Quotes, Paraphrases, and Allusions: Text Reuse in Sanskrit Commentaries and How to Encode It

Patrick McAllister

Electronic version
URL: https://journals.openedition.org/jtei/3324
DOI: 10.4000/jtei.3324
ISSN: 2162-5603

Publisher
TEI Consortium

Electronic reference
Patrick McAllister, "Quotes, Paraphrases, and Allusions: Text Reuse in Sanskrit Commentaries and How to Encode It", Journal of the Text Encoding Initiative [Online], Issue 13 | May 2020 - November 2022, Online since 10 April 2021, connection on 24 November 2022. URL: http://journals.openedition.org/jtei/3324 , DOI: https://doi.org/10.4000/jtei.3324

For this publication a Creative Commons Attribution 4.0 International license has been granted by the author(s) who retain full copyright.
Quotes, Paraphrases, and Allusions: Text Reuse in Sanskrit Commentaries and How to Encode It

Patrick McAllister

ABSTRACT

The Sanskrit literary genre of commentaries has several characteristics that are a challenge to the structural encoding of its texts. One particular tricky feature is the skillful and sophisticated reuse of text in such commentaries. This article examines practical examples of encoding these types of passages, drawing on the documents of SARIT\(^1\) (Search and Retrieval of Indic Texts) and the author’s own encoding projects.

INDEX

Keywords: scholarly Sanskrit, commentaries, quotations, linked texts
1. Introduction

This article explores possibilities for encoding the reuse of text in the works of the logico-epistemological school of Buddhism in South Asia (the texts of the pramāṇavādins). This school starts in about the fifth century CE in South Asia with Dignāga (ca 480 – ca 540 CE) and continues to produce texts up to around the thirteenth century. This corpus of Sanskrit literature consists of roughly 150 works, which were written by at least, but not many more than, forty authors. In terms of volume, again roughly speaking, the relevant section of the Tibetan canon, which includes translations of a significant number of these Sanskrit texts, is around 25 MB in size (encoded in ASCII). For comparison, a simple “text-only” measurement of the Folger Shakespeare edition results in approximately 8 MB of such raw text. These texts of the logico-epistemological school of Buddhism range in extent from a few hundred words to multivolume works. Not all texts survive in their original language, Sanskrit. Some are accessible only in translations, usually Tibetan ones, and others are known only through fragments surviving in other texts. Many of those texts that do survive in their original have not been critically edited.

Recent years have witnessed increased efforts to understand the various literary and social conventions that are characteristic of premodern South Asian literature, which this group of works belongs to. Freschi (2015), introducing a two-volume collection of specialist papers on text reuse in South Asian literature, provides a succinct overview of the expectations, difficulties, and hermeneutic considerations attached to the editing of these texts. Freschi (2015, 86–87) distinguishes three reasons for studying the reuse of text: first, it aids textual criticism, allowing the reconstitution of texts and increasing the certainty of editorial decisions; second, it facilitates a clearer understanding of the relations between works (and their authors); third, it provides an opportunity for reflecting on “historically conditioned notions of text, author, originality, copyright and so on” (Freschi 2015, 87). Most of these questions are outside the focus of this article, however. It aims one level below them: how to provide scholars and editors with the basis for making such judgments or decisions.

The term text reuse, furthermore, is broad. Even in the literature that deals with the texts under consideration, it subsumes such varied notions as citation, quote, quasi-quotations, report, passages that have been silently used, or secondary usage. Quotation is too narrow a category to subsume even the most basic cases of intratextual relations that are encountered in these texts.
4 To understand the phenomena that the term *reuse* is supposed to cover here, it is first necessary to delimit and characterize this group of texts. These texts fall mainly within two categories: some of them belong to the literary genre called *bhāṣya* (commentary). The others are independent treatises that are mostly dedicated to a particular topic, either refuting opposing views or establishing tenets of this Buddhist group as the author saw them. The works of both groups are composed in what is often referred to as a “nominal style” (Jacobi 1903), a style that avoids the complex verb forms that Sanskrit can boast of and favors formally simpler noun formations. This leads to a high degree of abstraction and compression.

5 More significantly for our purposes, however, these texts share a certain set of ways to make use of other texts. This method, or technique, consists of several acceptable conventions that can, at first glance, all be classified as quotes, either explicit or not, and that can show strongly diverging degrees of literal accuracy.

6 A formal system of analysis for these occurrences was developed by Ernst Steinkellner for printed critical editions. This system is the most differentiated that has been advanced for the classification of intertextual relations in the field of Buddhist studies. Its purpose is as follows:

   It is with a view to providing a fairly general system of categories for differentiating between
   the various materials available as witness for original Sanskrit texts in their degree of
   testimonial validity that I would like to present below a more detailed list of kinds of text.  
   (Steinkellner 1988, 112)

7 The system defines a set of sigla that are to be used for the characterization of reused text. The most important ones are these:

   $T$:  
   *textus*, the text being edited;

   $Ce$:  
   *citatum ex alio*, a quotation from another text in $T$;

   $Ce'$:  
   *Ce usus secundarii in T*: like $Ce$, but not explicitly acknowledged as a quote;
\textit{e:} \\

added to \textit{Ce} or \textit{Ce′:} signifying \textit{modus edendi}, that the reused text was modified by the person quoting it;

\textit{Ci:} \\

citatum \textit{in alio}, quotation of \textit{T} in another text, again with the possible addition of an apostrophe (‘) or an \textit{e}, yielding \textit{Ci′}, \textit{Cie}, and \textit{Ci′e};

\textit{R:} \\

citatum \textit{in alio modo referendi}, a span of text in \textit{T} that is referred to in a later work.

Each combination of sigla is usually followed by either a bibliographic reference to an edition of the related text, or, if the text has not been edited, the name of the related text.

The individual letters used in this system of characterization have separate and rather diverse functions. The initial capital, \textit{T}, \textit{C}, or \textit{R}, qualifies a span of text either, in the case of \textit{T}, as the aim of the editorial endeavor, or, in the other cases, as a passage that has been found to be reused.

The second character, \textit{i} or \textit{e}, indicates the direction of reuse, either in another text or from another text. In a recent explanation of the symbol \textit{e}, Steinkellner (2017, xii, n. 4) has made it explicit that it should not be understood to imply direct quotation: if the reuse of text \textit{A} in text \textit{B} is marked by an editor of text \textit{B} with the appropriate symbols (\textit{Ce, Ce′}, etc.), we should not conclude that the editor is asserting that the author of text \textit{B} had direct access to text \textit{A}.

This second character is, furthermore, defined from the perspective of the edited text. Any passage in text \textit{B} that is marked as a case of reuse of text \textit{A} (\textit{Ce} etc.) commits the editor to acknowledging a corresponding \textit{Ci} in the (perhaps hypothetical) edition of text \textit{A}. This \textit{Ci} in text \textit{A} is also fully dependent on the \textit{Ce} in text \textit{B}: obviously, nothing about a passage marked as \textit{Ci} indicates that it is reused in another text, and it is only by finding it reused in that other text that we can classify it in this way.

The decision whether this span of text is to be characterized as secondary usage, signalled by adding an apostrophe as the third character, is usually derived from the context in which the passage appears, the para-text surrounding the reused span of text. Any certainty in this regard is usually
due to the author explicitly saying that the current passage was said, thought, or held by someone. The addition of this apostrophe is a binary sign: either a passage is reused secondarily, or it is not. It allows no further differentiation.

13 The presence of the final letter, e (for modus edendi), is supposed to reflect the editor’s estimate of the difference between the reused text as it appears in a source and as it appears in the context of its reuse, insofar as one might reasonably assume that differences are due to an intentional change by the author of the text that incorporates the passage from another. Several editors (including this writer) have found it hard to give good reasons for this kind of decision in their editions, and have avoided the use of this category altogether. In any case, this siglum never was intended to characterize differences that could be due to simple mistakes in copying manuscripts, but to mark intentional changes on the part of the author reusing the text. This siglum is again a binary operator, and as such allows only for a basic distinction without further characterization.

14 Even though this system of characterizations has been criticized in recent years because of a basic impossibility of being certain about which copy of a text an author quoting another text might have had, it cannot be disputed that the distinctions in the conceptual space opened up by this analysis are significant and useful ones for any research concerned with these kinds of texts: C and R mark reused material as either literally accurate or only paraphrased or alluded to; e, together with a source description, makes it clear where the quote is from; the apostrophe (’) reflects a basic observation about the passage’s context in the text being edited (whether it occurs as an explicit quote or not). And the final, optional e represents the editor’s estimate of the quoted material’s accuracy as a representation of the source.

15 In the rest of this article, these categories will be applied to some sample passages using the conceptual framework of the TEI Guidelines. The purpose is to illustrate light-weight markup rules for some general cases of textual reuse in the works of this genre that allow for easy querying of the resulting structured information. Before doing that, however, it will be useful to briefly consider the current infrastructure that the version 4.1 of the Guidelines (TEI Consortium 2020) offers as to its general applicability to these phenomena.
1.1 Markup of quotations in the TEI P5 Guidelines version 4.1

The part of the TEI P5 Guidelines version 4.1 most relevant to our inquiry is section 3.3.3: Quotation (TEI Consortium 2020). It occurs within a longer section that “… deals with a variety of textual features, all of which have in common that they are frequently realized in conventional printing practice by the use of such features as underlining, italic fonts, or quotation marks, collectively referred to here as highlighting” (TEI Consortium 2020, 3.3: https://tei-c.org/Vault/P5/4.1.0/doc/tei-p5-doc/en/html/CO.html#COHQHighlighting and Quotation). Many of the texts under discussion here show features that are particular to South Asian manuscript culture. So, the categories of “conventional printing practice” (primarily European and American practice, that is) cannot be naively applied to the encoding of these texts: the devices employed in the manuscripts, and also the texts, for emphasizing, distancing oneself from, or quoting words or longer phrases are often quite different from those encountered in printed editions.

Before looking at some of these conventions, it will be useful to summarize the main points in section 3.3.3. of the TEI Guidelines. The elements presented there are as follows, grouped by their nearest model class (all quotes in the following list are from TEI Consortium 2020, 3.3.3):

- in the class of model.hiLike (and model.common)
  - <q>: the most generic marker for quotations;
- in the class of model.attributable
  - <said>: for quotations of what someone says or thinks;
    - with the boolean attributes @direct and/or @aloud
    - in the sub-class model.quoteLike
  - <quote>: for any passage that is “attributed by the narrator or author to some agency external to the text”;
    - allows the attribute @source, which points at a source;
  - <cit>, “cited quotation,” a container for <quote> in which it is associated with an explicit (in the text) reference to a source;
- in the class of model.emphLike
That the TEI Guidelines define elements and attributes largely through their inclusion in such classes means that these classes are relevant not only for the correct (as intended) employment of the elements in marking up documents, but also for the development of reliable software that queries the documents. For even though section 3.3.3 discusses the varying use cases for all the above-mentioned elements as candidates for analyzing material surrounded by quotation marks, querying for them efficiently presupposes the correct expectations of their various properties. And these are mostly a function of the classes that the elements belong to. For example, an XPath\(^{15}\) that collects all elements likely to be used for quotation marks would be
\[
//\text{q}|\text{said}|\text{quote}|\text{cit}|\text{mentioned}|\text{soCalled}
\]
(with duplication).\(^{16}\)

**Example 1.** XQuery for the local names of elements in `model.attributable` class (according to TEI Consortium 2020). The result is the list (`<said>`, `<quote>`, `<cit>`, `<floatingText>`).

```
declare default element namespace "http://www.tei-c.org/ns/1.0";

doc("https://www.tei-c.org/Vault/P5/4.1.0/xml/tei/odd/p5subset.xml")
//elementSpec[
  classes/memberOf[
    @key="model.attributable",
    //classSpec[
      classes/memberOf[@key="model.attributable"]
    ]
  ]/@ident/string()]
```

| Element | Nearest Model | Attributes linking to other occurrences |
|---------|--------------|----------------------------------------|
| `<q>`   | `model.hiLike` | `@who; @source; @corresp` |
It is noteworthy that the elements suggested for marking quotation in section 3.3.3 of the TEI Guidelines do not share one class: some belong to `model.attributable`, others to `model.emphLike`, and the most general marker—`<q>`—belongs to `model.hiLike`. This is a result of the presentation in section 3.3.3, which focuses on the use of quotation marks. Since modern printing does not, on the analysis proposed there, employ them to mark only related features, there need not be any expectation of a common superclass for the elements discussed here.

The quotation elements as described in the TEI Guidelines all share the `@source` and `@corresp` attributes. These two properties are expressive enough to provide the required intertextual links—that is, a link to a work, either in the abstract or as found in a certain edition, and to a passage in that edition. For this reason, they need not be discussed further here. What is not so easily solved, however, is how these various elements map onto the conceptual distinctions underlying Steinkellner’s model. Should one just use `<q>`, the most general form, and be done with it? Or can something be gained from differentiating among the elements for different cases? The next section will consider some examples to answer these questions.

### 2. Simple Quotations (and Their Context)

Example 2 shows a quotation in the fullest sense of the word.

Example 2. Simple quote with attribution (the target of `tei:quote/@source` is not shown).

```xml
<div>
  said
</div>
```
etena yad uktam udyotakareṇa

Through this also what Uddyotakara said,

If words do not denote anything, both the proposition and the reason [in your inference] are inconsistent. has been rejected.

Example 2 is from the ca. eighth-century “Tattvasaṅgrahapañjikā” (see Krishnamacharya 1926), the commentary by Kamalaśīla on Śāntarakṣita’s panorama, in verse, of contemporary philosophy, the “Tattvasaṅgraha” (see Krishnamacharya 1926).

There are three voices that must be distinguished in this passage:

1. The passage as a whole is spoken by Kamalaśīla, the commentator.
2. He is connecting a passage in the base text, the text he is commenting upon, to an objection by a non-Buddhist opponent. The base text is here not quoted but only pointed at, by saying Through this, that is, through what is said in the base text.
3. The opponent’s text is quoted verbatim, as a subphrase in a statement that rejects it. In this case, we can verify that it is a faithful quote, since the sentence is found in Uddyotakara’s work (“Nyāyabhāṣyavārttika” 312.21–22, in Thakur 1997).

The commentary here performs various functions:

1. The construction “Through this … has been rejected.” (“etena … pratyuktāṃ bhavati”): connection of the text commented upon and the provided quote.
2. *What was said* (*yad uktam*): introduction of the quote
3. Uddyotakara (*udyotakareṇa*): attribution of the quote (the spelling with a double *d* is preferable)
4. `<quote>`: repetition of what Uddyotakara said

Since it is generally useful to capture the context of a quote, it is embedded in a `<seg>` element, which carries little semantic baggage. One could also consider the employment of a `<cit>` element instead of it in this case. But in the absence of a “bibliographic reference to [the] source” (TEI Consortium 2020, 3.3.3) of the quotation, this would not be an appropriate use case for `<cit>`.

All of these functions can be tied to parts or segments of the sentences under consideration. Ideally, all these functions that the commentary performs should be easy to find by general queries run on the group of texts. A simple search for `<quote>` will catch this type of quote, and its context can be discovered by checking the parent (`//quote/parent::*`). Possible refinements that could be made by employing `@type` and/or `@ana` attributes, and which are not reflected in the structure of the markup, will not be discussed here.

In Steinkellner’s analysis, Uddyotakara’s words in this passage would be categorized as `Ce`, an explicit quotation that is exactly the same as in a source available to us today.

### 3. Quotes as References

Later in the same text, Kamalaśīla introduces verse 1061 of the text he is commenting on as shown in example 3.

**Example 3. Quotes as references, markup following typography.**

```xml
<div n="1061">
    <div xml:lang="sa">
        <div type="base-text">
            <lg xml:id="ts__1061">
                <l>agobhinṇam ca yad vastu tad aksair vyavasīyate /</l>
                <l>... </l>
            </lg>
        </div>
        <div type="commentary">
            <p>yac coktam <hi rend="bold">indriyair</hi> ityādi, tad asiddham iti</p>
        </div>
    </div>
</div>
```
The passage in example 3 shows a type of quote significantly different from the one we saw before. The commentator’s introduction of the passage from the base text, the text the commentary is written on, contains at least two elements (here still marked graphically by `<hi>` 18 that should be categorized as `<quote>` elements of some sort:

1. The first is `indriyair ityādi`. This refers to an earlier verse, Tattvāsaṅgraha 939 (see Krishnamacharya 1926), which itself is a quote, in the base text, of a passage by an opponent.

2. The second is `agobhinam ca ityādi`, which points to the beginning of the verse in the base text, Tattvāsaṅgraha 1061, in which the author of the base text answers the opponent’s claim that was reproduced in Tattvāsaṅgraha 939.
Both of these phrases have as their main function the indication of the passage that the commentator is going to speak about, and they indicate it by repeating a few words from the beginning of the passage. It is characteristic about this type of quotation that its content is usually irrelevant: it is a reference to a particular string of characters, or sequence of sounds, much like a lemma in a note of a critical edition indicates the place to which the note applies. Steinkellner (1988, 117) remarks that these cases should be categorized as “quotations from another text” (Ce), though in his own editions he usually marked these kinds of quotes by graphical means (bold face), probably so as not to make the apparatus overflow.

Example 4. Quotes as references, first attempt (translated in example 3; the element with the @xml:id of ‘ts__939’ is not shown).

```xml
<seg>yac coktam <quote type="lemma" source="#ts__939">indriyair</quote> ityādi, tad asiddham iti darśayann āha--<quote type="lemma" source="#ts__1061">agobhinnam ce</quote>tyādi.</seg>
```

A first solution, shown in example 4, is to use the <quote> tag for these repetitions, with @type set to "lemma" (in the sense used in textual criticism, not to be confused with the linguistic value as in the @lemma attribute). With this markup we can separate these <quote> elements from other types quite easily. The reference to the base text is encoded in a @source attribute. A possible drawback in this encoding is that the referring function of the passage in which the quotation appears does
not become evident. Ultimately, one would have to establish (in the <encodingDesc>) that the value "lemma" in the @type of a <quote> element means that this element has some features of a reference.21

The next suggested solution, in example 5, tries to make this feature explicit by using <ref> elements. However, this creates problems. Semantically, it is problematic because it is not the content of the quote that <ref> refers to: the commentary uses the phrase “that starting with xyz” to refer to the base text but only “xyz” is an actual quote. It also has the significant drawback that any query intended to catch all quote-like elements will now have to include certain <ref> elements. This would increase the complexity of queries significantly.

Example 5. Quotes as references, second attempt (translated in example 3; the element with the @xml:id of "ts__939" is not shown).

```xml
<div n="1061">
  <div type="base-text">
    <lg xml:id="ts__1061">
      <l>agobhinnaṃ ca yad vastu tad akṣair vyavasīyate /
      <l>...
    </lg>
  </div>
  <div type="commentary">
    <p>
      <seg>yac coktam <ref target="#ts__939" type="lemma">indriyair</ref>
      ityādi, tad asiddham iti darśayann āha—<ref target="#ts__1061" type="lemma">agobhinnan</ref> ce</seg>tyādi.</p>
  </div>
</div>
```

Example 6. Quotes as references, proposed solution (translated in example 3; the element with the @xml:id of "ts__939" is not shown).

```xml
<div n="1061">
  <div type="base-text">
    <lg xml:id="ts__1061">
      <l>agobhinnaṃ ca yad vastu tad akṣair vyavasīyate /
      <l>...
    </lg>
  </div>
</div>
```
32 It thus seems best to combine these solutions, and treat the quote together with its embedding phrase as a reference as shown in example 6. This expresses most of the features that we can see in these quotations: we can pick out these <quote> elements along with all the other types, and still easily differentiate them by examining whether they are embedded in a <ref> to the base text.

4. Quotes of Individual Words/Phrases for Elucidation

33 In a second passage that follows later in the commentary on the same verse, Kamalaśīla takes up the Sanskrit word ca (usually and or furthermore), and explains how it is to be understood in the verse: its function there is the inclusion of objects that are not explicitly mentioned in the verse, rather than the simple conjunction of this sentence with the previous one. This is not the same kind of referring quotation that we considered above in examples 4–6, because its main purpose is to comment on the significance or meaning of the term, not to indicate a particular point in the base text. A proposal for its markup with <mentioned> and <gloss> is shown in example 7.

Example 7. Quotes for explanation.

```xml
<div n="1061">
  <div xml:lang="sa">
    <div type="base-text">
      <lg xml:id="ts__1061">
        <l><seg xml:id="ts__1061_a">agobhinnaṃ</seg> <seg xml:id="ts__1061_b">ca</seg> yad vastu tad akṣair vyavasīyate /</l>
        <l>...</l>
      </lg>
    </div>
  </div>
</div>
```
In terms of query complexity, this has similar problems to the simple `<ref>` solution just discussed: any query for quotes in general will have to take the variation introduced by `<mentioned>` into account, and will therefore become more complicated. However, section 3.3.3 of the TEI Guidelines does suggest `<mentioned>` as the element to use for the phenomenon described here. The element `<mentioned>` therefore must be considered in well-constructed queries for reused text.

The primary function of repeating the word that is to be explained, *ca*, is not to refer to the text, but to say something about the meaning or content of that term. This explanation can be encoded as a `<gloss>` element. Both elements, `<mentioned>` and `<gloss>`, are loosely tied to the base text,
and not to each other, because there are variant forms of this phenomenon where either there is no clearly identifiable <gloss>, or the term which could be <mentioned> is not repeated in the text. As before, the context useful to understanding this occurrence is wrapped in an anonymous <seg> element.

Steinkellner marks text segments of this kind with T' (textus usus secundarii), though he notes that “[f]ormally this kind of text would correspond to the group of ‘Ce’-texts’... But in order to distinguish the group of quotation-texts in their hierarchical position more clearly I prefer the reference to the textus in the present definition” (1988, 116–17). This hierarchical position amounts to an expression of the value that an editor assigns to a text as a witness for some other text. In the present case, it declares the editor’s conviction that the direct commentary is a privileged witness that is second in reliability only to direct witnesses (usually manuscripts) for the text that is being commented upon. The markup solution proposed here (with <mentioned>) does not explicitly express this facet, since it captures only the typed structural correspondence between text passages. However, this facet can be defined by the editor in a systematic fashion through the annotation of the target of the @source attribute: the determination of the “hierarchical position” is something that the editor will usually make explicit in their organization and interpretation of the sources. The encoding proposed here thus has the benefit of keeping two unrelated things separate: the similarity (or, in some cases, identity) of sequences of characters in different texts, and the editor’s estimate of their value as witnesses for certain editorial decisions. For example, an explicit quote from the base text in a commentary (corresponding to Ce in Steinkellner’s schema) would receive the same markup as an explicit quote in any other text (<quote> with @source). If the editor decided that the commentary had greater reliability than the other text, this could be made explicit in any of the many ways that the TEI Guidelines provide for: for example, since a document encoded on the basis of witnesses will probably have a <listwit> element, one could at least add an @ana attribute to the <wit> elements contained therein, and derive the desired hierarchical position from a query of that attribute.
5. Silent Quotes and Allusions

The last case of text reuse to be discussed is that of “silent reuse.” In this variant, phrases, sentences, or even longer passages are reused without explicitly marking them as being due to another agency. This phenomenon is characterized as usus secundarii by Steinkellner. These passages can, in this process, also change to varying degrees, and it is a difficult philological problem to judge the differences between the “reused” and “original” texts. The markup outlined for passages in this category also applies to cases where a passage is only paraphrased or alluded to.

In practice, all three possible modes of reuse are encountered: a text might reproduce a passage from another by embellishing or otherwise extending it, it might summarize or condense the original passage, or it might reproduce it very faithfully. Often, one will find all three modes mixed together in the same text.

In many cases it is also not possible to determine through how many intermediary steps a passage has been received, especially when the work it is supposedly taken from has not survived as a whole.

Example 8. Encoding parallels, suggested solution with stand-off markup (<anchor> only). The three <div> elements could be in different documents. For the text content, see examples 11, 13, and 15. The internal markup of the text passages linked in this way is not relevant to this solution.

```xml
<div>
  <!-- source text -->
  <div type="text1">
    <p>..<anchor xml:id="j-1-s"/>tatraikas tāvad bāhyo ... dṛṣṭāv iva iti<anchor xml:id="j-1-e"/></p>
    <p>... atrāpy uktam ... <anchor xml:id="j-2-e"/>...</p>
  </div>
  <!-- First reuse: Ratnakīrti’s Apohasiddhi -->
  <div type="text2">
    <p>..<anchor xml:id="r-1-s"/>tatra bāhyo ’rtho ... dṛṣṭāv iva iti<anchor xml:id="r-1-e"/></p>
    <p>...atrāpy uktam ... <anchor xml:id="r-2-e"/>...</p>
  </div>
  <!-- Second reuse: Mokṣākaragupta’s Tarkabhāṣā -->
  <div type="text3">
    <p>...<anchor xml:id="m-1-s"/>bāhyārtho ... dṛṣṭāv iva iti<anchor xml:id="m-1-e"/></p>
  </div>
</div>
```
39 The source text referred to in example 8 is from a work by Jñānaśrīmitra, a Buddhist monk and scholar active at the beginning of the eleventh century CE in Bengal. The section of text is from his “Apohaprákāraṇa” (see example 11 for the full text). This section was reused twice, with various degrees of abridgements.

40 Its first reuse (encoded in example 13) is taken from the “Apohasiddhi” (see McAllister 2020) of Ratnakīrti, who lived in the first half of the eleventh century: he calls his texts “abridgements” or “summaries” (saṃkṣepa) of the much more extensive works of his teacher, Jñānaśrīmitra. The same passage is again reused a few centuries later by Mokṣākaragupta in his “Tarkabhāṣā” (see Iyengar 1952; encoded in example 15). This third manifestation of the passage is heavily abridged. This case illustrates how text passages were reused by several authors. It also contains a quote-within-a-quote, an explicit quotation of a partial verse from a foundational text for the logico-epistemological school of Buddhism, the “Pramāṇaviniścaya” (see Steinkellner 2007) (in the examples below, a modern partial edition of it is referred to by “#pvin1”; the @target “PVin.xml#v15” refers to a TEI XML edition containing the quote). One might hesitate about how to categorize this quote from the “Pramāṇaviniścaya”: it is, on the surface, an explicit quote, but it is unlikely that any of the two authors reusing the passage in which it occurs should be said to have quoted this contained item itself. It is more likely that the verse was transmitted by them in the form found in the passage surrounding it, which was being reused as a whole. This question has to be answered when one wishes to judge the value of the preserved version of the verse for an edition of the “Pramāṇaviniścaya.” The proposed markup should enable a researcher to easily collect information that allows them to answer that question.
Two solutions will be proposed here. The first employs only lightweight stand-off markup and might be suitable for cases where the markup of the source documents cannot easily be changed. The second solution has the benefit of simpler querying, but at the cost of deeper interventions in the source documents.

The first solution is to use encoding such as that shown in example 8. This proposal marks up parallel text passages in a way that is minimally invasive on the source documents. It adheres closely to the \texttt{<linkGrp>} solution proposed in section 16.5.2 of the TEI Guidelines, Alignment of Parallel Texts. This solution has at least two advantages: often, the interpretation of textual parallels is a rather subjective matter. The separation of this interpretative layer from the main text requires only \texttt{<anchor>} elements in the document, so that the impact on existing applications should be minimal. The \texttt{<LinkGrp>} elements can be easily shared, stored in separate files, and one can have several of them, so that they can accommodate conflicting interpretations by different researchers. A second benefit is that they also work when the quoted material does not align nicely with block-level elements, as is the case in example 8, where the reused material crosses over a paragraph border in the first case of its reuse.

A drawback of this solution is that the search for quotes will become more complicated. But I feel that in this situation this disadvantage is justified to a certain extent: all members of \texttt{model.attributable} are defined, after all, as having been explicitly marked as quotes in the source text. These “silent quotes” are not marked in such a way.

Example 9. Encoding parallels, solution with linked elements (the documents presupposed are shown in examples 11, 13, and 15).

\begin{verbatim}
<doc>
  <linkGrp type="silentreuse">
    <link target="#j1 #r1 #m1"/>
    <link target="#j2 #r2"/>
  </linkGrp>
</doc>
\end{verbatim}

The second proposed solution is as shown in example 9. This solution is superficially similar to the one shown in example 8, but it links whole elements to each other, not the start and end points of spans of text. It thus requires the identification of spans of texts (at least as \texttt{<seg>}) that
are to be linked, even in the case of the original text that is being reused. It might be preferable, however, since it facilitates easier querying of the document. It employs the mechanisms provided in `<att.global.linking>`, that is, the `@next` and `@prev` attributes as described in section 16.7 of the TEI Guidelines, *Aggregation.* This allows elements to be constructed around the hierarchical constraints of the XML tree. In this variant, the employment of `<seg>` (for a quote’s context) and `<quote>` (for the quoted material) would cohere closely to the suggestions offered above for the cases of explicit quotations. But to the extent that the employment of `model.attributable` elements is questionable, this solution might be inadequate. The main practical benefit is certainly that a simple “search for all quotes,” for example, `//quote`, would return these results as well as any quotes explicitly marked as such in the sources. In *examples* 13 and 15, I have differentiated the `<quote>` elements by setting `@type` to “silent”.

Both solutions could be simplified by using CTS URNs as suggested and discussed by Berti, Almas, et al. (2015) and Berti, Blackwell, et al. (2016). Such references would replace the simple internal references (like “#j1” in the examples). For the first solution, we would then not even have to introduce `<anchor>` elements, since we could link directly into the text. For the second solution, it would at least save us the necessary markup in the document containing the reused text.

6. Conclusion

All cases of text reuse analyzed so far can be quite clearly expressed within the framework of the TEI Guidelines. So in this sense, no particular changes are necessary to accommodate this facet of Sanskrit literature. The markup proposed here is expressive enough to allow for all of the distinctions that were developed expressly for the study and edition of these texts (Steinkellner 1988). However, it also allows for a cleaner separation of the raw textual evidence and the editor’s judgement, by encoding only the textual correspondences. The TEI Guidelines provide a rich set of mechanisms for attaching interpretations to the basic set of structural elements discussed here. A simple way would be the addition of `@ana` and/or `@type` attributes to the appropriate `<seg>` or `<quote>` elements. An alignment of a source text and a derived text, even through multiple derivations, is possible and could be opened for further automatic processing after adequate normalization of the `<linkGrp>` elements.

Example 10. List of passages connected through `<link>` elements (for documents as in example 8).
To summarize, the following (simplified) XPath expressions might be used to select and find the different categories of reuse discussed in the introduction:

- *Ce citatum ex alio*, the quotation from another text in the text being encoded (T):
  - A simple `//quote` is sufficient (see also example 1)

- *Ce' citatum ex alio usus secundarii*, a silent quote in T:
  - On the first solution (see example 8), an XPath is not as straightforward as in the cases of explicit quotes. Given the `<link>` elements used in example 8, one could list the corresponding passages as follows:
    ```
    <div xml:id="cases-of-silent-reuse">
      <p>...<anchor xml:id="j-1-s"/>tatraikas āvad bāhyo ... drṣṭāv iva iti<anchor xml:id="j-1-e"/>
      </p>
      <p>... atrāpy uktam ...
        <anchor xml:id="j-2-e"/>...
      </p>
      <p>...<anchor xml:id="r-1-s"/>tatra bāhyo 'rtho ... drṣṭāv iva iti<anchor xml:id="r-1-e"/>
      </p>
      <p>...atrāpy uktam ...
        <anchor xml:id="r-2-e"/>...
      </p>
      <p>...<anchor xml:id="m-1-s"/>bāhyārtho ... drṣṭāv iva iti<anchor xml:id="m-1-e"/>
      </p>
    </div>
    ```

  - The result is shown in example 10. This XQuery expression is not a general solution because the analysis of the links is very naive (matching on a partial string) and there is no loop over the elements of the `<linkGrp>` or grouping of results.

  - On the second solution proposed above (see example 9), the results of `//quote` would include these cases; one could filter on attribute values (not discussed here).
• *e* added to *Ce* or *Ce’*, signifying that the editor believes that the text was changed by the author of *T* (*modus edendi*)

- This is judged from the comparison of two passages; the editor’s decision can be recorded in an *@ana* attribute to all of the proposed elements, but should not be expressed in the structure of the markup.

• *Ci()* *citatum in alio (usus secundarii)*, (silent) quotation of *T* in another text

- This is a simple reversion of previous relations (i.e., a //quote or similar search on the target text).

• *R* *citatum in alio modo referendi*, a text of *T* referred to or paraphrased

- This case belongs to the ones discussed in section 5. An editor might express peculiarities of the reference or paraphrase with an *@ana* attribute that can be added to any of the given relations. In most cases, the “quotation” expressed through a <link> element will be the obvious place for such a specifier. In the solution that encodes silent quotes as <quote>, an example was given that would match //quote[@type="silent"].

**APPENDIXES**

Example 11. Source text from Jiñānaśrīmitra’s *“Apohaprakaraṇa”* (p. 208, in Thakur 1987; for a translation, see example 12). Markup as presupposed in example 9.

```xml
<div xml:id="jna">
  <p>... <seg xml:id="j1">tatraikas tāvad bāhyo ’dhyavasāyād eva vācyo vyavasthāpyate, na svalakṣaṇaparisphūrttyā katham iti cet? anubhavād eva. yathā hi pratyakṣacetasī desākālāvasthāniyatāni parisphutarūpāṇi svalakṣaṇāni pratibhānītānuhūyante, tathā na vikalpakāle. vikalpakāle hi vijātivyāvṛttameva parasparākārākīrṇam ivāsphuṭam iva pratyakṣāparicitaṁ kiñcid rūpam ābhāsayet anubhavaviṣayaḥ. <seg>yac <bibl corresp="#pvin1">chāstram</bibl> <quote source="PVin.xml#v15"> <lg>śabdenāvyāpṛtākṣasya buddhāv apratibhāsanāt</lg> </quote></seg></p>
</div>
```
Example 12. Translation by McCrea and Patil (2010, 60–61) of the text in example 11, with relevant markup added and some formatting removed.

Example 13. First instance of silently reused text in Ratnakīrti’s “Apohasiddhi” (pp. 53–54, in McAllister 2020; see the translation in example 14). Markup as presupposed in example 9.
And by the word “positive element” an external object that is distinguished from that of another nature is meant according to determination, and according to appearance a form of awareness [is meant]. Amongst these, the external object is defined as that to be expressed by a word only because of determination, not because of a particular’s appearance, since there is no appearance of a manifest particular that is limited to a [certain] place, time and condition as [there is in the case of] perception. [This is] what the treatise [says, too]:

Because an object does not appear due to a word in the awareness of [someone having] an inactive sense organ in the same way as [it does] in perception, [ ... ]

[Opponent:] There is a difference in the appearance of
[what is in fact] a single object, because the ways [of apprehending
an object, according] to the nature of a sense faculty [or] a word,
are different. [Proponent:] Also with regard to this it was said:
...<quote>...</quote>
</p>
</div>

Example 15. Second instance of silently reused text in Mokṣākaragupta’s “Tarkabhāṣā” (pp. 53–54, in
Iyengar 1952; see the translation in example 16). Markup as presupposed in example 9.
<div xml:id="moksa">
<p>... <seg>tasmāt sthitam etat, <quote xml:id="m1" type="silent">bāhyārtho
‘dhyavasāyād eva śabdavācye vyavasthāpyate; na tu
svalakṣaṇaparisphūrtyā; pratyakṣavad
desakālāravasthānaniyatapravyaktasvalakṣaṇāspurānt. <seg>yad āha
</seg>
<p>
</p>
</div>

Example 16. Translation of the text in example 15, with corresponding markup.
<div>
<p>... <seg>Therefore, this is established: <quote xml:id="m1" type="silent">An
external object is defined to be what a word denotes only on account
of determination; but not on account of an appearance with its
individual characteristic, because this object does not appear with
its individual characteristic that is fixed as to its location, time,
and state. <seg>Which <persName corresp="#dharmakīrti">the lord of
logic</persName> said— <quote source="PVin.xml#v15">
</quote>
</p>
</div>
BIBLIOGRAPHY

Berti, Monica, Bridget Almas, David Dubin, Greta Franzini, Simona Stoyanova, and Gregory R. Crane. 2015. “The Linked Fragment: TEI and the Encoding of Text Reuses of Lost Authors.” Journal of the Text Encoding Initiative 8. http://journals.openedition.org/jtei/1218; doi:10.4000/jtei.1218.

Berti, Monica, Christopher Blackwell, Mary Daniels, Samantha Strickland, and Kimbell Vincent-Dobbins. 2016. “Documenting Homeric Text-Reuse in the Deipnosophistae of Athenaeus of Naucratis.” Bulletin of the Institute of Classical Studies 59 (2): 121–39. doi:10.1111/j.2041-5370.2016.12042.x.

Bhatt, N. R., and T. Ramanujam, eds. 1958. Sphoṭa Siddhi. La démonstration du Sphoṭa par Maṇḍana Miśra. Translated by Madeleine Biardeau. Publications de l’Institut français d’indologie 13. Pondichéry: Institut français d’indologie.

Blackwell, Christopher, and Neel Smith. 2013. Homer Multitext Project: Documentation. Accessed May 23, 2019. http://www.homermultitext.org/hmt-doc/index.html.

krung go’i bod rig pa zhib ‘jug lte gnas kyi bka’ bstan dpe sdur khang ed. 1994–2009. bstan ‘gyur dpe bsdur ma. Beijing: krung-go’i bod kyi shes rig dpe skrun khang / China Tibetology Publishing House.

Freschi, Elisa. 2015. “The Reuse of Texts in Indian Philosophy: Introduction.” Journal of Indian Philosophy 43(2–3): 85–108. doi:10.1007/s10781-014-9232-9.

Freschi, Elisa, and Philipp A. Maas, eds. 2017a. Adaptive Reuse: Aspects of Creativity in South Asian Cultural History. Abhandlungen für die Kunde des Morgenlandes 101. Wiesbaden: Harrassowitz Verlag.

Freschi, Elisa, and Philipp A. Maas. 2017b. “Introduction: Conceptual Reflections on Adaptive Reuse.” In Freschi and Maas 2017a, 11–25.

Gnoli, Raniero. 1960. “Introduction.” In The Pramāṇavārttikam of Dharmakīrti. The First Chapter with the Autocommentary; Text and Critical Notes, xv–xl. Serie Orientale Roma 23. Rome: Istituto Italiano per il Medio ed Estremo Oriente.

Hellwig, Oliver. 2015. “Using Recurrent Neural Networks for Joint Compound Splitting and Sandhi Resolution in Sanskrit.” Human Language Technologies as a Challenge for Computer Science and Linguistics. 7th Language & Technology Conference: Proceedings, edited by Zygmunt Vetulani and Joseph Mariani, 289–93. Poznań, Poland: Fundacja Uniwersytetu im. Adama Mickiewicza w Poznaniu. http://ltc.amu.edu.pl/a2015/book/papers/LRL-1.pdf.
———. 2016. “Detecting Sentence Boundaries in Sanskrit Texts.” Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers, edited by Yuji Matsumoto and Rashmi Prasad, 288–97. Osaka, Japan: The COLING 2016 Organizing Committee. https://www.aclweb.org/anthology/C16-1028/.

Iyengar, H. R. Rangaswami. 1952. Tarkabhāṣa and Vādasthāna of Mokṣākaragupta and Jitāripāda. 2nd ed. Mysore: Hindusthan Press.

Jacobi, Hermann. 1903. “Über den nominalen Stil des wissenschaftlichen Sanskrits.” Indogermanische Forschungen 14: 236–51.

Krishnamacharya, Embar, ed. 1926. Tattvasaṅgraha of Śāntarakṣita: With the Commentary of Kamalaśīla. 2 vols. Gaekwad’s Oriental Series 30–31. Baroda: Central Library.

McAllister, Patrick. 2020. Ratnakīrti’s Proof of Exclusion. Beiträge zur Kultur- und Geistesgeschichte Asiens 98. Vienna: Austrian Academy of Sciences.

McCrea, Lawrence J., and Parimal G. Patil. 2010. Buddhist Philosophy of Language in India: Jñānaśrīmitra on Exclusion. New York: Columbia University Press.

Pollock, Sheldon. 2007. “Literary Culture and Manuscript Culture in Precolonial India.” In Literary Cultures and the Material Book, edited by Simon Elliot, Andrew Nash, and Ian Willison, 77–94. London: The British Library.

Preisendanz, Karin. 2008. “Text, Commentary, Annotation. Some Reflections on the Philosophical Genre.” Journal of Indian Philosophy 36 (5–6): 599–618. doi:10.1007/s10781-008-9036-x.

Sāṅkṛtyāyana, Rāhula. 1953. Pramāṇavārtikabhāshya or Vārtikālaṅkāraḥ of Prajñākaragupta (Being a Commentary on Dharmakīrti’s Pramāṇavārtikam). Tibetan Sanskrit Works Series 1. Patna: Kashi Prasad Jayaswal Research Institute.

Steinkellner, Ernst. 1988. "Methodological Remarks on the Constitution of Sanskrit Texts from the Buddhist pramāṇa-Tradition." Wiener Zeitschrift für die Kunde Südasiens und Archiv für indische Philosophie 32: 103–29.

Steinkellner, Ernst, ed. 2007. Dharmakīrti’s Pramāṇaviniścaya: Chapters 1 and 2. Sanskrit Texts from the Tibetan Autonomous Region 2. Beijing: China Tibetology Research Center; Vienna: Austrian Academy of Sciences.

Steinkellner, Ernst. 2017. Early Indian Epistemology and Logic: Fragments from Jinendrabuddhi’s Pramāṇasamuccayaṭīkā 1 and 2. Studia Philologica Buddhica Monograph Series 35. Tokyo: International Institute for Buddhist Studies of the International College for Postgraduate Buddhist Studies.

Steinkellner, Ernst, Helmut Krasser, and Horst Lasic, eds. 2005. Jinendrabuddhi’s Viśālāmalavatī Pramāṇasamuccayaṭīkā. Chapter 1. 2 vols. Sanskrit Texts from the Tibetan Autonomous Region 1. Vienna: Austrian Academy of Sciences, in cooperation with China Tibetology Publishing House.

TEI Consortium. 2020. TEI P5: Guidelines for Electronic Text Encoding and Interchange. Version 4.1.0. Last updated August 19. N.p.: TEI Consortium. https://tei-c.org/Vault/P5/4.1.0/doc/tei-p5-doc/en/html/.
Thakur, Anantalal, ed. 1987. Jñānaśrīmitranibandhāvali: Buddhist Philosophical Works of Jñānaśrīmitra. 2nd ed. Tibetan Sanskrit Works Series 5. Patna: Kashi Prasad Jayaswal Research Institute.

Thakur, Anantalal. 1997. Nyāyabhāṣyavārttika of Bhāradvāja Uddyotakara. Nyāyacaturgranthikā 2. New Delhi: Indian Council of Philosophical Research.

Tubb, Gary A., and Emery R. Boose. 2007. Scholastic Sanskrit: A Handbook for Students. Treasury of the Indic Sciences. New York: The American Institute of Buddhist Studies at Columbia University and Columbia University Press.

University of Chicago Press. 2010. The Chicago Manual of Style. 16th ed. Chicago: University of Chicago Press.

NOTES

1 https://sarit.indology.info/

2 The Shakespeare estimate is based on measuring the result of //text() on the XML files in http://www.folgerdigitaltexts.org/download/xml/FolgerDigitalTexts_XML_Complete.zip, downloaded 2018-09-11. So this includes any notes or additional material. The bstan ’gyur section of the Tibetan canon as downloaded from https://asianclassics.org/data/TENGYUR%20(Roman %20%20%20Letters).zip(md5sum: 4c6cb125c9eb8c6cfff0a3e3da232010, downloaded 2019-05-07) has a zipped size of 76 MB, and of 244 MB unzipped. It contains hardly any additional content like editorial comments or markup. The subdirectory TENGYUR (Roman Letters)/11 TSAD_MA, which contains many of the texts that belong to the group being discussed here in their Tibetan translation, is 25 MB in size. Other common measures of the size of collections do not easily transfer to the texts under consideration here: in the absence of a reliable tokenizer for Sanskrit combined with the characteristic lack of space between words, the number of distinct tokens in our collection remains a guess (see Hellwig [2015, 2016] for the state of the art and limitations to Sanskrit tokenizing and splitting). In a recent printed edition, the Tibetan section of pramāṇa texts (tshad ma) including minimal notes runs to approximately 18,000 pages in twelve volumes (see krung go’i bod rig pa zhib ’jug lte gnas kyi bka’ bstan dpe sdur khang 1994–2009, vols. 97–108).

3 Freschi and Maas (2017b, 13–21) develop the notion of “adaptive reuse,” where a significant and inseparable element of text reuse is the intention that is attached to it.

4 See Steinkellner (2017, xii–xiii) for the various phenomena considered, as well as Preisendanz (2008, 611) for the “quasi-quotations.”
An accessible overview to the bhāṣya style is given by Tubb and Boose (2007). Apart from the observations on the general purpose and techniques of commentaries in Sanskrit traditions, they discuss the most common ways in which a commentary can quote or paraphrase the base text for the purpose of explaining it.

This reference system was formalized by Steinkellner (1988, 116–18), and revised in later publications, e.g., Steinkellner, Krasser, and Lasic (2005, 1:lii–liv). The latest version is described and argued for by Steinkellner (2017, xx–xxv). Here, only the more common cases will be discussed.

The qualifier i is of somewhat limited use in the present investigation: if a passage in a given text is marked like this, it is mainly a helpful reminder to the reader that the passage in question is quoted somewhere else, and that this occurrence has been considered in editing the present text. It does not actually mark the passage as a quotation in the context of the text being edited.

This distinction becomes relevant when a passage is reused more than once by subsequent authors, or when it is contained in a passage that is itself a case of text reuse. See the discussion of example 8.

Steinkellner (2017, xxv) has recently maintained: “When comparing two texts, it is not always possible to determine the nature and origin of specific noticeable differences, yet the difference between changes due to a text being adapted to a new context and those due to scribal and/or editorial work is usually recognizable.” There is, as yet, no consensus on this point.

Steinkellner (2017, xxi–xxv) summarizes and responds to the main concerns that have been raised against his suggestions. Since the debate is, by and large, about the certainty that can be had about what these symbols posit about a span of text, the issues raised should be further discussed in publications pertinent to the corresponding field of research.

The modern editions of these texts generally do use the categories of modern printing. Insofar as texts are digitized not from manuscripts but from printed editions (as are many of the texts in the SARIT library, accessed October 10, 2020, https://github.com/sarit/SARIT-corpus), the categories are a better, though not perfect, fit. For further observations on the history of book printing in relation to South Asian manuscript culture, see Pollock (2007).

See also table 1 for an overview together with the attributes most useful for linking the segment of text to either a source or a parallel instance.

The element <cit> is used also in dictionaries, but this is not the issue here.
14 The *Chicago Manual of Style* (University of Chicago Press 2010, 7.58) describes the situation like this: “When a word or term is not used functionally but is referred to as the word or term itself, it is either italicized or enclosed in quotation marks.” It is for this kind of phenomenon that the \texttt{<mentioned>} tag is intended.

15 Unless indicated otherwise, all XPath and XQuery expressions shown in this paper operate in the default namespace, http://www.tei-c.org/ns/1.0.

16 A more elegant solution would be to derive the elements’ names from the corresponding documentation (or ODD file). See example 1 for a slightly less naive approach that returns the local names of the elements in \texttt{model.attributable} and its immediate subclasses. Even this simple example shows that “a search for all quotations” in a TEI document needs to take a schema as its basis, for it includes \texttt{<floatingText>} in its results, which is, however, differentiated from quotes in an important respect: “… the semantics of \texttt{<quote>} suggest that its content derives from a source external to the current text, \texttt{<floatingText>} carries no such implication …” ([TEI Consortium 2020, section 4.3.2](https://www.tei-c.org/tei-p5/doc/tei-xhtml/html/blackwell-2013-the-text-encoding-initiative-for-tei-professionals.html)).

17 At least if a well-construed system is used. A noteworthy example is the definition of the Uniform Resource Name (URN) defined for the “Canonical Text Services” (CTS) in the Homer Multitext Project (http://www.homermultitext.org/hmt-doc/cite/texts/ctsoverview.html see Blackwell and Smith 2013). Berti, Almas, et al. (2015) show that this reference format has proven useful even for organizing fragments, i.e., text passages that have been taken from works that did not survive as a whole. Applying this observation, Berti, Blackwell, et al. (2016) provide a detailed discussion and case-study of constructing valid URNs that are expressive enough to encode text reused in several ways, together with its referential aspects. Especially interesting here are their proposed solutions for encoding text that is either reused in a direct quotation, alluded to, or repeated for further explanation. They propose a very general system of aligning texts, mainly through the judicious use of three locators: one for the passage as found in a canonical edition, one for the passage as found in the work reusing it, and one for the new edition of the work as constituted by the collection of the reused passages.
This example is taken from a file produced by the project *Search and Retrieval of Indic Texts (SARIT)*. This project digitized Sanskrit texts from printed publications, and the first-level markup (as shown in example 3) consisted in marking the typographical features of the text encoded: in this case, `<hi>` elements with `@rend` set to "bold".

Since these kinds of reuse are intended to point at the beginning of a passage, they should usually not be translated, or the translation should just indicate which words are at the beginning of the translated passage. When the translation happens to follow the order of words in Sanskrit, this issue remains hidden. The translation in example 3 shows two of these solutions.

See, for example, Steinkellner, Krasser, and Lasic (2005, 1:lvii): “Words from the text commented upon—as *pratīka*, in grammatical or lexicographical explanations, or as used in an explanatory context—are in bold face.” (The term *pratīka* is here used to describe the quotations that are mainly pointers to the base text.)

One might consider encoding this information in a `@corresp` attribute. But this could lead to misunderstandings, as the TEI Guidelines (16.5.1: Correspondence) state: “the correspondence relationship is not from one to the other, but between the two objects. It is thus different from the `@target` attribute” https://tei-c.org/Vault/P5/4.1.0/doc/tei-p5-doc/en/html/SA.html#SACS1. In the current case, we would want to encode precisely such a “from” relationship: from the commentary to the base text, but not the other way around.

Silent reuse is very common across the genre of scholastic Sanskrit. See Gnoli (1960a, xix n. 2) for an example from Maṇḍana (“Sphoṭasiddhi,” 99.1–100.15, in Bhatt and Ramanujam 1958): the reused passage is not especially differentiated from the surrounding text, which is closed simply by the phrase *kevalam eṣa manorathah* (Sphoṭasiddhi, 100.15–16, my translation: “This is only wishful thinking.”).

Well-known examples are Ravigupta’s reworking of Prajinākaragupta’s “Pramāṇavarttikālaṅkārabhāṣya” (see Sāṅkṛtyāyana 1953), or Ratnakirti’s writings, discussed below.

https://tei-c.org/Vault/P5/4.1.0/doc/tei-p5-doc/en/html/SA.html#SACSAL

https://tei-c.org/Vault/P5/4.1.0/doc/tei-p5-doc/en/html/SA.html#SAAG
AUTHOR

PATRICK MCALLISTER

Patrick McAllister is currently a researcher at the Austrian Academy of Sciences’ Institute for the Cultural and Intellectual History of Asia. His main research interest is the development of Buddhist epistemological theories during the ninth to eleventh centuries (primarily in the works of Prajñākaragupta, Jñānaśrīmitra, and Ratnakīrti).