

『切韻』諸本テキスト一覧システムの構築について

鈴木 慎吾（大阪大学 言語文化研究科）

本稿は筆者が進める漢語音韻学の基本資料『切韻』諸本のデータベース化と公開システムについてその概要を報告する。このデータベースは句点を施した『切韻』諸本の全テキストをを階層構造によって一覧表示する。これにより『切韻』諸本の増補の状況を簡単に知ることができ、漢語中古音研究の基礎資料としての活用が可能である。

The construction of the database displaying the full text of the *Qieyun*

Suzuki Shingo (Graduate School of Language and Culture, Osaka University)

This paper presents an overview of the construction of the database which displays the text of the ancient Chinese rhyme dictionary, the *Qieyun*, and describes the web browsing system of the database. This database includes the fully punctuated text of the entire *Qieyun* manuscript in a hierarchical structure. Through this, the revision process of the *Qieyun* can be revealed, and as such, the database can provide fundamental support to research on Middle Chinese phonology.

1. まえがき

本稿は筆者が進める漢語音韻学の基本資料『切韻』諸本のデータベース化と公開システムについて、現状と課題を報告する。前半は『切韻』という書物について、後半はデータベース化と公開システムについて述べる。

2. 漢語音韻学について

漢語音韻学とは中国語の音声の歴史を探求する学問である。ただし、中国語は歴史的に、ほぼ一貫してもっぱら漢字のみによって表記されてきたことから、中国語の歴史的音声の研究とは、少なくとも現段階においてはほぼ漢字音の研究と言うに等しい。その漢字音の、共時的な多様性、また通時的な変遷がこの学問の主たる研究対象である。なお、漢語音韻学の知識は日本漢字音の研究においても必須であるため、漢語音韻学はわれわれ日本人にとっても重要な意味をもつ学問である。

漢語音韻学の対象は古代の音声であるから、その資料の第一は文献資料ということになる。その中心は韻書および韻図と呼ばれる書物である。このうち、韻書とは字音を基準に文字を配列した一種の字書である。これを分析することにより、その韻書が基礎としている音韻体系の全体像を知ることができる。

3. 『切韻』について

漢語音韻学で用いる韻書のうち、最も重要なものは宋・陳彭年等撰の『大宋重修廣韻（略して『広韻』）』（1008年）である。これは、本書が完

全な刊本の形で現存する最も古い韻書であるため、この『広韻』に反映する音韻体系は一般に中古音（Middle Chinese）と称される。中古音は、さらに古い時代、とくに先秦を中心とした上古音（Old Chinese）研究の起点であり、逆に宋～清代の音韻体系である近世音（Old Mandarin）や、また現代方言音も中古音を基礎に研究が行われるため、中古音はまさに漢語音韻学、方言学の枢要の位置にあると言える。

また、日本漢字音の主要な部分は古代に中古音が日本に伝わったものであるから、『広韻』は日本漢字音の研究においても中心的資料として扱われている。

ところで、この『広韻』は、そもそも隋唐代に通行していた『切韻』（601年）を底本に編まれたものである。『切韻』は隋・陸法言の撰で、唐代に幾度も改訂増補されたことが知られているが、残念ながら完全な形では残っていない。それでも、数多くの残巻が現代に伝わっており、それにより隋唐代における韻書の有様を知ることができ、何より『切韻』残巻を比較調査することで、『広韻』が内包する、原本『切韻』以降の増加部分を区別することができるので、中古音をより精密に研究する上で絶大な価値を有する。

しかしながら、これら『切韻』の残巻は、その大部分が20世紀初めに甘粛・新疆地域から出土した、いわゆる「敦煌・吐魯番文献」であり、他の大量の文献とともに世界各地の収蔵機関（フランス国立図書館、大英図書館、ベルリンブランデブルグ人文科学アカデミー、龍谷大学、ロシア科学アカデミーサンクトペテルブルク支局東洋学研究所、北京故宫博物院など）に分散している。

そのため、これらを研究の材料として用いるには相当な不便があった。

4. 先行研究

上述の事情により、『切韻』の研究は、まず世界中に散らばる諸本の模写・影印を収集するところから始められた。そのような研究で比較的大規模なものには劉復等『十韻彙編』[1]、姜亮夫『瀛涯敦煌韻輯』[2]、龍宇純「英倫藏敦煌切韻殘卷校記」[3]、潘重規『瀛涯敦煌韻輯新編』[4]、周祖謨『唐五代韻書集存』[5] などがある。また上田正『切韻殘卷諸本補正』[6] は、それまでの先行研究を集成し、誤りを正した研究である。近年、敦煌吐魯番文献の影印出版が進んだことを受けて、筆者は上記の研究で未収録の殘卷を収集し、内容の検討を行い、これによって現時点で可能な『切韻』殘卷の収集・整理作業はひとまず完了したと言ってよい[7, 8, 9]。

現在までに知られている『切韻』諸本には以下のものがある（番号は所蔵機関における資料番号）。

- ・フランス国立図書館所蔵
P.2017, P.2018, P.2019, P.2129, P.2638, P.2659, P.3693, P.3694, P.3695, P.3696, P.3798, P.3799, P.4746, P.4747, P.4871, P.4879, P.4917, P.5531
- ・大英図書館所蔵
S.2055 (切二), S.2071 (切三), S.2683 (切一), S.5980, S.6012, S.6013, S.6156, S.6176, S.6187, S.10720, S.11380, S.11383
- ・ベルリンブランデブルグ人文科学アカデミー所蔵
TID1055, THD1, Ch1072, Ch1106, Ch1150, Ch5555, TIV70+71, TIVK75, TIVK75-100a, TIVK75-100b, Ch79, Ch3605, Ch1577, Ch2917+343+323, Ch2492
- ・龍谷大学所蔵
大谷 3327+5465, 5395, 5397, 5401, 8107
- ・ロシア科学アカデミーサンクトペテルブルク支局東洋学研究所所蔵
Dx1372+3703, Dx3109+1267, Dx1466, Dx5596
- ・北京故宮博物院所蔵
宋跋本刊謬補缺切韻 (王三), 裴務齊正字本切韻 (王二)
- ・吳縣蔣斧氏旧蔵
唐韻殘卷

5. 『切韻』の構造

『切韻』のデータベース化について述べる前に、『切韻』自体の構造について説明しておく。

『切韻』を含む韻書という書物の構造を理解するには、さしあたり中国語の音節構造について知っておく必要がある。

中国語の音節は「声母」と呼ばれる頭子音と、それに続く「韻母」、また音節全体に関わる「声調」の三要素に分析される。たとえば「東」という字は中古音で [tuŋ^平] のように読まれたと推定されるが[10]、それは以下のような要素に分解される（「平」というのは声調の分類名である。当時の声調が具体的にどのようなトーンパターンを有していたかは不明であるため、声調は分類名で示すのが通例となっている）。

「東」 [tuŋ^平]
→ t (声母) + uŋ (韻母) + 平 (声調)

さて、『切韻』は全体が「声調」によって五巻に分けられ、各巻の内部はさらに「韻」（いくつかの韻母をまとめた組）によって分けられる。「韻」の代表字は「韻目」と呼ばれる。これら「韻」と「韻目」は、それぞれ部首によって分類された字書でいうところの「部」と「部首」に相当する。そして、各々の「韻」の内部はさらに「小韻」と呼ばれる同音字の組に分けられる。すなわち、『切韻』の中身は全体として以下のような階層構造からなっている。

[声調——韻——小韻——字]

たとえば「同」という文字は第一巻平声・東韻・同小韻の第1字目に収録されている、という具合である。

各小韻の第1字目（小韻代表字と呼ばれる）の注にはその小韻の字音（反切と呼ばれる表記法による）と小韻所属字数が記される。各字の注には意味を記した義注のほか、異体字注、また多音字の場合はその音注が記される。たとえば下記の「中」字の例において、「陟隆反」は [tʃuŋ^平] という音を示す反切表記、「景正」は義注、「又陟仲反」は別音（又音と呼ばれる） [tʃuŋ^平] を示す反切、「四」はこの小韻に収録されている字数を意味する。また「豊」字の注に見える「正作豊」は見出し字とは別の字体を説明している。なお、注におけるこれらの項目の出現順序は諸本によって違いがある（下の例はともに王三）。

中 陟隆反。景正。又陟仲反。四。

豊 敷隆反。多。正作豊。六。

6. 一覧システムの構築

近年、様々な字書が次々に電子化され、公開されている。それらがもたらす有益性についてははや贅言を要さない。しかし、『切韻』についてはこれまでそのテキストの全文が電子的に整理、公開されたものは存在せず、漢字音研究の基本インフラとして強く求められているところである。上述のように、必要な文献調査はすでに完了しており、電子化の機は熟している。ただし、単にレコードを表示させるだけのデータベースではあまり面白くない。今回構築しようというのは、音韻資料に適したデータベース化、言うなればデジタル技術による『十韻彙編』の増補版である。

今回の『切韻』のデータは、Excel を用いて基礎データを入力し、サーバに設置した MySQL にこれを格納し、PHP で検索・表示システムを構築している。

今回、『切韻』一覧システムをデザインするに当たっては、特に以下の点に重点を置いた。

- 1) 韻書の階層構造を分かりやすく表現する。
- 2) 『切韻』諸本のテキストを相互に比較できるようにする。
- 3) 原資料の字体をできる限り忠実に反映する。
- 4) 異体字による検索も可能とする。
- 5) 原資料の誤りに対する校正情報を表示する。
- 6) 注を内容別にマークアップする。

これらは、構築するシステムが単なる検索用途だけではなく、利用者が『切韻』や中古音そのものの仕組みを理解しやすいようにすることを意図したものである。以下、詳しく説明する。

6. 1 階層構造の表現

部首引きの字書が「部首」>「画数」という、字形から視覚的に理解可能な、比較的分かりやすい階層構造になっているのに対し、韻書の階層構造は声母、韻母、声調といった、音節の構成要素によっているため、中国語（しかも古代語）の音節構造にある程度慣れた者でないともそもも理解しにくいものとなっている。そこで本システムでは「声調」>「韻目」>「小韻」という階層構造が直感的に理解できるような表示方法を採用することにした。具体的には、タブの多段表示を使うことによってこれを表現している（図1上部）。各段の項目は暫時『広韻』の順序によって並べているが、本システムは『切韻』諸本の一覧であり、最終的には推定される原本『切韻』のものに変更する予定である。

6. 2 『切韻』諸本の一覧表示

上述のように『切韻』には非常に多くのテキストが伝わっており、これらを比較することによりテキスト増補の具体的なありさまや、写本間の関係を見ることが可能である。これは他の古文獻ではあまり見られない点で、文献学的にも極めて興

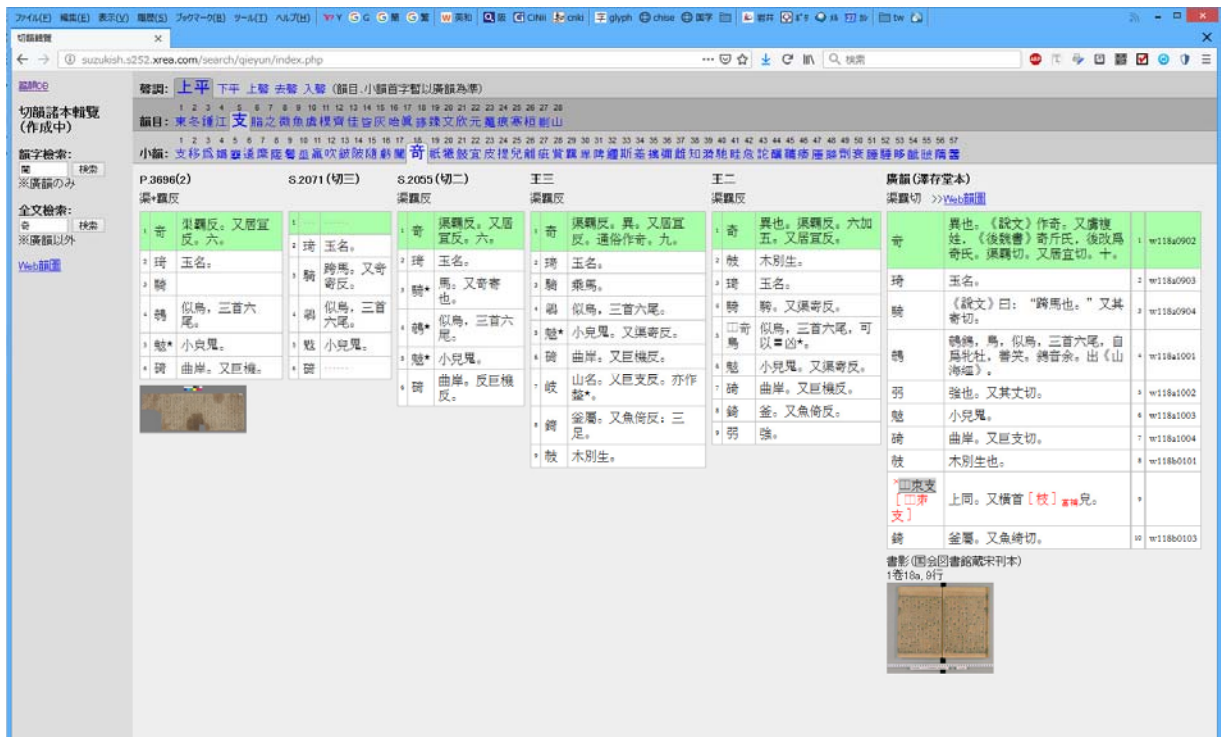


図1 表示画面

味深いものである。しかるに、『切韻』諸本は小断片も少なくなく、また各地に散らばっているためにそれらを相互比較することは大変面倒であった。そこで本システムでは小韻毎に諸本の内容を比較できるような表示方法を採用することにした(図1下部)。このような表示により、各小韻における所収字および注文の増補の状況を簡単に見て取ることができる(将来的には、小韻レベルの増補状況を表示することも考えている)。

また、『切韻』諸本の祖本である陸法言本はすでに存在しないが、各種テキストを比較することによってかなりの精度で復元が可能である。本システムはその復元作業の助けとなろう。

一つ例を挙げてみよう。たとえば、支韻「奇」字の諸本の状況は以下の通りである(図1から抜粋)。

- P.3696(2) 奇 渠羈反。又居宜反。六。
- S.2055 奇 渠羈反。又居宜反。六。
- 王三 奇 渠羈反 異 又居宜反 通俗作奇。九。
- 王二 奇 異也 渠羈反 六加五 又居宜反。
- 広韻 奇 異也。《説文》作奇。又虜複姓，《後魏書》奇斤氏，後改爲奇氏。渠羈切。又居宜切。十。

これらを比較すれば、原本『切韻』の内容は「奇、渠羈反。又居宜反。六。」であったことが分かる(「奇」の字体にも注意;次節も参照)。もっとも、原本の推定は単なる比較だけではなく、系統樹を考える必要があるが、ともかく本システムによって推定はずいぶん容易なものとなる。

6. 3 字体の表現

『切韻』諸本は当時の漢字字体の書写状況を知

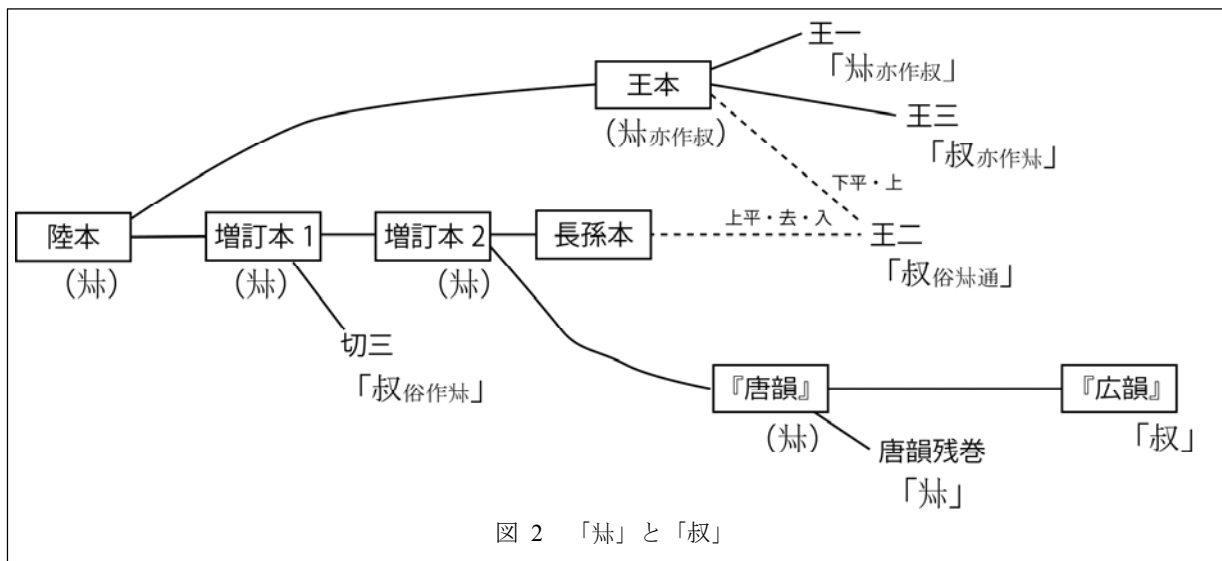


図2 「𪛗」と「叔」

る上でも貴重な資料である。とくに、前節で述べたように、『切韻』は数多くの写本が残り、またそれら写本間の関係もある程度分かっているため、唐代における字体規範の変遷を知る手掛かりとしても大変貴重である(図2は古い写本で「𪛗」と推定され、新しい写本で「叔」に置き換わっていく状況を示している)[11, 12]。このような用途に供するため、今回のデータベースでは原資料の字体をできる限り忠実に反映することにした。

ところで、刊本であり、その後もよく参照される『広韻』の収録字については、Unicode CJK 統合漢字・拡張 B までの範囲ではほぼ全ての字について入力が可能である。しかし『切韻』はその多くが写本であり、データ化の困難な字が数多く含まれる。もっとも、現在ではCJK 統合漢字・拡張 A~F が利用でき、以前に比べて状況はかなり改善した。

問題は今も Unicode に未収録の字である。これらは今回のデータベースにおいて、GlyphWiki [13] に登録されているならばそれを利用し、未登録のものは随時新規に登録している。

下は GlyphWiki の「屬」(id: zihai-066837) という字体を埋め込んだデータの例である。GlyphWiki の字体 id を glyph タグの id 属性として記述している。

```
<glyph id="zihai-066837">屬*-</glyph>。
```

データには一般的な字体である「屬」も入力されているので、「屬」による検索ももちろん可能である。

ブラウザで表示させる場合にはこれを PHP で次のように書き換えて出力している。

```
屬*</glyph>" class="glyph">
```

こうすることで画面上には GlyphWiki が公開している svg 画像を読み込んで表示している。図3はその表示例である。通常のテキストと混在させてもさほど違和感なく表示できている。



図3 「屬」の画面表示

img タグ出力時には alt 属性に元のデータを書き出すようにしてある。そのため、ブラウザでコピーすれば元のデータが得られるので、異体字の情報は失われることなく再利用が可能である。

6. 4 異体字による検索

本システムは Unihan Database で公開されている Unihan_Variants.txt のデータ[14]を加工した独自の異体字テーブルを搭載しており、検索は異体字であってもヒットする。たとえば入力「創」の場合、自動的に

創|創|創|創|創|創|創

の7字（簡体字を含む）による正規表現が生成され、それにより検索が実行される。なお、異体字テーブルは本システムの目的に合うよう随時更新を行っている。

ところで、本システムでは①見出し字、②全文、の二種類の検索を行うことができる。図4は「属」で全文検索を行った画面だが、入力「属」に対して「属|屬」の他、GlyphWiki の字体「屬」もヒットする（図4）。検索対象の文字はハイライトで強調表示しているが、GlyphWiki の svg 画像も背景色を変えることで同じようにハイライトさせているので他の字と見た目は変わらない。

6. 5 校正情報の表示

『切韻』諸本は大部分が写本であるために、書写時の誤りが非常に多い。これらに対し適切な注記を行わなければ、原本の誤りなのか、データ化の際の誤りなのかを区別できず、結果的にデータの信頼性を著しく損ねてしまうことになる。そこで今回は、原本の誤写には逐一注記を行った。

『切韻』諸本のテキストを校正する研究には相当な量の蓄積がある。主なものには周祖謨『廣韻



図4 「属」で全文検索

校勘記」[15]、龍宇純『唐寫全本王仁昉刊謬補缺切韻校箋』[16]、張涌泉主編『敦煌經部文獻合集』[17]などのほか、個別テキストに関するものがある。今回のデータ化に当たっては、それらの校正をできる限り反映させることにした。

誤りの種類には、衍字、脱字、誤字、誤倒などがある。それらに対して、以下のようなマークアップを行った。これらのタグの定義は TEI ガイドライン P5 に合わせている[18]。resp 属性には校正者の略称が入る。

衍字：<del resp="zz">A

脱字：<add resp="zz">B</add>

誤字：<corr sic="A" resp="zz">B</corr>（A が誤字。B に正す。）

誤倒：<corr sic="AB" type="rvs" resp="zz">BA</corr>（AB が誤倒。BA に正す。）

6. 6 注文のマークアップ

第5節で述べたように、『切韻』収録字の注の内容は義注、異体字注、反切、又音反切、小韻所属字数からなっている。これらに対して適切なマ

ークアップを行うことで、データのさらなる活用が期待できる。今回設定しているタグは以下の通りである（ただし、<f>以外はまだマークアップが完了していない）。

<var> 異体字注
type 「正、俗、又」など

<f> 反切

<yf> 又音反切
target 指し示す先の小韻を id で記述

以下はマークアップの例。

中 <f>陟隆反</f>。景正。<yf target="w3012 1">又陟仲反</yf>。四。

豊 <f>敷隆反</f>。多。正作<var type="正">豊</var>。六。

異体字注の var タグは、type 属性に「正、俗、又」などのラベルを記述する。こうしておくことで、異体字注を種類別に取り出すような使い方ができる。また又音注に target 属性を付与しておくことで、指し示す小韻にハイパーリンクを張るような利用法が考えられる。反切は上字（声母を指示する）と下字（韻母と声調を指示する）を区別してマークアップすることも考えられるが、データが繁雑になるので今はまだ着手していない。

7. 将来構想

韻書には、当然のことながらその背景としての音韻体系が存在する。

本システムは資料としての『切韻』を検索、表示するものだが、基礎となっている音韻体系の表示についてはすでに「Web 韻図」（<http://suzukis.h.s252.xrea.com/search/inkyō/index.php>）というサイトを構築済である。これら2つは合わせて中古音の「資料」と「音韻体系」の対をなしている。今後はさらにこれを上古音や近世音に展開することを計画している。そこでは中古音と同様に、「資料（上古音なら押韻資料と諧声字、近世音なら韻書）」と「音韻体系（韻図・諸家推定音）」が対となり、全体として中古音を軸とした歴史音韻総合データベースが構築される予定である（下表）。現在完成しているのは【 】で示した部分。『切韻』については、諸書に引用されているテキストも重要で、これは「切韻佚文データベース」としてすでに完成済である。

次は上古音のデータ化を予定している。少なくとも上古音と中古音の二つがそろって初めて

通時的な音韻変化を扱う工程に入ることができる。

	資料	音韻体系
上古音	押韻・諧声字	韻図・諸家推定音
中古音	【切韻諸本 ・切韻佚文】	【Web 韻図】 ・諸家推定音
近世音	韻書	韻図・諸家推定音

8. 謝辞

本研究は科学研究費基盤研究（C）16K02671「『切韻』系韻書総合データベースの構築」の助成により実施したものである。

参考文献

- [1] 劉復等：十韻彙編，3冊，國立北京大學研究院文史部，北京（1936）。
- [2] 姜亮夫：瀛涯敦煌韻輯，24卷，上海出版公司，上海（1955）。
- [3] 龍宇純：英倫藏敦煌切韻殘卷校記，慶祝董作賓先生六十五歲論文集（歷史語言研究所『集刊』外編第4種下冊），pp.803-825（1961）。
- [4] 潘重規：瀛涯敦煌韻輯新編，新亞研究所，香港（1972）。
- [5] 周祖謨：唐五代韻書集存，2冊，中華書局，北京（1983）。
- [6] 上田正：切韻殘卷諸本補正（東洋学文献センター叢刊第19輯），東京大学東洋文化研究所附属東洋学文献センター，東京（1973）。
- [7] 鈴木慎吾：『切韻殘卷諸本補正』未収の切韻殘卷諸本について，開篇，Vol.23，pp.76-90（2004）。
- [8] 鈴木慎吾：『切韻殘卷諸本補正』未収の切韻殘卷諸本—ベルリン本補遺，開篇，Vol.28，pp.125-136（2009）。
- [9] 鈴木慎吾：『切韻殘卷諸本補正』未収の切韻殘卷諸本—大谷本補遺，開篇，Vol.29，pp.22-25（2010）。
- [10] Karlgren, B.: Compendium of phonetics in ancient and archaic Chinese, Elanders Boktryckeri Aktiebolag (1954).
- [11] 鈴木慎吾：王仁昫切韻の異体字注記について，開篇，Vol.24，pp.14-36（2005）。
- [12] 鈴木慎吾：『切韻』諸本による陸本の復元について，漢デジ2014—デジタル翻刻の未来，京都大学人文科学研究所附属東アジア人文情報学研究センター，pp.45-59（2014）。
- [13] <http://glyphwiki.org/wiki/GlyphWiki>
- [14] <http://unicode.org/charts/unihan.html>
- [15] 周祖謨：廣韻校勘記，5冊，商務印書館，長沙（1938）。
- [16] 龍宇純：唐寫全本王仁昫切韻補缺切韻校箋，5冊，香港中文大學，香港（1968）。
- [17] 張涌泉主編：敦煌經部文獻合集，5-7，中華書局，北京（2008）。
- [18] <http://www.tei-c.org/>