Using TEI for digital Sanskrit editions containing commentaries
A study of Kālidas's *Raghuvaṃśa* with Mallinātha's *Sañjīvani*

Tanuja P. Ajotikar
Assistant Professor, The Sanskrit Library
Co-director Sanskrit Department India Programs,
Maharishi Vedic University, Vlodrop, The Netherlands
`tanuja@sanskritlibrary.org`

Ketaki Kaduskar
Research Assistant, The Sanskrit Library
`ketaki.kaduskar@gmail.com`

Peter M. Scharf
President, The Sanskrit Library
Adjunct Professor, IIIT Hyderabad
`scharf@sanskritlibrary.org`

Abstract

In the present project of making a TEI digital edition of Mallinātha's commentary *Sañjīvanī* on Kālidāsa's *Raghuvamīśa*, we encountered material that called for expansion of the procedures for creating XML editions of Sanskrit texts containing commentaries in accordance with the Text Encoding Initiative (TEI) guidelines. Ajotikar and Scharf (2023) previously described standardized procedures for digitizing the Sanskrit texts containing commentaries. To accommodate the additional material encountered in the present project we created several additional values of the `seg` element's `type` attribute: `alternate`, `def`, `paraphrase`, `syntax`, and `constituent`. These expand the categorization of glosses, refine categorization of derivational material, and deal more precisely with comments on compound constituents. We also verify, correct and make additional citations thereby making substantial contributions to textual research in the process of undertaking the higher-level encoding of the commentary.

Keywords: XML, TEI, Sanskrit, commentary, text-encoding

1 Introduction

Ajotikar and Scharf (2023) described standardized procedures for creating XML editions of Sanskrit texts containing commentaries in accordance with the Text Encoding Initiative (TEI) guidelines and discussed issues related to it that arose in their mark-up of Bhaṭṭi's *Rāvaṇavadha* with Mallinātha's commentary. In this paper, we present additional issues that arose in the TEI mark-up of Kālidas's *Raghuvamīśa* with Mallinātha's commentary *Sañjīvanī*. We completed mark-up of the first five contos of the *Raghuvamīśa* and the first three cantos of the *Sañjīvanī*.

2 File preparation procedure

The file preparation procedure includes three phases. In the first phase, the base text and the commentary are transcoded into the Sanskrit Library Phonetic ASCII encoding (SLP1), each in its own file. In the second phase, sandhi-analysis is conducted on the base text in a separate file. During this process, each word in the base text file is assigned an `xml:id` which serves as the key to co-ordinate the two files. The commentary is tagged in a third file which references words in the base text by referring to the `xml:id` of the word as the value of the `corresp` attribute.

3 Details of data preparation

Considerable effort is required to obtain quality character data of the base text and commentary. First of all, when creating a digital edition, it is necessary to consider the copyright issue. To

simply reproduce a printed edition in the digital medium without adding significant knowledge without permission of the holder of a valid copyright would violate intellectual property rights. However, no intellectual property rights would be violated by making a digital copy of an edition in the public domain. Nor would intellectual property rights be violated by producing a new edition with significant new knowledge added. The editions we considered are all in the public domain. We chose to tag Kale's (1922) edition. Although character data of the *Raghuvaṁśa* is easily found on the Web, character data of Mallinātha's commentary is not. To type the commentary would be time-consuming and expensive compared with deriving character data by optical character recognition (OCR). However, the OCR output available at `archive.org` produced several years ago was junk. Since then OCR software that produces relatively good output of a Devanāgarī printed page has become available on line. Initially we processed the scanned images of Kale (1922) at `https://ocr.sanskritdictionary.com`. Yet proof-reading this output revealed that it needed extensive correction which significantly affected the pace of the project. We then discovered a few digital editions of the *Sañjīvanī*, one of which with Devanāgarī character data is available at `https://www.ebharatisampat.in`. We transcoded into SLP1 the first five sargas of this edition, the source of which is not identified, and edited it to conform with Kale's (1922) edition. Thereby we obtained creditable digital character data of the *Sañjīvanī* efficiently.

3.1 The saṁhitā file

We refer to the file containing the continuous base text without sandhi analysis as the saṁhitā file. In this file, each verse of the base text is analysed metrically. Although the generic metrical pattern of each canto is primarily the same, identifying the meter-type in cantos composed in the Upajāti meter also recognized the specific subtype of each verse. This information is not documented in any previous digital or printed edition of the *Raghuvaṁśa*.

Figure 1 shows a sample of the mark-up of a verse, and Table 1 shows the metrical patterns employed in the first five cantos.

Figure 1

TEI encoding of metrical patterns in Kālidāsa's Raghuvaṁśa

```
<lg type='fdDiH' xml:id='s2.v1'>
 <l>
  <seg n='a'>aTa prajAnAmaDipaH praBAte</seg>
  <seg n='b'>jAyApratigrAhitaganDamAlyAm</seg>
  <pc>.</pc>
 </l>
 <l>
  <seg n='c'>vanAya pItapratibadDavatsAM</seg>
  <seg n='d'>yaSoDano Denumfzermumoca</seg>
  <pc>..2.1..</pc>
 </l>
</lg>
```

Table 1
Meter identification

sarga	verses	meters	subtype
1	95	अनुष्टुभ्, प्रहर्षिणी	None
2	75	उपजाति, मालिनी	ऋद्धि, भद्रा, उपेन्द्रवज्रा, इन्द्रवज्रा, जाया, माला, हंसी, कीर्ति, माया, सिद्धि, बाला, रा-मा, वाणी, शाला, बुद्धि, हैमी, आर्द्रा
3	70	वंशस्थ, हरिणी	None
4	88	अनुष्टुभ्, प्रहर्षिणी	None
5	76	उपजाति, वसन्तलतिका	मालिनी, पुष्पिताग्रा, हैमी, उपेन्द्रवज्रा, मा-या, इन्द्रवज्रा, जाया, बाला, कीर्ति, वाणी, भद्रा, शाला, आर्द्रा, ऋद्धि, बुद्धि, सिद्धि, रा-मा

3.2 The padapāṭha file

In the padapāṭha file, sandhi is analysed between words in the saṁhitā file, and each pada is set in a `w` element and provided with a unique `xml:id` attribute. Figure 2 shows an example of the markup in the padapāṭha file. Table 2 shows the word count in each of the first five cantos.

Figure 2
Word count in the first five sargas of Kālidāsa's Raghuvaṁśa

```
<lg n='1' xml:id='s1.v1'>
 <l n='1' xml:id='s1.v1.s1'>
  <w n='1' xml:id='s1.v1.w1'>vAgarTAviva</w>
  <w n='2' xml:id='s1.v1.w2'>saMpfktO</w>
  <w n='3' xml:id='s1.v1.w3'>vAgarTapratipattaye</w>
  <pc>.</pc>
  <w n='4' xml:id='s1.v1.w4'>jagataH</w>
  <w n='5' xml:id='s1.v1.w5'>pitarO</w>
  <w n='6' xml:id='s1.v1.w6'>vande</w>
  <w n='7' xml:id='s1.v1.w7'>pArvatIparameSvarO
   <m n='1' xml:id='s1.v1.w7.m1'>pArvatI</m>
   <m n='2' xml:id='s1.v1.w7.m2'>parameSvara</m>
  </w>
   <pc>..</pc>
 </l>
</lg>
```

Generally we do not need to analyse any compound into its constituents in the `padapāṭha` file. However, Mallinātha often dissolves compounds, and provides synonyms, derivations, references and other comments, not only on the compound as a whole but also on its constituents. In order to facilitate precise reference to the constituents commented upon in the commentary file, we split compounds into their constituents where the commentary analyzes the compound and provides a synonym, lexical reference or derivation for any of its constituents. Each compound

Table 2

Pada count

sarga	verses	padas
1	95	969
2	75	1039
3	70	1029
4	88	857
5	76	1060

constituent is set in an `m` element and provided with an `xml:id` attribute. Table 3 shows the count of the compound padas that are so analyzed into their constituents in the first three cantos, the percentage of padas analyzed, and number of segmented constituents.

Table 3

Tally of padas, analyzed compounds, and their constituents in cantos 1 and 2

sarga	verses	padas	analyzed padas	percentage of padas analyzed	segmented constitutents
1	95	969	88	9.08%	215
2	75	1039	99	9.53%	250
3	70	1029	71	7.00%	178

3.3 Coordinating analysis in the commentary file with the padapāṭha file

The commentary file consists of the digital character data of the *Sañjīvanī* in SLP1 encoding analyzed in accordance with the TEI guidelines in the same manner as Ajotikar and Scharf (2023) analyzed Bhaṭṭi's *Rāvaṇavadha*. Analytic elements `ab`, `seg`, `w`, etc. are coordinated with specific words, compound constituents, and morphemes in the padapāṭha file by referring to the `xml:id` of the `w` element containing the word, or the `m` element containing the compound constituent or morpheme in the padapāṭha file as the value of the `corresp` attribute of the analytic element in the commentary file. In the following subsections, we illustrate our method of tagging the constituents of compounds analyzed by Mallinātha, and how we supply information assumed as obvious by Mallinātha in order to make its relationship to the text commented upon explicit.

3.3.1 Compound constituent analysis

When Mallinātha supplies any information regarding a constituent of a compound, technically there is no word in the base text to which it corresponds. As mentioned above, in such cases, we split the compound pada in the padapāṭha file into its constituents and set each constituent in an `m` element with a unique `xml:id`. This `xml:id` occurs as a value of the `corresp` attribute in the analytic element in the commentary file. Below we provide four examples of the analysis of words into constituents. The first illustrates a simple compound analyzed into its constituents. The second analyzes a taddhita derivate into its constituent base and affix. The third and fourth present the analysis of compounds whose first and second constituent respectively undergo alteration in the compound.

1. In the very first verse, the word *pārvatīparameśvarau* (s1.v1.w7) is a dvandva compound. Mallinātha comments on each of the constituents of the compound as follows:

   पर्वतस्यापत्यं स्त्री पार्वती। 'तस्यापत्यम्' इत्यण्। 'टिड्ढाणञ्' इत्यादिना ङीप्। पार्वती च परमेश्वरश्च पार्वतीपरमेश्वरौ। परमशब्दः सर्वोत्तमत्वद्योतनार्थः।

The word *pārvatī* designates the female offspring of the mountain (*parvata*). The suffix *aṇ* is added (by *A.* 4.1.83 in the meaning of 'his offspring') by *A.* 4.1.92, after the word *parvata*. The feminine suffix *ṅīp* is added by *A.* 4.1.15 *ṭiḍḍhāṇañ* etc. (The dvandva compound) *pārvatīparameśvarau* is dissolved as *pārvatī* and *parameśvara*. The constituent *parama* (of the latter) denotes 'best of all'.

Here Mallinātha gives the derivation of the word *pārvatī* and supplies the sūtras for the same. Then he shows the paraphrase of the compound (*vigrahavākya*). In order to associate all the information about each constituent in the commentary file with its constituent in the padapāṭha file, we split the compound into two constituents in the padapAWa file as shown below:

```
<w n='7' xml:id='s1.v1.w7'>pArvatIparameSvarO
 <m n='1' xml:id='s1.v1.w7.m1'>pArvatI</m>
 <m n='2' xml:id='s1.v1.w7.m2'>parameSvara</m>
</w>
```

2. Mallinātha analyzes other derivates besides compounds into their constituents; these include *taddhita* derivates. For example, Mallinātha comments on the word *prasraviṇīm* (s2.v61.w17) as follows:

प्रस्रवः क्षीरस्रावो ऽस्ति यस्याः सा तां प्रस्रविणीं

The word *prasraviṇīm* means one who has a flowing forth, i.e. flow of milk.

Here Mallinātha does not merely supply the paraphrase but provides a synonym of the first constituent *prasrava*. To the word *prasrava* is added the possessive suffix *in*. Its first component is the word *prasrava*, and the second component is the affix *in*. In the padapāṭha file, we treat such cases similarly to the way we treat compounds, setting each component of the taddhita derivate in an m element. For the affix, we use the bare affix without any markers or supplementary sounds added for pronunciation. Thus here the second component is *in*, not *ini* as provided by Pāṇini *A.* 5.2.115 with *i* added for the sake of pronunciation.

```
<w n=`17' xml:id=`s2.v61.w17'>prasraviRIm
 <m n=`1' xml:id=`s2.v61.w17.m1'>prasrava</m>
 <m n=`2' xml:id=`s2.v61.w17.m2'>in</m>
</w>
```

3. In the process of dividing a compound into its constituents, there are a few instances of altered forms where a decision needs to be made regarding how to designate the base of the altered form. For example, for *anāsthā* (s2.v57.w15), which is a nañ-tatpuruṣa compound, Mallinātha provides the synonymous nañ-tatpuruṣa *anapekṣā*, and adds a lexical reference for the second constituent *āsthā* of the original compound. He states:

शरीरेष्वनास्था खल्वनपेक्षैव। 'आस्था त्वालम्बनास्थानयत्नापेक्षासु कथ्यते' इति विश्वः।

Indifference towards bodies is indeed detachment. Viśva provides four synonyms of the word *āsthā*: *ālambana*, *āsthāna*, *yatna*, and *apekṣā*.

In order to associate the lexical reference of the second constituent of *anāsthā*, we dissolve the compound into two constituents. However the first constituent *an* is not a free morpheme. It is an altered form of the negative particle *na*. Rather than using the altered form of the particle *an* before vowel-initial subsequent constituents, and *a* before consonant-initial subsequent constituents, we uniformly use the original unaltered particle *na*. In other words, we do not use the constituent of a compound as it occurs in it; rather we use the base word as lexicalized in the padapāṭha file, as shown below.

```
<w n=`15' xml:id=`s2.v57.w15'>anAsTA
 <m n=`1' xml:id=`s2.v57.w15.m1'>na</m>
 <m n=`2' xml:id=`s2.v57.w15.m2'>AsTA</m>
</w>
```

4. The previous example was of a compound in which the first constituent was altered. Here we provide an example of a compound in which it appears that the final constituent is altered. The word *kuṇḍodhnī* (s1.v84.w3) has two constituents: *kuṇḍa* and *udhas*. Mallinātha comments on the compound as follows:

कुण्डमिवोध आपीनं यस्याः सा कुण्डोध्री। ‘ऊधस्तु क्लीबमापीनम्’ इत्यमरः। ‘ऊधसो ऽनङ्’ इत्य-
नङादेशः। ‘बहुव्रीहेरूधसो ङीष्’ इति ङीष्॥

The one whose breast, i.e. udder, is like a pitcher. Amara states that the word *ūdhas* is neuter and its synonym is *āpīna*. The final sound of the word *ūdhas* is replaced by *anaṅ*, by *A.* 5.4.131 *ūdhaso 'naṅ*, in a *bahuvrīhi* compound of which the final constituent is *ūdhas*. The feminine suffix *ṅīṣ* is provided after the bahuvrīhi compound by *A.* 4.1.25 *bahuvrīher ūdhaso ṅīṣ*.

The provision of the samāsānta suffix *anaṅ* at the end of the compound whose final constituent is *ūdhas* results in the altered morpheme *ūdhan*, which after the addition of the feminine suffix *ṅīṣ* appears as *ūdhnī*. This morpheme is not a free morpheme, never occurs as an independent word, and is not lexicalized. Hence, we restore the second compound constituent to the base form *ūdhas* in the padapāṭha file and set it in an `m` element with an `xml:id` as shown below. In order to associate the passage of the lexical resource in the commentary file with the constituent in the padapāṭha file, this `xml:id` is referred to as the value of the `corresp` attribute of the `seg` element in which Mallinātha makes reference to the passage in Amara's lexicon that provides the synonyms for the constituent.

```
<w n=`3' xml:id=`s1.v84.w3'>kuRqoDnI
  <m n=`1' xml:id=`s1.v84.w3.m1'>kuRqa</m>
  <m n=`2' xml:id=`s1.v84.w3.m2'>UDas</m>
</w>
```

3.3.2 Inferred padas in the commentary file

One would expected that scholastic commentaries like Mallinātha's *Sañjīvanī* comment on almost every word of the base text. However Mallinātha does not always cite the exact pada of the base text. Even without explicitly repeating a compound or other word analyzed, he provides a compound-analysis or derivation, and just thereby refers to the original reading implicitly. In such cases, we add the whole pada analyzed to the commentary file in a `w` element provided with the attribute-value pair `type=`inferred''. This procedure is an adaptation in XML of the square brackets in which the word is inserted in the printed edition by the editor. Kale (1922) provides the inferred pada in square brackets before or at the end of the compound analysis. We invariably add the inferred pada at the beginning of the compound analysis. Sometimes Mallinātha omits a pada of the base text in his commentary. Usually it is a particle like *ca* or *eva* that is omitted. We deal with such an omission by including the particle at the appropriate place in the commentary file in a `w` element provided with the attribute-value pair `type=`supplied''. Just one such instance occurs in the first canto; none in the second or third. Table 4 shows the distribution of padas inferred in the first three cantos.

Table 4
Pada count

sarga	verse	padas	inferred padas	percentage
1	95	969	91	9.39%
2	75	1039	89	8.57%
3	70	1029	81	7.87%

4 Issues in tagging the commentary

4.1 Enhancement of attribute-value pairs

Although Ajotikar and Scharf (2023) standardized the procedure of tagging kāvya text with commentary, that sample study was based upon a small extent of text. The project of marking up the first five cantos of the *Raghuvaṁśa* with the *Sañjīvanī* extends the tagging of such texts with commentary significantly. While our procedures remain predominantly the same, the additional data required us to enhance the set of attribute-value pairs in order to record previously unencountered information in the commentary. The discussion hereafter focuses on reporting these enhancements with representative examples.

4.1.1 Synonymity

Ajotikar and Scharf (2023, p. 130) state

> Where Mallinātha supplies a word with a synonym, the synonym is put in a `seg` element as a sister to the `w` element containing the word. The `seg` element is supplied with the attribute-value pair `type='synonym'` and a `corresp` attribute with the value of the `xml:id` of the `w` element.

While observing his procedure in the current project, we noticed that Mallinātha treats compounds differently from simple words. When he provides a synonym of a simple word, he almost always supplies a reference to a thesaurus. For example, when he comments on the word *uḍupena* (s1.v2.w11) he supplies a reference to the *Amarakośa*.

> उडुपेन प्लवेन। ‘उडुपं तु प्लवः कोलः’ इत्यमरः।
> The word *uḍupena* 'by raft' means *plavena* 'by boat'. Amara records (two synonyms) *plava* and *kola* for *uḍupa*.

However, when Mallinātha provides a synonym of a compound of only occasional occurence, he makes no reference to a lexical resource because no lexical treatise includes the compound. In this case Mallinātha creates a synonym by replacing each constituent in the compound with a synonym. For example, he composes the synonym *īśvarakiṁkarasya* for the word *devānucara-sya* (s2.v52.w2). It is observable that he formulated the synonym by replacing the constituents *deva* and *anucara* with the synonyms *īśvara* and *kiṁkara* respectively. In both kinds of cases, whether he provides a lexical reference or not, we set the synonym in a `seg` element with the attribute-value pair `type='synonym'`.

 Different from the above cases of synonymity, we discovered three types of cases where Mallinātha glosses a word with a word that is not simply a synonym. In one case the gloss is a morphological alternate. In many cases, the gloss is a semantic explanation. These are of two types: an explanatory word that implicitly explicates grammatically known syntactico-semantic relations, and an explanatory word that merely explicates contextual semantics. We illustrate these types of glosses with an example of each below.

1. There is only one instance where Mallinātha provides an alternate morphological form in a gloss:

   > जरसा जरया विना (s1.v23.w9).

   The word *jarā* has the optional form *jaras* when a vowel-initial nominal termination follows. Here Kālidāsa uses the optional form *jarasā*. Mallinātha comments on it by supplying the other alternate form *jarayā*. This cannot be treated as a synonym since an optional morphological form of the same nominal base is not a synonym. We cover this instance by formulating a new attribute-value pair for the `seg` element `type='alternate'`.

2. In the many instances where Mallinātha provides a gloss that is a semantic explanation, we assign the value `def` (definition) to the attribute `type` in the `seg` element. Our first example illustrates the type of gloss by which he provides an explanatory word that implicitly

explicates grammatically known syntactico-semantic relations. For the word *upahāsyatām* (s1.v3.w4) 'the state of being to be ridiculed', Mallinātha provides the gloss *upahāsaviṣaya-tām* 'being an object of ridicule'. Here the explanatory word implicitly refers to grammatical rules that explicate the semantic structure of the glossed word. By stating *upahāsa-viṣaya*, Mallinātha implicitly refers to one of the senses in which the krtya affixes are provided after a root, namely, to denote the direct object of the action of the root, as stated by *A.* 3.4.70 *tayor eva krtyaktakhalārthāḥ.*

Our second example illustrates the type of gloss by which Mallinātha provides an explanatory word that merely explicates contextual semantics. He usually adds the phrase *ityarthaḥ* after such contextually explanatory glosses. The compound *ānākarathavartma-nām* (s1.v5.w6) 'the track of whose chariots reach heaven' is explained by providing the compound *indrasahacāriṇām* 'companions of Indra' concluded with the phrase *ityarthaḥ*. The statement that those whose chariots reach heaven are companions of Indra clearly does not provide a synonym because the gloss is contextually specific. Rather it states simply in ordinary terms what Kālidāsa poetically phrases in a term that could have different meanings in different contexts. The gloss does not denote the primary meaning of the glossed word.

4.1.2 Syntax

As is commonly done in scholastic commentaries, Mallinātha provides derivations of most of the words in the base text while commenting on a verse. Often these derivations are explicit grammatical comments relevant to construing the syntactic relation of the word to others in the sentence. In such cases, the `seg` element in which the statement is set is provided with the attribute-value pair `type=`syntax'`. Such notes may concern the syntax of the main verb of a clause, the participatory role a denoted object plays in an action (*kāraka*) denoted by a nominal, or any other relation. The following examples illustrate these points.

1. Regarding the verb *mucye* 'I am released' (s1.v72.w2), Mallinātha comments मुक्तो भवामि। कर्मणि लट्। 'I am freed. The *laṭ* affix is provided to denote the direct object (*karman*).' The first sentence provides a paraphrase of the finite verb form of the root *muc* 'free, liberate' with the past passive participle of the same root plus a finite verb of the root *bhū* 'be'. The second sentence makes an explicit grammatical statement concerning syntax, namely, that the *l*-affix *laṭ* denotes the direct object (*karman*). This explicit statement clarifies that the word *mucye* occurs in the present tense and is passive. The point is significant for the syntactic structure of the sentence.

2. Mallinātha glosses the word *locanābhyām* (s2.v19.w11) with the word *karaṇaiḥ*. Thereby he indicates that the word *locana* is the instrument of the action of drinking denoted by the root *pā* 'drink' in the finite verb form *papau* in the verse.

3. While explaining the word *ikṣvākūṇām* (s1.v72.w8), Mallinātha provides the reference to the grammatical rule *A.* 2.3.50 *ṣaṣṭhī śeṣe* which explains the use of its vibhakti stating, इद्वाकूणामिति शेषे षष्ठी। 'In the word *ikṣvākūṇām*, the sixth-triplet termination occurs in a remaining sense.'

4.1.3 Constituent

1. Generally Mallinātha reveals the constituents of a compound when he gives the paraphrase (*vigrahavākya*) but does not repeat those constituents outside the paraphrase. Occasionally, however, he mentions the constituents separately and comments on them in detail. For example, while commenting on the word *gūḍhākāreṅgitasya* (s1.v20.w3), Mallinātha describes the technical meaning of the words *ākāra* and *iṅgita*. His comments are as follows:

   शोकहर्षादिसूचको भ्रुकुटीमुखरागादिराकारः। इङ्गितं चेष्टितं हृदयगतविकारो वा। 'इङ्गितं हृद्रतो भावो बहिराकार आकृतिः' इति सज्जनः।

   The word *ākāra* 'shape' means facial expression by knitting the eyebrows etc.

which indicates (the emotions) sorrow, joy etc. The word *iṅgita* 'indication' means
a gesture or change of feeling. Sajjana says *iṅgita* means a feeling in the heart,
and *ākāra* means an external expression.

Here Mallinātha not only provides definitions of both constituents *ākāra* and *iṅgita* but also
supplies a lexical reference. In order to coordinate this information with the base text we
mention the constituents of the compound in separate `m` elements in the padapāṭha file. In
the commentary file, we introduce the new value `constituent` of the `type` attribute in the
`seg` element that has the `corresp` attribute with the value of the `xml:id` of the constituent
in the padapāṭha file. Figure 3 shows this segment marked up in the padapāṭha file.

Figure 3

*TEI encoding of Mallinātha's commentary
on the compound constituents* ākāra *and* iṅgita

```
<w n='2' type='inferred' xml:id='s1.v20.w3'>gUQAkAreNgitasya</w>
<seg type='def' corresp='s1.v20.w3.m2'>SokaharzAdisUcako BrukuwImuKarAgAdiH
 <seg type='constituent' corresp='s1.v20.w3.m3'>AkAraH.</seg></seg>
<seg type='constituent' corresp='s1.v20.w3.m3'>iNgitaM</seg>
<seg type='synonym' corresp='s1.v20.w3.m3'>cezwitam</seg>
<seg type='synonym' corresp='s1.v20.w3.m3'>hfdayagatavikAraH</seg> vA.
<seg type='lexicon'>
 <quote>iNgitaM hfdgato BAvo bahirAkAra AkftiH</quote>
  <note>
   <bibl corresp='xml/common/listbiblfile.xml#aufrecht.cc'>Aufrecht
<biblScope>1.687</biblScope></bibl>.
   <bibl corresp='xml/common/listbiblfile.xml#ncc'>NCC
<biblScope>37.160.a.1</biblScope></bibl>.
  </note>
  iti sajjanaH.</seg>
```

4.1.4 Paraphrase

There are many instances where Mallinātha elucidates a word in the base text with a paraphrase.
Such paraphrases are different from synonyms, which are single words, from vigrahavākyas,
which are paraphrases of compounds, and from definitions. In order to deal with this kind of
paraphrase we introduce a new value `paraphrase` of the `type` attribute of the `seg` element as
shown in the following examples:

1. Mallinātha explains the word *vanāya* (s2.v1.w6) (dative of *vana* 'forest') by paraphrasing
   it *vanaṁ gantum* 'in order to go to the forest'. He supplies the infinitive of the verb *gam*
   'go' and changes the dative to an accusative to show that the forest is the direct object (*ka-
   rman*) of the action of going. This paraphrase precisely reformulates the original expression
   in accordance with the ellipsis which Pāṇini deals with in *A.* 2.3.14 *kriyārthopapadasya ca
   karmaṇi sthāninaḥ.*

   ```
   <w n='10' xml:id='s2.v1.w6'>vanAya</w>
   <seg type='paraphrase' corresp='s2.v1.w6'>vanaM gantum</seg>
   ```

   Exactly similar is the paraphrase *yajñaṁ kartuṁ* 'in order to perform a yajña' for *yajñāya*
   'for a yajña' (s1.v26.w4).

2. In the compound *phalānumeyā* (s1.v20.w5) 'to be inferred by results', Mallinātha explains
   the constituent *anumeya* 'to be inferred' by paraphrasing it *anumātuṁ yogyā* 'fit to be
   inferred' which clarifies that the sense of the kṛtya affix *yat* occurs in the sense of being fit
   or suitable (*arha*) in accordance with *A.* 3.3.169 *arhe kṛtyatṛcaś ca.* We mark this in the
   commentary file as follows:

```
<w n=`10' xml:id=`s1.v20.w5'>PalAnumeyA</w>
…
<seg type=`paraphrase' corresp=`s1.v20.w5.m2'>anumAtuM yogyA</seg>
```

3. Similarly, Mallinātha explains the word *vivakṣuḥ* 'desirous to speak' (s2.v43.w9) by para-
phrasing it *vaktum icchuḥ* thereby showing that the word is formed from the desiderative
root *vivakṣa* derived from the primary root *vac*. The commentary file marks this paraphrase
as follows:

```
<w n=`9' xml:id=`s2.v43.w9'>vivakzuH</w>
<seg type=`paraphrase' corresp=`s2.v43.w9'>vaktum icCuH</seg>
```

4.1.5 Words supplied by the commentator

Occasionally Mallinātha supplies a word to fill out the syntax of the verse upon which he
comments, usually but not always by putting *iti śeṣaḥ* after the word. In such cases, we set the
word in a `w` element with the `type` attribute value `added`, and an `n` attribute given the value
101 or above. Several instances occur in the first three sargas (1.37, 1.68, 1.69, 1.83; 2.56; 3.6,
3.68).

4.2 Quotations

Ajotikar and Scharf (2023) included the procedure of tagging a quotation in a commentary:
add the reference in a `note` element which has the text identifier in a `bibl` element which in
turn has the location within the text in a `biblScope` element. However they did not identify
the quotations in their sample data. In this project, we tried to identify the source of every
quotation Mallinātha cites. We can categorise these quotations as follows:

1. quotations from lexical resources and metrical literature,
2. quotations from the grammatical resources which are mostly the sūtras quoted from the
   Aṣṭādhyāyī,
3. other quotations which are not from lexical and grammatical literature.

We refer to two editions of the *Raghuvaṁśa*, one by Nandargikar (1982) and one by Pandit
(1874). Nandargikar provides several appendices including a list of the works and authors quoted
by Mallinātha, and a list of unidentified quotations. Pandit includes among his appendices a list
of works and authors referred to by Mallinātha, and a list of unidentified quotations referred to
by Mallinātha. However, neither editor made any effort to verify the original source to which
Mallinātha attributes a quotation. We did. In our attempt to verify the original source of the
quotation, we found that there are many instances where the source to which the quotation is
attributed does not in fact contain the passage. Closer examination revealed that some of the
quotations occur in a different text from the one reported by Mallinātha. We traced the correct
source of some of these, and in addition some of the unidentified quotes. We provide the correct
source in our XML commentary file. The works which are not extant today but were available
to Mallinātha are provided with the reference to their entry in the *New Catalogus Catalogorum*
(*NCC*). Yet many quotations which we could not locate remain unidentified.

4.2.1 Quotations from lexical resources

Mallinātha refers heavily to lexical resources. He refers not only to the base text of such resources
but also occasionally to commentaries on them. As mentioned, we carefully attempted to verify
every quotation from a lexicon against the original source; doing so revealed numerous errors
in the printed editions. We categorize the quotations under three headings: *verified, corrected,*
and *unidentified.*

Verified: These quotations are found in the original texts cited. The following is a list of the
lexical resources (*kośas*) to which Mallinātha refers in the first three cantos:

1. *Anekārthasaṅgraha* of Hemacandra
2. *Anekārthasamucchaya* of Śāśvata
3. *Abhidhānaratnamālā* of Halāyudha
4. *Amarakośa* of Amara
5. *Kṣirasṭāmin*'s commentary on the Amarakośa
6. *Ekākṣararatnamālā* of Mādhava
7. *Nānārthārṇavasaṅkṣepa* of Keśava
8. *Vaijayantikosha* of Yādavaprakāśa
9. *Viśvaprakāśa* of Maheśvara
10. A lexicographical work of Sajjana no longer extant

For the last item, the lexical work composed by Sajjana referred to by Mallinātha (s1.v2.w11, s1.v20.w3), we provide the reference to the *NCC* (Dash 2015, 160a, first entry).

Corrected: When we discover the correct source of a quote which the printed edition of the *Sañjīvanī* incorrectly attributes to a different text, we provide the correct source. We do so by setting the incorrect source in the `sic` element, and the correct source in the `corr` element. In the first three cantos, there are eleven instances where we corrected the source. For example, while commenting on the word *dākṣiṇyarūḍhena*, for the constituent *dākṣi-ṇya*, Mallinātha provides the synonym *paracchandānuvartanam* 'following another's will'. Subsequently he states,

'दक्षिणः सरलोदारपरच्छन्दानुवर्तिषु' इति शाश्वतः।

According to Śāśvata, the word *dakṣiṇa* means 'straight-forward,' 'generous' and 'behavior as per another's will'.

Despite his claim that this quotation is in Śāśvata's *Anekārthasamucchaya*, it is not found in it. It is actually found in the *Viśvaprakāśakośa*. We correct it as follows:

```
<seg type=`lexicon' coresp=`s1.v31.w2.m1'>dakziRaH
saralodAraparacCandAnuvartizu
<quote>dakziRaH saralodAraparacCandAnuvartizu</quote>
…iti <sic>SASvataH</sic><corr>viSvaH</corr>
</seg>
```

The following is the list of all corrected instances in the first three cantos:

1. शाल (s1.v13.w3.m1): यादव corrected to शाश्वत
2. नेमीनाम् (s1.v17.13.m1): यादव corrected to हलायुध
3. वेलाः (s1.v30.w2.m1): विश्व corrected to शाश्वत
4. दाक्षिणयं (s1.v31.w2.m1): शाश्वत corrected to विश्व
5. सुरभिः (s2.v3.w7): विश्व corrected to शाश्वत
6. दावम् (s2.v8.w6): यादव corrected to शाश्वत
7. रागः (s2.v15.w8.m1): शाश्वत corrected to विश्व
8. रत्नषाम् (s2.v30.w7.m1): यादव corrected to विश्व
9. कङ्कु (s2.v31.w5.m3): विश्व corrected to एकाक्षररत्नमाला
10. अदह्यात (s2.v32.w7): यादव corrected to धनञ्जय (दशरूपक)
11. गुणाग्र्यवर्तिना (s3.v27.w10): विश्व corrected to शब्दरत्नसमन्वयः

Concerning 10 (s2.v32.w7), the reading in the edition, namely, 'अधिक्षेपाद्यसहनं तेजः प्राणा-त्ययेष्वपिऽ इति यादवः, suggests that the author of the cited passage is Yādava, the author of the *Vaijayantīkośa*. However, the passage occurs in the *Daśarūpaka* (Parab 1941, p. 42), not in any lexical resource.

Concerning 9 (s2.v31.w5.m3), Mallinātha mistakenly claims that the verse line कङ्कुः प-क्षिविशेषे स्याद्गन्नाकारे युधिष्ठिरे occurs in Viśva's *Viśvaprakāśakośa*. However, after careful examination, the quote is found in the *Ekākṣararatnamālā*. The first edition of the text by

Ramnikvijay (2019) identifies its title as *Ekākṣaraśabdamālā* and its author as Haritālarājā-mātyamādhava. These are not found as such in the *NCC*. Instead, the *NCC*, vol. 3, p.59b (the 17th entry in the column), edited by Raghavan (1967), reads

एकाक्षररत्नमाला by Mādhavācārya, son of Māyana, minister of Harihara.

Thus the *NCC* provides the title *Ekākṣararatnamālā* and identifies the author as Mādhavācārya, son of Māyana, minister of Harihara. What is given as the name of the author in the printed edition, Haritālarājāmātyamādhava, means 'Mādhava, a minister of Haritāla', which agrees to some extent with the information found in *NCC*. In this case, we provide references to both the printed edition and the *NCC* in the comentary file.

Unidentified: There are just two instances in which the quotation is unidentified:

1. At the end of verse 2.35, Mallinātha quotes the verse,
   पृथिवी सलिलं तेजो वायुराकाशमेव च।
   सूर्याचन्द्रमसौ सोमयाजी चेत्यष्टमूर्तयः ॥
   and claims that it is quoted from the *Vaijayantīkośa*. However the quote is not found in the *Vaijayantīkośa*.
2. भरणे पोषणे भर्म इति हैमः। (s3.v12.w6): When Mallinātha comments on the word *garbha-bharmaṇi*, he provides a lexical reference for the constituent *bharman* claiming that it is from Hemacandra's *Anekārthasaṅgraha*, but it is not found.

There is one interesting case where Mallinātha refers to a variant reading in the *Amarako-śa* which is noted by Liṅgayasūrin in his *Amarapadavivṛti* (Ramanathan 1978, p. 37). While commenting on the word *atrasta* (s1.v21.w3), he quotes the *Amarakośa* regarding the constituent *trasta* as त्रस्तो भीरुभीरुकभीलुकाः. In his citation, the quote contains the word *trasta* instead of *trasnu* which is found in the passage in most of the printed editions of the *Amarakośa* (अधीरे का-तरत्रस्नौ भीरुभीरुकभीलुकाः). Liṅgayasūrin in his *Amarapadavivṛti* registers a variant on this verse saying, त्रस्तो इति वा पाठः. Apparently Mallinātha knew this line in the *Amarakośa* as अधीरे कात-रत्रस्तौ भीरुभीरुकभीलुकाः. This is very significant from the point of view of tracing Mallinātha's sources.[1]

4.2.2 Quotations from treatises on meter

The verses of each canto are predominantly composed in a single metrical pattern; however, the pattern changes towards the end of the canto. Whenever the metrical pattern changes, Malli-nātha cites a characterization of the metrical pattern (*lakṣaṇa*) but never tells the source of the citation. In the first three cantos there are five occurrences where he identifies the meter and then cites its *lakṣaṇa* (s1.v95, s2.v1, s2.v75, s3.v1, s3.v70). The sources of these citations are not identified in the printed editions. We identify them. Four of them occur in Kedārabhaṭṭa's *Vṛttaratnākara*, and one occurs in Gaṅgādāsa's *Chandomañjarī*.

4.2.3 Quotations from grammatical texts

Mallinātha refers to the sūtras of the *Aṣṭādhyāyī* in most of the derivations and syntactic comments he provides. Occasionally he quotes from commentaries like the *Vyākaraṇamahābhā-ṣya* and *Kāśikāvṛtti*. We have verified all of these grammatical quotations except one. Mallinātha attributes the quotation, न केवलं श्रूयमाणैव क्रिया निमित्तं करणभावस्य। अपि तर्हि गम्यमानापि (s2.v34), to a work called the *Nyāsoddyota* which is not extant. For this quote we provide the reference to the entry for the text in the *NCC*.

One derivation is interesting because it reveals a discrepancy among grammatical texts, particularly regarding lists (*gaṇa*s). The derivation concerns the word *vārdhaka* (s1.v8.w5). Mallinātha derives it by adding the affix *vuñ* after the base *vṛddha* by *A.* 5.1.133 *dvandvamanojñādibhyaś*

---

[1]Mallinātha was from Andhra. Hence it is not surprising that he knew the variant noted by a southern commentator.

ca, stating, द्वन्द्वमनोज्ञादिभ्यश्च। इति वुञ्प्रत्ययः। Interestingly the word *vrddha* is not included in the *gaṇa* that begins with the word *manojña* as per the list in the *Kāśikāvrtti* on *A.* 5.1.133. If the reading in the *Kāśikāvrtti* is followed, then the word *vārdhaka* cannot be derived. However the *Gaṇaratnamahodadhi* does include the word *vrddha* in the gaṇa *manojñādi* (verse 409). We must suppose that either Mallinātha knew a reading in the *Kāśikāvrtti* which included the word *vrddha* in the gaṇa *manojñādi*, or he refered to the *Gaṇaratnamahodadhi*.

4.2.4 Micellaneous quotations

Apart from lexical and grammatical texts, Mallinātha quotes verses or passages from many other texts. As with the lexical quotations, we attempted to verify the source of these and categorize them according to whether they have been verified, corrected, or remain unidentified. We provide the reference in the critical edition of the text if available. The number of unidentified micellaneous quotations is greater than the unidentified lexical quotations. The list of these other texts quoted in the first three cantos is as follows:

1. *Agnipurāṇa*
2. *Aṣṭāṅgasaṅgraha*
3. *Āśvalāyanagrhyasūtra*
4. *Kīrātarjunīya*
5. *Gautamdharmasūtra*
6. *Cāṇakyanīti*
7. *Taittirīyasaṃhitā*
8. *Nītisāra*
9. *Parāśarasmrti*
10. *Brhajjātaka*
11. *Manusmrti*
12. *Mandāramaranda*
13. *Mahābhārata*
14. *Mānavagrhyasūtra*
15. *Meghadūta*
16. *Yājñavalkyasmrti*
17. *Rāmāyaṇa*
18. *Śaṅkhasmrti*
19. *Saṅgītamakaranda*
20. *Skandapurāṇa*
21. *Harivaṃśa*

4.2.5 Corrected Quotations

Below is the list of the quotations for which we provide the corrected source.

1. Commenting on the word *maunam* in s1.v22, Mallinātha states, यथाह कामन्दकः 'नान्योपतापि वचनं मौनं व्रतचरिष्णुता' इति। The cited verse line does not occur in the *Kāmandakanītisāra* but is found in the *Agnipurāṇa* (239.22).
2. At the end of his commentary on s1.v85, Mallinātha quotes the verse,

   आग्नेयं भस्मना स्नानमवगाह्यं तु वारुणम्।
   आपोहिष्ठेति च ब्राह्मं वायव्यं गोरजः स्मृतम्॥

   and claims that this verse occurs in the *Manusmrti* (उक्तं च मनुना). The verse does not occur in the *Manusmrti* but is found in the *Skandapurāṇa*.
3. Similarly, commenting on s2.v75 Mallinātha states,

   'यथेयं पृथिवी महुत्ताना गर्भमादधे।
   एवं त्वं गर्भमाधेहि दशमे मासि सूतवे॥'
   इत्याश्वलायनानां सीमन्तमन्त्रे स्त्रीव्यापारधारण आधानशब्दप्रयोगदर्शनादिति।

The verse does not occur in the *Āśvalāyanagrhyasūtra* but is actually found in the *Māna-vagrhyasūtra*.

4.2.6 Unidentified Quotations

Below is the list of the many unidentified quotations in the first three cantos.

1. आयुक्तकेभ्यश्चौरेभ्यः परेभ्यो राजवल्लभात्। पृथिवीपतिलोभाच्च नराणां पञ्चधा मतम्॥ (s1.v60)
2. ऋषं देवस्य यागेन ऋषीणां दानकर्मणा। संतत्या पितृलोकानां शोधयित्वा परिव्रजेत्॥ (s1.v71)
3. कामं पितरं प्रोषितवन्तं पुत्राः प्रत्याधावन्ति एवमेत (एवम् ह वैत-) मभ्रयः प्रत्याधावन्ति सश्रकलान्दारूनि-वाहरन् (s1.v49)
4. चतुर्थेऽनवलोभनम् इत्याश्वलायनः। (s3.v10.w6)
5. छिन्द्याद्वाहमपि दुष्टात्मनः (s1.v28)
6. त्रिंशद्धागात्मकं लग्नम् (s3.v13)
7. नासाकण्ठमुरस्तालुजिह्वादन्तांश्च संस्पृश्चन्। षड्भ्यः संजायते यस्मात्तस्मात्षड्ज इति स्मृतः (s1.v39)
8. निर्वाणोत्थानशयनानि त्रीणि गजकर्माणि (s1.v71)
9. पृथिवी सलिलं तेजो वायुराकाशमेव च। सूर्याचन्द्रमसौ सोमयाजी चेत्यष्टमूर्तयः॥ (s2.v35)
10. प्रतिपाद्यमहिम्ना च प्रबन्धो हि महत्तरः (s1.v2)
11. प्रोष्यागच्छतामाहिताग्नीनामभ्रयः प्रत्युद्यान्ति (s1.v49)
12. मांसलश्च (s3.v34.w3)
13. मृदं गां दैवतं विप्रं घृतं मधु चतुष्पथम्। प्रदक्षिणानि कुर्वीत विज्ञातांश्च वनस्पतीन्॥ (s1.v76.w6)
14. रविशास्तमयो योगो वियोगस्तूदयो भवेत् (s3.v13)
15. राजा त्वर्थान्समाहृत्य कुर्यादिन्द्रमहोत्सवम्। प्रीणितो मेघवाहस्तु महतीं वृष्टिमावहेत्॥ (s1.v26)
16. विषादश्चेतसो भङ्ग उपायाभावनाश्रयोः (s.v40.w1)
17. शक्तानां भूषणं क्षमा (s1.v22.w3)
18. शुभदो मो भुमिमयः (s1.v1)
19. स खलु पुत्रार्थिभिरुपास्यते (s1.v35.w2)
20. समिद्धेऽग्नावाहुतीर्जुहोति (s1.v53.w1)

Usually Mallinātha cites these passages with a vague reference to their source such as *iti vaca-nāt* or *iti smrteḥ*. Yet in a couple of instances, he attributes the quotation to a particular author. However these do not occur in the texts by those authors. On s1.v22, Mallinātha claims that the verse शक्तानां भूषणं क्षमा is by *Cāṇakya*. However we did not find it in the *Cāṇakyanīti*. Similarly the quote निर्वाणोत्थानशयनानि त्रीणि गजकर्माणि (s1.v71) is claimed to be by Pālakāpya, but we did not find it in his *Gajacikitsā*. Lastly मांसलश्च (s3.v34.w3) is claimed to be by a vrttikāra. We did not find it in the *Kāśikāvrtti* and do not know to which vrttikāra he refers. Since these have not been located in the presumed texts, we label them as well `unidentified`.

5 Conclusion

In the present project of making a TEI digital edition of Mallinātha's commentary *Sañjīvanī* on Kālidāsa's *Raghuvaṃśa*, we encountered material that called for expansion of the procedures for creating XML editions of Sanskrit texts containing commentaries in accordance with the Text Encoding Initiative (TEI) guidelines. To accommodate this material we created the following additional values of the `seg` element's `type` attribute: (1) `alternate`, (2) `def`, (3) `paraphrase`, (4) `syntax`, and (5) `constituent`. The first three expand the categorization of glosses. Where previously all glosses were categorized as synonyms, we now distinguish optional morphological forms, semantic explanations, and paraphrases from synonyms by the following values of the `seg` element's `type` attribute respectively: `alternate`, `def`, and `paraphrase`. Item (4) adds a refinement of the derivation category to distinguish an explicit grammatical comment relevant to construing syntactic relations from derivational material exclusively concerned with the morphology and semantics of the word. Item (5) allows one to relate comments regarding a compound constituent separated from the compound paraphrase to the compound constituent in the padapāṭha file.

In the present project, we also contributed substantially to the constitution and analysis of the text. We verified and corrected the sources of citations, and listed those that remain unidentified, added reference to the *NCC* for non-extant works, and added sources of citations where Mallinātha doesn't reveal them. The process of undertaking the higher-level encoding of a text reveals hidden textual problems with Sanskrit texts. The process of analyzing a commentary and precisely categorizing extents of it forces one to notice lacunae in prior work on the text which in turn gives one the opportunity to make significant philological contributions. The contemporary shift of the principal medium of knowledge transmission from the printed word to the digital medium is transforming philology into digital philology which is the future form of textual research.

6 Acknowledgements

References

Ajotikar, Tanuja P. and Peter M. Scharf (2023). "Development of a TEI standard for digital Sanskrit texts containing commentaries: A pilot study of bhaṭṭi's Rāvaṇavadha with Mallinātha's commentary on the first canto". In: *Proceedings of the Computational Sanskrit & Digital Humanities: Selected papers presented at the 18th World Sanskrit Conference.* Ed. by Amba Kulkarni and Oliver Hellwig. Canberra, Australia (Online mode): Association for Computational Linguistics, pp. 128–145. URL: `https://aclanthology.org/2023.wsc-csdh.9`.

Dash, Siniruddha, ed. (2015). *New Catalogus Catalogorum. An alphabetical register of Sanskrit and allied works and authors saṃyamanśreṇivicāra–samasyāhomaprayoga.* Madras: University of Madras.

Kale, Moreshvara Ramachandra (1922). *The Raghuvaṃśa of Kālidāsa. with the commentary the Sañjīvanī of Mallinātha Canto 1–10 edited with a literal translation into English, copious notes in Sanskrit and English and various readings.* Bombay: Gopal Narayan and co. Third revised edition.

Nandargikar, Gopal Raghunath (1982). *The Raghuvaṃśa of Kālidāsa.* Delhi: Motilal Banarasidas. Fifth edition.

Pandit, Shankar P. (1874). *The Raghuvaṃśa of Kālidāsa. with the commentary the Sañjīvanī of Mallinātha edited with notes, part III, cantos 14–19.* Bombay: Government central book depot.

Parab, Kashinath Pandurang, ed. (1941). *Śrīdhanañjayaviracitaṃ Daśarūpakam. Dhanikakrtayāvalokayā vyākhyayā bhāratīyanāṭyaśāstragatadaśanirūpaṇena ca sametam.* Mumbai: Ninrnaysagar Press.

Raghavan, V. (1967). *New Catalogus Catalogorum. An alphabetical register of Sanskrit and allied works and authors.* Ed. by Raja C. Kuhnan. Madras University Sanskrit Series 18. Madras: University of Madras.

Ramanathan, A. A. (1978). *Amarakośa. with the commentary the unpublished south indian commentaries Amaapadavivrti of Liṅgayasūrin Amaapadapārijāta of Mallinātha Amaapadavivaraṇa of Appayārya critically edited with introduction.* Madras: The Adyar library and research centre.

Ramnikvijay, Pannyas, ed. (2019). *Ekākṣaraśabdamālā.* Ahmedabad: Śrutajñāna saṃskārapīṭha.