

歴史的な中国語教材を対象とした オンラインデータベース構築について

氷 野 善 寛

On the Construction of an Online Database for Historical Chinese-language Teaching Materials

HINO Yoshihiro

In recent years the number of digitized books has been growing considerably. Yet on the other hand, this digitization has been carried on an institution-by-institution basis, resulting in a fragmentation of these information resources and a need for organizing them. This led us to the concept of a database of digital archives that would catalogue and collate these dispersed resources, and specifically the construction of a database specialized for Chinese-language teaching materials. This paper begins with a discussion of why such a database is needed, and then discusses the concept, systems design, and future aims of the database that was actually constructed.

Keywords: Chinese-language teaching materials, database, digital archives

0. はじめに

この数年、筆者は『清代民国漢語文献目録』（遠藤光暁・竹越孝主編、2011）や「尾崎・内田文庫蔵英華・華英辞典一覧」『近代英華華英辞典解題』（沈国威編、2011）、「東西学術研究所文献目録」（氷野善寛、2013）などを通じて日本、中国、欧米等で刊行された中国語教材の目録を編集し、公開している。とりわけ1870年頃から1950年頃までに刊行された中国語教材¹⁾の目録作成に力を入れている。これは当時どのような教材があったのかを明らかにすることにより、通時的な中国語の語彙教育に関する研究を進めるためである。当時の中国語教材並びに中国語教育分野の研究では、魚返善雄 1940「アメリカの支那語研究」（『中國文學』第68号）にはじまり、1980年以降には六角恒広 1994『中国語書誌』、1985『中国語関係書書目』（2001増補版）といった詳細かつ大規模な目録が作成されている。『中国語関係書書目』については、当時の教材を網羅的に集めることを主な目的にしているため、広告や他の目録に掲載されている書目についても収録し、また教材だけではなく雑誌や読み物といったものまで幅広く集め、刊行年順に掲載している。一方、現物が確認できないものもあり、また、紙幅の制約もあり、詳細な書誌や所蔵情報についてはまだまだ十分とは言えないものもあり、さらなる調査の必要がある。目録以外には六角恒広編『中国語教本類集成』や波多野太郎編『中国語文資料彙刊』、張美蘭『日本明治時期漢語教科書彙刊』など、当時の教材を影印出版したものもある。それぞれ数多くの教材が掲載されているが、どこに所蔵されている教材を影印したものかということについては必ずしも明らかにされておらず、実際これらを利用して研究を進めるにはその点について確認が必要である。印刷物以外に目を向けると、インターネット上には多くのデジタル化された資料が見られるようになってきた。一例をあげると、日本国立国会図書館近代デジタルライブラリー、関西大学CSACデジタルアーカイブズ、早稲田大学古典籍総合データベース、ハーバード大学、カリフォルニア大学、オーストラリア国立図書館などが挙げられる。これらの図書館や大学では機関ごとにデジタル化した書籍を公開しているが、当然中国語教材に特化して資料をデジタル化し公開しているわけではない。また、それぞれ異なるシステムを運用しているため、書籍を横断的に探すことはできない²⁾。さらにインターネット上には来歴不明のデジタル書籍も多くあり、ソース情報の不足を指摘できる。

そこで中国語教材という枠を設けた上で、この分野の書籍目録を再度整理し、

1) ここで述べる中国語教材とは、会話書・教科書・辞書・文法書などを指す。

2) アメリカのInternet Archivesのようにデジタル化された書籍の所蔵先について、詳細情報に記している例も見られる。

- ① これまで刊行された中国語教材の実態
- ② 書誌情報の整理
- ③ 所蔵先の調査
- ④ 影印版やデジタル版の有無
- ⑤ 影印版とデジタル版のソース情報の整理

といった様々な情報を網羅的に収録し、公開できる仕組みを持ったデータベースをインターネット上に整備する計画を立てた。そして、そのデータベースを通じて中国語教材を網羅的に検索し、より容易に書籍の情報や、デジタルデータにアクセスできるようにしたいと考えた。そこで本研究では、このデータベースを整備する上で、考慮しないといけない点、またデータベースのためのデータベース構築とも言えるこのデータベースの計画がなぜ必要なのかという点について具体例を挙げながら説明し、構築したシステムの全体像と収録の方針について報告する。

1. 中国語教材データベースの必要性——『官話指南』と『語言自邇集』を例に

今回構築しようとしている「中国語教材データベース」に求められることについて、清末の中国語会話書である『官話指南』と『語言自邇集』を例に考えてみる。『官話指南』は1882年上海で印刷刊行された北京官話の会話教材であり、現在の研究者にとっては19世紀の北京官話を研究するための原資料としての価値がある。しかし『官話指南』と名付けられた書籍について調査を進めていくと、これまでの目録ではほとんど知られていなかった書籍があることが判明した。また『官話指南』には派生本が多く、外国語に翻訳されたものや、中国の方言への翻訳されたもの、清末民国初期の南方非官話地域において中国人の北京官話の学習用にされた派生本も存在することも分かり、これまでの調査から少なくとも20数種類存在することが分かった。以下の表1～5はその一覧である。

まず表1に初版本及び初版本と極めて近い関係にある『官話指南』を列挙する。『官話指南』は上海で印刷され、一部が日本で、一部が上海で流通した。いずれも初版と同じテキストを用い、形態もほぼ同じであるが、初版本とされる#1を例にとってみても、国会図書館に所蔵されているものと、内田慶市氏が所蔵しているもの³⁾では、前者には日本の出版届が、後者

3) 『中国語教本類集成』の第1集に収録されている『官話指南』は初版とされているが、これは厳密には初版ではなく、#6の系統の版本である。

には英語母語話者による学習時の書き込みが確認できるなどの違いがある。また奥付に追記された出版届の有無で中国と日本のいずれに流通したものであるかも確認することができる。さらに#1が刊行された後、上海や福州で何度か重印されており、現在少なくとも7種類の版が確認できる。

表1 『官話指南』初版の流れを汲む版本一覧

書名	刊行年	冊数	編著者	出版地・出版社・印刷	主要所蔵先
1 『官話指南』	1882(明治15)年初版	4巻1冊	呉啓太・鄭永邦編著	上海:楊龍太郎発行・上海美華書館印刷	国会図書館 内田慶市
2 『官話指南』	1886(明治19)年再印	4巻1冊	呉啓太・鄭永邦編著	上海:楊龍太郎発行・上海脩文活版館印刷	国会図書館
3 『官話指南』	1900(光緒26)年重印	4巻1冊	呉啓太・鄭永邦編著	上海:楊龍太郎発行・上海美華書院印刷	尾崎實旧蔵
4 『官話指南』	1900(光緒26)年重印	4巻1冊	呉啓太・鄭永邦編著	福州:楊龍太郎発行・福州美華書院印刷	関西大学(泊園)
5 『官話指南』	1903年重印	4巻1冊	呉啓太・鄭永邦編著	上海:楊龍太郎発行・Kelly & Walsh Limited印刷	浙江図書館
6 『官話指南』	1881(明治14)年自序	4巻1冊	呉啓太・鄭永邦編著	不明:楊龍太郎発行	関西大学(内藤)
7 『官話指南』	刊行年不明	4巻1冊	呉啓太・鄭永邦編著	上海:楊龍太郎発行	韓国学大学院

1903年になると『官話指南』は、日本で改訂版が刊行された。その後、当時の中国語検定の中級レベルの参考書として位置付けられたこともあり、翻訳本や注釈本が刊行されるようになる。

表2 『官話指南』から派生した教材

書名	刊行年	冊数	編著者	出版地・出版社・印刷	主要所蔵先
8 『官話指南』	1903(明治36)年改訂版	4巻1冊	呉啓太・鄭永邦編著・金国璞改訂	東京:文求堂	所蔵多数
9 『官話指南總譯』	1905(明治38)年	1冊	呉泰寿訳	東京:文求堂	所蔵多数
10 『官話指南精解』	1939(昭和14)年	1冊	木全徳太郎	東京:文求堂:田中慶太郎	所蔵多数
11 『官話指南談論新編 応用問題並解答』	1931(昭和6)年	1冊	幸勉	大連:大阪屋號書店	大阪大学

一方、西洋や中国では、改訂版は流通せず、重印版が引き続き使われ、初版本を翻訳した英訳版は1920年以降の刷も確認されており、長い期間使われていたことが分かっている。また早い時期からフランス語にも翻訳されフランスと上海でその後出版された『京話指南』にも影響を

与えるなど影響の広がりが見られる。また大連で飯河道雄が執筆した『譯註 聲音重念附 官話指南自習書』のシリーズについても、底本としているのは、日本で出版された改訂版ではなく、L. C.Hopkinsが翻訳した英語版である。

表3 日本国外で利用された『官話指南』

	書名	刊行年	冊数	編著者	出版地・出版社・印刷	主要所蔵先
9	『官話指南』	1893年	4巻1冊	九江書會編	九江:九江書局	神戸市 外国語大学
10	『官話指南 <i>Koan-Hoa Tche-Nan Boussole du Language Mandarin</i> 』	1887年	1冊	Le Pere Nenri Boucher SJ	上海:Impromerie de la Mission Catholique	所蔵多数
11	『官話指南 <i>The guide to Kuan hua: a translation of the Kuan hua chih nan: with an essay on tone and accent in Pekinese and a glossary of phrases</i> 』	1895年	1冊	L. C. Hopkins	上海:Kelly & Walsh, LTD (Shanghai)	所蔵多数
12	『官話指南 <i>The Guide to Kuan Hua with English Translations</i> 』	1902年	1冊		上海:Commercial Press (Shanghai)	内田慶市
13	『官話指南 <i>Koan-Hoa Tche-Nan Boussole du Language Mandarin</i> 』		1冊	Le Pere Nenri Boucher SJ	上海:Impromerie de la Mission Catholique	所蔵多数
16	『譯註 聲音重念附 官話指南自習書 應對須知篇 使令通話篇』	1924(大正13)年	1冊	飯河道雄	大連:大阪屋號書店	関西大学 (東西研)
17	『譯註 聲音重念附 官話指南自習書 官商吐属篇』	1925(大正14)年	1冊	飯河道雄	大連:大阪屋號書店	関西大学 (東西研)
18	『譯註 聲音重念附 官話指南自習書 官話問答篇』	1926(大正15)年	1冊	飯河道雄	大連:大阪屋號書店	関西大学 (東西研)

『官話指南』は中国各地の方言にも翻訳されているが、これらについては中国人が利用したというより西洋人が各地の方言を学習するために利用したという側面が大きい。

表4 日本国外で利用された『官話指南』

	書名	刊行年	冊数	編著者	出版地・出版社・印刷	主要所蔵先
20	『土話指南 <i>T'ou-Wo Tse-Ne Bussole Du Language Mandarin Traduite et Romanisee en Dialecte du Chang-Hai</i> 』	1908年	1冊	師中董訳注	土山湾慈母堂	上海図書館
21	『滬語指南』	1897(光緒23)年	2巻1冊	曹鐘橙菊人甫訳	上海:上海美華書館	上海図書館
22	『粵音指南』	1895年	2巻2冊	H.R.Wells	文裕堂	
23	『訂正粵音指南』	1930年	3冊		香港:Wing Fat & Company	所蔵多数

さらに中国南方の非官話地域である福州を中心として『官話指南』は中国人自身が当時最新の北京官話を学習するための教材として利用されており、そこから以下のようなテキストが存在することが明らかになった。ここに記載しているのは現物が確認されたものだけで、実際にはこの他にも似たような書名を持つものが出版されていたと考えられる。

表5 日本国外で利用された『官話指南』

	書名	刊行年	冊数	編著者	出版地・出版社・印刷	主要所蔵先
24	『改良民国官話指南』	刊行年不明	4巻2冊	郎秀川重訂	開民書局	内田慶市
25	『訂正官話指南』	刊行年不明	4巻2冊	郎秀川重訂	不明	香港科技大学
26	『教科適用訂正官話指南』	1918年	4巻1冊	呉啓太・鄭永邦	広州:廣州科學書局	Harvard-Yenching Library
27	『官話指南』	1916年	1冊		廈門萃經堂印務公司	氷野善寛

以上のように27種類もの『官話指南』及び関連書が存在するが、現在の図書館のオンライン型の目録で「官話指南」と検索しても容易にこれらを区別することもできなければ、検索しても表示されないものも数多くある。当然『官話指南』の全体像や、『官話指南』に影響を受けた教材についても通常のデータベースからは全く見えてこない。

また、これだけある中でデジタル化され公開されている教材は#9『官話指南總譯』、#12『官話指南 *The guide to Kuan hua: a translation of the Kuan hua chih nan: with an essay on tone and accent in Pekinese and a glossary of phrases*』、#26『教科適用 訂正 官話指南』、#7『官話指南』の4冊のみで、影印されているものも#6『官話指南』、#9『官話指南總譯』、#17『譯註 聲音重念附 官話指南自習書 官商吐属篇』、#20『土話指南 *T'ou-Wo Tse-Ne Bussole Du Language Mandarin Traduite et Romanisee en Dialecte du Chang-Hai*』の3冊のみで初版本

を含む多くの『官話指南』にアクセスするのが容易でないことが分かる。また、初版本は2冊、また南京官話の資料とされる九江書會本『官話指南』、泊園書院蔵本『官話指南』、『民国改良官話指南』は現在各1冊しか現存する書籍が確認できておらず、積極的にデジタル化を進める必要がある。

次に、『語言自邇集』も『官話指南』と同様に北京官話を学ぶために執筆され、明治初期に日本に持ち込まれたものである。1867年に初版、1886年に第2版、1903年に第3版が出版されているが、翻刻されたものや写本が日本国内では多く利用された。『語言自邇集』の初版は、本編および英語版解説書である“*Key to the TZŪ ERH CHI*”の2冊の存在が一般に広く知られており、多くの目録では2冊セットの教材とされている。しかし正式にはこれに『平仄篇』と『漢字習寫法』を加えた4冊から成るが、この点について記載している目録はあまり見られず、デジタルデータを掲載している例も見られない⁴⁾。以下に初版から3版までの各巻のデジタル化の状況と影印出版の状況をまとめる。これから2版と3版は比較的デジタル化や影印が進んでおり、アクセスしやすい環境が整っているが、初版については一部しかデジタル化されていないことが分かる。またこれまで初版本については日本国内に何点か所蔵が確認されているが、書籍のサイズが少しずつことなり、こういった情報の蓄積も必要であると考えている。

表6 『語言自邇集』のデジタル化状況⁵⁾

書名	版本	デジタル化	影印
『語言自邇集』本編	1st	H、関	
『語言自邇集』KEY	1st	関	
『語言自邇集』平仄篇	1st		
『語言自邇集』漢字習寫法	1st		
『語言自邇集』Vol.1	2nd	T、関	○
『語言自邇集』Vol.2	2nd	T、関	○
『語言自邇集』Vol.3	2nd	T、関	○
『語言自邇集』Vol.1	3rd	T、関	○
『語言自邇集』Vol.2	3rd	T、関	○

さらに『語言自邇集』の底本となったと考えられる書籍や間接的あるいは直接的に影響を受

4) 2006年に「近代漢語文献データベース」を構築した際にはテキスト情報が多く含まれる部分を中心にデジタル化を行ったためこの部分については公開していない。

5) H = Harvard University、T = Toronto University、関 = 関西大学、早 = 早稲田大学、国 = 国会図書館、AS = National Library of Australiaを示す。

けたと考えられる書籍は数多くある。それらが実際にデジタル化、あるいは影印されているかという、近年の調査⁶⁾で明らかになったものもあり、まだ全体像が明らかになっているとは言い難い。『語言自邇集』と関連のある書籍とされるものを一覧にすると以下の表ようになる。

表6 『語言自邇集』のデジタル化状況⁵⁾

書名	版本	デジタル化	影印
『清語階梯語言自邇集』			○
『亜細亜言語集』	初版		
『亜細亜言語集』	再版	早	○
『増訂亜細亜言語集』		関、国	
『総譯亜細亜言語集』		関	○
『新校 語言自邇集 散語ノ部』		国	○
『北京官話 清國語學捷徑』			○
『參訂漢語問答篇國字解』		関、国	○
『問答篇』		As	
『登瀛篇』			
『自邇集平仄篇 四聲聯珠』			○
『華語跬歩』			
『語言自邇集抜碎』	写本		
『語言自邇集』	写本		
『語言自邇集 散語 問答』	写本		
『語言自邇集卷之一 問答十章』			

以上のことからデジタル化されたデータが分散しており、正確に全体像を把握しにくい現状がお分かりいただけるかと思う。

当時の教材を利用した研究を行うためには、正確な資料の把握と、なぜその資料を利用するのかという根拠が必要である。しかし現状として、中国語教材の目録整理やデジタル化はまだまだ発展途上であり、継続したデータの構築が必要である。

そこでインターネットというツールを使った新しい形の教材目録を構築し、紙に印刷された書籍とデジタル化された書籍とを結びつける必要があると考えた。目録についてはこれまで多くの研究者によって六角恒廣氏の『中国語関係書書目』が利用されているが、本書を含めて、これまで出版された目録類を再度調査し、整理する必要がある。

6) 氷野歩 2012 博士論文『『語言自邇集』の総合的研究』

そこでこの状況を整理し新たな教材目録を整理するために、教材に関する情報を再度整理すると共に、並行してインターネット上に中国語教材を専門に扱ったデータベースの構築を進めた。

2. 中国語教材データベースの概要

次に構築中のデータベースの概要を述べる。今回設計、構築している中国語教材に関するデータベース（以下「中国語教材データベース」）については『清代民国漢語文献目録』（遠藤光暁・竹越孝主編、2011）で筆者が担当した「日本的漢語教材（明治至昭和初期）」と2013年にまとめた「東西学術研究所文献目録」をサンプルデータとして用いて初期構築を行った。最終的には、教材が出版された時期や地域について特に限定せず、幅広く集めるつもりであるが、ひとはまず日本国内、あるいは日本ときわめて密接な関係にあった教材から目録整理を行っている。

次に、中国語教材データベースのシステムを構築する上で考えている基本的な設計理念は次の通りである。

- ①インターネット上で公開されるリレーショナル型のデータベースである。
 - ②閲覧利用にはアクセス制限など、一切の制限をかけない
 - ③インターネット上での共同編集を想定しており、項目の追加や情報の書誌情報の編集が可能である。（要ログイン）
 - ④PC、タブレット、スマートフォンなど複数のデバイスによる閲覧が可能で主要ブラウザに対応する。
 - ⑤位置情報、時間軸、キーワードなど様々な情報からの検索が可能である。
 - ⑥オープンソースを利用してプログラムを開発しているため、作成したソースコードについてもオープンソースとして公開する。
- 7書籍の所蔵情報やデジタル化された画像を表示するようにする。

以下に掲げる図1と2が中国語教材データベースの画面構成である。

「中国語教材データベース」では、中国語教材に特化して作成しており、上記で述べたとおり、サンプルとして日本と関連のある中国語教材を集中的にデータベース化している。特に現在は1870年から1950年頃までに刊行された中国語教材を対象としているが、最終的には、日本人が執筆した教材だけでなく、中国、西洋などあらゆる地域の教材の収録を進める方針である。書誌情報については、基本的な情報から、教材の分類、教材が扱う単語データを登録でき



図1 中国語教材データベースのトップ画面

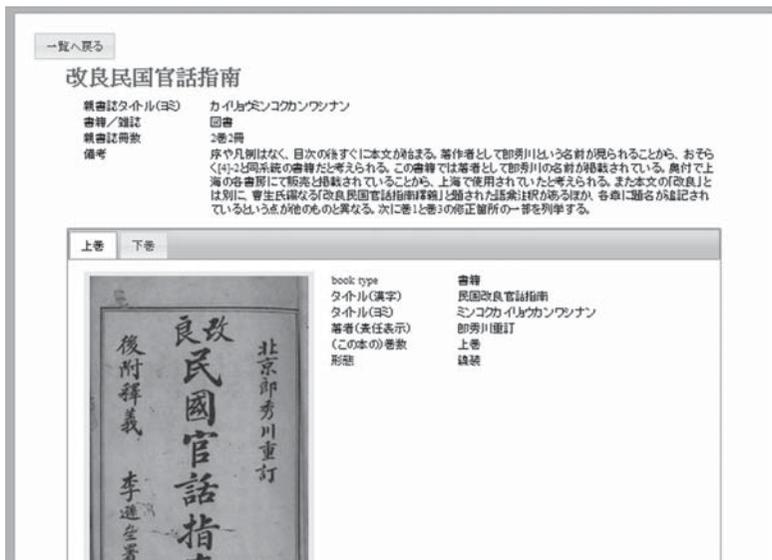


図2 教材データベースの詳細画面

ようになってきているがこの点については後述する。また収録対象としては文献だけではなく将来的には著作権の保護期間が切れた音声や画像データの収集も可能であれば行いたい。



図3 検索画面

5. 中国語教材データベースの設計について

まずデータ構造について説明する。このデータベースでは書籍単体の情報から各書籍が実際にどこに所蔵されているのか、デジタル化されているかといった情報を蓄積することが可能である。

教材データベースの構造は「論理セット」、「親論理冊子」、「子論理冊子」、「物理冊子」の4つのデータセットからなる。たとえば仮に〇〇シリーズに収録されている上下2巻の2冊からなるA本、同じく〇〇シリーズに収録された1巻1冊のC本、さらにいずれの叢書にも収録されていない1巻1冊からなるB本があったとする。この場合、親論理冊子となるのが〇〇シリーズのA本とB本、さらに1冊本のC本の3データとなる、さらに〇〇シリーズのA本は2冊本のため論理的には2つの子論理冊子が存在する。それ以外の書籍については「子論理冊子」は各1データと考える。ここではA本は上下巻が存在するのは確かだが実際には下巻しか所蔵がない、あるいは佚書となり存在しない、ということがある可能性を想定している。論理的には存在するが物理的には存在しないかもしれない本として「論理冊子」という名称を説明上用いている。次にそれぞれの「子論理冊子」に対して1つあるいは複数の「物理冊子」の情報がつながることがある。これは具体的にはある冊子対する所蔵情報を示し、実際にある書籍個別の情報となる。たとえば〇〇シリーズのA本の上巻は□□図書館と××図書館に1冊ずつ所蔵されているとする場合2つの「物理冊子」の情報が1つの同じ「子論理冊子」にぶら下がる。そして□□図書館が所蔵する書籍は第何刷の書籍で、誰かによる書込みがある、あるいはこういった蔵書印があるといった情報がここに記録される、××図書館に所蔵されている書籍はま

た別の書込みがあるといった実際の書籍に付随する個別の情報が記載されることになる。この「物理冊子」については仮にデジタル化されているデータがある場合は、Viewerボタンが表示され、データが掲載されているリンク先にジャンプし表示する。

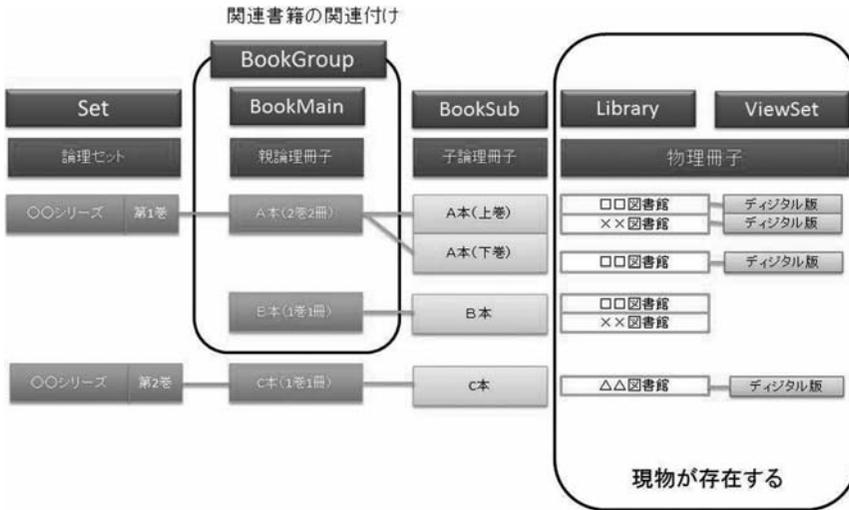


図4 「教材データベース」のデータ構造

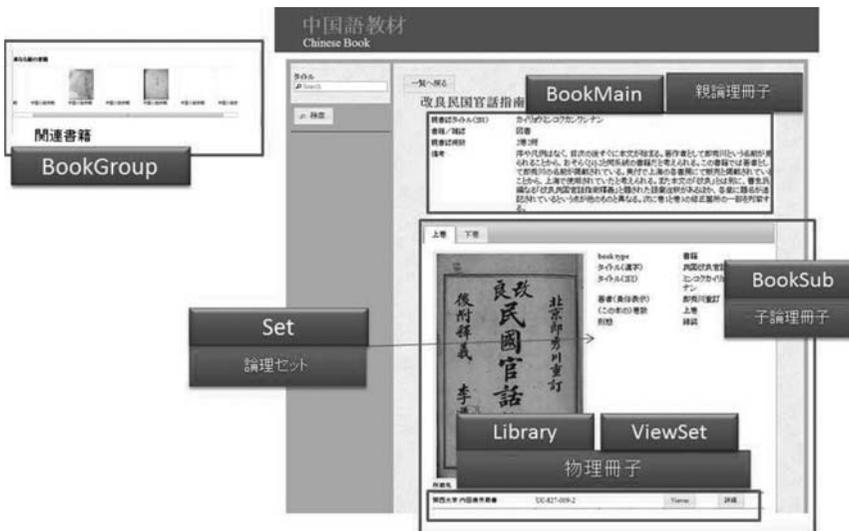


図5 「教材データベース」の表示画面とデータ構造図との関係

上記の構造設計に基づき、データベースに収録する項目を次のように整理した。

表8 BookMain (親論理冊子)

フィールド名	名称	データ型	フィールド名	名称	データ型
id	親書誌識別子	int(7)	mType	書籍／雑誌	tinyint(4)
group1	版次の異なるグループID	int(5)	edition	親書誌版次	archar(50)
group2	関連書籍のグループID	int(5)	sets	親書誌冊数	varchar(255)
mTitle	親書誌タイトル(代表名称)	varchar(255)	memo	備考	text
mRead	親書誌タイトル(ヨミ)	varchar(255)	research	研究文献	varchar(255)

表9 BookSub (子論理冊子)

フィールド名	名称	データ型	フィールド名	名称	データ型
id	id	int(7)	pubYear	刊行年	varchar(4)
bType	資料種別	tinyint(3)	volume	巻数	varchar(100)
bType2	資料分類	varchar(255)	partTitle	部分タイトル(巻次、部編番号)	varchar(255)
main	親書誌識別子	int(7)	pages	ページ数	varchar(100)
bTitle	タイトル(漢字)	varchar(255)	size	サイズ	varchar(50)
bRead	タイトル(ヨミ)	varchar(255)	attach	その他付属物	varchar(255)
bTitleEn	タイトル(欧文)	text	bind	形態	varchar(100)
bAlias	タイトル(別名)	varchar(255)	pubType	出版種別	varchar(10)
bSeries	叢書・シリーズ名1	varchar(255)	price	販売価格	varchar(255)
bSeriesR	叢書・シリーズ名1(ヨミ)	varchar(255)	lang	言語	varchar(255)
bSeriesEn	叢書・シリーズ名1(欧文)	varchar(255)	mokuji	目次	varchar(100)
bSeriesN	シリーズ巻数1	varchar(255)	ndc	NDC分類	varchar(30)
bSeries2	叢書・シリーズ名2	varchar(255)	ndlc	NDLC分類	varchar(255)
bSeriesR2	叢書・シリーズ名2(ヨミ)	varchar(255)	nsmarc	NSMARC番号	varchar(30)
bSeriesEn2	叢書・シリーズ名2(欧文)	varchar(255)	trcmarc	TRCMARC番号	varchar(30)
bSeriesN2	シリーズ巻数2	varchar(100)	nbn	全国書誌番号(JP番号)	varchar(30)
tabNom	タブ表示用	varchar(10)	lccn	LCカード番号	varchar(30)
author	著者(表示)	varchar(255)	isbn10	国際標準図書番号	varchar(20)
author2	著者(標目)	varchar(255)	isbn13	国際標準図書番号	varchar(20)
place	出版地	varchar(255)	issn	国際標準逐次刊行物番号	varchar(30)
pubPeople	出版社等(表示)	varchar(255)	kenri	著作権	tinyint(3)
pubFirm	出版社等(標目)	varchar(255)	note	備考	text
pubYjp	刊行年(和暦)	varchar(255)	contents	サムネイル画像	varchar(255)
pubYch	刊行年(中国暦)	varchar(255)	label	ラベル	varchar(255)
pubYen	刊行年(西暦)	varchar(255)	tag	タグ	varchar(255)

表10 Book Group (書籍グループの管理)

フィールド名	名称	データ型
id	親書誌識別子	int(5)
gType	グループタイプ	tinyint(3)
gTitle	グループ名称	varchar(255)
memo	備考	text
label	ラベル	varchar(255)
tag	タグ	varchar(255)

表11 c_Library (所蔵情報の管理)

フィールド名	名称	データ型	フィールド名	名称	データ型
id	id	int(7)	suri	刷	varchar(50)
dbid	データベースID	tinyint(4)	hako	個別の情報	text
sub	SubId	int(7)	digital	デジタル版	tinyint(3)
pMark	所蔵機関	int(9)	site	公開サイト	varchar(255)
kigou	請求記号	varchar(50)	url	デジタル化パス	varchar(255)
shiryo	資料ID	varchar(50)	note	備考	text

次にデータベースの特徴的な機能についていくつかとりあげ、解説する。

① デジタル化された画像と表紙画像の収集

著作権や所蔵権などの処理がまだ未処理で全てを公開できない場合は極力サムネイル画像として表紙画像を掲載するようにしている。

② 所蔵情報の表示

物理冊子がある場合は、その所蔵先を示し個別に特徴がある場合は特記事項として情報を入力する。

③ 研究文献の整理

『清代民国漢語文獻目録』（遠藤光暁・竹越孝主編、2011）では、該当する書籍の研究文献の有無についても調査を行った。そこで把握できている範囲のみ、原資料と研究成果のデータを関連付けて見られるように設計した。以下は研究論文に関して収集するデータ項目である。

表12 Research (研究文献)

フィールド名	名称	データ型	フィールド名	名称	データ型
id	id	int(9)	pageb	終了ページ	int(11)
type	タイプ	varchar(50)	pdate	出版年月	date
title	タイトル(日本語)	varchar(255)	publshj	発行者/出版社(日本語)	varchar(255)
titleen	タイトル(英語)	varchar(255)	publshe	発行者/出版社(英語)	varchar(255)
authorj	著者(日本語)	varchar(255)	pdf	PDF	varchar(255)
authore	著者(英語)	varchar(255)	url1	HPリンク1	varchar(255)
booksj	刊行物名(日本語)	varchar(255)	url2	HPリンク2	varchar(255)
bookse	刊行物名(英語)	varchar(255)	url3	HPリンク3	varchar(255)
kan	巻	varchar(50)	lang	記述言語	varchar(255)
no	号	varchar(50)	note	備考	text
pagea	開始ページ	int(11)			

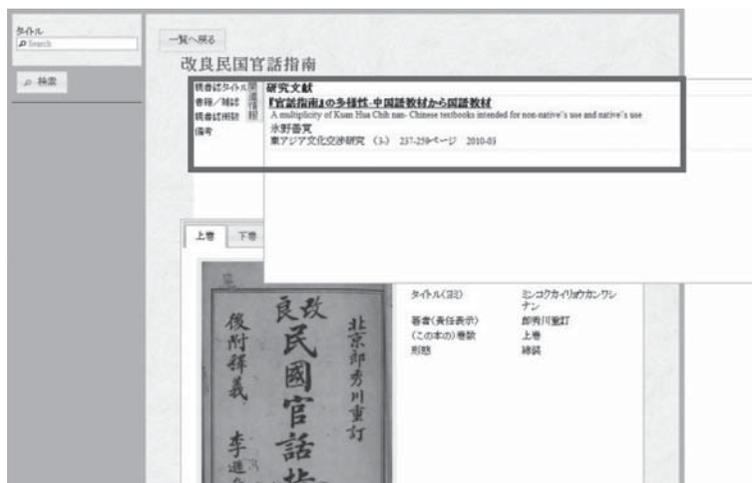


図6 研究文献の表示

④ 「Book Referデータベース」による関連書籍のグループ化

当時の中国語教材の全貌を明らかにするために、書籍の巻末に掲載されている広告情報を収集整理し、どの本にどの本の広告が掲載されているのかということデータベース化している。同時に『申報』など民国期に上海で刊行された新聞や日本で刊行された新聞に教材に関する広告がないかといった点についても情報収集をしており、これらの結果から当時のある教材に対する認知度がどれだけあったのかという調査をできればと考えている。この点についてはたとえば『官話指南』について調べてみると1882年8月1日の3版には

昨由西友送來官話指南一書披讀一過知係日本人吳君啓太鄭君永邦所合輯書分四類曰應對須知曰官商吐屬曰使令通話曰官話問答雖皆質直無文而口吻畢肖手此一編循誦宛如身入春明城與北人晤對名曰官話指南洵非虛也學者苟晨夕揣摩再得一京師人口授何患無成讀竟爰附識數語以多吳鄭二君之勤敏云(『申報』3322号3版、書官話指南後)

とあり、刊行されたばかりの1882年に、日本人である吳啓太と鄭永邦が編纂した書籍が友人から送られてきたことが記事になっている。また同じく『申報』では1888年7月10日から8月7日まで広告掲載されていたことが分かっている。

⑤ “Lang LINK” データベースとの関連付け

この機能については現段階では試作的なものである。Lang LINKには該当書籍の全文及び学

習対象となっている単語を登録することができる。こうしたデータを蓄積することにより、ある単語がどの教材で使われているのかといった情報の把握が可能となる。

⑥ 検索機能について

検索については一般的な書名や著者、刊行年からの検索を行える以外に様々な検索ができるように設計している。また検索と同時に、その中に含まれる情報をさらに再検索をかけ、画面

The screenshot shows a search results page for '中国語教材' (Chinese Book). The search criteria include 'タイトル' (Title) as '官話', '著者' (Author) as '官話', '著者国籍' (Author's Nationality) as '官話', and '出版年(中国暦)' (Publication Year in Chinese Calendar) as '官話'. The results list various editions of '官話指南' (Kuan Hua Guide) and related materials, such as '訂正官話指南' (Revised Kuan Hua Guide) and '官話指南精解' (Kuan Hua Guide Detailed Explanation). The results are organized into sections like 'デジタル化' (Digitalization), '所蔵先' (Collection locations), and '刊行年' (Publication Year). The page also includes a sidebar with search filters and a '検索' (Search) button.

図7 検索結果画面

右に、検索結果を対処とした二次検索を行った処理が表示されるようにした。たとえば図7の例では、「官話」と検索しているが、「官話」という文字列を含む書籍が中央に表示され、右にその中でデジタル化されている書籍が何件あるか、あるいはどこに所蔵されているかといった情報を一覧表示し、情報をさらに絞り込むことが可能となっている。

⑦デジタル画像の表示について

書籍閲覧用に作成したビューワーには、ビューワーデータベースに登録された画像や情報のうち、関西大学と中国語教材データベース自体が保有するデータについては、専用ビューワーで表示される。このビューワー本体はInternet ArchivesのプロジェクトであるOpen Libraryが開発し、オープンソースとして公開しているプログラムをベースに改造して構築したものである。自身が保有するデータ以外についても公開の許可を個別に取得したものや、各々のデータベースへの接続許可が得られたものについて直接データにアクセスできるようになっている。

6. おわりに

以上が中国語教材に特化したデータベースに関する簡単な解説である。今後、このデータベースで整理された情報により、より容易に正確な情報が付されたデジタル資料にアクセスできるようになることを願う。またデータベースやデジタルアーカイヴズをまとめあげるためのデータベースとして、このデータベースがどのように発展し、利用されていくか引き続き検証を及び構築作業を進めていきたい。

(謝辞) データベース構築に当たって、東西学術研究所所蔵の教材資料の広告データの入力では北京外国語大学院生の斎燦氏に、データベースの構築では北京外国語大学交換留学生の馬翼飛氏にご協力をいただきました。心より感謝いたします。

※本論文は2013年9月「“汉语国际传播历史”国际学术研讨会暨世界汉语教育史研究会第五届年会」で報告した内容を日本語に翻訳したのち、加筆修正を加えたものである。また平成25年度、文部科学省科学研究費若手研究(B)の研究成果の一部である。