日本の言語の多様性を記録・公開する

日琉諸語のデジタルアーカイブ

加藤幹治（ROIS-DS 人文学オープンデータ共同利用センター 特任研究員）

研究の目的

日本で話されている諸言語（日琉諸語）の実態を記録保存する

日本では日本語以外の言語が話されているが、その全ての実態が明らかになっているわけはない

目的①

- 日琉諸語の文法体系を記述する

目的②

- 辞書・文法記述・談話資料をデジタル化して広く公開する

日琉諸語とは？

日琉諸語 = 日本語 奄美語 沖縄語 宮古語 八重山語 与那国語

- 日本語と琉球（奄美・沖縄）の言語の総称
- 同一の先祖から発達した関係にある

	食べる	月
日本語	食べる	月
奄美語	コーユン	ツキュ
沖縄語	カミュン	チチ
宮古語	ファウ	クスクス

どれくらい違うの？

- 文法体系が異なる
- お互いの言葉を理解できない

目的① 日琉諸語の文法の記述

概要

文法記述とは？

- 言語の音声、語の形成方法、文の形成方法などを総称して文法という
- 言語の文法体系をまめたもの = 文法書
- 日琉諸語の言語で文法書を作成中

事例

・博多弁で「太郎が走る」「犬の走る」はどうして助詞が違う？

→主語名詞句の「有生性」による「助詞の交替」現象

※科研費特別研究員奨励費 19J20370 （代表 加藤）
「北琉球徳之島伊仙方言の記述文法書作成」
※科研費研究活動スタート支援 23K18667（代表 加藤）
「記述の危機言語の地点集中的記録保存：沖縄語糸満方言の記述文法書作成」

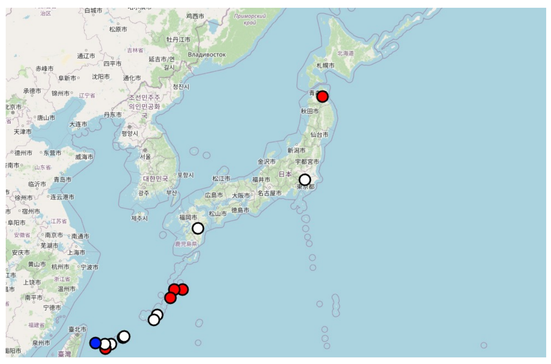
目的② デジタルアーカイブ化

LAJaR: 方言文法データセット

LAJaR:
Language Atlas of
Japanese and Ryukyuan

概要 日琉諸語の文法特徴をデータ化し、地図上に表示

新しいところ 初の日琉諸語文法データベース
それぞれの言語の特徴を視覚的に理解できる



	母音の数
日本語	5
奄美語	7
沖縄語	5
宮古語	4

※DS-JOINT #048RP2022 「日琉諸語の言語類型アトラスLAJaRの開発と分析」（代表: 筑波大学 宮川創）

研究の背景

■先行研究の課題

- 既存の文法データセットでは、日本語で書かれた文献が引用されていない
  →特定地域に特化したデータセットへの需要

データセットの特徴

■収録データ

約40地点の方言の文法特徴
-その方言にはどのような音があるか？
-その方言の語順は？
-その方言の助詞のつき方は？

■データ公開

- CC-BY 4.0（加工や再配布が容易）
- 地図上で分布を視覚化

何がわかるか？

■言語類型論への貢献

- 日本の言語はどこまで多様なのかが分かる
- 言語の文法概要を比較しやすくなる

NII
連絡先：ROIS-DS 人文学オープンデータ共同利用センター/国立情報学研究所 加藤幹治
URL：http://codh.rois.ac.jp/ Email：kanji_kato@nii.ac.jp