

英訳語を見出し語としたデジタル仏教語彙集作成の 可能性と問題点について

辻村優英[†]

近年、チベット仏教と近代科学の対話が進んでいる。そのなかで重要なのは、仏教概念と近代科学の概念のすり合わせであり、そこで用いられる共通言語は英語である。本発表では、分野横断的研究を視野に入れたデジタル仏教語彙集作成の可能性と問題点について報告する。

Making a Digital Glossary of Buddhist Terms, which has entries of English translation.

MASAHIDE TSUJIMURA[†]

Recently, Tibetan Buddhists and modern scientists have a dialogue. It is important, in that dialogue, to adjust the differences of concepts between Buddhist and scientist, based on English. In this presentation, I would like to show a possibility and challenges of making a digital glossary of Buddhist terms, taking interdisciplinary research into consideration.

1. はじめに

これまでサンスクリットやチベット語の仏教用語を英語によって説明する辞書は数多く作成され、仏教思想の内実を理解するための多大なる恩恵をもたらした。特に近年では、東京大学の斎藤明氏・高橋晃一氏らによって、俱舎論を中心とした様々な資料にもとづく仏教用語の定義的用例一瞥することのできる非常に有益な成果『俱舎論を中心とした五位七十五法の定義的用例集』¹⁾が刊行され、その web サイ

ト[a]も公開されている。筆者はこの研究の方法論に多大なる影響をうけ、ダライ・ラマ 14 世が用いる英訳語を中心としたリストを作成し、それにかんする論文とリストを高野山大学密教文化研究所紀要第 25 号にて発表した²⁾。このリストはアサンガのアビダルマサムツチャヤに説かれる五十一心所のチベット語訳を見出し語とし、チベット語の定義的用例およびダライ・ラマ 14 世らの英語訳とその定義的用例を辞書的に記載したものであった。

仏教用語を見出し語とした辞書や定義集は、それらの言葉がどのような意味内容を内包しているかを知るのに役立つ。仏教用語を現代語によって解するという方向であり、その出発点は仏教用語(サンスクリット・チベット語・漢語など)にある。

他方、これとは別の方向も考えることができる。例えば認知科学の研究者が、*feeling* という言葉について仏教思想ではどのように捉えられているか探る場合であり、その出発点は仏教用語ではなく、*feeling* という英語にある。

本発表では、分野横断的研究への貢献を視野にいれて、英訳語を出発点(見出し語)とした仏教用語の定義的用例集構築の可能性と問題点について検討したい。

2. 着想に至った背景について

上述のように現在、筆者はダライ・ラマ 14 世が用いる仏教用語の英訳語を収集しており、極めて簡単なものであるがそのデータベース化を進めている。本報告事例はその作業のなかから生まれたものである。その着想に至った背景は主に 5 つある。

- ①チベット仏教側に仏教思想の現代科学的説明を求める動きがある³⁾。
- ②西洋の科学者たちの間で仏教思想にヒントを得ようとする動きがある⁴⁾[b]。
- ③Mind & Life Institute[c]にてダライ・ラマを中心とするチベット仏教と現代科学との対話が 24 年間にわたって続けられている。
- ④分野横断的な研究(特に文理融合)において概念のすり合わせが必要だが、その基準となるのが英語である。
- ⑤仏教関連のジャーナルへの他分野からのアクセスがしにくい。その逆も同様である。したがって、双方が利用可能なプラットフォームとして機能する語彙集を作

a) http://www.l.u-tokyo.ac.jp/intetsu/html/akbh_75dharma.html

b) 例えば Science 誌上では以下のような論述がある。Marcia Barinaga "Studying the Well-Trained Mind"(Science 3 October 2003: Vol. 302 no. 5642 pp. 44-46), "The Mind: Buddhism and Biology"(Science 12 April 1991: Vol. 252 no. 5003 p. 206), Yudhijit Bhattacharjee "Neuroscientists Welcome Dalai Lama With Mostly Open Arms"(Science 18 November 2005: Vol. 310 no. 5751 p. 1104), Esther Sternberg "A Compassionate Universe?"(Science 3 February 2006: Vol. 311 no. 5761 pp. 611-612), Greg Miller "A Quest for Compassion"(Science 24 April 2009: Vol. 324 no. 5926 pp. 458-459)

c) <http://www.mindandlife.org/>

[†] 総合地球環境学研究所・高野山大学密教文化研究所
Research Institute for Humanity and Nature, the Research Institute of Esoteric Buddhist Culture, Koyasan University

ると便利だと考えられる。

3. チベット語文献電子データ化事例

チベット語の古典文献の中心となるものは、インド由来の経典類をチベット語に訳した「チベット大蔵経」、およびチベット人学僧の著した「蔵外文献」であろう。その資料形態はペチャ(dpe cha)[d]と呼ばれるチベット独特のもので、細長い紙の裏表に横書きのチベット文字が木版印刷されており、その一枚一枚は綴じずに1巻分重ねられ、大抵黄色い布に包まれて保管される。中国のチベット本土や亡命チベット人コミュニティのあるインドのダラムサラなどでは洋綴じ本の形式で出版され直されているが、それでもやはりチベット語古典文献の基本的な形態はペチャの形式である。他方、ダライ・ラマ14世や他の現代チベット人の著作はペチャの形式ではなく、洋綴じ本の形式で出版されている。このようにチベット語文献の資料形態の主なものにはペチャ形式と洋綴じ本形式の2種である。こうしたチベット語文献(特に古典チベット語文献)のデジタル化が進んでいる。以下にその代表的なものを示す。

①TBRC(Tibetan Buddhist Resource Center)[e]

ペチャのスキャン画像のPDFを公開している。木版印刷されたそのままの状態のものを閲覧できる利点がある一方、テキスト検索はできない。

②ACIP(Asian Classics Input Project)[f]

ペチャに記されたチベット文字をACIP独特の方式によってアルファベットに転写したテキストデータと、チベット文字そのままのPDFのデータを公開している。テキストデータは、出典箇所(ペチャの第**葉の裏表どちらかの第**行)を明示できるよう工夫されており、テキストデータの1行の文字列はペチャに書かれている1行の文字列と一致する。ただ、改行によって分断された単語は検索にかからない。

③THL(The Tibetan and Himalayan Library)[g]

チベット文字フォント(the Tibetan Machine Unicode font) [h]による表示。

④Tibetan-Sanskrit 構文対照電子辞書 eDic [i]

相場方式によってアルファベット転写されたテキストデータを公開している。一文

d) 本稿におけるチベット語は拡張ワイリー方式(後述)によって転写する。

e) <http://www.tbrc.org>

f) <http://www.asianclassics.org>

g) <http://www.thlib.org>

h) <http://www.thlib.org/tools/scripts/wiki/tibetan%20machine%20uni.html>

i) <http://suzuki.ypu.jp/edic>

単位で区切られている。

⑤『俱舍論を中心とした五位七十五法の定義的用例集』[j]

Wylie方式によって転写されたテキストデータを公開している。その名の通り定義的用例を集めたものなので、サンスクリットの仏教用語を見出し語とし、そのもとにサンスクリット・漢訳・チベット訳の定義的用例および、その英語訳・フランス語訳をひとつのまとまりとして記述している。

4. 入力データとアノテーションについて

4.1 入力データ

本事例において扱うデータは、以下のようなものである。

①英訳見出し語(ダライ・ラマ14世の著作に見られるもの)

②英訳見出し語の原語となるチベット語

③チベット語の定義的用例

④チベット語の定義的用例の出典:筆者,出版年,タイトル,出版社,出版地,ページ。

⑤英語の定義的用例

⑥英語の定義的用例の出典:筆者(原著者),出版年,タイトル,英訳者,出版社,出版地,ページ。

Mind & Life Instituteにおける議論が英語とチベット語であることと、データ入力時に特殊文字をもちいることなく記述可能であるという理由から、本事例では英語とチベット語に限定している。

4.2 チベット語入力時のアルファベット転写について

データ入力および検索の便宜上、チベット文字フォントを利用するのではなく、アルファベット転写したものを使用する。

複数ある転写方式のなかで、おそらく最も広く用いられており、かつ特殊文字を使用することなく記述できるのはWylie方式⁵⁾であろう。しかし、この方式ではサンスクリットのチベット文字音写を記述することができない。この弱点を補ったものが福田洋一氏による拡張ワイリー方式(Extended Wylie Method)である⁶⁾。したがって本事例では拡張ワイリー方式によって記述している。

4.3 アノテーションのフォーマット

データをXMLによって記述する際の、アノテーションのフォーマットとしては、

j) http://www.l.u-tokyo.ac.jp/intetsu/html/akbh_75dharma.html

人文系資料のアノテーションに際して国際的に広く用いられている7)TEI (Text Encoding Initiative)のガイドラインに沿うのが妥当だと思われる。高橋晃一氏による「XMLによる仏教重要語彙定義集の作成と課題」8)ではTEI P5 [k]に従った例が示されている。本事例でもそれにならないTEI P5にしたがって記述している。

4.4 アノテーション例 1

上述の入力データのアノテーションの基本型は以下のようなものである。

```
<entry>
  <form>
    <orth xml:lang="en">①英訳語見出し</orth>
  </form>
  <sense>
    <def xml:lang="bod">②英訳見出し語の原語となるチベット語</def>
    <cit type="TibetanExample" xml:lang="bo">
      <quote>③チベット語の定義的用例</quote>
      <bibl>④チベット語の定義的用例の出典：筆者，出版年，タイトル，出版社，出版地，ページ</bibl>
    </cit>
    <cit type="EnglishExample" xml:lang="en">
      <quote>⑤英語の定義的用例</quote>
      <bibl>⑥英語の定義的用例の出典：筆者（原著者），出版年，タイトル，英訳者，出版社，出版地，ページ</bibl>
    </quote>
  </cit>
</sense>
</entry>
```

②は①の原語であるとともにその意味を定義するものでもあるので<def>によってタグ付けた。③④⑤⑥の定義的用例と出典は@typeにより英語のもの(EnglishExample)とチベット語のもの(TibetanExample)を区別したうえで、<quote>タグのもとに定義的用例の引用を記述し、<bibl>タグのもとに出典情報を

記述した。出典情報の詳細にかんしては<bibl>タグにおいてひとまとめに記述したが、TEI Headerに沿ってより細分化することも考えられる。

4.5 アノテーション例 2

英訳語とチベット語が一对一対応するものだけでなく一つの英訳語に該当するチベット語原語が複数ある場合が考えられる。この場合、以下の例の、<sense n="1">、<sense n="2">のようにタグ付けして区別した。

```
<entry>
  <form>
    <orth xml:lang="en">①英訳語見出し</orth>
  </form>
  <sense n="1">
    <def xml:lang="bod">②英訳見出し語の原語となるチベット語 1</def>
    <cit type="TibetanExample" xml:lang="bo">
      <quote>③チベット語の定義的用例 1</quote>
      <bibl>④チベット語の定義的用例の出典 1：筆者，出版年，タイトル，出版社，出版地，ページ</bibl>
    </cit>
    <cit type="EnglishExample" xml:lang="en">
      <quote>⑤英語の定義的用例 1</quote>
      <bibl>⑥英語の定義的用例の出典 1：筆者（原著者），出版年，タイトル，英訳者，出版社，出版地，ページ</bibl>
    </quote>
  </cit>
  </sense>
  <sense n="2">
    <def xml:lang="bod">②英訳見出し語の原語となるチベット語 2</def>
    <cit type="TibetanExample" xml:lang="bo">
      <quote>③チベット語の定義的用例 2</quote>
      <bibl>④チベット語の定義的用例の出典 2：筆者，出版年，タイトル，出版社，出版地，ページ</bibl>
    </cit>
    <cit type="EnglishExample" xml:lang="en">
```

k) <http://www.tei-c.org/release/doc/tei-p5-doc/en/Guidelines.pdf>

<quote>⑤英語の定義的用例 2</quote>

<bibl>⑥英語の定義的用例の出典 2：筆者（原著者），出版年，タイトル，英訳者，出版社，出版地，ページ</bibl>

</quote>

</cit>

</sense>

</entry>

4.6 言語のアノテーションについて

IANA (Internet Assigned Numbers Authority)ではチベット語に関して次のようなものを用意している[1].

- ①Type: language, Subtag: bo, Description: Tibetan
- ②Type: language, Subtag: adx, Description: Amdo Tibetan
- ③Type: language, Subtag: khg, Description: Khams Tibetan
- ④Type: language, Subtag: otb, Description: Old Tibetan
- ⑤Type: language, Subtag: xct, Description: Classical Tibetan

古典文献からの引用は古典チベット語だが，ダライ・ラマ 14 世の著作では現代チベット語が用いられている．①によって両者を包括して `xml:lang="bo"` と記述することが考えられる．

次にチベット語を記述する文字が問題となるが，Tibetan and Himalayan Library におけるチベット文字による記述[m]も，「XML による仏教重要語彙定義集の作成と課題」における Wylie 方式のアルファベット転写による記述も，両者とも `xml:lang="bo"` との指定はあるが，`script` の指定はされておらず両者に区別はない．チベット語の `script` にかんして IANA で用意されているのは Type: script, Subtag: Tibt, Description: Tibetan のみである．TEI P5 ガイドラインの p.lv において zh-CN (Chinese in China, no script given), zh-Latn (Chinese transcribed in the Latin script) とする事例が示されている．チベット文字をアルファベット転写する場合は，これにならって `bo-Latn` とするのがよいだろうか．また転写方式の指定をどうするかという問題もある．

1) <http://www.iana.org/assignments/language-subtag-registry>
m)

<https://collab.itc.virginia.edu/wiki/toolbox/Rendering%20Tibetan%20Properly%20in%20Mixed%20Text%20Environments.html>

5. 可能性と問題点

本事例の可能性として以下のようなものが考えられる．

- ①英訳語を学問分野間のインターフェイスとみなし，それを中心にすることによって分野横断的研究のための概念のすり合わせに貢献する．
- ②ひとつの英訳語が複数の仏教概念を内包することがある．この場合，そのひとつの英訳語のもとに，複数の仏教概念を定義的用例とともに示すことによって，異なる仏教概念間の関連を見ることができる．

本事例の問題点として以下のようなものが考えられる．

- ①本発表事例における英訳語はダライ・ラマ 14 世が用いているものであるが，他のチベット人学僧や欧米の仏教学者が用いる英訳語をどのように扱うか．
- ②同じひとつの仏教概念に対して，複数の英訳語が与えられることがあるが，その場合どのように統一するか．

謝辞 ダライ・ラマ 14 世がもちいる仏教用語の英訳語リストアップにかんしてご助言賜った高野山大学教授室寺義仁氏，本発表の機会を与えてくださった総合地球環境学研究所准教授関野樹氏に謹んで感謝申し上げます．

参考文献

- 1) 斎藤明編：「俱舍論」を中心とした五位七十五法の定義的用例集，三喜房 (2011).
- 2) TSUJIMURA Masahide: The English translation of shes rab (prajñā) by the 14th Dalai Lama and the list of English words for the fifty-one mental factors (sems byung lnga bcu rtsa gcig), 高野山大学密教文化研究所紀要, 第 25 号, pp75-114 (2012).
- 3) 高松宏實(クンチョック・シタル): 近代におけるチベット仏教の直面する問題—チベット仏教と近代—, 現代密教, 第 21 号, 智山伝法院, pp. 159-174 (2010).
- 4) Lee, Joon: Cross-Cultural Consensus Between Buddhist Reality and Modern Science, International Journal of Buddhist Thought & Culture February, Vol.6, International Association for Buddhist Thought & Culture, pp.25-51 (2006).
- 5) Wylie, Turrell: A Standard System of Tibetan Transcription, Harvard Journal of Asiatic Studies, Volume 22, pp. 261-267 (1959).
- 6) Fukuda Yoichi: Extended Wylie Method of the transcription of Tibetan characters. <http://tibet.que.ne.jp/misc/EWylie1.html>
- 7) 永崎研宣: 人文学資料へのアノテーション～Text Encoding Initiative の挑戦, テキストアノテーションワークショップ・コンテスト予稿 (2012). http://nlp.nii.ac.jp/tawc/papers/W05_nagasaki.pdf
- 8) 高橋晃一: XML による仏教重要語彙定義集の作成と課題, 情報処理学会研究報告, Vol.2010-CH-86 No.2 (2010).

- 1) 斎藤明編: 「俱舍論」を中心とした五位七十五法の定義的用例集, 三喜房 (2011).
- 2) TSUJIMURA Masahide: The English translation of *shes rab (prajñā)* by the 14th Dalai Lama and the list of English words for the fifty-one mental factors (*sems byung lnga bcu rtsa gcig*), 高野山大学密教文化研究所紀要, 第 25 号, pp75-114 (2012).
- 3) 高松宏實(クンチョック・シタル): 近代におけるチベット仏教の直面する問題—チベット仏教と近代—, 現代密教, 第 21 号, 智山伝法院, pp. 159-174 (2010).
- 4) Lee, Joon: Cross-Cultural Consensus Between Buddhist Reality and Modern Science, *International Journal of Buddhist Thought & Culture* February, Vol.6, International Association for Buddhist Thought & Culture, pp.25-51 (2006).
- 5) Wylie, Turrell: A Standard System of Tibetan Transcription, *Harvard Journal of Asiatic Studies*, Volume 22, pp. 261-267 (1959).
- 6) Fukuda Yoichi: Extended Wylie Method of the transcription of Tibetan characters.
<http://tibet.que.ne.jp/misc/EWylie1.html>
- 7) 永崎研宣: 人文学資料へのアノテーション~Text Encoding Initiative の挑戦, テキストアノテーションワークショップ・コンテスト予稿 (2012).
http://nlp.nii.ac.jp/tawc/papers/W05_nagasaki.pdf
- 8) 高橋晃一: XML による仏教重要語彙定義集の作成と課題, 情報処理学会研究報告, Vol.2010-CH-86 No.2 (2010).