The Metronome Approach to Sanskrit Meter: Analysis for the Rigveda

Yuzuki Tsukagoshi
Univeristy of Tokyo / Tokyo, Japan
yuzuki@l.u-tokyo.ac.jp

Ikki Ohmukai
Univeristy of Tokyo / Tokyo, Japan
i2k@l.u-tokyo.ac.jp

Abstract

This study analyzes the verses of the Rigveda, the oldest Sanskrit text, from a metrical perspective. Based on metrical structures, the verses are represented by four elements: light syllables, heavy syllables, word boundaries, and line boundaries. As a result, it became evident that among verses traditionally categorized under the same metrical name, there are those forming distinct clusters. Furthermore, the study reveals commonalities in metrical structures, such as similar metrical patterns grouping together despite differences in the number of lines. Going forward, it is anticipated that this methodology will enable comparisons across multiple languages within the Indo-European language family.

1 Introduction

The oldest of the Vedic literature, the Rigveda, possesses a metrical structure. Metrical analysis involves examining the patterns of syllables, word boundaries, and poetic line boundaries, which are foundational elements in the study of ancient Indo-European poetry. Metronome analysis that is solely based on the metrical structure was demonstrated in Latin and some languages (Section 3). Similar to Latin, Vedic Sanskrit also considers the elements of syllable weight, word boundaries, and poetic line boundaries as components of its metrical structure (Section 2). By focusing on the metrical structure, the same analysis as Latin can be applied to Sanskrit, especially Vedicpoetry, as discussed in Section 4. Consequently, analyzing the Rigveda based on its metrical elements revealed that even when assigned the same metrical name, they clearly belong to different clusters (Section 5). This research enables us to uncover patterns and structures that are not immediately apparent through traditional philological or linguistic methods (Section 6).

2 Vedic Meter

Sanskrit, belonging to the Indo-Aryan branch of the Indo-Iranian subfamily within the Indo-European language family, has literature with a metrical structure. In particular, the Vedic literatures, in which Vedic Sanskrit is used, an ancient stage of the Sanskrit language, are known for their metrical composition. The calculation of meter in Sanskrit is syllabic; that is, Sanskrit poetry shapes its rhythm based on the number of syllables and their weight.

In this context, an overview of the meter in the Rigveda, which is used for analysis, is provided. The Rigveda is the oldest among the Vedic literature and consists of metrical verses. In the tradition of Rigveda studies, "meter" refers to the rhythmic pattern of entire verses composed of lines with a consistent rhythm. The meters of the Rigvedic verses are mentioned in traditional literature written in Sanskrit (Macdonell, 1886), Anukramaṇī.

Table 1 lists some primary meters. Taking the Triṣṭubh meter as an example, verses with this meter consist of four lines, each having eleven syllables. Additionally, in each line, the last four syllables repeat the pattern –heavy + light– twice. Generally, heavy syllables are represented as —, light syllables as ⏑, and when not specified as heavy or light, it is denoted as ⏒. Vedic metrical texts, including the Rigveda, are recited, and in the context of meter calculation, the last syllable of each line is pronounced long, whether it is heavy or light.

Although metrical patterns refer to the rhythmic patterns of entire verses in the Rigveda, the fundamental unit in actual structural patterns is the individual line. As shown in Table 1, when the number of syllables per line is determined, the cadence, which is the latter part of the line, tends to follow a certain syllabic pattern. Specifically, for an eight-syllable line, there is a repetition of the *light and heavy* pattern ⏑ — ⏑ ⏒, for an elevensylalble line, a repetition of the *heavy and light*

Proceedings of the 1st Workshop on Machine Learning for Ancient Languages (ML4AL 2024), pages 219–223

Table 1: Poplular meters

Meter	Syllables per line	Lines in a verse	Syllabic pattern in the cadence
Triṣṭubh	11	4	— ∪ — ⏑
Gāyatrī	8	3	∪ — ∪ ⏑
Jagatī	12	4	— ∪ — ∪ ⏑
Anuṣṭubh	8	4	∪ — ∪ ⏑

pattern — ∪ — ⏑, and for a twelve-syllable line, a repetition of the *heavy and light* pattern twice, followed by one syllable — ∪ — ∪ ⏑. However, despite the relatively strict regularity of cadence patterns, there are irregular lines that deviate from the presented patterns (Oldenberg, 1888). Additionally, for eight-syllable lines, such as Gāyatrī and Anuṣṭubh, there are cases where the pattern is reversed, resulting in a repetition of the *heavy and light* pattern — ∪ — ⏑. [1]

In addition to these, there exists a less strict regularity in lines with eleven and twelve syllables. The forepart of lines of these syllable counts are further divided into two parts: opening and caesura. This division often aligns with word boundaries. Moreover, the rhythm patterns in the initial part of caesura also exhibit some regularity (Oldenberg, 1888; Arnold, 1905).

The representative studies of meters in the Rigveda (Oldenberg, 1888; Kuryłowicz, 1927; Arnold, 1905; Gippert, 1997, 1999) have frequently focused on aspects such as syllable counts and syllable weight patterns in the cadence. This emphasis arises from the strong tendency in the Rigveda to strictly determine the rhythm based on syllable counts and syllabic weight patterns in the cadence.

3 Related Research

Nagy et al. (2023) analyzed Latin poetry using four elements known as the *metronome*: *light syllable*, *heavy syllable*, *word boundary,* and *poetic line boundary*. Following the metronome analysis code, we represent light syllables as 'w', heavy syllables as 'S', word boundaries as '.', and poetic line boundaries as '|'. By substituting these elements in Latin poetry, the verse is considered as a sequence of these four elements. Texts composed of these four metrical elements are similar to genes consisting of four types of base pairs, akin to genetic sequences.

The study of Vedic meter made significant advancements following the research of Oldenberg (Oldenberg, 1888). About a century later, Gippert pioneered the computer-based analysis of meters (Gippert, 1999). Recent research on Vedic meter (Ittzés, 2012; Beguš, 2015), including the studies mentioned above (Oldenberg, 1888; Gippert, 1997, 1999), focuses on the phonology of Vedic Sanskrit and the historical changes of sounds from Proto-Indo-European to Sanskrit. While these studies focus on phonological aspects and the development of phonological theories based on metrical patterns, there has been limited exploration of Vedic meters from a stylometric perspective. Our research aims to bridge this gap by applying a stylometric approach to the analysis of Vedic poetry, inspired by the metronome analysis of Latin poetry.

4 Method

This study follows the metronome analysis proposed by Nagy (2023); Nagy et al. (2023). Similar to their approach, it converts all verses in the Rigveda into four elements of the metronome, 'w' for a light syllable, 'S' for a heavy syllable, '.' for a word boundary, and '|' for a verse boundary, to perform metrical analysis. An example of the transformation of Rigveda text into the metronome is provided in Table 2[2]. The skeleton structure represents short vowels by V, long vowels by W, and consonants by C.

Table 2: Examples of transformation into metronome

Text	*agním īḷe puróhitaṃ*
Skeleton	VCCVC WCW CVCWCVCVC
metronome	Sw.SS.wSwS|

Specifically, the steps for the analysis are as follows. The electronic text of the Rigveda (Martínez García and Gippert, 1995) is utilized, and it undergoes a transformation into the metronome. The metronome sequences are then subjected to a score calculation using the Python module

[1] Known as Trochaic Gāyatrī.

[2] The detailed steps for transformation are outlined in Section A.

metronome[3]. Subsequently, a hierarchical clustering analysis is performed on the metronome sequences of Rigvedic verses[4]. For distance calculations, Euclidean distance, normalized Euclidean distance, Chebyshev distance, and Minkowski distance are employed. As for linkage methods, average, centroid, complete, median, single, and ward are used. All possible combinations of these distances and methods, resulting in twenty-four variants, are employed for clustering. By employing all possible combinations of these distances and linkage methods (resulting in twenty-four variants), the study aims to ensure robustness in the clustering results.

For clarity in the analysis results, the scope of the text is limited. Since the Rigveda comprises ten books with a total of over 10,000 verses, the analysis is conducted book by book. Notably, Books 2 to 7 of the Rigveda are known as *family books*, consisting of verses attributed to a single poetic family. Family books are considered to contain relatively more ancient verses of the Rigveda, making them particularly significant (Oldenberg, 1888; Arnold, 1905).

5 Result

First, we shall focus on books 2 to 7 of the Rigveda, known as family books, and also include the first and second halves of book 8. [5]

Figure 1 shows the result of clustering on Book 7, whose author is Vasiṣṭha[6]. The color threshold is set to 15. While the upper three groups (brown, purple, and red) are mainly Triṣṭubh verses, the green group contains Triṣṭubh and Jagatī verses. Distant from these, the bottom yellow group contains eight-syllable verses. The other family books also have the tendency that some of Triṣṭubh verses are more similar to Jagatī than other Triṣṭubh and eight-syllable verses are close despite their difference in the number of syllable.

Upon closer examination of eight-syllable verses, it can be seen that Gāyatrī, Anuṣṭubh, Paṅkti (= 8 syllables × 5 lines), and Bṛhatī (= 8 syllables × 2 lines followed by two-syllable line and ending with eight-syllable line) form a cohesive group (Figure 2). However, it is not possible to discern the difference between Gāyatrī in ⏑ — tone and Gāyatrī in — ⏑ tone here.

6 Conclusion

The analysis in this study revealed that some verses with the Triṣṭubh meter, traditionally considered to have rhythmical pattern similar to that of the Jagatī meter (Arnold, 1905; Van Nooten and Holland, 1994)., are distinctly different from standard Triṣṭubh verses. Even though Triṣṭubh verses were traditionally labeled as such, it was empirically known that some lines within these verses exhibited the rhythmic pattern of Jagatī meter lines in the cadence. In practice, scholars have sometimes categorized Triṣṭubh and Jagatī as *trimeter* (= opening + caesura + cadence) and Gāyatrī, Anuṣṭubh, and so on as *dimeter* (= opening + cadence), respectively, while at other times making clear distinctions based on syllable counts.

Regarding eight-syllable verses, Gāyatrī, Anuṣṭubh, Paṅkti all share the commonality of having eight syllables per line, with the only difference being the number of lines. Generally, it is believed that the rhythmic pattern is determined by the number of syllables per line, regardless of the number of lines. This study reveals that various eight-syllable verses do form cohesive clusters, indicating that the rhythmic pattern is indeed primarily influenced by the number of syllables per line. Additionally, a closer look at this cluster reveals that Gāyatrī and Bṛhatī meters each form distinct subclusters, showcasing rhythmic similarities based on differences in the number of lines. Unlike previous studies mentioned in the section 3 that focus solely on cadence, our research examines the entire verse, leading to these significant findings of the similarity of different syllable counts. This comprehensive approach highlights the importance of considering the whole verse.

This study contributes to Vedic philology by providing a nuanced understanding of the rhythmic structures within Rigvedic verses. By identifying distinct clusters and subclusters based on metrical patterns, our research contributes to a better understanding of the development of Vedic meters over time and can offer insights into the chronological aspects of Rigvedic verses.

While this study primarily focuses on the

[3] https://github.com/bnagy/metronome

[4] Scripts for converting the Rigveda into metronome and for a clustering analysis are accesible in https://github.com/Yuzki/metronome_veda

[5] Book 8 consists of the first half by the Kaṇva family and the second half by the Aṅgirasa family. Therefore both are often treated as ones of family books.

[6] Due to space constraints, it is not possible to present all patterns in the figure. Representative examples are shown.

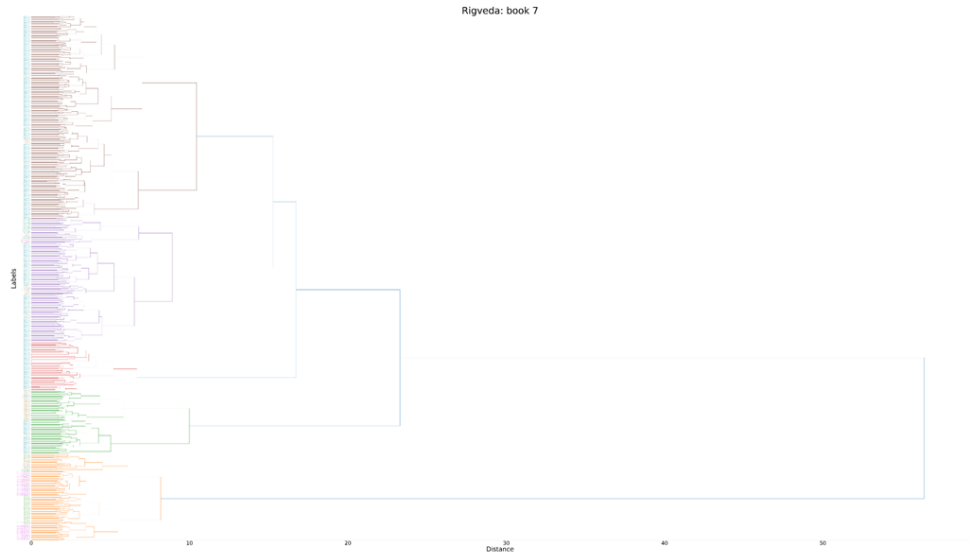

Figure 1: Rigveda Book 7, Euclidean distance, Ward's method

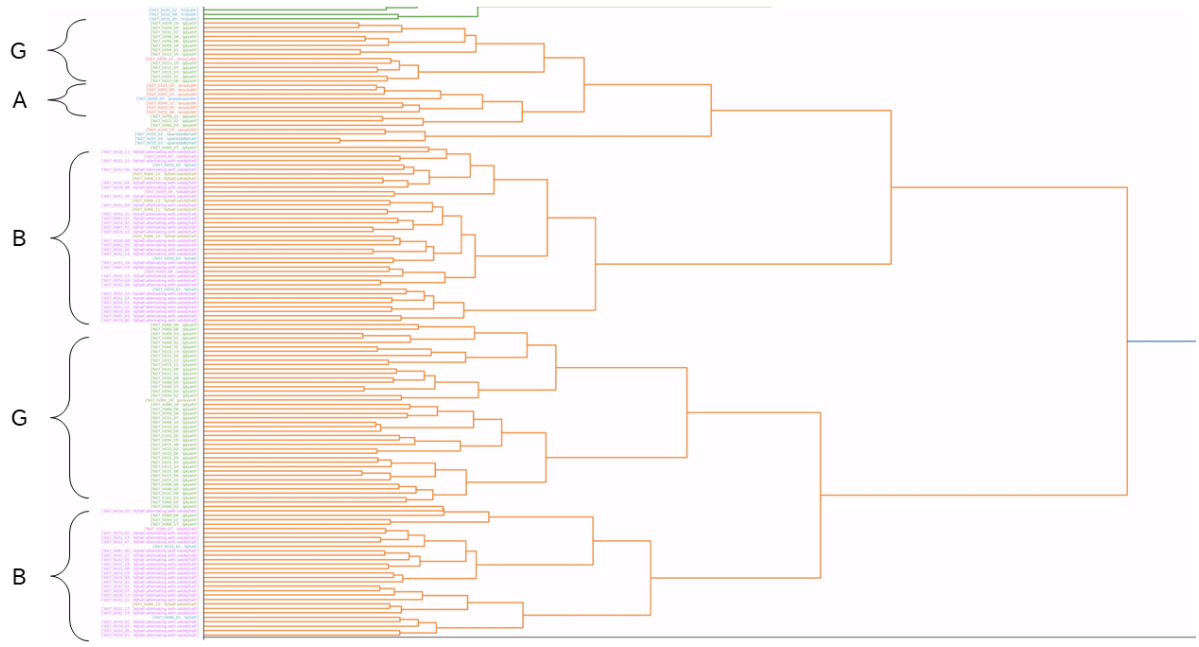

Figure 2: Enlarged view of the Gāyatrī section from Figure 1 (G: Gāyatrī, A: Anuṣṭubh, B: Bṛhatī)

Rigveda, the methodology and findings have broader implications for other Vedic texts and even extend to other Indo-European poetries. The rhythmic patterns and clustering identified in this study could serve as a model for analyzing similar meter structures in other Vedic texts, as well as later Sanskrit literatures.

Acknowledgments

This work was supported by JSPS KAKENHI Grant Number 23K18646.

References

Edward Vernon Arnold. 1905. Vedic metre in its historical development.

Gašper Beguš. 2015. A new rule in Vedic metrics. *Journal of the American Oriental Society*, 135(3):541–550.

Jost Gippert. 1997. Laryngeals and Vedic metre. In *Sound law and analogy. Papers in honor of Robert S.P. Beekes on the occasion of his 60th birthday*, pages 63–79. Amsterdam: Rodopi.

Jost Gippert. 1999. Neue Wege zur sprachwissenschaftlichen Analyse der vedischen Metrik. In *Compositiones Indogermanicae in memoriam Jochem Schindler*, pages 97–125. Praha: Enigma.

Máté Ittzés. 2012. Initial y in the Rigveda. In Roman Sukač and Ondřej Šefčik, editors, *The Sound of Indo-European 2. Papers on Indo-European Phonetics, Phonemics and Morphophonemics*, volume 41 of *LINCOM Studies in Indo-European Linguistics*, pages 86–115. LINCOM, München.

Jerzy Kuryłowicz. 1927. Les effets du ə en indoiranien. *Prace Filologiczne*, 11:201–243.

Arthur Anthony Macdonell. 1886. *Kâtyâyana's Sarvânukramanî of the Rigveda: with extracts from Shadgurusishya's commentary entitled Vedârthadîpikâ*. Anecdota oxoniensia: Aryan series. Clarendon Press, Oxford.

Francisco Javier Martínez García and Jost Gippert. Thesaurus indogermanischer text- und sprachmaterialien [online]. 1995.

Benjamin Nagy. 2023. Metre matters. In *DH2023 Workshop – SIG-DLS Seven Years on*.

Benjamin Nagy, Artjoms Šeļa, Mirella De Sisto, Wouter Haverals, and Petr Plecháč. 2023. Metronome: tracing variation in poetic meters via local sequence alignment. Poster presented at the CHR2023.

Hermann Oldenberg. 1888. *Metrische und textgeschichtliche Prolegomena*. Berlin: Verlag von Wilhelm Hertz.

Barend A. Van Nooten and Gary B. Holland. 1994. *Rig Veda: a metrically restored text with an introduction and notes*. Cambridge, Massachusetts: Department of Sanskrit and Indian Studies, Harvard University.

A Transformation Method to Metronome

This section demonstrates the method of transforming verses of the Rigveda into metronome format.

A.1 Metrical Calculation

First, we verify the metrical calculation method for Sanskrit, specifically clarifying the syllable structure and the definitions of light and heavy syllables.

Syllable formation spans word boundaries. For example, when a word ends with a consonant and the following word begins with a vowel, the final consonant of the first word and the initial vowel of the second word together form one syllable. In the case of *agním īḷe* as seen in Table 2, the final consonant *-m* of *agním* forms a syllable with the initial vowel *ī-* of the following word *īḷ*. Thus, dividing *agním īḷe puróhitaṃ* into syllables, we get the following:

ag	*ní*	*mī*	*ḷe*	*pu*	*ró*	*hi*	*taṃ* |
VC	CV	CW	CW	CV	CW	CV	CVC.

The structure in the second line is called the skeleton structure, consisting of consonants (C) and vowels (short vowel V, long vowel W).

The lightness or heaviness of a syllable is determined by the length of the syllable nucleus vowel and the presence or absence of a final consonant. A light syllable has a short vowel as its nucleus and no final consonant. A heavy syllable has a long vowel as its nucleus or a final consonant. Based on this, the skeleton structure text shown above has the following light ⏑ and heavy — syllable patterns:

ag	*ní*	*mī*	*ḷe*	*pu*	*ró*	*hi*	*taṃ* |
VC	CV	CW	CW	CV	CW	CV	CVC.

A.2 From Skeleton Structure to Metronome

Using the skeleton structure seen in the previous section, we perform the conversion to metronome. In the original text, when there is a word boundary between the nuclei of two adjacent syllables, this word boundary is noted between the corresponding skeleton structures. These word boundaries within the defined skeleton structure are provisional. This is because word boundaries in the original text and those in the skeleton structure do not necessarily match, as syllable formation spans word boundaries. Additionally, due to a phonological phenomenon called sandhi, sounds at the boundary of adjacent words may merge. In such cases, the merged entity is not separated back into the original individual words.

Using the method shown so far, the original text can be transformed into metronome format as follows. For readability, light and heavy syllables are indicated by ⏑ —, word boundaries by #, and verse boundaries by /.

ag	*ní*		*mī*	*ḷe*		*pu*	*ró*	*hi*	*taṃ*,	
VC	CV		CW	CW		CV	CW	CV	CVC,	
—	⏑	#	—	—	#	⏑	—	⏑	—	/.