



古今和歌集の現代日本語訳における追加率要素の分析

AUTHORS

Xudong Chen  

Bor Hodošček  

Hilofumi Yamamoto  

AFFILIATIONS

Tokyo Institute of Technology

Osaka University

Tokyo Institute of Technology

PUBLISHED

September 1, 2024

ABSTRACT

注釈書における翻訳における要素の追加を利用した古今和歌集の歌ことばのノンリテラル情報の可視化をこれまで行われてきたが、それら注釈における現代語訳や、現代語訳における追加要素に関する詳細の分析が示されていない。本稿の目的は、いままでの研究に用いられてきた古今集の現代語訳 10 種が、ノンリテラル要素の可視化に適している材料であるかについて詳細な分析を行うことである。まず、古今集の注釈における現代語訳について概観し、訳者の翻訳の意識についてまとめた。つぎに、直訳語をもたない歌ことばが原文に占める割合（不一致率）と現代語訳における追加要素の割合（追加率）の計算で、現代語訳の要素の追加と一致の実態を把握した。最後に、具体例を示し、現代語訳における追加要素の事例分析を行った。分析の結果、和歌の現代語訳の訳者の翻訳意識は、主観的に、作意中心・テキスト中心・読者中心と不明 4 類に分類できた。しかし、客観的には、翻訳意識が異なっても、計量的には同じ程度の追加要素を多く含めていた。B 分析事例においては、センテンスレベルの訳においても、要素の追加にコーパスレベルの共出現傾向に基づく推論があり、語の意味の志向性について補足要素を多く足していることが確認できた。以上により、現代語訳をノンリテラルレベルの可視化に適用できることを示した。 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 unmatched 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.

KEYWORDS

歌ことば, 現代語訳, ノンリテラル情報

1 はじめに

1.1 目的と結論の概要

本稿の目的は、原文との非一致率、訳文における追加率の計算と対応事例の追加要素の分析におけるという二つの側面から、和歌の現代語訳が、たとえ直訳・逐語訳に拘っているとしても、訳し方の意図が異なっても、解釈材料として、和歌における語の非字義的な情報を補足説明するために利用できることを示すことである。

本研究の主要な貢献は以下にある。

1. 材料として、著作権の消滅した 金子 (1933) における古今和歌集の現代語訳における初の語彙データを公開した。
2. 古今和歌集の口語訳の歴史的変遷を俯瞰し、特に20世紀における和歌の注釈書における現代語訳の訳し方の調査と整理を行った。
3. 古語と現代語訳の語レベルの非一致率、現代語訳における追加率の計算により、翻訳アプローチにかかわらず現代語訳における情報の追加が普遍的であることを示した。
4. 具体的な現代語訳の事例を提示し、対象語のコーパスレベルの共出現語、共出現語の全体的な志向性がセンテンスレベルの現代語訳それぞれに反映される可能性の有無を確認できた。

以上により、現代語訳をノンリテラル情報の可視化への応用が現実的であることを示した。

1.2 背景

1.2.1 和歌と和歌におけるノンリテラル要素

古今和歌集仮名序にあるように、歌人が「心に思ふもの」を見聞きした自然界の物事に託しており、直接明言することが珍しい。和歌のことば（以降「歌ことば」とする）を字義通り理解したところで、その「心に思ふもの」には直接アクセスしている

とはいえない。古今集の恋の歌を例にあげよう：

初雁の鳴きこそわたれ世の中の人心の秋しうければ (古今・恋五・貫之)

このテキストを読むとして、それぞれの語の字義的・リテラルな意味がわかれば、文字通り¹には理解できる。しかし、歌ことば辞典で「秋【あき】」(片桐 1983, 3)を調べると、リテラルな意味だけでは読まれていない情報があることに気付く。

¹ 初雁が鳴いて渡ってくるのだが、人の心の秋が来るのが悲しいので（筆者訳）。

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[…]

この記述に従い読み解いていけば、恋人の心が「飽き（秋）」で、歌人が「泣（鳴）」いている「失恋」の話であろうと推測される。このように「秋」から読み取る「飽きる恋」の情報は、原文の文脈（周辺語）ではノンリテラル的、つまりノンリテラルであるため、語の文字通りの意味だけでは読み取れない。厳密に語のノンリテラル情報の範囲を限定することが難しいが、ここでは限定として2つとりあげる：

1. 物理的にテキストに出現しない
2. 性質上ではその語由来の連想にかかわる説明か、その語の言語変種（地域方言・社会方言・時代・年代など）にかかわる説明

1.2.2 現代語訳の追加要素に基づく歌ことばのノンリテラル要素の可視化

このようなノンリテラル的要素の可視化に向けて、古今和歌集の10人の現代日本語訳をベースにした可視化システムは、筆者らがこれまでに開発を行ってきた (Yamamoto 2005, 2006; 陳, ホドシチュク, and 山元 2022; Chen, Yamamoto, and Hodošček 2024)。理論背景としてコミュニケーションモデルに立脚し、古語と現代語訳との相対化に見えてくる補足要素に歌ことばのノンリテラル情報が含まれていることとしている。

ここでの現代語訳との相対化というのは、現代語の要素（要素の決め方も重要な課題であり、本稿では対象としない）の集合から和歌の原文の要素を取り除き、現代語にしか存在しない要素を諸手法により引き立てることである。この考えが近藤 and 近藤 (2001), 近藤 (2011) では「引き算」と称している。しかし、引き算で残った要素、つまり現代語訳で追加された要素は、どのような性格であるか、ノンリテラル要素に捉えられるか、明確にする必要がある。

翻訳における要素の追加などの処理と操作に関して、Koller ([2004] 1979, 249) では翻

訳者による「介入 (Eingriff)」としている。「無害」とされる介入には、目的言語の読者の不足している背景知識や、デノテーション・コノテーション面の情報、そして、言語内的、社会文化的、間テキスト的な情報の損失を補うための注釈付き翻訳手法の結果としての追加が含まれていると指摘している (Koller [2004] 1979, 249)。そこで、翻訳が注釈・辞書同様に解釈資料としてノンリテラル的な情報の抽出に有効な材料と認識しうる。

ただし、翻訳における追加は読者の理解力を過大評価・過小評価した産物にもありうる (Eugene Albert Nida 1964, 155; Koller [2004] 1979, pp249–250)。この点においては、訳者による翻訳の追加要素のバリエーションは無視してはならない。古今和歌集の歌ことばの解釈資料としての現代語訳の有用性については、Chen, Yamamoto, and Hodošček (2024) では触れたものの、そのバリエーションについて十分説明されていない。材料として和歌の口語訳の歴史、使用してきた 10 人の現代語の背景や、訳者個人個人の翻訳の見方を踏まえ再検討する余地が大いにある。

したがって、本稿では材料として使用してきた 10 種類現代語訳の成立背景、追加要素の量的性質、質的性質について分析を行う。前述にあったように、翻訳における追加は読者の理解力の過大評価・過小評価に基づき生成された産物である可能性があるため、異なる訳者が和歌の情報をどこまで現代読者に開示しているか、あるいは開示できるか、それぞれ独自の考えがあるはずである。訳者の翻訳アプローチの意識を整理し、その異同を明確にする。そして、10 人の翻訳の実践としてはそれぞれ主観的に意識している翻訳アプローチを守っているかを客観的に分析することで、翻訳の普遍的な課題と限界を考察することができよう。

以上の知見を踏まえ、ノンリテラル的要素の可視化システムの根拠をより明確にでき、さらにその改善の方向性をも明確にできる。

1.3 問題意識と本稿の構成

ノンリテラルな要素の可視化の前提として、本稿では具体的に以下 4 つの面から再検討する：

1. 注釈書における現代語訳の背景の調査、翻訳アプローチの調査と分類
 - 訳者の翻訳の主観意識はどのように分類するか。
2. 現代語訳において原文の要素が削りが量的に少ないであることの明確化
 - 訳文に一致要素が存在しない原文の要素（不一致率）がどれほどあるか。
3. 現代語訳における追加要素が翻訳アプローチを問わず量的に十分であることの証明

- 訳文において追加要素がどれほどあるか、翻訳アプローチによる統計的に差が認められるか。

4. 現代語訳における追加要素がノンリテラル情報として扱えることの説明：

- コーパスレベルの共出現情報の全体的傾向が、翻訳の追加要素に反映される例があるか。

本稿は次のように構成される：[Section 2](#)では、材料 10 人の現代語訳の説明に兼ねて、20 世紀の和歌の口語訳の成立の背景をまとめる。[Section 3](#)では、本稿の 4 つのリサーチクエスションに対しそれぞれの方法の詳細を説明する。[Section 4](#)では方法と結果で得られた結果を提示し、[Section 6](#)では結果に対する考察と課題を行う。最後に、本稿の結論を述べる。

2 データ：古今集、注釈と現代語訳

本節では、材料である古今集の説明に併せ、和歌、古今集、注釈と、その現代語訳について、その歴史的な事情について概観する。

2.1 古今集

和歌は「歌」「詠む」と言われるだけあって、本来は、宮廷で、声を上げて歌われていたものであり、基本的に当時の話しことばである。古今集はその名前が示すように、それ以前の古代の歌と当時の歌を集めたものである。古代の和歌は、主に『万葉集』に収められた和歌を指し、7 世紀から 8 世紀の和歌も収録され、山部赤人、柿本人麻呂、額田王など和歌の歴史上重要な歌も仮名で収められている。

古今集は、その後の勅撰和歌集、二十一代集の冒頭に成立し、その後の勅撰和歌集の基本的な形式を確立したものである。その序章には、編者、紀貫之による古今和歌集仮名序が記載されており、和歌の理論書として位置付けられている。和歌文学の研究資料として、のちの本歌取りの元歌の前提とされている、「源氏物語」、「土佐日記」、「伊勢物語」ほか、数々の作品に掲載されている、など日本文学の基本的な資料としても重要である。また、前述のように、話しことばの特徴を持つため、日本語の歴史的変遷を知る上でも重要な資料となっている。

2.2 古今集の注釈と現代日本語訳

日本古典文学の歴史において、古今集は数多くの注釈書を生み出し、その注釈書はそれ自体の歴史ができるほどである (窪田 1960, 上中下:319)。注釈書に見られる多くの現代語訳は読者の和歌の理解を助けるために書かれており、翻訳家の翻訳目的は主に和歌の解釈を伝えるためである。近代については、必ずしも現代語とは言えないが、無視するわけにはいけないので、近代までの古今集の注釈の歴史を簡略に説明し、明治以後、注釈書掲載の現代語訳について考察する。

近代 (1868年) 以前は多くの古今集の注釈書が出版された。近代前 (1600 年–1868 年) の間に、70 以上の注釈書が出版されている (小島 and 新井 1989, 447–50)。表 [Table 1](#) は近代前の著名な注釈書名を 5 つ示す。

北村季吟の八代集抄は 108 巻 (50 冊) からなる八代集²の注釈書である。契沖 (1640–1701) の古今和歌集余材抄は最初の実践的な古今集の研究書である (小沢 1971, 36)。賀茂真淵 (1697–1767) による古今和歌集打聴は基本的に契沖のものを継承したものである。しかし、賀茂真淵は新しい歌よりも万葉集のような古今以前の歌に焦点を当てている。本居宣長の古今和歌集遠鏡はすべての和歌を当時の口語に翻訳する試みである³。香川景樹 (1768–1843) は「古今和歌集正義」でそれ以前の注釈書をかなり強い調子で批判している (小沢 1971, 36; 松田 1968, 上下:58)。とはいえ、1868年以來の多くの注釈書は上記の注釈書の影響を多かれ少なかれ受けていることは確かである。

² 八代集とは古今集、後撰集、拾遺集、後拾遺集、金葉集、詞花集、千載集、新古今集の 8 つの勅撰集である。

³ しかしながら、本居の翻訳意図は他の現代語の翻訳者のものとは少々異なっている。本居の意図は、古今集の注釈としてではなくて、純粹に人々に古今集がおもしろく価値のある文学であることを伝えようとしたところにあり、このことから本研究においてこの翻訳を他の翻訳者と同等に扱うことはできない。さらに、Shiozawa (1993) によると、遠鏡で使われている口語表現の実態、すなわち、どんな単語が使われているのか、どういう意味なのかがはっきりわかっているわけではないし、江戸時代の口語表現それ自身についても明らかになっているわけではないので、本研究で取り扱うことはできない。

現代になって重版、改訂などを含めて古今集の注釈書は 30 あまりと多数出版されているようである。

Author	Year	Annotation book title
Kitamura Kigin 北村季吟	1682	<i>Hachidaishūshō</i> 八代集抄
Keichū 契沖	1692	<i>Kokin Yozaishō</i> 古今余材抄
Kamo no Mabuchi 賀茂真淵	1784	<i>Kokinshū Uchigiki</i> 古今和歌集打聴
Motoori Norinaga 本居宣長	1793?	<i>Kokin Tōkagami</i> 古今集遠鏡

Table 1: 近代前（1600–1868）の古今集の代表的な注釈書

2.3 20 世紀の注釈における現代語訳とデータの概要

21 世紀に入っても現代語訳を含む注釈書が多数出版されているが、これまでの可視化システムの研究において整備できたものは、比較的近代以後に出版された古今集の現代語訳である [Table 2](#)。それぞれ、本稿の研究対象でありデータでもある。

データフォーマットは、Hodošček and Yamamoto (2022) で公開された space-delimited format に準拠している。トークンタイプの識別子（メタコード）として、旧分類語彙表番号が付与されている。データセットの特徴は、多義語に複数分類語彙表番号が付与されていることと、複合語にその下位分解を同時に提供していることがあげられる。複数の基準でフレキシブルな分析に対応できる。

ただし、これらの注釈書から文字化されたデータは和歌の現代語訳の著作権、校訂著作権、翻刻著作権などがあり、公開できない。本稿では、そのうち、著作権の消滅した 金子 (1933) のデータのみを公開する。

	略号	注釈書	底本	トークン		文書数
				数	タイプ数	
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	久曾神 (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

Table 2: 古今和歌集の短歌の20世紀の現代語訳 10 種：トークン・タイプ数の集計においては、複合表現の場合、その下位分解をカウントしていない。

3 方法

現代語訳からノンリテラル要素が効率よく抽出できるかの問題に対して、2つのステップで確認する。第一に、訳者の翻訳に対する主観的な認識を明確にし、主観的な認識においてノンリテラル情報の扱い方への言及を精査する。第二に、客観的にそれぞれの翻訳実践がその翻訳アプローチにの主観意識に従っているかいなかを調査する。

訳者の主観的意識・認識の調査には、注釈書における現代語訳の方針に関する言及の文献調査を行い、主に訳者の翻訳行動に際してのフォーカスを分類する。訳者の実践において、訳、とりわけ訳における要素の追加がその翻訳のアプローチに影響されているか、どの程度影響されているかについて、原文要素の不一致率、訳における追加率を算出し、統計的に翻訳アプローチによる差があるか明確にする。さらに、訳における追加要素がノンリテラル要素として認められるかについて、事例分析を行う。

具体的な方法の説明は次につづく。

3.1 文献調査による翻訳アプローチの分類

本節では 20 世紀の古今集の注釈書の著者の現代語訳執筆方針について文献調査を行い整理する。Yamamoto (2005: 102) では「訳者の理論的考えを踏まえつつ、それぞれ

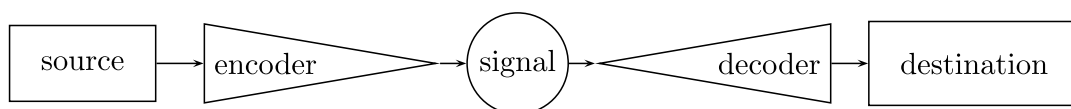
の存在する。Yamamoto (2005, 102) では、訳者の経験的関心と相関関係、これの現代語訳を実際に観察し、「逐語訳 (word-for-word)」 「作者の意思の尊重 (intension-oriented)」 「字句を補う (supplement for words)」 「語順・語法を変える (word change)」 「不詳 (not mentioned)」 のように分類しているが、「作者の意思の尊重 (intention-oriented)」 は翻訳の目標・フォーカスであり、その他は具体的な訳し方である、分類の視座が統一されていない。そのため、基準の統一した分類が必要である。ここでは、Yamamoto (2005) で採用された Schramm (1954) のコミュニケーションモデルの観点から翻訳の力点の置き方の分類を行う (Figure 1)。

現代語訳のコミュニケーションモデルは2つのサブプロセスを含めている。翻訳者が10世紀のサブプロセスの受信者でありながら、20世紀のサブプロセスの発信者である。2つのサブプロセスの経験野をコミュニケーションする役割を果たしている。

このモデルの考え方にに基づき、翻訳アプローチを3つに分類する：

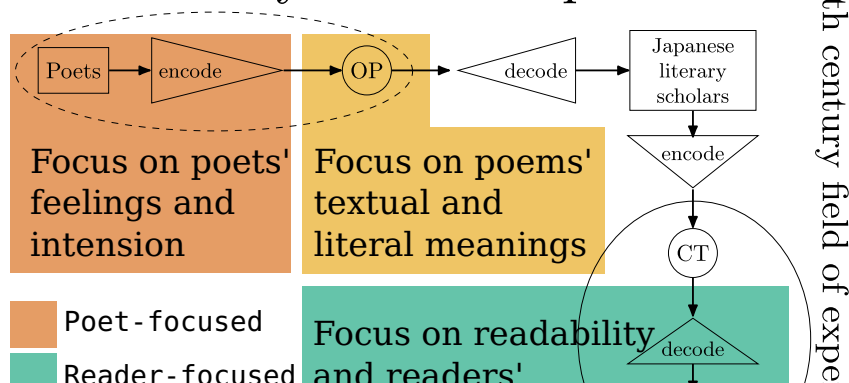
1. 歌人本位のアプローチ (Poet-focused approach): 10 世紀のサブプロセスの中のソースである歌人の作意をハイライトする
2. 原文本位のアプローチ (Text-focused approach): 10 世紀のサブプロセスにおけるシグナルであるテキストの文字通りの意味をハイライトする
3. 読者本位のアプローチ (Reader-focused approach): 20 世紀のサブプロセスの中のデスティネーションである読者の理解をハイライトする

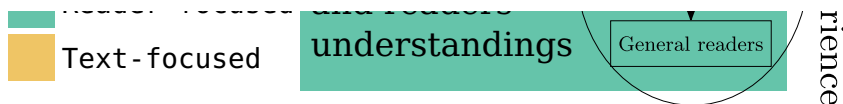
そのほかには、翻訳においては、訳自体のもつ読み物としての文学性、つまり 20 世紀の中のシグナルを重視するアプローチも想定できるが、注釈における訳であるため訳の面白さや文学性への重点的なが考えにくい。ここでは触れないようにする。



(a) コミュニケーションモデル

10th century field of experience





(b) 翻訳アプローチの分類とコミュニケーションモデルにおける位置づけ

Figure 1: コミュニケーションモデルから見る翻訳アプローチの分類

3.2 追加率・不一致率の計算

読者本位・歌人本位の訳文は解釈と推論の追加の可能性が高いため、ノンリテラル要素を多く抽出できそうであると思われる。しかし、訳者の主観的に読者本位・歌人本位と意識しても、翻訳実践と翻訳アプローチのずれが存在しないわけではない。この点について、客観的に追加率の調査で明確にしなければならない。それぞれの訳者が自身の翻訳アプローチに従い、どこまで一貫して訳しているかについて、実際の和歌と現代語訳の不一致率で計算する。不一致率の計算は、要素の順序と重複を考慮しない集合演算による方法（バッグ法）と、要素の順序と重複を考慮するアライメントによる方法（整列法）を考えている。手順については Yamamoto (2005) と Yamamoto and Hodošček (2019) にもあるが、説明の紙幅で明確に手順を示さなかった。本稿では先行研究で曖昧な部分を含め、計算のプロセスを明確にする。

Yamamoto (2005) と Yamamoto and Hodošček (2019) のアプローチは、分解可能でかつ語の複数の意味のメタコードを同時に保持して不一致率の計算を行っている。つまり和歌の語のデフォルトの分類語彙表番号、または、複合語において一致語がある場合、その下位分解やその他の意味の一致可能性について省略する。逆の場合は、その下位の分解や他の潜在的な意味における一致の可能性を考慮して探索する。したがって、和歌の述べ語数を計算する際、複合語単位で対応するか、単語のどの意味で一致するか、対象となる単位が結果から遡り動的に決められている。このアプローチのメリットは、複合表現対複合表現の一致を優先することと多義語の複数の意味のうち、いずれかの一致の可能性を計算の中間プロセスに含めることにある。実際の計算において、このアプローチには強い前提が設けられている：翻訳者が主観的な努力として 100% の情報を翻訳に含めようとしていると見なしている。そのため現代語において形式的に一致語が存在しない語であっても、原文のいずれかの要素と潜在的に 1 対 1 で対応しているとして、追加率を計算する際追加率からその分が引かれている（残った追加率を純粋なアノテーション (pure annotation) と呼んでいる）。しかし、明らかに歌ことばに対応語が存在しないことは、その語の情報の欠落を訳者が認識していることでもあるため、逐語で訳されている前提は理論的である。また、対応の単位が変動するため追加率の分母である訳語の数のカウントが変動し、個々の翻訳者の追加率の比較においては直感的でない問題もあるように思われる。

そこで本稿では、追加率の計算の説明の明確化と結果の再現性について補足を行う。計算における調整は、当初の結論を覆すものではない。本稿の計算では、単位の設定のデフォルトの分類語彙表の意味メタコードと、分解された最小単位に統一した上で行う。⁴

⁴ 計算のスク립トは、ほかの選択肢を与えている。

旧分類語彙表番号（旧 WLSP 番号）に基づく一致の層づけ

歌ことばと現代語訳語の「一致・不一致」は、次の式のように、二語のメタコード（旧 WLSP 番号）の最長共通部分列 (Longest common subsequence, LCS) ([Sankoff 1972](#); [Traum and Habash 2000](#)) の長さで U （不一致）, G （グループマッチ）, F （フィールドマッチ）, E （正確マッチ）の4つに層づけする。

$$\text{match}(s, t) \in \begin{cases} U, & \text{if } \text{LCS}(s, t) < 10 \\ G, & \text{if } 10 \leq \text{LCS}(s, t) < 13 \\ F, & \text{if } 13 \leq \text{LCS}(s, t) < 17 \\ E, & \text{if } \text{LCS}(s, t) \geq 17 \end{cases}$$

この式では、 s はソーステキスト（和歌）における語のメタコードであり、 t はターゲットテキスト（現代語訳）における語のメタコードである。 $\text{LCS}(s, t)$ は s, t 二語のメタコードの Longest common subsequence である。以下具体の分類例をあげる。

1. **Unmatch** ($\text{match}(s, t) \in U$; $\text{LCS}(s, t) = 4$):

- $s = [\text{BG-0}]1\text{-}5520\text{-}20\text{-}0401$ 梅 (plum)
- $t = [\text{BG-0}]8\text{-}0061\text{-}07\text{-}010\text{-}A$ の (of, genitive case)

2. **Group match** ($\text{match}(s, t) \in G$; $\text{LCS}(s, t) = 11$):

- $s = [\text{BG-01-}5520\text{-}]20\text{-}0401$ 梅 (plum)
- $t = [\text{BG-01-}5520\text{-}]19\text{-}115\text{-}A$ 秋萩 (autumn)

3. **Field match** ($\text{match}(s, t) \in F$; $\text{LCS}(s, t) = 11$):

- $s = [\text{BG-01-}2030\text{-}01\text{-}]0300$ 神 (god)
- $t = [\text{BG-01-}2030\text{-}01\text{-}]030\text{-}A$ 仏 (Buddha)

4. **Exact match** ($\text{match}(s, t) \in E$; $\text{LCS}(s, t) = 17$):

- $s = [\text{BG-01-}5520\text{-}20\text{-}040]1$ 梅 (plum)
- $t = [\text{BG-01-}5520\text{-}20\text{-}040]\text{-}A$ 梅 (plum)

バッグ法による不追加率の計算

バッグ法による追加率（現代語訳における和歌原文に対する追加要素の割合）は、2つのテキストを2つの集合と見做し、順序関係せず集合における一致語の数をもとに計算する。ソーステキスト（和歌） $S = (s_i)_{i \in \mathbb{Z}^+}$ とターゲットテキスト（現代語訳） $T = (t_i)_{i \in \mathbb{Z}^+}$ が与えられた時、2テキストのターゲットテキストにおける追加率の計算は下記に従う。

$$U_T(S, T) = 1 - \frac{1}{|T|} \text{agreement}_S \quad (1)$$

$$\text{agreement}_S = \sum_{s_i \in S} \mathbb{I}(\exists t_j \in T \text{ such that } \text{match}(s_i, t_j) \notin U) \quad (2)$$

agreement_S は、ソーステキストにおいて一致語をもつトークンの数である。

$\frac{1}{|T|} \text{agreement}_S$ を一致率（バッグ法）と定義する。分母はターゲットテキストのトークン数 $|T|$ にしている。 s_i, t_j はそれぞれソーステキストの i 番目と、ターゲットテキストの j 番目のトークンを意味する。追加率は1から一致率を引いて計算する。 \mathbb{I} は指示関数（indicator function）であり、インプットが真（ s_i が訳において一致語をもつ）の場合1を、偽（ s_i が訳において一致語をもたない）の場合0を返す。⁵

⁵ この一致率の計算の厳しさは \mathbb{I} 内の条件で調節できる。ここでは、メタコードが E, F, G のいずれかのレベルでマッチングすることを一致するとしている。

⁶ そのかわりに、本稿ではデータベースを利用して3パターンの単位基準の入力を提供している：多義性と複合単位を無視し短単位のみでの計算、多義性を無視し複合単位を優先する計算と、多義性の複合単位の潜在的構成要素をそのまま残す計算が選択できる。ただし、結果の報告では短単位で進める。

ここでの計算で注意されたいのは、この計算ではソーステキストの語について一致語をもつか、もたないかを、ソーステキストの視点から見ていることである。ターゲットテキストにおいて多対一で和歌の要素と一致しても、その複数の一語のみが一語と見做し、その他の語を追加要素にあてるという認識である。このようにする理由は処理の一貫性を重視するためである。例えば、「散る」を「散り乱れる」と訳している場合を多対一の訳とすべきかは曖昧であるので、一貫して「乱れる」のような部分を追加要素と見る。つまり、計算では一対一の訳を大前提としている。⁶

一方で、不一致率は、原文にある要素のうち、訳において一致語が存在しない割合として同様な計算を行う。ここではもちろん、分母をソーステキストの数にしている。

$$U_S(S, T) = 1 - \frac{1}{|S|} \text{agreement}_S \quad (3)$$

バッグ法で計算された追加率に、統計分析に用いて訳者の主観的に意識している翻訳アプローチによる影響が存在するかを検証する。不一致率については初歩的な記述統計で確認する。

整列法による追加率の計算

整列法による計算について説明する。整列法による計算は、文字通りにまずアライメントを推定してから、アライメントの2語の一致の数で計算を行う。集合としてではなく、シーケンスとして順序を考慮する計算である。

本稿では、分類語彙表番号を用いた動的計画法に基づきアライメントを行う。動的計画法を用いて2つのシーケンス間のアライメントを計算するためのスコア関数は以下のように定義する。

\$\$

$$S_{DP}(i, j) = \begin{cases} 0 & \text{if } i = 0 \text{ or } j = 0 \\ \max \begin{cases} S_{DP}(i-1, j) - \text{gap} \\ S_{DP}(i-1, j-1) + \text{weight}(s_i, t_j) \\ S_{DP}(i, j-1) - \text{gap} \end{cases} & i > 0 \text{ and } j > 0 \end{cases} \quad (4)$$

$$\text{weight}(s, t) = \begin{cases} -1, & \text{if } \text{match}(s, t) \in U \\ 10, & \text{if } \text{match}(s, t) \in G \\ 13, & \text{if } \text{match}(s, t) \in F \\ 17, & \text{if } \text{match}(s, t) \in E \end{cases} \quad (5)$$

$$\text{gap} = 0.01 \quad (6)$$

\$\$

ソーステキストのシーケンス S とターゲットテキストのシーケンス T の部分列を用いて、位置 i および j までの最適なアライメントスコア $S_{DP}(i, j)$ を再帰的に計算する。重みづけ関数 $\text{weight}(s, t)$ は LCS の値をそのまま用いる。Unmatch の場合は -1 のペナルティを課す。gap はギャップ（空白）を挿入する際に課されるペナルティ (gap penalty) であり、ここでは便宜上 0.01 と設定する。⁷ この式に基づき、最適なアライメントを探索する。整列した結果の例は次のようになる。

⁷ 単純に観測からして、和歌の翻訳で遠い対応が多く存在しているので、ギャップペナルティは 0.01 と設定することで、遠くの語の対応を許容させる。

>| 立田姫――手向ける――神のあれ　ばこそ――秋の木の葉の幣――と散る――らめー [Kokinshu 298]

>| 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 [@kubota1960Kokin]

この基礎の上で、一致率と追加率を計算する。整列後の S' と T' がある⁸として、整列法による一致率 $\text{agreement}'_T(S', T')$ の計算は次に示す。

⁸ ここでは、整列されたのでもちろん $|S'| = |T'|$ になっている。

$$U'_T(S, T) = 1 - \frac{1}{|T|} \text{agreement}'_S \quad (7)$$

$$\text{agreement}'_S = \sum_{s_i \in S'} \mathbb{I}(\text{match}(s'_i, t'_i) \notin U) \quad (8)$$

この式では、prime (') 記号がついているものは整列後のテキストにおけるものを意味する。指示関数 \mathbb{I} において、アライメントの2語が一致する場合1を返し、アライメントの2語が一致しない場合0を返す。ギャップが入っているアライメントは計算の対象としない。一致率・不一致率・追加率の計算において、分母は整列前のソーステキストのトークン数にする。

動的計画法は、和歌の翻訳における語順の交換などの site swap に弱く、句の順序、語順が入れ替えられる訳の場合アライメントがうまくできなくなる。さらに、整列法はアライメントの2語対が一致するという厳しめの前提が設けてあり、この前提と現実とはかなり異なる可能性が高い。したがって、整列法で計算される追加率は、全体的に統計的に分析せず、ケーススタディの説明においてのみ提示する。

3.3 翻訳の主観的意識と追加率の関係に関する統計モデリング

先行研究 (Yamamoto 2005; Yamamoto and Hodošček 2019) ではグループ・フィールド・同義レベルの一致率（言い換え率）、理論的・実験的追加率などの多くの指標を計算しているが、翻訳者別の追加率の平均と標準偏差しか提示していない。データの更新や、計算手順もすこし異なったので、結果の変化が生じる。そこで、新たに計算した追加率を対象に統計分析を行う。

訳者の翻訳アプローチが実践に移り、実際の現代語訳における追加率に影響を与えているかを確認するために、統計モデリングを行う。具体的には、翻訳アプローチ（Focus）が追加率（additional_rate）⁹ に与える影響を検討するために、ベータ分布¹⁰ に基づく回帰モデルを採用する。このモデルは、訳者（translator）と歌（poem）を変動を統制するためにランダム効果としてモデルに入れている。統計モデルは補足資料を参照されたい。

⁹ ここでは前述したようにバッグ法によって算出された追加率を用いる。

¹⁰ 計算された応答変数は、(0, 1) の区間にあり、二項分布を仮定して用いることも考えられる。ただし、二項分布を用いる場合は、現代語訳文の全語数と追加に相当する語数の情報が必要であり、本稿の計算は Yamamoto and Hodošek (2019) の計算結果を踏まえたため、全語数について情報が不足していることを前提としないベータ分布を用いる。

¹¹ ここでは、事後最狭信用区間 (Highest posterior density interval) を用いる。

モデルの推定は、Markov Chain Monte Carlo (MCMC) 法を用いる。具体的には、4 つのチェーンを使用し、各チェーンで 2000 回のイテレーションを行い、そのうち 1000 回をウォームアップ (burn-in) として設定する。R (R version 4.3.3 (2024-02-29)) (R Core Team 2024) パッケージ `brms` (2.21.0) (Bürkner 2017) を用いてモデルを実装し、事後分布の収束の評価は $\hat{\delta}$ 指標および有効サンプルサイズ (ESS) を確認する。モデルの推定結果として、モデルから各アプローチの追加率をサンプリングし、事後分布の中央値と 95% の信用区間 (Crl) ¹¹ を報告する。2 アプローチ間の追加率の差の事後分布も示す。

可視化と解釈について、Yu and Tamaoka (2020) を参考に、比較 2 群の差の事後分布の 95% の Crl が 0 をカバーしているかを観測するほか、2 群の差が 0 より大きい確率を同時に観測する：まず、95% の Crl が 0 をカバーしていない場合、2 群に差があると判断する；95% の Crl が 0 をカバーしていても、2 群の差が 0 より大きい確率が 95% より大きい、または 5% より小さい場合、傾向差があると認める。

3.4 「立田」歌の事例分析

和歌に対する現代語訳は、なんらかの言い換えと追加あったことが分っても、その追加要素や言い換えの要素の性格がわかっているわけではない。これら要素の性格は、アライメントにおける一致を確認しながら、コーパス言語学の視座から分析を試みる。本稿では J. M. Sinclair (1996) の拡張意味単位モデルの視座から示唆を得て、和歌の原文のコーパスレベルに観測されている傾向性が、センテンスレベルの対訳文でどのように処理されているかを明確にする。

拡張意味単位モデルは以下の 4 つのレベルから対象言語単位の記述を行っている。

1. コロケーション (collocation) : 他の語の共出現関係 (J. M. Sinclair, Jones, and Daley 1970, 15)
2. 類連結 (colligation) : 構文パターンや文法的要素との共出現関係 (cf., Firth and Palmer 1968, 183; J. M. Sinclair 1996, 11; J. Sinclair 2003, 171)
3. 意味的志向 (semantic preference) : 特定の意味の語群 (semantic set) との共出現関係 (J. Sinclair 2003, 178)
4. 談話韻律 (discourse prosody) : 拡張意味単位全体の評価・態度・語田論的音

コロケーションから談話韻律へと、直接的な観測ができなくなり、明示的でなくなっていく (Stubbs 2001, 87-88)。

和歌の性質とデータの量の少なさからして、確実に拡張意味単位を捉えることが難しいことがある。よって、本稿では単純に拡張意味単位の4つのレベルを4つの視座とみて、この4つの視座のうち、共出現の語、共出現の構文パターン、共出現の語の意味的まとまりで考えられる志向性から、10人の翻訳でどのように処理されているかを分析する。

Yamamoto (2005) では、現代語訳の紹介に歌枕「立田」の歌298番をとりあげている。本稿でも同じ歌をとりあげ、それが10人の翻訳、異なる翻訳の方針でどのように処理されているか、前掲の3つの視座から分析する。

コロケーションのレベルで確認すると、八代集全体における54首において、「立田」の文脈に頻出する内容語¹²には「山」(30; うち共出現語として15あり、「立田山」の複合表現として15ある)「紅葉づ」(17)「川」(「立田川」の複合表現として15ある)「秋」(15)「見る」(12)「紅葉葉」(9)「錦」(9)があげられる。頻度5以上の語には、「神無備」(5)「姫」(うち「立田姫」の複合表現として8ある)など神に関連する語と、「散る」(8)「流る」(5)「吹く」(6)「紅葉」(5)「黄葉」(5)など、落葉に関連する語が存在している。¹³

¹² 形態素解析システム Chasen を用いたため、Chasen の ID が 60 以前のもの

¹³ 詳細は補足資料を参照されたい。

共出現する構文パターンとして、動詞の終止形で終わる歌の少なさ(古今集12首の中で1首のみ)が観測される。また、古今集の中では、係り結びは9首観測されており、余韻の残し方に特徴があると考えられる。これら終わり方の構文パターンが、翻訳における処理について考察する。

共出現する語の性質は、内容語のコロケーションは基本的に、「立田川」「立田(の)山」など地名を構成する「山」「川」のグループのほか、秋の「神」に関連する関連語のグループが明瞭に見える。神聖なる場所の歌枕の性質が伺える。

298番歌の訳における追加要素のバリエーションを示しながら、以上のコーパスレベル・大局レベルの要素が298番歌のセンテンスレベルで組み込まれるか、どのように組み込まれているかを確認し、ノンリテラル要素と捉えられるかを論じる。

4 結果

4.1 現代語訳の整理と分類

注釈書の前書き、導入部に記されている翻訳アプローチに関する意識と認識に基づき、10人の翻訳アプローチをコミュニケーションの主体への視座の置き方によって3類に分類した：

1. テキストの字義を重視する方針：金子 (1933); 窪田 (1960); 片桐 (1998)
2. 作者の意図を重視する方針：奥村 (1978); 竹岡 (1976)
3. 読者の理解を重視する方針：小沢 (1971); 久曾神 (1979)

うち、明確に翻訳アプローチについて記されなかったのは松田 (1968); 小島 and 新井 (1989); 小町谷 (1982) である。

この中で、原文の字義を重視する翻訳アプローチは、もっとも逐語訳に拘っている。作者の意図・感受を重視する・読者の理解を重視する翻訳アプローチでは、多少の語と語順の入れ替え、語句の補いを許容しているように述べている。

Text-focused approach: KNK, KBT, KTGR

原文テキストの字義通りの解釈を重視する訳のグループは、原文に忠実に翻訳することに重みを置き、できるだけ語順や意味を変えずに翻訳を行う：金子 (1933) は「逐語訳」に徹し、語を加えたり、削除することを避けており、慎重に翻訳していると述べている；金子 (1933) は歌の意をそこねたり、調をあやまったりすることを恐れて、鏡花水月の訳法¹⁴に従い、きわめて小心に、一字一語の出入をもゆるがせにせぬことを期したことを述べている。また、文体的には、金子の翻訳調は例えば「くずをれる」のような古めかしいことばをわざと翻訳に入れるところなど、本居宣長の「遠鏡」に似ている。窪田 (1960) は、現代語訳の方針について、逐語訳にこだわり、語順を変えたり、新しい語を加えることを避けていると述べている。片桐 (1998) も同様に、逐語訳を行い、和歌の省略や長大な増補を避け、可能な限り忠実に現代語訳しているとしている。

¹⁴ あからさまに説明せず、ただその姿を眼前に思い浮かせるようにする漢文の表現法

Poet-focused approach: OKMR, TKOK

歌人の作意を重視する訳のグループでは、表面的な意味よりも、原文の意図と感受性を重視する群である。奥村 (1978, 7) は「口語訳は、原歌の言葉づかいや言い廻しを尊重し、一首の意味を厳密に解釈するより、基本的な作意をまず感じとってもらうことに重点をおいた」としている。原文の意図の理解を重視し、逐語訳よりも作意を優先して翻訳している。竹岡 (1976, 上下:21) は、訳(口語訳)は原文の「筋」の紹介や解説の類ではなく、作者の認識のしかた・感じ方をあらゆる面より明らかにしたのち、

それ現代語に存在する同じあるいは極力それに近似した認識のしかた、感じ方の表現にそっくりうつしかえることをいう。つまり、表面では、原歌の同義語に拘るものの、原歌における歌人の感受性の保持に重点を置いている。具体的な現代語訳の方針としては、竹岡 (1976, 上下:11-12) は、次のような独自の理論を提案している。

$$S = [(a + b + c + \cdots + n) \times X] \times Y$$

S は翻訳された歌であり、 a, b, c, \dots, n のそれぞれは歌における語、 X は訳に含まれる要素であり、 Y は訳に効果的に加えられた要素である。 a, b, c, \dots, n それぞれの一般的普遍的意義は辞書で理解でき、それぞれの $+$ のしかたに相当するものは文法書を参照すれば了解できる」としている。竹岡 (1976, 上下:11) は Y はふつうは 0 の値であるが、和歌にはいくつか効果的な要素が含まれる時、翻訳者は Y に何らかの値、例えば終助詞（わ、ね、よ）のようなものを付け加わえる。したがって、現代語訳作成のための重要な変数は X となる。その X は、語句の分割、対応する適切な現代語、語順、助詞・接続詞を含めた訳語、文の構造、話の流れ（談話）、場面の 7 つに分類されるとしている。このように、歌の筋より作者の意図を細部まで還元することを強調していると捉えられる。¹⁵

¹⁵ 契沖の「古今和歌集余材抄」以来の 7 種の注釈書の注を統合したはじめてのものである。竹岡はこの注釈書にて文学研究に分析的アプローチを組み込んでいる。竹岡は賀茂真淵、香川景樹の仕事を分析に根拠がないとして同意できないとしている一方で、契沖、本居宣長、富士谷成章 (1738-79) らの注釈を評価している。加えて、特に古典文法、語彙の観点から 7 種すべての古注間の違いについて慎重に議論している。

Reader-focused approach: OZW, KSJ

読者の理解を重視する訳のグループは、原文に忠実であることよりも、読者の内容の理解や解釈を読者に伝えることを重視している。そのため、語順を変えたり、新しい語を加えることもある。小沢 (1971) は原作の語順・語法を変更することをいとわず、読者が理解しやすいように内容を優先した翻訳を行っている。小沢 (1971, 46) は「口語訳はそれだけ独立しても意味がよくわかるように努めたので、時には原作の語順・語法を変えている」と述べており、逐語訳よりも内容の理解の方に力点が置かれている。久曾神 (1979) は、平易な口語訳を採用し、必要に応じて語句を補っており、解釈を重視した翻訳を行っている。現代語訳の方針としては「歌意は、歌句に基づく平易な口語訳としたが、必要と思われる場合は字句を補った」と述べている (久曾神 1979, 1-5:6)。

その他のアプローチ: KMCY, MTD, K&A

現代語訳の方針について、小町谷 (1982), 松田 (1968), 小島 and 新井 (1989) は明確に述べていない。

うち、松田 (1968) の現代語訳は他の 10 名の共同執筆者によって作られている。この注釈書冒頭の解説には、現代語訳は今までの研究にしたがって作られたとは述べているものの、10 名の共著者の間での翻訳の共通認識、作成方針は明らかに述べられていない。また、松田は古今集の和歌について、歌人の感情についてだけでなく、選者の編集意識についても、解釈にも含めようとしていると述べている (松田 1968, 上下:9)。小島 and 新井 (1989) の特色の一つは注釈は江戸期 (1600–1868) だけでなく、これまでの注釈書では無視されてきた中世期 (1392–1600) に作られた注釈も取り入れていることである (pp. 481–482)。また、万葉集の和歌と用語が対比されて引用されている (小島 and 新井 1989, 480)。各歌の注釈としては「歌番号」「大意」「語句の注」「参考事項」の順で述べられている。他の注釈本と比べてさほど量的に違いはないが、付録にはさまざまな資料が含まれており、付録だけで本の 30 % を占めているほどである。

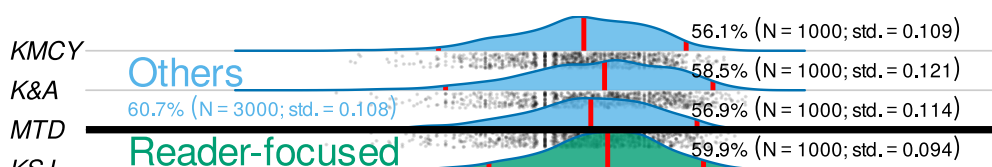
まとめ

それぞれの翻訳のアプローチを見る限り、ほとんどの作者が逐語訳を基本に原文の意味を変えない翻訳を試みているが、従来の注釈書の受け継ぎ方の特徴や情報伝達の力点の置き方がやはり異なっている。本稿では、「テキストの字義を重視」「作者の意図・感受の解釈を重視」「読者の理解を重視」「その他」のようにコミュニケーションモデルの観点から分類を行った。この中で、原文の字義を重視する翻訳アプローチは、もっとも逐語訳に拘っていた。作者の意図・感受を重視する・読者の理解を重視する翻訳アプローチでは、多少の語と語順の入れ替え、語句の補いを許容していた。ただし、この分類は互いに重なる部分もあるが予想される。

4.2 不一致率と追加率の推定結果

追加率の計算結果は、[Figure 2](#) を参照する。不一致率は 0–88 % であった [Table 3](#)。TKOK の現代語訳がもっとも不一致率が低く 17.2 % ($sd = 0.1$) であった [Figure 3](#)。一方、OZW の現代語訳がもっとも高く 23% ($sd = 0.112$) であった [Figure 3](#)。既に述べたように、OZW は語順を変更することを惜しまないのに対して、竹岡は自身の方針に遵守して翻訳した。それが如実にデータとしてあらわれていたといえよう。

10 種の現代語訳のうち、697 番歌は、88.2% ともっとも不一致率の高い和歌であった [Table 3](#)。つまり、和歌の中のほとんどの語が訳されていないと見てよい。このような歌については次の例で確認する。



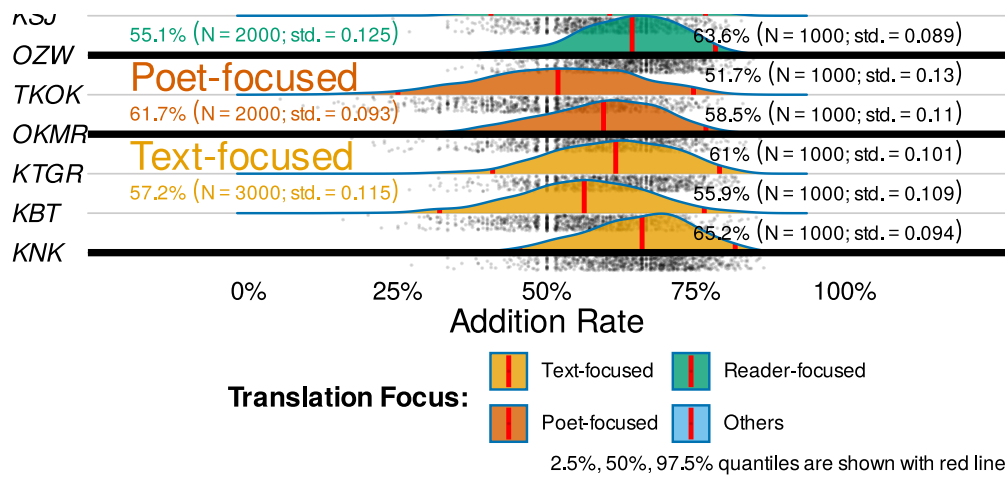


Figure 2: 訳者別の追加率の確率分布 Probability distribution of addition rates by translator

Picking joint bandwidth of 0.0242

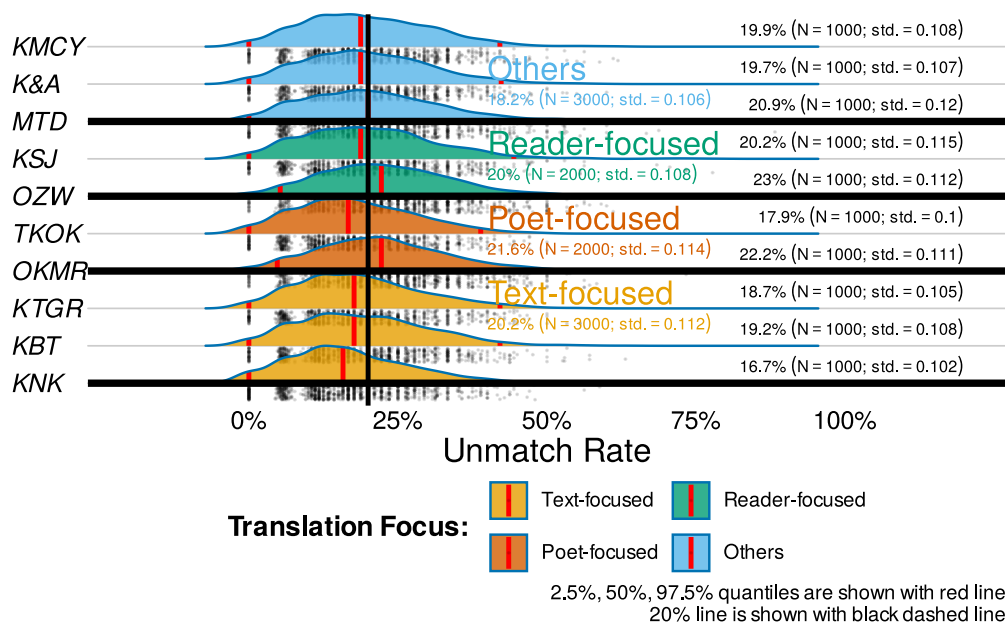


Figure 3: 訳者別の不一致率の確率分布（明確な対応をもたない要素が和歌原文を占める割合）
Probability distribution of unmatched rates by translator (unmatch rate is the proportion of elements in the original poem without agreement)

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

(a) 不一致率・追加率の概要 Summary of unmatch and addition rates

Translator	PoemID	UnmatchRate
MTD	697	0.882
KSJ	404	0.778
KSJ	886	0.733
OKMR	515	0.733

KSJ	515	0.733
-----	-----	-------

(b) 不一致率のもっとも高い歌・訳対 *Poems and translations with the highest unmatched rates*

Translator	PoemID	AdditionRate
TKOK	173	0.053
TKOK	343	0.091
TKOK	478	0.115
MTD	661	0.133
TKOK	744	0.143

(c) 不一致率のもっとも低い歌・訳対 *Poems and translations with the lowest unmatched rates*

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

(d) 現代語訳の不一致率のもっとも高い歌 *Poem with the highest unmatched 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

(e) 現代語訳の追加率のもっとも高い歌 *Poem with the highest addition rate in contemporary translations*

Table 3: 不一致率・追加率の概要 Summaries of unmatched and addition rates

不一致率の高い対訳

不一致率の高い歌の対訳対において、10 人の現代語訳に共通して欠落がみられたのは「序詞」「枕詞」「まわりくどい言い回し」の 3 種類あった。訳者の幾人かは意識的に「序詞」「枕詞」を訳さないで放置していることが観測された。例えば、松田 (1968) の現代語訳では、4 首は現代語訳の方がもとの歌より短くなっている。MTD (173) と MTD (665) の訳においては、枕詞「ひさかたの」「みつしほの」は省略されている。MTD (684) の MTD (697) の訳においては、枕詞も序詞も省略されている。これらの歌は当然不一致率も高い。特に、MTD (697) はもとの和歌よりもずっと短い¹⁶。MTD (697) では大和の枕詞「敷島の」が省略されている。元の歌は、「頃も」と「衣」が掛詞になっているが、その「衣」に掛かっていく序詞も省略されている。

¹⁶ KSJ (404) もまた極端に短く、序詞がやはり省略されている。

```
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%)
>| しきしまのやまとはあらぬからころもころもへずしてあふ——よし——もがな [697]
>| —————絶え間なく——会いたいものだなあ—— [MTD]
```

Figure 4: 697 番歌の 松田 (1968) による現代語訳：アライメントの結果を書き換えている；「しきしまのやまとはあらぬからころも」が省略されている。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.

和歌と現代語訳の間にある回りくどい表現の不一致は、翻訳における「欠落」というより本稿の一致の探し方（アルゴリズムとして一対一で探索する）に課題が存在しているといえるが、一例をとりあげる。OKMR (#346) は、奥村の現代語訳は和歌の語をほとんど使わず、あえてつぎのように回りくどい表現を使っている：

- 「とりそへて」=「さらにそえて」
- 「とどめおきて」=「のこしておきましょう」

特に整列法は、このような歌をうまくアライメントできなかった。

追加率の高い対訳

MTD (#629) は追加率が高い訳のひとつである。この歌は「あやなし」「まだき」などの古語、「なくに」のような現代語には使われない語法が含まれている。これらの語は、現代語と分類語彙表番号で対応づけることができなかった。また「名のたつ」と「たつたがは」の掛詞も存在する。

追加率の平均値上位 10 首のうち、ほかにも 705 番歌、669 番歌、617 番歌など追加率の高かった和歌がある。705 番歌は「身」と「雨」に「降（経）る」の掛詞が含まれている。669 番歌と 617 番歌は、それぞれ「海藻目」と「ながめ」の掛詞がある。

掛詞を含む和歌は曲線的でセーラー

```
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

>|-----あやなく-----て----- [629]
>|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けないも [MTD]

>|-----またきなき名のー立田河わたらてやまーんー物----- [629]
>|のではあるがわけもなくもう無実の評判が立っ-----てしまったーいずれ二人の恋は遂げ [MTD]

>|ーならーな-----くーに----- [629]
>|られないではいけないものではあるが----- [MTD]
```

Figure 5: 629 番歌の 松田 (1968) による現代語訳：アライメントの結果を書き換えている
Contemporary translation of Poem #629 by 松田 (1968): The alignment results have been modified.

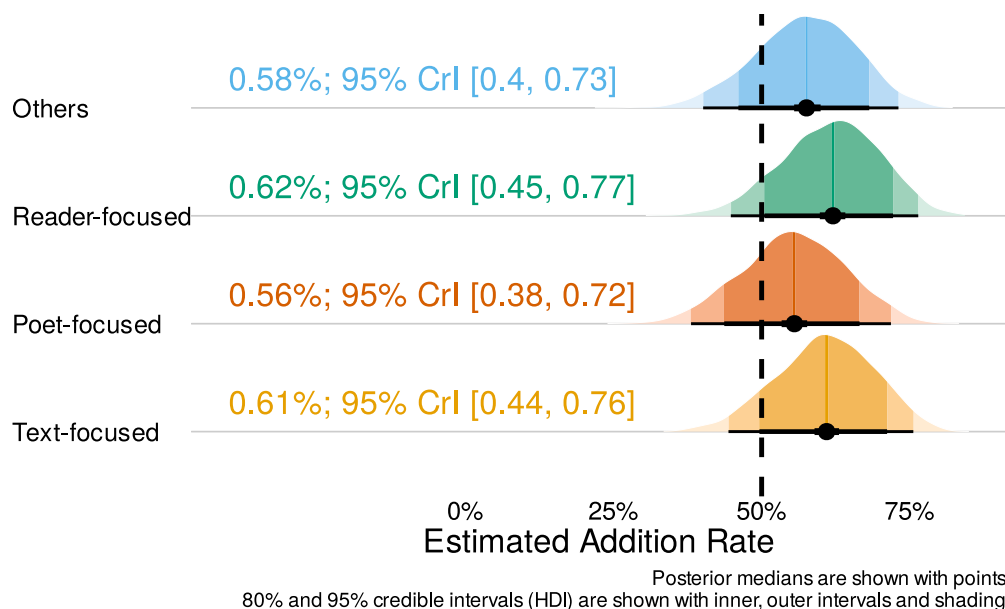
4.3 翻訳アプローチによる追加率の差

本節では、統計モデリングを用いた翻訳アプローチによる追加率の差を検証する。モデルの推定した効果のいずれの \hat{R} も 1 になっており、1.1 を下回り収束が良好であると示した (Brooks and Gelman 1998)。検証したい効果について、ESS がすべて 2000 程度を達した。

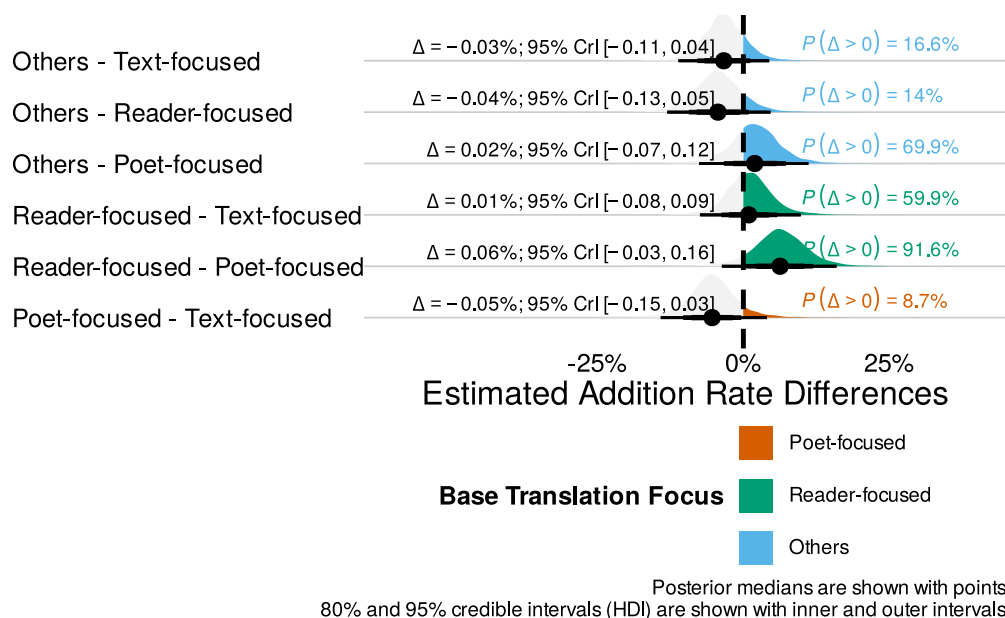
翻訳アプローチによる追加率の相違が小さい

Figure 6 は翻訳アプローチごとの追加率の事後分布である。この図から、以下の情報が得られた。

- 追加率は、翻訳アプローチを問わず 現代語訳の要素の 50% 前後が、追加的な要素であることがわかった。この点は、観測データからも確認された。
- 翻訳アプローチグループで比較した結果、2 グループの差の事後分布の 95% の CrI はすべて 0 をカバーしており、それぞれの 2 群の追加率は異なることが認められなかった。傾向としては、Reader-focused 群の追加率に比べ other 群の追加率が -4% 程度下回った (posterior median = -0.04%; 95% CrI = [-0.13, 0.05], $P(\Delta > 0) = 14\%$)。Reader-focused 群の追加率は、Poet-focused 群より 6% 程度下回った (posterior median = 0.06%; 95% CrI = [-0.03, 0.16], $P(\Delta > 0) = 91.6\%$) ことが観測されているが、傾向差の基準を満たさなかった。



(a) 予測された各翻訳アプローチの追加率の事後分布 Posterior distribution of addition rates for each translation approach



(b) 予測された翻訳アプローチによる追加率の相違の事後分布 Posterior distribution of differences in addition rates by translation approach

Figure 6: 予測された追加率の事後分布 Posterior distribution of addition rates

訳者による追加率の変動より歌による変動のほうが大きい

階層モデリングにおいて、訳者と歌をランダム切片としてモデル化している。訳者グループと歌グループがそれぞれ共有していると仮定した正規分布のパラメータである標準偏差の事後分布を可視化した [Figure 7](#)。訳者グループの標準偏差の分布が歌グループのより左寄りであった。訳者の間の変動よりも、歌の間の変動のほうが大きいことが伺える。つまり、ランダム効果の視点からは、歌が訳者に比べ追加率の変動に寄与していると考えられる。

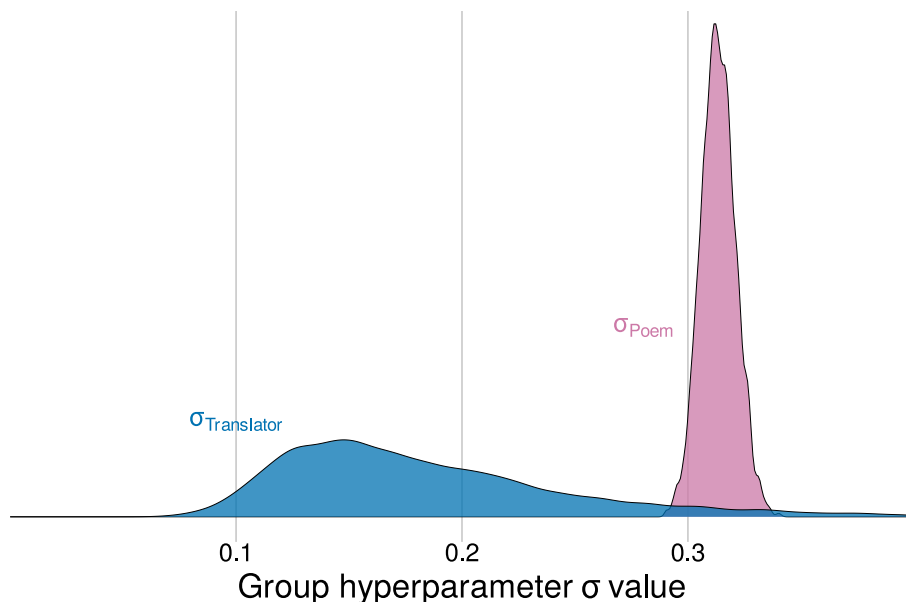


Figure 7: グループレベルのハイパーパラメータの事後分布 Posterior distribution of group-level hyperparameters

4.4 298 番「立田」歌の訳の事例分析

298 番「立田」の歌についての事例分析は、アライメントを示しながら進める。

298 番歌の不一致率、翻訳における追加率のまとめは [Table 4](#) に提示している。ここでの計算は整列法の結果である。歌のトークン数 19 の内、アライメントの 2 語対が一致する数が、OZW 以外に 14 を越えた。OZW はやはり、文の入れ替えを行っているため、うまくアライメントができなかった。原文の要素の不一致率が低い水準にキープされており、訳における追加率は OZW を取り除き 50% – 70% になっていた。1 語あたり 1.5 – 1.7 語で訳されていた。アライメントはプログラム出力を修正した上で [Figure 8](#) に整理した。修正前の出力は補足資料を参照されたい。さて、どのような追加要素になっているか確認していく。

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

Table 4: 「立田」歌 298 番の現代語訳の不一致率・追加率の概要 Overview of unmatched and addition rates in contemporary translations of “Tatsuta” Poem

Tatsutahime (the Goddess of autumn):

>	-----	立田姫	-----	[298]
>	-----	【秋をつかさどる】龍田姫	【が旅立ちにあたって】	----- [KTGR]
>	-----	竜田姫	【は】	----- [KBT]
>	-----	竜田姫	【には】	----- [MTD]
>	-----	【秋も終りに近づき秋の女神の】龍田姫	【がお帰りになる】	----- [OKMR]
>	-----	竜田姫	【が旅にあって】	----- [TKOK]
>	-----	立田姫	【は秋の神だが】	----- [KNK]
>	-----	竜田姫	【が】	----- [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】	----- [KMCY]
>	-----	【秋も終りに近づき】竜田姫	【がお帰りになる際に】	----- [KSJ]
>	-----	【もはや秋の終りで】龍田姫	【が帰り道にお着きになった】	----- [OZW]

Dedicate something to the god

>	-----	手向け	-----	る	-----	神	-----	の	-----	[298]
>	-----	手向け	-----	(をする)	-----	神	-----	が	-----	[KTGR]
>	-----	手向け	-----	(をするべき)	-----	神	-----	が	-----	[KBT]
>	【旅中】	-----	(供え物をささげ)	-----	る	-----	【道祖】神	-----	が	[MTD]
>	【姫が道中の安全を祈って】	-----	手向け	-----	(をなさる)	-----	神	-----	が	[OKMR]
>	-----	手向け	-----	る	-----	神	-----	が	-----	[TKOK]
>	【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】	-----	[KNK]
>	-----	手向け	-----	(をする)	-----	神	【さま】	-----	が	[K&A]
>	-----	手向け	-----	(をする)	-----	神	-----	が	-----	[KMCY]
>	【旅の安全を祈って】	-----	手向け	-----	(られる)	-----	神	-----	が	[KSJ]
>	【中略】	【道の】	(神様)	【にそれを】	(供えていらっしゃる)	*	-----	-----	-----	[OZW]

Exists, so that

Y	-----	あれ	-----	ば	-----	こそ	-----	[298]	
Y	-----	ある	-----	から	-----		-----	[KTGR]	
Y	-----	ある	-----	ので	-----		-----	[KBT]	
Y	-----	ある	-----	から	-----	こそ	-----	[MTD]	
Y	-----	ある	-----	から	-----	こそ	【だろう】	-----	[OKMR]
Y	-----	ある	-----	から	-----	こそ	-----	[TKOK]	
Y	-----	ある	-----	から	-----	こそ	-----	[KNK]	
Y	-----	(おられる)	-----	から	-----	こそ	-----	[K&A]	
Y	-----	ある	-----	から	-----	こそ	-----	[K&A]	
Y	-----	ある	-----	から	-----	こそ	-----	[KSJ]	
Y	【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) アライメント (前半) : たつたひめ 手向ける神の あればこそ Alignment (first half):
Tatsutahime tamukeru kami no areba koso ‘(There is) the god exits for Tatsutahime to
dedicate (things), so that ...’

Leaves of autumn

>	-----	秋の	-----	木の葉	-----	の	-----	[298]
>	-----	秋の	-----	木の葉	-----	が	-----	[KTGR]
>	【そのつかさどる】	-----	秋の	-----	木の葉	-----	が	[KBT]

Y	【このついでに】	秋の	木の葉	が	[MTD]
Y	-----	秋の【紅葉した】	— (葉)	-----	[MTD]
Y	-----	・	(紅葉)	-----	[OKMR]
Y	【それであるように】	秋の	木の葉	-----	[TKOK]
Y	【このように御自身お染めなされた】	秋の	木の葉	-----	[KNK]
Y	-----	秋の	木の葉【のもみじ】	-----	[K&A]
Y	-----	秋の	木の葉	-----	[KMCY]
Y	-----	秋の	(紅葉)	-----	[KSJ]
Y	-----	・	【山の】 (紅葉)	-----	[OZW]

Scattered as scared paper (which are ritual objects used in Shinto practices)

幣	と	散	る	らめ	[298]					
幣【を撒く】	(ように)	散	(っている)	の	一だろう	[KTGR]				
幣	(のように)	散	る	のであろう	【よ】	[KBT]				
幣	(のように)	散	(っている)	の	一だろう	[MTD]				
幣	(のように)	散	(っている)	【だろう*】		[OKMR]				
幣	(として)	散	(っている)	のであろう		[TKOK]				
【手向けの】	幣	(のように)	散	る	のであろう	[KNK]				
【手向け】	(もの)	(として)	散	る	の	でしょう	【ね】	[K&A]		
幣	(となつて)	散	(っている)	の	一だろう	[KMCY]				
幣	(となつて)	散	(っている)	のであろう		[KSJ]				
【色とりどりの】	幣	(となつて)	散	(っている)	【中略】	の	だ	・	【な】	[OZW]

(b) アライメント (後半) : 秋の木の葉の 幣と散るらめ 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 8: 10人の訳者のアライメント (筆者がアライメント推定プログラムの出力を修正して作成) : 「【】」は追加と判断する要素である ; 「()」は言い換えと判断する要素である。「*」は、句の順序の入れ替えで対応する文・語である。「・」は直訳されていないと判断する要素のプレースホルダーである。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.

共出現語の訳出

「立田」のコーパスレベルの共出現語を、内容語の「山」「川」「紅葉づ」「姫」「手向く」「紅葉葉」「神無備」「秋」「幣」「散る」の頻度が高い。298 番歌は典型的な「立田」の歌として、「姫」「手向く」「幣」「秋」を含めている。含まれている語のほとんどが、そのまま訳出されている。298 番には出現しない「紅葉葉」「紅葉づ」などは、298 番歌において「木の葉」が暗示しているため、次のように処理されている :

- 紅葉した葉 [MTD]
- 木の葉のもみじ [K&A]
- 紅葉 [KSJ, OZW, OKMR]

KBT、TKOK、KNK、KMCY、KTGR は「このは」の処理をしなかった。

また、コーパスレベルの共出現語「幣」について

- 【色とりどりの】 幣 [OZW]
- 【手向けの】 幣 [KNK]
- 幣【を撒く】（ように）散（っている） [KTGR]

のように追加は見られた。ただし、これらの追加が「立田」に帰属させるべきか、判断が難しい一面がある。「幣」に經由して間接的な情報と捉えて妥当であろう。

基本的に、コーパスレベルにおいて「立田」の共出現要素の処理は、そのまま残す方針か、その他の要素より多く追加する方針かになっていることが伺える。また、「このは」のような共出現ではないものについては、大局的なコロケーションを踏まえ「紅葉」であるとして訳出されるケースが見られた。訳出の方針は、訳者を問わず安定的な翻訳になっている。

訳と原文の差分において、コーパスレベルの共出現語に関連する重要な情報が見えてくると推測できる。

構文パターンの訳出

共出現する構文パターンとして、終止形で終わる歌の少なさ (古今集 12 首の中で 1 首のみ) が観測されている。特に、古今集の中では、係り結びは 12 首のうち 9 首が観測されている。これら文の終わり方などの構文パターンは、298 番歌の訳における処理について分析した。

298 番歌の係り結びは、強意の「こそ」と推量の「らむ」の活用形「らめ」で形成されていた。その訳として、10 人の中の 7 人が「こそ」を残しており、さらに 9 人が文末において「強調」の「ノ [であろう・でしょう]」構文を使用している。推量の「らむ」について 9 人が「だろう・でしょう」と訳した。唯一「だろう」で訳さない OZW は「だな」という詠嘆的な口調で訳されている。KBT, K&A は文末にそれぞれ「よ」「ね」の終助詞を追加した。

典型的な構文パターンとして、その訳出が非常に口バストなものになっていることが確認できた。ほかの係り受けを使わない「立田」歌の翻訳においても、この構文パターンが干渉して追加されているかいなか確認したところ、事例が見当たらなかった。つまり、特徴的な共出現の構文パターンの処理は、少なくともこの歌においては如実な訳であった。

その他文法的な要素、格助詞の「と」、動詞の基本形式「-u」（例えば「なる」）の訳は、コーパスレベルにおいては「立田」の特徴的な共出現ではないが、分析を試みた。その結果、次のようにそれぞれ多様な訳し方が確認された。

[…幣]「と」：

- […幣を撒く|幣の]ように [KTGR, KBT, MTD, OKMR, KNK]
- […幣|手向けもの]として [TKOK, K&A]
- […幣]となって [KMCY, KSJ, OZW]

[…散]「る」

- [散っ]ている [KTGR, MTD, OKMR, TKOK, KMCY, KSJ, OZW]
- [散]る [KBT, KNK, K&A]

[…手向け]「る」

- […手向けを]する [KTGR, K&A, KMCY]
- […供え物をささげ|手向け]る [MTD, TKOK]
- […手向けを|お手向け]なさる [OKMR, KNK]
- […手向けを]するべき [KBT]
- […手向け]られる [KSJ]
- […それを供え]ていらっしゃる [OZW]

このように、コーパスレベルの共出現構文要素より訳の追加・言い換えの揺れが大きかった。この中で、動詞フレーズの訳され方の揺れについては、理論よりも訳し方の揺れが大きい Yamamoto, Hodoscek, and Chen (2023) ことも報告されている。これらの追加要素と言い換えは、古語・現代語の言語変化・変異に由来したものと考えている。分類語彙表番号で一致と認定できず、現代語訳と原文の差分において抽出されるようになるが、ノンリテラル情報とは言い難い。

共出現語の意味傾向の訳出

「立田」のコーパスレベルでの共出現語の傾向としては、「立田川」「立田（の）山」など地名を構成する「山・川」のグループと、秋の「神」に関連する関連語のグループに属していることがあげられる。秋・紅葉が有名な神聖な場所の歌枕として使われている性質が伺える。この傾向性は、298 番の歌でどのように明確にされているかについては、やはり直接的な追加がなかった。ただし、間接的な現象として、10 人の訳では、「立田姫」「神」の周辺に追加要素が集中して出現していることが確認された。「立田姫」に対する要素の追加は、10 人のうち 7 人が行っている：

- 【秋をつかさどる】龍田姫【が旅立ちにあたって】 [KTGR]
- 【秋も終りに近づき秋の女神の】龍田姫【がお帰りになる】 [OKMR]
- 竜田姫【が旅にあって】 [TKOK]
- 立田姫【は秋の神だが】 [KNK]
- 【帰り道についた】龍田姫【が道中の無事を願って】 [KMCY]
- 【秋も終りに近づき】竜田姫【がお帰りになる際に】 [KSJ]

- 【もはや秋の終りで】 龍田姫【が帰り道にお着きになった】 [OZW]

KBT, MTD, K&A の訳では「立田姫」の周辺の追加が確認されなかった。「手向ける神」の文脈に確認できる追加は5人の訳から観測できた。

- 【旅中】（供え物をささげる）【道祖】神 [MTD]
- 【姫が道中の安全を祈って】手向け（をなさる）神 [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】 [KNK]
- 【旅の安全を祈って】手向け（られる）神 [KSJ]
- 【道の】（神様）【にそれを】（供えていらっしゃる）＊ [OZW]

K&A, KMCY, KTGR, KBT, TKOK は要素の追加を行わなかった。「立田姫」「手向く」「神」の情報について、ほかの要素よりも重点的に情報の提供と補足を行っている。「立田姫」が「秋の神」・秋の擬人化で、立田姫の「手向く」目的が「旅の安全」で、手向く対象の「神」が道の神、旅の神、道祖神であるといった、情報の補足が現代語訳と原文の差分で抽出できよう。

5 考察

5.1 追加率からみる翻訳アプローチと翻訳実践とのずれ

現代語訳における追加率の階層モデリングの結果、いずれの翻訳アプローチにおいても、相当の比例の追加要素が含まれていることが再現できた。また、訳者の間の変動より、歌による変動のほうが大きいと考えられる。これらの注釈書では、多くの語の解釈が注釈によってすでに詳しく解説されていながらも、追加率が低下しているとは意味しなかった。訳者の翻訳アプローチがテキストを重視するにせよ、作者の意図を重視するにせよ、読者の読み易さを重視するにせよ、基本的に追加せざるをえない要素があると推測できよう。

ただし、追加要素の中では、それぞれの訳を具体的に確認すると、OZW が語順の入れ替えを許容するなり、TKOK の訳が著者の本意を重視するなりの方針は確実にその実践に反映されている。個別の歌を精査することで明確な差が見えてくるものの、大局的にその差が傾向としかいえなかった。それらの要素は、つまり、翻訳アプローチ、訳者の個人差に左右されないものが量として十分ある。また、追加率の平均値上位10首の歌を観察したところ、掛詞を用いた歌が特徴的であった。掛詞の役割を明確するには、訳での補足が求められており、現代語・古代語の差分で残る要素が多いと考えられる。したがって、量的には現代語訳にトス、ハ、リ、ニ、リ 要素の可視化が現実的

えられる。したがって、量的には現代語訳によるノンリテラル要素の可視化が量的であった。

一方で、歌と現代語訳の差分によるノンリテラル情報の抽出の際に、訳されなかった語の存在がないことが望ましくないので、不一致率の高い歌について調査を行った。その結果、翻訳の「欠落」の多くは序詞、枕詞と回りくどい表現であることが判明した。現代語文と歌の差分では、枕詞と序詞などについてのノンリテラル要素の抽出が難しい一面が伺える。ただし、回りくどい表現に関しては、基本的に歌ことばとその回りくどい言い方の一部との一致を確定できるが、回りくどい表現全体との一致ができなかった。この問題を語の一致の課題だと述べたものの、最初に一对多の訳について言及したように一对多の訳と同様に、回りくどい表現は元の歌ことばに情報を付随して生成されることが多いため、回りくどい表現で対応で一致しない要素（つまり、訳のあまりもの）は、ノンリテラル要素を調べる目的からすれば、むしろ貴重なものである。

5.2 翻訳で開示できるノンリテラル要素の類型化

ケーススタディーの分析に基づき、拡張意味単位からヒントを得た3つの観点で訳における追加要素について、以下の類型化を提案する。

1. コーパスレベルでの共出現語に由来した追加要素

- 共出現語の直接的な追加：例えば、「木の葉」→「木の葉の紅葉」
- 共出現語に対する補足、つまり二次的・間接的な要素の追加：例えば、「幣」に「色とりどり」「撒く」

2. 構文・文法面での言語変化のギャップを埋めるための追加要素

- 分類語彙表番号で一致と認定できなかった助詞の変化：例えば、「と」の複数の訳し方
- 動詞基本形に対する要素の追加：「手向ける」の複数の訳し方

3. コーパスレベルの共出現語の意味群の全体的な傾向性をめぐる追加要素

- 共出現語の「秋」「神」の意味群にある語についての補足：「立田姫」「道祖神」の情報の補足

翻訳との差分に基づく可視化システムを構築する際にこれらの要素の区別が重要である。まず、「木の葉」に補足しての「紅葉」は、センテンスレベルではノンリテラルであるが、コーパスレベルではリテラルな共出現語であるといえる。差分による可視化でもアクセスできることが確認できたものの、その他の共出現分析に用いる指標でも十分抽出できる。そのため、差分による可視化の必要性を低くなる。

次に、構文の時代変化のギャップを埋めるための追加要素は、原文の語に由来するノンリテラル情報とは捉えられないため、除外する手続きを踏まなければならない。そして、共出現の語の全体的傾向性に関連する追加要素は、おそらく共出現語分析では抽出できない。翻訳との差分による可視化の独自の利点と考えられる。

最後に、本稿では触れなかった拡張意味単位の第4のレベルである談話韻律の視点について補足する。談話韻律に相当する感情面の評価と、語の社会言語学的な属性は、訳の追加要素では分からない情報と考えている。これらの属性を直接的な補足と追加で明示化することはない。一部は現代語にある connotative term で言い換えられており、これら機能が相当する語の入れ替えはあっても、その入れ替え自体の談話韻律の説明として直接は成立しない。よって、これらの評価は、ほかの方法を求めるべきである。

5.3 現代語訳の知識としての信憑性

10人の翻訳と原文の差分による可視化のシステムは、10人の個々の歌に対する知見を総合するという意味で、ある種のメタ分析や要約と見てよいと考えている。

しかし、現代語訳に含める知識の可視化は、どこまで歌ことばのノンリテラル情報にアクセスできるか、可視化としてどの程度妥当で信頼できるかは、本稿では触れなかった。翻訳において、過剰に意味を付けることは解釈として妥当性と整合性を失う。本来のテキストのシンプルな意味や意図が、過剰な意味付けによって曖昧になり、読者がテキストを正しく理解することが難しくなりうる。そのような翻訳を可視化に応用すると、可視化自体の意味の信憑性がなくなる。その可能性は考慮すべき一環であり、いかに訳者の訳語の中からロバストな部分を見出すかは翻訳に基づく可視化の重要な課題である。

5.4 翻訳研究の知見による解釈の必要性

本研究は、翻訳研究として位置づけるものではないが、翻訳における要素の追加や、原文における要素の喪失などの現象は、翻訳研究における仮説と理論によって解釈する余地は大いにあると考えられる。

例えば、翻訳普遍性仮説 (translation universal hypothesis; cf., Chesterman (2004); Edina (2016)) における長さ増加の普遍性 (law of lengthening, Vinay (1958))、明示化の普遍性 (law of explication, Baker (1996); Blum-Kulka and House (1996)) からしては、現代語訳の追加要素の存在が不思議なものではなく、むしろ翻訳における普遍的な現象である。特に明示化の普遍性は、本稿のいうノンリテラル要素の開示に関連性

か強い。翻訳における明示化について、Blum-Kulka and House (1986) では、翻訳者がソーステキストにはない cohesive marker¹⁷ をターゲットテキストで示すことが示唆されている。一方で、Baker (1996, 180) は、この問題を物事を暗黙的なままにしておくのではなく、明確に説明する傾向があるというように、より広く捉えられるようになった。「明示化」は、もともとは談話標識の訳文における増加傾向を指しており、本研究における要素の追加の一部の解釈としてありうる。とはいえ、すべての現象についてこの仮説で解釈されることが難しい¹⁸。例えば、談話標識や cohesive marker の追加を取り除いても、分析事例においてはかなりの追加要素があった。すべての要素の追加が翻訳普遍性に帰結できなかった。

¹⁷ Nunan (1993, 21) によると、書き手や話し手が文や発話の境界を越えて関係確立し、テキスト内の文を結びつけるのに役立つ語やフレーズのことを指す。例えば、「そして」「しかし」「それ」などがあげられる。

¹⁸ 実際、翻訳普遍性は、結局いずれも反例が存在するため、確率的な傾向、あるいは、普遍性とするのが妥当であるとされてる (cf., Pym 2008; Tymoczko 1998; Chesterman 2004, 2010)。とりもなおさず、一般性といいつつ例外が認めているわけである。

¹⁹ Chesterman (2011) によれば、this hypothesis claims that when processing a given text chunk, translator tend to start from a literal version of the target text, and then work towards a freer version.

他に、テキストの翻訳ではないが、翻訳の認知面の研究では、例えば「リテラル翻訳仮説」(literal translation hypothesis)¹⁹にも触れておくべきである。この仮説に基づき、認知処理と認知負荷の観点からのリテラル翻訳の可能性の度合い (translation literality) の操作化 (@ Carl and Schaeffer 2017) が 1. 原文と訳文の語順の類似性と 2. 可能な異なる翻訳表現の数、といったように考えている。いずれも、本稿でいうテキストとしての追加率、不一致率と異なった視点になっている。その中の 2 点目が筆者らのノンリテラル要素の可視化に関連する可能性があり、検討すべき尺度であった。

また、翻訳研究からの検討について、単位の一致の問題について、従来 translation equivalence (cf., Eugene A. Nida and Taber [1982] 1969) や translation unit (cf., Koller [2004] 1979; Malmkjær 2008) に関する議論が多かった。それらの視座から分析単位の影響を考察しなければならず、検討の余地は大いにある。

5.5 方法論の改善余地

本稿は完全にルールベースで一致率の計算であり、一致の探索の一貫性と簡素さの面では優れている方法であると考えられる一方で、整列法に関しては動的計画法によるアライメントには課題を残している。古語と現代語では、片方が低リソース言語ではあるものの、ニューラルベースな手法でのアライメントやマニュアルによるアライメントも視野に入れるべきである。例えば、Palladino, Shamsian, and Yousef (2022), Camilleri (2024) のように、Palladino, Shamsian, and Yousef (2022) で提案さ

れた [Ugarti](#) フラットフォーム を介したアライメントで、一対多の対応の分析、訳し方の共通パターンと分岐パターン、非対応の品詞情報の分析など、より精緻なコントロールが期待される。

5.6 分類語彙表番号の応用研究として

本稿における（不）一致率・追加率の計算では、言語処理のアライメントモデルではなく安定的な分類番号を用いた。とりわけ、バッグ法は、句や語の入れ替えに関係せず、意味の一致を検出できた。意味の分類体系が、同一の意味の語を階層的に対応づけることの有用性が古語研究で再確認できた。本研究では、まだ旧分類語彙表番号を用いているが、現在 Asahara et al. (2022) によって古語への分類語彙表番号が付与されていることによって広げられる研究の可能性が十分期待できるであろう。ただし、分類語彙表番号は、2024年現時点において概念レベルまでであって、類義語、同義語、同語異表記のメタコード・識別子として一致の検出への対応は、研究者各自で目標に沿ったアダプテーションが求められている。本稿で採用している意味付与もまた、その修正を重ねていく必要がある。この点において、意味体系の作り方に関して検討する余地があると考えている。

6 結論

本稿では、古今集の現代語訳から原文のノンリテラル要素の可視化の材料として有用である根拠を示すために、古今集の現代語訳 10 種類における追加要素の分析を行った。

具体的に、古今和歌集の現代語訳の概要について説明し、20 世紀における和歌の現代語訳の翻訳アプローチを、注釈書における訳者の記している内容に基づいた整理を行った。訳者の翻訳の意識は、コミュニケーションモデルの視点から、歌人の作意 (source) を重視する、テキストの文字通りの意味 (signal) を重視する、読者 (destination) を重視する、とその他明確でないもの 4 種類に分類できた。

それらの主観的な翻訳意識・翻訳アプローチが訳における追加と一致の客観的に指標に反映されているかについて、Yamamoto and Hodošček (2019) で行った追加率の計算手順を明示的にし、より明確な設定と更新データを用いた再計算を行った。古語と現代語訳の語レベルの非一致率、現代語訳における追加率の計算と統計モデリングにより、翻訳アプローチにかかわらず現代語訳における情報の追加が普遍的であり、訳者間の追加率の変動が歌間の追加率の変動より小さいことを示した。

追加要素が十分ノンリテラル要素として成立するかを検討するために、298 番「立田」歌の現代語訳の事例を 10 つ同時にアライメントとして提示し、対象語のコーパ

スレベルの共出現語、共出現構文パターン、共出現語の意味志向がセンテンスレベルの現代語訳でどのように処理されているかを調査した。その結果、コーパスレベルの共出現傾向をもつ単位の場合、センテンスレベルでの処理が10人の間で安定的な処理が見られた。コーパスレベルで共出現傾向にない要素の追加がよりバリエーションを示している。とりわけ注意すべき現象は、コーパスレベルの共出現語が訳の推論に用いられることと、コーパスレベルの共出現語の意味志向（全体的な意味の傾向）に関連する語が訳においてより多く補足的説明をもつことであった。一方で、構文面の追加は、言語変化のギャップを埋めるためのものであった。

以上により、和歌の現代語訳が、たとえ直訳・逐語訳に拘っているとしても、和歌辞典とは異なる解釈材料として、和歌のノンリテラル情報を補足説明するために利用できることを論じた。しかし同時に、言語変化のギャップを埋めるための要素の除外や、翻訳における信頼性の高い要素の取り立て方の工夫の重要性を示唆している。

データの公開利用

公開できるデータである金子 (1933) による現代語訳の分割語彙データは XXXXX にて参照する。

補足材料

補足資料は、[https://github.com/idiig/replication-test/blob/main/supplementary-materials/補足資料 Supplementary materials.pdf](https://github.com/idiig/replication-test/blob/main/supplementary-materials/補足資料%20Supplementary%20materials.pdf) を参照する。再現実験について、<https://github.com/idiig/replication-test/tree/main> を参照する。

謝辞

本研究は、XXXXX の助成を受けたものである。

References

- 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: Günter 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>.
- Bürkner, Paul-Christian. 2017. "**Brms** : An R Package for Bayesian Multilevel Models Using Stan." *Journal of Statistical Software* 80 (1). <https://doi.org/10.18637/jss.v080.i01>.
- 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 *Digital Humanities 2022 Conference Abstracts*, 647–48. Tokyo: The University of Tokyo.
- Koller, Werner. [2004] 1979. *Einführung in die Übersetzungswissenschaft*. 7., aktualisierte Aufl. Wiebelsheim: Quelle und Meyer.
- Malmkjær, Kirsten. 2008. "Unit of Translation." In *Routledge Encyclopedia of Translation Studies*, edited by Mona Baker and Kirsten Malmkjaer, 286–88. London: Routledge.
- 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>.
- R Core Team. 2024. "R: A Language and Environment for Statistical Computing." Vienna, Austria.
- 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.
- Shiozawa, Kazuko. 1993. "Motoori Norinaga, "Kokinshu Tokagami" Ni Okeru Keijo Jodoshi: 'Shiyaru', 'Sashiyaru' o Megutte (Motoori Norinaga's Humble Auxiliary Verb in the Kokinshu Tokagami: 'Shiyaru', 'Sashiyaru')." *Kindaigo Kenkyu (Study of Contemporary Vocabulary)* 9: 143–60.
- 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"* 99

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>.

久曾神昇. 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.

Reuse

[CC BY 4.0](#)

Citation

BibTeX citation:

```
@online{chen2024,
  author = {Chen, Xudong and Hodošček, Bor and Yamamoto, Hilofumi},
  title = {古今和歌集の現代日本語訳における追加率要素の分析},
  date = {2024-09-01},
  doi = {10.17928/jjadh.0.0_0},
  langid = {en},
  abstract = {注釈書における翻訳における要素の追加を利用した古今和歌集の歌ことばのノンリテラル情報の可視化をこれまで行われてきたが、それら注釈における現代語訳や、現代語訳における追加要素に関する詳細の分析が示されていない。本稿の目的は、いままでの研究に用いられてきた古今集の現代語訳
10
種が、ノンリテラル要素の可視化に適している材料であるかについて詳細な分析を行うことである。まず、古今集の注釈における現代語訳について概観し、訳者の翻訳の意識についてまとめた。つぎに、直訳語をもたない歌ことばが原文に占める割合（不一致率）と現代語訳における追加要素の割合（追加率）の計算で、現代語訳の要素の追加と一致の実態を把握した。最後に、具体例を示し、現代語訳における追加要素の事例分析を行った。
分析の結果、和歌の現代語訳の訳者の翻訳意識は、主観的に、作意中心・テキスト中心・読者中心と不明
4
類に分類できた。しかし、客観的には、翻訳意識が異なっても、計量的には同じ程度の追加要素を多く含めていた。B
分析事例においては、センテンスレベルの訳においても、要素の追加にコーパスレベルの共出現傾向に基づく推論があり、語の意味の志向性について補足要素を多く足していることが確認できた。以上により、現代語訳をノンリテラルレベルの可視化に適用できることを示した。
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 unmatched 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.}

}

For attribution, please cite this work as:

Chen, Xudong, Bor Hodošček, and Hilofumi Yamamoto. 2024. “古今和歌集の現代日本語訳における追加率要素の分析.” *Journal of Japanese Association for Digital Humanities*. September 1, 2024. https://doi.org/10.17928/jjadh.0.0_0.