, Takatsukasa 138, image 4 of the digitised version available at <https://shoryobu.kunaicho.go.jp/Toshoryo/Viewer/1000419380000/> (accessed on 13 August 2025).

[12] For a biography of Sōgi see Okuda 1998. Sōgi's surname is often given as Iio or Inoo 飯尾, but there is no firm evidence that he belonged to that or any other family; aside from him having been a Zen monk at Shōkokuji 相国寺 before he emerged in poetic circles in Kyoto, we know almost nothing about his early life, which is telling, given the extent of the prominence and fame he achieved in his own lifetime.

In 1490, Sōgi visited Sanetaka to ask a favor.

> 宗祇法師来話、京極黄門真筆色紙形（陽成院水無能河哥也）正真之由、予為証明可筆之由所望、更雖不可有信用可染筆之由領状了。
> The monk Sōgi came to chat. He requested I write something to confirm what he said was an authentic *shikishi*-sized slip in the hand of the Kyōgoku Middle Counselor [Teika] (Waka 'Minanogawa' by Cloistered Emperor Yōzei). It is not to be trusted at all, but I agreed to write it anyway.[13]

This is an apparent reference to *Hyakunin isshu* no. 13:

> つくばねのみねよりおつるみなのがは
> こひぞつもりてふちとなりぬる
> *Tsukubane no / mine yori otsuru / Minanogawa /*
> *koi zo tsumorite / fuchi to narinuru*[14]
> Falling from the peak of Mount Tsukuba, the Minano River:
> my longing has accumulated and become a pool.

Remarkably, the first clear mention of an Ogura *shikishi* reveals it as a forgery. It is possible that the *shikishi* was authentic and Sanetaka wrongly judged it as fake, but hardly likely. Although he does not say so, it is more likely that both Sōgi and Sanetaka knew the *shikishi* was a fake, whether they said so explicitly or not, and Sōgi rewarded Sanetaka for his work. In fact, in another entry, Sanetaka reports having received a *shikishi* by Teika from Sōgi as a gift, which he treasures.

During the sixteenth century, the Ogura *shikishi* surface again in the pages of records of tea gatherings (*chakaiki* 茶会記), records of who served what kind of tea with what kinds of furnishings to whom, where, and when. The tea master Takeno Jōō 武野紹鷗 (1502–1555) is believed to have been the first host to decorate the alcove (*tokonoma* 床の間) with one of the Ogura *shikishi*; up until that time, displays of calligraphy at tea gatherings had been limited to writings in classical Chinese, such as inscriptions by Zen monks. Eventually the Ogura *shikishi* found their way into the hands of various *daimyō*, including the collections of Akechi Mitsuhide 明智光秀 (1528–1582), Toyotomi Hideyoshi 豊臣秀吉 (1537–1598), Hosokawa Yūsai 細川幽斎 (1534–1610), Tokugawa Ieyasu 徳川家康 (1543–1616), Date Masamune 伊達政宗 (1567–1636), Kobori Enshū 小堀遠州 (1579–1647), and Matsudaira Fumai 松平不昧 (1751–1818), and were owned and displayed by cultural elites including the monk, painter, and calligrapher Shōkadō Shōjō 松花堂昭乗 (1584–1639).[15]

---

[13] *Sanetaka-kō ki*, entry of Entoku 2.11.29 (1490). Sanjōnishi 1931–44, vol. 2, 505–506.

[14] *Gosen wakashū* 後撰和歌集 no. 776 by Cloistered Emperor Yōzei 陽成院.

[15] Komatsu 1970, vol. 1, 577–583.

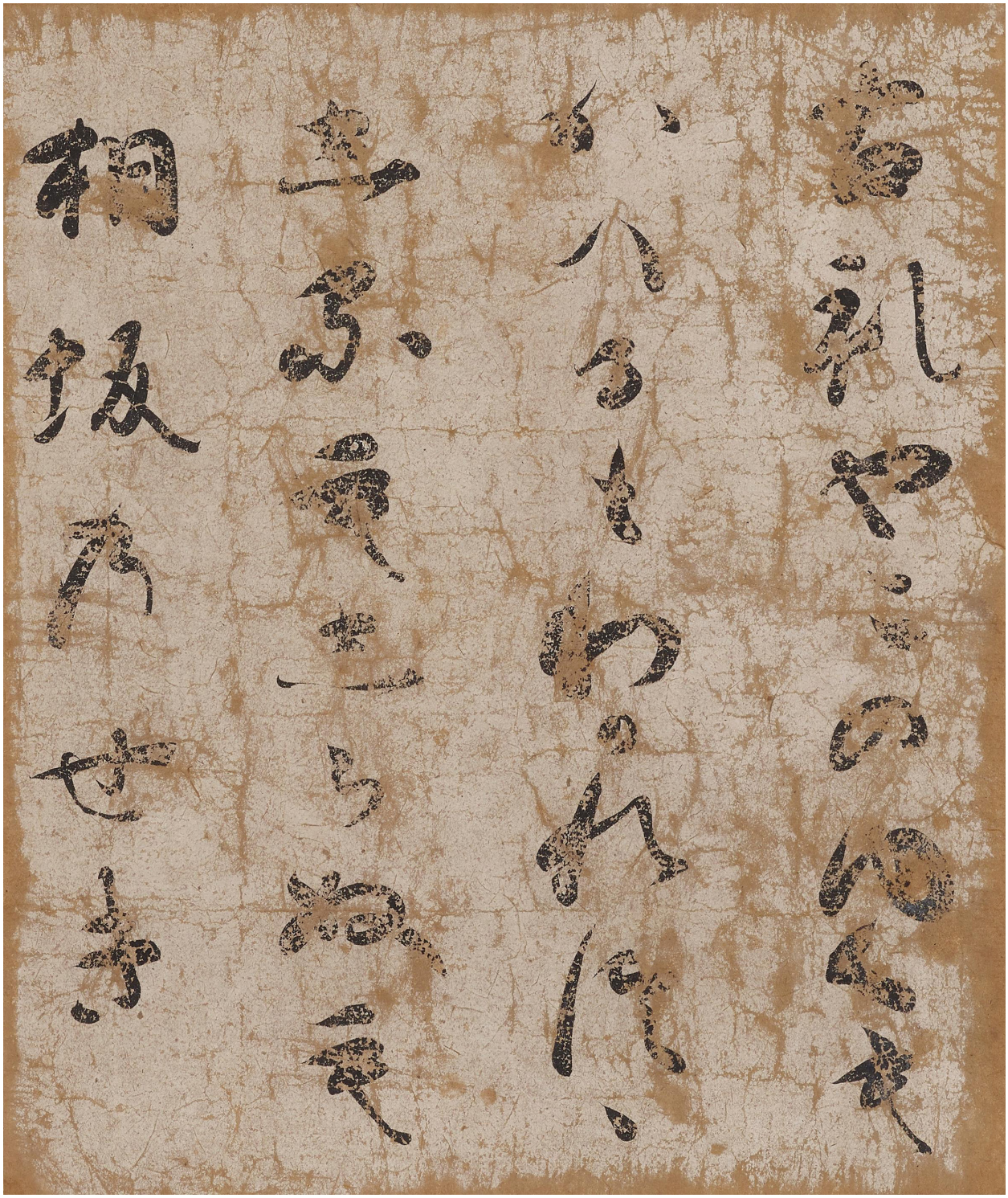

Fig. 2: Ogura *shikishi* ' Kore ya kono', attributed to the hand of Fujiwara no Teika; Tokyo, Tokyo National Museum, item no. B-3072.

1.2 Previous scholarship on the Ogura shikishi

Little more can be said about the putative origins of Ogura *shikishi*, so now let us turn to a review of the previous scholarship about them, with a focus on discussions of their authenticity. Most of the review will be devoted to modern works, but we should take a moment to address a remarkable work from the premodern era, a compendium of images of antiques in Japanese collections titled *Shūko jisshu* 集古十種 (*Ten Types of Antiques*), which we will return to later briefly in our description of the data. This multi-volume woodblock printed work was compiled by Matsudaira Sadanobu 松平定信 (1759–1829), a high-ranking official of the Tokugawa shogunate, and included illustrations by the noted artist Tani Bunchō 谷 文晁 (1763–1841). It devoted volumes to such categories as arms and armor, horse tack, musical instruments, writing implements, and paintings. Curiously, the Ogura *shikishi* merited their own booklet, a set of labeled illustrations with all of the extant Ogura *shikishi* that Matsudaira could arrange to view and Tani could copy. Comparison of Tani's copies with some of the extant Ogura *shikishi* show that he did a faithful job. Of note is a remark that Matsudaira appended at the end of the booklet:

> この色紙はその家々の重宝にしてながく伝えしことのなればもとより優劣の論にも及ばず、まいて憶見の取捨すべき事にもあらず。ただ温故のたよりになさむとてうるにしたがいて収入するのみ。
>
> These *shikishi* are the precious treasures of various families, which have transmitted them for a long time. As such, whether one is superior or inferior to another is beyond discussion and, what is more, it is not appropriate to engage in groundless speculation. I merely included them as I obtained them, with the intention of making them a means for cherishing the past.[16]

Among the illustrations included are two of the same poem, making it highly likely that at least one of them was a fake. Matsudaira was well aware that the quality of the calligraphy – to be specific, the degree to which it matched known samples of Teika's handwriting – varied greatly but, since he included the owners of the objects among his published notes, he took pains to discourage others from openly disparaging the authenticity of the objects.

Moving ahead to the Modern Era, let us review the range of opinions about the authenticity of Ogura *shikishi* that have been expressed over the last half century or so. The historian of Japanese calligraphy Haruna Yoshishige stated in 1969 that the high demand for Ogura *shikishi* and scarce supply had led to the creation of forgeries, and specifically mentioned the *shikishi* 'Konu hito wo', *Hyakunin isshu* no. 97, reproduced in *Shūko jisshu* as one of them, based on its clumsy calligraphic style. Furthermore, he claimed that 'Arashi fuku', *Hyakunin isshu* no. 69, was a forgery for similar reasons, and said that there were not a few forgeries or copies among the lot.[17] In 1972, Komatsu Shigemi also expressed general doubts about the authenticity of the Ogura *shikishi*, especially

---

[16] Colophon of *Shūko jisshu* volume titled *Hōjō: Teika-kyō shinseki Ogura shikishi* 法帖：定家卿真蹟小倉色紙 'Album of calligraphic exemplars: Ogura *shikishi* in the authentic hand of Lord Teika', twentieth-century reprint in the collection of the National Diet Library (Japan), item no. W991-H26, <https://dl.ndl.go.jp/pid/1881097/1/1> (accessed on 31 December 2025). Transcription by Paul S. Atkins, with punctuation and diacritics added.

[17] Haruna 1969.

those reproduced in *Shūko jisshu*.[18] Writing in 1987, in a catalog for an exceptional exhibition of Teika's calligraphy and works inspired by his style, Nagoya Akira said that there was no conclusion regarding the authenticity of the Ogura *shikishi*.[19] The same year, Sugitani Jurō acknowledged that many of the Ogura *shikishi* were mere copies or outright forgeries, but expressed a more favorable view of the *shikishi* 'Koi su te fu', *Hyakunin isshu* no. 41, with its verso inscription that was clearly in Teika's hand.[20] Writing in 1990, Ii Haruki discussed the two *shikishi* with verso text ('Koi su te fu', no. 41 and 'Sabishisa ni', no. 70) as well as the *shikishi* 'Ai mite no', no. 43. He certainly seemed to treat all three as authentic, although he did not say so explicitly.[21] Returning to the work of the curator Nagoya Akira, who is probably the leading expert on the Ogura *shikishi*, his 1994 essay is the most complete study of this topic, considering the Ogura *shikishi* in the broader context of Teika's calligraphy and examining documentary evidence, calligraphic style (especially in comparison to that seen in Teika's copy of the *Tosa nikki* 土佐日記, made the same year as the *shikishi* he inscribed for Yoritsuna, in 1235), and the type and quality of paper used. He dismissed quite a few, but singled out as potentially authentic these five: 'Tachiwakare', no. 16; 'Shinoburedo', no. 40; 'Koi su te fu', no. 41; 'Aimite no', no. 43; and 'Sabishisa ni', no. 70, because of their calligraphic similarity to Teika's copy of *Tosa nikki* and the similarity of the paper on which they were inscribed.[22] Writing in 1999, Yoshida Kōchi dismissed 'Momoshiki ya', no. 100, as inauthentic, but quoted a study by the scholar Katagiri Yōichi at length, in which Katagiri had carefully compared the calligraphic style of Teika's copy of *Tosa nikki* (*The Tosa Diary*, by Ki no Tsurayuki 紀貫之, 935 CE) to 'Aimite no', no. 43, and 'Hito mo oshi', no. 99, and pronounced them authentic.[23] To his credit, Katagiri noted that some of the *jibo* used in the *shikishi* he studied never appear in *Tosa nikki*, but unfortunately he explained them away by claiming that Teika altered his customary usage to suit the *shikishi* format. Yoshida then applied the same method to 'Ukarikeru', no. 40, and identified several disparities between its calligraphy and orthography and that of *Tosa nikki*, but dismissed them as well.[24] Katagiri subsequently revised his view of the Ogura *shikishi*, after having acquired extensive hands-on experience with materials in Teika's hand, and wrote in 2003 that he believed that all of the extant Ogura *shikishi* were forgeries made in a later age.[25]

Although it is not a scholarly work, it would be remiss to omit mention of the marvelous murder mystery *Hyakumai no Teika* 百枚の定家 by the historical novelist Azusazawa Kaname (the pen name of Nagata Michiko 永田道子), published in 1998. In it, a young curator at an art museum

[18] Komatsu 1972, 572.

[19] Nagoya 1987, 94.

[20] Sugitani 1987, 2. Sugitani wrote that he had had the opportunity to view the *shikishi* at the Tokugawa Museum when it was being remounted and had viewed the verso inscription himself; he stated that he felt no one but Teika would cut up his writing and reuse it for another purpose.

[21] Ii 1990.

[22] Nagoya 1994, 288.

[23] Yoshida 1999, 344–352.

[24] Yoshida 1999, 352–354.

[25] Katagiri 2003, 178–179.

near Tokyo is presented with the daunting task of attempting to authenticate a large cache of the Ogura *shikishi*, never before seen, that has been offered for sale to his museum. It is a thrilling and informational read, and includes a character who was unmistakably modeled on the scholar and curator Komatsu Shigemi.[26]

In 2013, Nagoya revisited the topic in a published dialogue with the waka scholar Kanechiku Nobuyuki, and said that he thought that most of the extant Ogura *shikishi* had been inscribed after the time of Teika.[27] He added that he thought that 'Ai mite no', no. 43, and 'Koi su te fu', no. 41, had been inscribed by the same person, possibly during Teika's lifetime.[28] The most recent significant work on the Ogura *shikishi* is an article published in 2020 by Tabuchi Kumiko, which focuses again on the *shikishi* that include verso text. Tabuchi noted the current consensus that a majority of the Ogura *shikishi* were inscribed during the Muromachi or Edo periods, and acknowledged the view that all were forgeries, but cited Nagoya's previous work, summarized above, in which he suggested that five of them were authentic. She then turned to 'Koi su te fu', no. 41, whose verso text appears to be a handful of poems by poets around Teika's time, and gives it extended consideration, concluding by concurring with Nagoya's affirmation of the five *shikishi*, speculating that Teika inscribed them as drafts, which would account for the reuse of old paper.[29] In closing, she affirmed them as authentic, justifying the rather atypical use of the *shikishi* format by Teika by explaining that he did so at the request of Yoritsuna. In a note, Tabuchi acknowledged the possibility that someone besides Teika obtained scraps of his handwriting, and faked his handwriting on the reverse sides, now the front of two or more Ogura *shikishi*, but dismissed it because she doubted whether such a person would choose writing with the work of relatively minor poets.[30] This unusual objection, however, presupposes that the scribe had a range of samples of Teika's handwriting to choose from, which seems unlikely, and further estimates that a would-be forger would sacrifice a work by Teika that was more valuable instead of a lesser one.

Such is the state of current scholarship on Ogura *shikishi*: some are dismissed outright as clumsy fakes, others are affirmed as authentic based largely on calligraphic style and the presence of verso text, and the rest are in an unknown category. The question is ripe for a definitive, objective treatment.

Before proceeding to a discussion of our data, methods, and analysis, it is worth reviewing briefly also the previous scholarship on *jibo* analysis, with a focus on Teika's usage. The curator and scholar Komatsu Shigemi, who founded the field of *kohitsugaku* 古筆学 ('study of pre-Muromachi Japanese manuscripts') in the postwar era, was the first to publish the discovery that scribes tended to have distinctive individual preferences in their use of hiragana and, by extension, *jibo*.[31] *Jibo*

[26] Azusazawa 1998.

[27] Nagoya and Kanechiku 2013, 143.

[28] Nagoya and Kanechiku 2013, 139.

[29] Tabuchi 2020, 46.

[30] Tabuchi 2020, 57 n. 14.

[31] Komatsu 1961. Komatsu's approach resembles that taken twenty years earlier by Ikeda Kikan in his philological study of *Tosa nikki*. See Ikeda 1941 and Kornicki 2008.

could be counted in a text containing hiragana and, by calculating the percentages or relative frequencies of *jibo* used to represent each sound, *jibo* choice could be compared across individual texts. For example, to express the sound A, one scribe might use the *jibo* A 安 half of the time, A 阿 the other half, and other versions never. Another might conceivably use A 安, A 愛 and A 悪 in equal proportions. Over a certain number of occurrences of the sound A recorded in a text of uncertain authorship, the pattern observed might be closer to one scribe than the other.

Komatsu first validated his approach by conducting a simple experiment with living calligraphers of his acquaintance, by asking them to copy out the same text with the *jibo* variants of their choosing, and comparing the results. He found sufficient variation among individuals and by gender and age to conclude that quantification of *jibo* frequency was a valid technique for assessing classical texts. Komatsu then applied his technique to a variety of premodern poetic texts attributed to different scribes, with the goal of determining whether a particular copy of *Gosen wakashū* 後撰和歌集, the second imperial anthology of waka, compiled in 951, was indeed in the hand of Asukai Masatsune 飛鳥井雅経 (1170–1221) as previously believed or, as Komatsu suspected, was by someone else. With painstaking work in counting and tabulation (the study was published in 1961), Komatsu presented a set of tables demonstrating that his hunch was correct. The percentages of the text in question matched more closely those of Fujiwara no Norinaga 藤原教長 (1109–1180).

Nonetheless, Komatsu did not apply any sort of statistical analysis to his figures, other than calculating the percentages of each *jibo* used to represent each sound. A decade later, Iura Miyuki applied a statistical analysis to relative *jibo* frequencies (in which the number of appearances of a *jibo* is divided by the total number of appearances of all *jibo* representing the same sound; a number between .00 and 1.00 is generated) for twenty-one manuscripts, most of which were attributed to the hand of Teika. She was able to divide the manuscripts into four clusters based on similarity in *jibo* usage: Teika; his scribes, including his children; his son Tameie and his wife Abutsu 阿仏; and a fourth group whose usage was very unlike Teika's.[32] Ue Kiyoko studied six manuscripts, five of which were attributed to Teika, with particular attention to the relationship between *jibo* usage and its position on the page (at the beginning or end of a line, for example).[33] Toyoda Naoko returned to the problem of *jibo* usage in Teika's manuscripts and line position. Although previous research had claimed that Teika deliberately avoided using the same *jibo* at the head of consecutive lines (presumably to prevent accidentally skipping or swapping lines during the copying process), Toyoda found that this practice depended on the nature of the text, namely whether it possessed a mainly practical or mainly aesthetic value.[34]

More recently, Teika's *jibo* usage has been studied by Saitō Tetsuya, who has published extensively on statistical analysis of *jibo* usage in classical Japanese manuscripts, and is leading the subfield. He used hierarchical cluster analysis (HCA) and principal cluster analysis (PCA) to sort manuscripts attributed to Teika by the year of copying and was able to identify a cluster around

[32] Iura 1971.

[33] Ue 1979.

[34] Toyoda 1992.

the Tenpuku era (1233–1234) as well as point out significant *jibo*.[35] In another study, he used the same statistical techniques to differentiate Teika's writing from that of others, with some success, and to further explore differentiating Teika's manuscripts by date.[36] Sakamoto Kiyoe has analyzed the *kana* usage, including *jibo*, of heads of the Mikohidari 御子左 family, from Shunzei 俊成 to Tamesuke 為相, including Teika, and found differences among them. She noted that Teika's *kana* usage changed after he copied *Tosa nikki* in 1235.[37]

The literature review above shows that there is a range of opinions about the authenticity of the Ogura *shikishi*, and there is a fair amount of research about Teika's use of *jibo*, although its use of statistical methods is limited and it does not examine manuscripts at the level of the individual line or poem. The present study attempts to address the question of the authenticity of the Ogura *shikishi* individually, not as a whole, and, in order to do so, we have had to adapt previous methods and propose new ones, as discussed below.

1.3 Methodological contributions of the present study: a non-technical explanation

The present study makes contributions to the previous scholarship on two levels. First, it contributes to the previous scholarship on the Ogura *shikishi* and, by extension, Fujiwara no Teika, *Hyakunin isshu*, medieval Japanese waka, and the history of Japanese calligraphy by establishing, via objective quantitative methods, that the probability that Teika brushed any of the known Ogura *shikishi* is very low.

Second, we make important methodological innovations in the statistical analysis of *jibo* that should enable and facilitate future studies. These new methods are explained in detail later in the technical discussion, but a concise, non-technical overview is provided here for the lay reader.

The first innovation is the recognition that *jibo* must be analyzed using a special way of handling statistics called compositional data analysis (CoDa). Basic *jibo* analysis requires counting individual *jibo* in a text, summing those counts by sound, and then dividing the individual *jibo* counts by the summed sound counts in order to derive percentages for each *jibo*. These percentages must always sum to 100% or 1.0. If the percentage of one *jibo* increases, the other *jibo* representing that sound must necessarily decrease. Data that exhibits this type of relationship is called compositional data and appears in a range of scientific fields, including geology, in which the various elements comprising a rock must always add up to 100%. Compositional data cannot be handled as is using standard statistical methods. It must first be transformed using a common technique called log-ratio transformation, in which numbers are converted to exponential form and only the powers to which the original numbers are raised are used. Only then may standard statistical methods be employed. None of the previous scholarship on *jibo*, which dates back to 1961 or earlier, has recognized the necessity of this approach, which should be used in all future statistical analyses of *jibo*.

[35] Saitō 2016.

[36] Saitō 2018.

[37] Sakamoto 2020.

As a second methodological innovation, we introduce a method of evaluating very short inscriptions, called the probability rating (PR). The previous scholarship on *jibo* has conducted statistical analyses of *jibo* in comparing relatively long manuscripts. But it is readily apparent to scholars of Japanese calligraphy and literature that many important specimens of premodern writing exist in brief fragments, as short as the thirty-one or fewer characters used to inscribe a waka poem, as we see in the case of the Ogura *shikishi*. Given a longer manuscript that can be used as a reference, the PR can be used to calculate the consistency of the brief inscription with the statistical trends in *jibo* usage exhibited by the reference manuscript and the brief inscription combined. The previous scholarship does not include anything like the PR, which not only multiplies the relative frequencies but normalizes the result for length of inscription, so that inscriptions of varying lengths may be compared. The PR has potential as a very powerful tool to evaluate fragments that appear in calligraphic albums or mounted on scrolls. It can also be used to compare longer manuscripts to one another or to evaluate a single manuscript for internal consistency to detect a change in scribes.

Thus, the takeaway from this article should not be only that Teika probably did not brush any of the Ogura *shikishi*: that has been suspected for decades, if not centuries, and we have merely settled the matter by confirming some of the previous scholarship. Perhaps the greater epiphany is the immense potential that statistical analysis of *jibo* has for reevaluating the vast corpus of premodern Japanese manuscripts inscribed in hiragana, using traditional methods of compositional data analysis and our newly proposed probability rating.

2. Data acquisition and preparation

At last, we should like to describe how we acquired the data used for this study and prepared it for analysis. We have three data sets: one comprises the *jibo* for all known Ogura *shikishi*, including both the extant exemplars as well as those that survive only in form of transcriptions; one comprises *jibo* from a set of seven manuscripts that are believed with high confidence to have been inscribed by Teika; and one comprises *jibo* from two manuscripts that were inscribed by persons other than Teika (see Fig. 3 for an example of Teika's handwriting that is not in the dataset).

Regarding the Ogura *shikishi*, the *jibo* data were generated relying on published images in print and online. Because our approach to authentication is not based on material analysis, examination of the actual items was superfluous; clear images sufficed. Indeed, because our approach relies on orthographical features, not calligraphic style, we did not even require photographs or digital images; any copy, even one done by hand, that faithfully rendered the shape of the characters such that the underlying *jibo* was detectable was sufficient. In some cases, *jibo* data were extracted from images no larger than one inch square. In fact, *jibo* data were produced for Ogura *shikishi* whose whereabouts are currently unknown and may no longer be extant. Thirty-three instances of the Ogura *shikishi* appear in *Shūko jisshu*, discussed above, all of which were considered. It is apparent that these are faithful copies because some of the originals are still extant, and match the copies perfectly. The *jibo* data for the Ogura *shikishi* are given in Appendix A.

As for the reference manuscripts, those regarded as having been inscribed by Teika, we used the following seven:

(i) *Kokin wakashū* 古今和歌集 (*Date-bon* 伊達本, manuscript formerly owned by the Date family) (=KODA)

(ii) *Kokin wakashū* 古今和歌集 (*Karoku-bon* 嘉禄本, manuscript copied during the Karoku era, 1225–1227) (=KOKA)

(iii) *Gosen wakashū* 後撰和歌集 (=GOSE)

(iv) *Shūi wakashū* 拾遺和歌集 (=SIHU)

(v) *Sarashina nikki* 更級日記 (=SARA)

(vi) *Tosa nikki* 土佐日記 (=TOSA)

(vii) *Kindai shūka* 近代秀歌 (*jihitsu-bon* 自筆本, autograph manuscript) (=KIND)

In addition to the seven manuscripts believed with high confidence to have been inscribed by Teika, we generated *jibo* data from two additional manuscripts that are certainly not in the hand of Teika, to use as points of reference. The first is the famous *Shōwa-gire* 昭和切 (=SHOWA), a portion of *Kokin wakashū* amounting approximately to the first half of the anthology, in the hand of Teika's father, Fujiwara no Shunzei 藤原俊成. We extracted the waka from the *Shōwa-gire* and derived the *jibo* data from them, totaling 468 poems.

The second is *Hyakunin shūka* 百人秀歌 (=HNSK), a set of 101 poems, almost all of which overlap with *Hyakunin isshu*. *Hyakunin shūka* has some connection to *Hyakunin isshu*, but whether one text is derived from the other and, if so, which from which, or whether they share a common ancestor is unknown. We used the Reizei 冷泉 family manuscript, which is the ancestor of the only other two copies from the premodern era; its date is unknown but was probably inscribed no later than the middle of the fifteenth century.[38] Previous research has shown that the Reizei *Hyakunin shūka* was inscribed by an amateur whose identity was unknown, but it certainly was not Teika.[39] *Jibo* data was extracted from all 101 poems.

Further information about the manuscripts and their direct sources are given in Appendix B. All of the manuscripts contain substantial numbers of waka and some are composed entirely of waka. Only the texts of waka were extracted from these manuscripts, omitting prose, prefaces, chapter titles, the names of poets, and so forth, in order to generate the most suitable comparison with the

[38] Commentary by Kamijō Shōji in Reizei-ke Shiguretei Bunko 1996.

[39] Takei 2013.

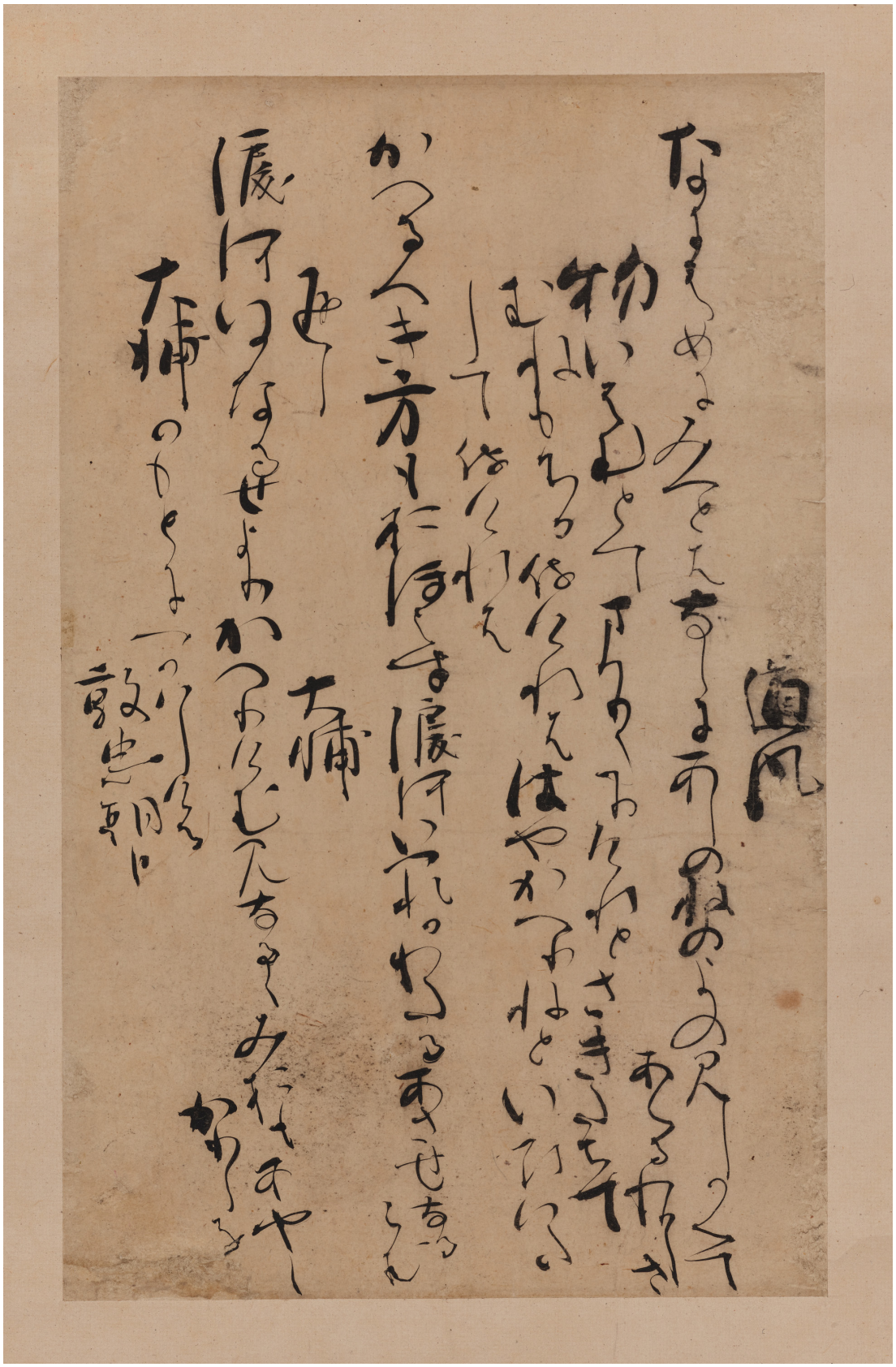

Fig. 3: Fragment from vol. 12 of a copy of *Gosen wakashū* 後撰和歌集, the second imperial collection of Japanese poetry, attributed to the hand of Fujiwara no Teika and formerly owned by the renga master Satomura Jōha 里村紹巴 (1525–1602); Tokyo, Tokyo National Museum, item no. B-3537.

jibo of the waka inscribed in the Ogura *shikishi*. The data from manuscripts ascribed to the brush of Teika total about 5,000 waka or lines and the number of characters totals over 100,000.

Data acquisition and preparation are notoriously time-consuming tasks. We have been very fortunate that *jibo* data and transcriptions of the seven manuscripts by Teika and transcriptions of the two others have been published by Shibuya Ei'ichi, professor emeritus of Takachiho University in Tokyo in his online database, *Fujiwara no Teika no chosaku to Heian-chō kotenseki no shosha kōkan ni kansuru sōgō deetabeesu* 藤原定家の著作と平安朝古典籍の書写校勘に関する総合データベース ('Comprehensive database related to the writings of Fujiwara no Teika and his copying and collation of Heian Era classical texts').[40] We are deeply grateful to Professor Shibuya for making his data freely available with the sole restrictions that any use be not-for-profit and any adaptations be indicated, and we outline our modifications of the data and its processing in the discussion below.[41] This study would not have been possible without Professor Shibuya's valuable data.

First Shibuya's *jibo* data for each manuscript, composed entirely of *kanji*, was aligned with its corresponding transcription, composed of a mixture of *kanji* and *kana*. Professor Shibuya's transcriptions are exceptionally faithful, recording cross-outs, emendations, and interlinear notes, and each line was numbered and coded depending on whether it was poetry or prose. These code numbers were used to extract only the waka poems from each text, by means of a custom formula written with Microsoft Excel. Then the *kana* in the waka transcriptions were converted to romaji, using another formula. In this way, the *kanji* would remain as *kanji*, and the kana would be readily distinguished. This was necessary because some *jibo* are homographic, that is, the same *jibo* might represent different sounds, and therefore needed to be counted separately. For example, the *jibo* 野 typically represents the sound NO, but it can also represent YA. Once the aligned *jibo* and their pronunciations were placed into separate cells, and then concatenated so that each *jibo* and its pronunciation occupied a single cell and each poem a single line, it was relatively simple to generate relative frequencies for each *jibo* in each text, as well as relative frequencies for each *jibo* across all the waka in all seven manuscripts attributed to the hand of Teika (see Figs 15 and 16 in Appendix C). In the process of concatenating *jibo* and pronunciations into individual cells, occasional errors in transcription were inevitably detected and have been corrected.

Rather than generating a list of the *jibo* directly out of the cells that contained each *jibo* and its pronunciation, we used the standard reference work *Kuzushiji yōrei jiten* by Kodama Kōta.[42] In his surveys of classical texts Kodama identified about 300 *jibo*, which formed our list.

[40] Shibuya 1998.

[41] *Riyō kitei* 利用規定 ('Regulations for use'), item 4, <http://genjiemuseum.web.fc2.com/kenkyukai.html> (accessed on 13 August 2025).

[42] Kodama 1981, 1237–1294.

3. Methodology

Previous research has examined Teika's usage of *jibo* at the level of long manuscripts, each manuscript containing thousands of characters. However, since there is no credible evidence that the Ogura *shikishi* constitute a single unified set of inscriptions produced by the same person at about the same time, we must approach them at the level of the individual poem. Specifically, we would like to determine whether each poem in the Ogura *shikishi* conforms to Teika's known patterns of *jibo* usage.

This approach incurs the limitation of sparsity, that is, whether each string of thirty or so *jibo* will be long enough for conclusions that cannot be explained in terms of typical fluctuations in Teika's use of *jibo*. Moreover, it is known that Teika used amanuenses, so the notion of a genuine autograph manuscript (*jihitsu-bon* 自筆本) by Teika has been thrown into question. Finally, our approach assumes for the purposes of argument that Teika's choice of *jibo* was independent – that he did not choose one *jibo* or another based on factors other than personal preference – but previous research shows convincingly that such was not the case. We return to these limitations and others at the end of our analysis, but wish to acknowledge them here at the outset of the description of the methodology.

Under the tentative assumption of independence, we are using the so-called 'bag-of-words' approach (perhaps more accurately termed in this case 'bag of characters'). That is, different orderings of *jibo* in the same poem are treated as equivalent, so the data of each poem becomes a simple count of different *jibo*, without reference to the sequence in which they appear. Let us begin as an example with just the first line of a poem, 'Koi su te fu' こひすてふ (*Hyakunin isshu*, no. 41).[43] Given the string of *jibo* 己飛寸天不, we produce the following string of pronunciations, *jibo*, and counts: KO 己, 1; HI 飛, 1; SU寸, 1; TE 天, 1; FU 不, 1.

In order to distinguish scribeship, we assume that the *jibo* frequencies of scribes differ from those of others but are consistent among their own works. Thus, the numbers of *jibo* used to represent the same sound reflect the probabilities with which the *jibo* may be expected to appear. For example, if the number of times that each *jibo* used to represent the sound HA by Shunzei in his *Shōwa-gire* is as follows: HA 波, 103; HA 者, 288; HA, 八, 361, then we can use these absolute frequencies (AF) to generate relative frequencies (RF), i.e. the ratio of occurrences of each individual *jibo* relative to the sum of the occurrences of all *jibo* representing the same sound. The resulting AFs and RFs of this example are shown in Table 1.

Sound	*Jibo*	AF	RF
HA	波	103	0.14
HA	者	288	0.38
HA	八	361	0.48
(sum)		752	1.00

Table 1: Absolute and relative frequencies of *jibo* for the sound HA in *Shōwa-gire*.

[43] *Koi su te fu / wa ga na wa madaki / tachinikeri / hito shirezu koso / omoisomeshika* こひすてふわがなはまだきたちにけり　ひとしれずこそおもひそめしか ('Already have I gained / the name of one / who is in love / even though my feelings grew / unknown to anyone else.') *Shūi wakashū*, no. 621, by Mibu no Tadami 壬生忠見.

Then, the *jibo* variants for HA would appear with these probabilities:

P(波|HA) = 0.14
P(者|HA) = 0.38
P(八|HA) = 0.48

where the expression P(*jibo*|*sound*) denotes the conditional probability that the individual *jibo* appears among all *jibo* used to represent the same sound. We refer to these probabilities as *relative frequencies* of a given corpus. Although they did not employ this term, relative frequencies have served as the cornerstone of previous analyses, beginning with Komatsu, and continuing through the work of Ue, Iura, and Saitō. Because RFs reflect the orthographic habits of individual scribes, as shown in the previous research, it follows that significant differences in RFs imply differences in scribeship.

An important observation is that RFs for a fixed sound can be viewed as probabilities summing up to one. Data with this particular structure is called *compositional data*, and certain data transformations are necessary for the adaptation of standard statistical techniques. Furthermore, while RFs are good indicators of scribeship at the level of entire manuscripts, metrics for the scribeship of *individual* poems are largely absent from the existing literature. We factor in both of these concerns and propose the *probability rating* (PR), which measures how likely the choices of *jibo* in a given poem are to appear given a scribe's RFs, normalized by the length of the poem. The PR provides a new quantitative criterion for assessing scribeship of individual classical Japanese inscriptions.

The rest of this part on methodology is organized as follows. In Section 3.1, we introduce compositional data analysis and the necessary mathematical background. In Section 3.2, we review principal component analysis, a common dimensionality reduction technique, and outline our method for exploratory data visualization. In Sections 3.3 and 3.4, we introduce the probability rating (PR) of a poem, discuss its probabilistic meaning, and the technique to normalize the PR with respect to poem length. In Section 3.5, we discuss limitations of our approach and directions for improvement.

3.1 Compositional data analysis

Our data is *compositional* in nature. Namely, we are especially concerned not with the exact numbers of *jibo* but rather with their relative proportions. Moreover, the RFs for a given sound are subject to various constraints. Each is always a value between 0 and 1, and they always sum to 1. For example, the *jibo* used to represent the HA sound in Teika's *Date-bon Kokinshū* 伊達本古今集 have the following counts and relative frequencies:

Sound	*Jibo*	AF	RF
HA	波	53	0.036
HA	者	947	0.651
HA	磐	10	0.007
HA	八	446	0.306
(sum)		1456	1.000

Table 2: Absolute and relative frequencies of *jibo* for the sound HA in *Date-bon Kokinshū*.

The difference between the RFs in *Date-bon Kokinshū* and *Shōwa-gire* is clear by comparison with Table 1. Note that RFs in the final column should remain stable no matter how we scale the first row, and will always sum to 1. Data that behaves in this fashion is known as compositional data (CoDA).

Compositional data analysis is a pivotal technique used when data are parts of a whole and convey relative significance. For example, in geochemistry, the composition of elements in a rock sample is analyzed to determine its origin or depositional environment.[44] These data require specific statistical treatment because classical statistical methods can lead to spurious conclusions if the relative nature of the data is ignored.[45] None of the previous scholarship has recognized the necessity of treating *jibo* data as compositional data, and this discovery is a principal contribution of the present article.

Due to their constraints, compositional data cannot be correctly represented by points in the usual Euclidean geometry. The geometric spaces on which these points reside are standard *n*-simplices, in which the use of standard statistical tools may lead to erroneous conclusions. Instead, we need compositional data analysis to study their distribution.

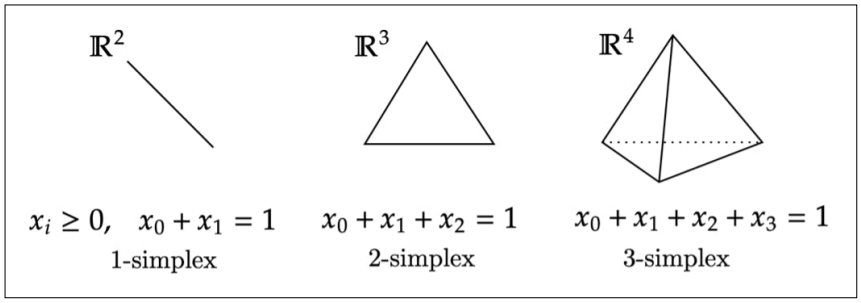

Fig. 4: Standard *n*-simplices in $\mathbb{R}^{n+1}$ as sampling spaces of compositional data.

The *Aitchison geometry* is the standard geometry for CoDA and uniquely satisfies the CoDA axioms.[46] In it, compositional data are treated as points on the open *n*-simplex, an *n*-dimensional

[44] Sadeghi et.al. 2024.

[45] For a technical introduction to compositional data analysis, see Egozcue and Pawlowsky-Glahn 2018 and 2019. See Greenacre 2021 for a survey.

[46] Aitchison 1982.

subset of the Euclidean space $R^{(n+1)}$. See Fig. 4 for an illustration. The simplex serves as the sample space, where each point strictly contains non-negative components summing to a constant, typically 1. The Aitchison geometry is equipped with specialized operations such as perturbation and powering, facilitating the appropriate algebraic manipulation of compositional data.[47]

Distances in Aitchison geometry are given by the inner product

$$\langle x,y\rangle = \sum_{i=0}^{n} log\frac{x_i}{g(x)} \quad log\frac{y_i}{g(y)},$$

where x, y are points on the n-simplex, and $g(x)$ denotes the geometric mean $\left(\prod_{i=0}^{n} x_i\right)^{\frac{1}{n+1}}$ of the components of x. This is a well-defined inner product on the n-simplex because it is scaling-invariant on the arguments x, y.

We now consider the sampling space of *jibo* in poems. If a sound has *n jibo* variants, then the sampling space for that sound is the standard *(n-1)*-simplex. There are 48 sounds used in classical Japanese orthography.[48] Therefore, the sampling space for each of the 48 sound categories is a standard simplex of the right dimension, termed the coordinate simplex. Under the assumption of independence, the sampling space for *jibo* of all sounds is the Cartesian product space of all 48 simplices. Then, each poem is represented by the choice of a point in each coordinate simplex. As an example, a poem beginning '*aware*...' あはれ with the *jibo* string 阿波礼... is represented by a tuple of red dots in the appropriate coordinate simplices (see Fig. 5).

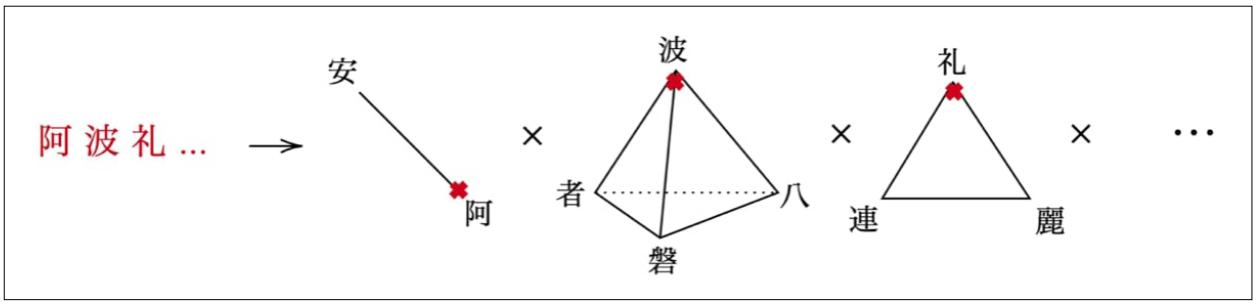

Fig. 5: An example of a poem in sampling space.

To study the distribution of poems in the 48-fold product space of simplices, we must transform the simplices into an honest Euclidean space, while maintaining the interpretability of the coordinates. The technique that achieves this is the central log-ratio transform. Let Δ^n be the open n-simplex embedded in $(n+1)$-dimensional space. The *central log-ratio transform* (CLR) is a smooth map $clr:\Delta^n \to R^{n+1}$ which sends a composition $x = (x_0,...,x_n)$ to

$$clr(x) = \left(log\,\frac{x_0}{g(x)}, log\,\frac{x_1}{g(x)}, ..., log\,\frac{x_n}{g(x)}\right),$$

[47] Aitchison 1982.

[48] Including N无 as a separate sound, about which we have reservations.

where $g(x)$ is the geometric mean of coordinates of x. Since traditional statistical methods like variance analysis, regression models, and hypothesis testing are constructed over Euclidean spaces, we need to apply the CLR transform to compositional data so that statistical tools become available.

It is important to note that the CLR transform has the open simplex as its domain, so values on the boundary – namely, *jibo* with zero occurrences – behave badly under this transformation. The common practice is to add very small constants to each of the zeroes to ensure that all poems transform properly. In particular, we use zero-padding with a Dirichlet prior.[49]

After applying the CLR transform to each of the coordinate simplices, the total sampling space becomes a product of Euclidean spaces. The transformed poems become scattered points in this Euclidean space, and we may now apply our preferred statistical methods on them. Since this space has very large dimension, it is natural to first apply principal component analysis, a common dimension reduction technique.

3.2 Data visualization via principal component analysis

After the CLR transform, the data is now in a form that allows correct application of common statistical methods. This section contains exploratory data visualization as a first step towards the analysis. We apply principal component analysis to reduce the dimensionality of the data and facilitate interpretation thereof. We provide visual evidence that the RFs of a poem distinguish scribeship.

Principal component analysis (PCA) is a statistical procedure that has been widely utilized across various fields for reducing the dimensionality of data sets.[50] This technique enhances interpretability while minimizing information loss. The reduction is achieved by identifying the principal components (PCs) which account for the most variance within the dataset.

PCA has applications in fields such as bioinformatics, image processing, and quantitative finance, where high-dimensional data is common. In image processing, PCA can compress images while retaining most of the original visual information. In genetics, PCA is used to identify genetic markers in dimensionally reduced spaces that are otherwise too complex to analyze.[51] PCA has been considered for CLR transformed data ever since the advent of compositional data analysis.[52]

Mathematically, PCA involves the singular value decomposition of a covariance matrix, typically after mean centering and normalizing the data. The PCs are the eigenvectors of the covariance matrix, and they are orthogonal to each other; the first PC has the highest possible variance, and each succeeding component has the highest variance possible under the constraint that it is orthogonal to the preceding components. The scores of the PCs are the coordinates of the data points in this new basis.

---

[49] See Pawlowsky-Glahn-Bucciatini 2011, 43–58.

[50] See Johnson and Wichern 2007, Chapter 8 on principal components.

[51] See Greenacre et al. 2022 for a survey of applications of PCA in various disciplines.

[52] Aitchison 1983.

One major advantage of PCA is that we can visualize data of very high dimension using 2D plots. However, to make sense of this technique, we need to first transform the compositional data into a Euclidean space. The steps are as follows:

(i) Transform each poem into a counting vector where coordinates are all *jibo* variants.
(ii) Pad the zero values with small values to make the CLR well-defined.
(iii) Perform a CLR transform on *jibo* variants of each sound category.
(iv) Perform a 5-component PCA to reduce dimensionality.
(v) Generate 2D plots of pairs of principal components.

The results of the principal component analysis (PCA) are visualized in Fig. 6, in an array of two-dimensional scatter plots for all six pairwise combinations of the first four principal components (PCs). The data for the seven manuscripts attributed to the hand of Teika are represented by different shades of blue, with the *Shōwa-gire* in orange, the *Hyakunin shūka* in purple, and the Ogura *shikishi* in red. Within each subplot, the distribution of the manuscripts is represented using Kernel Density Estimation (KDE).[53] A filled contour is plotted at the eighty-fifth percentile of each group's density distribution, a method that summarizes the primary location and spread of the data in the PC space while avoiding overplotting.

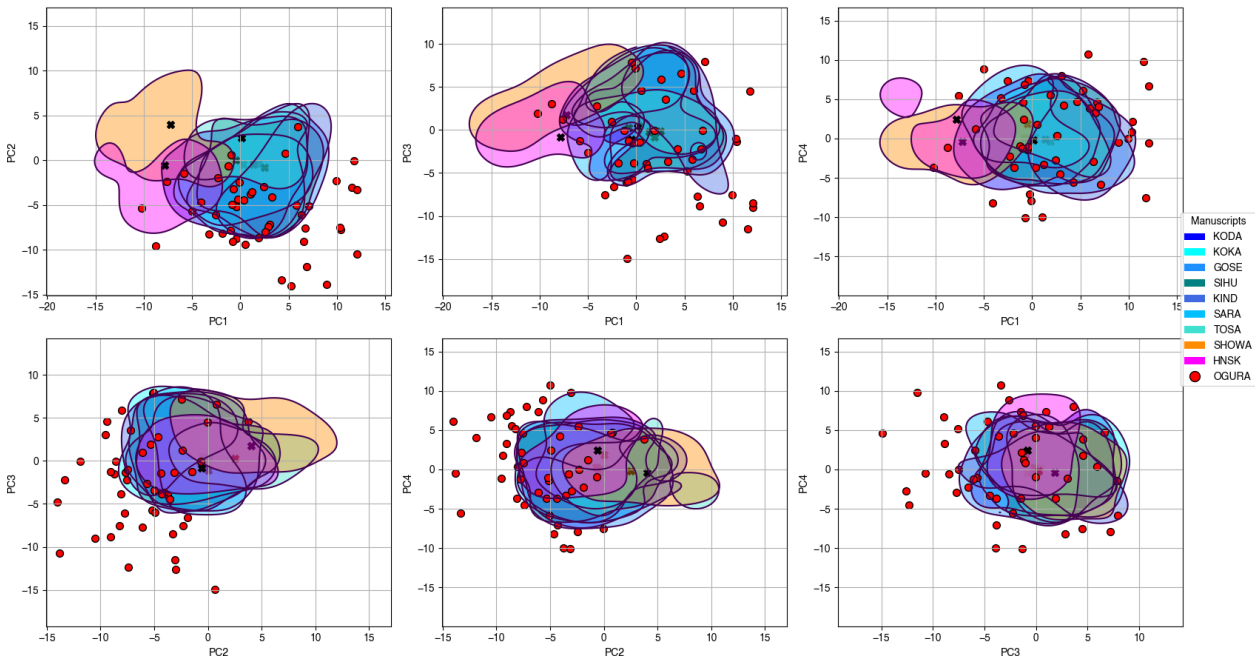

Fig. 6: Plots of leading PCs colored by manuscript.

[53] See Chen 2017.

We observe that the contours of Teika (shades of blue) greatly overlap with each other, while the contours of *Shōwa-gire* (orange), and *Hyakunin shūka* (purple) bulge out of the blue contours, seen most evidently in plots in the first row. Hence, our method of visualization distinguishes different manuscripts by the *jibo* frequencies of different scribes. The red dots which represent the Ogura *shikishi* are more sparsely distributed, many of which are outside of the blue contours representing Teika's scribeship.

3.3 Analysis of individual poems and proposal of a probability rating (PR)

We introduce a new statistic on poems called the probability rating (PR). Our motivation comes from this simple observation: if the text of a poem consists of several *jibo* with low relative frequencies derived from the work of a given scribe, then it is less likely to have been produced by that scribe as a string of *jibo* with the same sounds. We illustrate this fact with the following example.

Say we are given the relative frequencies:

P(已|KO) = 1.00
P(飛|HI) = 0.17
P(寸|SU) = 0.87
P(天|TE) = 1.00
P(不|HU) = 0.58

for the poem fragment KO HI SU TE FU こひすてふ. Under the assumption of independence, the joint probability of a poem's specific *jibo* choices among all possible *jibo* combinations is simply the product of all relative frequencies. Thus the combined probability that 已飛寸天不 appears as a combination of *jibo* for the sounds KO HI SU TE HU equals

P (已飛寸天不 | KO HI SU TE FU)
= P (已|KO) · P (飛|HI) · P (寸|SU) · P (天|TE) · P (不|HU)
= 1.00 · 0.17 · 0.87 · 1.00 · 0.58
= 0.08578

and this combined probability provides a palpable measure of how likely the given poem is to appear across all other poems with the same length in a given manuscript.

Formally, we view each poem as a finite-length string and introduce notation as follows. We represent a poem p as an ordered list of *jibo*, say

$$p = (j_1, j_2, \ldots, j_N),$$

where $N = l(p)$ is the length of the poem. For a sound s, let $jibo(s)$ be the list of *jibo* that represent s. For a $jibo\ j \in jibo(s)$, let $RF(j)$ be the relative frequency of j defined as

$$RF(j) = RF_{\text{Teika}}(j) := \frac{\text{count}_{\text{Teika}}(j)}{\sum_{j' \in jibo(s)} \text{count}_{\text{Teika}}(j')},$$

where $\text{count}_{\text{Teika}}(j)$ denotes the number of occurrences of j across all Teika volumes. Then, for a poem with a fixed length N, the above probability calculation can be summarized as

$$P(p) = \prod_{k=1}^{N} RF(j_k).$$

These probabilities have the following Monte-Carlo interpretation. Given the relative frequencies of one or more manuscripts, we may sample *jibo* choices for a string of sounds independently according to those frequencies. Then, $P(p)$ is the probability of the particular combination of *jibo* occurring among all *jibo* choices for the same sequence of sounds.

3.4 Normalizing the probability rating

Because relative frequencies lie between 0 and 1, PR is necessarily smaller for longer poems unless it is normalized. Therefore, it is only appropriate to compare $P(p)$ across poems of the same length. To make PR comparable across poems of arbitrary length, it is necessary to normalize with respect to $N = \ell(p)$. The simplest normalization is taking the arithmetic mean, i.e. dividing $P(p)$ by $\ell(p)$. However, taking arithmetic means is appropriate only on Euclidean data, and *jibo* data becomes Euclidean after the CLR transform. Thus, we would like to divide the CLR transformed $P(p)$by $\ell(p)$, which, under the inverse transform, is tantamount to taking the geometric mean on the original $P(p)$. Furthermore, the probability $P(p)$ is a multiplicative, so the *geometric mean* is therefore the natural averaging operation.

Furthermore, note that this probability becomes 0 whenever there is a *jibo* in the poem with zero RF, one that Teika never uses throughout the known volumes. The presence of a single zero-frequency *jibo* collapses the PR to zero, nullifying the contribution of all other *jibo* to the product. To deal with this issue, we again appeal to CoDA. We perform the same zero-padding with Dirichlet prior at $\alpha = 0.5$ (see Section 3.1) on the CLR-transformed RFs. Then, we map the padded values back to RF format via an inverse CLR transform. For example, the zero *jibo* A愛 gets assigned a Dirichlet-padded RF of 3.802e-13 – still extremely low, but no longer forcing the poem-level probability to zero. Zero-padding allows the PR to detect subtle differences between poems even with the presence of zero-frequency *jibo*.

With the above modifications, we define the *probability rating* PR of a poem $p = (j_1,...,j_N)$ to be the normalized probability

$$PR(p) = PR_{\text{Teika}}(p) := \left(\prod_{k=1}^{N} RF_{\text{Teika}}(j_k) \right)^{\frac{1}{N}}. \qquad (3.1)$$

It follows that the PR of each poem will be a number between 0 and 1. The value of PR can be thought of as the probability of the specific *jibo* choices of a poem appearing in a given manuscript, assuming the RFs of the manuscript and normalizing every poem to have length 1. Lower PR indicates that the observed combination of *jibo* is less likely under the reference RFs. We propose the following outline that ranks the likelihood of a poem within a given manuscript:

(i) Count and record the number of *jibo* in each poem.
(ii) Calculate the relative frequency RF(j) for each *jibo* j.
(iii) Calculate the PR of each poem according to (3.1).
(iv) Sort the poems in increasing order of their PR and examine ones with the lowest likelihoods. Fit a logit-normal distribution and analyze p-value thresholds.

See Section 4 for detailed analyses of the proposed method and the dataset.

3.5 Limitations of the proposed approach

Our approach is not without its limitations, including those not immediately apparent to us. For the time being, we acknowledge these five potential weaknesses:

(1) *Assumption of independence*. We assumed independence of *jibo* choices, whereas Teika is known to have followed certain conventions in his choices of *jibo*.[54] For example, he avoided using the same *jibo* at the head of two adjacent lines. This suggests dependencies within Teika's choice of *jibo*, which could be evidence against a normal distribution. This could be improved by considering a Markov model and other more advanced methods. Nevertheless, we observe that such dependencies are infrequent, and that it is harmless as a first attempt to make the simplifying assumption that all choices of *jibo* in each poem are independent.

(2) *Sparsity of data at the level of the poem*. A waka poem typically comprises 31 syllables, and is typically written with one or more *kanji* instead of *kana*, reducing the average number of *jibo* per poem to between 25–30. This is at the very low range of a significant sample size. This sparsity may lead to performance issues with our methods, most notably PCA. On the other hand, the PR is more robust to sparse data due to its simplicity and normalization. Hence, we detail in Section 4 the analysis of Ogura *shikishi* using the PR.

(3) *Teika's use of amanuenses*. It is known that Teika employed other persons to copy texts for him, including one retainer who Teika said could imitate his handwriting.[55] As a result, the entire concept of an autograph copy (*jihitsu-bon*) has been questioned, most prominently in the work of Ieiri Hironori. Nonetheless, the seven manuscripts that we have used as a baseline to establish the parameters of Teika's *jibo* usage are not among those questioned by Ieiri.[56]

[54] For example, see Toyoda 1992.

[55] Atkins 2017, 197.

[56] Ieiri 2010.

(4) *Changes in Teika's jibo usage over time*. Teika's use of *jibo* may have changed slightly over time, as shown by Saitō Tetsuya.[57] However, it remains to be seen how much the patterns changed, and whether the degree of changes exceeds that between the seven manuscripts by Teika considered in this study and the individual poems of the Ogura *shikishi*.

(5) *Overlaps between jibo usage of Teika and other scribes*. Our own analysis of the *Shōwa-gire* using the PR method indicates some overlap between *jibo* usage of Teika and that of his father Shunzei, which is to be expected from a culture in which writing was taught in the home, and from a son who deeply admired his father and sought to emulate him in many ways. In our case, this overlap would be relevant if we were attempting to distinguish the *jibo* usage of Teika from his relatives (such as his son Tameie) or other contemporaries (like his friend Fujiwara no Ietaka), but that is not the case in this instance; the choices are strictly between Teika and someone else, anyone else. Moreover, the distances between *jibo* choice in the Ogura *shikishi* and the seven reference manuscripts attributed to the hand of Teika are greater than the distances between Shunzei's *Shōwa-gire* and the seven reference manuscripts.

4. Discussion

This section presents the results and validation of the PR framework introduced in Sections 3.3 and 3.4. In Section 4.1, we validate the PR method using the Teika volumes, *Shōwa-gire*, and *Hyakunin shūka*. We then analyze the PR of poems in the Ogura *shikishi*, report significance thresholds, and conduct sensitivity analysis. In Section 4.2, we discuss PRs of a few selected individual poems and address the intriguing case of the verso inscriptions (*shihai moji* 紙背文字). In Section 4.3, we address the issue of positional dependence among the *jibo* variants. In Section 4.4, we compare the baseline PR method to weighted PR models.

4.1 Probability ratings of poems

We now assess the scribeship of the Ogura *shikishi* using relative frequencies and probability ratings. Even before formal modeling, distinctive patterns from the distribution of RFs across different manuscripts already suggest a difference in the scribeship of the Ogura *shikishi* and the known Teika manuscripts. Fig. 14 in Appendix C contains line plots of RFs of manuscripts considered in this study, ranked according to Teika's overall RFs in descending order. We used a rolling average for smoothing. The top plot displays the raw RF values, and the bottom plot displays log values for better discernment of rare *jibo* tendencies. It is evident from these plots that RFs are consistent across the manuscripts attributed to Teika, while the *Shōwa-gire* and the *Hyakunin shūka* deviate from Teika slightly. RFs of the Ogura *shikishi*, on the other hand, deviate from Teika's distribution in a clearly significant way.

We observe the internal consistency of RFs across the seven Teika manuscripts. However, we see clear differences between the lines representing the Teika manuscripts (blue) vs. Ogura *shikishi*

[57] Saitō 2016 and 2018.

(red) in the second figure. There is clear clustering of the PR calculated for the Teika poems, in shapes of unimodal distributions centered near the probability rating of 0.8. We hypothesize that the Teika PR dataset follows a *logit-normal distribution*. The logit-normal distribution is appropriate for continuously distributed probabilities between 0 and 1. The Ogura *shikishi* in red appear as outliers to the Teika cluster. They do not cluster to form a unimodal distribution, with every poem having a PR between 0 and 0.6, apart from four that appear very close to 0.8, the mode of the Teika distributions. These four poems happen to belong to the verso inscriptions, the analysis of which is deferred to Section 4.2.

We move on to validating our proposed methods in two ways. First, we use *Hyakunin shūka* and *Shōwa-gire*, the two other non-Teika volumes (the underlying *positives*), to perform PR analysis. We calculate test metrics including precision, recall, and the F1-score, demonstrating that the PR test distinguishes known non-Teika volumes from Teika volumes. Second, we perform 10-fold cross-validation within the Teika manuscripts (the underlying *negatives*). We show that PRs within the Teika volumes are internally consistent, and that the test does not tend to flag known Teika poems as false positives.

We begin by calculating test metrics on the underlying positives. The performance of a test is calculated from the confusion matrix, which summarizes the actual versus predicted results of a hypothesis test. The typical structure of a confusion matrix is shown in Table 3.

	Significant (Predicted)	**Not Significant (Predicted)**
Significant (Actual)	True Positive	False Negative
Not Significant (Actual)	False Positive	True Negative

Table 3: The structure of a confusion matrix.

Precision is the ratio of true positives to the total number of positives produced by the test. Recall is the ratio of true positives to the total number of underlying positives. The F1-score is the harmonic mean of precision and recall. Higher F1-scores indicate higher test performance. We randomly split the Teika data into halves, fit a normal distribution to the first half, and use the second half together with either the *Hyakunin shūka* or the *Shōwa-gire* as testing data. The results are shown in Tables 4 and 5.

	Significant (Predicted)	Not Significant (Predicted)
Significant (Actual)	83	18
Not Significant (Actual)	3	98

Precision: 0.965. Recall: 0.822. F1-score: 0.888. Accuracy: 0.896.

Table 4: Confusion matrix for *Hyakunin shūka*.

	Significant (Predicted)	Not Significant (Predicted)
Significant (Actual)	85	387
Not Significant (Actual)	29	467

Precision: 0.746. Recall: 0.180. F1-score: 0.290. Accuracy: 0.570.

Table 5: Confusion matrix for *Shōwa-gire*.

While the test performs well on distinguishing the scribeship of Teika and the *Hyakunin shūka*, it does worse on distinguishing the scribeship of Teika and the *Shōwa-gire*. The high precision and low recall of the PR method are consistent with type II errors. A type II error occurs when the test produces a false negative, i.e., classifying a positive sample as negative. It is therefore considered an error of omission, meaning that the test tends to be more lenient so as to not distinguish outliers from the authentic. That it performs well on the *Hyakunin shūka* but makes the type II error on the *Shōwa-gire* also agrees with our observation that Shunzei (scribe of the *Shōwa-gire*, father of Teika) has a closer orthographic style to Teika than the scribe of the *Hyakunin shūka*, who is presumably not directly related to Teika.

The histogram in Fig. 7 illustrates the fact that the *Shōwa-gire* appears closer to Teika than the *Hyakunin shūka*.

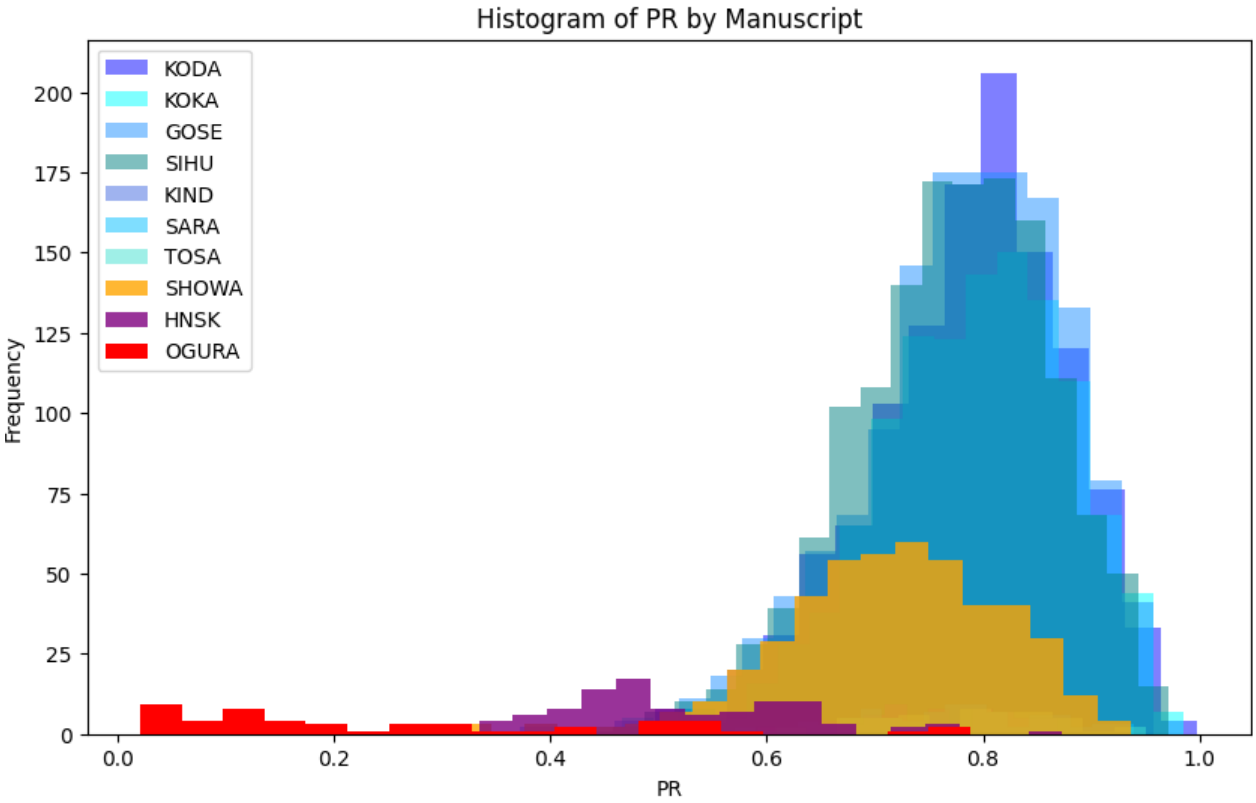

Fig. 7: Histogram of PR for poems in manuscripts attributed to the hand of Teika with high confidence, *Shōwa-gire* (=SHOWA), *Hyakunin shūka* (=HNSK), and the Ogura *shikishi* (=OGURA).

Next, we perform a 10-fold cross-validation within the Teika volumes. We partition the Teika dataset into 10 equally sized subsets, referred to as the 'folds'. In each iteration, one fold is held out as a test set while the remaining 9 folds serve as training data. This process repeats 10 times, with each fold used exactly once as the test set. The results are then averaged across all 10 iterations. Specifically, we iteratively train the logit-normal model on 90% of Teika poems, then count how many of the remaining 10% are false positives. Under a well-calibrated model with $\alpha = 0.05$, we expect approximately 5% of true Teika poems to fall below the significance threshold by chance. The results are displayed in Table 7 below. The observed FP rate of 3.67% is below the nominal $\alpha = 0.05$, suggesting the proxy corpus is internally coherent and, if anything, conservative in its calibration. This concludes the validation phase of our tests.

Metric	Value
Mean False Positive Rate	3.67%
Standard Deviation	1.02%
Expected (under α = 0.05)	5.00%

Table 7: Results of the 10-fold cross-validation.

Having validated the test, we continue with PR analysis of the Ogura *shikishi*. We fit a 1-dimensional logit normal distribution to the Teika data and test the Ogura *shikishi* against the fitted distribution. We set our p-value threshold to be 0.05, following common practices. We dub poems with *p*-values of less than .01 as *inconsistent*, those between .01 and .05 as *marginal*, and those greater than .05 as *consistent*. We display the thresholds calculated from the Teika distribution as follows:

Consistent: PR $\geq$ 0.575

Marginal: 0.488 $\leq$ PR < 0.575

Inconsistent: PR < 0.488

The results are shown in Tables 8–10.

Manuscript	Size	Mean PR	SD PR	5th %ile	95th %ile
KODA	1177	0.787	0.090	0.623	0.921
KOKA	1175	0.789	0.091	0.630	0.923
GOSE	1425	0.782	0.094	0.607	0.920
SIHU	1434	0.770	0.092	0.607	0.911
KIND	68	0.740	0.083	0.602	0.852
SARA	88	0.765	0.087	0.631	0.908
TOSA	59	0.749	0.125	0.508	0.894
OGURA	51	0.269	0.209	0.034	0.654

Table 8: Distributions of PR across manuscripts. Here we treat the Ogura *shikishi* as a single manuscript in order to display the statistics, but we still perform the analysis on individual poems.

Ogura Classification	Count	Percentage
Outside 99% band (Inconsistent)	40	78.4%
Between 95–99% (Marginal)	8	15.7%
Within 95% band (Consistent)	3	5.9%

Table 9: The classification of Ogura poems with respect to Teika's RFs.

DOI: 10.15460/s8e93666

Serial Number	PR	*p*-value	Classification
OGURA074	0.028	3.95e-17	Inconsistent
OGURA047	0.028	6.36e-17	Inconsistent
OGURA089	0.031	2.66e-16	Inconsistent
…	…	…	…
OGURA035	0.494	9.64e-3	Marginal
OGURA004	0.535	0.019	Marginal
OGURA070x2	0.583	0.042	Marginal
OGURA070x1	0.724	0.251	Consistent
OGURA070x3	0.771	0.402	Consistent
OGURA041ax	0.788	0.469	Consistent

Table 10: Statistics of Ogura poems ranked in increasing order of PR.

We see that out of the 51 poems in the Ogura *shikishi*, 48 returned significant *p*-values. Consequently, the statistical evidence strongly refutes the traditional attribution of the Ogura *shikishi* to Teika's own hand. Compare Table 10 also with Fig. 8 for a plot of all of the Ogura *p*-values. In particular, the PRs of the entire Ogura *shikishi* differ significantly from the *Tosa nikki*, which dates the closest to the putative date of the creation of the Ogura *shikishi*, 1235. Even though the *Tosa nikki* has been used as a basis by Nagoya to show similarity of calligraphic styles with the Ogura *shikishi*, the significant difference in their PRs shows that the latter is statistically unlikely to be inscribed by Teika.[58]

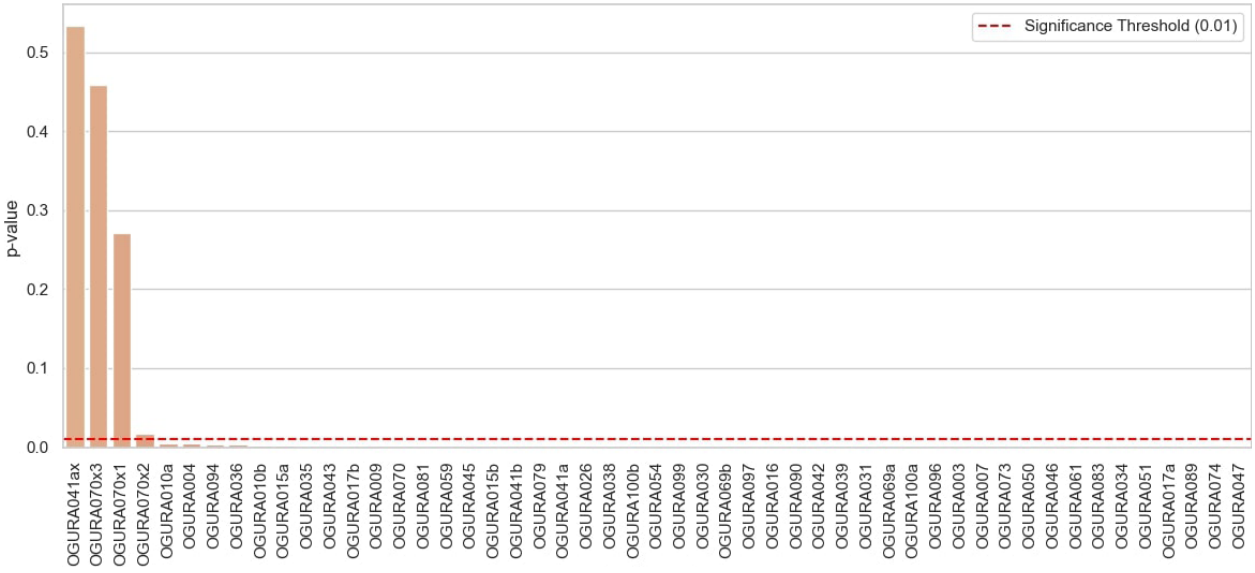

Fig. 8: *p*-values of Ogura *shikishi* poems. Verso inscriptions are leftmost.

[58] Nagoya and Kanechiku 2013.

For sensitivity analysis, we considered leaving out one of the Teika manuscripts and rerunning the tests. Table 11 below shows model statistics and the number of significant Ogura poems when each Teika manuscript is excluded for the purposes of calculating the Teika RFs. The mean of the logit-normal has a range of [1.3431, 1.3831], stable within ±0.02. The standard deviation of the logit-normal has a range of [0.5866, 0.5963], stable within ±0.01. The 48 significant Ogura poems are completely stable across all trials.

Excluded Manuscript	μ (logit)	σ (logit)	OGURA Significant
KODA	1.3831	0.5958	48
KOKA	1.3777	0.5963	48
GOSE	1.3457	0.5872	48
SIHU	1.3431	0.5866	48
KIND	1.3593	0.5908	48
SARA	1.3590	0.5903	48
TOSA	1.3602	0.5904	48
Full corpus	1.3593	0.5907	48

Table 11: Sensitivity analysis of the PR test.

Notice that the four Ogura poems with the highest PRs all have an 'x' in their serial numbers. These are exactly the verso inscriptions (*shihai moji*), which one may recall from Section 1.2. If we mark all 4 verso inscriptions in the Teika distribution, we see that at least 3 of them do not appear to be outliers to Teika's distribution at all. In Fig. 9, the green dotted lines representing these verso inscriptions lie right in the middle of the blue cluster on the histogram. This strongly suggests that scribeship of the Ogura *shikishi* is likely not homogeneous, and that the verso inscriptions have a higher likelihood to be authentic by Teika. We will return to a more detailed discussion of these individual poems in Section 4.2.

All in all, with the only exceptions being the verso inscriptions, the results of the PR test have been validated. There is little chance that any of the Ogura *shikishi* is Teika's own calligraphy.

4.2 Discussion of probability ratings of individual poems

Returning to the level of the individual poems, we would like to reconsider a poem we have been citing only in part and examine it more closely.

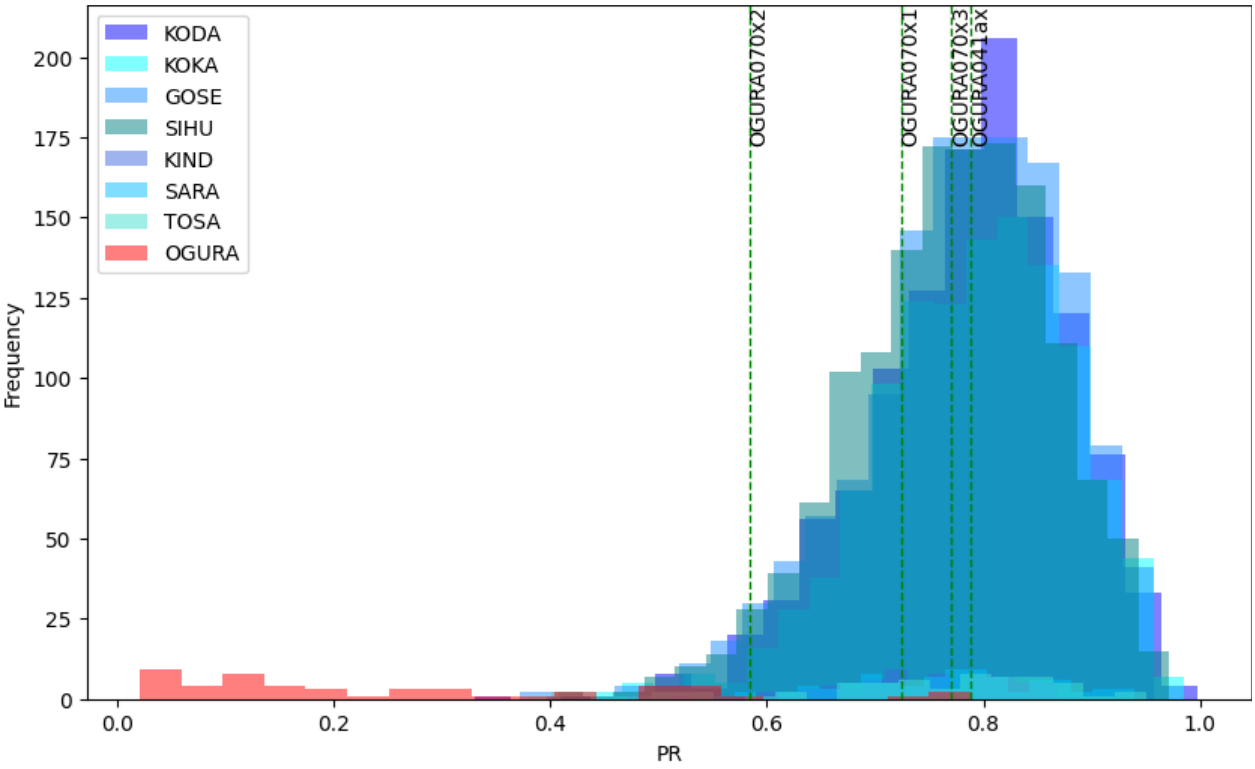

Fig. 9: Histogram of PR with verso inscriptions highlighted.

Koi su te fu wa ga na wa madaki tachinikeri hito shirezu koso omoisomeshika
恋すてふ我が名はまだき立ちにけり人知れずこそ思ひ初めしか
Already I have gained
the name of one
who is in love
even though my feelings grew
unknown to anyone else.

The poem appears in *Shūi wakashū* (no. 621), which is one of the seven manuscripts attributed to Teika that we used as our reference corpus. In that manuscript, Teika uses the following *jibo*:

KO己 HI飛 SU春 TE天 HU不 WA和 KA可 名名 HA者 MA満 TA多 KI幾 立立 NI尓 KE个 RI利 人人 SI志 RE礼 SU寸 KO己 SO曽 思思 SO曽 ME女 SI之 KA可

(In our transliterations, *dakuten* 濁点 voicing marks are omitted and *kanji* and repetition symbols are indicated by doubled characters.)

This poem also appears in *Hyakunin isshu*, no. 41. An extant Ogura *shikishi* now in the collection of the Tokugawa Museum bears its inscription, with these *jibo*:

KO己 HI飛 SU寸 TE天 FU不 WA和 KA可 NA那 HA者 MA満 TA多 KI幾 TA堂 TI知 NI尓 KE介 RI里 HI比 TO止 SI之 RE礼 SU春 KO己 SO處 O於 MO毛 HI飛 SO楚 ME免 SI之 KA加

Based on the relative frequencies we derived from the seven manuscripts inscribed by Teika, we calculate the non-normalized PR of the Tokugawa Ogura *shikishi* at 0.000000000000000000, largely due to the presence of the *jibo* SO楚, which never appears in more than 5,000 poems included in the seven reference manuscripts, but appears four times in the fifty or so poems that appear among the Ogura *shikishi*. Even with the zero RF *jibo*, there are other poems that are even less probable than this one, so it does not have the lowest PR of all, but is not far from the bottom: it ranks the 30th lowest of 5,479 poems. In contrast, we calculate the PR of the *Shūishū* poem inscribed by Teika. It ranks no. 502 of 5,479 poems, just below the tenth percentile, but still well above the highest-ranked recto Ogura *shikishi* at no. 123 (see Fig. 2). Its relatively low ranking may be explained in part by Teika's use in this poem of the *jibo* NA那, which he uses to represent this sound only about ten percent of the time.

The Tokugawa *shikishi* formerly belonged to Tokugawa Ieyasu, who made a copy of it in his own hand, which is also extant, and was transmitted through the Owari branch of his family. Due to the eminence of its owner, it appears at the very beginning of the *Shūko jisshu* volume of Ogura *shikishi* (see Fig. 10).

A comparison with images of the extant object shows that this is indeed a true copy, and the *jibo* match perfectly (see Fig. 11). 'Koi su te fu' is one of the most famous Ogura *shikishi*, in part due to its illustrious provenance, in part due to its colorful paper, the verso inscription, and its lovely, tenuous calligraphy. The exquisite uncertainty evinced in the character *hi* ひ at the top of the third line from the right indeed suggests the trembling heart of one who has begun to fall in love in secret.

This *shikishi* is one of two Ogura *shikishi* bearing verso inscriptions. It was not uncommon for scribes to reuse paper, as it was precious, and in some cases the verso inscriptions had some connection to the material inscribed on the recto; Teika himself was known to copy out his diary on the backs of old letters that had been sent to him around the same time as the events covered in that portion of his diary.

The inscription (Fig. 12), which is incomplete, and therefore left untranslated here, reads:

yo to tomo ni nagarete taenu o / shita ni ya fukaki fuchi to

世止ゝ毛尓流天多衣奴於志多尓也布可幾布知止

We interpreted the *jibo* of this fragment as follows:

世世　TO止　ゝゝ　MO毛　NI尓　流流　TE天　TA多　E衣　NU奴　O於　SI志　TA多　NI尓　YA也　HU布　KA可　KI幾　HU布　TI知　TO止

and evaluated it using the same method as above. Its PR ranks 2,744 out of 5,479 poems, placing it in a very respectable position among poems from manuscripts known to have been inscribed by Teika. Even if we normalized the PR to account for the relatively short length of the fragment, the results should not change much; the rarest *jibo* used in it is SI志, with an RF of 0.18. In this instance, the probability rating of the verso inscription strongly aligns with the Teika baseline, leading us to judge it as authentic.

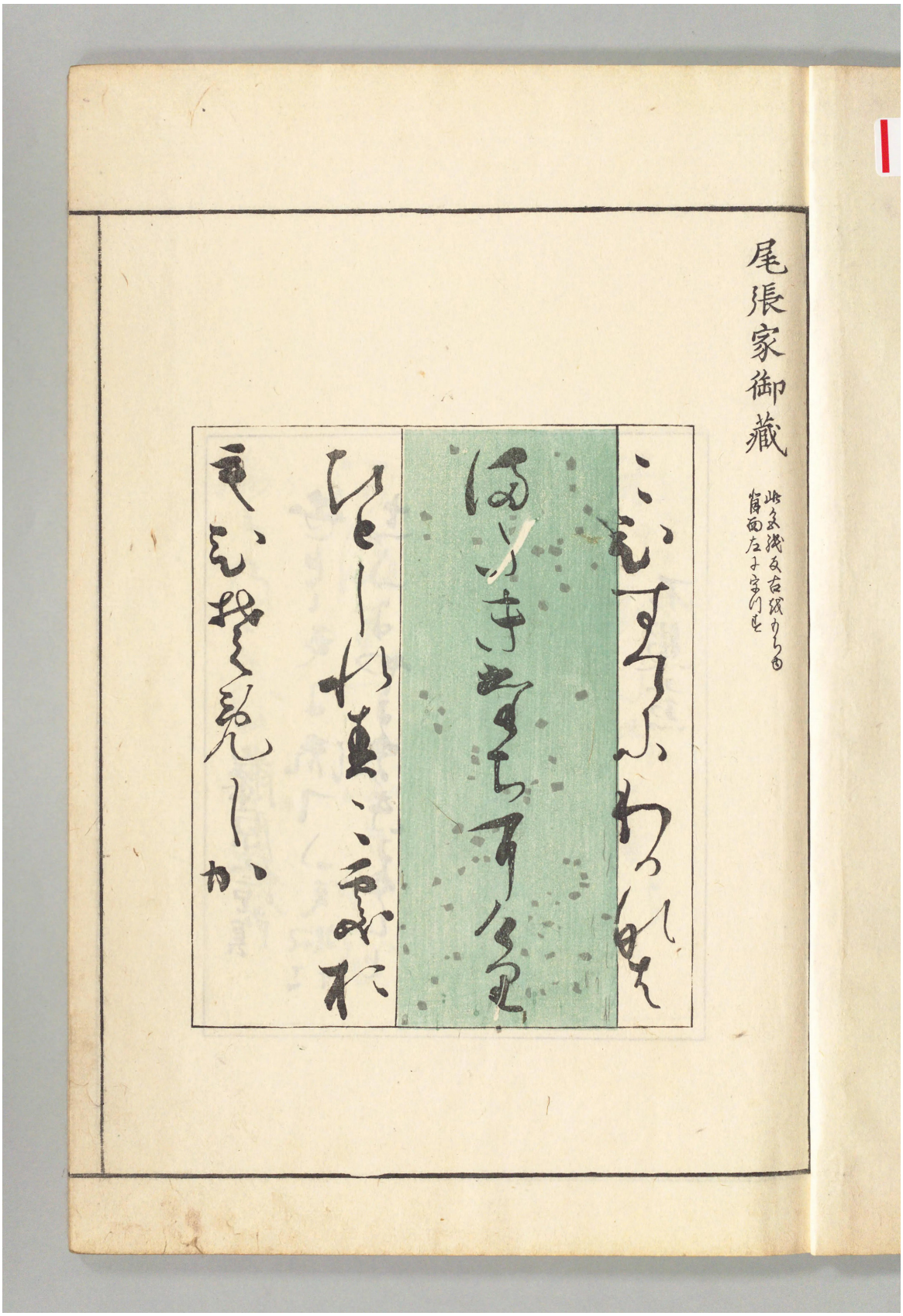

Fig. 10: Copy of the Ogura *shikishi* 'Koi su te fu', no. 41, from *Shūko jisshu*, twentieth-century reprint in the collection of the National Diet Library (Japan), item no. W991-H26.

DOI: 10.15460/s8e93666

Fig. 11: Recto of Ogura *shikishi* 'Koi su te fu', attributed to the hand of Fujiwara no Teika; Nagoya, Tokugawa Art Museum, item no. TAM000053.

DOI: 10.15460/s8e93666

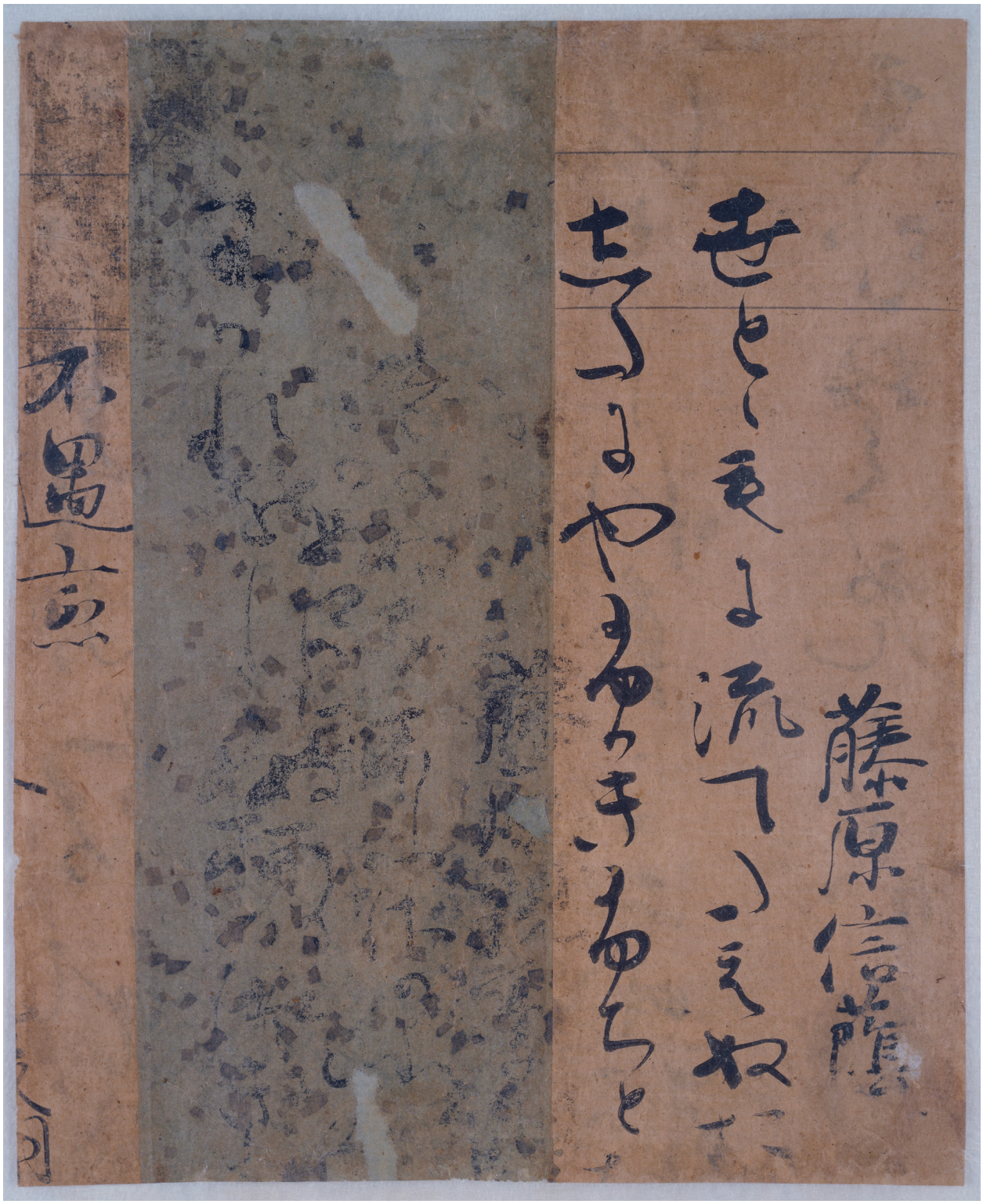

Fig. 12: Verso of Ogura *shikishi* 'Koi su te fu', attributed to the hand of Fujiwara no Teika; Nagoya, Tokugawa Art Museum, item no. TAM7286.

DOI: 10.15460/s8e93666

The other example of a verso inscription in *Shūko jisshu* is a longer fragment on the back of the *shikishi* 'Sabishisa ni', no. 70. Unfortunately, its current whereabouts are unknown. The recto reads:

sabishisa ni yado wo tachiidete nagamuredo izuko mo onaji aki no yūgure
さびしさに宿を立ち出でて眺むれどいづこも同じ秋の夕暮れ
Out of loneliness
I emerge from my dwelling
and gaze out
but it is the same everywhere,
twilight in autumn.

We interpret the *jibo* as follows:

SA佐 HI飛 SI之 SA左 NI尓 YA也 TO登 WO遠 TA堂 TI知 I以 TE天 ゝゝ NA奈 KA可 MU武 RE礼 TO止 I伊 TU徒 KO己 MO毛 O於 NA奈 SI之 秋秋 NO乃 YU遊 HU布 KU具 RE礼

The verso inscription reads as follows, untranslated as it is also a set of three fragments. We divide it into three segments for ease in processing:

70x1: O於 MO毛 HI飛 NE祢 MO毛 KA加 HI比 NA奈 KI幾 KU久 SA佐 I以 TU徒 KA可 MI美 YA也 KO己 WO遠 YU由 ME女 NI尓 TA多 NI尓

70x2: YU遊 KI幾 TO止 MA満 RU留 TO止 KO己 RO呂 TO止 TE天 YA也 HA者 O於 HA者 NA那 KA可 MO毛

70x3: TO止 WO遠 YA也 TO止 ゝゝ SA佐 TA多 SA佐 NO乃 ゝゝ WO遠 KA加 YU由 KU久 KA可 TA多 TO止 WO遠 KI幾

The recto inscription returns a normalized PR of 0.372, ranking 39 out of 5,479. In contrast, the three verso lines return the following PRs and rankings:

Line	PR	Ranking
OGURA70x1	0.724	1,395/5,479
OGURA70x2	0.584	211/5,479
OGURA70x3	0.771	2,352/5,479

Table 12: Probability ratings and rankings of verso texts 70x1, 70x2, and 70x3.

In this case as well, the recto inscription is more likely to be a forgery whereas the verso inscription is more likely to be authentic. Note that OGURA70x2 is borderline, but still above the highest-ranked recto Ogura *shikishi* (ranked at no. 123).

The Ogura *shikishi* that lack verso inscriptions in the hand of Teika could have been created for a variety of reasons: experimentation, personal enjoyment, homage, or even for votive purposes, as Teika was almost venerated as a saint in some circles in Medieval Japan. Therefore, we have been avoiding the term 'forgery' thus far because it implies the intent to deceive. But in the case of the two examples that include verso inscriptions, the persons who inscribed the rectos deliberately sacrificed an authentic piece of Teika's handwriting to create an inscription that looked like Teika's but was not, and this deliberate act certainly seems to qualify as a forgery.

4.3 Analysis of positional dependence

In this section, we address the assumption of independence of *jibo* choices in the dataset across all positions in a poem. For simplicity, we consider the position of a *jibo* in the whole poem as opposed to its lexical position within a lexeme. Recall that for a poem $p=(j_1, j_2,..., j_N)$, the baseline PR is calculated as in (3.1):

$$R(p)=\left(\prod_{j_k \in poem} RF(j_k)\right)^{\frac{1}{N}}.$$

We introduce positional dependence to the model by defining the relative frequency for *jibo* $j \in jibo(s)$ of sound s at position $p \in \{1,2,\ldots,31\}$ as

$$RF^{(p)}(j) := \frac{\text{count}_{\text{Teika}}\,(j \textit{ at position } p)}{\sum_{j' \in jibo(s)} \text{count}_{\text{Teika}}\,(j' \textit{ at position } p)}.$$

Then, we define the position-dependent probability rating as

$$PR^{pos}(p) := \left(\prod_{j_k \in poem} RF^{(p)}(j_k)\right)^{\frac{1}{N}}. \qquad (4.1)$$

We display some of the results below. First, we discuss patterns in positional RFs. While most *jibo* show less-than-two-percent positional variation in the RF, specific combinations like NI尓 at position 1 show substantial deviation. However, this represents rare edge cases. See Table 13 for examples.

Jibo	Position	Positional RF	Global RF	Deviation (%)
NO乃	1	0.9999	0.9920	+0.80%
NO乃	4	0.9836	0.9920	−0.85%
NO乃	5	0.9747	0.9920	−1.75%
NI尓	1	0.5714	0.9598	−40.46%
NI尓	3	1.0000	0.9598	+4.19%
NI尓	7	0.9891	0.9598	+3.05%

Table 13: Positionally-conditioned RFs of NO 乃 and NI 尓.

To quantify positional variation for each *jibo*, we compute the *normalized entropy* across positions:

$$H_j = \frac{-1}{log(N)} \sum_{p=1}^{P} RF^{(p)}(j) log RF^{(p)}(j),$$

where P=31 is the number of positions. The above entropy is a measure of uncertainty or randomness in a probability distribution. Values close to 1 indicate uniform usage across positions. Values close to 0 indicate strong positional preference. The average normalized entropy across all positions is calculated as 0.286. We conclude that while moderate positional dependence exists, most *jibo* maintain consistent usage patterns.

Next, we report the comparison of the positional model with the baseline PR model in Table 14.

Metric	Value
Mean absolute difference in log PR	0.0605
Standard deviation of difference	0.1171
Maximum absolute difference	1.5573
Correlation (baseline vs positional model)	0.9270
T-statistic	8.9035
p-value	7.05×10^{-1}

Table 14: Comparison between baseline and positional PR models.

While the difference between global and positional models is statistically significant with $p<10^{-18}$, the correlation of 0.927 means that 92.7% of the variance is shared between models. These results confirm empirical observations that Teika – and scribes more generally – tended to use different *jibo* depending on the position within a poem (see Section 1). However, for practical classification purposes, the models would yield very similar results.

We now study the impact of positional dependence on the Ogura classification problem. Table 15 shows a comparison of Ogura poems significantly different from the Teika RF distributions according to the baseline and the positional models.

Metric	Value
Total OGURA poems	51
Significant (Baseline model, $p < 0.05$)	48
Significant (Positional model, $p < 0.05$)	47
Classifications changed	1

Table 15: Impact of positional dependence on Ogura classification.

Only 1 of 51 Ogura poems change classification when positional dependence is modeled, and this poem (OGURA070x2) is a borderline case, with its *p*-value changing from 0.042 (significant) to 0.074 (not significant at the 0.05 level, but still atypical). It is, of course, one of the verso inscriptions. We display OGURA070x2 below for the reader's reference – note the rare *jibo* YU遊 and NA那. The remaining 47 significant poems maintain strong significance under both models.

Raw poem text for OGURA070x2:
YU遊 KI幾 TO止 MA満 RU留 TO止 KO己 RO呂 TO止 TE天 YA也 HA者 O於 HA者 NA那 KA可 MO毛 TO止 WO遠 YA也 TO止 ゝゝ SA佐 TA多
We conclude that positional dependence exists as predicted but has minimal impact on final classifications. At the same time, position-specific effects are localized. The extreme case of NI 尓 at position 1, which caused 40% deviation, affects only seven poems in the entire corpus, demonstrating that positional effects, while present, are sparse and localized. The baseline model, despite not taking the full effects of positional dependence and other potential dependencies into account, still provides an excellent approximation with *substantially simpler computation*. Therefore, we conclude that the baseline PR model remains appropriate for distinguishing scribeship, even with the assumption of total statistical independence in the *jibo* choices.

4.4 Comparison with weighted models

The baseline PR model regards all sounds as equally important, whereas in practice, certain sounds are more common than others, and some sounds cause more variation than others in the dataset. Although these factors can affect the baseline method, both kinds can be addressed by adding weights. In this section, we consider more sophisticated models weighted by count and entropy.

We extend the PR formula to incorporate *jibo*-level weights as follows. For a poem *p*, we define the *weighted probability rating* with weights *w* as

$$\log PR^{(w)}(p) := \sum_{j \in p} {}_{wj} \cdot n_j \cdot logRF(j), \qquad (4.2)$$

where n_j is the count of the *jibo j* in *p*, and w_j is the weight assigned to *j*. We consider the following two natural candidates for *w*.

Count-based weighting. Sounds with more variants receive heavier weights. If $j \in jibo(s)$ for the sound *s*, then we define

$$w_j^{var} := log(|jibo(s)|).$$

Entropy-based weighting. Entropy here measures how 'spread out' Teika's preferences are among variants for a given sound. High entropy means Teika uses multiple variants roughly equally. Low entropy means Teika strongly prefers one variant. We give more weight to sounds with lower entropy, since sounds with higher entropy have a more uniform distribution of *jibo* and hence carry less discriminative power. We define

$$w_j^{ent} := H_{max} - H_{jibo(s)}$$

where

$$H_{jibo(s)} := -\sum_{j' \in jibo(s)} RF(j')\ logRF(j')$$

is the entropy of the sound s, and H_{max} is the maximum possible entropy across all sounds.

The results of the two weighted models are displayed in Tables 16–18 and Fig. 13.

Sound	Entropy Weight	Variant Weight
A	0.349	0.843
HA	1.998	1.375
HI	4.095	1.064
HO	4.384	0.843
KI	0.100	1.235
MA	3.555	1.064
N	0.000	0.532
NO	0.153	1.235
RI	0.577	1.235

Table 16: Two kinds of weights for selected sounds.

Method	μ (logit)	σ (logit)	OGURA Significant
Unweighted PR	1.3593	0.5907	48/51
Entropy-Weighted	0.3550	0.515	41/51
Variant-Weighted	1.2929	0.5953	47/51

Table 17: Impact on Ogura classification.

Weighting Scheme	Correlation (r)
Entropy-Weighted	0.7797
Variant-Weighted	0.9942

Table 18: Correlation with baseline PR model.

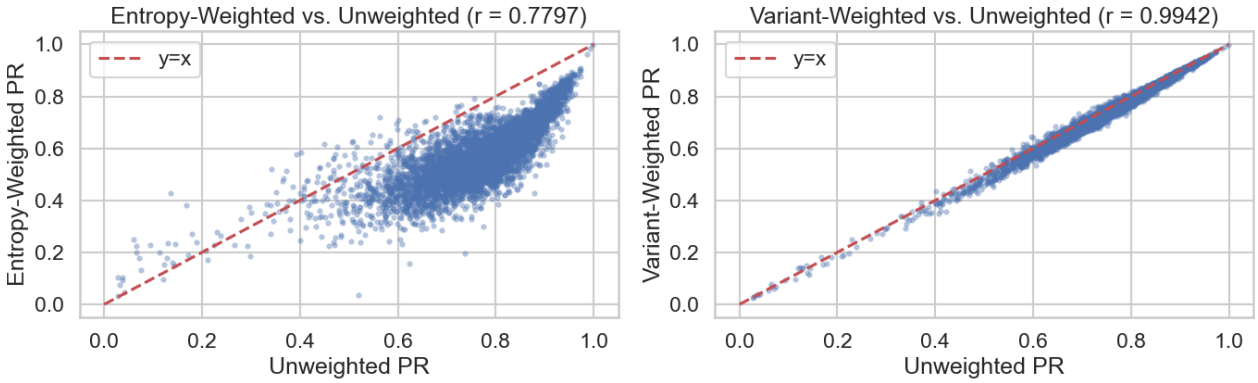

Fig. 13: Correlation of the weighted models with the baseline PR model.

The variant-weighted PR shows a correlation of 0.9942 with unweighted PR, indicating that accounting for variant counts produces essentially the same rankings. This validates the robustness of the unweighted approach. Furthermore, the only Ogura poem that flips classification is again the marginal OGURA070x2. Recall from Section 4.3 that the positional model also flipped the classification of this single poem, confirming its borderline nature.

The entropy-weighted PR has correlation of 0.78 with the unweighted PR model. However, it still identifies 41/51 Ogura poems as significant. The seven poems that lose significance are the borderline cases. The classification of these seven poems relies heavily on sounds for which Teika uses near-exclusive variants with entropy close to zero, so weighting by entropy downweighs their contribution significantly. Below are the six poems that lost significance apart from the previously discussed verso inscription OGURA070x2. Note that all six poems are classified by the baseline PR model as *marginal*.

OGURA004: 田田 子子 NO乃 浦浦 NI尓 U宇 TI地 出出 TE天 見見 RE礼 HA者 白白 妙妙 NO乃 HU布 SI之 NO能 TA堂 KA可 NE祢 NI尓 YU由 KI幾 HA者 HU布 RI利 TU川 ゝゝ (田子の浦に打ち出でて見れば白妙の富士の高嶺に雪はふりつゝ – Poem 4)

OGURA010a: KO古 RE礼 YA也 KO己 NO乃 YU由 KU久 MO毛 KA加 HE部 RU累 MO毛 WA和 KA可 RE礼 TU川 ゝゝ SI志 RU留 MO毛 SI志 RA良 NU奴 MO毛 相相 坂坂 NO乃 SE世 KI幾 (これやこのゆくも帰るも別れつゝ知るも知らぬも逢坂の関 – Poem 10, version a)

OGURA010b: KO古 RE礼 YA也 KO己 NO乃 YU由 KU久 MO毛 KA加 HE部 RU留 MO毛 WA和 KA可 RE礼 TU徒 ゝゝ SI志 RU累 MO毛 SI志 RA良 NU奴 MO毛 相相 坂坂 NO乃 SE世 KI幾 (これやこのゆくも帰るも別れつゝ知るも知らぬも逢坂の関 – Poem 10, version b)

OGURA015a: 君君 KA可 TA堂 ME免 春春 NO乃 野野 NI尓 I伊 TE天 ゝゝ WA和 KA可 NA奈 TU川 MU武 我我 衣衣 手手 NI仁 YU由 KI幾 HA者 HU布 RI利 TU川 ゝゝ (君がため春の野に出でて若菜摘むわが衣手に雪はふりつゝ – Poem 15, version a)

OGURA035: 人人 HA者 I以 SA左 KO己 KOゝ RO呂 MO毛 SI之 RA良 SU須 HU布 RU留 郷郷 HA八 HA者 NA那 SO曽 MU武 KA可 SI之 NO乃 KA加 NI尓 ゝゝ HO本 HI日 KE遣 RU留 (人はいさ心も知らずふるさとは花ぞ昔の香ににほひける – Poem 35)

OGURA036: 夏夏 NO乃 夜夜 HA八 MA末 TA太 宵宵 NA奈 KA加 RA良 A安 KE計 NU奴 RU累 WO遠 雲雲 NO乃 I以 TU徒 KO己 NI尓 月月 YA也 TO止 RU留 RA良 MU武 (夏の夜はまだ宵ながら明けぬるを雲のいづこに月やどるらむ – Poem 36)

OGURA094: MI見 YO与 SI之 NO乃 ゝゝ YA也 MA末 NO乃 A安 KI幾 KA加 SE勢 SA左 YO与 HU不 KE遣 TE天 HU布 RU留 郷郷 SA左 MU武 KU久 KO己 RO呂 MO毛 U宇 TU徒 NA奈 RI利 (みよしのの山の秋風さ夜ふけてふるさと寒く衣うつなり – Poem 94)

5. Conclusions

In this paper, we have addressed the authenticity of the Ogura *shikishi* from a single perspective: whether their *jibo* usage is consistent with that found in manuscripts attributed with high confidence to Fujiwara no Teika.

We have built on previous approaches that calculate the relative frequencies of individual *jibo* in a manuscript and compared them to the same figures in other manuscripts, in order to cluster manuscripts and draw conclusions about scribeship or date of inscription. However, by the very nature of the question posed, this study has had to apply these methods to inscriptions that are the length of individual short poems, pushing the boundaries of statistical significance. In order to carry out this approach, we proposed a new index, termed probability rating (PR), in which each *jibo* in a poem is replaced with its relative frequency in a corpus, the relative frequencies are multiplied, and that product is normalized for the number of *jibo* by dividing by the *n*th root, when *n* equals the number of *jibo*. Then the poems are ordered by PR, and the clusters are analyzed.

We validated our approach and our choice of reference manuscripts using PCA, F1-scores, and cross-validation, which showed that *jibo* choice was consistent among seven manuscripts attributed to the hand of Teika, and differed from that observed in a manuscript attributed to the hand of his father, Shunzei, and from that in a manuscript in the hand of an unknown scribe.

Our comparison of the PR of approximately fifty poems reproduced in the Ogura *shikishi* with over 5,000 poems from manuscripts believed to have been inscribed by Teika shows that every analyzed recto Ogura *shikishi* inscription has a PR inconsistent with the Teika baseline. If these inscriptions were consistent with the Teika baseline, we would expect these fifty poems to be

evenly distributed among the 5,000 reference poems, but they are clustered toward the bottom. Over thirty of them rank at the absolute bottom of the list, and none rank higher than the bottom 3.5%. This indicates that it is highly unlikely, under the modeled *jibo* frequencies, that any of the analyzed recto Ogura inscriptions was inscribed by Teika.

Finally, we obtained verso inscriptions from the backs of two Ogura *shikishi*, one known to be extant and the other unknown. Verso inscriptions have been adduced in the previous scholarship as evidence of possible authenticity. We subjected the same analysis to the verso inscriptions and found that the PR of the verso inscriptions was highly consistent with *jibo* usage in manuscripts believed to have been inscribed by Teika, in sharp contrast to the inscriptions on their rectos. Therefore, we believe that the Ogura *shikishi* containing verso inscriptions were inscribed on the back of authentic manuscripts in the hand of Teika. The use of a relatively valuable manuscript in the hand of Teika by other persons to create a secondary manuscript in Teika's style of writing strongly suggests that they are both forgeries.

Our conclusions extend beyond the Ogura *shikishi* to *Hyakunin isshu* itself, an extraordinarily influential poetic text with a complicated history. Based on our analysis, the Ogura *shikishi* cannot be treated as evidence substantiating an attribution of the compilation of *Hyakunin isshu* to Teika.

The statistical analysis of *jibo* usage has great potential for further use, as previous scholarship, including the most recent scholarship by Saitō Tetsuya, clearly demonstrates.[59] As in any data science project, obtaining good data is a major obstacle. Machine learning appears to have great potential for generating *jibo* data in large quantities with little effort, using digitized images of manuscripts. Analysis of *jibo* using statistical methods holds immense potential to aid curators, librarians, dealers, and collectors ascertain scribeship, and to help scholars rewrite the histories of classical Japanese language, literature, and art, exposing hidden connections between disparate manuscripts and revealing truths and falsehoods hitherto unknown.

Statement on data availability

The data used to conduct the analyses described in this article have been uploaded to Zenodo and are available at the following link: 10.5281/zenodo.18793911.

Acknowledgements

Versions of this paper were presented at Japan Women's University (Tokyo); Asian Studies Conference Japan (Tokyo); Japanese Association for Digital Humanities annual conference (University of Osaka); the Marco Institute for Medieval & Renaissance Studies (University of Tennessee, Knoxville); Walter Chapin Simpson Center for the Humanities (University of Washington, Seattle); the Tateuchi East Asia Library (University of Washington, Seattle), and as the 28th Yamagata Bantō Prize Commemorative Lecture (Osaka Prefecture). The authors are

[59] See Saitō 2016 and 2018.

grateful to all the organizers and audiences. Advice at an early stage of this project was provided by the UW Statistical Consulting Service. Moreover, the authors thank Professors Sakamoto Kiyoe, Saitō Tetsuya, and Kanechiku Nobuyuki and Dr. Herman Chau for especially valuable advice and encouragement. The authors are grateful to the two anonymous peer reviewers and editor Dr. Thies Staack for their many helpful suggestions and corrections.

Appendix A: List of all known Ogura *shikishi*, extant as objects or in form of transcriptions, with *jibo* data

For reference, the number of the corresponding poem in *Hyakunin isshu* (HNIS) is given. Versions of the same poem with different *jibo* data are indicated after the poem number by the letters a, b, and c. The letter x after the poem number indicates that the *jibo* data derives from a verso inscription. The letter X in the transcription indicates unreadable or missing text.

No.	HNIS	*Jibo*
1	3	安之飛幾乃也滿登李能乎乃志堂里於農奈可、、之夜遠飛止利加毛祢武
2	4	田子乃浦尓宇地出天見礼者白妙乃布之能堂可祢尓由幾者布利川、
3	6	加佐、幾乃和多世留者之尓X久新毛乃志呂起X見礼者与楚不X耳个累
4	7	安満乃者羅布里佐遣見礼波加数可奈累美可左能山耳伊帝之月加毛
5	9	花乃色者宇川利尓遣利奈以堂徒良尓和可身世仁布留奈加免世志満仁
6	10a	古礼也己乃由久毛加部累毛和可礼川、志留毛志良奴毛相坂乃世幾
7	10b	古礼也己乃由久毛加部留毛和可礼徒、志累毛志良奴毛相坂乃世幾
8	15a	君可堂免春乃野尓伊天、和可奈川武我衣手仁由幾者布利川、
9	15b	幾見可堂免者留乃野耳伊天、和加奈徒武我己呂毛手仁由幾者布里徒、
10	16	堂知和可礼伊奈葉乃山能美祢耳於布類満川止之幾可波以末加部李己武
11	17a	知者也布累加美世裳幾可須多徒田河加羅久礼奈井耳見春具、類止八
12	17b	千者也不留神代毛幾加寸龍田河加良久礼奈為尓水具、累止波
13	26	越久羅也満美祢乃毛美知葉己、呂安良波伊末飛止多比乃美遊幾満多奈无
14	30	晨明乃徒連奈久見盈之和可礼与里安可川幾者加李宇起乃波奈
15	31	朝保良遣有明乃月止三累末天尓芳乃、佐登耳不礼留志羅由幾
16	34	耳勢无太可満徒毛武可之乃裳奈良奈久尓
17	35	人者以左己、呂毛之良須布留郷八者那曽武可之乃加尓、本日遣留
18	36	夏乃夜八末太宵奈可良安計奴累遠雲乃以徒己尓月也止留良武
19	38	王数良累、身遠波於毛者須知可飛天志飛止乃伊能知乃於之具毛安累可那

20	39	安佐地不乃遠能、新能者羅志乃布礼止阿満里天奈止加飛登乃己飛之幾
21	40	志乃布礼止以呂尓以天耳気里和可己比波毛乃也於毛不止人能登不満天
22	41a	己飛寸天不和可那者満多幾堂知尓介里比止之礼春己處於毛飛楚免之加
23	41ax	世止、毛尓流天多衣奴於志多尓也布可幾布知止
24	41b	己飛寸天不和可那者満多幾堂知尓介里比止之礼寸己處於毛飛楚免之加
25	42	知幾梨幾那可多美尓楚天遠志保里徒、数衛乃末川山奈美己佐之止波
26	43	安飛見天乃、知能己、呂仁久良布礼者武可之者毛乃毛於毛者佐里計利
27	45	安者礼止毛伊不遍幾飛止者於毛保盈帝身乃伊多川良尓奈里奴部幾可那
28	46	遊羅乃止遠和多類布奈比止加知越多盈由久衛毛新羅怒己飛乃美知可那
29	47	也遍武久羅新介礼類屋戸能佐飛之起耳飛止己處見盈祢安幾者幾尓介礼
30	50	可免於之加羅佐李新伊乃知佐遍奈可久毛可那登於毛比奴累哉
31	51	加久登堂耳盈也葉伊不木乃佐新裳草左之毛志羅志那毛遊累於毛比遠
32	54	王数礼志乃由久寸部万天波加堂介礼者遣不遠可幾李乃伊能知止毛可那
33	59	也数良者天祢奈満之毛乃遠佐夜布計帝加堂婦久満天乃徒幾遠見之可那
34	61	伊尓之遍農奈良乃美也己能也部佐久羅幾不古、乃部尓耳保飛奴類可那
35	69a	阿羅之婦具美武呂乃山乃毛美地葉、堂徒太乃加者乃耳之幾奈利介里
36	69b	嵐吹見武呂乃山能毛美知葉者太徒太乃加者農錦奈利遣利
37	70	佐飛之左尓也登遠堂知以天、奈可武礼止伊徒己毛於奈之秋乃遊布具礼
38	70x1	於毛飛祢毛加比奈幾久佐以徒可美也己遠由女尓多尓
39	70x2	遊幾止満留止己呂止天也者於者那毛止遠也止、佐多
40	70x3	佐乃、遠加由久可多止遠幾
41	73	堂可佐己乃於能遍乃佐久羅佐幾耳介李登也末乃加須美太、数毛安良奈武
42	74	宇加里遣類飛登遠葉徒勢乃山乎呂志与波計之閑礼止者移農羅奴毛乃遠
43	76a	和多乃者良己幾出天美連八久可多乃雲井耳万可不沖徒之良浪
44	76b	和堂乃者羅古幾伊天帝見 礼X飛XXX乃久毛XXXXXXXXXXX
45	76c	和堂乃X羅古幾伊天帝見礼波飛佐加太乃久毛井X万加不於幾徒志良奈美
46	79	秋風尓太奈比久雲能多盈万与利毛礼以津累月乃加計乃左也介佐
47	81	保止、幾須奈幾徒留加多遠奈可武礼者堂、安里阿計乃徒幾處能己礼累
48	83	世中与美知己處奈遣礼於裳飛以留也満乃奈可耳裳志可楚那久奈類
49	89	堂満農遠与田江那者多衣祢那可羅邉波志乃布累己登能与者梨毛處数留

50	90	見勢者也那遠之満乃安末乃楚天太耳毛奴礼耳曽奴礼之伊呂者加者良数
51	94	見与之乃、也末乃安幾加勢左与不遣天布留郷左武久己呂毛宇徒奈利
52	96	者那佐曽婦安羅新乃仁者乃由幾奈良天布里行物者我身也介里
53	97	古奴飛止遠満川本乃宇羅能由不那幾尓也久也裳之保乃身毛己可礼徒、
54	99	飛止裳於之比止毛宇良免之安地幾那具世遠於毛不遊部耳毛乃思布身波
55	100a	百之起屋布累幾乃支者乃忍尓毛猶阿満利安留武可之奈利介留
56	100b	百敷也布累幾軒者乃志乃布耳毛猶安末李安留武加之奈利介留

Appendix B: List of manuscripts used to produce relative frequencies of *jibo* in this study

Items 1–7 are attributed to the hand of Teika with confidence; items 8–9 to other persons. All sites accessed on 13 August 2025.

1. *Kokin wakashū* 古今和歌集 (*Date-bon* 伊達本)
Jibo data: <http://genjiemuseum.web.fc2.com/koda2.html>
Transcription: <http://genjiemuseum.web.fc2.com/koda1.html>
Source: Kyūsojiin (ed.) 1991.

2. *Kokin wakashū* 古今和歌集 (*Karoku-bon* 嘉禄本)
Jibo data: <http://genjiemuseum.web.fc2.com/koka2.html>
Transcription: <http://genjiemuseum.web.fc2.com/koka1.html>
Source: Reizei-ke Shiguretei Bunko (ed.) 1994.

3. *Gosen wakashū* 後撰和歌集
Jibo data: <http://genjiemuseum.web.fc2.com/gose3.html>
Transcription: <http://genjiemuseum.web.fc2.com/gose2.html>
Source: Reizei-ke Shiguretei Bunko (ed.) 2004.

4. *Shūi wakashū* 拾遺和歌集
Jibo data: <http://genjiemuseum.web.fc2.com/sihu3.html>
Transcription: <http://genjiemuseum.web.fc2.com/sihu2.html>
Source: Kyūsojiin (ed.) 1990.

5. *Sarashina nikki* 更級日記
Jibo data: <http://genjiemuseum.web.fc2.com/sara3.html>
Transcription: <http://genjiemuseum.web.fc2.com/sara2.html>
Sources: *Gyobutsu Sarashina nikki*, 1981; and *Sarashina nikki*, 2004.

6. *Tosa nikki* 土佐日記
Jibo data: <http://genjiemuseum.web.fc2.com/tosa2.html>
Transcription: <http://genjiemuseum.web.fc2.com/tosa3.html>
Source: No specific edition is cited, but see *Teika-bon Tosa nikki*, 1928.

7. *Kindai shūka* 近代秀歌 (*jihitsu-bon* 自筆本)
Jibo data: <http://genjiemuseum.web.fc2.com/kind2.html>
Transcription: <http://genjiemuseum.web.fc2.com/kind3.html>
Source: No specific edition is cited, but see Furuya 1979.

8. *Shōwa-gire* 昭和切 (first ten books of *Kokin wakashū*, 468 waka, in the hand of Fujiwara no Shunzei)
Jibo data: *Shōwa-gire*, 1935.
Transcription: <http://genjiemuseum.web.fc2.com/shouwa1.html>
Source: *Shōwa-gire*, 1935.

9. *Hyakunin shūka* 百人秀歌 (anthology of 101 waka poems, largely overlapping with *Hyakunin isshu*; in the hand of an anonymous amateur, certainly not Teika or Shunzei)
Jibo data: derived by Paul S. Atkins
Transcription: <http://genjiemuseum.web.fc2.com/hyaku2.html>
Source: Reizei-ke Shiguretei Bunko (ed.) 1996, pp. 607–652 (Chinese numerals). A discussion by Kamijō Shōji is provided on pp. 92–96 (Arabic numerals).

Appendix C

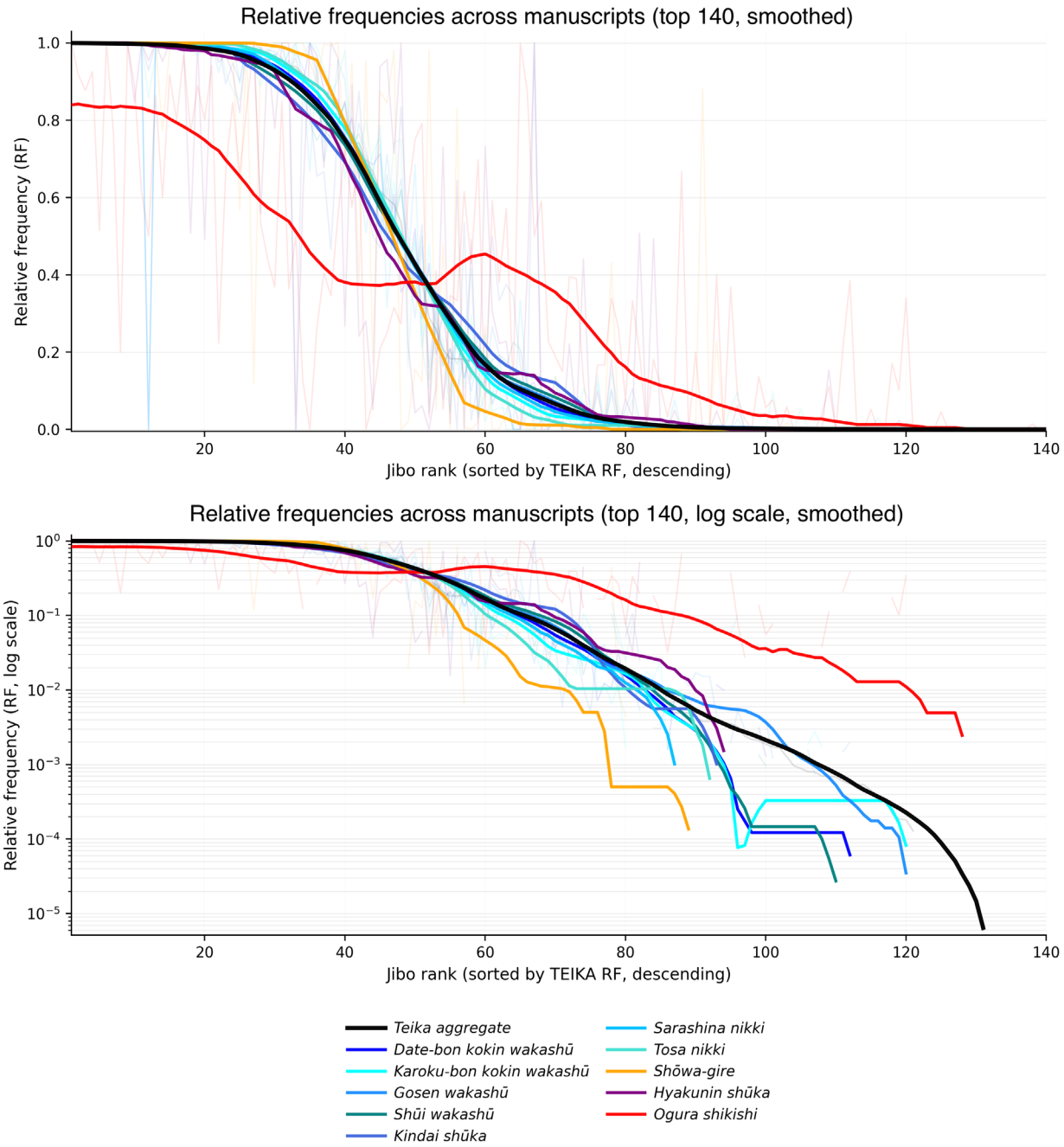

Fig. 14: Comparison of RFs among manuscripts considered in this study, in descending order of Teika's RFs. The top plot displays raw RFs and the bottom plot displays RFs after a log transformation.

Relative Frequencies of Fujiwara no Teika

Orig.	Sound	Jibo	TEIKA RF
1	A	安	0.990515
2	A	阿	0.009485
6	I	以	0.941398
7	I	伊	0.058602
12	U	宇	0.999038
15	U	右	0.000962
19	E	衣	0.990132
20	E	江	0.004934
22	E	盈	0.004934
26	O	於	1
27	KA	加	0.258827
28	KA	可	0.741173
45	KI	幾	0.995715
46	KI	起	0.001875
49	KI	木	0.00241
55	KU	久	0.995666
56	KU	具	0.004334
61	KE	計	0.176338
62	KE	个	0.6948
63	KE	遣	0.128862
67	KO	己	0.998957
68	KO	古	0.000695
74	KO	子	0.000348
75	SA	左	0.617699
76	SA	佐	0.381416
77	SA	散	0.000885
84	SI	之	0.816387
85	SI	志	0.183426
87	SI	新	0.000186
92	SU	寸	0.872351
93	SU	春	0.105954
95	SU	須	0.018163
96	SU	数	0.003532
98	SE	世	0.996296
99	SE	勢	0.003704
102	SO	曽	0.999199
104	SO	所	0.000801
107	TA	太	0.02322
108	TA	多	0.826625
109	TA	堂	0.150155
113	TI	知	0.955163
114	TI	地	0.044158
115	TI	千	0.000679
121	TU	川	0.817546
122	TU	徒	0.181755
123	TU	津	0.000699
126	TE	天	0.997561
127	TE	帝	0.002091
128	TE	亭	0.000348
135	TO	止	0.999403
136	TO	登	0.000597
147	NA	奈	0.980788
148	NA	那	0.018419
151	NA	名	0.000792
153	NI	仁	0.036413
154	NI	尓	0.959783
156	NI	耳	0.000181
157	NI	二	0.003623
161	NU	奴	1
166	NE	祢	0.96071
167	NE	年	0.003802
169	NE	根	0.003802
173	NE	子	0.031686
174	NO	乃	0.992022
175	NO	能	0.00785
179	NO	野	0.000129
180	HA	波	0.050845
181	HA	者	0.740797
182	HA	盤	0.00166
183	HA	八	0.206246
185	HA	葉	0.000453
190	HI	比	0.536014
191	HI	飛	0.174989
192	HI	日	0.288997
199	HU	不	0.582757
200	HU	布	0.414016
201	HU	婦	0.003227
203	HE	部	0.998907
204	HE	遍	0.000546
207	HE	辺	0.000546
214	HO	保	0.403141
215	HO	本	0.596859
220	MA	未	0.182977
221	MA	万	0.666363
223	MA	満	0.15066
230	MI	美	0.792463
231	MI	見	0.050955
232	MI	三	0.156051
235	MI	身	0.000531
237	MU	武	0.990446
238	MU	無	0.00828
240	MU	无	0.001274
244	ME	女	0.918463
245	ME	免	0.081537
250	MO	毛	1
257	YA	也	1
262	YU	由	0.9875
263	YU	遊	0.0125
268	YO	与	0.991311
271	YO	夜	0.002172
272	YO	世	0.006517
275	RA	良	1
280	RI	利	0.956999
282	RI	里	0.043001
286	RU	留	0.957219
287	RU	流	0.004902
288	RU	類	0.000223
289	RU	累	0.037656
290	RE	礼	0.994933
291	RE	連	0.005067
293	RO	呂	1
301	WA	和	0.979345
302	WA	王	0.020655
306	WI	為	0.93865
307	WI	井	0.06135
313	WE	恵	0.901478
314	WE	衛	0.098522
315	WO	遙	0.985572
316	WO	越	0.01271
317	WO	乎	0.001718
321	N	无	0.998476

Fig. 15: A list of RFs of *jibo* used by Fujiwara no Teika in selected manuscripts.

DOI: 10.15460/s8e93666

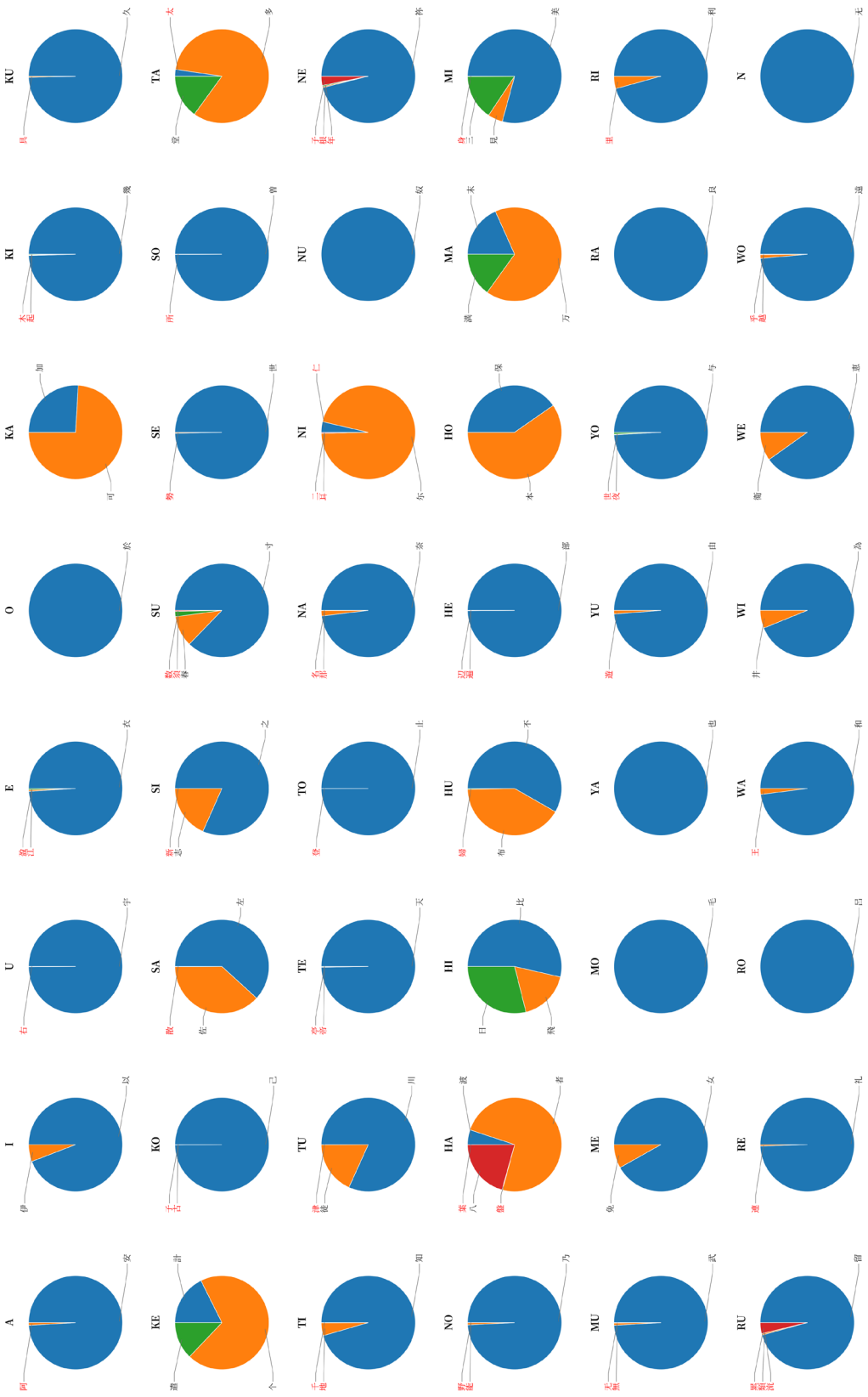

Fig. 16: A visualization of Teika's RFs categorized by sound.

Abbreviations

GOSE: *Gosen wakashū* 後撰和歌集

HNSK: *Hyakunin shūka* 百人秀歌

KIND: *Kindai shūka* 近代秀歌 (*jihitsu-bon* 自筆本, autograph manuscript)

KODA: *Kokin wakashū* 古今和歌集 (*Date-bon* 伊達本, manuscript formerly owned by the Date family)

KOKA: *Kokin wakashū* 古今和歌集 (*Karoku-bon* 嘉禄本, manuscript copied during the Karoku era, 1225–1227)

OGURA: Ogura *shikishi* 小倉色紙

SARA: *Sarashina nikki* 更級日記

SHOWA: *Shōwa-gire* 昭和切

SIHU: *Shūi wakashū* 拾遺和歌集

TOSA: *Tosa nikki* 土佐日記

Manuscripts

Dazaifu, Kyushu National Museum
Item no. B46

Nagoya, Tokugawa Art Museum
Item no. TAM000053
Item no. TAM7286

Tokyo, Kunaichō Shoryōbu Toshoryō Bunko 宮内庁書陵部図書寮文庫 (Imperial Household Agency, Archives and Mausolea Department)
Item no. 503: 236
Takatsukasa 138

Tokyo, Tokyo National Museum
Item no. B-3072
Item no. B-3537

References

Aitchison, John (1982), 'The Statistical Analysis of Compositional Data', *Journal of the Royal Statistical Society: Series B (Methodological)*, 44: 139–160.

—— (1983), 'Principal Component Analysis of Compositional Data', *Biometrika*, 70/1: 57–65.

Akisada Yayoi 秋定弥生 (2002), 'Sōgi to "Ogura shikishi": Asukai Eiga kara no jōto wo megutte' 宗祇と「小倉色紙」: 飛鳥井栄雅からの譲渡をめぐって, *Mukogawa kokubun* 武庫川国文, 60: 117–137.

Atkins, Paul. S. (2017), *Teika: The Life and Works of a Medieval Japanese Poet*, Honolulu: University of Hawai'i Press.

Azusazawa Kaname 梓澤要 (1998), *Hyakumai no Teika* 百枚の定家, Tokyo: Shin jinbutsu ōraisha.

Chen, Yen-Chi (2017), 'A Tutorial on Kernel Density Estimation and Recent Advances', *Biostatistics & Epidemiology*, 1/1: 161–187.

Egozcue, Juan José and Vera Pawlowsky-Glahn (2018), 'Modelling Compositional Data: The Sample Space Approach,' in B.S. Daya Sagar, Qiuming Cheng and Frits Agterberg (eds), *Handbook of Mathematical Geosciences: Fifty Years of IAMG*, Springer International Publishing, 81–103.

—— (2019), 'Compositional Data: The Sample Space and its Structure', *TEST*, 28: 599–638.

Furuya Minoru 古谷稔 (comm.) (1979), *Kamakura: Fujiwara no Teika: Kindai shūka (Jūyō Bunkazai)* 鎌倉: 藤原定家: 近代秀歌 (重要文化財) (Nihon meiseki sōkan 日本名跡叢刊, 33), Tokyo: Nigensha.

Greenacre, Michael (2021), 'Compositional Data Analysis', *Annual Review of Statistics and Its Application*, 8: 271–299.

Greenacre, Michael, Patrick J. F. Groenen, Trevor Hastie, Alfonso Iodice D'Enza, Angelos Markos and Elena Tuzhilina (2022), 'Principal Component Analysis', *Nature Reviews Methods Primers*, 2: 100, <https://doi.org/10.1038/s43586-022-00184-w>.

Gyobutsu Sarashina nikki 御物更級日記, (Kasama Eiin sōkan 笠間影印叢刊, 2), Tokyo: Kasama shoin, 1981.

Haruna Yoshishige 春名好重 (1969), 'Ogura shikishi' 小倉色紙, in *Kohitsu jiten* 古筆辞典, Tokyo: Sansaisha, 72–74.

Hibino Junzo 日比野純三 (1986), 'Ōsen Keisan-sen "Hyakunin isshu" ni tsuite oyobi "Kōhon Hyakunin isshu (kō)" tsuketari hairetsu ichiran' 横川景三撰『百人一首』について及び『校本百人一首（稿）』付排列一覧, in Yano Kan'ichi 矢野貫一 and Nagatomo Chiyoji 長友千代治 (eds), *Nihon bungaku setsurin* 日本文学説林, Osaka: Izumi shoin, 237–258.

Ieiri Hironori 家入博徳 (2010), *Chūsei shosharon: Shunzei, Teika no shosha to shakai* 中世書写論: 俊成・定家の書写と社会, Tokyo: Bensei shuppan.

Ii Haruki 伊井春樹 (1990), 'Hyakunin isshu no seiritsu' 百人一首の成立, in Ōtsubo Toshikinu 大坪利絹, Kamijō Shōji 上條彰次, Shimazu Tadao 島津忠夫 and Yoshikai Naoto 吉海直人 (eds), *Hyakunin isshu kenkyū shūsei* 百人一首研究集成, Osaka: Izumi shoin, 223–236.

Ikeda Kikan 池田亀鑑 (1941), *Koten no hihanteki shochi ni kansuru kenkyū* 古典の批判的処置に関する研究, 3 vols, Tokyo: Iwanami shoten.

Iura Miyuki 井浦美幸 (1971), 'Chūkinsei ni okeru hiragana no jitai: Fujiwara no Teika no hiragana' 中近世における平仮名の字体：藤原定家の平仮名, *Tōkyō Joshi Daigaku 'Nihon bungaku'* 東京女子大学『日本文學』, 37: 48–65.

Johnson, Richard A. and Dean W. Wichern (2007), *Applied Multivariate Statistical Analysis*, 6th edition, Upper Saddle River, NJ: Pearson Prentice Hall, 149–167.

Katagiri Yōichi 片桐洋一 (2003), 'Hyakunin isshu zatsudan: sono ato' 百人一首雑談：その後, in Ōtsubo Toshikinu 大坪利絹, Kamijō Shōji 上條彰次, Shimazu Tadao 島津忠夫, and Yoshikai Naoto 吉海直人 (eds), *Hyakunin isshu kenkyū shūsei* 百人一首研究集成, Osaka: Izumi shoin, 175–185.

Kodama Kōta 児玉幸多 (1981), *Kuzushiji yōrei jiten: fukyū-ban* くずし字用例辞典：普及版, Tokyo: Tōkyōdō shuppan.

Komatsu Shigemi 小松茂美 (1961), *Gosen wakashū: kōhon to kenkyū* 後撰和歌集：校本と研究, 2 vols, Tokyo: Seishin shobō.

—— (1970), *Nihon shoryū shi* 日本書流史, 2 vols, Tokyo: Kōdansha.

—— (1972), *Kohitsu* 古筆, Tokyo: Kōdansha.

Kornicki, Peter (2008), 'Ikeda Kikan and the Textual Tradition of the *Tosa Nikki*: European Influences on Japanese Textual Scholarship', *Revue d'histoire des textes*, 3: 263–282.

Kyūsojiin Hitaku 久曽神昇 (ed.) (1990), *Shūi wakashū: Fujiwara no Teika-hitsu* 拾遺和歌集：藤原定家筆, 2 vols, Tokyo: Kyūko shoin.

—— (ed.) (1991), *Fujiwara no Teika-hitsu Kokin wakashū* 藤原定家筆：古今和歌集, 2 vols, Tokyo: Kyūko shoin.

Mostow, Joshua S. (1996), *Pictures of the Heart: The 'Hyakunin Isshu' in Word and Image*, Honolulu: University of Hawai'i Press.

—— (2024), *Hyakunin'shu: Reading the 'Hundred Poets' in Late Edo Japan*, Honolulu: University of Hawai'i Press.

Nagoya Akira 名児耶明 (1987), 'Teika shikishi no juyō' 定家色紙の受容, in *Teika-yō* 定家様, Tokyo: Gotō bijutsukan, 94–108.

—— (1994), 'Teika-yō to Ogura shikishi' 定家様と小倉色紙, in 'Waka bungaku ronshū' henshū iinkai 『和歌文学論集』編集委員会 (ed.), *Hyakunin isshu to shūkasen* 百人一首と秀歌撰 (Waka bungaku ronshū 和歌文学論集, 9), Tokyo: Kazama shobō, 257–296.

—— and Kanechiku Nobuyuki 兼築信行 (2013), 'Taidan: Ogura shikishi to Hyakunin isshu: Fujiwara no Teika no nazo ni semaru' 対談：小倉色紙と百人一首：藤原定家の謎にせまる, *Yuriika* ユリイカ, 44/16: 137–151.

Okuda Isao 奥田勲 (1998), *Sōgi* 宗祇, Tokyo: Yoshikawa kōbunkan.

Pawlowsky-Glahn, Vera and Antonella Buccianti (eds) (2011), *Compositional Data Analysis: Theory and Applications*, Chichester: Wiley.

Reizei-ke Shiguretei Bunko 冷泉家時雨亭文庫 (ed.) (1994), *Kokin wakashū: Karoku 2-nen-bon, Kokin wakashu: Jōō 2-nen-bon* 古今和歌集：嘉禄二年本・古今和歌集：貞応二年本 (Reizei-ke Shiguretei bunko sōsho 冷泉家時雨亭文庫叢書, 2), Tokyo: Asahi shinbunsha.

—— (ed.) (1996), *Godai kan'yō, Teika kagaku* 五代簡要・定家歌学 (Reizei-ke Shiguretei sōsho 冷泉家時雨亭叢書, 37), Tokyo: Asahi shinbunsha.

—— (ed.) (2004), *Gosen wakashū: Tenpuku 2-nen-bon* 後撰和歌集:天福二年本, (Reizei-ke Shiguretei sōsho 冷泉家時雨亭叢書, 3), Tokyo: Asahi Shinbunsha.

—— (ed.) (2018), *Honkoku Meigetsuki* 翻刻明月記, vol. 3, Tokyo: Asahi Shinbunsha.

Sadeghi, Behnam, Hossein Molayemat, and Vera Pawlowsky-Glahn (2024), 'How to choose a proper representation of compositional data for mineral exploration?', *Journal of Geochemical Exploration* 259, <https://doi.org/10.1016/j.gexplo.2024.107425>.

Saitō Tetsuya 齊藤鉄也 (2016), 'Kana jibo no shutsugen hindoritsu ni motozuku Fujiwara no Teika shosha shiryō no nendai suitei' 仮名字母の出現頻度率に基づく藤原定家書写資料の年代推定, *Jinmonkon 2016 ronbunshū* じんもんこん 2016 論文集, 197–202.

—— (2018), 'Kana jibo no shutsugen keikō o mochiita Fujiwara no Teika shosha shiryō no chōsa' 仮名字母の出現傾向を用いた藤原定家書写資料の調査, *Jōhō shori gakkai ronbunshi* 情報処理学会論文誌, 59/2: 315–322.

Sakamoto Kiyoe 坂本清恵 (2020), 'Mikohidari-ke rekisei to kana yōhō: Shunzei, Teika, Tameie soshite Tamesuke' 御子左家歴世と仮名用法—俊成・定家・為家そして為相—, *Nihon bungaku kenkyū jyaanaru* 日本文学研究ジャーナル, 16: 21–36.

Sanjōnishi Saneyoshi 三条西実義 (ed.) (1931–44), *Sanetaka-kō ki* 実隆公記, 6 vols, self-published.

Sarashina nikki 更級日記 (Nihon meihitsu sen 日本名筆選, 43), Tokyo: Nigensha, 2004.

Shibuya Ei'ichi 渋谷栄一 (1998), *Fujiwara no Teika no chosaku to Heian-chō kotenseki no shosha kōkan ni kansuru sōgō deetabeesu* 藤原定家の著作と平安朝古典籍の書写校勘に関する総合データベース. Online database, <http://genjiemuseum.web.fc2.com/kenkyukai.html> (accessed on 13 August 2025).

Shōwa-gire 昭和切 (Kana meiseki shūsei 仮名名跡集成, 1), Tokyo: Kōbunsha, 1935.

Sugitani Jurō 杉谷寿郎 (1987), 'Ogura shikishi no shihai' 小倉色紙の紙背, *Gakusō* 学叢, 42: 1–2.

Sugiura Taeko 杉浦妙子 (2002), 'Shikishi no teisai no henka ni tsuite no ikkōsatsu' 色紙の体裁の変化についての一考察, *Shogaku shodōshi kenkyū* 書学書道史研究, 12: 33–43.

Tabuchi Kumiko 田渕句美子 (2020), 'Ogura shikishi to "Saga Chūin shōji shikishigata": shihai to seiritsu wo chūshin ni' 小倉色紙と「嵯峨中院障子色紙形」:紙背と成立を中心に, *Kagami*, 50: 36–59.

Takei Kazuto 武井和人 (2013), '"Shin Hyakunin isshu" seiritsu kō, zokuchō'『新百人一首』成立攷・続貂, *Kokugakuin zasshi* 國學院雑誌, 114/11: 540–554.

Teika-bon Tosa Nikki 定家本土佐日記 (1928), 2 vols (Sonkeikaku sōkan 尊経閣叢刊), Tokyo: Ikutoku Zaidan, <https://dl.ndl.go.jp/pid/1191112> (accessed on 13 August 2025).

Toyoda Naoko 豊田尚子 (1992), 'Fujiwara no Teika jihitsu no hiragana-bun ni okeru kana no yōhō ni tsuite' 藤原定家自筆の平仮名文における仮名の用法について, *Kokubungaku kō* 国文学攷, 136: 14–24.

Ue Kiyoko 植喜代子 (1979), 'Fujiwara no Teika no hentaigana no yōhō ni tsuite' 藤原定家の変体仮名の用法について, *Kokubungaku kō* 国文学攷, 82: 1–24.

Yoshida Kōichi 吉田幸一 (1999), *Hyakunin isshu: Tameie-bon, Son'en shinnō-bon kō* 百人一首：為家本・尊円親王本考, Tokyo: Kasama shoin.

Picture Credits

Fig. 1: © ColBase database <https://colbase.nich.go.jp/collection_items/kyuhaku/B46> (accessed on 24 March 2026).

Fig. 2: © ColBase database <https://colbase.nich.go.jp/collection_items/tnm/B-3072> (accessed on 24 March 2026).

Fig. 3: © ColBase database <https://colbase.nich.go.jp/collection_items/tnm/B-3537> (accessed on 24 March 2026).

Figs 4–9: © Michael R. Zeng

Fig. 10: © National Diet Library (Japan) <https://dl.ndl.go.jp/pid/1881097/1/3> (accessed on 31 December 2025).

Figs 11–12: © The Tokugawa Art Museum Image Archives/DNPartcom.

Figs 13–16: © Michael R. Zeng