

# TEI を用いた『リグ・ヴェーダ』のマークアップについて

塚越柚季<sup>1,a)</sup>

概要：『リグ・ヴェーダ』(RV)をはじめヴェーダ文献はマークアップにより構造化されたデジタルテキストがほぼ存在しない。そこで RV のマークアップを行い、各詩節の韻律およびその詩を作成したと伝承されている詩人の情報を加えた。韻律詩である RV は、伝承されるテキストが想定される韻律に当てはまらない場合がある。そのような詩は、インド・ヨーロッパ祖語やインド・イラン祖語の再建により韻律が復元される。これら伝承されるテキストおよび韻律の復元されたテキストをまとめて構造化することに関しても議論をする。

## Markup of “Rigveda” using TEI

### 1. サンスクリット文献のデジタルテキスト

#### 1.1 古典サンスクリット

サンスクリットは、ヴェーダ文献に現れる言語と、それ以降の時代の、1つの規範的な文法に則った言語とに大別される。前者がヴェーダ期サンスクリット、後者が古典サンスクリットである。

現在、古典サンスクリットで表された文献およびヴェーダ期サンスクリットで表された文献の一部のデジタルテキストが利用可能である。TITUS [7] や GREITL [5] は幅広い種類の文献のデジタルテキストを公開している。しかし、いずれもフラットなテキストファイルしか存在しない。

一方、SARIT [6] は TEI によってマークアップされたファイルを公開している。ただし、文献の種類は前述の TITUS, GREITL より少なく、記述の詳細の程度もまちまちである。

#### 1.2 ヴェーダ期サンスクリット

ヴェーダ文献のデジタルテキストも古典の文献同様、多くは TITUS, GREITL で利用が可能である。しかし、SARIT には『リグ・ヴェーダ』(RV)をはじめとしたヴェーダ文献が存在しない。

RV 他いくつかの文献に関して、TITUS には構造化されたテキストデータが存在するように見える。ウェブページ上のテキスト中の語から、その語の語形(名詞・形容詞の

語幹・性・数・格や動詞の語根・時制・法・人称など)情報にアクセスすることができる。さらに RV 中で同じ語形が現れる他の箇所への行の情報も提供する。難点としてこれらはオンラインでのみ利用可能で、各文献のデジタルテキストの構造のしくみが明らかではないため、特に量的研究における応用の可能性が低い。

韻律復元をした RV は、デジタルテキストが書籍 [4] に付随している。書籍では韻律復元にあたって変更した箇所を斜体にする事で明記している。しかしながら、デジタルテキストでは、変更を加えた箇所もそれ以外の箇所も同一に打ち込まれている。そのため、テキスト中の韻律復元した箇所が明らかでないという問題がある。

### 2. 『リグ・ヴェーダ』のテキストの構造化

#### 2.1 『リグ・ヴェーダ』について

『リグ・ヴェーダ』(RV)とはヴェーダ文献の中でも最古の文献である。つまり、RV で用いられる言語は、現状、最も古いサンスクリットである。RV は、韻律を持つ詩であり、10巻から構成され、それぞれの巻はいくつかの讃歌から成る。特に2巻~7巻は、主にそれぞれ一つの詩人家系が作成したとされている。また、これらの巻の一部および他の巻は、讃歌ごとなし讃歌を構成する詩節ごと、詩節を構成する詩連ごとに様々な詩人が作成したとされる。

RV は初め口頭によって代々伝承され、後の時代になって固定・編纂された。この時代間の言語の差異のために、現在利用可能な RV には、想定される本来の形とは異なる部分が含まれていると考えられる。その中に RV の韻律

<sup>1</sup> 東京大学大学院人文社会系研究科

<sup>a)</sup> juzki13@gmail.com

の復元がある。RV の韻律は規則的であり、伝承されたテキストの韻律が想定される韻律とは異なる場合は、祖語形に遡って韻律を復元することがある。韻律の復元の方法は様々で、また韻律の乱れた同一の箇所に対して異なる復元が提唱されることもある。

## 2.2 韻律研究にあたって必要な情報

RV の韻律はひとまとまりの詩節をもとに各詩行の音節数、音節の軽重が想定されるため、詩行のみならず詩節も一つの単位とする必要がある。詩節に含まれる詩連の韻律は、例えば、Gāyatrī (8 音節 ×3 行)、Triṣṭubh (11 音節 ×4 行)、Anuṣṭubh (12 音節 ×4 行) のような韻律の類型に分類される。加えて、多くは詩節単位で詩人が異なるため、ある詩節に対してそれを作成した詩人が対応する。

本稿では、「語ごとの韻律の復元方法の適用可否と該当する行を含む詩連を作成した詩人との間に関係がある」という執筆者の仮説を検討するために必要なテキストの作成を試す。そのために上記の詩節ないし詩連ごとの韻律と詩人の情報が必要となる。また、伝承されるテキストの各詩行の音節の軽重は当然ながら必要である。

## 2.3 作成したテキストについて

前節のとおり、詩行に対しては音節の軽重の情報を、詩連・詩節に対しては作成した詩人の情報を与える。公開されている RV のデジタルテキスト [2] および、詩連・詩節と詩人との対応表によって自動的にマークアップが可能である。

例として RV 第 1 巻の冒頭を挙げる。詩節を作成した詩人は、<note> を用いて本文に入れず注釈として追加する。また RV の校訂版 [1] や韻律復元版 [4]、その他翻訳に倣い、その詩の讃える対象の神も詩人同様に追加する。

```
<text xml:id="RV1">
  <group>
    <text xml:id="RV1.1">
      <note type="poet">
        <name type="poet"
          nymRef="#madhucchandasa_vaisvamitra">
          madhucchandasa vaiśvāmitra
        </name>
      </note>
      <note type="god">
        <name type="god"
          nymRef="#agni">
          agni
        </name>
      </note>
    </body>
```

```
      :
    </group>
  </text>
  RV 本文については @real によって行全体の音節の軽重を示す。
  <body>
    <lg n="1" xml:id="RV1.1.1">
      <l real="+-+-+--+" xml:id="RV1.1.1a">
        agnīm iḍe puróhiṭam
      </l>
      :
    </lg>
    :
  </body>
```

## 3. 問題点

### 3.1 連声

サンスクリットには、文中における語の隣接部で特定の音連続があった場合、隣接部の音が変化する連声がある。この変化は伝承されるテキストにそのまま反映される。

連声規則の中に、文中において語の隣接部が母音である場合、つまり母音が連続する場合、これらの母音が融合するという規則がある。例えば、RV 1.2.1 の a, b 行を見る。

- (1) vāyav ā yāhi darśatēmé sómā āraṃkṛtāḥ
- (2) vāyav ā yāhi darśata imé sómā āraṃkṛtāḥ

(1) は連声が適用された形で伝わる RV (サンヒターパータ) の 1 行で、(2) は下線部の連声を解除して連声適用前の形に戻したものである。

### 3.2 連声にかかる問題

前節で示したように、2 単語に分ける必要がある場合も母音の融合により、分けることが困難になる。RV の構造化テキストの作成にあたってこれが問題となる。サンヒターパータにおいて、行をまたいで母音融合の連声がある場合、韻律上、根拠を持って多くは連声を起こさない形に復元ができる。

RV の校訂版 [1] との差異を明示することを踏まえると、連声が適用された語と語に対して、韻律復元した語を対応させてマークアップする必要がある。連声に関する場合以外もいずれの場合も、該当箇所の語の祖語形や祖語からサンスクリットにかけての音変化などを考慮して韻律の復元を検討せねばならない。

### 3.3 韻律復元のマークアップ法

連声を含んだ問題は当然ながら既に議論が行われており、マークアップの方法も提案されている [3]。<choice>

エレメントを用いて、連声適用後の語形と連声適用前の語形を並列して記述するとしている。この方法ではさらに、後の末尾に着目して続く語とどのように切り離されるかについてのレベルを記述する。

韻律復元に際する連声解除は、文中の語の分析による連声解除とは異なる点がある。韻律の復元にあたって、例えば、祖語形で含まれていた音素の存在ゆえに RV の時代にもそれが残存し、そもそも連声を引き起こさないという場合がある。このことを考慮に入れると、語の末尾にのみ着目するのでは不十分であり、その語に続く語の先頭も着目する必要がある。実際に、RV の韻律復元版 [4] では次のように語の末尾と先頭がともに斜体によって表される。

darśata imé (校訂版 [1] では darśatemé \*<sup>1</sup>)

このように、韻律分析のための、TEI を用いたマークアップによる RV の構造化においては、連声に関わる語すべてに対して、そして語レベルのみならず音レベルでの詳細な記述も必要となる。

## 参考文献

- [1] Aufrecht, T.: *Die Hymnen des Rigveda*, (1861).
- [2] Martínez García, F. J. and Gippert, J.: Plain text retrieval, Thesaurus Indogermanischer Text- und Sprachmaterialien, 入手先 (<http://titus.fkidg1.uni-frankfurt.de/private/texte/indica/vedica/rv/pp/rvarpp.txt>) (2019.04.10).
- [3] Smith, John D.: The ‘end of word’ problem in Sanskrit: report of the workgroup, 入手先 (<https://teic.org/Vault/Workgroups/CE/cew12.pdf>) (2019.4.11).
- [4] van Nooten, B. A. and Holland G B.: *Rig Veda, a metrically restored text with an introduction and notes*, (1994).
- [5] Göttingen register of electronic texts in Indian languages, 入手先 (<http://gretel.sub.uni-goettingen.de>) (2019.04.10).
- [6] Search and retrieval of Indic texts, 入手先 (<http://sarit.indology.info>) (2019.04.10).
- [7] Thesaurus Indogermanischer Text- und Sprachmaterialien, 入手先 (<http://titus.uni-frankfurt.de/indexe.htm>) (2019.04.10).

---

\*<sup>1</sup> ś と ṣ はローマナイズの方法が違うだけであり、実際は同一の音素