

国立国語研究所学術情報リポジトリ

『日本言語地図』データベースの概要

| | |
|-------|---|
| メタデータ | 言語: Japanese 出版者: 公開日: 2020-03-18 キーワード (Ja): キーワード (En): 作成者: 熊谷, 康雄 メールアドレス: 所属: |
| URL | https://doi.org/10.15084/00002698 |

『日本言語地図』データベースの概要

熊谷 康雄
(国立国語研究所)

1. はじめに

『日本言語地図』(以下 LAJ) の目的や基本的な考え方は、再録した「『日本言語地図』のデータベース化」(p.165) および本報告書の熊谷 (p.111) に記してある。また、『日本言語地図』データベース (LAJDB) の詳細については、別途に「『日本言語地図』データベース 解説」を作成する。ここでは上述の記述への補足を含め、簡単に概要を示す。

2. データベースの構築状況

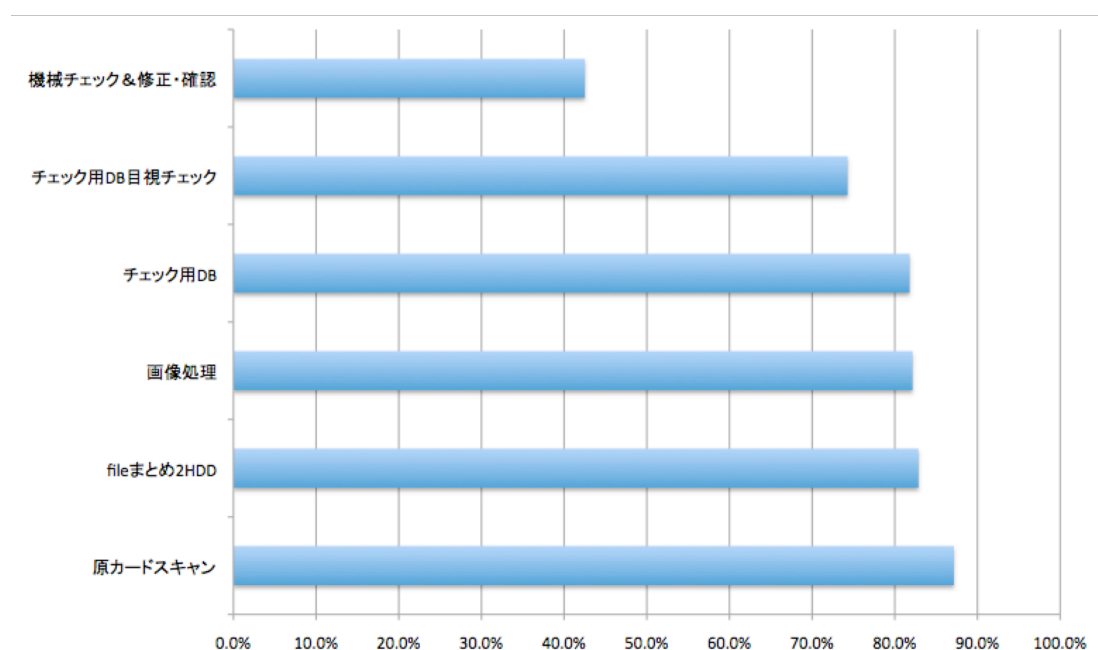


図1 『日本言語地図』データベースの構築状況
(調査項目数に対する各処理段階の済んだ項目の割合)

LAJDB は、1999 年頃から本格的に取り組みはじめた。現在の状況は上の通りであり、整備完了の項目は 119 項目、全調査項目数に対する割合は 43%、LAJ として地図化、刊行された項目 240 項目に対しては 49%である。

3 画像データベース

山カードの表面、裏面、語彙カードの表面、裏面の画像と対応する文字データをデータベース化してある。LAJDB の画像データベースの文字データの項目は、「カード番号」、「項目番号」、「地点番号」、「面記号」、「見出し語形数」、「県名」、「備考」、「項目名」、「語形 (LAJ 凡例)」、「画像ファイル名」である。「備考」には、データやデータベース化の作業に関わる注記が入力してある。

画像データベースはファイルメカを利用して作成してある。データベースの提供形態や利用方法の詳細は、後述の LAJDB の公開ページ（および別途に作成する「『日本言語地図』データベース 解説」）に記す。

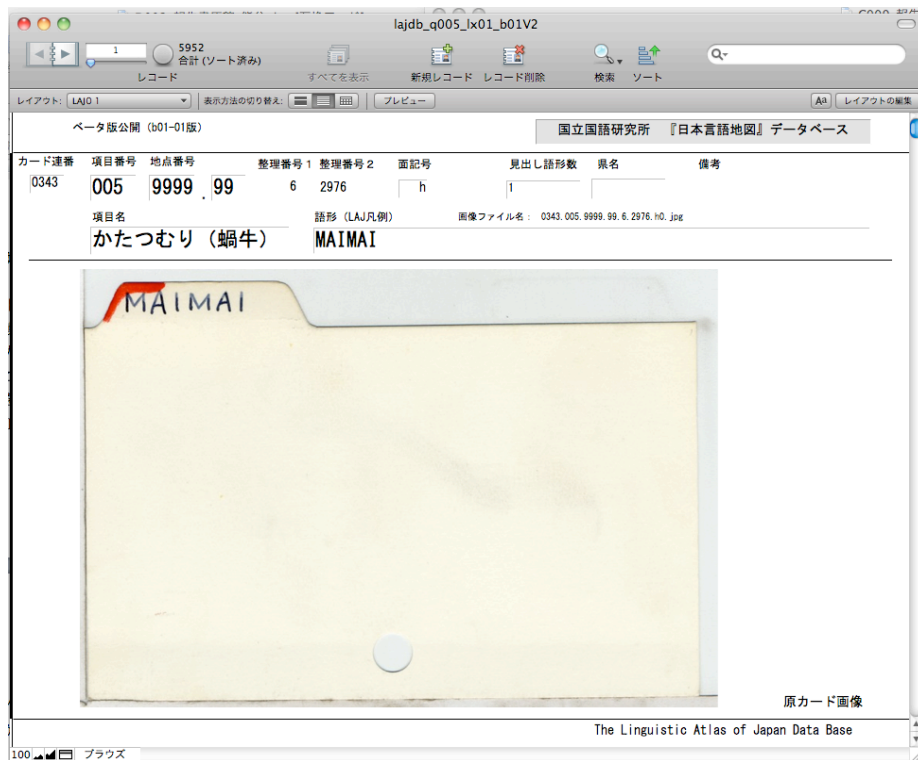


図2 山カード サンプル画面

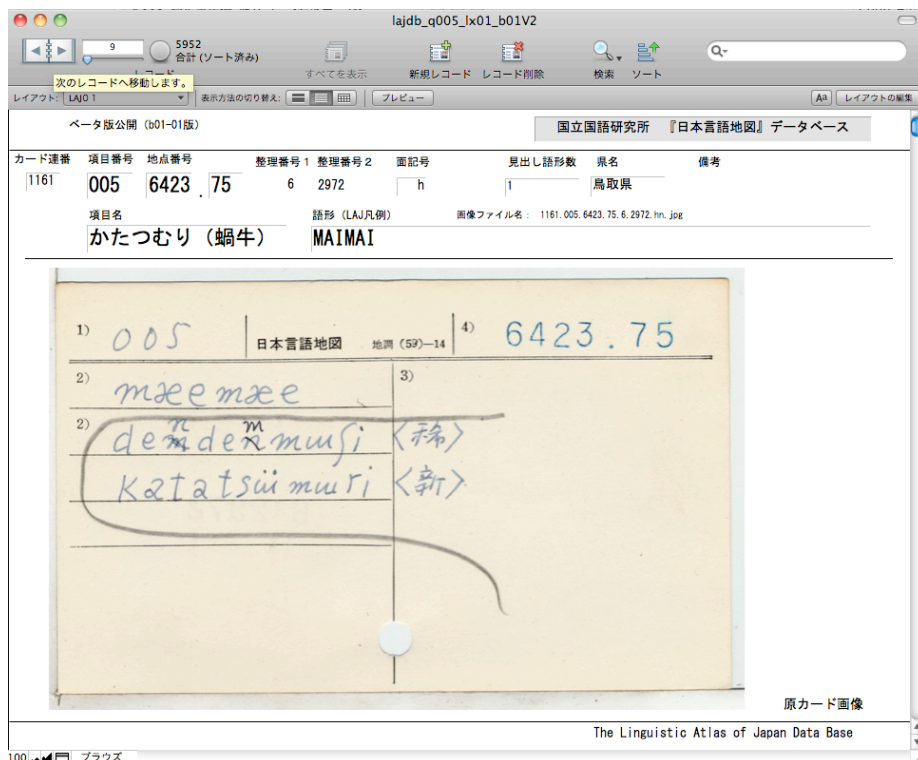


図3 語彙カード サンプル画面

4. 文字データ (EXCEL ファイル)

文字データは目的に合わせて使いやすいうように、形式1と形式2の2通りの形で作成した。

形式1：1地点、1レコード、複数回答の場合は、語形欄に複数回答

形式2：1地点、複数レコード、複数回答は回答毎に1レコード。回答欄は1語形のみ

| | A | B | C | D | E | F | G | H | I | J | K |
|----|------|---------|-------|-------|--------|--------|-----------|------|-----------------------|-------|---|
| 1 | no | item_id | loc_a | loc_b | code_a | code_b | 項目名 | 県名 | 語形 | 併用語形数 | |
| 2 | 2319 | 005 | 0228 | 96 | 6 | 2542 | かたつむり(蛸牛) | 鹿児島県 | CINNYA(N) | 1 | |
| 3 | 2320 | 005 | 0237 | 79 | 6 | 2536 | かたつむり(蛸牛) | 鹿児島県 | CINNYAMA | 1 | |
| 4 | 2321 | 005 | 0237 | 84 | 6 | 2721 | かたつむり(蛸牛) | 鹿児島県 | CINMYAA | 1 | |
| 5 | 2322 | 005 | 0238 | 55 | 6 | 2535 | かたつむり(蛸牛) | 鹿児島県 | CINNYAMA | 1 | |
| 6 | 2323 | 005 | 0246 | 48 | 6 | 2491 | かたつむり(蛸牛) | 鹿児島県 | CINMO | 1 | |
| 7 | 2324 | 005 | 0246 | 97 | 6 | 2500 | かたつむり(蛸牛) | 鹿児島県 | CINDARI | 1 | |
| 8 | 2325 | 005 | 0247 | 31 | 6 | 2499 | かたつむり(蛸牛) | 鹿児島県 | CINDARI | 1 | |
| 9 | 2326 | 005 | 0247 | 56 | 6 | 2541 | かたつむり(蛸牛) | 鹿児島県 | CINNYA(N) | 1 | |
| 10 | 2327 | 005 | 0248 | 00 | 6 | 2533 | かたつむり(蛸牛) | 鹿児島県 | CIN YAMA | 1 | |
| 11 | 2328 | 005 | 0249 | 17 | 6 | 2544 | かたつむり(蛸牛) | 鹿児島県 | CUNNYAAGUU | 1 | |
| 12 | 2329 | 005 | 0256 | 08 | 6 | 2498 | かたつむり(蛸牛) | 鹿児島県 | CINDARI | 1 | |
| 13 | 2330 | 005 | 0256 | 76 | 6 | 2503 | かたつむり(蛸牛) | 鹿児島県 | CUNDARI | 1 | |
| 14 | 2331 | 005 | 0256 | 89 | 6 | 2502 | かたつむり(蛸牛) | 鹿児島県 | CUNDARI | 1 | |
| 15 | 2332 | 005 | 0257 | 12 | 6 | 2497 | かたつむり(蛸牛) | 鹿児島県 | CINDARI | 1 | |
| 16 | 2333 | 005 | 0257 | 43 | 6 | 2496 | かたつむり(蛸牛) | 鹿児島県 | CINDARI | 1 | |
| 17 | 2334 | 005 | 0265 | 96 | 6 | 0056 | かたつむり(蛸牛) | 鹿児島県 | CINMYANDERA | 1 | |
| 18 | 2335 | 005 | 0275 | 36 | 6 | 0054 | かたつむり(蛸牛) | 鹿児島県 | CINNYANDEERA | 1 | |
| 19 | 2336 | 005 | 0275 | 97 | 6 | 2540 | かたつむり(蛸牛) | 鹿児島県 | CINNYA(N) | 1 | |
| 20 | 2337 | 005 | 0276 | 50 | 6 | 2539 | かたつむり(蛸牛) | 鹿児島県 | CINNYA(N) | 1 | |
| 21 | 2338 | 005 | 0294 | 66 | 6 | 2494 | かたつむり(蛸牛) | 鹿児島県 | CINTAI | 1 | |
| 22 | 2339 | 005 | 0294 | 93 | 6 | 2493 | かたつむり(蛸牛) | 鹿児島県 | CINTAI | 1 | |
| 23 | 2340 | 005 | 0340 | 00 | 6 | 2725 | かたつむり(蛸牛) | 鹿児島県 | CUNMYAAYUU | 1 | |
| 24 | 0001 | 005 | 0716 | 46 | 6 | 1726 | かたつむり(蛸牛) | 北海道 | DENDENMUSI | 1 | |
| 25 | 0002 | 005 | 0724 | 12 | 6 | 1725 | かたつむり(蛸牛) | 北海道 | DENDENMUSI | 1 | |
| 26 | 0003 | 005 | 0724 | 58 | 6 | 1724 | かたつむり(蛸牛) | 北海道 | DENDENMUSI | 1 | |
| 27 | 0004 | 005 | 0724 | 95 | 6 | 0468 | かたつむり(蛸牛) | 北海道 | KATACUMURI.DENDENMUSI | 2 | |
| 28 | 0005 | 005 | 0737 | 94 | 6 | 1723 | かたつむり(蛸牛) | 北海道 | DENDENMUSI | 1 | |
| 29 | 0006 | 005 | 0747 | 70 | 6 | 0201 | かたつむり(蛸牛) | 北海道 | NAMEKUZI.KATACUMURI | 2 | |
| 30 | 0007 | 005 | 0776 | 88 | 6 | 1722 | かたつむり(蛸牛) | 北海道 | DENDENMUSI | 1 | |
| 31 | 0008 | 005 | 0779 | 03 | 6 | 1721 | かたつむり(蛸牛) | 北海道 | DENDENMUSI | 1 | |
| 32 | 0009 | 005 | 0789 | 95 | 6 | 1720 | かたつむり(蛸牛) | 北海道 | DENDENMUSI | 1 | |

図4 エクセルファイル (形式1)

| | A | B | C | D | E | F | G | H | I | J | K |
|----|------|---------|-------|-------|--------|--------|-----------|------|--------------|-------|-----------------------|
| 1 | no | item_id | loc_a | loc_b | code_a | code_b | 項目名 | 県名 | 語形 | 併用語形数 | 語形(併用バターン) |
| 2 | 2319 | 005 | 0228 | 96 | 6 | 2542 | かたつむり(蛸牛) | 鹿児島県 | CINNYA(N) | 1 | CINNYA(N) |
| 3 | 2320 | 005 | 0237 | 79 | 6 | 2536 | かたつむり(蛸牛) | 鹿児島県 | CINNYAMA | 1 | CINNYAMA |
| 4 | 2321 | 005 | 0237 | 84 | 6 | 2721 | かたつむり(蛸牛) | 鹿児島県 | CINMYAA | 1 | CINMYAA |
| 5 | 2322 | 005 | 0238 | 55 | 6 | 2535 | かたつむり(蛸牛) | 鹿児島県 | CINNYAMA | 1 | CINNYAMA |
| 6 | 2323 | 005 | 0246 | 48 | 6 | 2491 | かたつむり(蛸牛) | 鹿児島県 | CINMO | 1 | CINMO |
| 7 | 2324 | 005 | 0246 | 97 | 6 | 2500 | かたつむり(蛸牛) | 鹿児島県 | CINDARI | 1 | CINDARI |
| 8 | 2325 | 005 | 0247 | 31 | 6 | 2499 | かたつむり(蛸牛) | 鹿児島県 | CINDARI | 1 | CINDARI |
| 9 | 2326 | 005 | 0247 | 56 | 6 | 2541 | かたつむり(蛸牛) | 鹿児島県 | CINNYA(N) | 1 | CINNYA(N) |
| 10 | 2327 | 005 | 0248 | 00 | 6 | 2533 | かたつむり(蛸牛) | 鹿児島県 | CIN YAMA | 1 | CIN YAMA |
| 11 | 2328 | 005 | 0249 | 17 | 6 | 2544 | かたつむり(蛸牛) | 鹿児島県 | CUNNYAAGUU | 1 | CUNNYAAGUU |
| 12 | 2329 | 005 | 0256 | 08 | 6 | 2498 | かたつむり(蛸牛) | 鹿児島県 | CINDARI | 1 | CINDARI |
| 13 | 2330 | 005 | 0256 | 76 | 6 | 2503 | かたつむり(蛸牛) | 鹿児島県 | CUNDARI | 1 | CUNDARI |
| 14 | 2331 | 005 | 0256 | 89 | 6 | 2502 | かたつむり(蛸牛) | 鹿児島県 | CUNDARI | 1 | CUNDARI |
| 15 | 2332 | 005 | 0257 | 12 | 6 | 2497 | かたつむり(蛸牛) | 鹿児島県 | CINDARI | 1 | CINDARI |
| 16 | 2333 | 005 | 0257 | 43 | 6 | 2496 | かたつむり(蛸牛) | 鹿児島県 | CINDARI | 1 | CINDARI |
| 17 | 2334 | 005 | 0265 | 96 | 6 | 0056 | かたつむり(蛸牛) | 鹿児島県 | CINMYANDERA | 1 | CINMYANDERA |
| 18 | 2335 | 005 | 0275 | 36 | 6 | 0054 | かたつむり(蛸牛) | 鹿児島県 | CINNYANDEERA | 1 | CINNYANDEERA |
| 19 | 2336 | 005 | 0275 | 97 | 6 | 2540 | かたつむり(蛸牛) | 鹿児島県 | CINNYA(N) | 1 | CINNYA(N) |
| 20 | 2337 | 005 | 0276 | 50 | 6 | 2539 | かたつむり(蛸牛) | 鹿児島県 | CINNYA(N) | 1 | CINNYA(N) |
| 21 | 2338 | 005 | 0294 | 66 | 6 | 2494 | かたつむり(蛸牛) | 鹿児島県 | CINTAI | 1 | CINTAI |
| 22 | 2339 | 005 | 0294 | 93 | 6 | 2493 | かたつむり(蛸牛) | 鹿児島県 | CINTAI | 1 | CINTAI |
| 23 | 2340 | 005 | 0340 | 00 | 6 | 2725 | かたつむり(蛸牛) | 鹿児島県 | CUNMYAAYUU | 1 | CUNMYAAYUU |
| 24 | 0001 | 005 | 0716 | 46 | 6 | 1726 | かたつむり(蛸牛) | 北海道 | DENDENMUSI | 1 | DENDENMUSI |
| 25 | 0002 | 005 | 0724 | 12 | 6 | 1725 | かたつむり(蛸牛) | 北海道 | DENDENMUSI | 1 | DENDENMUSI |
| 26 | 0003 | 005 | 0724 | 58 | 6 | 1724 | かたつむり(蛸牛) | 北海道 | DENDENMUSI | 1 | DENDENMUSI |
| 27 | 0004 | 005 | 0724 | 95 | 6 | 0468 | かたつむり(蛸牛) | 北海道 | KATACUMURI | 2 | KATACUMURI.DENDENMUSI |
| 28 | 0005 | 005 | 0724 | 95 | 6 | 0468 | かたつむり(蛸牛) | 北海道 | DENDENMUSI | 2 | KATACUMURI.DENDENMUSI |
| 29 | 0006 | 005 | 0737 | 94 | 6 | 1723 | かたつむり(蛸牛) | 北海道 | DENDENMUSI | 1 | DENDENMUSI |
| 30 | 0007 | 005 | 0747 | 70 | 6 | 0201 | かたつむり(蛸牛) | 北海道 | NAMEKUZI | 2 | NAMEKUZI.KATACUMURI |
| 31 | 0008 | 005 | 0747 | 70 | 6 | 0201 | かたつむり(蛸牛) | 北海道 | KATACUMURI | 2 | NAMEKUZI.KATACUMURI |
| 32 | 0009 | 005 | 0776 | 88 | 6 | 1722 | かたつむり(蛸牛) | 北海道 | DENDENMUSI | 1 | DENDENMUSI |

図5 エクセルファイル (形式2)

5. 凡例語形の文字データのコーディング

ASCII で入力できる LAJ の凡例欄の表記はそのままであるが、一部の音声記号による表記の部分は基本的に GAJ のコーディングと同様のシステムに従いながら、必要に応じて一部に改変を加えて入力した。必要に応じて表記に関する注記を加え、データの README 等に記載した。これらは、まとめた上で「LAJDB 解説」に記載する。

6. 基礎図・参考図の電子地図

LAJDB のデータを利用して分布図を作成、分析するために利用するため、基礎図の電子化と参考図（近代道路図、地勢図、藩領図）を電子化した。ファイル形式は GIS 等で標準的に用いられるシェープファイル形式と Adobe Illustrator 形式（AI 形式）のものを作成した。

シェープファイルは、LAJ の制作年を考慮して、日本測地系で作成した。シェープファイルは、基礎図、参考図合わせて、海岸線、都道府県支庁界、湖、mesh_小、mesh_大、調査 PT（位置調整済み）、poi2mesh（地点番号から機械的に展開した場合）、都市、道路、藩都、藩領、地勢_山系、地勢_水系、河川名を作成した。詳しくは「LAJDB 解説」で説明するので、ここでは、一部のイメージを示すのみに留める。（図 6）

AI ファイルは、シェープファイルを元に作成し、北海道、奄美・沖縄などをカットして配置したものである。LAJ の地図の配置に近づけた形で作成してある。GAJ のプラグインの利用を考えたものであるが、各調査地点には 6 桁の地点番号を持たせてあり、GAJ の白地図にあるような 2 桁の地点コードを介在させず、6 桁の地点番号で直接対応づけるようにしてある。図 7 に基礎図に対応するもの、図 8 に近代道路図のレイヤーの一部を示す。

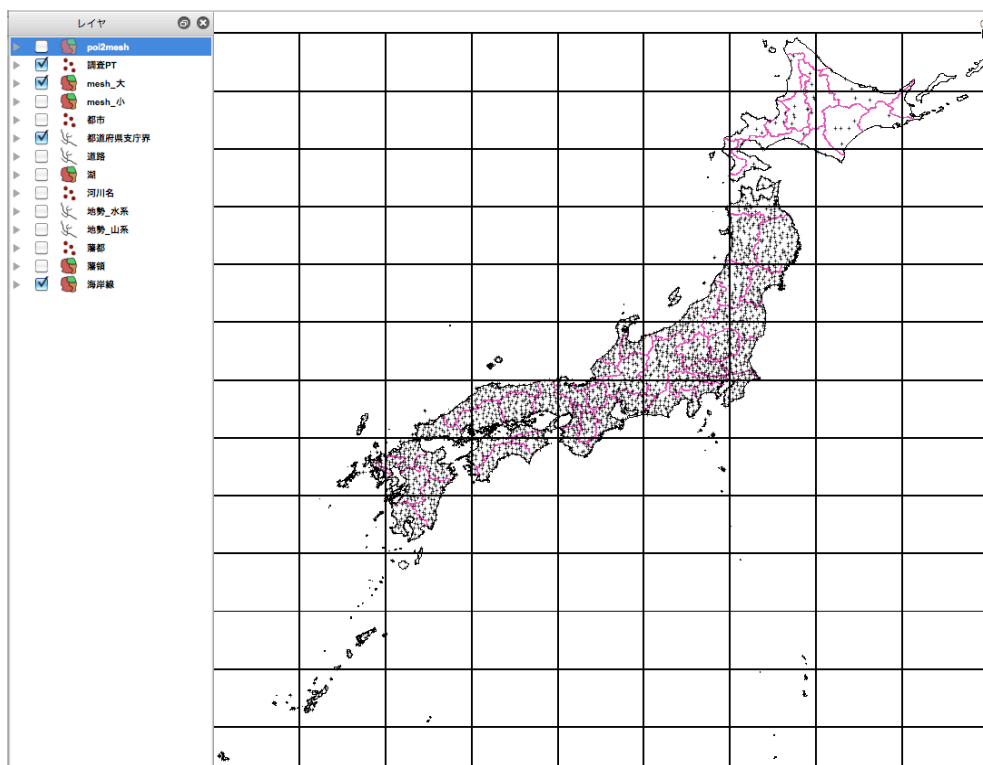


図 6 シェープファイルの表示（mesh_大、海岸線、都道府県支庁界、調査 PT）



図7 Adobe Illustrator 形式（AI 形式）の基礎図（白地図）の表示（レイヤーを選択表示）

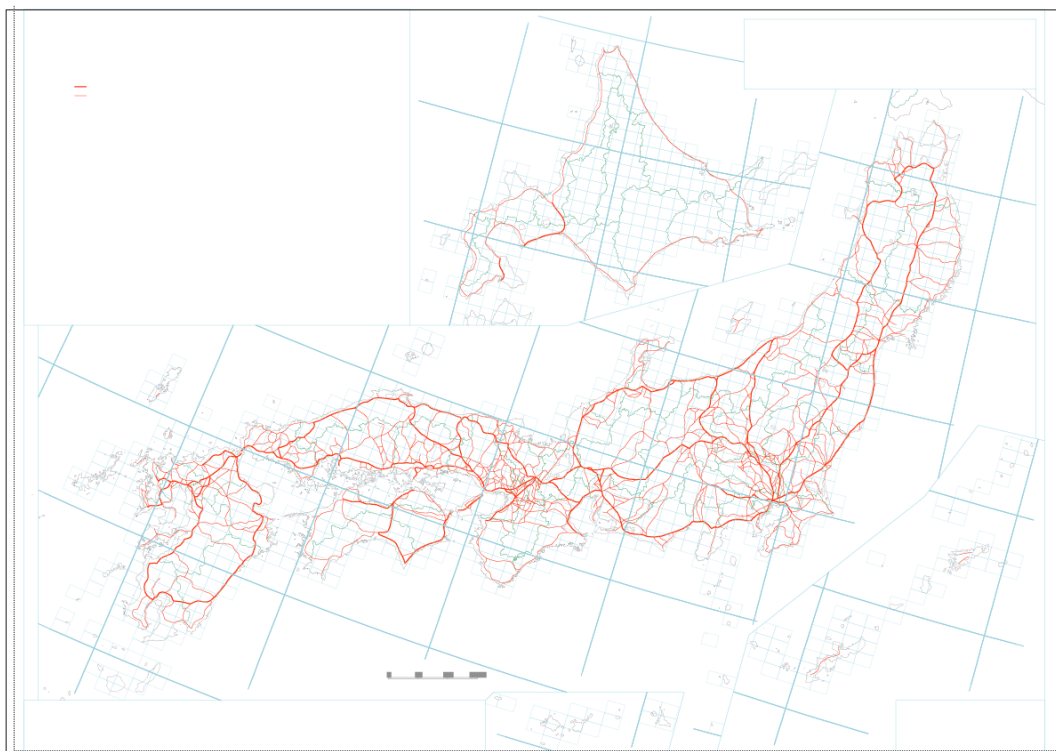


図8 Adobe Illustrator 形式（AI 形式）の近代道路図の表示（レイヤーを選択表示）

7. 関連情報のデータ

『解説』所載の表を入力したデータを作成してある。

(1) 地図番号、調査項目番号の対応表、(2) 調査地点一覧にある被調査者のフェースシート情報（個人名無）、(3) 調査項目の履歴と分類、(4) 調査地域表（後期計画で調査する地域しない地域を分けた項目と地域の関係）、(5) 行政単位別調査地点数、(5) 調査票質問文、(6) 調査者リスト

8. 公開済みの LAJ 関連のデータ

LAJ の関連データとして、過去に国立国語研究所「日本語情報資料館」のコンテンツとして作成し、現在も、国立国語研究所 HP 上で公開しているものに次のものがある。いずれも、画像を PDF ファイルで公開した。(LAJDB の公開ページからもリンクする。)

(1) 『日本言語地図』地図画像、(2) 『日本言語地図』解説、(3) 『日本言語地図』調査票、(4) 『日本言語地図語形索引』、(5) 『日本言語地図注記一覧』

9. LAJDB 処理済み 119 項目一覧

データベースとして利用可能になった 119 項目の一覧を示す。

表 1 LAJDB 処理済み 119 項目一覧（項目番号順）

| | | | |
|---------------------------|---------------------------|-----------------------|---------------------------|
| LAJ_Q001 : かまきり | LAJ_Q083 : たけうま | LAJ_Q165 : まわた | LAJ_Q233 : くるぶし (踝) |
| LAJ_Q005 : かたつむり | LAJ_Q089 : かくれんぼ | LAJ_Q166 : いと | LAJ_Q235 : すてる |
| LAJ_Q006 : なめくじ | LAJ_Q090 : おかね (貨幣) | LAJ_Q167 : きぬいと | LAJ_Q237 : おそろしい (恐ろしい) |
| LAJ_Q007 : おたまじゃくし | LAJ_Q091 : おつり (釣銭) | LAJ_Q169 : はたいと (機糸) | LAJ_Q238 : なのか (七日) |
| LAJ_Q012 : とかげ (蜥蜴) | LAJ_Q092 : かぞえる (お金を) | LAJ_Q173 : こめ (米) | LAJ_Q239 : ここのか (九日) |
| LAJ_Q031 : あたま | LAJ_Q095 : やる | LAJ_Q174 : うるち | LAJ_Q240 : ひまご |
| LAJ_Q032 : つむじ | LAJ_Q103 : きのう | LAJ_Q176 : はんまい (飯米) | LAJ_Q241 : やしやご |
| LAJ_Q036 : ものもらい | LAJ_Q104 : おととい | LAJ_Q179 : めか (糠) | LAJ_Q244 : いえ (家) |
| LAJ_Q037 : はな (鼻) | LAJ_Q105 : さきおととい | LAJ_Q182 : あぜ | LAJ_Q248 : ふすま |
| LAJ_Q038 : におい (芳香) | LAJ_Q108 : あした | LAJ_Q184 : とりおどし | LAJ_Q250 : きれいだ |
| LAJ_Q039 : におい (悪臭) | LAJ_Q109 : あさって | LAJ_Q185 : かかし | LAJ_Q252 : とうがらし |
| LAJ_Q042-7 : においをかぐ (前部分) | LAJ_Q110 : しあさって | LAJ_Q186 : じゃがいも | LAJ_Q253 : おいしい |
| LAJ_Q042-8 : においをかぐ (後部分) | LAJ_Q111 : やのあさって | LAJ_Q187 : さといも | LAJ_Q261 : おんぶする (幼児を負う) |
| LAJ_Q047 : くちびる | LAJ_Q113 : あしたのばん | LAJ_Q188 : さつまいも | LAJ_Q264 : かつく (材木) |
| LAJ_Q048 : した | LAJ_Q116 : つき | LAJ_Q190 : とうもろこし | LAJ_Q265 : かつく (天秤棒) |
| LAJ_Q051 : うすい | LAJ_Q117 : あめ (雨) | LAJ_Q191 : かぼちゃ | LAJ_Q266 : かつく (二人で) |
| LAJ_Q052 : あまい | LAJ_Q118 : つゆ (梅雨) | LAJ_Q192 : すみれ | LAJ_Q268 : いる |
| LAJ_Q056 : ほほ | LAJ_Q119 : ゆうだち | LAJ_Q194 : つくし | LAJ_Q280 : カガミ (鏡) の-G の音 |
| LAJ_Q057 : かお | LAJ_Q122 : いなずま | LAJ_Q195 : すきな | LAJ_Q282 : なす |
| LAJ_Q059 : あざになる | LAJ_Q124 : にじ | LAJ_Q200 : とげ (いばら等の) | LAJ_Q284 : とんぼ (蜻蛉) |
| LAJ_Q060 : ほくろ (小) | LAJ_Q125 : ゆき | LAJ_Q214 : おうま | |
| LAJ_Q061 : ほくろ (大) | LAJ_Q127 : こおる : 水が凍る | LAJ_Q215 : めうま | |
| LAJ_Q063 : おやゆび | LAJ_Q129 : つらら | LAJ_Q216 : こうま | |
| LAJ_Q064 : ひとさしゆび | LAJ_Q131 : ごみ (目にはいるもの一塵) | LAJ_Q217 : たてがみ | |
| LAJ_Q066 : くすりゆび | LAJ_Q132 : ごみ (掃除の対象一塵芥) | LAJ_Q218 : うし | |
| LAJ_Q067 : こゆび | LAJ_Q134 : ごみ (川のごみ一塵芥) | LAJ_Q219 : おうし | |
| LAJ_Q068 : しもやけ | LAJ_Q135 : じしん | LAJ_Q220 : めうし | |
| LAJ_Q072 : すわる (座る) | LAJ_Q148 : たく | LAJ_Q221 : こうし | |
| LAJ_Q074 : あか (垢) | LAJ_Q149 : にる | LAJ_Q222 : もうもう | |
| LAJ_Q075 : ふけ | LAJ_Q153 : ゆけ (飯の場合) | LAJ_Q223 : もぐら | |
| LAJ_Q076 : うろこ | LAJ_Q155 : すりばち | LAJ_Q228 : すずめ | |
| LAJ_Q080 : おとこ | LAJ_Q156 : すりこぎ | LAJ_Q229 : ちゅんちゅん | |
| LAJ_Q081 : おんな | LAJ_Q157 : せともの | LAJ_Q231 : はげる (禿げる) | |

10. LAJDB の公開

本プロジェクト（「大規模方言データの多角的分析」）のホームページ

(<http://www.ninjal.ac.jp/research/project/b/daikibo/>)

に公開に関する情報を掲載する。

(付記) LAJDB は科学研究費研究成果公開促進費（データベース）〔平成 13,14,15,16,17,20 年、『日本言語地図』データベース（研究代表者：熊谷康雄）〕の補助を得た。