古今和歌集の現代日本語訳における追加率要素の分析

Xudong Chen

Bor Hodošček

Hilofumi Yamamoto

2024-09-01

Abstract

Previous research has explored the visualization of non-literal information in the Kokinshū’s poetic words (歌語) through the addition of elements in translations found in annotation books. However, detailed analysis of the contemporary Japanese translations in these annotations and the additional elements within them has not been conducted. This paper aims to provide a comprehensive analysis of whether the ten contemporary Japanese translations of the Kokinshū used in earlier studies are suitable for visualizing non-literal elements. First, the paper reviews the contemporary Japanese translations used in the annotations of the Kokinshū and examines the translators’ approaches to their work. It then calculates the unmatch rate—the proportion of poetic words without direct equivalents—and the addition rate—the proportion of additional elements in the translations—to clarify how elements were added and matched. Lastly, specific examples are provided, along with a detailed case analysis of the additional elements in these translations. The analysis reveals that translators’ subjective approaches can be categorized into four types: intention-focused, text-focused, reader-focused, and unclear. However, from both quantitative and statistical perspectives, despite these differing approaches, the translations include a similar amount of additional elements. In case analysis, it was confirmed that at the sentence level, the addition of elements relied on co-occurrence tendencies at the corpus level, with many supplementary elements added to refine the directional meaning of words. These findings demonstrate that contemporary Japanese translations can be effectively used to visualize non-literal aspects of the original poetic words.

Table of contents

|  |
| --- |
| Note |
|  |

# 1. Introduction

## 1.1 Overview of purpose and findings

This paper aims to demonstrate, from two perspectives—the calculation of the unmatch rate between the original poem and its translation, and the analysis of additional elements in the translation—that contemporary Japanese translations of classic poetry can serve as interpretive tools to supplement and explain non-literal information in poetic words, regardless of the translator’s intent, even when focused on literal or word-for-word translation.

The main contributions of this study are as follows:

1. As source material, the paper presents the first publicly available lexical data of contemporary Japanese translations of the *Kokinshū* (古今和歌集) from 金子 (1933), whose copyright has expired. The paper provides a historical overview of colloquial translations of the Kokinshū, focusing particularly on the translation approaches used in annotation books in the 20th century.
2. By calculating the unmatch rate at the word level between the original poems and the contemporary Japanese translations, as well as the addition rate in the translations, the study demonstrates that the addition of information in contemporary translations remains consistent across different translation approaches.
3. Specific examples of contemporary translations confirm whether co-occurring words at the corpus level and their collective semantic preferences are reflected at the sentence level in each translation. These findings show that applying contemporary Japanese translations to the visualization of non-literal aspects of the original text is a plausible approach.

## 1.2 Background

### 1.2.1 Non-literal elements in Classical Japanese Poetry

As described in the preface of the *Kokinshū*, poets often entrust their inner feelings (“what the heart thinks” 心に思ふもの) to natural phenomena. However, understanding the literal meaning of a poem does not necessarily grant direct access to these deeper emotions. For instance, consider the following love poem from the *Kokinshū*:

初雁の鳴きこそわたれ世の中の人の心の秋しうければ (Kokinshū, Love-5, by Tsurayuki)

Roman: hatsukarino nakikoso watare yononakano hitono kokorono akishi ukereba  
  
Morph: hatsu+kari=no nak-i=koso watar-e yo+no+naka=no hito=no kokoro=no aki=si u-kere=ba  
  
Gross: first+geese=GEN cry-THM=EMP cross-THM world+of+inside=GEN peaple=GEN heart=GEN autumn=EMP sorrow-THM-REASON

If we grasp the literal meaning of each word, we may only interpret it as a straightforward description[[1]](#footnote-24). However, by consulting a poetic word dictionary on the word “autumn” (aki 秋) (片桐 1983, 3), we realize that there is more non-literal information beyond the surface meaning.

秋【あき】 […]「秋」と「飽」を掛け、過ぎ去ってゆく秋と過ぎ去ってゆく愛を惜しむことが多かった。【脚注】[雁の]「鳴く」を人が「泣く」と同列にしか把握しない […]

Authors’ translation: […] “Aki” is often associated with *akiru* “to tire (あきる),” and the passing of autumn is frequently used to lament the fading of love. [In footnote] The cry of the wild geese may be interpreted as a metaphor for human weeping, linking it to the idea of loss. zFollowing this interpretation, we infer that the poem speaks of a lover’s heart “tiring” (aki 秋/飽き) and the poet “weeping” (naku 鳴く/泣く) over the loss of love. The association of autumn with the end of love is implicit, or non-literal, and cannot be fully understood from the literal meaning of the words alone.

### 1.2.2 Visualizing non-literal elements in poetic words based on additional elements in contemporary Japanese translations

The development of a system for visualizing non-literal information based on ten contemporary Japanese translations of the Kokinshū has been a primary focus of the authors’ previous studies (Yamamoto 2005, 2006; 陳, ホドシチェク, and 山元 2022; Chen, Yamamoto, and Hodošček 2024). This research, grounded in a communication model (Schramm 1954), argues that the additional elements identified through the comparison between classical Japanese and contemporary translations carry non-literal information embedded within the poetic words.

The process of relativizing the original text with the contemporary translation involves subtracting the elements present in the original from those found in the translation, thereby highlighting the elements unique to the translation. Scholars such as 近藤 and 近藤 (2001) and 近藤 (2011) refer to this method as “subtraction” (sabun 差分). However, it is crucial to examine the nature of these subtracted elements—specifically, the additional elements in the translation—and determine whether they can be classified as non-literal information.

Regarding the addition of elements in translation, Koller ([2004] 1979, 249) describes these modifications as “interventions” (Eingriff) by the translator. Such “harmless” interventions, as Koller explains, compensate for the target audience’s lack of background knowledge, denotative and connotative nuances, and for any linguistic, socio-cultural, or intertextual references lost in translation (Koller [2004] 1979, 249). Therefore, translations, much like annotations or dictionaries, can be valuable sources for extracting non-literal information.

However, additions in translation can also stem from overestimating or underestimating the reader’s level of comprehension (Eugene Albert Nida 1964, 155; Koller [2004] 1979, 249–50). It is important not to overlook the variations in the additional elements introduced by different translators. While Chen, Yamamoto, and Hodošček (2024) touched on the effectiveness of contemporary translations as interpretative tools for the Kokinshū, the study did not fully address the variability in the added elements. There remains significant scope to re-evaluate the ten translations used, taking into account the translators’ individual backgrounds and perspectives.

This paper will analyze the historical context of these ten contemporary translations, focusing on both the quantitative and qualitative aspects of the additional elements they contain. As mentioned earlier, the additions in translation may result from different degrees of overestimation or underestimation of the reader’s comprehension, leading each translator to disclose or emphasize varying amounts of information to the contemporary audience. Each translator follows their own approach in determining how much of the original poem’s information should be accessible to modern readers. By organizing and examining these varying approaches, we can objectively assess whether the translators adhere to their intended strategies, providing insight into the broader challenges and limitations of translation as a tool for visualizing non-literal information.

Finally, by building on these findings, we will refine the theoretical basis of the visualization system for non-literal elements and propose directions for its improvement.

## 1.3 Research questions and structure

As a foundation for visualizing non-literal elements, this paper re-evaluates the topic from the following four perspectives:

1. Investigation of the background of contemporary Japanese translations in annotation books and the classification of translation strategies: How can we categorize the translators’ approaches?
2. Demonstration that the elements of the original text are minimally altered in contemporary translations: To what extent do elements from the original text lack corresponding counterparts in the translations?
3. Demonstration that additional elements in contemporary translations are consistently present, regardless of translation strategy: How many additional elements are introduced, and are there statistical differences across different translation approaches?
4. Evidence that additional elements in contemporary translations embed non-literal information: Are there instances where corpus-level and collective co-occurrence data are reflected in the added elements of the translations?

The paper is structured as follows:

[Section 2](#sec-materials) offers an overview of the ten contemporary Japanese translations used as source material, alongside the historical background of colloquial translations of classical poetry in the 20th century. [Section 3](#sec-methods) details the methodology for addressing each of the four research questions. [Section 4](#sec-results) presents the findings derived from this methodology, and [Section 6](#sec-conclusions) discusses the results and the potential challenges they pose. Finally, the paper concludes with a summary of the overall findings.

# 2. The Kokinshū, annotation books, and contemporary translations: data and background

In this section, we present an overview of the *Kokinshū*, along with a historical survey of its annotation books and the contemporary translations found within them.

## 2.1 The *Kokinshū*

As the terms “songs” (*uta* 歌) and “to recite” (*yomu* 詠む) suggest, classical Japanese poetry was originally performed aloud in the imperial court, deeply rooted in the spoken language of the time. The *Kokinshū*—as its title[[2]](#footnote-30) implies—brings together both ancient and contemporary poems from its era. The ancient poems mainly refer to those included in the *Man’yōshū* (万葉集), which contains works from the 7th and 8th centuries by notable poets in Japanese literary history, such as Yamabe no Akahito (山部赤人), Kakinomoto no Hitomaro (柿本人麻呂), and Nukata no Ōkimi (額田王).

The *Kokinshū* holds a prominent place as the first of the twenty-one imperial anthologies (chokusen wakashū 勅撰和歌集), setting the foundational structure for all subsequent collections. Its preface, written by the editor Ki no Tsurayuki (紀貫之), is regarded as a cornerstone of Japanese poetic theory. As a crucial resource for the study of classical poetry, the *Kokinshū* became a source of inspiration for later literary works that employed “honkadori” (本歌取り)[[3]](#footnote-31), such as *The Tale of Genji* (Genji monotatari 源氏物語), *Tosa Diary* (Tosa nikki 土佐日記), *The Tale of Ise* (Ise monogatari 伊勢物語), and many other key texts in Japanese literature. Additionally, its connection to the spoken language of its time makes it an invaluable source for understanding the historical evolution of Japanese.

## 2.2 Annotation books and their contemporary Japanese translations

Throughout the history of Japanese classical literature, the *Kokinshū* has inspired a vast number of annotation books, which have developed a significant scholarly tradition and history of their own (窪田 1960, 上中下:319). Many of these annotation books include contemporary translations intended to help readers interpret the original poems, with the translators primarily focused on conveying their understanding of the poetry. Although some translations from the early modern period may not fully align with contemporary Japanese, they remain important and cannot be disregarded. Therefore, this section provides a brief overview of annotation books up to the early modern period and examines the contemporary translations included in annotation books.

Before the modern period (1868), many annotation books on the *Kokinshū* were published. More than 70 annotation books were produced between 1600 and 1868 (小島 and 新井 1989, 447–50). [Table 1](#tbl-annotation) lists five of the most notable annotation books from this period.

Kitamura Kigin’s *Hachidaishūshō* (北村季吟『八代集抄』) is a comprehensive annotation covering 108 volumes (50 books) on the eight imperial anthologies[[4]](#footnote-33). Keichū’s (1640–1701) *Kokin Yozaishō* (契沖『古今余材抄』) is recognized as the first practical research book on the *Kokinshū* (小沢 1971, 36). Kamo no Mabuchi (1697–1769) expanded upon Keichū’s work with his own *Kokinshū Uchigiki* (賀茂真淵『古今和歌集打聴』), although his primary focus was on earlier poems, such as those in the *Man’yōshū*. Motoori Norinaga’s *Kokin Tōkagami* (本居宣長『古今集遠鏡』) attempted to translate all of the poems into the colloquial language of his time[[5]](#footnote-34). Kagawa Kageki (1768–1843), in his *Kokinshū Seigi* (香川景樹『古今和歌集正義』), offered strong critiques of earlier annotation books (小沢 1971, 36; 松田 1968, 上下:58).

After 1868, over 30 annotation books on the *Kokinshū*—including reprints and revisions—were published. It is evident that many of the annotation books published after 1868 were influenced, to varying degrees, by these earlier works.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table 1: Representative Annotation Books of the *Kokinshū* Before the Modern Period (1600–1868)   | Author | Year | Annotation book title | | --- | --- | --- | | Kitamura Kigin 北村季吟 | 1682 | *Hachidaishūshō* 八代集抄 | | Keichū 契沖 | 1692 | *Kokin Yozaishō* 古今余材抄 | | Kamo no Mabuchi 賀茂真淵 | 1784 | *Kokinshū Uchigiki* 古今和歌集打聴 | | Motoori Norinaga 本居宣長 | 1793? | *Kokin Tōkagami* 古今集遠鏡 | | Kagawa Kageki 香川景樹 | 1832 | *Kokinshū Seigi* 古今和歌集正義 | |

## 2.3 Summary of data: Contemporary translations in 20th century annotation books

In the 21st century, many annotation books with contemporary translations continue to be published. However, the data used in previous research on visualization systems primarily includes contemporary translations of the *Kokinshū* from the modern era onwards [Table 2](#tbl-CT-data), which form the core focus and dataset for this study.

The data format adheres to the space-delimited format introduced by Hodošček and Yamamoto (2022). Each token is linked to an identifier based on the old-version Word List of Semantic Principles (WLSP) (Nakano et al. 1994). A key feature of the dataset is its ability to assign multiple lexicon numbers to polysemous words and provide detailed decomposition for compound words, supporting more flexible and nuanced analysis.

However, due to copyright constraints on these contemporary translations—covering translation rights, editorial rights, and reprint rights—the dataset cannot be fully disclosed. For this paper, we are able to release only the data from 金子 (1933), which is now in the public domain.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table 2: 10 Contemporary Translations of the *Kokinshū* from the 20th Century: The token and type counts do not account for the decomposition of compound expressions.   |  | Abbr. | Reference | Manuscript | Token count | Type count | Document count | | --- | --- | --- | --- | --- | --- | --- | | 1 | KNK | 金子 (1933) | Teika | 42,439 | 3,356 | 1,000 | | 2 | KBT | 窪田 (1960) | Teika | 32,210 | 2,701 | 1,000 | | 3 | MTD | 松田 (1968) | Teika | 31,860 | 3,007 | 1,000 | | 4 | OZW | 小沢 (1971) | Teika | 36,173 | 3,384 | 1,000 | | 5 | TKOK | 竹岡 (1976) | Teika | 29,844 | 2,861 | 1,000 | | 6 | OKMR | 奥村 (1978) | Teika | 32,321 | 3,153 | 1,000 | | 7 | KSJ | 久曽神 (1979) | Teika | 34,050 | 2,770 | 1,000 | | 8 | KMCY | 小町谷 (1982) | Teika | 30,869 | 2,692 | 1,000 | | 9 | K&A | 小島 and 新井 (1989) | Teika | 33,867 | 2,955 | 1,000 | | 10 | KTGR | 片桐 (1998) | Teika | 36,362 | 2,882 | 1,000 | |  | Total |  |  | 339,995 | 8,252 | 10,000 | |

# 3. Methods

To assess whether non-literal elements can be efficiently extracted from contemporary translations, this study adopts a two-step approach. First, we clarify the translators’ subjective awareness of their translation approaches and examine how they handle non-literal information in their translations. Second, we investigate whether their actual translation practices align with their stated intentions.

To explore the translators’ subjective awareness, we conduct a review of their stated translation policies as found in annotation books. We categorize the approaches taken by each translator.

For the analysis of translation practices, we calculate the unmatch rate (the proportion of original text elements that are absent in the translation) and the addition rate (the proportion of added elements in the translation) to assess whether these additions are influenced by the translators’ stated intentions. We also examine whether there are statistically significant differences based on their translation strategies. Additionally, we conduct case studies to determine whether the added elements in the translations can reasonably be considered non-literal information.

## 3.1 Categorization of translation approaches

This section organizes and reviews the subjective awareness of translation approaches among the ten authors of *Kokinshū* annotation books discussed in the previous section. In Yamamoto (2005, 102), the author analyzes each contemporary translation based on the translators’ theoretical perspectives and classifies them into categories such as “word-for-word,” “intention-oriented,” “word-modified,” and “not specified.” However, while “intention-oriented” reflects the overall translation objective, the other categories focus on specific strategies, leading to inconsistencies in classification criteria. To address this, a unified framework for classification is necessary. Here, we categorize the translation strategies using the communication model by Schramm (1954), as employed in Yamamoto (2005), and investigate how translators approach paraphrasing and the addition of elements ([Figure 1](#fig-schramm-schema)).

The communication model for contemporary translations includes two sub-processes. The translator acts as both the recipient in sub-process 1 and the sender in sub-process 2, connecting the two fields of experience in communication. Based on this model, translation approaches can be grouped into the following categories:

1. **Poet-focused**: Emphasizing the poet’s intent, which highlights the source in the 10th-century communication sub-process.
2. **Text-focused**: Prioritizing the literal meaning of the text, treating it as the signal in the 10th-century communication sub-process.
3. **Reader-focused**: Focusing on readability for the modern audience, which highlights the destination in the 20th-century communication sub-process.

Another important aspect involves the literary quality of the translation itself, focusing on the signal within the 20th-century context. However, since these translations are intended for annotations, focusing on literary style or entertainment value is less likely. For this reason, such possibilities are excluded from this analysis.

|  |  |  |
| --- | --- | --- |
| |  | | --- | | (a) Basic communication model | |  |

|  |  |
| --- | --- |
| |  | | --- | | (b) Potential categories of translation approaches | |

Figure 1: Classification of translation approaches based on the communication model

## 3.2 Calculating Addition Rate and Unmatch Rate

When extracting non-literal elements from contemporary translations, it may seem that translations focused on enhancing reader comprehension allow for the extraction of more non-literal elements. However, discrepancies often arise between the translator’s stated intent and their actual translation practice. To address this, we objectively examine the addition rate to clarify this issue. The unmatch rate between the original poetic words and their translations will also be calculated to determine how consistently each translator adheres to their stated translation approaches.

We offer two methods for calculating the unmatch rate and addition rate: the “bag method,” which conducts set operations without considering the order or repetition of elements, and the “alignment method,” which accounts for both the order and repetition of elements. Although these procedures are described in Yamamoto (2005) and Yamamoto and Hodošček (2019), some descriptive ambiguities remain. This paper aims to clarify the calculation process, addressing the unclear aspects from previous studies.

The approach outlined in Yamamoto (2005) and Yamamoto and Hodošček (2019) calculates the unmatch rate while preserving the polysemy of words and the decomposability of compound words. This means that when agreement exists for the default sense and compound words, the calculation excludes the compounds’ decompositions or alternative senses. Conversely, when agreement does not exist for the default sense and compound words, the calculation considers decompositions and other potential sense identifiers. Thus, when counting agreements within a poem, the target unit—whether it is the compound word as a whole or a specific meaning of a word—is determined dynamically. The strength of this approach lies in its ability to account for as many potential agreements as possible during intermediate stages of the calculation.

However, this approach operates on a strong premise during the calculation. In Yamamoto (2005) and Yamamoto and Hodošček (2019), the assumption is that 100% of the information from the original is captured in the translation[[6]](#footnote-50). Even when no direct agreement exists in the translation, each element in the poem is assumed to correspond to some latent element in the translation, which is excluded from the addition rate calculation (the remaining portion is termed “pure addition”). However, when it is clear that a given poetic word has no agreement in translation, a weaker assumption acknowledges that some information may be missing.

A further issue arises from the use of compound words by certain translators, which influences the denominator when calculating the number of words in the translation (i.e., the basis for the addition rate), making direct comparisons between individual translators’ addition rates less intuitive.

This paper provides further clarification on the calculation and reproducibility of the addition rate. The adjustments made during the calculations do not alter the original conclusions. All calculations in this paper follow a unified approach, utilizing default identifiers and decomposition units for consistency. [[7]](#footnote-51)

#### Layering of agreement based on WLSP code

The agreement (“match/unmatch”) between poetic words and their contemporary translations is categorized into four layers— (Unmatch), (Group Match), (Field Match), and (Exact Match)—determined by the length of the Longest Common Subsequence (LCS) (Sankoff 1972; Traum and Habash 2000) between the two words’ metacodes (old-version WLSP codes), as outlined in the following formula.

In this formula, represents the metacode of a word in the source text (the poem), and represents the metacode of a word in the target text (the contemporary translation). refers to the LCS between the two metacodes and . Below, we provide examples for each classification.

1. **Unmatch** (; ):

* [BG-0]1-5520-20-0401 梅 (plum)
* [BG-0]8-0061-07-010-A の (of, genetive case)

1. **Group match** (; ):

* [BG-01-5520-]20-0401 梅 (plum)
* [BG-01-5520-]19-115-A 秋萩 (autumn)

1. **Field match** (; ):

* [BG-01-2030-01-]0300 神 (god)
* [BG-01-2030-01-]030-A 仏 (Buddha)

1. **Exact match** (; ):

* [BG-01-5520-20-040]1 梅 (plum)
* [BG-01-5520-20-040]-A 梅 (plum)

#### Calculation of addition rate using the bag method

We calculate the addition rate, which represents the proportion of additional elements in the contemporary translation relative to the original poem, using the bag method. In this approach, we treat both texts as sets and base the calculation on the number of agreements without considering their order. Given the source text (poem) and the target text (contemporary translation) , we calculate the addition rate for the target text as follows:

Here, represents the number of tokens in the source text that have non-repetitive agreements with tokens in the target text. The term gives the agreement rate (bag method) for the translation, where the denominator is the total number of tokens in the target text. We use to denote the total number of tokens in the target text, and , to represent the th and th tokens in the source and target texts, respectively. We calculate the addition rate by subtracting the agreement rate from 1. The indicator function returns 1 if the input condition is true (i.e., if has an agreement in the translation), and 0 if it is false (i.e., if has no agreement in the translation). [[8]](#footnote-53)

One key point is that the numerator of the agreement rate reflects the number of agreements in the source text, not the number of agreements in the translation. This approach allows us to recognize one-to-many translations in the target text, where part of the translation corresponds to elements in the poem, while the rest is considered additional. For example, translating “chiru” 散る (scatter) as “chirimidareru” 散り乱れる (scatter wildly) raises the question of whether this should be treated as a one-to-many translation. To ensure consistency, we assume a one-to-one translation as the default for simplex units. [[9]](#footnote-54)

We calculate the unmatch rate similarly, as the proportion of elements in the source text that have no agreement in the translation, using the following formula:

We use the addition rate calculated by the bag method for statistical analysis to examine the influence of the translator’s subjective awareness of their translation approach, as detailed in the previous section. We review the unmatch rate using basic descriptive statistics.

#### Calculation of addition rate using the alignment method

We now explain the calculation method using the alignment approach. Unlike the bag method, which ignores word order, the alignment method estimates aligned word pairs and calculates the number of agreements within the alignment, treating the poem and its translation as sequences rather than sets.

In this paper, we perform the alignment using dynamic programming. The score function for calculating the alignment between two sequences is defined as follows:

$$
\text{S}\_{\text{DP}}(i, j) =
\begin{cases}
0 & \text{if } i = 0 \text{ or } j = 0 \\
\max
\begin{cases}
\text{S}\_{\text{DP}}(i-1, j) - \text{gap} \\
\text{S}\_{\text{DP}}(i-1, j-1) + \text{weight}(s\_i, t\_j) \\
\text{S}\_{\text{DP}}(i, j-1) - \text{gap}
\end{cases} & i > 0 \text{ and } j > 0
\end{cases}\\
$$

The matching weight function is based on the LCS value of the two word sense identifiers, i.e., the WLSP codes. For an unmatch, we apply a penalty of :

Using the subsequences of the source text and the target text , we recursively calculate the optimal alignment score for positions and . The gap penalty (for inserting gaps) is set to 0.01, which allows for distant word alignments, a common occurrence in translations of poetry. [[10]](#footnote-56) Based on this formula, we search for the most likely alignment. An example of the aligned result is shown below:

>| Tatsutahime give god exist Reson emp autumn tree leaf ?? scatter inference  
>| 立田姫ー手向けるーーーー神のあれ　ばこそーーーーーーー秋の木の葉の幣ーーーと散るーーーらめー [298]  
>| 竜田姫は手向けをするべき神があるのでーーそのつかさどる秋の木の葉が幣のように散るのであろうよ [KBT]  
  
>| Tatsuta+hime.....tamukeru............kami=no.are..=ba=koso.................aki=no.ki=no.ha=no.nusa=to.......chiru.....=rame [298]  
>| Tatsuta+hime=ha..tamuke=wo.suru=beki.kami=ga.aru=node.....sono.tsukasadoru.aki=no.ki=no.ha=ga.nusa=no.yo=ni.chiru=no=dearou [KBT]

With such alignment, we can calculate the agreement and addition rates. Given the aligned sequences and [[11]](#footnote-57), the agreement rate using the alignment method is calculated as follows:

In this formula, any term with a prime () refers to the aligned text. The indicator function returns 1 if the alignment is an agreement and 0 if it is not. Alignments with gaps are excluded from the calculation. For the calculation of the agreement, unmatch, and addition rates, we use the total number of tokens in the source text (before alignment) as the denominator.

Dynamic programming is sensitive to issues such as word/chunk order changes (site swaps) in translations of poems, where phrases and word order are often rearranged. Additionally, the alignment method operates on a strict assumption that the aligned word pairs are exact matches, which may not hold in many cases. For this reason, we do not use the addition rate calculated by the alignment method for general statistical analysis but instead present it as part of case study explanations.

## 3.3 Statistical modeling of the relationship between translation approach and the addition rate

Previous studies (Yamamoto 2005; Yamamoto and Hodošček 2019) have calculated various metrics such as agreement rates (paraphrase rates) at the group, field, and synonym levels, as well as theoretical and practical addition rates. However, they only provided the mean and standard deviation of the addition rate by translator. Given the updates to the data and calculation procedures, the results may differ. Therefore, we conduct a new statistical analysis using the recalculated addition rates.

To assess whether the translators’ intended approaches influence the addition rate in contemporary translations, we perform statistical modeling. Specifically, we examine the effect of translation approach on the addition rate[[12]](#footnote-60). We employ a hierarchical model based on a beta distribution[[13]](#footnote-61) to analyze this effect. Translator and poem are included as random effects to account for variability across these factors. Details of the statistical model are available in the appendix.

We estimate the model’s coefficients using the Markov Chain Monte Carlo (MCMC) method, running 4 chains with 2000 iterations each, with the first 1000 iterations used as warm-up (burn-in). The model is implemented using the brms package in R (R version 4.3.3 (2024-02-29)) [@???], and we evaluate the convergence of the posterior distribution using the R-hat statistic and effective sample size (ESS). The results will reveal how translation focus influences the addition rate, and we will report the posterior median along with the 95% highest posterior credible interval (HDI).

## 3.4 Case Study of the “Tatsuta” Poem, No. 298

While we know that contemporary translations of classical Japanese poetry often involve paraphrasing and adding elements, the precise nature of these additions and paraphrases remains unclear. To clarify the characteristics of these elements, we analyze the alignments and examine them from a corpus linguistics perspective. Drawing on Sinclair’s Extended Unit of Meaning model (J. M. Sinclair 1996), this paper explores how the tendencies observed at the corpus level of the original text are reflected in the sentence-level translations.

The extended unit of meaning model describes linguistic units across four levels:

1. **Collocation**: The tendency of words to co-occur with other words (J. M. Sinclair, Jones, and Daley 1970, 15).
2. **Colligation**: The tendency of words to co-occur with specific syntactic patterns or grammatical elements (cf., Firth and Palmer 1968, 183; J. M. Sinclair 1996, 11; J. Sinclair 2003, 171).
3. **Semantic preference**: The tendency of words to co-occur with particular semantic sets (J. Sinclair 2003, 178).
4. **Discourse prosody**: The specific attitudinal, emotional, or evaluative tendency expressed by the extended unit of meaning (J. M. Sinclair 2004, 174).

From the level of collocation to that of discourse prosody, the elements become less explicitly observable and more implicit (Stubbs 2001, 87–88).

Given the nature of classical poetry and the limited data available, capturing extended units of meaning with certainty is often challenging. Therefore, we treat the four levels of the extended unit as distinct perspectives. Through these perspectives—co-occurring words, syntactic patterns, and semantic sets—we analyze how the ten different translations handle the tendencies present in the original text.

Yamamoto (2005) uses poem #298 as an example to illustrate how the word “Tatsuta” is processed in contemporary translations. We analyze the same poem, focusing on how it is handled in the ten translations, each employing different translation approaches, based on the three perspectives mentioned above.

At the level of collocation, we observe that in the *Hachidaishū* dataset, the following content words[[14]](#footnote-63) frequently appear in the context of “Tatsuta”: 山 (*yama*, “mountain”) (30 as collocates, 15 as a decomposition of 立田山 *tatsuta-yama*, “Mt. Tatsuta”), 紅葉づ (*momidzu*, “(leaves) turning red”) (17), 川 (*kawa*, “river”) (15 as a decomposition of 立田川 *Tatsuta-gawa*, “River Tatsuta”), 秋 (*aki*, “autumn”) (15), 見る (*miru*, “to see”) (12), 葉 (*ha*, “leaves”) (9), and 錦 (*nishiki*, “brocade”) (9). Words related to gods, such as 神無備 (*Kamnabi*, “Kamnabi (proper noun)”) (5) and 姫 (*hime*, “goddess/princess”) (8 as a decomposition of 立田姫 *Tatsuta-hime*, “Goddess Tatsuta”), also frequently appear, along with words associated with falling leaves, such as 散る (*chiru*, “scattering”) (8), 流る (*nagaru*, “flowing”) (5), 吹く (*fuku*, “blowing”) (6), 紅葉 (*momidzi*, “red leaves”) (5), and 黄葉 (*momidzi*, “yellow leaves”) (5).[[15]](#footnote-64)

Regarding co-occurring syntactic patterns, we find that poems ending in the conclusive form of a verb are rare (only one in the *Kokinshū* out of twelve poems). Additionally, nine poems in the *Kokinshū* utilize the emphatic grammatical structure “kakarimusubi” (係り結び), which emphasizes a lingering resonance at the poem’s conclusion. We will examine how these distinctive syntactic patterns are handled in the translations.

As for the semantic properties of the co-occurring words, the collocates of content words typically fall into two main groups: those related to place names (e.g., 立田川 *tatsuta-gawa*, “River Tatsuta”, 立田の山 *tatsuta-no-yama*, “Mountain of Tatsuta”) and those associated with the autumn goddess. These suggest that the poetic word “Tatsuta” is closely tied to a sacred location known for its autumnal colors.

Finally, we examine how the corpus-level and collective elements described above are incorporated into the sentence-level translations of poem #298, highlighting the variations in added elements across the ten translations.

# 4. Results

## 4.1 Classification of contemporary translation approaches

Based on the prefaces and introductions in the annotation books, where the translators express their awareness, understanding, and intent regarding translation approaches, the approaches of the ten translators are classified into three categories according to their focus in the communication process:

1. **Focus on the text’s literal meaning (signal)**: 金子 (1933); 窪田 (1960); 片桐 (1998)
2. **Focus on the author’s intent (source)**: 奥村 (1978); 竹岡 (1976)
3. **Focus on the reader’s comprehension (destination)**: 小沢 (1971); 久曽神 (1979)

The translation approach was not clearly stated in the works of 松田 (1968), 小島 and 新井 (1989), and 小町谷 (1982).

Among these, the translators who focus on the literal meaning of the original text tend to adhere most closely to word-for-word translation. On the other hand, those who prioritize the author’s intent or the reader’s understanding tend to allow for some reordering of words and the supplementation of expressions.

### Text-focused approach: KNK, KBT, KTGR

The group of translations that emphasize a literal interpretation of the original text focuses on translating as faithfully as possible, minimizing changes in word order or meaning. 金子 (1933) states that he follows a strict word-for-word approach, avoiding the addition or omission of words, and carefully translating to maintain the integrity of the original. He states that, fearing to distort the meaning or rhythm of the poems, he adhered to the *kyōkasuigetsu* (鏡花水月) method[[16]](#footnote-67), paying meticulous attention to every word without altering even a single word. Stylistically, Kaneko’s translation resembles Motoori Norinaga’s *Kokin Tōkagami* in his deliberate inclusion of archaic expressions.

窪田 (1960) also emphasizes a strict adherence to word-for-word translation, avoiding changes in word order or word additions. Similarly, 片桐 (1998) follows the same approach, performing a literal translation while avoiding omissions or extensive additions, aiming to remain as faithful as possible to the original poems.

### Poet-focused approach: OKMR, TKOK

The group of translations that emphasize the poet’s intent focuses more on the underlying meaning and sensibility of the original text than on its surface meaning. 奥村 (1978, 7) states that the contemporary translation respects the wording and phrasing of the original poem, but rather than strictly interpreting the meaning of each poem, the emphasis is placed on conveying the basic intent of the poet.

In this approach, understanding the original intent takes precedence over a literal word-for-word translation. 竹岡 (1976, 上下:21) explains that translation (into contemporary colloquial Japanese) is not a mere introduction or explanation of the plot of the original text, but rather aims to clarify the poet’s way of perceiving and feeling in all aspects, and then transposes these expressions into modern equivalents as faithfully as possible. While synonyms in the original poem are considered important, the emphasis is on preserving the poet’s intent and feelings. As a specific translation strategy, 竹岡 (1976, 上下:11–12) proposes the following unique theory:

In this formula, represents the translated poem, are the individual words in the poem, represents the elements included in the translation, and represents additional elements effectively included in the translation. Takeoka explains that the general universal meaning of can be understood through a dictionary, and the way they are combined (denoted by ) can be clarified by referencing grammar books. 竹岡 (1976, 上下:11) also notes that typically has a value of 0, but when effective elements are included in a poem, the translator adds a value to , such as a final particle (e.g., わ *wa*, ね *ne*, よ *yo*). Therefore, the most important variable for creating a contemporary translation is . Takeoka divides into seven categories: word segmentation, corresponding modern vocabulary, word order, particles and conjunctions, sentence structure, discourse, and scene. This approach emphasizes fully conveying the poet’s intent rather than just the surface meaning of the poem’s structure.[[17]](#footnote-69)

### Reader-focused approach: OZW, KSJ

The group of translations that prioritize the reader’s understanding focuses more on conveying the meaning and interpretation of the original poem to the reader, rather than strictly adhering to the original text. This often involves changing the word order or inserting additional words. 小沢 (1971) does not avoid modify the word order and grammar of the original, prioritizing a translation that is easier for the reader to understand. 小沢 (1971, 46) states that he made sure that the translation could stand alone and still be understood, even if it meant changing the word order or grammar of the original text. This highlights a focus on comprehension over literal word-for-word translation. Similarly, 久曽神 (1979) adopts a plain and accessible translation, supplementing expressions where necessary to prioritize interpretation. The translation approach is explained that the intent of the poem is expressed in a highly-readable translation based on the original poems, but additional words were added where necessary (久曽神 1979, 1–5:6).

### Approaches that were not specified: KMCY, MTD, K&A

The translation approach is not explicitly stated in the works of 小町谷 (1982), 松田 (1968), and 小島 and 新井 (1989).

Among these, the contemporary translation in 松田 (1968) was produced by ten co-authors. Although the preface of the annotation book mentions that the translation was created based on previous research, it does not clarify the shared understanding or guidelines for translation among the ten co-authors. Additionally, Matsuda states that his interpretation of the *Kokinshū* not only includes the poets’ emotions but also the editors’ intentions in compiling the anthology (松田 1968, 上下:9). Following each poem, the annotation book includes sections titled “General Explanation (通釈 *Tsūshaku*),” “Word Explanation (語釈 *Goshaku*),” “Classical Commentaries (古注 *Kochū*),” and “Evaluations (評 *Hyō*),” providing interpretation and appreciation of the poems.

One distinguishing feature of 小島 and 新井 (1989) is that it incorporates annotations from the medieval period (1392–1600), which were often overlooked in previous annotation books, alongside annotations from the Edo period (1600–1868) (pp. 481–482). Additionally, comparisons and references to *Man’yōshū* poems and terms are frequently made (小島 and 新井 1989, 480).

### Summary

When examining the translation approaches, most translators aim to maintain the original meaning without altering it, primarily adhering to word-for-word translation, though the emphasis differs slightly. In this paper, we categorized these approaches from the perspective of the communication model into the following groups: text-focused approach, poet-focused approach, reader-focused approach, and approaches that were not specified. Among these, the text-focused approach was the most committed to word-for-word translation. On the other hand, approaches that prioritized the author’s intent or the reader’s understanding allowed for some reordering of words and the supplementation of phrases. However, there is likely overlap among these categories. It should also be noted that these approaches are based on the translators’ subjective awareness of their own translation strategies.

## 4.2 Results of unmatch and addition rate estimates

Referring to [Figure 2](#fig-data) for the calculation results of the addition rate, we can observe that the unmatch rate ranges from 0% to 88% [Table 3](#tbl-data-review). TKOK’s translation had the lowest unmatch rate at 17.2% (sd = 0.1) [Figure 3](#fig-unmatch-rate), while OZW’s translation recorded the highest at 23% (sd = 0.112) [Figure 3](#fig-unmatch-rate). As mentioned earlier, OZW showed greater flexibility in changing word order, whereas Takeoka adhered strictly to his translation principles, a difference clearly reflected in the data.

Of the 10 contemporary translations analyzed, Poem #697 exhibited the highest unmatch rate at 88.2% [Table 3](#tbl-data-review), indicating that most of the words in the poem were not literally translated. We will revisit these poems in the “Discussion” section for further analysis.

|  |
| --- |
| Figure 2: Probability distribution of addition rates by translator |

Picking joint bandwidth of 0.0242

|  |
| --- |
| Figure 3: Probability distribution of unmatch rates by translator (unmatch rate is the proportion of elements in the original poem without agreement) |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | (a) Summary of unmatch and addition rates   | variable | min | max | median | mean | sd | | --- | --- | --- | --- | --- | --- | | UnmatchRate | 0.000 | 0.882 | 0.188 | 0.198 | 0.110 | | AdditionRate | 0.053 | 0.864 | 0.600 | 0.587 | 0.114 | | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | (b) Poems and translations with the highest unmatch rates   | Translator | PoemID | UnmatchRate | | --- | --- | --- | | MTD | 697 | 0.882 | | KSJ | 404 | 0.778 | | KSJ | 886 | 0.733 | | OKMR | 515 | 0.733 | | KSJ | 515 | 0.733 | | |  | |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | (c) Poems and translations with the lowest unmatch rates   | Translator | PoemID | AdditionRate | | --- | --- | --- | | TKOK | 173 | 0.053 | | TKOK | 343 | 0.091 | | TKOK | 478 | 0.115 | | MTD | 661 | 0.133 | | TKOK | 744 | 0.143 | | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | (d) Poems with the highest unmatch rate in contemporary translations   | PoemID | variable | mean | sd | | --- | --- | --- | --- | | 718 | UnmatchRate | 0.481 | 0.078 | | 703 | UnmatchRate | 0.472 | 0.109 | | 576 | UnmatchRate | 0.467 | 0.070 | | 150 | UnmatchRate | 0.447 | 0.084 | | 515 | UnmatchRate | 0.440 | 0.164 | | |  | |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | (e) Poems with the highest addition rate in contemporary translations   | PoemID | variable | mean | sd | | --- | --- | --- | --- | | 669 | AdditionRate | 0.773 | 0.046 | | 420 | AdditionRate | 0.773 | 0.065 | | 705 | AdditionRate | 0.771 | 0.052 | | 749 | AdditionRate | 0.768 | 0.052 | | 143 | AdditionRate | 0.767 | 0.059 | | |

Table 3: Summaries of unmatch and addition rates

## 4.3 Results from statistical modeling: relationship between translation approach and addition rate

The model sampled the addition rates for each translation according to the translation approach [Figure 4](#fig-poster). The $R^{\hat}$ values for all estimated effects were 1, indicating good convergence, as they remained below the threshold of 1.1 (cf., Brooks and Gelman 1998). Additionally, the effective sample size (ESS) reached approximately 2000 in all cases.

### No Significant Differences in Addition Rate by Translation Approach

For the visualization and interpretation of results from the model, following a method similar to Yu and Tamaoka (2020), we observed whether the 95% CrI of the posterior distribution of the difference between two groups included 0, and we assessed the probability that the difference between the two groups was greater than 0. First, if the 95% CrI does not include 0, we conclude that there is a significant difference between the two groups. Even if the 95% CrI includes 0, if the probability that the difference is greater than 0 exceeds 95%, or is lower than 5%, we consider it a notable trend.

The results were as follows: (a) Regardless of the translation approach, approximately 50% of the elements in the translation were additional elements, a finding corroborated by descriptive statistics. (b) When comparing the translation approach groups, the 95% CrI of the posterior distribution for the difference between groups included 0 in all cases, indicating no statistically significant differences in addition rates between the groups.

A trend was observed where the addition rate of the Other group was approximately lower than that of the Reader-focused group (, ). The addition rate for the Reader-focused group was approximately lower than that of the Poet-focused group (, ), although this did not meet the threshold for a significant trend.

|  |  |  |
| --- | --- | --- |
| |  | | --- | | (a) Posterior distribution of addition rates for each translation approach |  |  | | --- | | (b) Posterior distribution of differences in addition rates by translation approach |   Figure 4: Posterior distribution of addition rates |

### Variation in Addition Rates Is Greater Between Poems Than Between Translators

In hierarchical modeling, translators and poems are modeled as random intercepts. The posterior distribution of the standard deviations, which are parameters of the normal distribution shared by each group level, has been visualized [Figure 5](#fig-hyperparameter). The standard deviation distribution for the translator group was greater than that of the poem group. This suggests that the variation between poems is greater than the variation between translators. In other words, from the perspective of random effects, poems contribute more to the variation in addition rates than translators.

|  |
| --- |
| Figure 5: Posterior distribution of group-level hyperparameters |

## 4.4 Case analysis of the translation of “Tatsuta” Poem #298

The case analysis of the “Tatsuta” Poem #298 begins by illustrating the alignment.

[Table 4](#tbl-tatsuta-review) presents the unmatch rate and addition rate for the translations of “Tatsuta” Poem #298, calculated using the alignment method. Out of the 19 tokens in the poem, all translations except OZW aligned more than 14 word pairs. Due to its sentence reordering, OZW struggled to align effectively. The unmatch rate for elements in the original text remains low, while the addition rate for the translations (excluding OZW) ranges from 50% to 70%. On average, translators used 1.5 to 1.7 words per word in the original. Next, we examine the types of elements added.

[Figure 6](#fig-alignment) shows the refined alignment from the alignment script output. For the raw output, refer to the supplementary materials (**?@sec-app-tatsuta**).

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table 4: Overview of unmatch and addition rates in contemporary translations of “Tatsuta” Poem   | Translator | PoemID | Method | TotalMatch | UnmatchRate | AdditionRate | | --- | --- | --- | --- | --- | --- | | KNK | 298 | Alignment | 17 | 0.105 | 0.691 | | KTGR | 298 | Alignment | 16 | 0.158 | 0.568 | | K&A | 298 | Alignment | 15 | 0.211 | 0.500 | | KMCY | 298 | Alignment | 17 | 0.105 | 0.622 | | KBT | 298 | Alignment | 15 | 0.211 | 0.500 | | KSJ | 298 | Alignment | 16 | 0.158 | 0.628 | | MTD | 298 | Alignment | 14 | 0.263 | 0.562 | | OKMR | 298 | Alignment | 14 | 0.263 | 0.682 | | OZW | 298 | Alignment | 7 | 0.632 | 0.851 | | TKOK | 298 | Alignment | 16 | 0.158 | 0.500 | |

|  |  |  |
| --- | --- | --- |
| |  | | --- | | # Tatsutahime (the Goddess of autumn): >| ーーーーーーーーーーーーーーーーー立田姫ーーーーーーーーーーーーーーーーーーーーーー [298] >| ーーーーーーーー【秋をつかさどる】龍田姫【が旅立ちにあたって】ーーーーーーーーーーー [KTGR] >| ーーーーーーーーーーーーーーーーー竜田姫【は】ーーーーーーーーーーーーーーーーーーー [KBT] >| ーーーーーーーーーーーーーーーーー竜田姫【には】ーーーーーーーーーーーーーーーーーー [MTD] >| ーー【秋も終りに近づき秋の女神の】龍田姫【がお帰りになる】ーーーーーーーーーーーーー [OKMR] >| ーーーーーーーーーーーーーーーーー竜田姫【が旅にあって】ーーーーーーーーーーーーーー [TKOK] >| ーーーーーーーーーーーーーーーーー立田姫【は秋の神だが】ーーーーーーーーーーーーーー [KNK] >| ーーーーーーーーーーーーーーーーー竜田姫【が】ーーーーーーーーーーーーーーーーーーー [K&A] >| 【秋の末近くなって帰り道についた】龍田姫【が道中の無事を願って】ーーーーーーーーーー [KMCY] >| ーーーーーーー【秋も終りに近づき】竜田姫【がお帰りになる際に】ーーーーーーーーーーー [KSJ] >| ーーーーーーー【もはや秋の終りで】龍田姫【が帰り道にお着きになった】ーーーーーーーー [OZW]   # Dedicate something to the god >| ーーーーーーーーーーーーーーーーーーー手向けーーーーーーるーーーーーー神ーーーーのー [298] >| ーーーーーーーーーーーーーーーーーーー手向けーーー（をする）ーーーーー神ーーーーがー [KTGR] >| ーーーーーーーーーーーーーーーーーーー手向けー（をするべき）ーーーーー神ーーーーがー [KBT] >| 【旅中】ーーーーーーーーーー（供え物をささげ）ーーーーーるーー【道祖】神ーーーーがー [MTD] >| 【姫が道中の安全を祈って】ーーーーーー手向けーー（をなさる）ーーーーー神ーーーーがー [OKMR] >| ーーーーーーーーーーーーーーーーーーー手向けーーーーーーるーーーーーー神ーーーーがー [TKOK] >| 【それすら暮れて行かれる折には】【お】手向けーーー（なさる）ー【道の】神【様】ーがー [KNK] >| ーーーーーーーーーーーーーーーーーーー手向けーーー（をする）ーーーーー神【さま】がー [K&A] >| ーーーーーーーーーーーーーーーーーーー手向けーーー（をする）ーーーーー神ーーーーがー [KMCY] >| 【旅の安全を祈って】ーーーーーーーーー手向けーーー（られる）ーーーーー神ーーーーがー [KSJ] >| 【中略】【道の】（神様）【にそれを】（供えていらっしゃる）＊　　　　　　　　　　　　 [OZW]  # Exists, so that >| ーーーーーあれーーーーーーーーーーーーーーーーーーーーーーーーばーこそーーーーーーー [298] >| ーーーーーあるーーーーーーーーーーーーーーーーーーーーーーーからーーーーーーーーーー [KTGR] >| ーーーーーあるーーーーーーーーーーーーーーーーーーーーーーーのでーーーーーーーーーー [KBT] >| ーーーーーあるーーーーーーーーーーーーーーーーーーーーーーーからーこそーーーーーーー [MTD] >| ーーーーーあるーーーーーーーーーーーーーーーーーーーーーーーからーこそ【だろう】ーー [OKMR] >| ーーーーーあるーーーーーーーーーーーーーーーーーーーーーーーからーこそーーーーーーー [TKOK] >| ーーーーーあるーーーーーーーーーーーーーーーーーーーーーーーからーこそーーーーーーー [KNK] >| ーー（おられる）ーーーーーーーーーーーーーーーーーーーーーーからーこそーーーーーーー [K&A] >| ーーーーーあるーーーーーーーーーーーーーーーーーーーーーーーからーこそーーーーーーー [K&A] >| ーーーーーあるーーーーーーーーーーーーーーーーーーーーーーーからーこそーーーーーーー [KSJ] >| 【中略】（のは）【姫が道の神様にそれを供えていらっしゃる】（のだ）【な】＊　　　　　 [OZW]  (a) Alignment (first half): Tatsutahime tamukeru kami no areba koso ‘[There is the] god exits [for] Tatsutahime to dedicate [things], so that …’ | |  |

|  |  |
| --- | --- |
| |  | | --- | | # Leaves of autumn >| ーーーーーーーーーーーーーーーーー秋のーーーーーー木の葉ーーーーーーのーーーーーーー [298] >| ーーーーーーーーーーーーーーーーー秋のーーーーーー木の葉ーーーーーーがーーーーーーー [KTGR] >| 【そのつかさどる】ーーーーーーーー秋のーーーーーー木の葉ーーーーーーがーーーーーーー [KBT] >| ーーーーーーーーーーーーーーーーー秋の【紅葉した】ー（葉）ーーーーーがーーーーーーー [MTD] >| ーーーーーーーーーーーーーーーーー・・ーーーーーー（紅葉）ーーーーーがーーーーーーー [OKMR] >| 【それであのように】ーーーーーーー秋のーーーーーー木の葉ーーーーーーがーーーーーーー [TKOK] >| 【このように御自身お染めなされた】秋のーーーーーー木の葉ーーーーーーがーーーーーーー [KNK] >| ーーーーーーーーーーーーーーーーー秋のーーーーーー木の葉【のもみじ】がーーーーーーー [K&A] >| ーーーーーーーーーーーーーーーーー秋のーーーーーー木の葉ーーーーーーがーーーーーーー [KMCY] >| ーーーーーーーーーーーーーーーーー秋のーーーーーー（紅葉）ーーーーーがーーーーーーー [KSJ] >| ーーーーーーーーーーーーーーーーー・・ーー【山の】（紅葉）ーーーーーがーーーーーーー [OZW]  # Scattered as scared paper (which are ritual objects used in Shinto practices) >| ーーーーーーーー幣ーーーーーーーーとー散ーーーーるーーーーーーーーらめーーーーーーー [298] >| ーーーーーーーー幣【を撒く】（ように）散（っている）ーーーーのーだろうーーーーーーー [KTGR] >| ーーーーーーーー幣ーーーー（のように）散ーーーーるーーーーーのであろうーーーー【よ】 [KBT] >| ーーーーーーーー幣ーーーー（のように）散（っている）ーーーーのーだろうーーーーーーー [MTD] >| ーーーーーーーー幣ーーーー（のように）散（っている）ーーーーー【だろう＊】ーーーーー [OKMR] >| ーーーーーーーー幣ーーーーー（として）散（っている）ーーーーのであろうーーーーーーー [TKOK] >| ー【手向けの】ー幣ーーーー（のように）散ーーーーるーーーーーのであろうーーーーーーー [KNK] >| ー【手向け】（もの）ーーーー（として）散ーーーーるーーーーーのでしょうーーーー【ね】 [K&A] >| ーーーーーーーー幣ーーーー（となって）散（っている）ーーーーのーだろうーーーーーーー [KMCY] >| ーーーーーーーー幣ーーーー（となって）散（っている）ーーーーのであろうーーーーーーー [KSJ] >| 【色とりどりの】幣ーーーー（となって）散（っている）【中略】のだーー・ーーーー【な】 [OZW]  (b) Alignment (second half): aki no konoha no nusa to chiru rame ‘leaves of autumn may be scattered as scared papers [which is dedicated to the god]’ | |

Figure 6: Alignment of translations by 10 translators (the author corrected the output of the alignment estimation program): Elements in “【】” are judged as additions, those in “（ ）” are considered paraphrases, “＊” indicates phrases or words aligned through sentence reordering, and “・” represents elements judged not to be translated directly.

#### Additional elements regarding co-occurring words

This section examines the translation process for co-occurring words at the corpus level. In the “Tatsuta” poem, content words such as “mountain” (*yama* 山), “river” (*kawa* 川), “to turn colored” (*momidzu* 紅葉づ), “princess/goddess” (*hime* 姫), “to dedicate offerings to the gods” (*tamuku* 手向く), “colored leaf” (*momiji-ba* 紅葉葉), “Kamnabi” (proper noun, *kamnabi* 神無備), “autumn” (*aki* 秋), “ritual paper” (*nusa* 幣), and “to be scattered” (*chiru* 散る) frequently appear. As a typical “Tatsuta” poem, Poem #298 includes “princess/goddess,” “to dedicate offerings to the gods,” “ritual paper,” and “autumn.” Most of the words from the original poem are translated directly. Words like “colored leaf” (*momidzi-ba* 紅葉葉) and “to get colored” (*momidzu* 紅葉づ), which do not explicitly appear in Poem #298, are implied by “leaves” (*konoha* 木の葉) and are translated as follows:

* 【紅葉した】葉 *[momiji shita] ha* (“[colored] leaves”) [MTD]
* 木の葉【のもみじ】 *konoha [no momiji]* (“[coloring of] leaves”) [K&A]
* 紅葉 *[momiji]* (“[colored leaves]”) [KSJ, OZW, OKMR]

Translators TKOK, KNK, and KMCY did not address the translation of “leaves.”

For the corpus-level co-occurrence of “ritual paper,” the following translations added extra elements:

* 【色とりどりの】幣 *[irotoridori no] nusa* (“[colorful] ritual paper”) [OZW]
* 【手向けの】幣 *tamuke no nusa* (“ritual paper [offered to the gods]”) [KNK]
* 幣【を撒くように】 *nusa wo maku yo ni* (“ritual paper [scattered like]”) [KTGR]

These additions raise some challenges in determining whether they should be attributed specifically to “Tatsuta.” However, interpreting them indirectly within the context of “ritual paper” seems reasonable.

Overall, the process of co-occurring elements in “Tatsuta” either retains the original content or adds more compared to other factors. In cases where words like “leaves” do not explicitly co-occur, translators opted for “colored leaves,” considering broader collocations. Regardless of the translator, the approach to translation remained consistent.

The differences between the translations and the original text suggest that co-occurring words at the corpus level may offer important insights into translation choices.

#### Translation of grammatical patterns

Co-occurring grammatical patterns in the *Kokinshū* reveal a noticeable scarcity of poems ending in the conclusive form (only 1 out of 12), along with a frequent use of *kakarimusubi* (9 out of 12 poems), which contributes to a lingering tone at the end of the poems. This analysis focuses on how these grammatical patterns were handled in the translation of Poem #298.

In Poem #298, the *kakarimusubi* is formed by the emphasis particle “こそ” (*koso*) and the inferential mood auxiliary verb “らむ” (*ramu*) in its conjugation “らめ” (*rame*). Of the 10 translators, 7 retained *koso*, and 9 used the EMPHASIS “の[でしょう]” (*no*) construction at the end of the sentence. Regarding inferential *ramu*, 9 translators rendered it as “だろう” (*darō*) or “でしょう” (*deshō*) (“likely modality/inferential mood”). OZW, the only translator who did not use “だろう” (*darō*), translated it with the exclamatory “だな” (*da na*). KBT and K&A added the final particles “よ” (*yo*) and “ね” (*ne*), respectively, at the sentence’s end.

This confirms that the translation of this typical grammatical pattern is highly robust. In translations of other “Tatsuta” poems that do not use *kakarimusubi*, no instances were found where this grammatical pattern was added. Thus, the handling of these distinctive co-occurring grammatical patterns was consistently reflected in the translations.

Other grammatical elements, such as the particle “と” (*to*) and the translation of the *-u* verb conjugation, were also observed, though they are not characteristic co-occurring elements of “Tatsuta” at the corpus level. The following variations in translation were noted:

**幣「と」 (*nusa “to”*):**

* 幣を撒く「ように」 *nusa wo maku [youni]* (“[like] scattering ritual paper”) [KTGR, KBT, MTD, OKMR, KNK]
* 幣「として」 *nusa [toshite]* (“[as] ritual paper”) [TKOK, K&A]
* 幣「となって」 *nusa [tonatte]* (“[becomes] ritual paper”) [KMCY, KSJ, OZW]

**散「る」 (*chir-[u]*):**

* 散っ「ている」 *chir-[teiru]* (“scatter + [progressive aspect]”) [KTGR, MTD, OKMR, TKOK, KMCY, KSJ, OZW]
* 散「る」 *chir-[u]* (“scatter”) [KBT, KNK, K&A]

**手向け「る」 (*tamuker-[u]*):**

* 手向けを「する」 *tamuke wo suru* (“[do] dedicating”) [KTGR, K&A, KMCY]
* 供え物をささげ「る」 *sonaemono wo sasager-[u]* | 手向け「る」 *tamuker-[u]* (“dedicate”) [MTD, TKOK]
* 手向けを「なさる」 *tamuke wo nasaru* | お手向けなさる *o-tamuke nasaru* (“[make+(honorific)] dedicating”) [OKMR, KNK]
* 手向けを「するべき」 *tamuke wo suru [beki]* (“[do + likely mood/modality] dedicating”) [KBT]
* 手向け「られる」 *tamuke [rareru]* (“dedicate + [honorific]”) [KSJ]
* それを供えて「いらっしゃる」 *sore wo sonaete [irassharu]* (“dedicate + [progressive aspect + honorific]”) [OZW]

The co-occurring grammatical elements at the corpus level show significant variation in the addition and paraphrasing of translations. The variability in how naked verb phrases are translated has also been reported in Yamamoto, Hodoscek, and Chen (2023). These additions and paraphrases likely stem from linguistic change and variation between classical and contemporary Japanese. While they cannot be identified using WLSP codes for identifying agreements, they emerge from the differences between the contemporary translations and the original text. But they are not considered non-literal information.

#### Translation of semantic preferences

At the corpus level, the co-occurring words in “Tatsuta” tend to belong to groups like “mountain” and “river” (*yama* and *kawa*), which form place names such as “Tatsutagawa” (Tatsuta River) and “Tatsuta(no)yama” (Mountain of Tatsuta). Additionally, words related to the “gods” of autumn are prominent. These reflect the nature of *utamakura* (poetic place names), which represent sacred locations associated with autumn and autumn leaves. Although this tendency is not directly emphasized in Poem #298 through explicit additions, it was observed indirectly that additional elements often appeared around “Tatsutahime” (Goddess Tatsuta) and “kami” (god) in the translations by the 10 translators.

Seven of the 10 translators added elements related to “Tatsutahime”:

* 【秋をつかさどる】龍田姫【が旅立ちにあたって】 *[aki wo tsukasadoru] Tatsutahime [ga tabidachi ni atatte]* (“[Autumn goddess] Tatsutahime [embarks on her journey]”) [KTGR]
* 【秋も終りに近づき秋の女神の】龍田姫【がお帰りになる】 *[aki mo owari ni chikazuki aki no megami no] Tatsutahime [ga okaeri ni naru]* (“[As autumn nears its end, the autumn goddess] Tatsutahime [returns home]”) [OKMR]
* 竜田姫【が旅にあって】 *Tatsutahime [ga tabi ni atte]* (“Tatsutahime [is on a journey]”) [TKOK]
* 立田姫【は秋の神だが】 *Tatsutahime [wa aki no kami da ga]* (“Tatsutahime [is the god of autumn]”) [KNK]
* 【帰り道についた】龍田姫【が道中の無事を願って】 *[kaerimichi ni tsuita] Tatsutahime [ga dōchū no buji wo negatte]* (“[As she reached her way home,] Tatsutahime [prays for a safe journey]”) [KMCY]
* 【秋も終りに近づき】竜田姫【がお帰りになる際に】 *[aki mo owari ni chikazuki] Tatsutahime [ga okaeri ni naru sai ni]* (“[As autumn nears its end,] Tatsutahime [returns home]”) [KSJ]
* 【もはや秋の終りで】龍田姫【が帰り道にお着きになった】 *[mohaya aki no owari de] Tatsutahime [ga kaerimichi ni otsuki ni natta]* (“[With autumn already at its end,] Tatsutahime [has arrived on her journey home]”) [OZW]

In the translations by KBT, MTD, and K&A, no additional elements were added around “Tatsutahime.”

For the context of “the god to whom offerings are made,” additional elements appeared in 5 translations:

* 【旅中】（供え物をささげる）【道祖】神 *[tabi naka] (sonaemono wo sasageru) [dōso] kami* (“[During the journey] (offering a gift) [to the road] god”) [MTD]
* 【姫が道中の安全を祈って】手向け（をなさる）神 *[hime ga dōchū no anzen wo inotte] tamuke (wo nasaru) kami* (“[The goddess prays for safety during the journey] as she offers to the god”) [OKMR]
* 【それすら暮れて行かれる折には】【お】手向け（なさる）【道の】神【様】 *[sore sura kurete ikareru ori ni wa] [o]tamuke (nasaru) [michi no] kami[sama]* (“[At the time of her departure as the day ends,] the offering to the god [of the road]”) [KNK]
* 【旅の安全を祈って】手向け（られる）神 *[tabi no anzen wo inotte] tamuke (rareru) kami* (“[Praying for a safe journey,] the god receives the offering”) [KSJ]
* 【道の】（神様）【にそれを】（供えていらっしゃる）＊ *[michi no] (kamisama) [ni sore wo] (sonaete irassharu)* (“To the god [of the road,] (she is offering) [that]”)＊ [OZW]

K&A, KMCY, KTGR, KBT, and TKOK did not add elements in these contexts. It can be inferred that more emphasis is placed on providing additional information related to “Tatsutahime,” “offerings,” and “gods” than other elements. The differences between the original and contemporary translations suggest that “Tatsutahime” is the “autumn god” or a personification of autumn, the purpose of “offering” by Tatsutahime is for “safe travels,” and the god receiving the offering is a road god, a travel god, or *dōsojin* (a deity protecting travelers).

# 5. Discussion

## 5.1 Gap between translation approaches and practices

The statistical modeling of addition rates in contemporary translations shows that a substantial proportion of additional elements is present, regardless of the translation approach. It appears that variations between poems contribute more to the addition rate than variations between translators. Despite the detailed explanations of many words in annotation books, the addition rate has not significantly decreased. Whether translators focus on the text itself, the poet’s intent, or making the translation more accessible to readers, there seem to be elements that inevitably require additions.

Closer examination of individual translations reveals clear reflections of each translator’s approach. For instance, OZW allows for changes in word order, while TKOK’s translations emphasize the author’s intent. While analyzing specific poems uncovers distinct differences, these differences, when viewed collectively, tend to emerge as general trends rather than significant variations. These factors seem to go beyond the personal viewpoint and style of the translator, leaving room to further clarify the nature of these elements. The non-literal elements in contemporary translations could potentially be visualized more effectively.

## 5.2 Many “omissions” in contemporary translations involve *jokotoba* and *makurakotoba*

When extracting non-literal information through translation differences, it is preferable for indirect translations to be included, but it becomes problematic when certain words are entirely absent. A high unmatch rate between the original and the translation may indicate that some elements are incorporated in some form, but it is also likely that many elements have been omitted. Several such poem-translation pairs have been identified.

Upon reviewing these pairs, three common omissions were observed across the 10 contemporary translations: *jokotoba* (prefatory phrases), *makurakotoba* (pillow words), and complex expressions. Some translators consciously chose to omit *jokotoba* and *makurakotoba* entirely. For example, in MTD’s translations, four of them are noticeably shorter than the original poems. In MTD (#173) and MTD (#665), the *makurakotoba* “hisakata no” and “mitsushio no” are omitted. In MTD (#684) and MTD (#697), both *makurakotoba* and *jokotoba* are omitted. Naturally, these omissions result in a high unmatch rate. Notably, MTD (#697) is much shorter than the original poem[[18]](#footnote-121). In MTD (#697), the *makurakotoba* “Shikishima no,” referring to “Yamato” (a proper noun), is omitted. In the original poem, there is a *kakekotoba* (pun) on “koromo” (meaning either “time” or “clothes”), but the *jokotoba* linked to “clothes” is also omitted.

Thus, extracting non-literal elements such as *makurakotoba* and *jokotoba* from the differences between contemporary translations and the original poems proves to be a challenging task.

mode=2; level=3; elements of poem=16; elements of translation=6;  
bag (E=1, F=0, G=0, U=15, T=1, AddRate=83.33%, UnmatchRate=93.75%);  
alignment (E=1, F=0, G=0, U=15, T=1, AddRate=83.33%, UnmatchRate=93.75%)  
>| しきしまのやまとにはあらぬからころもころもへずしてあふーーよしーーーもがな [697]  
>| ーーーーーーーーーーーーーーーーーー絶え間なくーー会いたいものだなあーーー [MTD]

Contemporary translation of Poem #697 by 松田 (1968): The alignment has been revised; the *jokotoba* “Shikishima no Yamato ni wa aranu kara koromo” has been omitted.

It may seem desirable to align the complex expressions found in the original poem with those in the contemporary translation, but achieving this is technically challenging. For example, in OKMR (#346), Okumura’s contemporary translation uses very few words from the original poem, instead opting for the following complex expressions:

* とりそえて *torisoete* = さらにそえて *sarani soete* (“additionally”)
* とどめおきて *todome okite* = のこしておきましょう *nokoshite okimashō* (“let’s leave it”)

The alignment method struggles to handle expressions like this effectively. Creating a dictionary that pairs a single poetic word with a complex expression of equivalent meaning and achieves a perfect match is difficult. While it is sometimes possible to align certain poetic words with parts of complex expressions, it is rarely feasible to align the entire expression. This issue has been described as a problem of word alignment, and as previously mentioned in relation to one-to-many translations, complex expressions often add additional information to the original poetic words. As a result, the unmatched elements from these complex expressions (i.e., the leftover parts of the translation) are valuable when examining non-literal elements.

## 5.3 Many additions in contemporary translations involve kakekotoba (pun)

MTD (#629) shows that most words did not align well with the original poem using the alignment method. This poem includes archaic words such as あやなし *ayanashi* (“without reason”) and まだき *madaki* (“early”), along with grammatical constructions like なくに *naku ni* (negation with interjection nuance), which are no longer used in contemporary Japanese. These terms cannot be mapped to their modern equivalents using WLSP codes. Additionally, the poem contains a *kakekotoba* (pivot word) involving 名の立つ *na no tatsu* (“to have a rumor”) and 立田川 *Tatsutagawa* (“Tatsuta River”).

Among the top 10 poems with the highest average addition rates, other poems such as Poem #705, Poem #669, and Poem #617 also have high addition rates. Poem #705 features a *kakekotoba* involving 身 *mi* (“body”) and 雨 *ame* (“rain”), with ふる (降る/経) *furu* serving a dual meaning: “(rain) to fall” and “(body) to age.” Poem #669 and Poem #617 contain *kakekotoba* involving 海松目/見る目 *mirume* (a pun on “seaweed” and “to meet”) and 眺め/長雨 *nagame* (a pun on “gaze/reverie” and “long rain”), respectively. To clarify the role of *kakekotoba*, supplemental information is often required in translations, and many elements likely remain unaligned when comparing contemporary and classical Japanese.

mode=2; level=3; elements of poem=15; elements of translation=60;  
bag (E=1, F=2, G=6, U=6, T=9, AddRate=85.00%, UnmatchRate=40.00%);  
alignment (E=1, F=0, G=6, U=8, T=7, AddRate=88.33%, UnmatchRate=53.33%)  
args: translator:matsuda; poem No.0629; weight:(U=-1, G=10, F=13, E=17); gap penalty: 0.01  
  
# Without any reason,  
1| ーーーーーーーーあやなくーーーてーーーーーーーーーーーーーーーーーーーーーーーーーー　[629]  
1| わけもなくもう無実の評判が立ってしまったいずれ二人の恋は遂げられないではいられないも [MTD]  
# soon rumor has spreaded, the River Tatsuta, without crossing, [I cannot] stop,  
2| ーーーーーーーーーーーーまたきなき名のー立田河わたらてやまーーんーー物ーーーーーーー　[629]  
2| のではあるがわけもなくもう無実の評判が立っーーーーーてしまったーいずれ二人の恋は遂げ [MTD]  
# It is not possible...  
3| ーーならーーなーーーーーーーくーーにーーーーーーーーーーーーーーーーーーーーーーーー [629]  
3| られないではいられないものではあるがーーーーーーーーーーーーーーーーーーーーーーーー　[MTD]

Contemporary translation of Poem #629 by 松田 (1968): The alignment results have been modified.

## 5.4 Categories of non-literal elements in the case study

Based on the concept of extended meaning units, we analyzed additional elements in translations from three perspectives and propose the following categorization:

1. Additional elements derived from co-occurring words at the corpus level:
   * Direct additions of co-occurring words: e.g., *momiji* (autumn leaves)
   * Supplementary or secondary/indirect additions to co-occurring words: e.g., “colorful” or “scattering” in reference to *nusa* (ritual paper)
2. Additional elements that fill gaps caused by language change, particularly grammatical shifts:
   * Changes in grammatical structures that cannot be matched using WLSP codes: e.g., the translation of *to* (particle)
   * Extensions of bare verb forms: e.g., the translation of *tamukeru* (to dedicate)
3. Additional elements related to the semantic preference of target words:
   * Further supplementation of words based on their semantic preference at the corpus level: e.g., additional information surrounding *kami* (god)

Distinguishing these elements is essential when constructing a visualization system. For instance, supplementing “autumn leaves” (*momiji*) to “leaves” (*konoha*) may be considered non-literal at the sentence level but literal at the corpus level, as it represents co-occurring words. This can be identified through difference-based visualization, though word association measures are also effective. As a result, the necessity for difference-based visualization in this case is reduced.

Next, additional elements that bridge gaps caused by grammatical shifts over time should be excluded, as they do not represent non-literal information derived from the original text.

Moreover, elements related to the collective semantic preference of co-occurring words may not be easily captured through word association measures, which is where difference-based visualization provides unique advantages.

Lastly, this paper did not cover the fourth level of extended meaning units—discourse prosody—in detail. Emotional values related to discourse prosody and the sociolinguistic characteristics of words are considered elements that cannot be captured through additional elements in translation. These attributes are not explicitly conveyed through direct supplementation or additions. In some cases, connotative terms in contemporary Japanese may paraphrase these aspects, but even when such terms are functionally equivalent, they do not provide direct explicitation. Therefore, alternative methods should be considered to evaluate these elements.

## 5.5 Credibility of Knowledge in Translation

Although this paper does not address the credibility in detail, the system for visualizing differences between the 10 translations and the original text can be seen as a form of meta-analysis, in the sense that it systematically synthesizes the insights of the 10 individual translators regarding each poem.

However, the paper does not explore how much the knowledge included in contemporary translations allows access to the non-literal aspects of the poetic language, or to what extent the translation is valid and reliable. For example, adding excessive interpretation in translation can undermine both validity and coherence. The original text’s simple meaning or intent may become ambiguous due to over-interpretation, making it difficult for readers to accurately understand the text. If such translations are applied to visualization, the visualization itself loses its credibility. This is a concern that must be considered. Finding robust elements within the translators’ choices is therefore a critical challenge for translation-based visualization.

## 5.6 The Need for Interpretation through Translation Studies Insights

This paper does not primarily position itself within translation studies, but phenomena such as the addition of elements in translation or the omission of elements from the original text can be interpreted through various hypotheses and theories in the field.

For example, the translation universal hypothesis (cf. Chesterman (2004); Edina (2016)) introduces concepts like the law of lengthening (cf. Vinay (1958)) and the law of explicitation (cf. Baker (1996); Blum-Kulka and House (1986)), which help explain why additional elements appear in contemporary translations. These phenomena are seen as universal within translation. The law of explicitation, in particular, is highly relevant to the non-literal elements discussed in this paper. Blum-Kulka and House (1986) suggests that translators often introduce cohesive markers absent from the source text when translating into the target text[[19]](#footnote-126). Meanwhile, Baker (1996, 180) expanded the definition of explicitation, describing it as a broader tendency to explain things explicitly, rather than leaving them implicit. Originally, explicitation referred to an increase in discourse markers in translations, and this concept can help interpret some of the added elements discussed in this paper. However, not all phenomena fit into this hypothesis[[20]](#footnote-127).

For example, while not all added elements involve discourse markers, they still account for a significant proportion of the additional elements observed in the analyzed cases. Thus, not all added elements can be attributed solely to translation universals.

In addition to this, research focusing on the cognitive aspects of translation, such as the literal translation hypothesis[[21]](#footnote-128), offers additional insights. This hypothesis suggests that translation literality can be operationalized through 1) word-order similarity between the source and target texts, and 2) the number of potential translation renderings available for a given text (cf. Carl and Schaeffer 2017). These factors provide a different perspective from the addition and unmatch rates discussed in this paper. The second point, in particular, may provide useful metrics for visualizing non-literal elements and is worth further exploration.

Moreover, translation studies have long discussed the issue of analytical units, focusing on translation equivalence (e.g., Eugene A. Nida and Taber [1982] 1969) and translation units (cf. Koller [2004] 1979; **Malmkjaer1998Unit?**). The impact of these analytical units requires careful consideration and offers significant possibilities for further investigation.

## 5.7 Room for Methodological Improvement

While this paper’s rule-based approach to calculating the agreement rate offers advantages in simplicity and consistency, challenges remain, particularly in alignment using dynamic programming methods. Although historical Japanese from certain periods qualifies as a low-resource language, it would be valuable to explore neural-based alignment methods alongside manual alignment. For instance, as suggested by Palladino, Shamsian, and Yousef (2022) and Camilleri (2024), the Ugarti platform introduced by Palladino, Shamsian, and Yousef (2022) offers greater precision. This platform enables detailed analysis of one-to-many and many-to-many correspondence patterns in translations, as well as part-of-speech analysis for non-aligned elements.

## 5.8 As an application using the WLSP semantic category

In this paper, the calculation of the unmatch and addition rates did not rely on an alignment model from natual language processing but instead used a stable semantic classification metacode system, WLSP. The bag method, in particular, can detect semantic agreements without being influenced by word order changes. The utility of a hierarchical classification system that groups words with the same meaning was reaffirmed in the study of old Japanese. While this study still employs the old-version WLSP codes, the potential for expanded research is promising with the assignment of classification codes to classical Japanese by Asahara et al. (2022) in the Corpus of Historical Japanese. However, as of 2024, the WLSP codes are assigned only at the group level, meaning the CHJ-WLSP cannot currently be used as direct token identifiers. Researchers will need to apply additional tagging to detect semantic agreement, including identifiers for synonyms, homonyms, and orthographic variants, depending on their research goals. The WLSP codes used in this paper will also require ongoing refinement. There remains significant potential to further explore how to build a more effective semantic category code system.

# 6. Conclusion

In this paper, we analyzed the additional elements in 10 different contemporary translations of the *Kokinshū* to demonstrate their usefulness in visualizing the non-literal elements of the original text.

We specifically outlined the background of contemporary translations of the *Kokinshū* and categorized the various translation approaches of the 20th century based on commentary provided by the translators. These approaches can be grouped into four categories from a communication model perspective: emphasis on the poet’s intent (source), emphasis on the literal meaning of the text (signal), emphasis on the reader (destination), and approaches that were not clearly defined.

To determine whether these subjective awarenesses of translation approaches were reflected in the objective measures of additions and agreement, we recalculated the addition rates using a clearer methodology and updated data, following the procedure established in Yamamoto and Hodošček (2019). Statistical modeling of the unmatch rates between the original poems and the contemporary translations, along with the addition rates, revealed that information was universally added regardless of the translation approach, and that variations in addition rates between translators were smaller than variations between individual poems.

To assess whether the added elements qualify as non-literal elements, we aligned the contemporary translations of Poem #298 “Tatsuta” across 10 translations and examined how co-occurring words at the corpus level, co-occurring grammatical patterns, and the semantic preferences of these words were handled in the sentence-level translations. The results showed that for units with a strong co-occurrence tendency at the corpus level, sentence-level processing was relatively consistent across the 10 translators. In contrast, elements that did not exhibit a co-occurrence tendency at the corpus level showed greater variation in the additional elements. Notable findings included the use of co-occurring words at the corpus level to inform inference in translation, as well as the tendency to provide supplementary explanations for words related to the semantic preferences of co-occurring words. Meanwhile, additions related to grammatical patterns were primarily used to bridge gaps caused by linguistic changes.

This study demonstrates that even contemporary translations of classical Japanese poetry, which may adhere strictly to literal or word-for-word translation, can serve as valuable interpretive tools, distinct from dictionaries, to supplement non-literal information embedded in poetic language. It also highlights the importance of excluding elements added to address linguistic gaps and emphasizes the need to carefully select reliable elements in translation for further study.

# Appendix

## Statistcal model for addition rate

Here[[22]](#footnote-134):

* refers to the addition rate and is assumed to follow a beta distribution.
* and correspond to the shape parameters of the beta distribution.
* is the mean parameter of the beta distribution, obtained by applying a logistic transformation to the linear predictor .
* , where , represents the fixed effect for the four translation approaches, modeling a distinct effect for each translation approach.
* , where , represents the random effect term (random intercept) for the translators.
* , where , represents the random effect term (random intercept) for the poems (PoemID).
* is the precision parameter of the beta distribution and is modeled by the random effect term for the poems (PoemID).
* and are the population-level intercepts for the part and the part, respectively.

Regarding the prior distributions, we generally follow the default settings in brms: - The intercepts , , and the fixed effects for each translation approach are assumed to follow Student-t prior distributions (the default prior distribution in brms). - , , and represent the standard deviations of the random effect terms for the translators, poems, and precision parameters, respectively, and are assumed to follow Student-t distributions (the default prior distribution in brms).

Refer to the following table for the default priors:

Warning: Argument 'autocor' should be specified within the 'formula' argument.  
See ?brmsformula for help.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table 5: Default proirs   | prior | class | coef | group | resp | dpar | nlpar | lb | ub | source | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | |  | b |  |  |  |  | a |  |  | default | |  | b | Intercept |  |  |  | a |  |  | default | | student\_t(3, 0, 2.5) | sd |  |  |  |  | a | 0 |  | default | |  | sd |  | PoemID |  |  | a |  |  | default | |  | sd | Intercept | PoemID |  |  | a |  |  | default | |  | sd |  | Translator |  |  | a |  |  | default | |  | sd | Intercept | Translator |  |  | a |  |  | default | |  | b |  |  |  |  | b |  |  | default | |  | b | FocusOthers |  |  |  | b |  |  | default | |  | b | FocusPoetMfocused |  |  |  | b |  |  | default | |  | b | FocusReaderMfocused |  |  |  | b |  |  | default | |  | b | FocusTextMfocused |  |  |  | b |  |  | default | | student\_t(3, 0, 2.5) | Intercept |  |  |  | phi |  |  |  | default | | student\_t(3, 0, 2.5) | sd |  |  |  | phi |  | 0 |  | default | |  | sd |  | PoemID |  | phi |  |  |  | default | |  | sd | Intercept | PoemID |  | phi |  |  |  | default | |

Family: beta   
 Links: mu = logit; phi = log   
Formula: AdditionRate ~ a + b   
 a ~ 1 + (1 | Translator) + (1 | PoemID)  
 b ~ 0 + Focus  
 phi ~ 1 + (1 | PoemID)  
 Data: . (Number of observations: 10000)   
 Draws: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;  
 total post-warmup draws = 4000  
  
Multilevel Hyperparameters:  
~PoemID (Number of levels: 1000)   
 Estimate Est.Error l-95% CI u-95% CI Rhat Bulk\_ESS Tail\_ESS  
sd(phi\_Intercept) 0.43 0.02 0.38 0.47 1.00 1821 2719  
sd(a\_Intercept) 0.31 0.01 0.30 0.33 1.00 1183 2180  
  
~Translator (Number of levels: 10)   
 Estimate Est.Error l-95% CI u-95% CI Rhat Bulk\_ESS Tail\_ESS  
sd(a\_Intercept) 0.19 0.08 0.10 0.40 1.00 1668 2287  
  
Regression Coefficients:  
 Estimate Est.Error l-95% CI u-95% CI Rhat Bulk\_ESS  
phi\_Intercept 3.73 0.02 3.69 3.77 1.00 3284  
a\_Intercept 0.27 1.06 -1.78 2.31 1.00 1706  
b\_FocusTextMfocused 0.17 1.06 -1.83 2.26 1.00 1708  
b\_FocusPoetMfocused -0.06 1.07 -2.07 2.01 1.00 1739  
b\_FocusReaderMfocused 0.21 1.06 -1.83 2.30 1.00 1804  
b\_FocusOthers 0.03 1.06 -2.02 2.09 1.00 1741  
 Tail\_ESS  
phi\_Intercept 3297  
a\_Intercept 1980  
b\_FocusTextMfocused 1940  
b\_FocusPoetMfocused 1990  
b\_FocusReaderMfocused 2066  
b\_FocusOthers 2008  
  
Draws were sampled using sample(hmc). For each parameter, Bulk\_ESS  
and Tail\_ESS are effective sample size measures, and Rhat is the potential  
scale reduction factor on split chains (at convergence, Rhat = 1).

## “Tatsuta” poem #698

Document frequency = 12  
01:000108 | 花のちることやわひしき春霞立田の山の鴬のこゑ  
01:000283 | 龍田川紅葉みたれてなかるめりわたらは錦中やたえなん  
01:000284 | 龍田川紅葉はなかる神なひのみむろの山に時雨ふるらし  
01:000294 | ちはやふる神代もきかす立田川から紅に水くゝるとは  
01:000298 | 立田姫たむくる神のあれはこそ秋のこのはのぬさとちるらめ  
01:000300 | 神なひの山を過行秋なれは龍田川にそぬさはたむくる  
01:000302 | もみちはのなかれさりせは立田川水の秋をはたれかしらまし  
01:000311 | としことに紅葉はなかす龍田川みなとや秋のとまり成らん  
01:000314 | 龍田川錦をりかく神無月しくれの雨をたてぬきにして  
01:000629 | あやなくてまたきなき名の立田河わたらてやまん物ならなくに  
01:000994 | 風ふけはおきつ白浪たつた山夜半にや君かひとりこゆらん  
01:000995 | たかみそきゆふつけとりか唐衣たつたの山におりはへてなく

High frequency content collocates of tatsuta:  
 1. 30 BG-01-5240-05-0100|山  
 2. 17 BG-02-5020-01-1201|紅葉づ  
 3. 15 BG-01-5250-01-0102|川+  
 4. 15 BG-01-1624-04-0100|秋  
 5. 13 BG-01-5240-05-0100|山+  
 6. 12 BG-02-3510-01-0800|見る\*  
 7. 12 BG-02-3090-02-0100|見る\*  
 8. 12 BG-02-3060-05-1000|見る\*  
 9. 12 BG-02-1556-01-0202|見る  
 10. 9 BG-01-5530-10-0305|紅葉葉  
 11. 9 BG-01-5530-09-0100|葉+  
 12. 9 BG-01-4201-03-0100|錦  
 13. 9 BG-01-2590-02-0600|唐+  
 14. 8 BG-02-3420-01-0100|す  
 15. 8 BG-02-1552-01-0100|散る  
 16. 8 BG-01-2030-01-1001|姫+  
 17. 7 BG-02-5150-03-0700|時雨る  
 18. 7 BG-01-5151-02-0100|嵐  
 19. 7 BG-01-5020-03-0400|白+  
 20. 6 BG-02-5150-05-0100|吹く\*  
 21. 6 BG-02-3394-07-0100|吹く  
 22. 6 BG-02-1527-01-0102|来  
 23. 6 BG-01-5130-03-0201|水  
 24. 6 BG-01-4000-01-0800|物\*  
 25. 6 BG-01-2020-01-0200|人\*  
 26. 6 BG-01-2000-06-0800|人  
 27. 6 BG-01-1000-03-0201|物  
 28. 5 CH-29-0000-00-1303|神奈備\*  
 29. 5 CH-29-0000-00-1300|神奈備  
 30. 5 BG-03-1200-02-0200|無し  
 31. 5 BG-02-1523-06-0102|流る  
 32. 5 BG-01-5530-10-0300|黄葉+  
 33. 5 BG-01-5530-10-0300|紅葉+  
 34. 5 BG-01-5152-09-0400|春霞  
 35. 5 BG-01-5152-09-0100|霞+  
 36. 5 BG-01-4220-02-0200|衣+  
 37. 5 BG-01-1624-02-0100|春+

High frequency function collocates of tatsuta:  
 1. 111 BG-08-0061-07-0100|の  
 2. 42 BG-08-0061-05-0100|に  
 3. 29 BG-08-0071-01-0100|の+  
 4. 27 BG-08-0065-07-0100|は  
 5. 20 BG-08-0061-10-0100|を  
 6. 19 BG-08-0065-08-0100|も  
 7. 16 BG-09-0050-01-0100|なり\*  
 8. 16 BG-09-0010-04-0300|けり  
 9. 16 BG-09-0010-02-0700|なり  
 10. 16 BG-08-0065-14-0100|や  
 11. 16 BG-08-0064-16-0100|て  
 12. 15 BG-09-0010-01-0100|ず\*  
 13. 15 BG-08-0061-04-0100|と  
 14. 15 BG-03-1200-02-0800|ず  
 15. 13 BG-08-0064-26-0100|ば  
 16. 11 BG-09-0010-03-0200|ぬ\*  
 17. 11 BG-09-0010-01-0101|ぬ\*  
 18. 11 BG-03-1200-02-0900|ぬ  
 19. 10 BG-08-0065-12-0100|ぞ\*  
 20. 10 BG-08-0065-05-0100|ぞ  
 21. 9 BG-09-0010-02-0103|らむ  
 22. 8 BG-09-0010-02-0102|む\*  
 23. 8 BG-03-3012-03-2600|む  
 24. 7 BG-08-0065-14-0400|か  
 25. 5 BG-09-0010-02-0400|らし

Running script with poemID=0298 and translator=katagiri  
Poem content: 立田姫たむくる神のあれはこそ秋のこのはのぬさとちるらめ  
Translation content: 秋をつかさどる龍田姫が旅立ちにあたって手向けをする神があるから、あのように、秋の木の葉が幣を撒くように散っているのだろう。  
output file: ./artifacts/calc\_results/0298-katagiri.csv  
statistics: mode=3; level=3  
OP=19; CT=37;  
bag (E=10, F=1, G=7, U=1, T=18, AddRate=51.35%, UnmatchRate=5.26%);  
alignment (E=10, F=1, G=5, U=3, T=16, AddRate=56.76%, UnmatchRate=15.79%)  
args: translator:katagiri; poem No.0298; weight:(U=-1, G=10, F=13, E=17); gap penalty: 0.01  
 +------- pair No.  
 | +---- matching level (E=Exact, F=Field, G=Group)  
 | | +- pos No.  
 | | | OP decomposition (+) ----------+ +--------------- CT decomposition  
 | | | OP potential sense (\*) ------+ | | +------------- CT potential sense  
 | | | OP token No. --------------+ | | | | +----------- CT token No.  
 | | | OP token ---------------+ | | | | | | +-------- CT token  
 | | | OP WLSP code ---+ | | | | | | | | + CT WLSP code  
 | | | | | | | | | | | | |  
 1 - - - - - <-> 1 秋 BG-01-1624-04-010-A  
 2 - - - - - <-> 2 を BG-08-0061-10-010-A  
 3 - - - - - <-> 3 つかさどる BG-02-3600-03-010-A  
 4 G 11 CH-29-0000-00-1800 立田 1 + <-> + 4 立田 CH-29-0000-14-010-A  
 5 F 02 BG-01-2030-01-1001 姫 1 + <-> + 4 姫 BG-01-2030-01-101-A  
 6 - - - - - <-> 5 が BG-08-0061-02-010-A  
 7 - - - - - <-> + 6 旅 BG-01-3371-01-010-A  
 8 - - - - - <-> + 6 立つ BG-02-1521-06-020-A  
 9 - - - - - <-> 7 にあたって BG-08-0063-18-010-A  
10 E 02 BG-01-5730-02-0100 手 2 + <-> + 8 手 BG-01-5730-02-010-A  
11 E 47 BG-02-1700-01-0400 向ける 2 + <-> + 8 向ける BG-02-1700-01-040-A  
12 - - - - - <-> 9 を BG-08-0061-10-010-A  
13 - - - - - <-> 10 する BG-02-3420-01-010-A  
14 E 02 BG-01-2030-01-0300 神 3 <-> 11 神 BG-01-2030-01-030-A  
15 G 61 BG-08-0061-07-0100 の 4 <-> 12 が BG-08-0061-02-010-A  
16 E 47 BG-02-1200-01-0102 あれ 5 <-> 13 ある BG-02-1200-01-010-A  
17 G 64 BG-08-0064-26-0100 は 6 <-> 14 から BG-08-0064-02-010-A  
18 - - - - - <-> 15 あの BG-03-1000-03-010-A  
19 - - - - - <-> 16 よう BG-03-1300-01-010-A  
20 U 65 BG-08-0065-01-0100 こそ 7 <-> 17 に BG-08-0072-02-010-A  
21 E 02 BG-01-1624-04-0100 秋 8 <-> 18 秋 BG-01-1624-04-010-A  
22 U 61 BG-08-0061-07-0100 の 9 <-> 19 の BG-08-0071-01-010-A  
23 E 02 BG-01-5510-02-0101 こ 10 <-> + 20 木 BG-01-5510-02-010-A  
24 U 61 BG-08-0061-07-0100 の 11 <-> + 20 の BG-08-0071-01-010-A  
25 E 02 BG-01-5530-09-0100 は 12 <-> + 20 葉 BG-01-5530-09-010-A  
26 G 61 BG-08-0061-07-0100 の 13 <-> 21 が BG-08-0061-02-010-A  
27 E 02 BG-01-4500-02-1700 ぬさ 14 <-> 22 幣 BG-01-4500-02-170-A  
28 - - - - - <-> 23 を BG-08-0061-10-010-A  
29 - - - - - <-> 24 撒く BG-02-1552-02-010-A  
30 - - - - - <-> 25 よう BG-03-1300-01-010-A  
31 G 61 BG-08-0061-04-0100 と 15 <-> 26 に BG-08-0061-05-010-A  
32 E 47 BG-02-1552-01-0100 ちる 16 <-> 27 散る BG-02-1552-01-010-A  
33 - - - - - <-> 28 て BG-08-0064-16-010-A  
34 - - - - - <-> 29 いる BG-02-1200-02-010-A  
35 - - - - - <-> 30 の BG-07-0010-03-010-A  
36 - - - - - <-> 31 だ BG-09-0050-01-030-A  
37 E 74 BG-09-0010-02-0103 らめ 17 <-> 32 う BG-09-0010-02-010-A  
Running script with poemID=0298 and translator=kubota  
Poem content: 立田姫たむくる神のあれはこそ秋のこのはのぬさとちるらめ  
Translation content: 竜田姫は、手向けをするべき神があるので、そのつかさどる秋の木の葉が、幣のように散るのであろうよ。  
output file: ./artifacts/calc\_results/0298-kubota.csv  
statistics: mode=3; level=3  
OP=19; CT=30;  
bag (E=10, F=1, G=7, U=1, T=18, AddRate=40.00%, UnmatchRate=5.26%);  
alignment (E=10, F=1, G=4, U=4, T=15, AddRate=50.00%, UnmatchRate=21.05%)  
args: translator:kubota; poem No.0298; weight:(U=-1, G=10, F=13, E=17); gap penalty: 0.01  
 +------- pair No.  
 | +---- matching level (E=Exact, F=Field, G=Group)  
 | | +- pos No.  
 | | | OP decomposition (+) ----------+ +--------------- CT decomposition  
 | | | OP potential sense (\*) ------+ | | +------------- CT potential sense  
 | | | OP token No. --------------+ | | | | +----------- CT token No.  
 | | | OP token ---------------+ | | | | | | +-------- CT token  
 | | | OP WLSP code ---+ | | | | | | | | + CT WLSP code  
 | | | | | | | | | | | | |  
 1 G 11 CH-29-0000-00-1800 立田 1 + <-> + 1 立田 CH-29-0000-14-010-A  
 2 F 02 BG-01-2030-01-1001 姫 1 + <-> + 1 姫 BG-01-2030-01-101-A  
 3 - - - - - <-> 2 は BG-08-0065-07-010-A  
 4 E 02 BG-01-5730-02-0100 手 2 + <-> + 3 手 BG-01-5730-02-010-A  
 5 E 47 BG-02-1700-01-0400 向ける 2 + <-> + 3 向ける BG-02-1700-01-040-A  
 6 - - - - - <-> 4 を BG-08-0061-10-010-A  
 7 - - - - - <-> 5 する BG-02-3420-01-010-A  
 8 - - - - - <-> 6 べし BG-04-3110-02-140-A  
 9 E 02 BG-01-2030-01-0300 神 3 <-> 7 神 BG-01-2030-01-030-A  
10 G 61 BG-08-0061-07-0100 の 4 <-> 8 が BG-08-0061-02-010-A  
11 E 47 BG-02-1200-01-0102 あれ 5 <-> 9 ある BG-02-1200-01-010-A  
12 - - - - - <-> 10 ので BG-01-1113-01-130-A  
13 - - - - - <-> 11 その BG-03-1000-02-010-A  
14 - - - - - <-> 12 つかさどる BG-02-3600-03-010-A  
15 - 64 BG-08-0064-26-0100 は 6 <-> - - -   
16 - 65 BG-08-0065-01-0100 こそ 7 <-> - - -   
17 E 02 BG-01-1624-04-0100 秋 8 <-> 13 秋 BG-01-1624-04-010-A  
18 U 61 BG-08-0061-07-0100 の 9 <-> 14 の BG-08-0071-01-010-A  
19 E 02 BG-01-5510-02-0101 こ 10 <-> + 15 木 BG-01-5510-02-010-A  
20 U 61 BG-08-0061-07-0100 の 11 <-> + 15 の BG-08-0071-01-010-A  
21 E 02 BG-01-5530-09-0100 は 12 <-> + 15 葉 BG-01-5530-09-010-A  
22 G 61 BG-08-0061-07-0100 の 13 <-> 16 が BG-08-0061-02-010-A  
23 E 02 BG-01-4500-02-1700 ぬさ 14 <-> 17 幣 BG-01-4500-02-170-A  
24 - - - - - <-> 18 の BG-08-0071-01-010-A  
25 - - - - - <-> 19 よう BG-03-1300-01-010-A  
26 G 61 BG-08-0061-04-0100 と 15 <-> 20 に BG-08-0061-05-010-A  
27 E 47 BG-02-1552-01-0100 ちる 16 <-> 21 散る BG-02-1552-01-010-A  
28 - - - - - <-> 22 の BG-07-0010-03-010-A  
29 - - - - - <-> 23 だ BG-09-0050-01-030-A  
30 - - - - - <-> 24 ある BG-02-1200-01-010-A  
31 E 74 BG-09-0010-02-0103 らめ 17 <-> 25 う BG-09-0010-02-010-A  
32 - - - - - <-> 26 よ BG-08-0069-30-010-A  
Running script with poemID=0298 and translator=matsuda  
Poem content: 立田姫たむくる神のあれはこそ秋のこのはのぬさとちるらめ  
Translation content: 竜田姫には、旅中、供え物をささげる道祖神があるからこそ、秋の紅葉した葉が、幣のように散っているのだろう。  
output file: ./artifacts/calc\_results/0298-matsuda.csv  
statistics: mode=3; level=3  
OP=19; CT=32;  
bag (E=7, F=2, G=7, U=3, T=16, AddRate=50.00%, UnmatchRate=15.79%);  
alignment (E=7, F=2, G=5, U=5, T=14, AddRate=56.25%, UnmatchRate=26.32%)  
args: translator:matsuda; poem No.0298; weight:(U=-1, G=10, F=13, E=17); gap penalty: 0.01  
 +------- pair No.  
 | +---- matching level (E=Exact, F=Field, G=Group)  
 | | +- pos No.  
 | | | OP decomposition (+) ----------+ +--------------- CT decomposition  
 | | | OP potential sense (\*) ------+ | | +------------- CT potential sense  
 | | | OP token No. --------------+ | | | | +----------- CT token No.  
 | | | OP token ---------------+ | | | | | | +-------- CT token  
 | | | OP WLSP code ---+ | | | | | | | | + CT WLSP code  
 | | | | | | | | | | | | |  
 1 G 11 CH-29-0000-00-1800 立田 1 + <-> + 1 立田 CH-29-0000-14-010-A  
 2 F 02 BG-01-2030-01-1001 姫 1 + <-> + 1 姫 BG-01-2030-01-101-A  
 3 - - - - - <-> 2 に BG-08-0061-05-010-A  
 4 - - - - - <-> 3 は BG-08-0065-07-010-A  
 5 - - - - - <-> 4 旅 BG-01-3371-01-010-A  
 6 - - - - - <-> 5 中 BG-01-1652-01-010-A  
 7 - - - - - <-> + 6 供える BG-02-3750-03-010-A  
 8 - - - - - <-> + 6 物 BG-01-1010-04-030-A  
 9 - - - - - <-> 7 を BG-08-0061-10-010-A  
10 - 02 BG-01-5730-02-0100 手 2 + <-> - - -   
11 U 47 BG-02-1700-01-0400 向ける 2 + <-> 8 ささげる BG-02-1540-03-160-A  
12 F 02 BG-01-2030-01-0300 神 3 <-> 9 道祖神 BG-01-2030-01-470-A  
13 G 61 BG-08-0061-07-0100 の 4 <-> 10 が BG-08-0061-02-010-A  
14 E 47 BG-02-1200-01-0102 あれ 5 <-> 11 ある BG-02-1200-01-010-A  
15 G 64 BG-08-0064-26-0100 は 6 <-> 12 から BG-08-0064-02-010-A  
16 E 65 BG-08-0065-01-0100 こそ 7 <-> 13 こそ BG-08-0065-01-010-A  
17 E 02 BG-01-1624-04-0100 秋 8 <-> 14 秋 BG-01-1624-04-010-A  
18 U 61 BG-08-0061-07-0100 の 9 <-> 15 の BG-08-0071-01-010-A  
19 U 02 BG-01-5510-02-0101 こ 10 <-> 16 紅葉 BG-01-5530-10-030-A  
20 - - - - - <-> 17 する BG-02-3420-01-010-A  
21 - - - - - <-> 18 た BG-09-0010-04-010-A  
22 - 61 BG-08-0061-07-0100 の 11 <-> - - -   
23 E 02 BG-01-5530-09-0100 は 12 <-> 19 葉 BG-01-5530-09-010-A  
24 G 61 BG-08-0061-07-0100 の 13 <-> 20 が BG-08-0061-02-010-A  
25 E 02 BG-01-4500-02-1700 ぬさ 14 <-> 21 幣 BG-01-4500-02-170-A  
26 - - - - - <-> 22 の BG-08-0071-01-010-A  
27 - - - - - <-> 23 よう BG-03-1300-01-010-A  
28 G 61 BG-08-0061-04-0100 と 15 <-> 24 に BG-08-0061-05-010-A  
29 E 47 BG-02-1552-01-0100 ちる 16 <-> 25 散る BG-02-1552-01-010-A  
30 - - - - - <-> 26 て BG-08-0064-16-010-A  
31 - - - - - <-> 27 いる BG-02-1200-02-010-A  
32 - - - - - <-> 28 の BG-07-0010-03-010-A  
33 - - - - - <-> 29 だ BG-09-0050-01-030-A  
34 E 74 BG-09-0010-02-0103 らめ 17 <-> 30 う BG-09-0010-02-010-A  
Running script with poemID=0298 and translator=okumura  
Poem content: 立田姫たむくる神のあれはこそ秋のこのはのぬさとちるらめ  
Translation content: 秋も終りに近づき、秋の女神の龍田姫がお帰りになる。姫が道中の安全を祈って手向けをなさる神があるからこそだろう、紅葉が幣のように散っている。  
output file: ./artifacts/calc\_results/0298-okumura.csv  
statistics: mode=3; level=3  
OP=19; CT=44;  
bag (E=10, F=0, G=8, U=1, T=18, AddRate=59.09%, UnmatchRate=5.26%);  
alignment (E=7, F=1, G=6, U=5, T=14, AddRate=68.18%, UnmatchRate=26.32%)  
args: translator:okumura; poem No.0298; weight:(U=-1, G=10, F=13, E=17); gap penalty: 0.01  
 +------- pair No.  
 | +---- matching level (E=Exact, F=Field, G=Group)  
 | | +- pos No.  
 | | | OP decomposition (+) ----------+ +--------------- CT decomposition  
 | | | OP potential sense (\*) ------+ | | +------------- CT potential sense  
 | | | OP token No. --------------+ | | | | +----------- CT token No.  
 | | | OP token ---------------+ | | | | | | +-------- CT token  
 | | | OP WLSP code ---+ | | | | | | | | + CT WLSP code  
 | | | | | | | | | | | | |  
 1 - - - - - <-> 1 秋 BG-01-1624-04-010-A  
 2 - - - - - <-> 2 も BG-08-0065-08-010-A  
 3 - - - - - <-> 3 終り BG-01-1651-02-050-A  
 4 - - - - - <-> 4 に BG-08-0061-05-010-A  
 5 - - - - - <-> 5 近づく BG-02-1560-11-010-A  
 6 - - - - - <-> 6 秋 BG-01-1624-04-010-A  
 7 - - - - - <-> 7 の BG-08-0071-01-010-A  
 8 - - - - - <-> 8 女神 BG-01-2030-01-100-A  
 9 - - - - - <-> 9 の BG-08-0071-01-010-A  
10 G 11 CH-29-0000-00-1800 立田 1 + <-> + 10 立田 CH-29-0000-14-010-A  
11 F 02 BG-01-2030-01-1001 姫 1 + <-> + 10 姫 BG-01-2030-01-101-A  
12 - - - - - <-> 11 が BG-08-0061-02-010-A  
13 - - - - - <-> 12 お BG-05-0070-01-010-A  
14 - - - - - <-> 13 帰り BG-01-1528-02-020-A  
15 - - - - - <-> 14 に BG-08-0061-05-010-A  
16 - - - - - <-> 15 なる BG-02-1220-01-030-A  
17 - - - - - <-> 16 姫 BG-01-2020-05-120-A  
18 - - - - - <-> 17 が BG-08-0061-02-010-A  
19 - - - - - <-> 18 道中 BG-01-1652-01-050-A  
20 - - - - - <-> 19 の BG-08-0071-01-010-A  
21 - - - - - <-> 20 安全 BG-03-1340-01-010-A  
22 - - - - - <-> 21 を BG-08-0061-10-010-A  
23 - - - - - <-> 22 祈る BG-02-3041-05-020-A  
24 - - - - - <-> 23 て BG-08-0064-16-010-A  
25 E 02 BG-01-5730-02-0100 手 2 + <-> + 24 手 BG-01-5730-02-010-A  
26 E 47 BG-02-1700-01-0400 向ける 2 + <-> + 24 向ける BG-02-1700-01-040-A  
27 - - - - - <-> 25 を BG-08-0061-10-010-A  
28 - - - - - <-> 26 なさる BG-02-3420-02-030-A  
29 E 02 BG-01-2030-01-0300 神 3 <-> 27 神 BG-01-2030-01-030-A  
30 G 61 BG-08-0061-07-0100 の 4 <-> 28 が BG-08-0061-02-010-A  
31 E 47 BG-02-1200-01-0102 あれ 5 <-> 29 ある BG-02-1200-01-010-A  
32 G 64 BG-08-0064-26-0100 は 6 <-> 30 から BG-08-0064-02-010-A  
33 E 65 BG-08-0065-01-0100 こそ 7 <-> 31 こそ BG-08-0065-01-010-A  
34 - - - - - <-> 32 だ BG-09-0050-01-030-A  
35 - - - - - <-> 33 う BG-09-0010-02-010-A  
36 - 02 BG-01-1624-04-0100 秋 8 <-> - - -   
37 - 61 BG-08-0061-07-0100 の 9 <-> - - -   
38 - 02 BG-01-5510-02-0101 こ 10 <-> - - -   
39 - 61 BG-08-0061-07-0100 の 11 <-> - - -   
40 G 02 BG-01-5530-09-0100 は 12 <-> 34 紅葉 BG-01-5530-10-030-A  
41 G 61 BG-08-0061-07-0100 の 13 <-> 35 が BG-08-0061-02-010-A  
42 E 02 BG-01-4500-02-1700 ぬさ 14 <-> 36 幣 BG-01-4500-02-170-A  
43 - - - - - <-> 37 の BG-08-0071-01-010-A  
44 - - - - - <-> 38 よう BG-03-1300-01-010-A  
45 G 61 BG-08-0061-04-0100 と 15 <-> 39 に BG-08-0061-05-010-A  
46 E 47 BG-02-1552-01-0100 ちる 16 <-> 40 散る BG-02-1552-01-010-A  
47 - - - - - <-> 41 て BG-08-0064-16-010-A  
48 - - - - - <-> 42 いる BG-02-1200-02-010-A  
49 - 74 BG-09-0010-02-0103 らめ 17 <-> - - -   
Running script with poemID=0298 and translator=takeoka  
Poem content: 立田姫たむくる神のあれはこそ秋のこのはのぬさとちるらめ  
Translation content: 竜田姫が、旅にあって手向ける神があるからこそ、それであのように秋の木の葉が、幣として散っているのであろう。  
output file: ./artifacts/calc\_results/0298-takeoka.csv  
statistics: mode=3; level=3  
OP=19; CT=32;  
bag (E=11, F=1, G=7, U=0, T=19, AddRate=40.62%, UnmatchRate=0.00%);  
alignment (E=11, F=1, G=4, U=3, T=16, AddRate=50.00%, UnmatchRate=15.79%)  
args: translator:takeoka; poem No.0298; weight:(U=-1, G=10, F=13, E=17); gap penalty: 0.01  
 +------- pair No.  
 | +---- matching level (E=Exact, F=Field, G=Group)  
 | | +- pos No.  
 | | | OP decomposition (+) ----------+ +--------------- CT decomposition  
 | | | OP potential sense (\*) ------+ | | +------------- CT potential sense  
 | | | OP token No. --------------+ | | | | +----------- CT token No.  
 | | | OP token ---------------+ | | | | | | +-------- CT token  
 | | | OP WLSP code ---+ | | | | | | | | + CT WLSP code  
 | | | | | | | | | | | | |  
 1 G 11 CH-29-0000-00-1800 立田 1 + <-> + 1 立田 CH-29-0000-14-010-A  
 2 F 02 BG-01-2030-01-1001 姫 1 + <-> + 1 姫 BG-01-2030-01-101-A  
 3 - - - - - <-> 2 が BG-08-0061-02-010-A  
 4 - - - - - <-> 3 旅 BG-01-3371-01-010-A  
 5 - - - - - <-> 4 に BG-08-0061-05-010-A  
 6 - - - - - <-> 5 ある BG-02-1200-01-010-A  
 7 - - - - - <-> 6 て BG-08-0064-16-010-A  
 8 E 02 BG-01-5730-02-0100 手 2 + <-> + 7 手 BG-01-5730-02-010-A  
 9 E 47 BG-02-1700-01-0400 向ける 2 + <-> + 7 向ける BG-02-1700-01-040-A  
10 E 02 BG-01-2030-01-0300 神 3 <-> 8 神 BG-01-2030-01-030-A  
11 G 61 BG-08-0061-07-0100 の 4 <-> 9 が BG-08-0061-02-010-A  
12 E 47 BG-02-1200-01-0102 あれ 5 <-> 10 ある BG-02-1200-01-010-A  
13 G 64 BG-08-0064-26-0100 は 6 <-> 11 から BG-08-0064-02-010-A  
14 E 65 BG-08-0065-01-0100 こそ 7 <-> 12 こそ BG-08-0065-01-010-A  
15 - - - - - <-> 13 それ BG-01-1000-01-050-A  
16 - - - - - <-> 14 で BG-08-0061-03-010-A  
17 - - - - - <-> 15 あのように BG-03-1140-01-100-A  
18 E 02 BG-01-1624-04-0100 秋 8 <-> 16 秋 BG-01-1624-04-010-A  
19 U 61 BG-08-0061-07-0100 の 9 <-> 17 の BG-08-0071-01-010-A  
20 E 02 BG-01-5510-02-0101 こ 10 <-> + 18 木 BG-01-5510-02-010-A  
21 U 61 BG-08-0061-07-0100 の 11 <-> + 18 の BG-08-0071-01-010-A  
22 E 02 BG-01-5530-09-0100 は 12 <-> + 18 葉 BG-01-5530-09-010-A  
23 G 61 BG-08-0061-07-0100 の 13 <-> 19 が BG-08-0061-02-010-A  
24 E 02 BG-01-4500-02-1700 ぬさ 14 <-> 20 幣 BG-01-4500-02-170-A  
25 U 61 BG-08-0061-04-0100 と 15 <-> 21 として BG-08-0063-13-010-A  
26 E 47 BG-02-1552-01-0100 ちる 16 <-> 22 散る BG-02-1552-01-010-A  
27 - - - - - <-> 23 て BG-08-0064-16-010-A  
28 - - - - - <-> 24 いる BG-02-1200-02-010-A  
29 - - - - - <-> 25 の BG-07-0010-03-010-A  
30 - - - - - <-> 26 だ BG-09-0050-01-030-A  
31 - - - - - <-> 27 ある BG-02-1200-01-010-A  
32 E 74 BG-09-0010-02-0103 らめ 17 <-> 28 う BG-09-0010-02-010-A  
Running script with poemID=0298 and translator=kaneko  
Poem content: 立田姫たむくる神のあれはこそ秋のこのはのぬさとちるらめ  
Translation content: 立田姫は秋の神だが、それすら暮れて行かれる折には、お手向けなさる道の神様があるからこそ、このように、御自身お染めなされた秋の木の葉が、手向けの幣のように散るのであろう。  
output file: ./artifacts/calc\_results/0298-kaneko.csv  
statistics: mode=3; level=3  
OP=19; CT=55;  
bag (E=11, F=1, G=7, U=0, T=19, AddRate=65.45%, UnmatchRate=0.00%);  
alignment (E=11, F=1, G=5, U=2, T=17, AddRate=69.09%, UnmatchRate=10.53%)  
args: translator:kaneko; poem No.0298; weight:(U=-1, G=10, F=13, E=17); gap penalty: 0.01  
 +------- pair No.  
 | +---- matching level (E=Exact, F=Field, G=Group)  
 | | +- pos No.  
 | | | OP decomposition (+) ----------+ +--------------- CT decomposition  
 | | | OP potential sense (\*) ------+ | | +------------- CT potential sense  
 | | | OP token No. --------------+ | | | | +----------- CT token No.  
 | | | OP token ---------------+ | | | | | | +-------- CT token  
 | | | OP WLSP code ---+ | | | | | | | | + CT WLSP code  
 | | | | | | | | | | | | |  
 1 G 11 CH-29-0000-00-1800 立田 1 + <-> + 1 立田 CH-29-0000-14-010-A  
 2 F 02 BG-01-2030-01-1001 姫 1 + <-> + 1 姫 BG-01-2030-01-101-A  
 3 - - - - - <-> 2 は BG-08-0065-07-010-A  
 4 - - - - - <-> 3 秋 BG-01-1624-04-010-A  
 5 - - - - - <-> 4 の BG-08-0071-01-010-A  
 6 - - - - - <-> 5 神 BG-01-2030-01-030-A  
 7 - - - - - <-> 6 だ BG-09-0050-01-030-A  
 8 - - - - - <-> 7 が BG-08-0064-04-010-A  
 9 - - - - - <-> 8 それすら BG-01-1990-05-051-A  
10 - - - - - <-> 9 暮れる BG-02-1600-02-040-A  
11 - - - - - <-> 10 て BG-08-0064-16-010-A  
12 - - - - - <-> 11 行く BG-02-1527-03-010-A  
13 - - - - - <-> 12 れる BG-02-1110-02-010-A  
14 - - - - - <-> 13 折 BG-01-1612-02-020-A  
15 - - - - - <-> 14 に BG-08-0061-05-010-A  
16 - - - - - <-> 15 は BG-08-0065-07-010-A  
17 - - - - - <-> 16 お BG-05-0070-01-010-A  
18 E 02 BG-01-5730-02-0100 手 2 + <-> + 17 手 BG-01-5730-02-010-A  
19 E 47 BG-02-1700-01-0400 向ける 2 + <-> + 17 向ける BG-02-1700-01-040-A  
20 - - - - - <-> 18 なさる BG-02-3420-02-030-A  
21 - - - - - <-> 19 道 BG-01-4710-01-010-A  
22 - - - - - <-> 20 の BG-08-0071-01-010-A  
23 E 02 BG-01-2030-01-0300 神 3 <-> + 21 神 BG-01-2030-01-030-A  
24 - - - - - <-> + 21 様 BG-07-0070-01-010-A  
25 G 61 BG-08-0061-07-0100 の 4 <-> 22 が BG-08-0061-02-010-A  
26 E 47 BG-02-1200-01-0102 あれ 5 <-> 23 ある BG-02-1200-01-010-A  
27 G 64 BG-08-0064-26-0100 は 6 <-> 24 から BG-08-0064-02-010-A  
28 E 65 BG-08-0065-01-0100 こそ 7 <-> 25 こそ BG-08-0065-01-010-A  
29 - - - - - <-> 26 この BG-03-1000-01-010-A  
30 - - - - - <-> 27 よう BG-03-1300-01-010-A  
31 - - - - - <-> 28 に BG-08-0072-02-010-A  
32 - - - - - <-> 29 御 BG-05-0070-01-030-A  
33 - - - - - <-> 30 自身 BG-01-2010-01-030-A  
34 - - - - - <-> 31 お BG-05-0070-01-010-A  
35 - - - - - <-> 32 染め BG-01-3842-02-010-A  
36 - - - - - <-> 33 なされる BG-02-3420-02-080-A  
37 - - - - - <-> 34 た BG-09-0010-04-010-A  
38 E 02 BG-01-1624-04-0100 秋 8 <-> 35 秋 BG-01-1624-04-010-A  
39 U 61 BG-08-0061-07-0100 の 9 <-> 36 の BG-08-0071-01-010-A  
40 E 02 BG-01-5510-02-0101 こ 10 <-> + 37 木 BG-01-5510-02-010-A  
41 U 61 BG-08-0061-07-0100 の 11 <-> + 37 の BG-08-0071-01-010-A  
42 E 02 BG-01-5530-09-0100 は 12 <-> + 37 葉 BG-01-5530-09-010-A  
43 G 61 BG-08-0061-07-0100 の 13 <-> 38 が BG-08-0061-02-010-A  
44 - - - - - <-> + 39 手 BG-01-5730-02-010-A  
45 - - - - - <-> + 39 向ける BG-02-1700-01-040-A  
46 - - - - - <-> 40 の BG-08-0071-01-010-A  
47 E 02 BG-01-4500-02-1700 ぬさ 14 <-> 41 幣 BG-01-4500-02-170-A  
48 - - - - - <-> 42 の BG-08-0071-01-010-A  
49 - - - - - <-> 43 よう BG-03-1300-01-010-A  
50 G 61 BG-08-0061-04-0100 と 15 <-> 44 に BG-08-0061-05-010-A  
51 E 47 BG-02-1552-01-0100 ちる 16 <-> 45 散る BG-02-1552-01-010-A  
52 - - - - - <-> 46 の BG-07-0010-03-010-A  
53 - - - - - <-> 47 だ BG-09-0050-01-030-A  
54 - - - - - <-> 48 ある BG-02-1200-01-010-A  
55 E 74 BG-09-0010-02-0103 らめ 17 <-> 49 う BG-09-0010-02-010-A  
Running script with poemID=0298 and translator=kojimaarai  
Poem content: 立田姫たむくる神のあれはこそ秋のこのはのぬさとちるらめ  
Translation content: 竜田姫が手向けをする神さまがおられるからこそ、秋の木の葉のもみじが手向けものとして散るのでしょうね。  
output file: ./artifacts/calc\_results/0298-kojimaarai.csv  
statistics: mode=3; level=3  
OP=19; CT=30;  
bag (E=8, F=2, G=8, U=1, T=18, AddRate=40.00%, UnmatchRate=5.26%);  
alignment (E=8, F=2, G=5, U=4, T=15, AddRate=50.00%, UnmatchRate=21.05%)  
args: translator:kojimaarai; poem No.0298; weight:(U=-1, G=10, F=13, E=17); gap penalty: 0.01  
 +------- pair No.  
 | +---- matching level (E=Exact, F=Field, G=Group)  
 | | +- pos No.  
 | | | OP decomposition (+) ----------+ +--------------- CT decomposition  
 | | | OP potential sense (\*) ------+ | | +------------- CT potential sense  
 | | | OP token No. --------------+ | | | | +----------- CT token No.  
 | | | OP token ---------------+ | | | | | | +-------- CT token  
 | | | OP WLSP code ---+ | | | | | | | | + CT WLSP code  
 | | | | | | | | | | | | |  
 1 G 11 CH-29-0000-00-1800 立田 1 + <-> + 1 立田 CH-29-0000-14-010-A  
 2 F 02 BG-01-2030-01-1001 姫 1 + <-> + 1 姫 BG-01-2030-01-101-A  
 3 - - - - - <-> 2 が BG-08-0061-02-010-A  
 4 E 02 BG-01-5730-02-0100 手 2 + <-> + 3 手 BG-01-5730-02-010-A  
 5 E 47 BG-02-1700-01-0400 向ける 2 + <-> + 3 向ける BG-02-1700-01-040-A  
 6 - - - - - <-> 4 を BG-08-0061-10-010-A  
 7 - - - - - <-> 5 する BG-02-3420-01-010-A  
 8 F 02 BG-01-2030-01-0300 神 3 <-> 6 神さま BG-01-2030-01-031-A  
 9 G 61 BG-08-0061-07-0100 の 4 <-> 7 が BG-08-0061-02-010-A  
10 G 47 BG-02-1200-01-0102 あれ 5 <-> 8 おる BG-02-1200-02-020-A  
11 - - - - - <-> 9 れる BG-02-1110-02-010-A  
12 G 64 BG-08-0064-26-0100 は 6 <-> 10 から BG-08-0064-02-010-A  
13 E 65 BG-08-0065-01-0100 こそ 7 <-> 11 こそ BG-08-0065-01-010-A  
14 E 02 BG-01-1624-04-0100 秋 8 <-> 12 秋 BG-01-1624-04-010-A  
15 U 61 BG-08-0061-07-0100 の 9 <-> 13 の BG-08-0071-01-010-A  
16 E 02 BG-01-5510-02-0101 こ 10 <-> + 14 木 BG-01-5510-02-010-A  
17 U 61 BG-08-0061-07-0100 の 11 <-> + 14 の BG-08-0071-01-010-A  
18 E 02 BG-01-5530-09-0100 は 12 <-> + 14 葉 BG-01-5530-09-010-A  
19 - - - - - <-> 15 の BG-08-0071-01-010-A  
20 - - - - - <-> 16 もみじ BG-01-5530-10-030-A  
21 G 61 BG-08-0061-07-0100 の 13 <-> 17 が BG-08-0061-02-010-A  
22 - - - - - <-> + 18 手 BG-01-5730-02-010-A  
23 - - - - - <-> + 18 向ける BG-02-1700-01-040-A  
24 - - - - - <-> 19 もの BG-01-1000-03-010-A  
25 - 02 BG-01-4500-02-1700 ぬさ 14 <-> - - -   
26 U 61 BG-08-0061-04-0100 と 15 <-> 20 として BG-08-0063-13-010-A  
27 E 47 BG-02-1552-01-0100 ちる 16 <-> 21 散る BG-02-1552-01-010-A  
28 - - - - - <-> 22 の BG-07-0010-03-010-A  
29 - - - - - <-> 23 です BG-09-0050-01-030-A  
30 E 74 BG-09-0010-02-0103 らめ 17 <-> 24 う BG-09-0010-02-010-A  
31 - - - - - <-> 25 ね BG-08-0069-16-010-A  
Running script with poemID=0298 and translator=komachiya  
Poem content: 立田姫たむくる神のあれはこそ秋のこのはのぬさとちるらめ  
Translation content: 秋の末近くなって帰り道についた龍田姫が、道中の無事を願って手向けをする神があるからこそ、秋の木の葉が幣となって散っているのだろう。  
output file: ./artifacts/calc\_results/0298-komachiya.csv  
statistics: mode=3; level=3  
OP=19; CT=45;  
bag (E=12, F=1, G=6, U=0, T=19, AddRate=57.78%, UnmatchRate=0.00%);  
alignment (E=12, F=1, G=4, U=2, T=17, AddRate=62.22%, UnmatchRate=10.53%)  
args: translator:komachiya; poem No.0298; weight:(U=-1, G=10, F=13, E=17); gap penalty: 0.01  
 +------- pair No.  
 | +---- matching level (E=Exact, F=Field, G=Group)  
 | | +- pos No.  
 | | | OP decomposition (+) ----------+ +--------------- CT decomposition  
 | | | OP potential sense (\*) ------+ | | +------------- CT potential sense  
 | | | OP token No. --------------+ | | | | +----------- CT token No.  
 | | | OP token ---------------+ | | | | | | +-------- CT token  
 | | | OP WLSP code ---+ | | | | | | | | + CT WLSP code  
 | | | | | | | | | | | | |  
 1 - - - - - <-> 1 秋 BG-01-1624-04-010-A  
 2 - - - - - <-> 2 の BG-08-0071-01-010-A  
 3 - - - - - <-> 3 末 BG-01-1651-06-010-A  
 4 - - - - - <-> 4 近い BG-03-1920-06-070-A  
 5 - - - - - <-> 5 なる BG-02-1220-01-030-A  
 6 - - - - - <-> 6 て BG-08-0064-16-010-A  
 7 - - - - - <-> + 7 帰る BG-02-1527-04-010-A  
 8 - - - - - <-> + 7 道 BG-01-4710-01-010-A  
 9 - - - - - <-> 8 に BG-08-0061-05-010-A  
10 - - - - - <-> 9 つく BG-02-1560-02-010-A  
11 - - - - - <-> 10 た BG-09-0010-04-010-A  
12 G 11 CH-29-0000-00-1800 立田 1 + <-> + 11 立田 CH-29-0000-14-010-A  
13 F 02 BG-01-2030-01-1001 姫 1 + <-> + 11 姫 BG-01-2030-01-101-A  
14 - - - - - <-> 12 が BG-08-0061-02-010-A  
15 - - - - - <-> 13 道中 BG-01-1652-01-050-A  
16 - - - - - <-> 14 の BG-08-0071-01-010-A  
17 - - - - - <-> 15 無事 BG-03-1340-01-040-A  
18 - - - - - <-> 16 を BG-08-0061-10-010-A  
19 - - - - - <-> 17 願う BG-02-3041-04-010-A  
20 - - - - - <-> 18 て BG-08-0064-16-010-A  
21 E 02 BG-01-5730-02-0100 手 2 + <-> + 19 手 BG-01-5730-02-010-A  
22 E 47 BG-02-1700-01-0400 向ける 2 + <-> + 19 向ける BG-02-1700-01-040-A  
23 - - - - - <-> 20 を BG-08-0061-10-010-A  
24 - - - - - <-> 21 する BG-02-3420-01-010-A  
25 E 02 BG-01-2030-01-0300 神 3 <-> 22 神 BG-01-2030-01-030-A  
26 G 61 BG-08-0061-07-0100 の 4 <-> 23 が BG-08-0061-02-010-A  
27 E 47 BG-02-1200-01-0102 あれ 5 <-> 24 ある BG-02-1200-01-010-A  
28 G 64 BG-08-0064-26-0100 は 6 <-> 25 から BG-08-0064-02-010-A  
29 E 65 BG-08-0065-01-0100 こそ 7 <-> 26 こそ BG-08-0065-01-010-A  
30 E 02 BG-01-1624-04-0100 秋 8 <-> 27 秋 BG-01-1624-04-010-A  
31 U 61 BG-08-0061-07-0100 の 9 <-> 28 の BG-08-0071-01-010-A  
32 E 02 BG-01-5510-02-0101 こ 10 <-> + 29 木 BG-01-5510-02-010-A  
33 U 61 BG-08-0061-07-0100 の 11 <-> + 29 の BG-08-0071-01-010-A  
34 E 02 BG-01-5530-09-0100 は 12 <-> + 29 葉 BG-01-5530-09-010-A  
35 G 61 BG-08-0061-07-0100 の 13 <-> 30 が BG-08-0061-02-010-A  
36 E 02 BG-01-4500-02-1700 ぬさ 14 <-> 31 幣 BG-01-4500-02-170-A  
37 E 61 BG-08-0061-04-0100 と 15 <-> 32 と BG-08-0061-04-010-A  
38 - - - - - <-> 33 なる BG-02-1220-01-030-A  
39 - - - - - <-> 34 て BG-08-0064-16-010-A  
40 E 47 BG-02-1552-01-0100 ちる 16 <-> 35 散る BG-02-1552-01-010-A  
41 - - - - - <-> 36 て BG-08-0064-16-010-A  
42 - - - - - <-> 37 いる BG-02-1200-02-010-A  
43 - - - - - <-> 38 の BG-07-0010-03-010-A  
44 - - - - - <-> 39 だ BG-09-0050-01-030-A  
45 E 74 BG-09-0010-02-0103 らめ 17 <-> 40 う BG-09-0010-02-010-A  
Running script with poemID=0298 and translator=kyusojin  
Poem content: 立田姫たむくる神のあれはこそ秋のこのはのぬさとちるらめ  
Translation content: 秋も終りに近づき竜田姫がお帰りになる際に、旅の安全を祈って手向けられる神があるからこそ、秋の紅葉が幣となって散っているのであろう。  
output file: ./artifacts/calc\_results/0298-kyusojin.csv  
statistics: mode=3; level=3  
OP=19; CT=43;  
bag (E=10, F=1, G=7, U=1, T=18, AddRate=58.14%, UnmatchRate=5.26%);  
alignment (E=10, F=1, G=5, U=3, T=16, AddRate=62.79%, UnmatchRate=15.79%)  
args: translator:kyusojin; poem No.0298; weight:(U=-1, G=10, F=13, E=17); gap penalty: 0.01  
 +------- pair No.  
 | +---- matching level (E=Exact, F=Field, G=Group)  
 | | +- pos No.  
 | | | OP decomposition (+) ----------+ +--------------- CT decomposition  
 | | | OP potential sense (\*) ------+ | | +------------- CT potential sense  
 | | | OP token No. --------------+ | | | | +----------- CT token No.  
 | | | OP token ---------------+ | | | | | | +-------- CT token  
 | | | OP WLSP code ---+ | | | | | | | | + CT WLSP code  
 | | | | | | | | | | | | |  
 1 - - - - - <-> 1 秋 BG-01-1624-04-010-A  
 2 - - - - - <-> 2 も BG-08-0065-08-010-A  
 3 - - - - - <-> 3 終り BG-01-1651-02-050-A  
 4 - - - - - <-> 4 に BG-08-0061-05-010-A  
 5 - - - - - <-> 5 近づく BG-02-1560-11-010-A  
 6 G 11 CH-29-0000-00-1800 立田 1 + <-> + 6 立田 CH-29-0000-14-010-A  
 7 F 02 BG-01-2030-01-1001 姫 1 + <-> + 6 姫 BG-01-2030-01-101-A  
 8 - - - - - <-> 7 が BG-08-0061-02-010-A  
 9 - - - - - <-> 8 お BG-05-0070-01-010-A  
10 - - - - - <-> 9 帰り BG-01-1528-02-020-A  
11 - - - - - <-> 10 に BG-08-0061-05-010-A  
12 - - - - - <-> 11 なる BG-02-1220-01-030-A  
13 - - - - - <-> 12 際 BG-01-1612-02-030-A  
14 - - - - - <-> 13 に BG-08-0061-05-010-A  
15 - - - - - <-> 14 旅 BG-01-3371-01-010-A  
16 - - - - - <-> 15 の BG-08-0071-01-010-A  
17 - - - - - <-> 16 安全 BG-03-1340-01-010-A  
18 - - - - - <-> 17 を BG-08-0061-10-010-A  
19 - - - - - <-> 18 祈る BG-02-3041-05-020-A  
20 - - - - - <-> 19 て BG-08-0064-16-010-A  
21 E 02 BG-01-5730-02-0100 手 2 + <-> + 20 手 BG-01-5730-02-010-A  
22 E 47 BG-02-1700-01-0400 向ける 2 + <-> + 20 向ける BG-02-1700-01-040-A  
23 - - - - - <-> 21 られる BG-02-1110-02-020-A  
24 E 02 BG-01-2030-01-0300 神 3 <-> 22 神 BG-01-2030-01-030-A  
25 G 61 BG-08-0061-07-0100 の 4 <-> 23 が BG-08-0061-02-010-A  
26 E 47 BG-02-1200-01-0102 あれ 5 <-> 24 ある BG-02-1200-01-010-A  
27 G 64 BG-08-0064-26-0100 は 6 <-> 25 から BG-08-0064-02-010-A  
28 E 65 BG-08-0065-01-0100 こそ 7 <-> 26 こそ BG-08-0065-01-010-A  
29 E 02 BG-01-1624-04-0100 秋 8 <-> 27 秋 BG-01-1624-04-010-A  
30 - 61 BG-08-0061-07-0100 の 9 <-> - - -   
31 - 02 BG-01-5510-02-0101 こ 10 <-> - - -   
32 U 61 BG-08-0061-07-0100 の 11 <-> 28 の BG-08-0071-01-010-A  
33 G 02 BG-01-5530-09-0100 は 12 <-> 29 紅葉 BG-01-5530-10-030-A  
34 G 61 BG-08-0061-07-0100 の 13 <-> 30 が BG-08-0061-02-010-A  
35 E 02 BG-01-4500-02-1700 ぬさ 14 <-> 31 幣 BG-01-4500-02-170-A  
36 E 61 BG-08-0061-04-0100 と 15 <-> 32 と BG-08-0061-04-010-A  
37 - - - - - <-> 33 なる BG-02-1220-01-030-A  
38 - - - - - <-> 34 て BG-08-0064-16-010-A  
39 E 47 BG-02-1552-01-0100 ちる 16 <-> 35 散る BG-02-1552-01-010-A  
40 - - - - - <-> 36 て BG-08-0064-16-010-A  
41 - - - - - <-> 37 いる BG-02-1200-02-010-A  
42 - - - - - <-> 38 の BG-07-0010-03-010-A  
43 - - - - - <-> 39 だ BG-09-0050-01-030-A  
44 - - - - - <-> 40 ある BG-02-1200-01-010-A  
45 E 74 BG-09-0010-02-0103 らめ 17 <-> 41 う BG-09-0010-02-010-A  
Running script with poemID=0298 and translator=ozawa  
Poem content: 立田姫たむくる神のあれはこそ秋のこのはのぬさとちるらめ  
Translation content: もはや秋の終りで龍田姫が帰り道にお着きになった。山の紅葉が色とりどりの幣となって散っているのは、姫が道の神様にそれを供えていらっしゃるのだな。  
output file: ./artifacts/calc\_results/0298-ozawa.csv  
statistics: mode=3; level=3  
OP=19; CT=47;  
bag (E=5, F=1, G=10, U=3, T=16, AddRate=65.96%, UnmatchRate=15.79%);  
alignment (E=3, F=1, G=3, U=12, T=7, AddRate=85.11%, UnmatchRate=63.16%)  
args: translator:ozawa; poem No.0298; weight:(U=-1, G=10, F=13, E=17); gap penalty: 0.01  
 +------- pair No.  
 | +---- matching level (E=Exact, F=Field, G=Group)  
 | | +- pos No.  
 | | | OP decomposition (+) ----------+ +--------------- CT decomposition  
 | | | OP potential sense (\*) ------+ | | +------------- CT potential sense  
 | | | OP token No. --------------+ | | | | +----------- CT token No.  
 | | | OP token ---------------+ | | | | | | +-------- CT token  
 | | | OP WLSP code ---+ | | | | | | | | + CT WLSP code  
 | | | | | | | | | | | | |  
 1 - - - - - <-> 1 もはや BG-03-1650-03-030-A  
 2 - - - - - <-> 2 秋 BG-01-1624-04-010-A  
 3 - - - - - <-> 3 の BG-08-0071-01-010-A  
 4 - - - - - <-> 4 終り BG-01-1651-02-050-A  
 5 - - - - - <-> 5 で BG-08-0061-03-010-A  
 6 G 11 CH-29-0000-00-1800 立田 1 + <-> + 6 立田 CH-29-0000-14-010-A  
 7 F 02 BG-01-2030-01-1001 姫 1 + <-> + 6 姫 BG-01-2030-01-101-A  
 8 - 02 BG-01-5730-02-0100 手 2 + <-> - - -   
 9 - 47 BG-02-1700-01-0400 向ける 2 + <-> - - -   
10 - 02 BG-01-2030-01-0300 神 3 <-> - - -   
11 G 61 BG-08-0061-07-0100 の 4 <-> 7 が BG-08-0061-02-010-A  
12 U 47 BG-02-1200-01-0102 あれ 5 <-> + 8 帰る BG-02-1527-04-010-A  
13 - - - - - <-> + 8 道 BG-01-4710-01-010-A  
14 U 64 BG-08-0064-26-0100 は 6 <-> 9 に BG-08-0061-05-010-A  
15 - - - - - <-> 10 お BG-05-0070-01-010-A  
16 - - - - - <-> 11 着く BG-02-1521-08-010-A  
17 U 65 BG-08-0065-01-0100 こそ 7 <-> 12 に BG-08-0061-05-010-A  
18 - - - - - <-> 13 なる BG-02-1220-01-030-A  
19 - - - - - <-> 14 た BG-09-0010-04-010-A  
20 - - - - - <-> 15 山 BG-01-5240-05-010-A  
21 - 02 BG-01-1624-04-0100 秋 8 <-> - - -   
22 U 61 BG-08-0061-07-0100 の 9 <-> 16 の BG-08-0071-01-010-A  
23 U 02 BG-01-5510-02-0101 こ 10 <-> 17 紅葉 BG-01-5530-10-030-A  
24 G 61 BG-08-0061-07-0100 の 11 <-> 18 が BG-08-0061-02-010-A  
25 U 02 BG-01-5530-09-0100 は 12 <-> + 19 色 BG-01-5020-01-010-A  
26 - - - - - <-> + 19 とりどり BG-03-1300-07-070-A  
27 U 61 BG-08-0061-07-0100 の 13 <-> 20 の BG-08-0071-01-010-A  
28 E 02 BG-01-4500-02-1700 ぬさ 14 <-> 21 幣 BG-01-4500-02-170-A  
29 E 61 BG-08-0061-04-0100 と 15 <-> 22 と BG-08-0061-04-010-A  
30 - - - - - <-> 23 なる BG-02-1220-01-030-A  
31 - - - - - <-> 24 て BG-08-0064-16-010-A  
32 E 47 BG-02-1552-01-0100 ちる 16 <-> 25 散る BG-02-1552-01-010-A  
33 - - - - - <-> 26 て BG-08-0064-16-010-A  
34 - - - - - <-> 27 いる BG-02-1200-02-010-A  
35 - - - - - <-> 28 の BG-07-0010-03-010-A  
36 - - - - - <-> 29 は BG-08-0065-07-010-A  
37 - - - - - <-> 30 姫 BG-01-2020-05-120-A  
38 - - - - - <-> 31 が BG-08-0061-02-010-A  
39 - - - - - <-> 32 道 BG-01-4710-01-010-A  
40 - - - - - <-> 33 の BG-08-0071-01-010-A  
41 - - - - - <-> + 34 神 BG-01-2030-01-030-A  
42 - - - - - <-> + 34 様 BG-07-0070-01-010-A  
43 - - - - - <-> 35 に BG-08-0061-05-010-A  
44 - - - - - <-> 36 それ BG-01-1000-01-050-A  
45 - - - - - <-> 37 を BG-08-0061-10-010-A  
46 - - - - - <-> 38 供える BG-02-3750-03-010-A  
47 - - - - - <-> 39 て BG-08-0064-16-010-A  
48 - - - - - <-> 40 いらっしゃる BG-02-1200-02-030-A  
49 - - - - - <-> 41 の BG-07-0010-03-010-A  
50 - - - - - <-> 42 だ BG-09-0050-01-030-A  
51 - - - - - <-> 43 な BG-08-0069-10-010-A  
52 - 74 BG-09-0010-02-0103 らめ 17 <-> - - -

Asahara, Masayuki, Nao Ikegami, Tai Suzuki, Taro Ichimura, Asuko Kondo, Sachi Kato, and Makoto Yamazaki. 2022. “CHJ-WLSP: Annotation of ‘Word List by Semantic Principles’ Labels for the Corpus of Historical Japanese.” In *Proceedings of the Second Workshop on Language Technologies for Historical and Ancient Languages*, edited by Rachele Sprugnoli and Marco Passarotti, 31–37. Marseille, France: European Language Resources Association.

Baker, Mona. 1996. “The Challenges That Lie Ahead: Corpus-based Translation Studies.” In *Terminology, LSP and Translation: Studies in Language Engineering in Honour of Juan C. Sager*, edited by Harold Somers, 175. Benjamins Translation Library. John Benjamins Publishing Company. <https://doi.org/10.1075/btl.18.17bak>.

Blum-Kulka, Shoshana, and J. House. 1986. “Shifts of Cohesion and Coherence in Translation.” In *Interlingual and Intercultural Communication*, 17–35. Tübingen: Gűnter Narr.

Brooks, Stephen P., and Andrew Gelman. 1998. “General Methods for Monitoring Convergence of Iterative Simulations.” *Journal of Computational and Graphical Statistics* 7 (4): 434–55. <https://doi.org/ggfntd>.

Camilleri, Gabriele. 2024. “Evaluating Word Alignment Strategies in a Japanese-Italian Translation Corpus.” *Proceedings of JADH Conference* 2024: 26–28.

Carl, Michael, and Moritz Schaeffer. 2017. “Measuring Translation Literality.” In *Translation in Transition: Between Cognition, Computing and Technology*, 81–105. John Benjamins Publishing Company. <https://doi.org/10.1075/btl.133.03car>.

Chen, Xudong, Hilofumi Yamamoto, and Bor Hodošček. 2024. “Translation-Based Connotation Visualization for Classical Poetic Japanese Vocabulary of the Kokin Wakashū Ca. 905.” *Journal of Computational Literary Studies* 2 (1): 1–32. <https://doi.org/10.48694/jcls.3596>.

Chesterman, Andrew. 2004. “Hypotheses about Translation Universals.” In *Benjamins Translation Library*, edited by Gyde Hansen, Kirsten Malmkjaer, and Daniel Gile, 50:1–13. Amsterdam: John Benjamins Publishing Company. <https://doi.org/10.1075/btl.50.02che>.

———. 2010. “Why Study Translation Universals?” In *Kiasm*, edited by Ritva Hartama-Heinonen and Pirjo Kukkonen, 38–48. Acta Translatologica Helsingiensia (ATH). Helsinki: Helsingfors universitet, Nordica.

———. 2011. “Reflections on the Literal Translation Hypothesis.” In *Methods and Strategies of Process Research: Integrative Approaches in Translation Studies*, edited by Cecilia Alvstad, Adelina Hild, and Elisabet Tiselius, 23–35. Benjamins Translation Library. John Benjamins Publishing Company. <https://doi.org/10.1075/btl.94.05che>.

Edina, Robin. 2016. “Translation Universals Revisited.” *FORUM* 15 (January). <https://doi.org/10.1075/forum.15.1.03rob>.

Firth, J. R., and F. R. Palmer. 1968. *Selected Papers, 1952-59*. Harlow: Prentice Hall Press.

Hodošček, Bor, and Hilofumi Yamamoto. 2022. “Development of Datasets of the Hachidaishū and Tools for the Understanding of the Characteristics and Historical Evolution of Classical Japanese Poetic Vocabulary.:” In *ADHO 2022 - Tokyo*.

Koller, Werner. [2004] 1979. *Einführung in die Übersetzungswissenschaft*. 7., aktualisierte Aufl. Wiebelsheim: Quelle und Meyer.

Nakano, Hiroshi, Ooki Hayashi, Hisao Ishi, Makoto Yamazaki, Masahiko Ishii, Yasuhiko Kato, Tatsuo Miyazaki, and Akio Kirioka. 1994. *Bunrui Goi Hyo Furoppi Ban/Word List by Semantic Principles, Floppy Disk Version*. Vol. 5. Kokuritsu Kokugo Kenkyujo Gengo Shori Deta Shu/National Language Research Institute Language Data. Tokyo: Dainippon shoten.

Nida, Eugene Albert. 1964. *Toward a Science of Translating: With Special Reference to Principles and Procedures Involved in Bible Translating*. Brill Archive.

Nida, Eugene A., and Charles R. Taber. [1982] 1969. *The Theory and Practice of Translation: With Special Reference to Bible Translating*. Reprint edition. Leiden: Brill Academic Pub.

Nunan, David. 1993. *Introducing Discourse Analysis*. 1. publ. Penguin English Applied Linguistics. London: Penguin English.

Palladino, Chiara, Farnoosh Shamsian, and Tariq Yousef. 2022. “Using Parallel Corpora to Evaluate Translations of Ancient Greek Literary Texts. An Application of Text Alignment for Digital Philology Research.” *Journal of Computational Literary Studies* 1 (1). <https://doi.org/10.48694/jcls.100>.

Pym, Anthony. 2008. “On Toury’s Laws of How Translators Translate.” In *Beyond Descriptive Translation Studies: Investigations in Homage to Gideon Toury*, edited by Anthony Pym, Miriam Shlesinger, and Daniel Simeoni, 311–28. Benjamins Translation Library. John Benjamins Publishing Company. <https://doi.org/10.1075/btl.75.24pym>.

Sankoff, D. 1972. “Matching Sequences Under Deletion-Insertion Constraints.” *Proceedings of the National Academy of Sciences of the United States of America* 69 (1): 4–6. <https://doi.org/10.1073/pnas.69.1.4>.

Schramm, Wilbur Lang. 1954. *The Process and Effects of Mass Communication*. Urbana: University of Illinois Press.

Sinclair, John. 2003. *Reading Concordances*. First Edition. London: Pearson ESL.

Sinclair, John McHardy. 1996. “The Search for Units of Meaning.” *Textus* 9: 75–106.

———. 2004. *Trust the Text: Language, Corpus and Discourse*. Edited by Ronald Carter. 1st ed. London ; New York: Routledge. <https://doi.org/10.4324/9780203594070>.

Sinclair, John McHardy, S. Jones, and R. Daley. 1970. *English Lexical Studies: Report to OSTI on Project C/LP/08. Final Report for Period January 1967 - September 1969*. Birmingham: Department of English, University of Birmingham.

Stubbs, Michael. 2001. *Words and Phrases: Corpus Studies of Lexical Semantics*. Oxford ; Malden, MA: Blackwell Publishers.

Traum, David, and Nizar Habash. 2000. “Generation from Lexical Conceptual Structures.” In *NAACL-ANLP 2000 Workshop: Applied Interlinguas: Practical Applications of Interlingual Approaches to NLP*.

Tymoczko, Maria. 1998. “Computerized Corpora and the Future of Translation Studies.” *Meta : journal des traducteurs / Meta: Translators’ Journal* 43 (4): 652–60. <https://doi.org/10.7202/004515ar>.

Vinay, Jean-PaulDarbelnet. 1958. *Comparative Stylistics of French and English*. *Btl.11*. John Benjamins Publishing Company.

Yamamoto, Hilofumi. 2005. “A Mathematical Analysis of the Connotations of Classical Japanese Poetic Vocabulary.” PhD thesis, Canberra: The Austrilian National University.

———. 2006. “Extraction and visualisation of the connotation of classical Japanese Poetic Vocabulary -How to construct semantic space by graph-.” *Proceedings of Symposium for Computers and the Humanities Symposium 2006 (Jinmoncom 2006)* 2006 (December): 21–28.

Yamamoto, Hilofumi, Bor Hodoscek, and Xudong Chen. 2023. “Development of a Dataset for Comparison Between Predicate Verb Phrases in the Kokinshu and Their Contemporary Translations.” In *Proceedings of JADH Conference.*, 2023:64–67. Tokyo, Japan: The Japanese Association for Digital Humanities (JADH).

Yamamoto, Hilofumi, and Bor Hodošček. 2019. “An Analysis of the Differences Between Classical and Contemporary Poetic Vocabulary of the Kokinshu.” In *The 9th Conference of Japanese Association for Digital Humanities (JADH2019) “Localization in Global DH”*, 68–71. Japanese Association for Digital Humanities.

Yu, Shaoyun, and Katsuo Tamaoka. 2020. “Trade-Off Effect in the Processing of Korean Case-Drop Sentences: An Eye Tracking Investigation.” *Human Behaviour and Brain*, June. <https://doi.org/10.37716/HBAB.2020010203>.

久曽神昇. 1979. *古今和歌集 全注釈*. Vol. 1–5. 講談社学術文庫. 東京: 講談社.

奥村恆哉. 1978. *古今和歌集*. 新潮日本古典集成. 東京: 新潮社.

小島憲之, and 新井栄蔵. 1989. *古今和歌集*. 東京: 岩波書店.

小沢正夫. 1971. *古今和歌集*. 13th ed. 日本古典文学全集. 東京: 小学館.

小町谷照彦. 1982. *古今和歌集：現代語訳対照*. 旺文社文庫. 東京: 旺文社.

松田武夫. 1968. *新釈古今和歌集*. Vol. 上下. 東京: 風間書房.

片桐洋一. 1983. *歌枕歌ことば辞典*. 東京: 角川書店.

———. 1998. *古今和歌集全評釈*. Vol. 上中下. 東京: 講談社.

窪田空穂. 1960. *古今和歌集評釈*. Vol. 上中下. Tokyo: Tokyodo.

竹岡正夫. 1976. *古今和歌集全評釈: 古注七種集成*. Vol. 上下. 東京: 右文書院.

近藤泰弘. 2011. “平安時代の漢文訓読語の分類.” *訓点語と訓点資料 = Diacritical language and diacritical materials / 訓点語学会 編* 127 (September): 120–31.

近藤泰弘, and 近藤みゆき. 2001. “N-gramの手法による言語テキストの分析方法–現代語対話表現の自動抽出に及ぶ.” *漢字文献情報処理研究 / 漢字文献情報処理研究会 編*, no. 2 (October): 50–55.

金子元臣. 1933. *古今和歌集評釈: 昭和新版*. 東京: 明治書院.

陳旭東, ホドシチェクボル, and 山元啓史. 2022. “単語アライメントの誤り対応を用いた歌ことばのコノテーション検出.” *じんもんこん2022論文集* 2022 (December): 111–18.

1. The first geese cry as they fly across the sky, but sorrow comes with the autumn of people’s hearts (author’s translation). [↑](#footnote-ref-24)
2. Collection of Japanese Poems of Ancient and Modern Times [↑](#footnote-ref-30)
3. the technique of alluding to an older poem within a new one [↑](#footnote-ref-31)
4. The eight imperial anthologies include the *Kokinshū*, *Gosen Wakashū*, *Shūi Wakashū*, *Goshūi Wakashū*, *Kin’yō Wakashū*, *Shika Wakashū*, *Senzai Wakashū*, and *ShinKokinshū*. [↑](#footnote-ref-33)
5. However, Norinaga’s goal differed from that of other translators. Rather than creating a standard commentary, he sought to highlight the literary value and appeal of the *Kokinshū*. Consequently, his translation is not analyzed in the same way as others in this study. Additionally, as noted by (**shiozawa1993?**), the specific vocabulary and meanings of the colloquial expressions used in *Kokin Tōkagami* remain unclear, making it challenging to include in the visualization project. [↑](#footnote-ref-34)
6. This premise aligns with the translator’s subjective awareness of their work. [↑](#footnote-ref-50)
7. The calculation scripts offer alternative configurations. [↑](#footnote-ref-51)
8. We can adjust the strictness of the agreement rate calculation through the conditions within . In this case, the function returns 1 if a match exists at any of the , , or levels. [↑](#footnote-ref-53)
9. We also provide three input unit options in the database: one that ignores polysemy and focuses only on simplex units, one that prioritizes compound units while ignoring polysemy, and one that retains potential decompositions of compound units. For this paper, we use the simplex unit approach. [↑](#footnote-ref-54)
10. The gap penalty is set to 0.01 to account for the fact that translations of poems often involve distant agreements between words. [↑](#footnote-ref-56)
11. Naturally, in alignment, . [↑](#footnote-ref-57)
12. In this analysis, we use the addition rate calculated via the bag method. [↑](#footnote-ref-60)
13. Since the response variable lies within the interval, we considered using a binomial distribution. However, this would require detailed information on the total number of words in the translation and the number of additional elements. Based on Yamamoto and Hodošček (2019), we use a beta distribution, which does not assume this information. [↑](#footnote-ref-61)
14. Content words are identified by chasen-based PoS number below 60. [↑](#footnote-ref-63)
15. See the appendix for details. [↑](#footnote-ref-64)
16. A classical Chinese expression that suggests presenting the essence of something without explicit explanation, allowing the reader to imagine it. [↑](#footnote-ref-67)
17. This method represents the first integrated analysis of the seven major commentaries on the *Kokinshū* since Keichū’s *Kokinshū Yozaishō*. Takeoka adopts an analytical approach in his literary studies, while rejecting the work of Kamo no Mabuchi and Kagawa Kageki as lacking in analytical foundation. However, he values the annotations of Keichū and Motoori Norinaga. Additionally, he carefully discusses the differences among all seven classical annotation books from the perspectives of the studies of classical Japanese grammar and vocabulary. [↑](#footnote-ref-69)
18. KSJ (#404) is also significantly shorter, with the *jokotoba* similarly omitted. [↑](#footnote-ref-121)
19. According to Nunan (1993, 21), cohesive markers are “words and phrases which enable the writer or speaker to establish relationships across sentence or utterance boundaries, and which help to tie the sentences in a text together.” [↑](#footnote-ref-126)
20. Translation universals tend to allow for exceptions, making them more probabilistic tendencies rather than strict universals (cf. Pym 2008; Tymoczko 1998; Chesterman 2004, 2010). While they claim universality, exceptions are recognized. [↑](#footnote-ref-127)
21. According to Chesterman (2011), this hypothesis claims that when processing a given text chunk, translators tend to start with a literal version of the target text, then work towards a freer version. [↑](#footnote-ref-128)
22. Note that while the model is written in Centered Parameterization, its implementation in brms uses Non-Centered Parameterization. [↑](#footnote-ref-134)