

訓点データベースと点図の自動判別

堤 智昭^{1,a)} 田島 孝治² 小助川 貞次³ 高田 智和⁴

受付日 2021年5月18日, 採録日 2021年11月2日

概要: 本稿では, 訓点研究の成果である点図を収録した訓点データベースについて論じる. 点図をデータベース化することで, 訓点研究の過程で移点や, 釈文, 書き下し文の作成を行うときの情報検索が容易になる. また, 訓点研究の知を集約化し, 計算機を用いて解析を行うことで, これまで手作業だった訓点の比較研究を効率化する. 訓点データベースの活用事例として, 計算機による点図の分類を試みた. 実際の訓点資料から作成された点図のうち, これまでの訓点研究において第1群から第8群に分類された正解の分かっている点図に対して, 計算機を用いて自動的に分類を判定し, その精度を検証した. その結果, 重要位置9カ所にある「・」形状の点に着目して計算を行うことで, 95%を超える正答率を得られることが分かった.

キーワード: 訓点資料, 漢文訓読, ヲコト点, 構造化記述方式

Automatic Recognition of *Wokototen* Table Using *Kunten* Database

TOMOAKI TSUTSUMI^{1,a)} KOJI TAJIMA² TEIJI KOSUKEGAWA³ TOMOKAZU TAKADA⁴

Received: May 18, 2021, Accepted: November 2, 2021

Abstract: In this paper, we discuss a *Kunten* Database that contains point diagrams that are the results of *Kunten* study. *Kunten* Database offers easy searching of *Kunten* information and *Wokototen* Table. It assists comparative studies of *Kunten* that were previously hand worked. As an example of using the *Kunten* database, we tried to automatically recognition of *Wokototen* table by computer. We used *Kunten* data in *Kunten* Database to automatically determine the classification of the *Wokototen* table, and verified the classification accuracy. As a result, the correct answer rate was 95% or more can be obtained by focusing on the shape and position of *Kunten*.

Keywords: *Kunten* material, vernacular reading, *Wokototen*, digitalization

1. はじめに

日本を含む東アジア諸国においては, 共通の書記言語として漢文が用いられてきた時代がある. そのため, 日本語

史の研究をはじめ, 歴史研究では漢文で書かれた古典籍が研究対象となっている. その中に, 訓点資料と総称される, 漢籍, 仏典, 国書などの漢文に本文とは別のアノテーションがついた書籍が存在する.

このアノテーションのことを訓点と呼ぶ. 訓点は漢文を自言語として理解するための注釈であり, 漢籍, 仏典, 国書など奈良・平安・鎌倉時代を中心にそれ以降も広く用いられてきた. 訓点は, 時代や利用していた人々の流派によって様々な種類が存在し, これまでに, 吉澤義則 [1], [2], [3], 中田祝夫 [4], 大坪併治 [5] 築島裕 [6], [7], 小林芳規 [8], [9], 春日政治 [10] など多くの研究によって分類されてきた.

現在でも, 訓点研究を専門とする研究者によって, 訓点

¹ 筑波大学
University of Tsukuba, Tsukuba, Ibaraki 305-8577, Japan
² 岐阜工業高等専門学校
National Institute of Technology, Gifu College, Motosu, Gifu 501-0495, Japan
³ 富山大学
University of Toyama, Toyama 930-8555, Japan
⁴ 国立国語研究所
National Institute for Japanese Language and Linguistics, Tachikawa, Tokyo 190-8561, Japan
^{a)} tsutsumi.tomoaki.gn@u.tsukuba.ac.jp

資料の解説や、資料の歴史的背景の研究が行われている。しかし一般的な訓点研究は、いまだに研究者が訓点資料に付加されたヲコト点や仮名点を目視で解説して日本語文、いわゆる書き下し文を作成する方式が多くとられている。この方式は、訓点資料に精通した研究者であっても容易ではなく、また時間がかかる方式である。そこで本研究では、これらの工程を電子化しコンピュータによる作業支援を行うために訓点データベース（以下訓点 DB）を開発した。また、訓点 DB を活用した訓点研究事例として、訓点資料を解説する過程で作成したヲコト点図が、中田祝夫 [4]、築島裕 [6]、[7] によって分類された第 1 群から第 8 群のいずれと近いものであるか自動判別を試み、その精度を評価した。

2. 訓点研究

2.1 訓点資料とヲコト点図

かつての東アジア諸国は、中国語である漢文を広く受容し、自言語とは異なる漢文を自言語によって直接理解するために、漢文訓読という読書法を考案した。このときに用いられたのが訓点であり、訓点の書き入れがなされた資料が訓点資料である。訓点資料は主に日本語の表記、音韻、語彙、文法などの日本語史の研究に利用されるが、そこに記載された内容を研究することで、当時の人々がどのように漢文資料を扱っていたかを理解でき、文学、歴史学、仏教学、文化学など広く史的研究を行ううえでも重要な研究対象となっている。

訓点の中にはヲコト点と呼ばれる、助詞や助動詞、活用語尾、漢字の読みなどを表す符号が存在する。これは、漢字の字画内部や周辺につけられる符号であり、その形状（・、一、┐、└ など）と付けられた位置によって読みが決まる。たとえば漢字の右上に「・」の形状を持つ点が記載されていたら「ヲ」と読み、同様の形状でも右上よりも下の位置にあれば「コト」と読む、など様々なものが存在する。どの位置にどのような形状の点を付与するかは、ヲコト点を使用する流派や加点者によって様々な種類が存在する。

喜多院点や西墓点といった訓点の種類ごと、もしくはある訓点資料に付与されたヲコト点を網羅し帰納的にまとめたものをヲコト点図という。ヲコト点図の例を図 1 に示す。ヲコト点図にはヲコト点に分類されない訓点、たとえば、アクセントを表す声点、漢字の派生義を表す破音、段落を表す科段点などが記載されることもあるが、通常はヲコト点の形状ごとに、漢字に見立てた四角形（「壺」と呼ぶ）の周辺に、その読みとともに記載していく。そのためヲコト点図は、ヲコト点が記載された壺を 1 つ、もしくはヲコト点の形状ごとに複数の壺を集めた集合体であるといえる。

現在までに確認されているヲコト点の種類は 200 種を超えている。中田祝夫 [4] は、代表的なヲコト点 26 種（＝壺の集合体としてのヲコト点図、以下、主要ヲコト点 26 種）をヲコト点の歴史的変遷・発達を考慮し 8 つの群（第 1 群

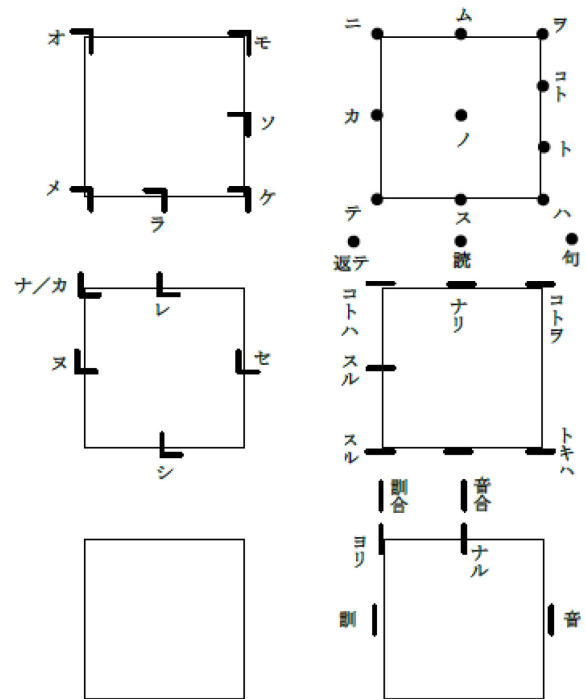


図 1 東洋文庫蔵重要文化財古文尚書ヲコト点図（小助川貞次作図）
Fig. 1 Wokototenzu of Kobunshosho owned by Toyobunko (Example of Wokototenzu created by Teiji Kosukegawa).

点、第 2 群点、第 3 群点、第 4 群点、第 5 群点、第 6 群点、第 7 群点、第 8 群点）に分類している。主要ヲコト点 26 種のヲコト点図は築島裕 [7] にまとめられている。これら主要ヲコト点には、喜多院点、西墓点などの名称が古くより付されており、訓点を読み解くための資料として編纂されてきたヲコト点図といえる。

一方、このような歴史的に伝承されたヲコト点図とは別に、実際の訓点資料に使われているヲコト点から、その資料を研究、解釈するために作成される新しいヲコト点図も存在する。こういった訓点資料研究、資料解釈の過程で作成されるヲコト点図は当時の訓点使用者が作成したもの以外にも存在し、現在も新しいものが作られている。たとえば、訓点研究者が実際の訓点資料に記載されたヲコト点をまとめた点図などがそれにあたる。これらは、同じ訓点資料から作成されるヲコト点図であっても、研究者の解釈によって異なることもある。

2.2 訓点資料研究におけるヲコト点図

訓点資料を対象とした主要な研究には、中田祝夫 [4]、大矢透 [11]、春日政治 [12]、大坪併治 [13]、築島裕 [14]、太田次男・小林芳規 [15] などがある。これらの研究では、訓点資料研究の成果を発表する方式として、漢文資料に付加されたヲコト点や仮名点を解説して日本語文にする書き下し文方式が多くとられている。この解説の過程で、先に述べたような訓点資料に記載されているヲコト点を一覧にした

ヲコト点図が作成される。

2.3 訓点資料電子化の研究

情報システムを用いて訓点資料の解説を支援する取り組みには、田中勝ら [16] がある。田中らは、訓点資料の翻刻を支援するシステムを開発している。このシステムでは、実際の資料画像を用いて、文字領域や訓点の座標を抽出し、電子的な翻刻作業のサポートを実現している。田中らは、訓点の読みの自動推定も行っている。訓点の打点位置判定で用いる座標系は、字画の中を 3×3 マスで分け、さらにその周辺上下左右を囲むように 8 マスを設定し、判定している。田中らの取り組みでは、資料に記載された訓点の種類は事前に判明し、位置と形状、読みの情報がデータベース化されている訓点資料を対象としている。実際の資料に加点された訓点の読みを推定するものであり、我々が行っている点図の種類を推定するものとは対象が異なる。

訓点の電子化を行う場合、漢字のどの部分に加点されているかを表すために、座標軸を定義する必要がある。訓点の記載位置の座標化は、韓国における口訣点の研究においても行われている [17], [18], [19], [20], [21], [22]。朴鎮浩 [17]、呉美寧 [21] では訓点の打点位置として、字画を中心に 5×5 マスの座標系を定義している。我々は、同様の座標系を日本の訓点に適用するために、字画を中心として 7×7 マスの座標を定義し電子化を試みた [23]。本稿でも、同様の座標系を用いている。

3. 訓点研究の電子化

本稿で述べる訓点 DB は、2.2 節に示した訓点研究の過程で作成された点図を、コンピュータを用いて効率化し、電子的に研究成果を扱えるようにする。それにより研究成果の比較・共有を容易にすることを、システム設計の目標とした。訓点 DB により、これまでは個別の研究者が研究を積み重ね、弟子に相伝されていくことが多かった訓点研究の知を集合化する。

3.1 開発システムの概要

訓点 DB は、これまでの訓点研究の成果として作成された点図を収集したものであり、この DB に対して、作成したヲコト点図などの訓点情報の入力や訓点の検索ができるようにすることで、移点や釈文、書き下し文の作成をするときの情報検索を支援する。訓点 DB に収録されたデータは、ヲコト点データベース検索システムから検索する。訓点 DB にデータを入力したり、他の研究者と共有したりするための点図データを生成するには、点図入力アプリケーションを用いる。訓点 DB の情報は、WebAPI からも取得可能であり、訓点 DB を用いた別システムの開発も容易になるよう実装した [24]。これにより、訓点研究の情報インフラとして機能する。

3.2 訓点 DB の設計

3.2.1 点図の種類

点図には、ヲコト点のほかに、区切りを示す句読、アクセントを示す声点・漢字の派生義を表す破音、熟語を示す合符なども記載される。これらの訓点に分類されるアノテーションを、本 DB は集約、管理する。

本 DB には大きく分けて 2 種類の点図が入力される。1 つ目は主要ヲコト点 26 種のヲコト点図である。これは、訓点を使用していた人々が伝えてきた知識、情報を、訓点研究者が研究し主な使用者や流派によって分類、まとめたものである。2.2 節の研究手順で示したような、特定の資料に書かれていた訓点をまとめたものではなく、歴史的に伝承されたヲコト点図を網羅したものである。そのため、実際の訓点資料とは直接結び付かない。よって 26 種の点図名をもとに、分類、管理する必要がある。

2 つ目は実際の訓点資料に基づいて作成された点図である。これは 2.2 節で説明したように、訓点資料の内容（漢文）を解説する過程においてその訓点資料に記載されていた訓点を書き出した点図である。そのため訓点資料や、書き出した人物の解釈によって出来上がる点図が異なる。たとえば、同様の内容が記載された写本が 2 つあったとしても、その写本に記載される訓点は訓点を付与した人物の流派や解釈によって異なる。また 1 つの写本に、複数の年代、複数の人物によって加点される場合もあり、訓点資料ごとに記載されている訓点の内容は異なる。よって、これらの点図は作成に用いられた訓点資料と作成者の情報とともにデータ化する必要がある。

3.2.2 訓点 DB のデータテーブル設計

DB にまとめる点図は、上述のとおり主要ヲコト点 26 種と訓点資料に基づいて作成された点図で大きく意味が異なる。そのため、それぞれに適したテーブルを作成し別に管理する設計とした。

(1) 主要ヲコト点 26 種のデータ構造

主要ヲコト点 26 種は、「ヲコト点テーブル」と「ヲコト点の種類テーブル」の 2 つを持つこととした。ヲコト点テーブルには、ヲコト点図に記載されたヲコト点 1 つを 1 単位として保存する。ヲコト点の種類テーブルには、ヲコト点の種類（東大寺点、喜多院点など）を 1 単位として保存する。

ヲコト点テーブルが持つ項目は、表 1 に示す 8 項目とした。ヲコト点の構成要素である「種類」「位置」「形状」「読み」については、それぞれ点の種類 ID、X 軸の位置・Y 軸の位置、形状、読みの項目に保存する。さらに、ヲコト点図の何壺目に記されていた点であるかを壺数の項目に保存した。主キーとなる点の ID は、ユニークな整数を作成し割り振った。

ヲコト点の種類テーブルが持つ項目は、表 2 に示す 3 項目とした。点の種類 ID の項目には「喜多院点」や「西墓

表 1 ヲコト点テーブルの項目
Table 1 Items of Wokototen Table.

項目名	説明
点の ID	ヲコト点 1 つ 1 つに振った固有の ID を保存する.
形状	ヲコト点を持つ形状 (「・」「 」など) を保存する.
読み	ヲコト点を持つ読み (「ヲ」「コト」など) を保存する.
点の種類 ID	ヲコト点の種類データベースと連携するためのデータ (1 から 26 までの番号) を保存する.
X 軸の位置	-3 から 3 まで, X 軸方向のどの位置に点が存在したかを保存する.
Y 軸の位置	-3 から 3 まで, Y 軸方向のどの位置に点が存在したかを保存する.
壺数	何壺目に記載されていた点であるかの情報を保存する.

表 2 ヲコト点の種類テーブルの項目
Table 2 Items of Wokototen type Table.

項目名	説明
点の種類 ID	点の種類 1 つ 1 つに振った固有の ID を保存する. ヲコト点テーブルから点の種類テーブルを参照する時に使用する.
点の種類名	点の種類名 (「喜多院点」「西墓点」など) を保存する.
群番号	所属する群の番号 (1 群から 8 群) を保存する.

点」といった点の種類名を保存する. 群番号は, 築島裕 [7] に記載された各ヲコト点の群分類に基づいて, 1 から 8 までの数字を保存する. 点の種類 ID は各ヲコト点の種類 1 つずつに, 1 から 26 までの値を重複しないように割り振り保存する. この値は, ヲコト点テーブルが持つ点の種類 ID と対応しており, ヲコト点テーブルからヲコト点の種類テーブルのデータを参照可能である.

(2) 訓点資料に基づいて作成された点図のデータ構造

訓点資料に基づいて作成された点図は, 3 つのテーブルから構成される. 1 つ目は「位置」「形状」「読み」などの訓点情報を保存する「訓点テーブル», 2 つ目は訓点資料の書誌情報を保存する「書誌情報テーブル», 3 つ目は点図作者に関する情報を保存する「点図作成者テーブル」である. 主要ヲコト点 26 種のデータと連携して利用することを想定し, 訓点テーブルは (1) のヲコト点テーブルに, 書誌情

表 3 書誌情報テーブルの項目
Table 3 Items of the bibliographic information table.

項目名	説明
書誌情報 ID	書誌情報 1 つ 1 つに割り振った固有の ID を保存する. ヲコト点テーブルから点の種類テーブルを参照する時に使用する.
訓点資料名	訓点資料の名称を保存する.
巻次	巻次情報を保存する.
所蔵	所蔵されている場所の情報を保存する.
群の分類	所属する群の情報を保存する.

報テーブルは (1) のヲコト点の種類テーブルに対応するように設計し, 必要な情報を追加する形で実装した.

訓点テーブルは, 主要ヲコト点 26 種のヲコト点テーブルと同様に, 訓点 1 つを 1 レコードとして入力する. 項目には表 1 に示したものに加えて, 資料記載の「点の色」, 句読, 合符などどのような訓点なのかを保存する「訓点記号カテゴリ」の項目を持つテーブルとした. また, “書誌情報テーブル” と連携するための書誌 ID, および “点図作成者テーブル” と連携するための作成者 ID も追加した.

「訓点記号カテゴリ」は 2 桁の数字を割り当てることとした. カテゴリの割当てを表 4 に示す. 1 桁目は大枠の種類を, 2 桁目は小分類とする方針とした. 具体的には, ヲコト点, 科段点, 句読点といった漢文の読解を表す訓点については 0 から始まるコードを付与した. 校符などの補助的な書き入れは 9 から始まるコードを付与した. 仮名点をはじめ, 今後入力を予想される訓点は 2~8 のコードを 1 桁目に割り当てることとした.

書誌情報テーブルは, 1 つの書誌を 1 レコードとし, 「書誌情報 ID」「訓点資料名」「巻次」「所蔵」「群」の 5 つの項目を持つこととした. このテーブルでは「訓点資料名」「巻次」「所蔵」と資料を特定するためのデータと, その訓点資料に記載されている訓点は何の「群」に分類されるものかを保存する. 訓点資料に基づいて作成された点図の場合, すべてが 1~8 群に分類された主要ヲコト点 26 種とは異なり, 特殊点と呼ばれるような分類不能の点が記載されている場合もあり, 必ずしも 1~8 群に分類できるとは限らない. そのため, 数字での保存ができず, 文字列を保存することとなった. そのため表 2 に示した「群番号」という項目名ではなく, 表 3 にあるように「群の分類」とした. また, 「書誌情報 ID」は数字とし, ユニークな値を割り振った.

点図作成者テーブルは, 作成者 1 人を 1 レコードとし, 「点図作成者 ID」「点図作成者名」の 2 つの項目を持つこととした. 「点図作成者 ID」は数字とし, ユニークな値を割り振った.

表 4 訓点記号カテゴリ
Table 4 Category of Kuntten.

コード	カテゴリ名	内容	備考
01	科段	科段	段落等
02	句読	句，読，切	区切り
03	声点・破音	四声，軽重，濁，派生義	アクセント，意義
04	名詞句	人名，地名，国名，書名	固有名
05	合符	音合，訓合	熟語
06	語順符	返，一二，レ，行待	返り点
07	音読符・訓読符	音読符・訓読符	読み方（音訓区別）
08	ヲコト点	ヲコト点	読み方（語形表示）
90	校符	見消，補入	本文校正
98	読み記載なし	∅	∅を記載している，読み記載のない点に割り当てる
99	その他	その他	分類不能

*仮名点等は一部省略

3.3 訓点 DB 収録データ

本 DB には現在，主要ヲコト点 26 種と訓点資料から作られたヲコト点図 258 種のデータが入っている．後者 258 種の内訳は，築島裕 [7] 記載の点図 192 種，中田祝夫 [4] 記載の点図 55 種および小助川作成の点図 11 種である．築島裕 [7] 記載の点図のうち，1 群から 8 群までに分類されている点図は 159 種，1 群から 8 群に分類不能な特殊点と呼ばれる点図は 33 種である．

収録データは，Web 上の検索ページ*1から検索可能である．主要ヲコト点 26 種の検索では，訓点の形状，読み，位置，名称の項目から訓点を検索可能である．訓点資料から作られたヲコト点図データは，訓点の形状，読み，位置，訓点資料名，所蔵，点図作成者から訓点の検索が可能である．検索項目が異なるため，これらは検索ページを分けてそれぞれ作成した．これら項目のうち，内容の記載がない場合は「∅」を入力することとした．

主要ヲコト点 26 種の検索ページから，訓点の検索を行った結果を図 2 に示す．図 2 の例は，喜多院点の星点「・」を検索し表示した例である．検索結果のページでは，該当する訓点を一覧にした表と，それを 7×7 マスの座標軸上に図示したものが表示される．表の項目は，点名，読み，形状，X，Y，壺数である．7×7 マスの座標への図示は，該当するヲコト点が存在した座標のセルに形状と読みが表示

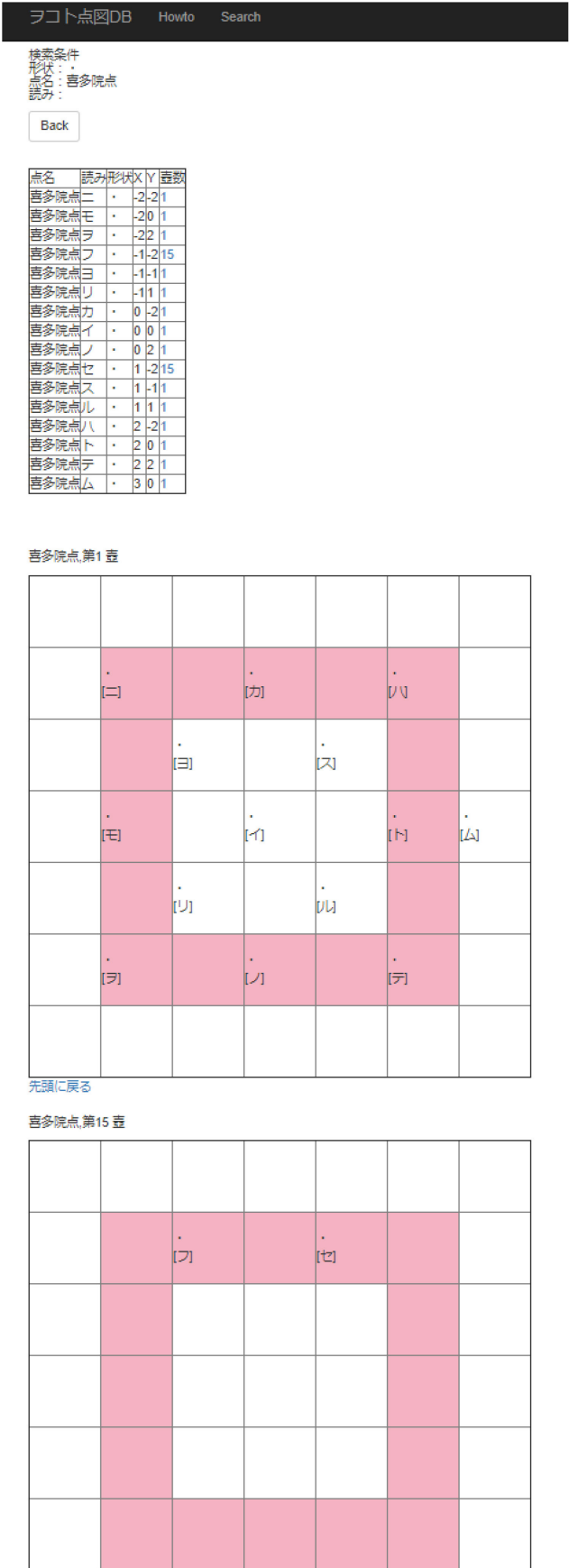


図 2 訓点 DB の検索結果例
Fig. 2 An example of search result of Kuntten DB.

*1 <https://cid.ninjal.ac.jp/wokototen/b>

される。実際の点図と近い形で表示されるように、検索結果は壺ごとに表示することとした。字画の境界にあたる、中心5マスは薄紅色で表示し、ヲコト点図の四角形を視覚的に表現した。

4. 訓点データベースを用いた点図自動分類の試み

訓点研究の過程では、資料に記載されている訓点が、既存のもののうちいずれの種類のものに近いかを判断することがある。このとき多くの場合基準になるのは中田祝夫[4]の主要ヲコト点26種とそれらの1~8群の分類である。たとえば、この訓点資料に加点されている訓点は、東大寺点の流れをくむものであり第3群点に分類される、というように判断する。それによって、従来の点図情報とも照らし合わせながら、訓点の読みを解読していく場合もある。

本研究では訓点DBを利用して、この主要ヲコト点26種との照らし合わせ作業を自動化し、類似する訓点種類の自動推定を試みた。実験には、主要ヲコト点26種のデータと、訓点DBに集約した点図データのうち築島裕[7]記載の1群から8群までに分類されている点図159種を用いた。築島の分類を正解とし、判定結果の精度を検証した。

4.1 判定方法

判定では、159種の点図それぞれに対して、主要ヲコト点26種とそれぞれどの程度類似性があるかを判定する。最も類似していると判定された主要ヲコト点が属する群が、その点図の判定結果となる。たとえば、ある資料から作られた点図が主要ヲコト点26種のうち第2群点である喜多院点に最も一致していると判定された場合、その点図は第2群点であると判定される。

判定には、点図の特徴をとらえた判定方法が必要であると考え、適切な判定方式を検討するため、比較する形状や位置を変化させて複数の方法で実験を行った。形状を変更した実験では6種類の計算方法で判定を試みた。また、位置による重みを変更した実験では4種類の計算方法で判定を試みた。

(1) 比較する形状

点図には、多くの読みと形状が記載される。これまでの我々の研究でも、主要ヲコト点26種の中でもそれぞれの点図に記載される点の数は大きな差があることが分かっている。主要ヲコト点26種のうち、最も点の個数が多いものは宝幢院点の263個であり、最も少ないものは智証大師点の32個である。このように、ひと口に点図といってもそこに記載された訓点の個数には大きな差がある。この傾向は、実際の訓点資料から点図を作成した場合でも同様に起こりうると考え、すべての点を総当たりで比較する場合と、頻出度の高い点に絞って比較を行った場合で精度を検証した。

実験では、形状を「・」「ㄣ」「|」「一」に限定して比較することとした。これらを選んだ理由は、主要ヲコト点26種のすべてに記載がある形状であり、26種すべての中で多く用いられる点であるためである[25]。これら4つは、主要ヲコト点26種の中では「・」点が最も多く登場し、ついで「ㄣ」が多く「|」「一」は同数という順になっている。さらに、1つの形状のみでは判定できない場合を想定し、「・」「ㄣ」「|」「一」4種すべてを用いて判定した場合、およびすべての形状を対象に判定した場合の実験を行った。

(2) 比較する位置

訓点は、漢字のどの位置に付いているかによっても、その点の持つ意味が異ってくる。特に、これまでの訓点研究において、字画の四隅に付いている点は重要な意味を持つとされており、中田祝夫[4]では字画の四隅に注目して8つの群に分類を行っている。築島裕[26]では中田とは異なる分類を論じているが、やはり字画の四隅に着目していることは変わらない。また、我々のヲコト点図の計量的研究においても四隅が最も多く点が付与される位置であることが分かっている[25]。この計量では、四隅の次に多く点が付く位置は、上辺中央(0, -2)、下辺中央(0, 2)、左辺中央(-2, 0)、右辺中央(2, 0)、および中央の(0, 0)であることが分かっている。よって、字画の四隅とこれら5つの位置にある点には重みを付けて計算を行うこととした。以下これらの位置を重要位置9カ所と呼ぶ。

(1)をふまえ、まずは対象とする形状を変更し、以下の6種類の計算方法で類似性を点数化した。比べている点図間に、同じ位置と読み、形状の点があればその位置では一致していると見なし1ポイントとして点数化することとした。なお、点数化した結果、同点のものが複数あり1つの群に絞れなかった場合、一致した点の数が0の場合を除き、複数の候補のうちに築島の分類と群の分類が等しい点図があれば一致とした。

- i. すべての形状の点を対象とする
- ii. 「・」のみを対象とする
- iii. 「ㄣ」のみを対象とする
- iv. 「|」のみを対象とする
- v. 「一」のみを対象とする
- vi. 「・」「ㄣ」「|」「一」を対象とする

次に、(2)をふまえ、位置に応じて4種類の方法で重みを付け、類似性を点数化した。(1)をふまえ形状は星点「・」に限ることとした。比べている点図間に、同じ位置と読みの点があればその位置では一致していると見なし、一致した位置(マス)の個数に応じて点数化することとした。なお、同じ位置に複数の読みが記載されている場合、1つでも一致していれば、一致とした。①の条件はiiと同様のものである。

- ① すべての座標(49マス)を平等に1ポイントとする
- ② 重要位置9カ所のうち、四隅を3ポイント、中心を2

表 5 判定結果 1 (形状の違い)

Table 5 Classification result 1 (Difference of shape).

判定方法	築島分類と合致した点図数 (正解判定数/総数)	判定候補が複数群であった点図数
i	144/159	5
ii	152/159	6
iii	28/159	27
iv	33/159	52
v	25/159	33
vi	151/159	11

表 6 判定結果 2 (位置による違い)

Table 6 Classification result 2 (Difference of position).

判定方法	築島分類と合致した点図数 (正解判定数/総数)	判定候補が複数群であった点図数
①	152/159	6
②	153/159	4
③	153/159	4
④	41/159	32

ポイント, 重要位置 9 カ所以外を 1 ポイントとする

- ③ 重要位置 9 カ所のみ 1 ポイント, それ以外は 0 ポイントとする
- ④ 重要位置 9 カ所のみ 0 ポイント, それ以外を 1 ポイントとする

4.2 判定結果

4.1 節で示した条件 i から vi, および①から④の計算方法で 159 種の点図を判定した結果, 表 5, 表 6 のような結果となった. 形状による違いは, i, ii, vi で 90%を超える正答率を得た. 特に, ii, vi は 95%に近い正答率となっている. 一方で「・」以外の形状では 20%程度の低い正答率となり, 判別できない点も多く存在した.

位置による違いは, ①②③で 95%を超える高い正答率を得た. 一方で, 重要位置 9 カ所を用いない④では 25%程度の低い正答率となった. すべての座標を対象とした場合と, 重要位置 9 カ所を条件に判定した場合で, いずれも近い結果となった. ①と②で結果が異なった点は 1 つであり, 聖語蔵の四未曾有経であった. この点図は築島の分類では第 4 群点に分類されるものであるが, ①の判定方法では第 2 群点となり, ②の判定方法では第 4 群点となった.

4.3 考察

4.3.1 点図の機械的な群判定

形状による点図の群判定の結果から, 主要ヲコト点 26 種すべてに記載のある形状のうち, 「・」を用いた場合に高い正答率が得られることが分かる. 「ㄣ」「|」「—」をそれぞれ用いた場合の正答率が低く, 判定不能な点図也多

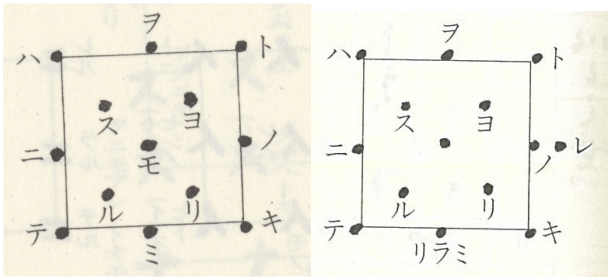


図 3 東南院点, 大乘大集地蔵十輪經「・」点. 文献 [7] より引用
Fig. 3 Tonaninten and Daijodaijuzojurinkyo.

くなったことから, 「・」点が訓点の種類を判別するためには重要な点であるといえる. さらに, すべての形状を対象とした場合よりも, 「・」点のみを判定に用いたほうが, 精度が高い. そのため, これまでの訓点研究と同様の群分類を自動で判定する場合, 「・」点を判定に用いる形状とすることが有効であると考えられる.

位置による点図の群判定の結果から, 重要位置 9 カ所を用いると高い精度で従来の分類と同様の判定結果を得られることが分かる. 重要位置 9 カ所に重みを付けて判定した場合, すべての位置を同様の重みで判定する場合に比べて精度が下がることはなく, 正答数が 1 増え, 判定候補が複数群だった個数が 2 減り, 若干の精度向上が見られたことから, 重要位置 9 カ所を重視した判定が有効であると考えられる.

4.3.2 訓点 DB を用いた訓点研究の可能性

4 種類の形状を用いた場合に複数の群が候補に上がる点図の個数が増えたことと, すべての形状を用いた場合に正答率が若干下がったことから, 形状「・」の点以外に従来の群分類とは異なる共通点がある可能性が示唆される. 重要位置 9 カ所以外の位置に付与された点を対象に判定を試みた④の結果でも, 判定候補が複数の群だったものが増えていることから, 重要位置 9 カ所以外の位置にある点に, これまで着目されていなかった共通点があると考えられる.

そこで今回は, 重要位置 9 カ所以外の「・」点に着目し, ④の方法で判定した結果のうち, 一致した個数の多かった点図を抽出した. ④の方法では, 主要 26 種ヲコト点と一致した点の数は 0~4 個の範囲であった. そこで, 3 個以上一致した点図を抽出した結果, 表 7 のような結果が得られた.

表 7 記載の点図の多くが, 第 2 群点である喜多院点と第 3 群点である東南院点と一致しており, 築島の群分類とは異なる結果となるものも多い. これら 2 つの点図の「・」点は図 3, 図 4 の左の点図のようになっている. 表 7 から一例として, 喜多院点に近いと判定された金光明經文句と, 東南院点に近いと判定された大乘大集地蔵十論經の「・」点を図 3, 図 4 の右側に示す. これらを見比べると, 字画の中にある点 4 つ, 「ヨ」「ス」「ル」「リ」と読む点が一

表 7 ④の判定結果の抜粋

Table 7 Excerpt of classification result of ④.

訓点資料名	所蔵	書写年代	加点点代	築島分類	類似判定された 26種ワコト点	群分類	一致数
金光明經文句	園城寺	唐時代	寛平 [889-898 年]頃	第1群点	喜多院点	第2群点	4
建立曼荼羅護摩 儀軌	石山寺	平安時代中期	平安時代中期	第3群点	喜多院点	第2群点	4
成唯識論述記	知恩院	平安時代初期	長保 [999- 1004年]頃	第3群点	喜多院点	第2群点	4
大毗盧遮那成仏 經〔第二次点〕	国立国会図 書館	奈良時代	平安時代中期	第3群点	喜多院点	第2群点	4
華嚴經探玄記 〔第二次点〕	聖語蔵	奈良時代	平安時代初期	第4群点	喜多院点	第2群点	4
大乘大集地蔵十 論經	東大寺図書 館, 聖語蔵	平安時代初期	元慶 7[883]年	第3群点	東南院点	第3群点	4
神供法	醍醐寺	平安時代中期	平安時代中期	第3群点	東南院点	第3群点	4
大唐故三蔵玄奘 法師行狀	東寺	平安時代後期	嘉保 3[1096] 年	第3群点	東南院点	第3群点	4
四未曾有經	聖語蔵	平安時代初期	長保 [999- 1007年]頃	第4群点	喜多院点	第2群点	3
華嚴經探玄記	聖語蔵	奈良時代	平安時代初期	第3群点	喜多院点	第2群点	3
大毗盧遮那成仏 經〔第三次点〕	国立国会図 書館	奈良時代	治安 2 [1023] 年	第2群点	喜多院点	第2群点	3
大般涅槃經	西福寺	天平宝字 7[763]年	嘉保 3[1096] 年	第1群点	喜多院点	第2群点	3
彌勒上生經疏	五島美術館	平安時代中期	平安時代中期	第3群点	東南院点	第3群点	3
往生要集	聖徳寺	長徳 2[996]年	長徳 2[996]年	第5群点	仁都波迦点	第1群点	3

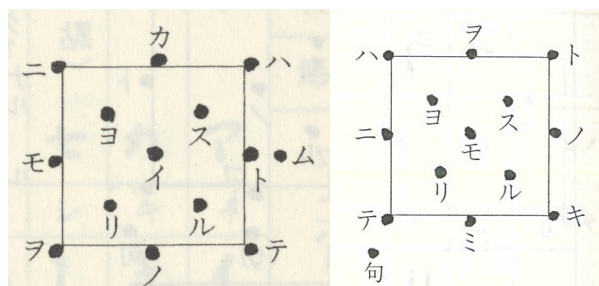


図 4 喜多院点, 金光明經文句「・」点. 文献 [7] より引用

Fig. 4 Kitanointen and Konkomyokymongu.

ていることが分かる. 表 7 に記載された, 喜多院点, および東南院点に分類された資料の点図はすべて同様の特徴を示しており, これらには何らかの関係があると考えられる.

従来の訓点研究では字画の四隅に付与されている「・」点の重要視されていた. 実際の訓点資料にはこれらの位置, 形状が多く存在していたためであるが, それと同時に計算機の導入が遅れていた訓点研究では大量の点図を比較検討することが困難であり, 自然と重要な点に注力して研究が

行われていたためであると考えられる. 計算機上で機械的に処理が可能な訓点 DB を実現したことにより, これまでに比較が困難であった大量の点図間の比較が容易になり, このような新たな研究の発展も見込めると考えられる.

5. おわりに

本稿では, 訓点研究の過程で作成された点図を電子的にまとめた訓点 DB を実現した. これにより, 移点や, 釈文, 書き下し文の作成を行うときの情報検索が容易になった. また, これまでは個別の研究者が研究を積み重ね, 弟子に相伝されていくことが多かった訓点研究の知を集合化することができるようになった.

さらに, 訓点 DB を用いて訓点資料を解読する過程で作成した点図が, 中田祝夫 [4] による第 1 群から第 8 群までの群分類のいずれと近いものであるか自動判別を試みた. 結果として, 「・」点を用いて, 四隅や字画の中心などの重要位置 9 カ所に注目して判定することで, 機械的にも高い精度で従来の群分類と同様の結果が得られることを明らか

にした。これにより、似ている点図を表示することや、点図にあるべき点の記載がない場合に提示するなど、訓点研究の効率化が見込める。このような機能を持つ訓点研究支援システムを開発することが今後の課題の1つである。また、点図の自動判定は、訓点研究において必要となる書き下し文の自動生成を行うために必要な機能の1つであるといえる。今回の成果をもとに、今後書き下し文の自動生成の実現可能性を探る予定である。

また、計算機上で大量の点図を機械的に処理することで、手作業では比較検討がしにくい訓点研究が可能となった。これまでの訓点研究で注目されていた「・」点や字画の四隅以外、たとえば字画の内部に関する特徴を抽出することが可能になり、さらなる訓点研究の発展が見込める。今後は、訓点研究者とも連携し、訓点 DB に収録する点図をさらに増やし、言語学的な訓点 DB を用いた言語学的な研究も進めていく。

謝辞 本研究は JSPS 科研費 19K13167 および 20K00654 の助成を受けたものである。また、本研究は、人間文化研究機構広領域連携基幹研究プロジェクト「異分野融合による総合書物学」の国語研ユニット「表記情報と書誌形態情報を加えた日本語歴史コーパスの精緻化」による成果の一部である。

参考文献

- [1] 吉澤義則：国語国文の研究，岩波書店 (1927)。
- [2] 吉澤義則：点本書目，岩波講座日本文学，岩波書店 (1931)。
- [3] 吉澤義則：国語説鈴，立命館大学出版部 (1931)。
- [4] 中田祝夫：古点本の国語学的研究（総論編・訳文編），講談社 (1954)。
- [5] 大坪併治：訓点語の研究，風間書房 (1961)。
- [6] 築島 裕：平安時代の漢文訓読語につきての研究，東京大学出版会 (1963)。
- [7] 築島 裕：平安時代訓点本論考，汲古書院 (1986)。
- [8] 小林芳規：平安鎌倉時代に於ける漢籍訓読の国語史的研究，東京大学出版会 (1967)。
- [9] 小林芳規：角筆文献の国語学的研究，汲古書院 (1987)。
- [10] 春日政治：古訓点の研究，風間書房 (1956)。
- [11] 大矢 透：仮名遣及仮名字体沿革史料，国定教科書共同販売所 (1909)。
- [12] 春日政治：西大寺本金光明最勝王経古点の国語学的研究，斯道文庫 (1942)。
- [13] 大坪併治：訓点資料の研究，風間書房 (1968)。
- [14] 築島 裕：興福寺本大慈恩寺三蔵法師伝古点の国語学的研究（訳文篇・索引篇・研究篇），東京大学出版会 (1965-1967)。
- [15] 太田次男，小林芳規：神田本白氏文集の研究，勉誠社 (1982)。
- [16] 田中 勝，村川猛彦，宇都宮敬吾：訓点資料を対象とした翻刻支援システムの構築，情報処理学会論文誌，Vol.59, No.2, pp.288-298 (2018)。
- [17] 朴 鎮浩：周本華嚴経卷第 36 点吐口訣の解説，日韓漢字・漢文受容に関する国際学術会議論文集，pp.127-158 (2003)。
- [18] 張 景俊：『瑜伽師地論』点吐釈読口訣の「指示線」について，日韓漢字・漢文受容に関する国際学術会議論文集，pp.159-177 (2003)。
- [19] 李 丞宰：京都国立博物館蔵『華嚴経』卷第十七の訓点，第 3 回国際学術会議「輯漢文読法とアジアの文字」発表論文集，pp.251-283 (2005)。
- [20] 朴 鎮浩：文字生活史の観点から見た口訣，文学，Vol.12, No.3, pp.169-181 (2011)。
- [21] 呉 美寧：韓国の口訣資料および口訣研究の現況について，日韓漢文訓読研究，pp.289-332，勉誠出版 (2014)。
- [22] 鄭 在永，安 大鉉：漢文読法と口訣，野間秀樹編『韓国語教育論講座』第 3 卷，pp.335-370，くろしお出版 (2018)。
- [23] 堤 智昭，土井裕絵，田島孝治，高田智和，小助川貞次：ヲコト点データベースと検索システムの試作，人文科学とコンピュータシンポジウム「じんもんこん 2017」，じんもんこん 2017 論文集，Vol.2017, pp.253-258 (2017)。
- [24] 堤 智昭，田島孝治，高田智和，小助川貞次：訓点研究支援のための基盤システムの設計・実装，人文科学とコンピュータシンポジウム「じんもんこん 2020」，じんもんこん 2020 論文集，Vol.2020, pp.89-94 (2020)。
- [25] 堤 智昭，田島孝治，小助川貞次，高田智和：訓点資料の構造化記述方式と計算機を用いた基礎計量，情報処理学会論文誌，Vol.59, No.2, pp.278-287 (2018)。
- [26] 築島 裕：平安時代語新論，東京大学出版会 (1969)。



堤 智昭 （正会員）

筑波大学。2010 年東京農工大学工学部情報工学科卒業。2012 年同大学院工学府情報工学専攻博士前期課程修了。2015 年同大学院工学府電子情報工学専攻博士後期課程修了。東京電機大学情報環境学部助教を経て、現在、筑波大学人文社会学系助教。モバイルネットワークエミュレータ，時刻情報応用システム，自律分散型インターネットセキュリティ基盤に関する研究に従事する一方，漢字・訓点の情報処理，通時コーパスの構築・応用に関する研究にも従事。日本語学会会員，博士（工学）。



田島 孝治 （正会員）

岐阜工業高等専門学校。2007 年東京農工大学大学院工学府情報コミュニケーション工学専攻博士前期課程修了。2010 年同大学院工学府電子情報工学専攻博士後期課程修了。2011 年より岐阜工業高等専門学校電気情報工学科で，情報通信技術，計算機技術の教育研究や，人文学と情報学の複合的な研究に従事。現在，同校電気情報工学科准教授。電子情報通信学会，社会言語科学会，日本語学会各会員。博士（工学）。



小助川 貞次

富山大学. 1981 年北海道大学文学部卒業. 1986 年同大学大学院文学研究科修士課程修了. 1987 年同大学文学部助手の後, 1993 年より富山大学人文学部助教授を経て, 現在, 同大学同学部教授. 自言語による古典語文献読解

に関する比較研究, 日本国内現存漢籍訓点資料の解説・記述研究に従事. 日本語学会, 訓点語学会各会員. 文学修士.



高田 智和 (正会員)

国立国語研究所. 1975 年生まれ. 1999 年北海道大学文学部卒業. 2004 年同大学大学院文学研究科博士後期課程修了. 2005 年独立行政法人国立国語研究所研究員, 2009 年から大学共同利用機関法人人間文化研究機構国立

国語研究所准教授を経て, 現在, 同所教授. 漢字字体の研究, 漢字資料の文献学的研究, 漢字情報処理, 文字コード標準化に従事. 日本語学会, 計量国語学会, 訓点語学会各会員. 博士 (文学).