

国立国語研究所学術情報リポジトリ

『日本言語地図』をデータベース化する： 問題と方法

メタデータ	言語: jpn 出版者: 公開日: 2021-06-25 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/10.15084/00003368

『日本語地図』をデータベース化する：問題と方法

熊谷康雄（言語変異研究領域）

1. 『日本語地図』のデータベース化 (LAJDB)

■日本語地図とは：『日本語地図』（国立国語研究所，調査1957-1965，刊行1966-1974，全6巻）は本格的な言語地理学的調査に基づく日本で初めての全国的な言語地図である。昭和30年代の全国的な方言分布が一望できる基礎資料である。

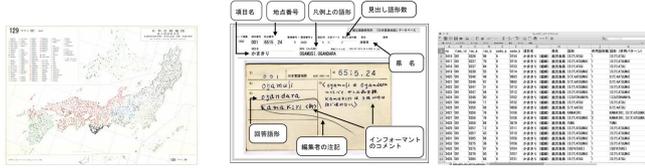


図1 地図の例(かかと) 図2 LAJDB (1) 画像DB (2)文字DB

『日本語地図』地図画像：https://mmsrv.ninjal.ac.jp/laj_map/

LAJDB 試験公開 <https://www.lajdb.org/>

■LAJを利用する上での問題：(1)『日本語地図』は印刷物：全国2400調査地点の言語地図から，個々に凡例記号を読み取って，データを作成することは困難(2)『日本語地図』は「解釈地図」：個々の地図(項目)ごとの基準により回答語形の変種を統合して凡例語形にまとめられている。個々の実際の回答語形を知るには原資料を見るが必要で，参照が難しい。(3)注記情報の記載なし(原則)：調査時に得られた付随的情報は原資料に記載：言語地図の解釈や分析に有用な，調査時に得られた被調査者のコメントや調査者の注記などの全体は原資料に記されているが地図上にはなく，参照が難しい(『日本語地図』資料の刊行は実現せず)。

■データベース化の目的・方法(『日本語地図』データベースLinguistic Atlas of Japan Database: LAJDB)：(1)目的：臨地調査による原資料のカード(約54万枚)の保存と『日本語地図』の利用の高度化(2)方法：(a)地図の語形(凡例語形)の分布情報のみならず，個々の回答語形と原資料をリンクさせ，原資料のカードと併せて，全体をデータベース化する。(b)データベース化は印刷した地図ではなく，地図編集の原資料のカードに残された編集時の秩序(地図に対応する)を利用し，原資料から作成する。[データベース化は資料校訂，資料批判に関わる]

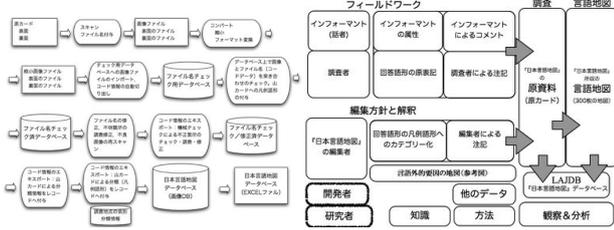


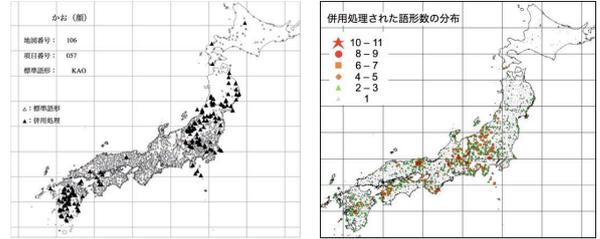
図3 『日本語地図』のデータベース化の工程(概念図) 図4 『日本語地図』データベースの開発と活用に関わる要素

新たなデジタルデータによる分析，原資料を参照した新たな視点による地図の作成や分析，過去の分析の検証，資料批判など新たな展開が可能となる。『日本語地図』を作成した研究者達が意図し，様々な仕掛けを用意しながら，時代的，技術的制約で十分に取り組みなかった情報や課題もLAJDBや様々な電子データにより改めて生かすこともできる。

©LAJDBの構築に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)、『日本語地図』データベース(研究代表者熊谷康雄)の補助を得ました。本稿には，国立国語研究所共同研究プロジェクト「大規模方言データの多角的分析」(リーダー：熊谷康雄，2010-2012)、「消滅方言の調査・保存のための総合的研究」(リーダー：木部陽子，2013-2015)、「日本の消滅危機言語・方言の記録とドキュメンテーションの作成」(リーダー：木部陽子，2016)およびJSPS科研費JP26370555(基礎研究(C)「大規模方言分布データの計量的分析の発展」，研究代表者：熊谷康雄，2014-2017)の成果を含みます。

2. 『日本語地図』データベースによる展開

■原資料のカード情報の参照(話者/インフォーマントの内省情報)



(1) 標準語形と併用処理の分布(顔) (2) 併用処理された語形数の分布

図5 併用された標準語の回答として削除(併用処理)された語形

原カード上に記された話者の内省と編集記録を参照した例。編集基準により地図に描かれていない標準語形の復元。分布がある。

■印刷物の言語地図を元には難しかった精密な地点別集計による観察

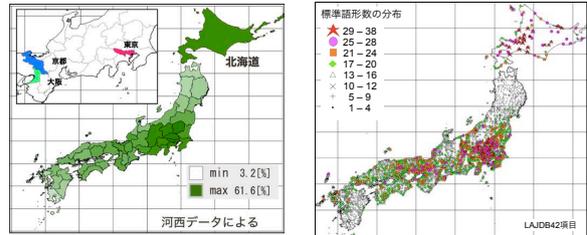


図6 県別の標準語形の分布 図7 地点別の標準語形数の分布

(地図を読み取り作成した先行研究)(LAJDBの2400の地点別集計)

■言語外的情報との関係の詳細な観察と分析(全国2400地点)

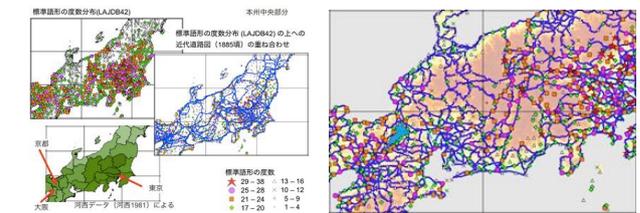


図8 標準語形数の分布と近代道路 図9 地形，現代の道路網と図との重ね合わせ(本州中央部分) 標準語形数の分布(本州中央部分)

■方言の類似度に関する量的な観察(共有度で測った類似度の例)

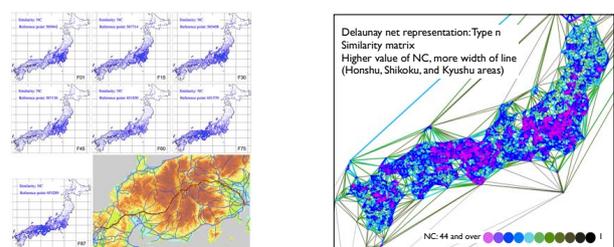


図10 街道沿い(中山道)の類似度の分布パターンの変化(赤：参照中心) 図11 隣接する調査地点間のネットワーク(delaunay net)上の類似度

文献

国立国語研究所(1966-1974)(全6巻)，東京：大蔵省印刷局。(縮刷版1981-1985)
熊谷康雄(2017)『『日本語地図』と『日本語地図』データベース：データベース化(LAJDB)による多角的分析に向けて』『方言の研究』3, pp.29-51
Kumagai, Yasuo(2016)“Developing Linguistic Atlas of Japan Database and advancing analysis of geographical distributions of dialects.”In Marie-Hélène Côté, Remco Knooihuizen & John Nerbonne (eds.), *The future of dialects*. Berlin: Language Science Press, 333-362. DOI:10.17169/langsci.b81.159