

黄河・汾河流域諸方言データベース説明書

沈力・川崎廣吉・鄭弯弯・孫雪婷・星英仁

2026年3月31日

1. データベース作成の要旨

黄河・汾河流域諸方言データベースは、同志社大学言語生態科学研究センターにおける学際的研究成果の一部である。本研究の目的は、黄河流域および汾河流域における晋方言地域（北部）と中原官話地域（南部）の間に見られる言語変化の段階性を捉え、両地域間における言語伝播の実態を解明するための基礎データを提供することである。

本データベースは、日本学術振興会科学研究費助成金・基盤研究 B（海外学術調査）「黄河流域方言混合地帯における言語伝播の実態解明—地理情報科学の手法を用いて—」（研究代表：沈力、課題番号 15H05156、研究期間 2015 年～2018 年）および、中国国家社会科学基金一般項目「GIS を用いた汾河流域方言混合地帯における言語伝播の実態解明」（研究代表：沈力、課題番号 18BY042、研究期間 2018 年～2021 年）の支援を受けて構築されたものである。

調査対象地域（県・市）は、晋方言と中原官話の混合地帯に属する以下の地域である。

(1) 黄河沿岸（県以下の方言名は省略）

- a. 陝西省側（西）：神木、佳県、綏徳、呉堡、清澗、延川
- b. 山西省側（東）：興県、臨県、柳林、石楼、永和、大寧

(2) 汾河沿岸（県以下の方言名は省略）

- a. 汾河流域北部：文水、汾陽、孝義、祁県、平遥、介休
- b. 汾河流域中部：靈石、汾西、霍州
- c. 汾河流域南部：洪洞、臨汾、稷山、新絳、侯馬、襄汾、河津、万榮

本説明書では、まず黄河・汾河沿岸諸方言データベース作成の経緯を述べ、その後本データベースの利用方法について説明する。

2. 方言データベース作成の回顧

上記（1）、（2）に示したように、黄河流域の方言混合地帯には、陝西省側および山西省側にそれぞれ 6 県市、汾河流域の方言混合地帯には北部に 6 県市、中部（靈石高地）に 3 県市、南部に

8 県市があり、合計 29 の県市が調査対象となっている。当初は、方言区分の数も県市数とほぼ同程度であると予想していたが、実際には方言の種類ははるかに多様であった。本研究で扱う黄河流域の方言だけでも 51 種類（山西側 32 種類、陝西側 19 種類）に及ぶ。さらに、汾河流域では 125 種類（北部 58 種、中部〔高地〕20 種、南部 47 種）に達している。

黄河・汾河流域における言語伝播のメカニズムを解明するためには、方言データベースの構築が不可欠である。本研究では、多数の協力者による調査結果を基に、これらのデータを体系的に整理し、データベースとしてまとめた。

本データベースには、①『漢語方言調査字表』（中国社会科学院語言研究所）に基づく調査項目 3,810 字、②それらに対応する『廣韻』の音韻的特徴、③北京語における 3,810 字の声母・韻母・声調が併記されており、合計 11,430 項目に及ぶ。さらに、黄河秦晋沿岸の 51 種類および汾河沿岸の 125 種類の方言について、それぞれの声母・韻母・声調の情報を、歴史的情報および現代語情報とともに付加している。

データベース全体の文字数は、総計 757,539 字に達する。本データベースでは、例えば声母を検索する場合、以下の対応関係を参照することができる。

- 各方言における 3,810 字の声母に対応する『廣韻』の声母および北京語の声母
- 各方言における 3,810 字の韻母に対応する『廣韻』の韻母および北京語の韻母
- 各方言における 3,810 字の声調に対応する『廣韻』の声調および北京語の声調

本データベースを活用することにより、黄河秦晋沿岸および汾河流域南北における諸方言の歴史的・地域的变化を体系的に分析することが可能となるよう設計されている。

本科研の研究期間は、2015 年度-2021 年度であり、すべて予定通りに完成したが、データベースの再確認のあと、2025 年度に「同志社大学言語生態科学研究センター」のホームページで公開することになった。さらに、公開したデータを利用して、黄河秦晋沿岸・汾河沿岸諸方言のセグメント、スブラセグメントだけではなく、一層幅広い言語変化の実態の解明を期待する。

3. データベースの使用方法について

3.1 データベースの構造

黄河・汾河流域諸方言データベースは、「漢字検索」と「方言検索」の二つのモジュールから構成される。

3.1.1 漢字検索

漢字検索モジュールは、検索条件設定エリアと字表データ表示エリアの二つの部分から構成される。上

部は検索条件設定エリア、下部は条件に合致した字表データの表示エリアである。

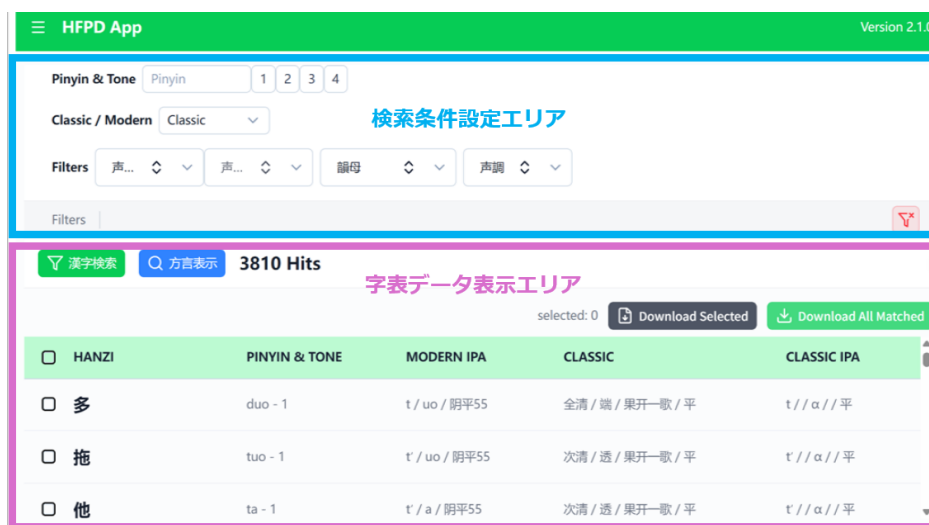


図 1：漢字検索画面

(1) 検索条件設定エリア

本エリアでは三種類の独立した検索方式を提供し、以下の通りである。

- ① 拼音・声調による検索
- ② 現代漢語国際音声記号（IPA）による検索
- ③ 中古音条件による検索

各検索方式は相互に独立し、それぞれ異なる音韻的側面から字表データを抽出するために用いられる。

(2) 字表データ表示エリア

本エリアでは、検索条件に合致する漢字項目を表示する。現在、データベースには計 3810 件の字表データが収録され、その配列は『漢語方言調査字表』に基づく。各データ項目には、漢字、拼音表記、現代漢語の国際音声記号（IPA）表記、中古音対応形式および中古音再構（王力体系参照）が順に示される。字表の右上部には検索結果のデータダウンロード機能が提供されている。

3.1.2 方言検索

方言検索モジュールは、方言点選択エリアと方言データ表示エリアの二つの部分から構成される。

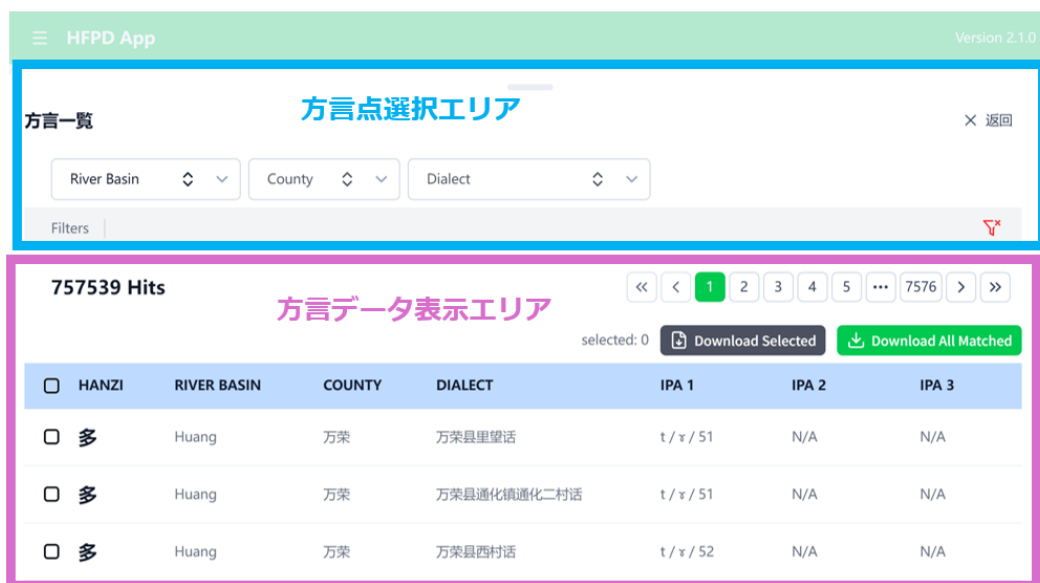


図 2 : 方言検索画面

(1) 方言点選択エリア

本エリアは方言点の段階的選択を行うためのものである。第一段階では流域区分に基づき、黄河流域・汾河流域の別を選択する。第二および第三段階では、これを基礎として行政単位情報を用い、方言点を段階的に絞り込むことで、目的とする方言点の特定を可能とする。

(2) 方言データ表示エリア

本エリアでは、選択された方言点に対応する方言データを表示する。データは 2015 年から 2018 年にかけて収集され、総計 75739 件が収録されている。各データ項目には、漢字、所属流域区、所属県名、方言名称（方言分類）および方言音の国際音声記号（IPA）表記が順に示される。データ表示エリアには検索結果のページ閲覧およびデータダウンロード機能が提供されている。

3.2 データベースの使用方法的説明

3.2.1 漢字検索

(1) 拼音・声調による検索

数値 1~4 は現代漢語普通話の第一声から第四声に対応する。検索欄に対象漢字の拼音を入力し、声調を指定した上で「漢字検索」ボタンをクリックすると、条件に合致する漢字が下の字表データ表示エリアに表示される。具体例は図 3 を参照。

声調の選択は複数選択が可能である。声調の数値ボタンをクリックすると黒色に変化し、選択された状態を示す。再度クリックすると白色に戻り、選択が解除される。

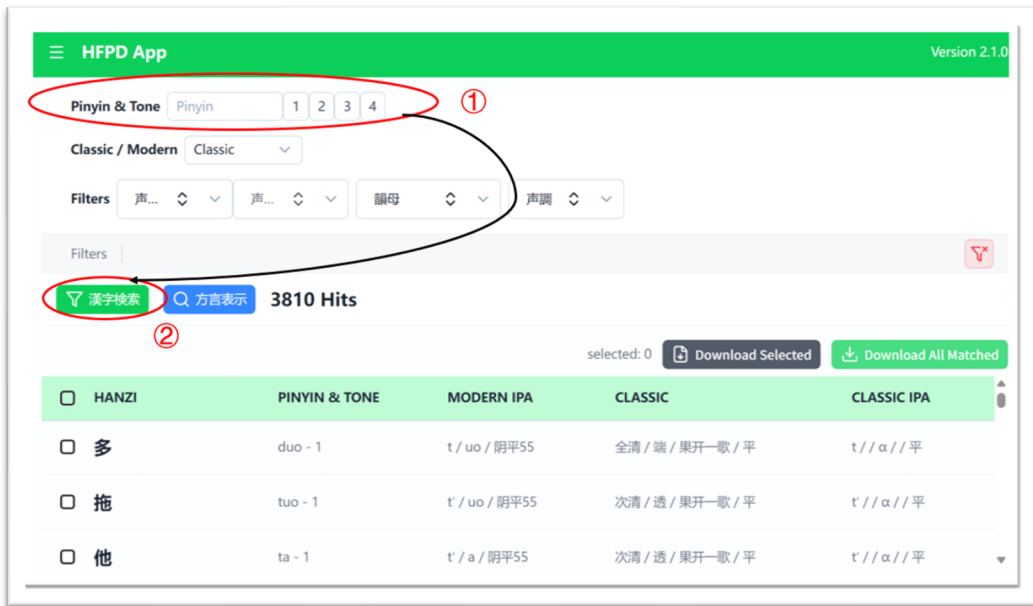


图 3 : 拼音·声调检索例

(2) 現代漢語國際音声記号 (IPA) による検索

検索条件設定エリアにおいて「Modern」を選択すると、「Filters」に声母・韻母・声調の選択項目が表示される。各項目を選択後、「漢字検索」ボタンをクリックすることで、該当する漢字が表示される。具体例は図 4 を参照。

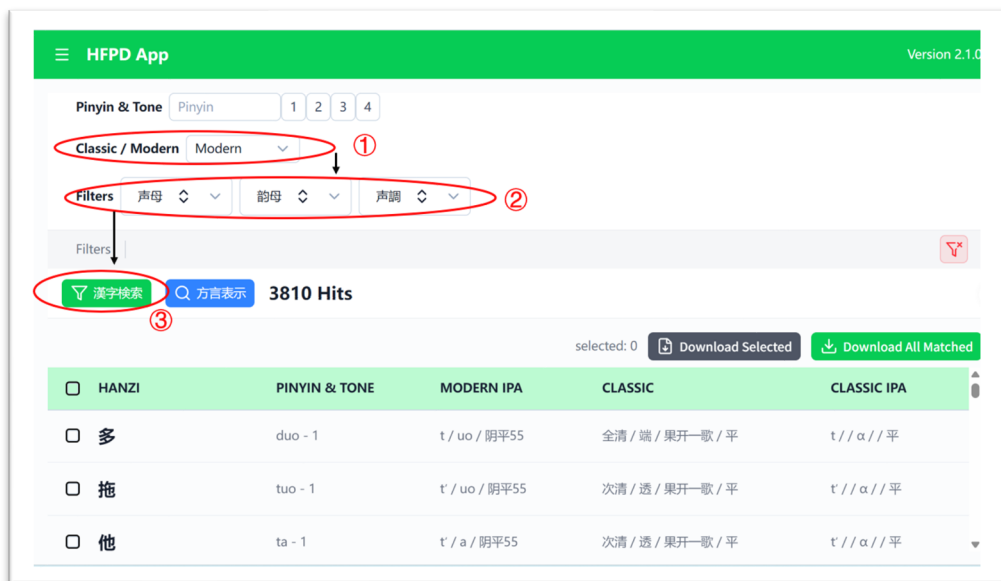


图 4 : 現代漢語國際音声記号 (IPA) 检索例

(3) 中古音条件による検索

検索条件設定エリアにおいて「Classic」を選択すると、「Filters」に声母 1、声母 2、韻母、声調の四種類の選択項目が表示される。声母 1 は声母の清濁など音韻的属性による分類であり、声母 2 は具体的な古声母名称の選択である。声母 2 の選択肢は多いため、選択の利便性を考慮し、声母 1 と声母 2 の間には制約関係が設定され、声母 2 を選択する際には事前に声母 1 を選択する必要がある。なお、多様な検索条件に対応するため、声母 1 のみを選択し声母 2 を選択しない操作も可能である。韻母および声調については独立して選択することができ、いずれか単独でも漢字の抽出が可能である。具体的な操作手順は図 5 を参照。

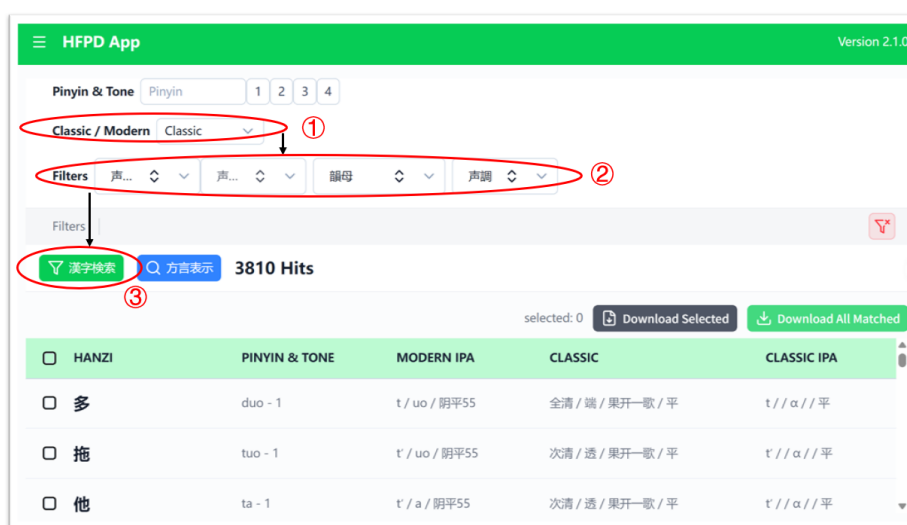


図 5 : 中古音検索例

3.2.2 方言検索

(1) 「漢字検索」から「方言検索」への遷移

条件に合致する漢字の検索が完了した後、「方言表示」ボタンをクリックすることで方言データ画面へ遷移する。この際、初期状態では抽出された全ての漢字が対象として設定される。操作ボタンは図 6 に示す通りである。

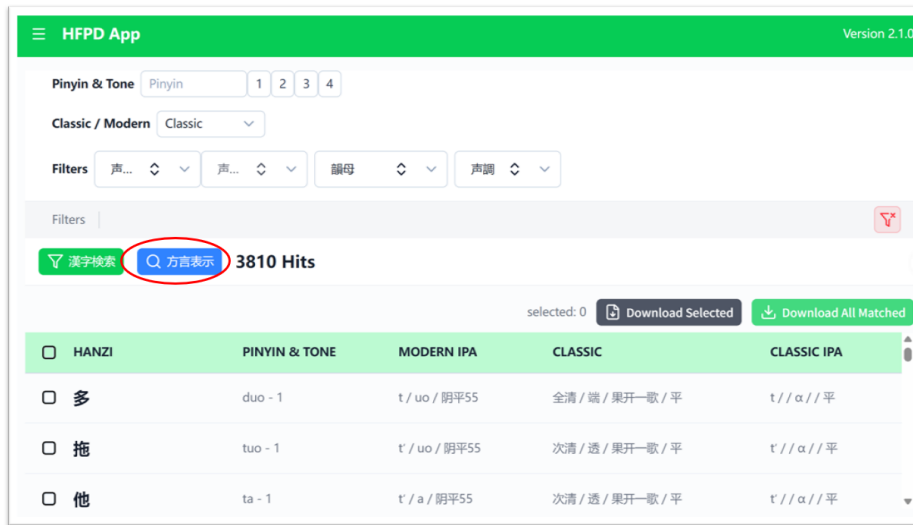


図 6：漢字検索から方言検索への遷移操作例

また、抽出された漢字に対してさらに選択を行った上で方言データを確認することも可能である。字表データ表示エリアにおいて、漢字列左側のチェックボックスを用いて対象漢字を選択する。単一選択および複数選択のいずれにも対応する。一括選択を行う場合は、「HANZI」左側のチェックボックスをクリックする。一度に選択される件数の初期値は 500 件である。500 件を超えて選択する場合は、右側のスクロールバーを操作して表示範囲を更新した後、再度「HANZI」左側のチェックボックスをクリックする。具体的な操作手順は図 7 を参照。

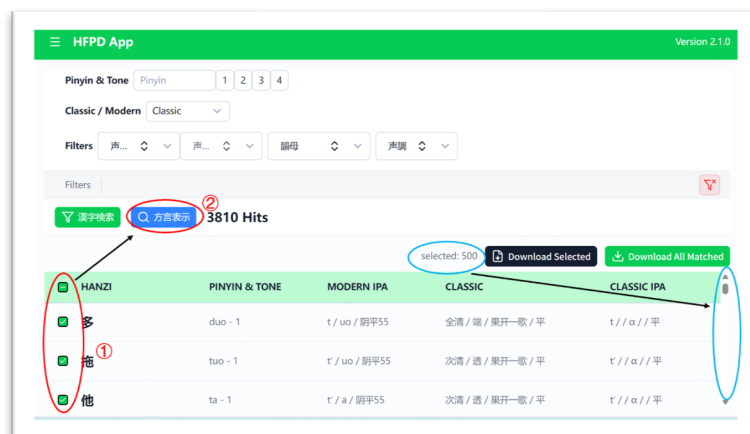


図 7：漢字の再選択による方言検索遷移操作例

(2) 方言点の選択

フィルター項目は三種類あり、方言点数が多いため、効率的な検索を行うためには左から右の順に段階的に条件を設定する必要がある。検索条件の設定が完了した後、キーボードの「Enter」キーを押すと、

下部の字表データ表示エリアに検索結果が表示される。汾河流域および黄河流域のデータは統合して検索することが可能である。操作手順については図 8 を参照。

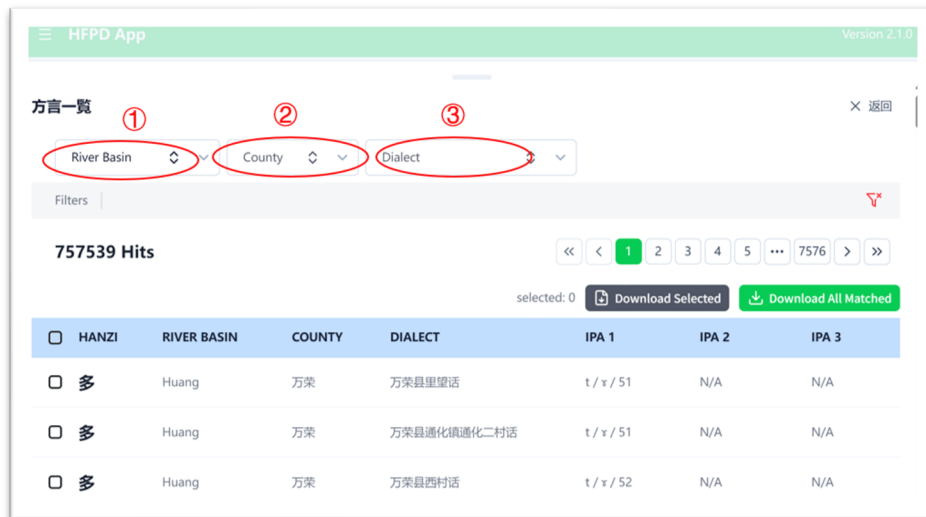


図 8 : 方言点選択操作例

3.3.その他の補足説明

(1) データダウンロード機能

漢字検索画面および方言検索画面の結果表示エリア右上にはダウンロードボタンが設置されている。選択したデータのみをダウンロードする場合は「DownloadSelected」、表示結果全体をダウンロードする場合は「DownloadAllMatched」を使用する。

(2) 注釈表示

注釈が付与された漢字の右側には緑の識別マークが表示される。当該漢字にカーソルを合わせることで注釈内容が表示される。

(3) ページ閲覧機能

漢字検索画面では、下部の字表データ表示エリアを右側のスクロールバーにより閲覧する。

方言検索画面では、字表データ表示エリア上部のページネーションにより閲覧を行う。

(4) フィルターの削除

漢字検索画面および方言検索画面では、フィルター設定エリア下部の灰色欄に選択済みの条件がリアルタイムで表示される。各条件の右上に表示される削除アイコンをクリックすることで、当該条件のみを個別に解除することができる。また、灰色欄右端のフィルターアイコンをクリックすることで、選択中のすべての条件を一括で解除することが可能である。

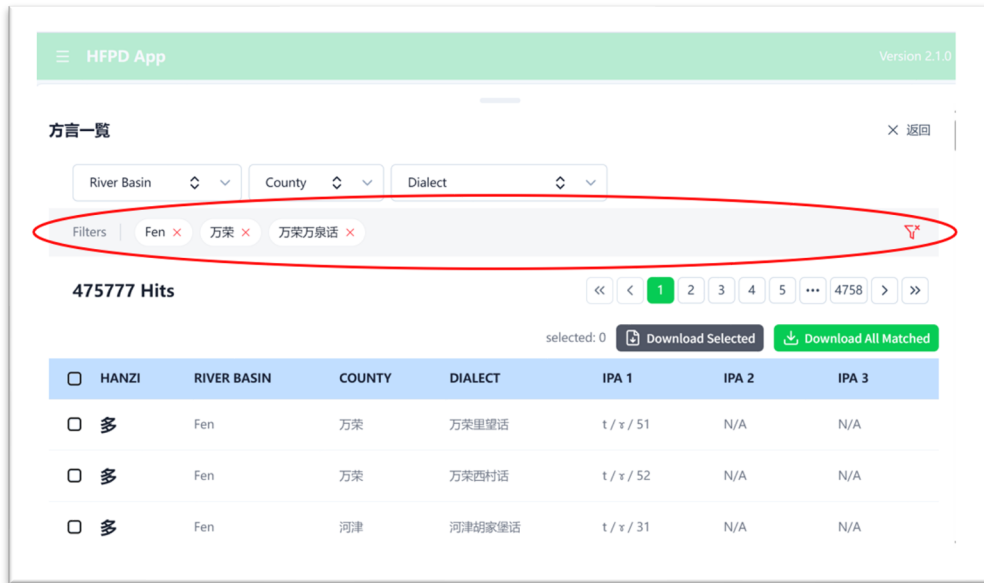


図 9 : Filters 欄操作例

(5) 「方言検索」画面から「漢字検索」画面への復帰

方言検索画面右上の「返回」ボタンをクリックすることで、漢字検索画面へ戻ることができる。画面左上のブラウザの戻るボタンは使用しないこと。これを使用するとブラウザのホーム画面に戻り、直前の検索結果が失われる可能性がある。例は図 10 を参照。

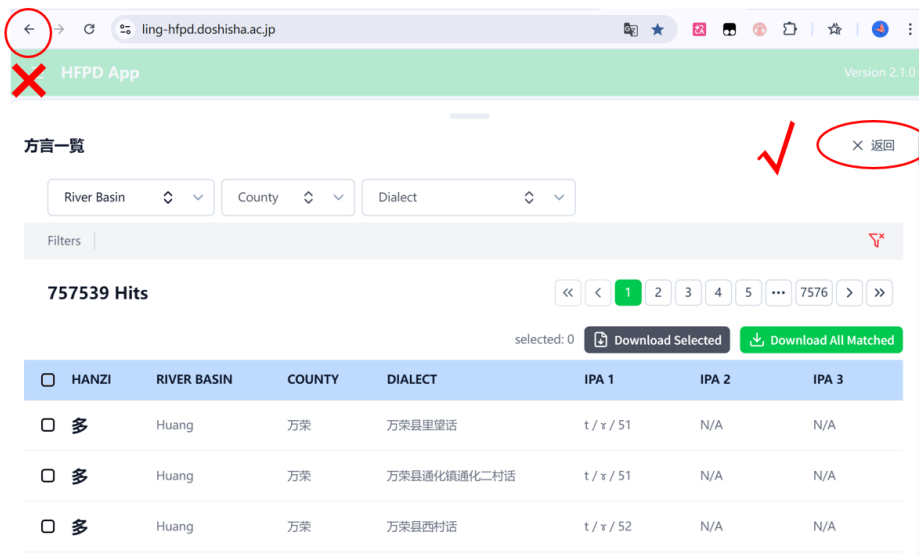


図 : 10 方言検索画面から漢字検索画面への復帰操作例