

Web 版神道用語グロッサリーの制作

—オンライン文献の有効なプレゼンテーションの提案—

井上順孝 ハイヴンズ・ノルマン 黒崎浩行

國學院大學日本文化研究所

要旨 本研究所では、神道・日本宗教関連英文・英訳刊行物の Web 公開を進めてきたが、研究・学習での活用可能性を高めるため、神道用語を起点に文献間をリンクしたグロッサリーを制作した。その過程と、XML や全文検索など現状・将来の Web 技術との比較考察を報告する。

The Production of the Web Glossary on Shinto Names and Terms:

A Suggestion on the Effective Presentation of Online Documents

INOUE Nobutaka, Norman HAVENS, KUROSAKI Hiroyuki

Institute for Japanese Culture and Classics, Kokugakuin University

Abstract: Our Institute released the online glossary which linked to the documents already published online via the names and terms on Shinto and Japanese religion, in order to raise utility of learning and studying. In this paper we report the process of the production and the comparison with the other Web technologies such as XML and the full text retrieval.

1 制作の目的と背景

國學院大學日本文化研究所では、平成 9 年から現在にかけて、研究所がそれまで刊行してきた研究成果のうち、とくに英文のものについて、インターネット公開を進めている (表参照)。これは、研究所設立当初から、海外の神道・日本宗教研究者への研究情報発信と相互交流を推し進めてきた活動の流れの一環として、インターネットという国境を超えたコミュニケーションメディアの普及状況に対応する試みであった。

そして、英文刊行物のオンライン公開が一段落した平成 12 年からは、公開したコンテンツを教育・研究に有効活用するための手段を検討する段階に移った。もちろん、公開されたテキストは Google などの一般的なサーチエンジンや國學院大學内のサーチエンジンを通じて、適切なキーワ

タイトル	刊行年	Web 版
<i>Basic Terms of Shinto</i>	1985 (1958)	1997
<i>Cultural Identity and Modernization in Asian Countries</i>	1983	1999
<i>Contemporary Papers on Japanese Religion series:</i>		
1. <i>Matsuri</i>	1998	1997
2. <i>New Religions</i>	1991	1997
3. <i>Folk Beliefs in Modern Japan</i>	1994	1997
4. <i>Kami</i>	1998	2000
<i>Globalization and Indigenous Culture</i>	1997	2001

表: 英文刊行物の Web 公開

ードによって容易に探し出せる状態にある。また、ひとつの完結した文献として、いかようにでも参照することができる。だがわれわれは、それだけではコンテンツの活用可能性が十分に引き出せないと考えた。

それは、神道・日本宗教に関する英文・英訳の刊行物、ということにかかわっている。そこには、当然のことながら神道や日本宗教に関するさまざまな専門用語の英訳が含まれている。これらの訳語はいずれも、ヘイヴンズによる継続的な翻訳作業の積み重ねのなかで、定訳として学界に定着していくことを目標に、十分な文献調査のうえで吟味、選択されてきたものである。

こうした翻訳作業の蓄積を広めるための手段としては、辞典を編纂して刊行するのがもっとも効果的であろう。それは本研究所編『神道事典』(平成6年刊)の翻訳として以前から進められている。だが、こちらは9部構成のうち1部(「第二部 神」)がようやく平成13年に刊行されたという状況である。そうした本格的な取り組みと並行して、あくまで神道・日本宗教用語の簡易な和英対照表といった位置づけで、なおかつハイパーテキストの利点を活かしながら、適切な訳語選択を支援することをめざして、Web版神道用語グロッサリーを制作することにした。

なお、制作にあたっては、文書検索・リンクづけとWebページ生成にかかわるプログラム(Rubyスクリプト)を黒崎が作成したほかは、現時点で先端的・実験的と思われるような技術や作業環境はあまり採り入れず、実作業を担当する非常勤スタッフの負担が内容の質的向上に集中できるよう心がけた。また、利用者側の環境も、最低限のインターネット接続環境と一般に普及しているWebブラウザで閲覧することを考慮した。

2 グロッサリーの概要

2.1 全体構成

グロッサリーの全体構成は、次のようになっている(図1参照)。

- 用語集。アルファベット順索引と各項目からなる。各項目は項目IDをもち、ローマ字表記、日本語表記、英訳、説明文、当該語句が出現する文献へのリンクからなる。頭文字のアルファベットごとに別ページになっており、http://www.kokugakuin.ac.jp/ijcc/wp/glossary/def_頭文字.html#項目ID というURLで参照できる。
- 文献群。従来から公開していた内容に、用語集の項目へのリンク(@ボタン)を加えている。@ボタンを選択すると、別ウィンドウが開いて、用語集の該当項目の記述が表示される。

各文献のページは、段落ごとにIDをつけている。本来ならばp要素にid属性をつけて、それを参照すればすむところだが、Microsoft Internet Explorerではid属性のある箇所にはジャンプしない。そのため、p要素内の先頭の子要素として空のa要素を加え、これにname属性をつけることで対処した。

例: `<p> ... </p>`

これにより、用語集の各項目から、文献ページ中の当該語句が出現する最初のパラグラフへ直

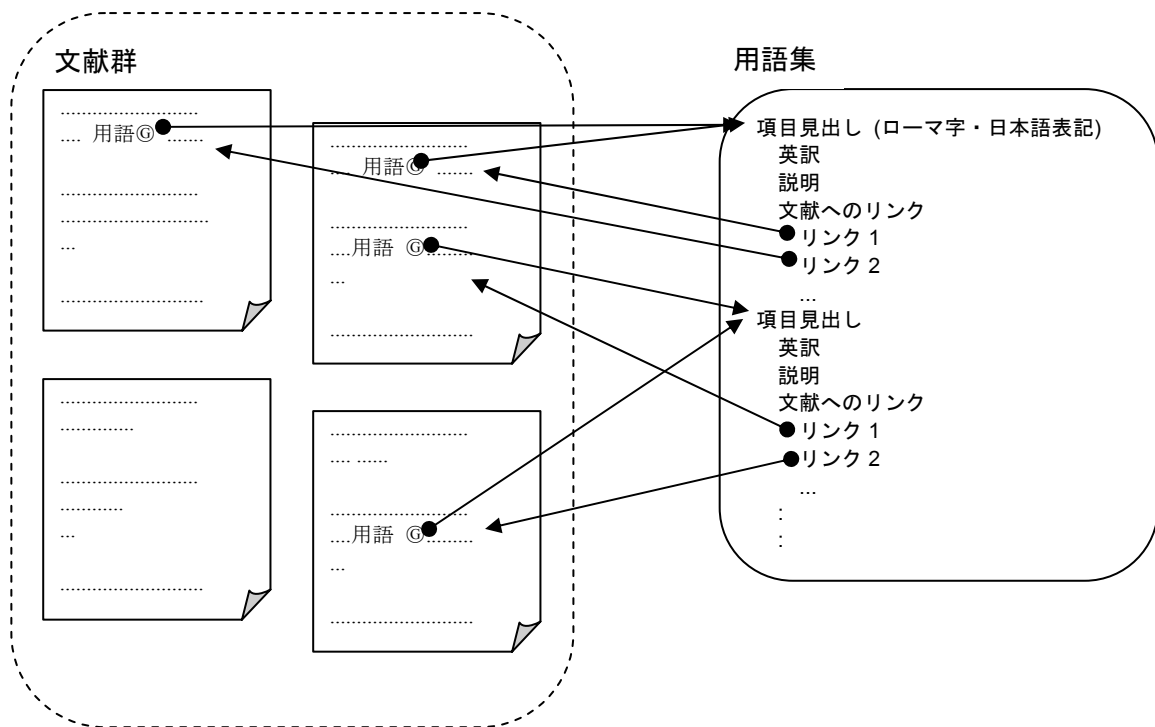


図 1: グロッサリーの全体構成

接リンクを張ることができ、リンクを選択するとすぐに該当パラグラフにたどりつける。

2.2 利用のイメージ

主として二とおりの利用方法を想定している (図 2 参照)。

2.2.1 イメージ 1: 用語集で語句を探す

知りたい語句があらかじめ決まっているときは、用語集のアルファベット順索引から目的の語句を探せばよい。項目に記述された内容で満足すればそれで終わりとなる。もし、さらに詳細を知りたい場合は、リンク先の各文献を順次たどって読むことになる。

2.2.2 イメージ 2: 文献から用語集へ、さらに別の文献へ

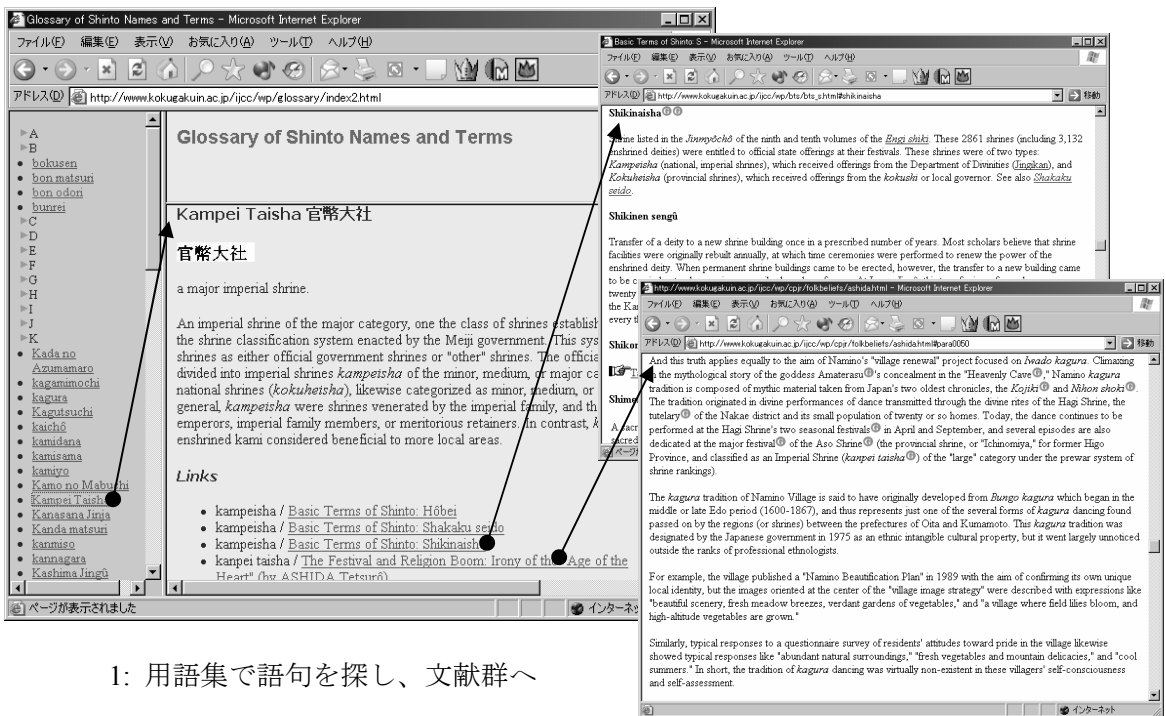
それぞれの文献をオンラインで読んでいるとき、G ボタンのある箇所の語句に興味があれば、ボタンをクリックして用語集を表示する。そこに、同じ語句が出現する別の文献もリストアップされているとき、興味があればそれらの文献も読んでみる。

3 制作過程

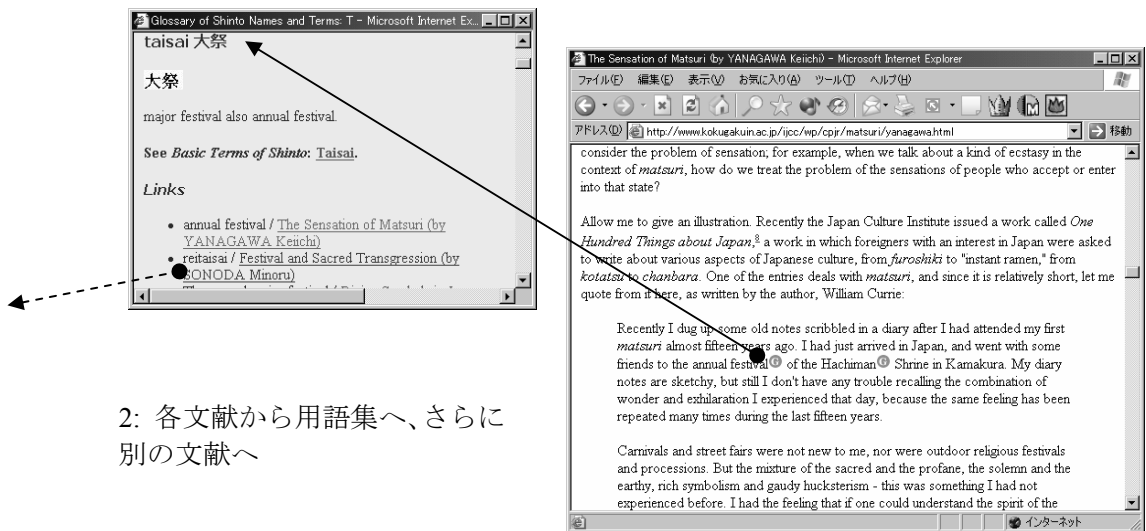
3.1 全体の流れ

簡易な英和対照表という水準とはいえ、制作にあたっては専門知識をとまなういくつかのステップを要した。

1. 項目の選定。英文刊行物のうち、*Contemporary Papers on Japanese Religion* シリーズに収録された論文の原文 (日本語) から、重要な固有名詞と、説明を要すると思われる神道・日本宗教関連の専門用語をピックアップした。そこから *Basic Terms of Shinto* および『神道事典』の収録語などを参考に重要度のランクづけを行い、最終的には 353 項目に絞り込んだ。



1: 用語集で語句を探し、文献群へ



2: 各文献から用語集へ、さらに別の文献へ

図 2: 利用のイメージ

2. 説明文の執筆。ごく簡潔な説明文を執筆した。
3. 用語と説明文の英訳。英訳文献を参照しながら行った。最終的にヘイヴンズが校訂した。
4. 項目から文献へのリンクの作成。後述。
5. 用語集の Web ページ生成。ここまでの作業を記録したデータファイルをもとに、自作した Ruby スクリプトにより静的な Web ページを生成した。
6. 各文献へのリンクタグ挿入。公開済みの各文献ページを加工し、文献の該当箇所から用語集にジャンプするためのタグを挿入した。自作のプログラム (Ruby スクリプト) で行った。
7. 検証、公開。

3.2 用語集から文献へのリンク作成

この節では、「リンク」の語は HTML コードとしてのリンクではなく、用語集の各項目から文献内の出現箇所への対応づけ、という意味で用いる。

このリンクは、次のような手順で作成した。

1. 項目となる語を選定する段階 (前節の手順 1) で、データファイル中に参照した文献 ID を記号で記録しておく。
2. 各項目のローマ字表記と英訳 (前節の手順 3) の 2 つのフィールドをもとに全文書を検索し、マッチした結果を出力するプログラムを実行する。出力レコードは、リンクのタイプ (`expected`, `unexpected`, `notfound` のいずれか) と、マッチした文字列@文献 ID#パラグラフ ID の形式からなる。また、検索結果に参照文献が含まれているかもチェックし、参照文献が含まれていない場合は、ダミーの出力レコードを `notfound` タイプとして加える。
3. 前の手順で生成されたデータファイルを HTML 形式に変換し、ブラウザでリンク先をたどって、適切なリンクかどうかを選別する。たとえば、`gyô` (行) の訳語は `religious practice` だが、これにマッチした箇所のなかには、日本語の原文では「行」ではないものもある。適切なリンクは `expected`、除外するリンクは `unexpected` としてデータファイルに記録する。
4. 項目選定時に参照した文献が先の手順による検索結果に含まれていない場合 (`notfound`) について、手作業で拾い、データファイルを修正する。

当初、各文献の HTML ファイルの日本語 (のローマ字表記) 部分はあらかじめ `em` 要素としてタグづけしており、そのタグを検索条件に加えることでふるい分けは容易になると予想していた。だが実際には、英訳語への検索が必要であったため、用語に対する検索結果で `unexpected` に分類されるケースは 2 割 (全出力件数 1099 中 229 件) に及んだ。したがって、作業にはプログラムを介しているものの、実態としてはほぼ手作業といえる。

4 考察

4.1 XML との比較

今回作成したグロッサリーは、中間データファイルの形式として一部 XML を用いてはいるものの、実は XML で実現可能な技術をあえて独自プログラムで加工・生成した HTML (ウィンドウのサイズ調整に JavaScript も使用) と、一般に普及している Web ブラウザで代替したものといえる。これは冒頭で述べたように、現時点で先端的・実験的と思われるような技術や作業環境はあまり採り入れない、という方針にもとづくものであった。

グロッサリーのリンク構造は、概念的には XLink の拡張リンクに相当するものである。ただし、最終的に HTML で出力したものを公開しているため、XLink の次のような特性は失われている。

- ロケータによって文書内のどの要素でもリソースとして指示できる。
- リソース間のアークを記した文書をリンク元文書とは別に作成でき、リンク元文書には手を加える必要がない。

つまり、用語集と各文献の双方向リンク、また用語集を介した文献間の相互リンクを HTML のみで実現するには、元の文献ページに次のような加工をしなければならなかった。

- 用語集から文献へ、パラグラフ単位でリソースを指示できるようにするため、各文献ページにパラグラフ ID をあらかじめ挿入しておく。それも本来ならば `p` 要素の `id` 属性でよいところだが、Microsoft Internet Explorer の仕様にあわせて `p` 要素内に空の `a` 要素を作り、これに `name` 属性をつけた。

- 文献から用語集へリンクするため、元の文献ページの該当語句初出箇所に a 要素を挿入し、href 属性で用語集の項目 ID を指示する。

現時点で XLink 対応ブラウザが普及していれば、こうした加工をするために独自の Web ページ生成プログラムを介することなく、汎用的なフォーマットの XML データの作成のみでリンクが構築できたと思われる。

また、他の研究機関との交流の一環として、グロッサリーからリンクする文献の範囲を、他の研究機関が公開しているものにまで広げることが考えられる。これも、XLink のようなリンク技術や、RDF のようなメタデータの集積・交換技術が普及すれば、よりスムーズに実現することができるだろう。

4.2 全文検索との比較

今回、用語集から各文献へのリンクは、相当な手作業を伴った。また、収録項目数も基本的な 353 項目に絞った。

利用者側からすれば、語句の意味や用法を知るには、各文献が全文検索できさえすればよく、またそのほうがより網羅的であるととらえられるかもしれない。

しかし、これは目的が異なる。冒頭に述べたように、グロッサリー制作の目的は、そうした検索を正しく行うためにも暗黙のうちに必要とされる、専門用語の和英対訳についての知識の蓄積を提示することにあつた。

たとえば、全文検索エンジンの検索語入力に対して英和・和英自動翻訳機能を加え、その辞書としてグロッサリーのデータを取り込むことができれば、両方の要求を融合することができるだろう。

いずれにしても、さまざまな副産物を含むような研究蓄積の有効な提示方法としては、現状の WWW のハイパーテキスト技術にもまだ大きな可能性が残っている。今後は、XML などの技術がこれを後押ししていく方向を見据えながら、普及状況をにらみつつ取り組んでいきたい。

[参考文献]

ヘイヴンズ・ノルマン「『神道事典』の英訳『KAMI』の神道研究における位置」『國學院大學日本文化研究所報』222 (2001): 3-5。

The World Wide Web Consortium, *XML Linking Language (XLink) Version 1.0*. W3C Recommendation 27 June 2001, <<http://www.w3.org/TR/xlink/>>.

國學院大學日本文化研究所

<http://www.kokugakuin.ac.jp/ijcc/>

オンライン刊行物

<http://www.kokugakuin.ac.jp/ijcc/wp/>

Glossary of Shinto Names and Terms

<http://www.kokugakuin.ac.jp/ijcc/wp/glossary/>