Vol.2022-CH-129 No.10
2022/5/21

近世における呉音・漢音の勢力と混読現象の展開
—『日本語歴史コーパス』を用いて—

大島英之 [1]

概要：本研究では，『日本語歴史コーパス』（CHJ）の「キリシタン資料」「狂言」「近松浄瑠璃」「洒落本」「人情本」に現れる一字漢語・二字漢語を対象に，それらに付与されている「書字形」「語形代表表記」「語形」「振り仮名」といった情報を活用し，発表者作成の「呉音漢音音形対立表」のデータと照合することで，漢字の音読みが呉音であるか漢音であるかを判断する自動的にプログラムを作成した．プログラムでは，原本における漢語表記の在り方を考慮して，分析を行うことを可能とするような工夫も試みた．この実行結果を基に，近世において，呉音・漢音の比率と，二字漢語において呉音と漢音を混ぜて用いる語（混読語）の割合とがどう推移するのかを分析した．その結果，①異なり数ではなく延べ数において漢音拡大の傾向が観察されること，②混読語の割合は増加するが延べ数と異なり数で遅速があることを明らかにした．

キーワード：日本語歴史コーパス，呉音，漢音

The Spread and Blending of *Go-on* and *Kan-on* in Early Modern Japanese
— Using *the Corpus of Historical Japanese* —

HIDEYUKI OHSHIMA[†1]

1. はじめに

日本語における漢字では，例えば「言」の字を「ゲン」とも「ゴン」と読むように，一字に二つ以上の音読み（字音）がある場合が少なくない．このような例においては，一方（ゴン）は呉音で，もう一方（ゲン）は漢音であることが多い．漢音は，奈良時代から平安時代にかけて，遣唐使や留学生らにより体系的にもたらされた字音であり，唐代長安の発音を反映していると考えられている．貴族による漢籍の学習に用いられ，当初は正しい音として奨励もされた. しかしそれ以前に, 日本には呉音が伝えられており，既に一定の定着を見せていた．呉音は南方系の発音を反映しているとされており，漢音とは異なる点が大きかった．仏教界においては呉音による経典読誦の伝統が守られたため，中国文献の学習という場に限っても，漢籍は漢音で，仏典は呉音で読むという複層的な在り方が存続することとなった．

江戸時代に入ると，儒教の普及によって漢音が尊重されるようになり, また多くの漢語が一般に普及するにつれて，各漢字の読み方が呉音・漢音のどちらか一方に定まる傾向が強まった結果，一語内で呉音と漢音を混ぜて読む（以下「混読」）漢語が増加したと説かれている[1]．しかし，その実態が計量的に確認されたことは無いようである．

筆者は以前に, 院政期の古辞書である『色葉字類抄』と，1603 年刊行の『日葡辞書』とに見られる，二字漢語の語形を分析し，中世を通じて，呉音漢音を混読する二字漢語の割合が相対的に増加していることを明らかにした[2]．しかし，対象が辞書資料であることもあって，異なり語数における割合のみを根拠としていた．そこで本研究では，国立国語研究所によって構築が進んでいる『日本語歴史コーパス』（以下「CHJ」）を用いて，漢音尊重の趨勢を計量的に確かめるとともに，近世における混読現象について，異なり語数・延べ語数の，双方からの分析を行うこととした．

CHJ には, 各語彙素に対して語種情報が付与されており，特定の作品に出現する漢語を一括して取得できるという利点がある．また，室町・江戸時代編では，翻刻テキストがコーパスの底本となっている作品も多く，それらにおいては原表記の情報も保存されている．これらを活用すれば，膨大な量の漢語に対して，一定の厳密性を担保しつつ，呉音漢音を自動的に判定することが可能となる．本研究は，漢字音研究資料としての，CHJ の有効性と限界性をうかがう試みでもある．

2. 漢語の表記と CHJ におけるマークアップ

2.1 漢語の表記

日本語文中における漢語の表記を，漢字一字を取り出してみると，次の三通りに分類できる．

1 東京大学大学院人文社会系研究科
Graduate School of Humanities and Sociology, University of Tokyo
a4.ohshima@gmail.com

ⓒ2022 Information Processing Society of Japan

①漢字に対して，その字の字音による振り仮名がある（付訓）．
②漢字の字音が仮名書きされている．
③漢字で書かれており，振り仮名が無い（無訓）．

このうち，特定の漢字と字音との結び付きを厳密に論じ得るのは①のみである．そのため日本漢字音研究では，①の例に富む古辞書や古訓点資料が主に利用されてきた．

②は，語形は確定できるが，表記を一意に定められない場合があることが問題となる．「仮名書き法華経」のような原表記が明らかな文献を除いては，補助的な利用が多かったといえよう．キリシタン資料のローマ字本もこれに準じるものと考えられる．

③は，語形を一意に定められないことが問題となる．場合によっては，漢語ではない可能性，すなわち，訓読みした可能性を排除できないことすらある．

以上を要するに，漢字音研究においては，①が主たる分析対象であり，また②も補助的に利用されてきたが，③は直接的な分析対象には据えがたいということになる．そのため，CHJ において③を排除する方法の検討が重要となる．

続いて，二字漢語について考えると，①～③の重複組合せで 6 通りとなる（図 1）．

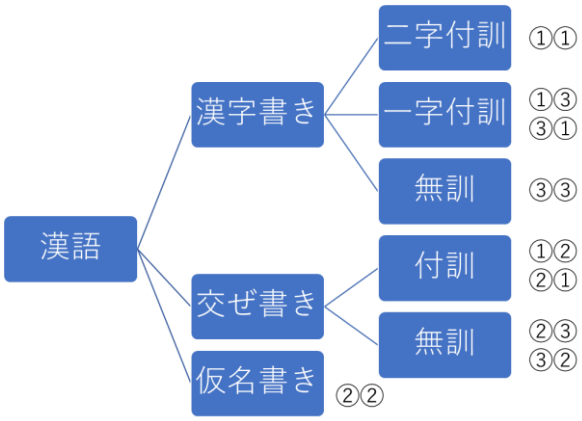
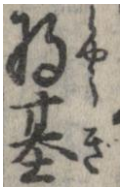


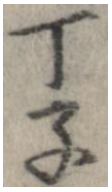
図 1　二字漢語の表記の分類

洒落本『通言総籬』(国立国語研究所蔵本の画像データ[3]より引用）に見られる実例を以下に掲げる（図 2）．

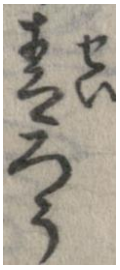


漢字書き二字付訓
将基｛しやうぎ｝
15 オ

漢字書き一字付訓
腹｛ふく｝中
24 ウ

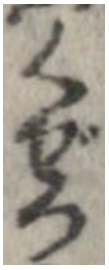


漢字書き無訓
丁子
3 オ



交ぜ書き付訓
青｛せい｝ろう
4 ウ

交ぜ書き無訓
つう言
5 ウ



仮名書き
くぜつ
37 ウ

図 2　洒落本『通言総籬』にみる二字漢語の表記

このように，洒落本では表記の在り方が多様であり，6 通り全てが現れる．もっとも，実際には「漢字書き二字付訓」「漢字書き無訓」「仮名書き」の三者が大多数を占める．

またこのほか，「歌妓｛げいしや｝」や「串戯｛じようだん｝」のように，漢字表記に対して，その字の字音とは異なる漢語（「芸者」・「冗談」）の語形が，「漢字書き二字付訓」の形式で現れることがある（以下「別漢語付訓」と呼ぶ）．別漢語付訓は特に人情本に目立つが，用いられる漢字表記は必ずしも漢語としての実例があるものには限らないようである[4]．そのため本研究では，別漢語付訓については，振り仮名に示される語形を「仮名書き」（②②）と見なすことにするが，漢字表記については対象外とする．

次節では，これらが CHJ においてどのようにマークアップされているのかを確認し，③の排除法を検討する．

2.2 CHJ におけるマークアップ

CHJ では，活用のない語における，語形または表記に関わる情報として，次の 9 種が用意されている（表 1）．以下，「『日本語歴史コーパス 室町時代編』形態論情報規程集 Ver. 1.0」[5]（以下「規程集」）を参照しつつ，それぞれについてみていく．

表 1　CHJ における語形・表記に関わる付与情報

語形に関わる情報	語彙素読み，語形，仮名形出現形，発音形出現形
表記に関わる情報	語彙素，語形代表表記，書字形
底本の表記に関わる情報	原文文字列，振り仮名

「語彙素読み」は，「同一語の活用変化・音の転化・ゆれ・省略・融合等によって生じた異形態や送り仮名の違い等の異表記をグループ化するための情報」（規程集 p. 103）と定義されており，カタカナで表される．「語彙素」には，「語彙素読み」に対する代表的な国語の表記が一つ選定されている．

「書字形」は，原則として出現形がそのまま採用されている（規程集 p.144）．ただし，「原文文字列」の誤りや不

ⓒ2022 Information Processing Society of Japan

審例に対しては，最低限の校訂が施されているように見受けられる．字体も，原則として新字体に統一されている．「原文文字列」がカタカナである部分も，ひらがなに統一されている．

「語形」は，「書字形」の読み方である．「原文文字列」や「振り仮名」から読み方が特定できる場合は，概ねそれが採用されており，カタカナ・現代仮名遣いで表される．原文が仮名書きの場合，「仮名形出現形」に原文の仮名遣いが登録されるようである．また「発音形出現形」では，「語形」における長音が長音符（ー）で表されている．

「語形代表表記」には，各語形における代表的な表記が示される．漢語においては，ごく少数を除き漢字表記であり，「書字形」に仮名を含む場合に活用できる．また，漢語の字数判定にも使用できる．

3.1 に掲げた六例のマークアップ状況は，次のようになっている（表 2）．

表 2　図 2 の例の CHJ におけるマークアップ

語彙素読み	語彙素	語形	語形代表表記	書字形	仮名形出現形	発音形出現形	原文文字列	振り仮名
ショウギ	将棋	ショウギ	将棋	将基	ショウギ	ショーギ	将基	しやうぎ
フクチュウ	腹中	フクチュウ	腹中	腹中	フクチュウ	フクチューー	腹中	ふく
チョウジ	丁子	チョウジ	丁子	丁子	チョウジ	チョージ	丁子	
セイロウ	青楼	セイロウ	青楼	青ろう	セイロウ	セーロー	青ろう	せい
ツウゲン	通言	ツウゲン	通言	つう言	ツウゲン	ツーゲン	つう言	
コウゼツ	口舌	クゼツ	口舌	くぜつ	クゼツ	クゼツ	くぜつ	

注意を要するのが，「語形」「書字形」と「語彙素読み」「語彙素」との差違である．「語彙素読み」「語彙素」は，同一語の異形態を包摂する概念であるが，その包摂基準には「呉音・漢音・慣用音等の差異」も含まれている（規程集 p.140）ため，例えば表 2 最終行の「口舌」では，「語形」は「クゼツ」であるものの，「語彙素読み」は「コウゼツ」となっている（「口」は呉音ク／漢音コウ）．そのため本研究では，「語彙素読み」でなく「語形」に従って，呉音漢音を判断する必要がある．

また，同一語の異表記も包摂されるため，「書字形」と「語彙素」「語形代表表記」とで，漢字表記が異なることがある．例えば表 2 最初の行では，書字形すなわち原文における出現形は「将基」であるが，「語彙素」「語形代表表記」では「将棋」表記となっている．漢字の違いも呉音・漢音の違いに関わることがあるため，判断に際しては「書字形」を利用することとする．ただし，「書字形」が仮名書き・交ぜ書きの場合には，基本的に「語形代表表記」を利用する．

2.1 節の図 1 に示した漢語の表記分類のうち，「漢字書き」か「交ぜ書き」か「仮名書き」かについては，「書字形」からほぼ自動的に判断することができる．更に「振り仮名」列に文字があるか否かによって，付訓の有無を判断できる．この情報を用いれば，ある程度までタイプ分けが可能となる（図 3）．

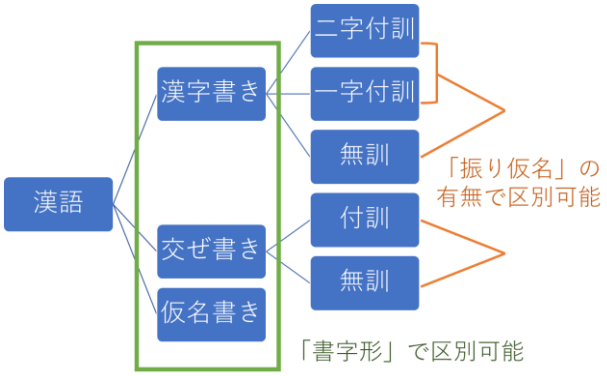


図 3　CHJ の付与情報による漢語の表記タイプの判別

しかし，「漢字書き二字付訓」と「漢字書き一字付訓」を自動的に区別するのは難しい．「漢字書き一字付訓」は無訓字を含むため除外したいところであるが，例はさほど多くないこともあり，本研究では「二字付訓」と区別せず「漢字書き付訓」として，分析対象とした．

また，「別漢語付訓」の「書字形」は，洒落本では漢字表記を，人情本では振り仮名を反映する例が多いようである（表 3）．前者の場合，書字形を参照して「仮名書き」に分類することが行えず，本研究では分析不能となっている．

表 3　別漢語付訓の CHJ におけるマークアップの一例

コーパス	語形	語形代表表記	書字形	原文文字列	振り仮名
洒落本	ゲイシャ	芸者	歌妓	歌妓	げいしや
人情本	ゲイシャ	芸者	げいしや	歌妓	げいしや

3.　対象資料

本研究では，近世の漢字音を分析するため，室町時代末期から江戸時代の作品が収められている 5 つのコーパス（表 4）の，コアデータ（人手修正を経たデータ）において，語種が「漢語」であるものを対象とする．

ⓒ2022 Information Processing Society of Japan

表 4　対象とするコーパス

コーパス名	収録作品	時代	底本	漢語数
室町時代編Ⅰ 狂言	虎明本狂言集	室町末	大蔵虎明能狂言集翻刻註解	18,738
室町時代編Ⅱ キリシタン資料	天草版平家物語・伊曾保物語	室町末	翻刻テキスト	11,267
江戸時代編Ⅰ 洒落本	通言総籬ほか 30 作品	江戸後期	洒落本大成	16,880
江戸時代編Ⅱ 人情本	春色梅児与美ほか 8 作品	江戸後期	翻刻テキスト	6,902
江戸時代編Ⅲ 近松浄瑠璃	曾根崎心中ほか 24 作品	江戸前期	新編日本古典文学全集	22,179

3.1 室町末期資料

『虎明本狂言集』は，1642 年の書写にかかる狂言台本であるが，そこに書かれる言葉は室町時代の上方語を基盤としているとされる．原本にはほとんど振り仮名が無いが，CHJ の底本である『大蔵虎明能狂言集翻刻註解』[6]では，通読の便宜のためにルビが振られており，CHJ の「振り仮名」列にはそのルビが反映されている．ただし，ごく稀に存する原本の振り仮名・振り漢字については，翻刻書では「〈　〉」で括られており，CHJ でも同様の処理がなされているため，付訓の有無を判断することは可能である，

「キリシタン資料」として収録されている『天草版平家物語・伊曾保物語』は，1592-93 年刊のキリシタン版で，数字を除き全文ローマ字書きである．底本のテキストは国立国語研究所ウェブサイトに公開されている[7]．なお本研究では全て「仮名書き」相当とみなす．

3.2 江戸後期資料

「洒落本」は，会話文体による遊里小説で，CHJ には 1757 年から 1836 年までの作品が収録されている．底本の『洒落本大成』は，原本の表記に忠実な翻刻であるため，CHJ における「振り仮名」の情報をそのまま活用することができる．2 節で見たように，漢語表記の在り方が多様である．

「人情本」は，滑稽本の後を受けて登場した，町人の恋愛や葛藤を描いた小説群で，CHJ には 1821 年から 1864 年までの作品が収録されている．底本のテキストは国立国語研究所ウェブサイトに公開されている[8]．全体的に，洒落本よりも振り仮名に富み，語形を確定できる例が多い．

なお両資料には，2.1 節末で述べた「別漢語付訓」を多く含む．これらは別漢語の「仮名書き」と見なすが，2.2 節末に述べた事情により，洒落本では処理できていない．

3.3 江戸前期資料

近松門左衛門による世話浄瑠璃（1703 年～1722 年上演）24 作品が収録されている．ただし，底本となっている『新編日本古典文学全集』の本文は，原本の表記が改められているため，CHJ の形態論情報を用いることには慎重を要する．とはいえ，当時の語形や用字法はある程度考慮されて校訂・注釈されていると考えられるので，細かい問題は含むかもしれないが，時代の空隙を埋めるためにも，分析の対象とすることにした．

なお CHJ では同時代資料として「随筆・紀行」（芭蕉作品）も公開されているが，他と比べて言語量が小さいため，本研究では扱わなかった．

4. 呉音・漢音データの作成

4.1 呉音漢音の認定

各字について呉音と漢音をどう定めるかは，難しい問題である．かつては，隋唐時代における漢字音（中古音）の枠組みを利用して，演繹的に呉音形と漢音形を導くことが行われ，その立場によって呉音・漢音を判断した漢和辞典も今なお多く存在している．しかしこの方法を採ると，実際には古い文献に確認できない音形が，呉音や漢音と認定され，逆に古くから用いられてきた音形が，呉音にも漢音にも分類できず慣用音に分類されるといったことが，多数生じる．また，特に呉音については，内部に複層性があると考えられており，一元的な把握が難しい側面がある．

そこで，日本語学では，古い仏典や漢籍の訓点資料や字書類に用いられている音形を重視して，実例から帰納的に呉音・漢音を認定することが行われてきた．本研究では，この帰納的な立場による字音認定が行われている『三省堂五十音引き漢和辞典 第二版』（以下『五十音引き』）[9]を用いた．

4.2 呉音漢音音形対立表の作成

呉音漢音が同形である字については，字音体系の判定を下すことができない．例えば，「利益」の場合，語形が「リヤク」ならば「益」は呉音，「リエキ」ならば漢音と判定できるが，「利」は呉音も漢音もリなので，判定不能である．

また，二字漢語について，呉音同士で読む語を「呉音読語」，漢音同士で読む語を「漢音読語」，呉音漢音を混ぜ読みする語を「混読語」と呼ぶと，「利益」の場合「利」の字音体系は不明なので，いずれにも分類できないことになる．

このような問題を回避するため，字音体系の判定対象となる字は，「益」のように，呉音・漢音が明確な音形対立をなす字に限定することとした．『五十音引き』において一方の字音を欠く字（例：「勉」呉音なし／漢音ベン）や，一方にのみ特有の字音がある字（例：「容」呉音ヨウ・ユウ／漢音ヨウ）なども，対象から除くこととした．

この他，次のような，日本語音韻史上，呉音／漢音間の音形差が明確な形で対立していたとは必ずしも言えないようなケースがある．

- 清濁のみで対立（例：「大」呉音ダイ／漢音タイ）
- 長短のみで対立（例：「諷」呉音フ／漢音フウ）
- サ行直拗のみで対立（例：「朱」呉音ス／漢音シュ）
- t 韻尾の―チ／―ツ表記のみで対立（例：「蜊」）
- 呉音シュ／漢音シウ（シュウ）で対立（例：「周」）
- オ段長音の開合のみで対立（例：「窓」呉音ソウ／漢音サウ）

上記のような字を除外して，残った 1128 字について，呉音形と漢音形をまとめた表データである「呉音漢音音形対立表」を作成した（図 4）．

	kanji	go	go_dakuonka	kan	kan_dakuonka
378	劇	ギャク,ギャッ		ケッ,ケキ	ゲキ
379	撃	キャッ,キャク	ギャク	ケッ,ケキ	ゲキ
380	激	キャッ,キャク	ギャク	ケッ,ケキ	ゲキ
381	歌	カッ,カツ	ガツ	ケツ,ケッ	ゲツ
382	竭	カッ,カツ	ガツ	ケツ,ケッ	ゲツ
383	月	ガチ,ガツ,ガッ		ゲッ,ゲツ	
384	建	コン	ゴン	ケン	ゲン
385	軒	カン	ガン	ケン	ゲン

図 4　呉音漢音音形対立表の一部

本表の「go」「kan」という列には，『五十音引き』に記載されている呉音・漢音の，現代仮名遣いに基づく音形が直接入力されているほか，「―チ」「―ツ」「―キ」「―ク」「―フ（現代仮名遣いでは―ウ）」で終わる字（入声字）には，その促音形を，コンマ区切りで追加している．

さらに，漢語では 2 字目に濁音化（いわゆる「連濁」）・半濁音化が生じることがあるため，呉音・漢音が清音ではじまる字については，「go_dakuonka」「kan_dakuonka」という列に濁音形・半濁音形（ハ行のみ）を追加している．

促音化や連濁はある程度規則的に起こった側面があるので，プログラムによる自動判定ができるよう表に加えた．この他にも，個別的には促音の脱落（例：「乞食」コツジキ→コジキ），短呼化（「音頭」オンドウ→オンド）などがあり，このような例でも呉音か漢音かを判断することが可能な場合はあるが，『五十音引き』にその音形が掲載されない限りは，表には加えなかった．

5. 分析用プログラムの作成

5.1 プログラムの概要

コーパス検索アプリケーション「中納言」を利用すれば，CHJ の検索結果を csv 形式でダウンロードすることができる．そこで，各コーパスのコアデータにおいて，語種を「漢語」に指定して取得した検索結果（以下「コーパスデータ」）と，筆者作成の「呉音漢音音形対立表」（以下「対立表」）とを，Python の Pandas ライブラリを用いて読み込み，「コーパスデータ」に「表記タイプ」「分析用表記」「判定」の三列を新たに追加し，判定結果を格納していくという処理を，全ての行について繰り返し行うというプログラムを作成した．以下に詳述する．

5.2 表記タイプの判定と分析用表記の決定

漢語の表記タイプは，洒落本・人情本については，2.2 の図 3 で示したように，「コーパスデータ」の各行における「書字形」の文字種と「振り仮名」の有無で判断した．文字種の判別には正規表現を用い，仮名は[ぁ-ゟ]，漢字は[一-龜々〆]で判断した．ただし，「／＼」「〳〵」は，漢字に対しても仮名に対しても用いられており，「書字形」でもそのまま現れるので，文字種の判別に先立ち，(^.+?)[／〳〵]＼→ ¥1¥1 という置換処理を行った．虎明本狂言については，〈　〉がある場合のみ「振り仮名」情報を用いた．キリシタン資料については全て「仮名書き」とした．

また，「交ぜ書き無訓」については，「仮名＋漢字」の場合と「漢字＋仮名」の場合とを区別した．

分析用表記については，「漢字書き」では書字形をそのまま代入し，「仮名書き」「交ぜ書き」では語形代表表記を代入した．ただし「交ぜ書き」の漢字部分において，語形代表表記と書字形の表記が異なる場合は，その漢字部分のみ書字形の表記で置換した．例えば，書字形が「気しき」で語形代表表記が「景色」の場合，分析用表記は「気色」とした．

5.3 字音体系の判定

分析用表記が 1 字ないし 2 字の場合を対象とし，3 字以上については「対象外」とする．

1 字の場合で説明すると，まず当該字が「対立表」に含まれているかを判断し，含まれていれば，当該字の音形リストを呉音・漢音それぞれ作成する．「語形」と完全一致する音形が含まれていれば，「判定」の列に「呉音」ないし「漢音」を格納する．例えば分析用表記「客」語形「キャク」であれば，「対立表」から「客」字の呉音リスト[‘キャク’, ‘キャッ’]と漢音リスト[‘カク’, ‘カッ’]を作成し，前者に語形が含まれるため，「判定」を「呉音」とする．一致する音形がリストにない場合は「その他」とし，そもそも漢字が対立表に含まれていない場合には「不明」とした．

分析用表記が 2 字の場合は，1 字目と 2 字目それぞれについて「語形」を「対立表」と対照し，「呉音,呉音」「呉音,漢音」……「不明,その他」「不明,不明」の 16 通りの判定結果を「判定」の列に格納した．なお，2 字目に限っては，4.2 で述べた濁音化・半濁音化の音形をリストに追加した上で判定している．例えば，「請」は呉音ショウ・漢音セイだが，「勧請」という語では濁音の「ジョウ」で現れ，これも「呉音」と判定できるようにしている．ただし，もし「請

（ジョウ）」で始まる漢語があったとしても，1 字目については濁音化を考慮しないので，「不明」と判定される．

5.4 集計

以上の処理によって，新たに追加された「コーパスデータ」の「判定」列における，字単位での「呉音」「漢音」の個数と，二字漢語における「呉音,呉音」「漢音,漢音」「呉音,漢音」「漢音,呉音」の個数を集計した．この結果を元に，次節では，具体的な数値と，コーパス間での割合の変化を確認する．なお集計は，表 5 に示した基準により，「表記タイプを考慮する場合」（6.1.1 節，6.2.1 節）と，「表記タイプを考慮しない場合」（6.1.2 節，6.2.2 節）との，二通りに分けた．

表 5　二通りの集計方法

表記タイプ	考慮する場合	考慮しない場合
漢字書き付訓	○	○
漢字書き無訓	×	○
交ぜ書き付訓	○	○
交ぜ書き無訓	△(後述)	○
仮名書き	○	○

「表記タイプを考慮する場合」において，表記タイプが「交ぜ書き無訓」である場合は．字単位の集計では，仮名書きの字のみ集計し，漢字書き（無訓）の字は集計しなかった．例えば書字形が「こん日」であれば，「今（コン）」を呉音として集計するが，「日」は集計しなかった．二字漢語の集計では，対象外とした．

6. 結果と考察

6.1 呉音漢音の勢力の変遷

本節では，一字漢語・二字漢語における呉音・漢音の総数を，異なり数・延べ数双方について見ていく．

6.1.1 表記タイプを考慮する場合

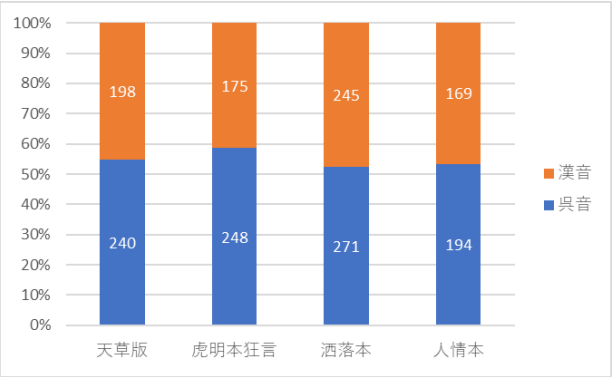


図 5　呉音漢音の勢力の変遷（異なり数・表記考慮）

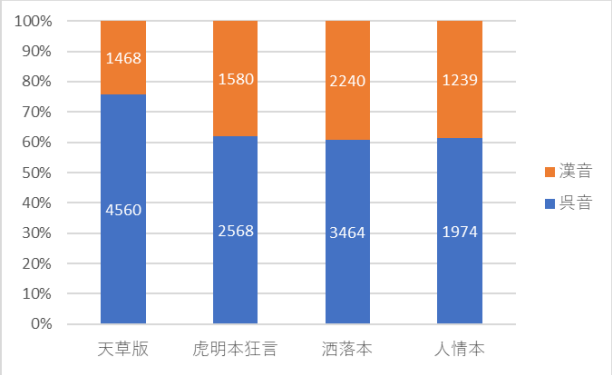


図 6　呉音漢音の勢力の変遷（延べ数・表記考慮）

異なり数（図 5）では，いずれの作品においても呉音の方が多いものの，50-60%に収まっており，漢音が伸張しているという傾向は特にうかがえない．

延べ数（図 6）で見ると，天草版では 3 倍以上の差で呉音が優勢となっているが，同時代の虎明本狂言では差は 2 倍以下に縮まっている．1 節で述べた「漢音尊重の趨勢」はうかがえるものの，虎明本狂言の呉音率が低いようにも思われる．

6.1.2 表記タイプを考慮しない場合

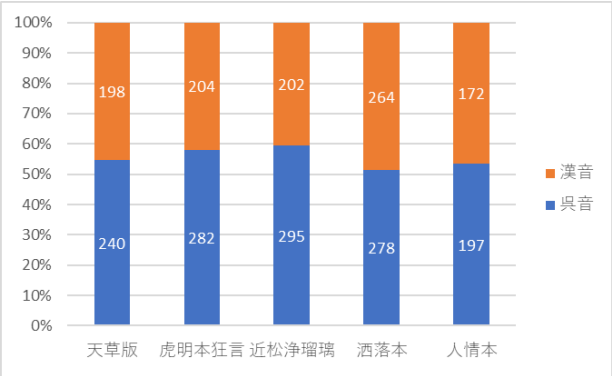


図 7　呉音漢音の勢力の変遷（異なり数・表記不考慮）

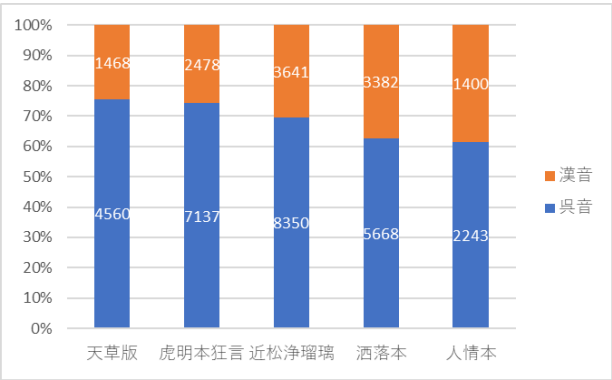


図 8　呉音漢音の勢力の変遷（延べ数・表記不考慮）

6.1.1 の表記タイプを考慮する場合と比較した場合，異なり数の在り方（図 7）は図 5 とさほど変わらない．しかし

延べ数においては違いがあり，図 6 では洒落本・人情本と同程度であった虎明本狂言の呉音率が，図 8 では天草版と同程度にまで上昇している．結果として，図 8 は江戸時代に入ると漢音が勢力を漸次的に拡大していることをよく示す推移となっている．

虎明本狂言においては，付訓例は僅か 55 例にとどまり，仮名書きないし漢字書き無訓が大多数を占めるため，図 6 と図 8 の違いは，漢字書き無訓の例が，漢音よりも呉音で読む（と判断された）字に目立つことを示している．そこで，虎明本狂言において呉音形で現れる頻度が 150 以上である字について，「漢字書き無訓」「仮名書き」のそれぞれの頻度を調べると，「漢字書き無訓」に顕著に偏る例の多いことが分かる（表 6）．例えば「日」字は，仮名書き例は「にっちう（日中）」1 例のみであり，残りの 267 例はすべて漢字で書かれる．

表 6　虎明本狂言で呉音の頻度が 150 以上の字の表記

漢字	呉音	漢字書き	仮名書き	合計
御	ゴ	579	68	647
人	ニン	619	19	638
日	ニチ・ニッ	267	1	268
今	コン	193	41	234
内	ナイ	120	69	189
二	ニ	154	4	158
下	ゲ	143	14	157
無	ム	45	109	154
細	サイ	115	37	152

表 6 から，使用頻度が多く，字画が少ない字は漢字で書かれることが多く，そのような例は漢音よりも呉音に多かったものと推測される．これらの例が呉音で読まれたとするのは校訂者・コーパス作成者の判断であるため，呉音以外で読まれた可能性のある例も含むと思われるが，多くはないであろう．このことは，語形を確定できる例に絞ると，かえって実態を歪めてしまう可能性あることを示唆する．

とはいえ，全体として「漢音尊重の趨勢」と大きく矛盾するような結果とはならず，特に延べ数にはおいては，漢音の拡大する傾向が，概ね認められる．

6.2 呉音漢音混読現象の変遷

続いて，二字漢語において「判定」が「呉音,漢音」「漢音,呉音」「呉音,呉音」「漢音,漢音」である例を抽出し，前二者（混読語）の割合（混読率）がどの程度増加しているのかに注目することで，混読現象の変遷を捉えていく．

なお，ここで示す「異なり数」は「異なり表記数」であり，同一の語彙素であっても表記が揺れる場合は別々に計上されていることに注意されたい．

6.2.1 表記タイプを考慮する場合

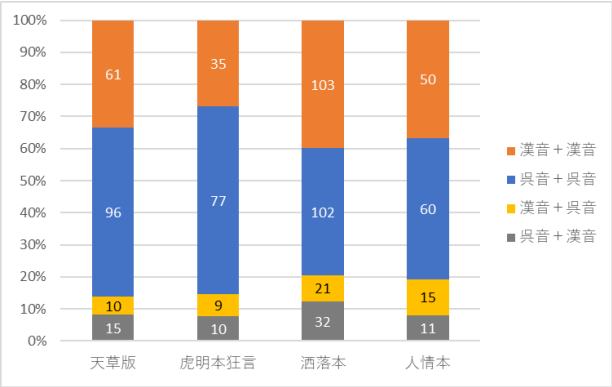


図 9　二字漢語の字音体系の変遷（異なり数・表記考慮）

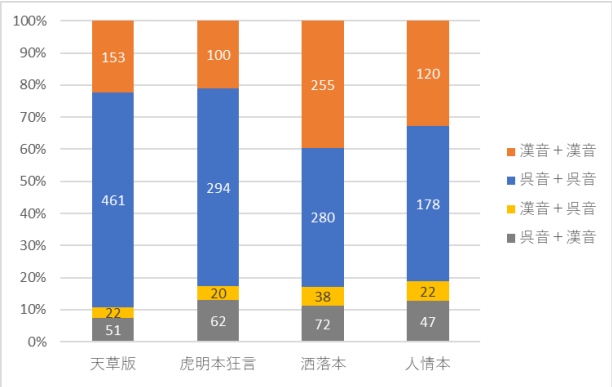


図 10　二字漢語の字音体系の変遷（延べ数・表記考慮）

混読語の割合は，異なり数（図 9）では，室町末期から江戸後期にかけ，15%程度から 20%程度に推移しており，増加が認められる．「呉音＋呉音」に対する「漢音＋漢音」の割合も増えており，室町末期では「呉音＋呉音」が優勢であるのに対し，江戸後期では拮抗している．

延べ数（図 10）でみた場合．虎明本狂言の混読率が高くなっているが，6.1.2 で述べたように，漢字書きされやすい字に呉音が多く含まれることの影響を受け，「呉音＋呉音」が実態よりもかなり少なく現れている可能性に注意する必要がある．

6.2.2 表記タイプを考慮しない場合

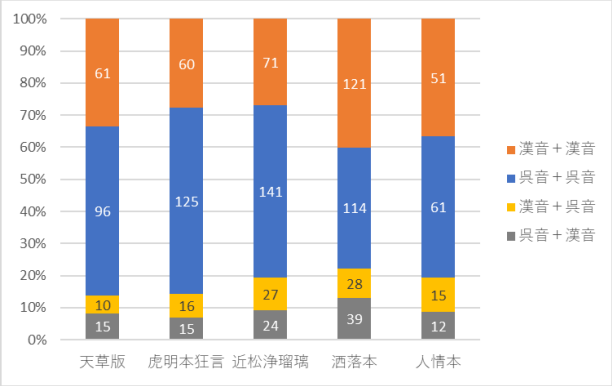


図 11　二字漢語の字音体系（異なり数・表記不考慮）

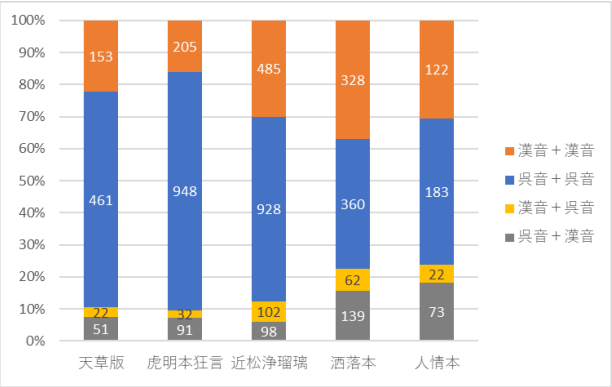


図 12　二字漢語の字音体系（延べ数・表記不考慮）

異なり数の在り方（図 11）は図 9 とさして変わらないが，「近松浄瑠璃」では混読率が既に 20%近くに達しており，江戸前期に混読語の種類が増えていることが推察される．

延べ数（図 12）では，先述した理由により，図 10 と比べて「呉音＋呉音」の例数が大きく増えており，それに従い虎明本狂言の混読率が大きく下がっている（17.2%→9.6%）．また「近松浄瑠璃」の混読率は，異なり数（図 11）より低く，12.4%に留まっており，洒落本・人情本が二割を超えているのと違いが際立っている．

異なり語と延べ語の間における変化の遅速は，江戸前期には混読語の種類が増加したものの新しい語が多く使用頻度は低かったが，江戸後期には定着し使用頻度も高まったという状況をうかがわせる．例えば呉音＋漢音である「病気」は，「近松浄瑠璃」では 14 例，「洒落本」では 5 例だが，言語量の最も少ない「人情本」では 25 例に増加する．その他の具体例など，詳しい検討は今後の課題としたい．

7. 結論と今後の課題

本研究では，CHJ の各種情報と「呉音漢音音形対立表」を照合させることで，字音体系の自動判定を行い，その結果を利用して，近世における呉音・漢音の勢力争いや，呉音漢音混読現象の拡大について分析した．前者については，「表記タイプを考慮しない場合」の特に延べ数において，漢音が拡大する傾向がうかがえた．後者については，異なり語数の割合が延べ語数に先立って増加していることなどが確認された．

作成したプログラムの不十分な点としては，2.2 に述べたように，「漢字書き二字付訓」と「漢字書き一字付訓」の区別を行っていない点や，「別漢語付訓」を処理できていない点が挙げられる．また，5.2 で述べたように，「仮名書き」の分析用表記に「語形代表表記」を用いているが，この表記は，現代語を基準に定められている「語彙素」の表記と一致することが多いため，近世当時の一般的な表記とは異なる場合がある．その結果，分析用表記に「書字形」を用いた「漢字書き」の例との間で，同一語が別々に計上されている可能性が残る．

6.1.2 節では，虎明本狂言においては，仮名書きされるか漢字書きされるかという表記上の問題が，直接の因果関係は想定しにくいものの，呉音・漢音の比率と関わることが明らかとなった．語形が確定できる例のみを集めて分析すると，かえって実態から遠ざかる危険性が想定された．

また．本研究では諸変化を時代差のみによって解釈しているが，今後，様々な位相差を考慮して慎重に考察する必要がある．この点についても，CHJ に付与されている「本文種別」や「話者情報」が活用できると思われる．

この他にも，呉音・漢音の対立がある字の一部しか検討されていない点や，三字以上の漢語や混種語・固有名詞が対象となっていない点など，種々の課題を残しているが，今後検討していきたい．

謝辞

本稿の執筆にあたっては，大向一輝氏，永崎研宣氏から数々の有益なご助言を頂きました．記して感謝を申し上げます．なお本研究は，JSPS 科研費 JP21J20167 の助成を受けております．

参考文献

[1] 柏谷嘉弘．漢語の変遷．講座日本語学 4 語彙史．明治書院, 1982, pp.63-66.
[2] 大島英之．中世における呉音漢音混読現象の展開―『色葉字類抄』と『日葡辞書』の漢語語形の比較―．計量国語学会第 65 回大会予稿集, 2021, pp.13-18.
[3] “日本語史研究資料[国立国語研究所蔵]”．https://dglb01.ninjal.ac.jp/ninjaldl/bunken.php?title=soumagaki. (参照 2022-4-27).
[4] 矢野準．人情本の漢字．漢字講座 7 近世の漢字とことば．明治書院, 1987, pp.199-218.
[5] 国立国語研究所 言語変化研究領域（片山久留美）編．『日本語歴史コーパス 室町時代編』形態論情報規程集 Ver. 1.0. https://clrd.ninjal.ac.jp/chj/morph_muromachi_v1_0.pdf.pdf (参照 2022-4-27)
[6] 大塚光信編．大蔵虎明本狂言集翻刻註解（上下）．清文堂出版, 2006.
[7] “日本語史研究用テキストデータ集” https://www2.ninjal.ac.jp/textdb_dataset/. (参照 2022-4-28)
[8] [7]に同じ．
[9] 沖森卓也編．三省堂五十音引き漢和辞典 第二版．三省堂, 2014.

ⓒ2022 Information Processing Society of Japan

正誤表

下記の箇所に誤りがございました．お詫びして訂正いたします．

訂正箇所	誤	正
5 ページ左 1-2 行目	……日本語音韻史上，呉音／漢音間の音形差が明確な形で対立していたとは必ずしも言えないようなケースがある．	……日本語音韻史上，呉音／漢音間の音形差が明確な形で対立していたとは必ずしも言えないようなケースや，音韻の合流により近世には対立が失われてしまったケースがある．
5 ページ左 5 行目	長短のみで対立	長短のみで対立（※ただし「由」など呉音ユ／漢音イウの類は除外せず判定の対象とする）
6 ページ左 2 行目	……「不明」と判定される	……「その他」と判定される
6 ページ右 図 6 中	（虎明本狂言・呉音）2568 （人情本・呉音）1974	（虎明本狂言・呉音）2569 （人情本・呉音）1975
6 ページ右 図 8 中	（虎明本狂言・呉音）7137 （近松浄瑠璃・呉音）8350 （人情本・呉音）2243	（虎明本狂言・呉音）7139 （近松浄瑠璃・呉音）8352 （人情本・呉音）2244
7 ページ左 14 行目	例えば「日」は……	例えば「日（ニチ・ニッ）」は……
7 ページ右 図 9 中	（虎明本狂言・呉音＋漢音）10	（虎明本狂言・呉音＋漢音）11
7 ページ右 図 10 中	（虎明本狂言・呉音＋漢音）62	（虎明本狂言・呉音＋漢音）63
7 ページ右 図 11 中	（虎明本狂言・呉音＋漢音）15	（虎明本狂言・呉音＋漢音）16
8 ページ左 図 12 中	（虎明本狂言・呉音＋漢音）91	（虎明本狂言・呉音＋漢音）93
8 ページ左 7-8 行目	……虎明本狂言の混読率が大きく下がっている（17.2%→9.6%）	……虎明本狂言の混読率が大きく下がっている（17.4%→9.8%）
8 ページ左 11 行目	異なり語と延べ語との……	異なり数と延べ数との……
8 ページ右 参考文献	（1 点記載漏れ有り）	国立国語研究所．日本語歴史コーパス（バージョン 2022.03，中納言バージョン 2.5.2）https://clrd.ninjal.ac.jp/chj/ (参照 2022-4-14)

ⓒ2022 Information Processing Society of Japan