Vol.2017-CH-114 No.2
2017/5/13

仏教文献における注釈構造の可視化に関する予備的研究 —パーリ語仏教文献を事例として—

青野道彦[†1]　永崎研宣[†2]　下田正弘[†3]

概要：仏教文献には数多くの注釈書が存在しているが，それらは仏教研究の貴重な資料となっている．それらを厳密かつ効率的に参照できるようにするためには，内容に即して構造化し，原典と関係づける作業が必要となる．今日，仏教文献の多くは大正新脩大藏經テキストデータベース，ACIP (Asian Classics Input Project)，VRI (Vipassana Research Institute) などによりデジタルテクストが整備されており，それらを用いることで柔軟に構造化し，文献の特質に即して原典と注釈書とを関係付けることが可能である．筆者は，現在，パーリ語仏教文献について，原典と注釈文献とを TEI P5 ガイドラインに準拠してマークアップすることで，両者を相互参照可能なテクストとして構造化することを進めているが，本稿ではその際に直面した課題について論じる．

キーワード：TEI P5，パーリ語，仏教文献，註釈

Preliminary Research on the Method for Marking up the Structure of the Commentaries on the Buddhist Scriptures — Mainly based on the Pali Buddhist Literatures —

MICHIHIKO AONO[†1]　KIYONORI NAGASAKI[†2]
MASAHIRO SHIMODA[†3]

Abstract: There are a huge number of commentaries on the Buddhist scriptures, which have played an important role in the Buddhist studies in that the commentaries are helpful to reveal the meaning of the scriptures. It formerly required some effort to make reference to the commentaries and gain exact access to the partinent part of them. Recently, the SAT Daizōkyō Text Database, the Asian Classics Input Project, and the Vipassana Research Institute release a voluminous amount of the digitized texts of the Buddhist scriptures and commentaries, which make it possible to perform full-text search. However, it remains difficult for non-specialists to access the relevant part of commentaries, partly because they lack the basic knowledge of the unfamiliar rules and complicated structure of commentaries. In order to seek an effective way to smooth the difficulties, I am currently encoding the commentaries on the Pali Buddhist scriptures and marking up their structure according to the TEI P5 Guideline. In this short paper, I will point out some concrete problems which I encountered when trying to apply the Guideline in the case of the Pali Buddhist scriptures and their commentaries.

Keywords: TEI P5, Pāḷi, Buddhist Literature, Commentary

1. はじめに

仏教文献研究において注釈文献はきわめて貴重な資料である.原典を解明する際に主要な参照資料となるとともに，原典に対する解釈の伝統を検討する際に主要な考察対象となるからである．

注釈文献の仏教文献に占める割合を測ることは容易ではないが，その分量は膨大である．筆者が研究対象としているパーリ語仏教文献では，原典であるパーリ (pāḷi) に対して，一次的な注釈文献であるアッタカター (aṭṭhakathā) と二次的な注釈文献であるティーカー (ṭīkā) が存在する．更に，それらにパーリ語と現地諸語（シンハラ語，ミャンマー語，タイ語など）とを用いたバイリンガルの注釈文献が加わる[a]．それらをひっくるめると，注釈文献の分量は原典を遥かに凌ぐ．

従来，このパーリ語注釈文献を参照し，目当ての注釈箇所にアクセスするためには，当該のテクストに対する一定の理解と慣れが必要であった．その後，テクストのデジタル化が進み，注釈文献の参照手段が大きく変わった．VRI (Vipassana Research Institute) などによりデジタルテクストが制作・公開され，全文検索が可能となり，注釈箇所へのアクセスが容易になった[b]．その結果，注釈文献を参照する人々の裾野は広がり，当該テクストについて十分な知識を有さずともアクセスできるようになった．

しかし，同じく仏教文献を研究対象としていても，研究領域を異にする研究者にとって，専門分野から外れる注釈

†1 東京大学文学部助教
Research Associate, The University of Tokyo
†2 人文情報学研究所主席研究員
Senior Fellow, International Institute for Digital Humanities
†3 東京大学大学院人文社会系研究科教授
Professor, The University of Tokyo, Graduate School of Humanities and Sociology.

[a] 参考文献 [1] を参照.
[b] 参考文献 [2] を参照.

ⓒ2017 Information Processing Society of Japan

文献の参照には依然として制約がある．注釈文献には独自の注釈技法なり注釈規則なりがあるためである．それは専門家にとっては暗黙の裡に了解されていることであっても，それ以外の人々にとっては常識ではない．このため，全文検索が可能となったからと言って，目当ての注釈箇所に必ず辿り着くことができるというわけではないのである．

この他，検索結果の利用にも制約がある．目当ての注釈箇所が見つかったとしても，当該箇所が注釈文献中のどの部分に含まれ，どの様な文脈で述べられているのか把握するのは門外漢にとって決して容易なことではない．注釈文献の構造を捉えるのには一定の労力がかかり，デジタル化により利便性が向上したにもかかわらず，その利用には二の足を踏む研究者が少なからずいるようである．

以上の様な制約は本人の努力により取り除かれるべき側面もあるが，当該テクスト，当該分野の専門家の方でもその制約を取り払う努力をする必要があろう．特に，専門性が高く，研究の担い手が少ないテクストにおいては，異分野の研究者を巻き込み, 新たな知見を取り込むためにも，必要であろう．

その方法であるが，一つは訳注研究であり，それは文脈を捉える上できわめて有用である．しかし，注釈文献は分量が膨大であるため, 訳注研究は遅々として進んでいない．この点でより現実的と思われるのは，既存のデジタルテクストを活用し, XML により註釈構造を可視化することである．それは，訳注研究より作業負担が少なく，膨大なテクストを処理していく上でより効率的である．また，XML を採用することで注釈対象である原典とリンクすることができ，原典と注釈文献の関係を明確化する上でより有用であろう．

この様な見方に立ち，現在，筆者はパーリ語仏教文献の注釈構造を XML により可視化し，最終的に注釈文献を原典とリンクさせ，相互参照を可能なものとする方法を模索している．この作業において基本となるのはテクストのマークアップであるが，準拠したのは人文学資料を構造化する際のデファクトスタンダードとなっている TEI P5 ガイドラインである[c]．基本的には，パーリ語仏教文献の構造分析にもこのガイドラインに示されたタグを用いることができる．しかし，TEI が欧米の文献を前提にして作られたものであるためか，そのタグセットをそのまま利用するだけではテクストに忠実なマークアップができない場合がある．以下，筆者が TEI P5 ガイドラインをパーリ語仏教文献に適用する際に直面した課題のうち，基礎的なものを取り上げ，その対処案を提示したい．

2. パーリ語注釈文献の特徴とそこから見える課題

ここで，パーリ語の注釈文献アッタカターの特徴を簡単に説明し，問題の所在を明らかにしておこう．

2.1 バーシュヤという注釈形式

先ず，パーリ語注釈文献を特徴づけるものとしてバーシュヤ (bhāṣya) という注釈形式がある[d]. これはサンスクリット語や中期インド語で書かれた注釈文献及びその漢訳・チベット語訳等に広く見られるものであり，注釈対象の語又は句を提示して，語の意味を同義語や類義語に置き換えることで明示するものである．それだけで注釈が終わる場合もあるが，それに続けて内容について解説する場合が多い. その説明は一行にも満たない短いものの場合もあれば，数行，数頁にも及ぶ長いものの場合もある．

問題となるのは長いものの場合である．通常，注釈箇所をマークアップする際に用いるタグは<gloss>であるが，長い註釈においてはそれが使えないときがあるのである．バーシュヤ形式の注釈はしばしば他の文献から文章を引用したり，韻文を提示したりして，それを典拠として内容について解説することがある．その様な場合には，引用文を示す<quote>や韻文を示す<l>，<lg>を使うことが求められるが，<gloss>はそれらのタグを内包することができないのである．

そもそも“gloss” (Gk. glossa, Lat. glossa) とは本来的には本文の行間や余白に記される簡単な語句説明を指す語である[e]．TEI P5 ガイドラインのマークアップ例を見ても，このタグは本文に添えられた注記，又は，本文に組み込まれた注釈をマークアップするためのものである．原典と独立して存在するパーリ語注釈文献をマークアップする上で<gloss>は適当ではなく，それに代わる別のマークアップ方法を検討する必要があろう．

問題はそれにとどまらず，<gloss>の対象をマークアップする方法についても存在する．通常，注釈対象語は<term>でマークアップされるが，それが適切でない場合があるのである．パーリ語注釈文献は原典から句を引用して，そのうちの一部の語句についてしか注釈しない場合があり，この様な場合にその句には注釈対象語句以外のものも含まれる．それを<term>でマークアップし，<gloss>と関係づけるのはテクストに忠実ではない．そこで，引用と注釈対象語句とを別個にマークアップする方法について検討することが求められる．

2.2 アプッバパダヴァンナナーという注釈規則

もう一つ，パーリ語注釈文献を特徴付けるものとしてアプッバパダヴァンナナー (apubbapadavaṇṇanā) という注釈

[c] 参考文献 [3].
[d] バーシュヤ形式については参考文献 [4] の p. 171 以降に詳しい.
[e] 参考文献 [5] [6] の“gloss”の項目を参照.

規則がある[f]．アプッパパダヴァンナナーとは「前にない語句の説明」を意味し，基本的には原典における初出の語句にのみ説明をし，同一語句が原典の後続箇所で再び現れた場合には説明を省くという規則である．この注釈規則は全体の分量を減らし，伝承の負荷を軽減するのに役立ってはいるが，読み手は目当ての注釈を原典の前の部分に遡って探すことが求められる．

一見，これは簡単な作業に見えるかもしれないが，全文検索が可能な今日においても，目当ての注釈を探し出すのは厄介である．後続箇所で現れる語句が前出の語句と数・格が同じであるとは限らず，結果として，同じ語句でも文字列としては異なっていて検索してもヒットしない場合があるからである．さらに，語として単独で注釈されず，句として注釈される場合もあるからである．これらは検索を難しくし，目当ての注釈箇所へのアクセスを阻害する要因となる．この阻害要因を除きアクセスを容易なものとするための方策を検討する必要があろう．

3. マークアップ試論

前項で二つの課題を提示したが，以下ではそれらに対処する方法について検討したい．

3.1 バーシュヤ形式の注釈文献をマークアップする手法

TEI P5 ガイドラインによると，注釈箇所をマークアップするタグは<gloss>であり，注釈対象をマークアップする代表的なタグは<term>であるとされる．そして，@xml:id と @target という属性を用いて，<gloss>と<term>とを関連付けるべきであるという．

ところで，このマークアップの手法はパーリ語注釈文献にはそのまま適用することができない．先述の通り，当該の注釈が他の要素を含んだりする場合に，<gloss>というタグが使えないことがあるからである．また，原典と切り離されたかたちで存在し，原典の一部を引用して，注釈するテクストにおいて<gloss>を使うのは一般的ではないからである．

そこで，<gloss>に代わるものが必要となるが，ここで<note>というタグの使用を提案したい．これは注記一般について用いられ，厳密さを欠くので，@type を付して，<note type="bhasya">というタグを用いる．ただし，このような場合には，専用タグを用意し，その要素によってマークアップできるようになることが望ましい．

次に，注釈対象についてであるが，先述の通り，パーリ語注釈文献では原典から句を引用して，そのうちの一部の語句についてのみ注釈を施す場合がある．その場合，その引用は注釈対象ではない語句も含むので，<term>でマークアップするのは正確ではない．そこで，引用を示す<quote>を用いて，原典からの引用箇所をマークアップして，その枠内において注釈対象語句を更にマークアップしていくことを提案したい．これにより，引用箇所と注釈対象語句の二つを明確化することができることになる．

3.2 省略された注釈を参照可能にするためのマークアップの手法

先述の通り，パーリ語注釈文献は，アプッパパダヴァンナナーという注釈規則を採用しているため，前に注釈した語句については繰り返しの注釈を行わない．そのため，目当ての注釈をテクストの先行箇所に遡って探す必要があるが，以下では，その負担を軽減するための方法について検討したい．

(1) 注釈対象語句と注釈範囲のマークアップ

目当ての注釈箇所を探す負担を軽減するために先ず行うべきことは，注釈文献の注釈対象語句を網羅的にマークアップすると共に，それら語句に対する注釈箇所をマークアップすることである．これにより，注釈対象語句に限定して検索することが可能となると共に，註釈箇所を正確に把握することができるようになる．

先述の通り，通常，注釈対象語句をマークアップするタグは<term>である．しかし，見出し語を設け，注釈対象語を検索しやすくするために，<term>の代わりに<w>を用いることを提案したい．現行のガイドラインでは，見出し語を表すための属性である@lemma は<term>という要素のもとでは用いることができないからである．以下に<w>を用いたマークアップ例を示す．

```
<quote type="basetext" xml:id="A01" corresp="B01 ">
  <w lemma="assosi" xml:id="A01-1">Assosi</w>
  <w lemma="kho" xml:id="A01-2 ">kho</w>
  <w lemma="verañja" xml:id="A01-3">verañjo</w>
  <w lemma="brāhmaṇa" xml:id="A01-4">brāhmaṇo</w>
</quote>
<note target="#A01">ti
  <note target="#A01-1">assosī ti . . . . </note>
  <note target="#A01-2">Kho ti padapūraṇamatte . . . . </note>
  <note target="#A01-3">Verañjāyaṃ jāto . . . . </note>
  <note target="#A01-4">Brahmaṃ aṇatīti . . . . </note>
</note>
```

図 1　注釈対象語のマークアップ

注釈対象が語ではなく句である場合には，<w>ではなく<phr>を用いることを提案したい．そして，見出しを<orig>

[f] 参考文献 [2] の p. 214, 230, 256, 293 を参照．

というタグで示すことにしたい. 以下に<phr>を用いたマークアップ例を示す.

```
<quote type="basetext" xml:id="A02 " corresp="B02 ">
  <phr>
    <choice>
      <reg>Taṃ kho panā</reg>
      <orig>taṃ kho pana</orig>
    </choice>
  </phr>
</quote>
<note type="bhasya" target="#A02">
ti itthambhūtākhyānatthe . . . .
</note>
```

図 2　注釈対象句のマークアップ

(2) **原典と注釈文献のリンク**

以上の作業を行った上で，更に二つの作業を行うことで，目当ての註釈箇所を探す負担を大幅に軽減することができる．

一つ目は注釈文献を原典とリンクさせることである．その方法であるが，注釈文献の引用語句とそれに対応する原典の被引用語句とを@xml:id と@correp を用いて関連付けることを提案したい．以下は，図 2 と対応する原典のマークアップ例である．

```
<phr xml:id="B02" corresp="A02">Taṃ kho pana</phr>
bhavantaṃ gotamaṃ evaṃ . . . .
```

図 3　原典のマークアップ 1

これにより注釈文献の引用箇所と原典の被引用箇所とがリンクされる．しかし，これだけではアプッパパダヴァンナナーという注釈規則への対処法として不十分である．原典の後続箇所で引用箇所と同一語句が現れた時にも参照できるようにする必要があり，そのために@correp によって前出の引用箇所と関連付けることを提案したい．例えば，図 3 と同じ句が原典の後続箇所で現れた場合，以下の様にマークアップする．

```
<phr corresp="B02">Taṃ kho pana</phr>bhavantaṃ gotamaṃ
evaṃ . . . .
```

図 4　原典のマークアップ 2

以上の二つの作業を行うならば，原典と注釈文献がリンクされ，相互参照が可能となると共に，原典の後続箇所からも当該の註釈箇所を確実に参照できるようになる．ただし，二つ目の作業は多大な労力が必要となると共に，内容理解が必要であり，専門性が求められるものである．

4. おわりに

以上，パーリ語注釈文献を専門外の研究者でも参照できるようにするためのマークアップ手法について，バーシュヤ形式の注釈技法とアプッパパダヴァンナナーという注釈規則の二点に焦点を当てて検討してきた．

先ず，バーシュヤ形式の注釈技法に対処するために，従来のマークアップの方法に修正を加え，注釈対象語句を<term>ではなく<w>と<phr>によりマークアップし，それに対応する注釈箇所を<gloss>ではなく<note type="bhasya">を用いてマークアップする方法を提示した．このうち，<note type="bhasya">については，属性ではなく，要素によってマークアップする方法を模索すべきことを指摘した．

次に，アプッパパダヴァンナナーという注釈規則に対処するために，注釈文献にある引用語句を媒介として原典と注釈文献を@xml:id と@correp を用いてリンクさせる方法を提示した．また，原典の後続箇所に現れる同一語句を@corresp を用いて初出の語句とリンクさせる方法を提案した．これにより，目当ての注釈箇所を確実に参照することが可能となるが，注釈文献と原典とのリンクを完全に達成するには相当の労力を要する上，専門家でなければ難しい作業であることを指摘した．

上記のマークアップ手法はパーリ語仏教文献のみならず, 他の仏教文献においても応用できるものである. 先ず, バーシュヤ形式の注釈文献はサンスクリット語や中期インド語の注釈文献及びその翻訳資料にも見られ，ここで提示したマークアップの方法はそれら文献においても応用可能である．次に，アプッパパダヴァンナナーはパーリ語注釈文献に特有の注釈規則であるが，繰り返しの注釈を避け，分量を削減する努力は他の仏教注釈文献にも見られるものである．その様なテクストのマークアップにも，ここで提示した手法を応用することができる．

なお，実際にこのマークアップを研究者が活用できるようにするためには, このマークアップに対応した Viewer が必要である．現在, Javascript を用いて Web ブラウザ上で動作するものを開発中であり，それに関しては次の機会に報告したい．

謝辞　本研究は JSPS 科研費 JP16K16693 と JP15H05725 の助成を受けたものです．

参考文献

[1] von Hinüber, Oskar. *Handbook of Pali Literature*. Walter de Gruyter, 1996.
[2] “Pāḷi Tipiṭaka”. http://www.tipitaka.org/, (cited 2017-04-16).
[3] “TEI P5: Guidelines for Electronic Text Encoding and Interchange 3.1.0”. http://www.tei-c.org/release/doc/tei-p5-doc/en/html/index.html, (cited 2017-04-16).
[4] Gary A. Tubb and Emery. R. Boose. *Scholastic Sanskrit: A Manual*

ⓒ2017 Information Processing Society of Japan

for Students. The American Institute of Buddhist Studies at Columbia University in the City of New York, 2007.
[5] "Oxford English Dictionary". http://www.oed.com/, (cited 2017-04-16).
[6] "Catholic Encyclopedia". http://www.newadvent.org/cathen/, (cited 2017-04-16).

ⓒ2017 Information Processing Society of Japan