様　式　Ｃ－１９、Ｆ－１９－１、Ｚ－１９　（共通）

科学研究費助成事業　　研究成果報告書

令和　５ 年　６ 月　２ 日現在

機関番号：１７１０２
研究種目：若手研究
研究期間：2020～2022
課題番号：２０Ｋ１３０５０

研究課題名（和文）訓点資料本文データベースの作成

研究課題名（英文）Construction of a database of the text of kunten materials

研究代表者
蛭沼　芽衣（Hirunuma, May）

九州大学・人文科学研究院・助教

研究者番号：２０８０７１７７

交付決定額（研究期間全体）：（直接経費）　　1,300,000円

研究成果の概要（和文）：　本研究の目的は、訓点資料本文データベースを作成するための基礎的研究および入力のためのシステムを構築していくことである。訓点資料のもつ膨大な情報を、いかに効率よく、また利便性の高い形でデータ化していけるかを検討するものである。
　本研究では、入力用のフォームを作成し、一方でデータの元となる資料のテキスト化を行い、実際に訓点資料のデータベースを作成しながら、入力システムやデータ内容の調整をしていった。実際の資料を見ながら作成した、データベースの構造とつき比べていったことで、データ化した点図にない点の処理や、別筆による点の扱いなど、さらに検討するべき課題があることも明らかになった。

研究成果の学術的意義や社会的意義
訓点資料は、上代末から中古～中世前期にかけてさかんにおこなわれた、漢文訓読の成果である。後世の写本でしか伝わっていない和文資料にくらべ、加点時の原本が残っているため、当時の言葉としての信頼性が高い。現存資料の点数も膨大で、内容的にも、語彙、語法、字音声調やアクセント、統語構造、仮名字体の変遷など、多種多様な情報をそなえた資料群であるといえる。しかし、調査や解読が難しく、ほとんどデータ化されていないため、現在の日本語研究に充分活用されているとはいいがたい。本研究でデータベース化の方針を立てることによって、訓点データベースが作成できれば、新たな研究可能性が開けると予想される。

研究成果の概要（英文）：The puepose of this study is to conduct basic research to create a database of Kunten materials and to build a system for their input. The Kunten materials have a lot of information, and I will examine how I can convert this information into data in an efficient and convenient manner.
In this study, I created a form to input Kunten data, and also created a text file of the material to create a database. While creating the database using these materials, I modified the input method and the content of the data.Creating the database while looking at the real materials made it clear that there were issues that needed to be further examined.

研究分野：国語学

キーワード：訓点資料　データベース

※科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属します。

様　式　Ｃ－１９、Ｆ－１９－１、Ｚ－１９（共通）

１．研究開始当初の背景

訓点資料とは、漢文を訓読する際に加えられた「訓点」をもつ資料群のことである。訓点にはヲコト点や仮名点、返り点、声点などがあり、漢文の読み方として語彙や語法、統語構造、声調やアクセント、仮名字体の変遷などがわかる、実に情報量の多い資料である。訓点の加点は上代末から始まり、中古～中世前期にかけて盛んにおこなわれた。後世の写本でしか残存していない和文資料とは異なり、加点時の原本が残っている。そのため、文体の差は留意しなければならないが、当時のことばとしての信頼性が高い。資料の点数も膨大で、未調査の資料も各所に眠っていると考えられる。これらの資料のもつ情報をデータ化すれば、日本語史の研究資料としての有用性が高い。しかし、データ化はおろか、出版・公開されているものも少なく、充分に研究活用されているとはいいがたい。そこで、訓点資料に現れる様々な情報を兼ね備えた訓点データベースを構築したいと考えた。

２．研究の目的

訓点データベースを構築するにあたって、そのためのシステムづくりを目指す。訓点資料は上述のとおり情報量が多い。入力を少しでも簡便にし、データ収集の効率化をはかる必要がある。一方で、日本語史研究に活用できる形でなければならない。そこで、データベースそのもののデザインと、データ入力用のフォームデザインとをあわせておこなっていく必要がある。

本研究では、Excel のマクロ機能（ユーザーフォーム）を使用する。比較的どのパソコンでも利用でき、簡単なプログラミングで目的に応じたフォームが作れるためである。Excel を使うことで、表の形式で入力でき、データの見やすさや並べ替えなどの利点もある。

本研究では、これまでの研究で作成してきた入力用のフォームを使用して、実際の訓点資料をデータ化し、そのために必要な作業と、データ化の中で必要となった調整を行っていく。

３．研究の方法

1. 試験用データ作成準備

入力用フォームのテストと、データベースの有効性の確認を兼ねて、試験的に訓点資料のデータ化を行った。それに先立って、データ化する資料のテキストデータの作成と、フォームと連動する点図情報のデータ化を行った。

点図情報は、築島裕編の『訓点語彙集成第一巻』にまとめられた 26 図を基におこなった。もともとフォーム作成時点で、これらの点図を基に、点の位置と形の情報を収集していたが、点図データを整えたことで、フォームに点の形と位置を入力すると、自動的に読みが提示されるようになる。これらを活用することで、解読の難しい点の読みの候補が見つけやすくなった。

また、データ化する資料に即したテキストデータを作成すれば、データベース用のフォーマットに変換できるように、テキストデータからデータベース形式に変換するマクロも作成した。

2. 試験用データの作成

フォームを利用して実際に資料をデータ化していった。訓点は漢字単位で加点されるため、データは単字ごとのものになる。ただし必ずしもすべての漢字に訓点がついているものではないため、フォームの出し入れが頻繁となる。また、漢字ごとに差される情報（ヲコト点だけなのか、仮名点だけなのか、またはその両方なのかなど）が異なるため、ヲコト点、仮名点など、点の種類ごとのフォームも作成したのだが、字ごとにフォームが異なったりして、かえって煩雑であった。また、上記で準備した点図データにない点の存在も判明し、フォームそのものの改良が必要となることがわかった。

3. 声点データの作成

訓点データ利用のケーススタディ用に、作成中のフォームを使用して、儀軌訓点資料から、声点のみのデータベースを作成した。儀軌とは密教の儀式次第を説明した仏典の一種である。その儀式の中で唱える呪文（真言・陀羅尼）は、梵語を漢字で音訳して表わされている。その読み方の注記として声点や仮名点などが差されることがある。九州大学所蔵の儀軌資料 9 点から、この真言・陀羅尼のみを取り出し、声点と仮名点のデータを作成して、真言への加点の性格について検討してみることにした。ただしこの研究では、これまで用意してきた声点に関する項目（墨色、形、単点複点の別、声点の種類、漢語か梵語の音訳漢字かの別）のほか、原梵語や漢字の原音声調などの情報も必要となったため、別にそれらの項目も設けている。原語の推定には、『真言事典』にあるものはそれを基に、『真言陀羅尼』、大正新脩大蔵経所収の悉曇文字、Monier Williams Sanskrit Dictionary、梵語文法などを用いておこなった。そのため、これまで作成してきたデータベースから単純に声点情報だけを抜き出したものではなくなった。

４．研究成果

データ形式

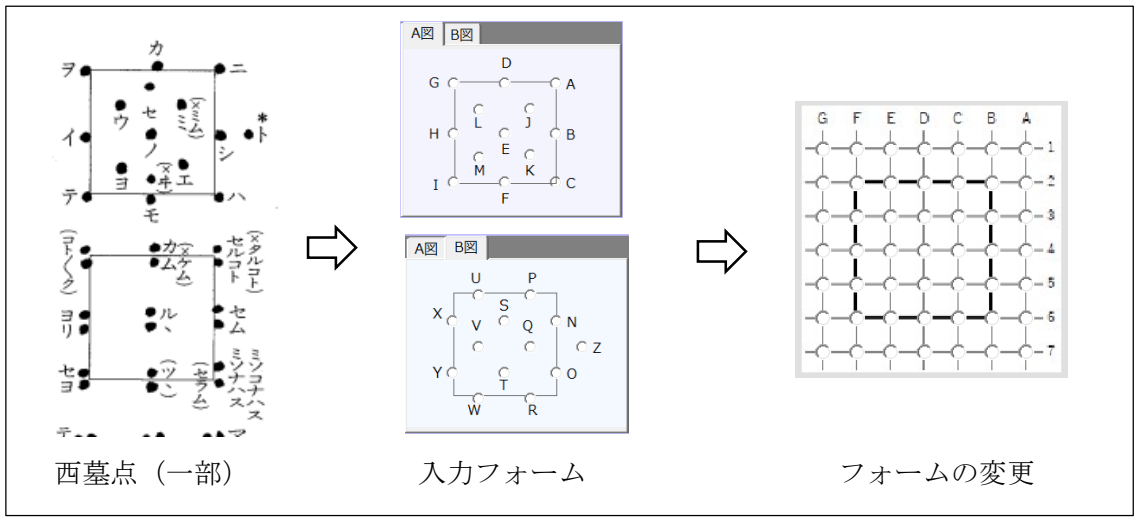
1. テキストデータ

訓点データベースを作成するためには、元となる資料のテキストデータが必要である。そのため、テキストデータ化も同時にすすめた。ただし、後の検索の便も考慮すると、資料の表記そのままではなく、漢字表記の方法について方針を定める必要が生じる。そこで、大正新脩大蔵経データベースに倣い、異体字などを用いず旧字体で統一することにした。しかし一方で、元資料の表記との乖離が生じてしまう。この問題を解消するためには、別に「元漢字」項目が必要となる。ただし、どの程度までを異体字とするかが問題となる。例えば、(本文)「経」→(データ)「經」の場合、元漢字は「経」でよいが、(本文)「𠮟」→(データ)「剛」のように Unicode にもない字の場合、どのように扱うべきかといったものである。このほかにも「麽、摩、磨」など区別なく使われる字や、「唵」のように Unicode になく、代替文字もないような特殊な漢字なども同様に検討していかなければならない。検索の便宜と資料の再現性とを考慮しなければならず、今後の課題である。

2. 入力フォームとデータ

入力フォームの作成段階では仮入力していた点図データを完備させた。これによってさまざまな資料に対応できるようになった。ヲコト点の形と位置をフォーム上で指定すると、自動的にその読みが入力されるようにしてあるので、訓点の位置や形が不確かな場合でも、入力を試してみながら、文脈にあわせて検討していくことができる。訓点と漢文についての基礎的な理解があれば、ある程度データが集められるようにしている。それでも解決できない点もあるが、データにしておけば、後から類似の点を集めて再検討が可能である。データ化のメリットは、訓読文作成のための課題が振り返りやすくなるところにもある。

研究方法で述べたとおり、点図データには現在知られている点図集をもとに作成した。それらに収められている図から、必要なヲコト点の位置と形を整理し、特に位置に関しては、アルファベットの記号をあてて示すようにしていた。国立国語研究所の点図データベースでは座標を用いて示しているが、検索の便を考慮し、あえて記号化したのである。しかし、実際の資料を用いてデータを作成していくと、点図にはない点が現れてくる。そのため、ヲコト点の位置情報の扱



西墓点（一部）　入力フォーム　フォームの変更

いも再検討する必要が生じた。その結果、国語研に倣って座標をとることにした。ただし、数値による座標（0, 0）ではなく、横軸にアルファベット、縦軸に数字を配し、その組み合わせを用いるという Excel 式によって示すことに改めた。すべての位置に対応でき、検索もしやすいと考

えられる。この仕組みによって改めて資料をデータ化しながら、有用性について検討する必要がある。

このように、実際のデータを作成していったことによって、さらなる検討課題や改善すべき点を見つけることができた。

3. 声点データ

訓点データ利用のために、作成中の入力フォームを活用して、儀軌資料から真言のみをとりあげ、そこに差された声点および仮名点のデータセットを作成した。9 資料の真言部分の総漢字数は 19,125 字、うち訓点をもつものは 8,960 字である。

声点の差声は上声点が最も多く、入声点はほとんどあらわれない。漢字の原音声調と比べると、声点の調類と一致するものは、全体の 3 割程度しかなく、真言の声点は漢字の原音を表したものではないことがわかる。そこで、比較の範囲を梵語原語の単語単位に広げて調べてみた。そのために、データ上には「梵語」の項目が必要となった。梵語と比較してみることで、例外はあるものの、それぞれの単語にはある決まった差声がなされていることがわかった。

また、仮名点による真言の読みを見てみると、やはり漢字音をそのまま反映したものでないことがわかる。「薩嚩（sarva）サラバ」のように、字音体系に合致しない読みもみられるところから、梵語の理解があったものと考えられる。

これまでの研究では、漢字の声点は漢字の字音に関連した音調を表すものとしてみなされてきた。しかし音訳真言は、漢字を使っているとはいえ、梵語を書きあらわすための手段であって、「漢字（漢語）」ではない。当時の僧侶たちにとって、漢字で書かれた「梵語」という認識があったということがわかる。訓点のデータを作成し、資料を超えて総合的にみることで、これまであまり研究されてこなかった音訳真言の訓点に新たな知見を得ることができた。

ただし、この結論は上述のように、声点のデータだけではなく、補足的な情報があって初めて成り立つものである。「語彙」情報の問題はこのような梵語だけでなく漢語でもいえることである。「菩薩」のように、明らかに二字以上で一語となっているものもあるため、「薩は」だけをデータとして集めても意味がない。しかしこのような情報を補うには、資料を表面上でもとらえられればデータになる訓点とは異なり、漢語や梵語などの知識が必要となる。本研究の目的であるデータ入力の簡便化に反してしまう。とはいえ、研究利用上重要な項目でもあるため、どのようにデータに取り入れていくべきか、さらなる検討が必要である。これも今後の課題としたい。

５．主な発表論文等

〔雑誌論文〕　計1件（うち査読付論文　0件／うち国際共著　0件／うちオープンアクセス　0件）

１．著者名 蛭沼芽衣	４．巻 135
２．論文標題 音訳真言の訓点について―悉曇学と比較してー	５．発行年 2023年
３．雑誌名 語文研究	６．最初と最後の頁 62-49
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 －

〔学会発表〕　計1件（うち招待講演　0件／うち国際学会　0件）

１．発表者名 蛭沼芽衣
２．発表標題 音訳真言の訓点からみる音韻
３．学会等名 2022年度九州大学国語国文学会
４．発表年 2022年

〔図書〕　計0件

〔産業財産権〕

〔その他〕

－

６．研究組織

	氏名 （ローマ字氏名） （研究者番号）	所属研究機関・部局・職 （機関番号）	備考

７．科研費を使用して開催した国際研究集会

〔国際研究集会〕　計0件

８．本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関