

日琉諸語諸方言の電子辞書の構築—沖縄語伊平屋方言の電子辞書の構築を例に—¹カルリノ・サルバトーレ²下地理則³

1. 研究背景

本プロジェクトの目的は日琉諸語諸方言のオンライン電子辞書を構築することである。後述するように、この辞書は、様々な方言のデータを csv ファイルで入力することで、1つのウェブサイト内で多数の方言の辞書項目が比較可能になる仕様になっている。その第一弾のサンプルとして、沖縄県伊平屋村で話されている伊平屋方言の辞書を作成中であり、本発表ではこのサンプルを用いたデモンストレーションを行う。

辞書の作成は、記述文法書、談話資料とともに、ある言語を記述するために必要なものとされている(Evans and Dench 2006)。近年では日琉諸語諸方言の記述研究が進み、一定の数の記述文法書が作成されてきている(例えば宮古語伊良部方言(Shimoji 2008, 2017)、八重山黒島方言(原田 2015)奄美語湯湾方言 Niinaga (2014)、沖縄語伊平屋方言(Carlino 2019))。琉球諸語諸方言では紙媒体の辞書の編纂が数十年前から進んでおり、与論方言辞典(2005)、沖縄語辞典(国立国語研究 2001)、や竹富方言辞典(2011)、宮古伊良部方言辞典(2013)のような一万以上の項目からなる大辞書が存在する。これらと異なり、本プロジェクトではオンライン電子辞書として作成していく。紙媒体の辞書に比した場合のメリットは数多い。作成する側から見ると、随時、簡単に方言データや個別項目の追加、更新ができ、紙媒体より搭載可能な情報が多くなる。使用する側から見ると、電子辞書なら検索可能であり、音声が再生できる点が挙げられる。現在、通方言的なオンライン電子辞書は他に存在するが、次節で示すように多くの問題を抱えている。本発表でまず既存のオンライン電子辞書の問題点を指摘した後、本プロジェクトで作成中の辞書の特徴を説明する。

2. 既存の辞書とその諸問題

現在オンラインデータベース形式の方言電子辞書は他に存在する。複数の方言のデータを含むものとして国立国語研究所の危機言語データベース (<http://kikigengo.ninjal.ac.jp/>)と沖縄言語

¹ 本プロジェクトは以下の支援を受けている:外国人特別研究員奨励プロジェクト番号 20F2000. 本辞書で使用しているデータの一部は次のプロジェクトの支援で収集した(2017~2019)「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」(プロジェクトリーダー 木部暢子) 科研費基盤研究(B)「比較言語学的方法による日本語・琉球諸語諸方言の祖語の再建および系統樹の構築」(研究代表者 五十嵐陽介)

² CARLINO Salvatore(九州大学/日本学術振興会外国人特別研究員) nanajuu@gmail.com

³ しもじ みちのり(九州大学)

研究センターの琉球語の音声データベース (<http://ryukyu-lang.lib.u-ryukyu.ac.jp/>)が挙げられるが、2021年3月26日時点で琉球語の音声データベースはリニューアル閉鎖中であるため、危機言語データベースを取り上げる。危機言語データベースでは多くの地点のデータを含めており、日本列島の言語の記録と保存のため重要な役割を果たしている。ただし、まだ多くの問題を抱えている。まず、検索機能が限られており、標準語、方言(IPA, カナ)と備考の情報だけが検索対象可能になっている。項目情報は少なく、標準語、カナ表記、音声表記、アクセント型に限られている。さらに、語彙は基礎語彙に限られている。例文に関しては翻訳だけがついており、グロスがない。最後に、一つの大きな問題は検索結果がダウンロードできないことである。検索結果がダウンロードできないということが、データベースがオフラインになった時には無論調べられなくなることを意味する。2020年にコロナの影響で図書館の閉鎖によって紙媒体へのアクセスも難しくなり、ネット上にある資源の重要性がより明らかになったが、琉球語の音声データベースと同様、データベースも一時的とはいえアクセス不可能になることもあり、それらに収められているデータがダウンロード可能にする必要があることが明白である。

3. 本辞書の特徴

3.1 検索画面

検索は項目に入力されているすべての情報(すなわち音素、音声、カナ表記、品詞情報、意味分類情報、アクセント型、音節構造)を検索可能(1)としている。

図 1 検索の画面

ようするに、語彙だけではなく直接アクセント型を入力して検索、または調べたい音節構造(例えば語頭で重音節で始まる語彙)が検索できる。日本語からの検索は主に漢字からである。正規表現も使用可能になっている(2)ため、特定の分節音を特定の環境という条件で検索できるなど(例えば、鼻音で終わる語彙や正門閉鎖音で始まる語彙など)、詳細な検索が可能である。検索結果は

ダウンロード可能である(3)。ユーザーは検索結果をダウンロードし、オフラインでも検索できる。ファイルはデータベースのものと csv にもとづいて.txt ファイルとして保存される。表示設定(4)では情報を無表示にすることができる。このオプションは一部の情報を必要としない、一般の方を念頭に入れている。

3.2 辞書項目の画面

辞書項目のページは図 3 のようである。音声の他、画像も搭載可能である。辞書項目情報として音素、カナ、地域、音声表記、意味、意味分類、品詞、アクセント型、音節構造がある。辞書項目の情報にクリックすれば、同じ条件の辞書項目が表示される検索結果の画面に戻る。以下はアクセント型をクリックした例を挙げている(5)。項目情報とグロスが無表示にできるオプションがある。例文にグロスを付与する(グロスについて 4.4.3 を参照)。辞書にこのようにグロスをつけるというのは、研究者がそのまま引用して使用できるメリットがある。そして、グロスから対応する辞書項目にリンクされ、例文に現れて使用者が関心を持った項目に直接アクセスできるようになっており、研究者が一つ一つ用例を探さなくても、その項目の用例が全部見られる便利さがある。例文は 3 つまで入れると想定しており、音声も搭載可能である。

6

一覧へ戻る

表示設定

All ON | All OFF

☒ IPA ☒ 意味領域 ☒ 品詞 ☒ アクセント型 ☒ 音節構造 ☒ 例文かな ☒ 例文言葉

☒ 形態素 ☒ グロス ☒ 例文訳

čibu[ru
ちぶ[る

▶ 0:00 / 0:01

🔊

地域: [沖縄伊平屋](#)

IPA: [tɕibu\[ru](#)

意味: [頭](#)

意味領域: [人体](#)

品詞: [名詞](#)

アクセント型: [ㇿ](#)

音節構造: [CV.CVCV](#)

例文1

かな: [ちぶる](#) [やめー](#)

言葉: [čiburu](#) [jamee](#)

形態素: [čiburu](#) [jam-ee](#)

グロス: [頭](#) [痛い-SEQ](#)

訳: [頭が痛い](#)

▶ 0:00 / 0:01

🔊

5

一覧へ戻る

jamu[n
やむん

地域: [沖縄伊平屋](#)

IPA: [jamu\[n](#)

訳: [痛い](#)

意味領域: [動詞](#)

品詞: [動詞](#)

アクセント型: [b](#)

音節構造: [CV.CVC](#)

▼ 1ページ辺りの表示件数100

| IPA | 意味 | 品詞 | 意味領域 | アクセント型 |
|-------------|-----|----|------|--------|
| tɕibu[ru | 頭 | 名詞 | 人体 | c |
| hana]zi[tɕi | 髪の毛 | 名詞 | 人体 | c |
| [ma:]tɕi | つむじ | 名詞 | 人体 | c |
| iri[tɕi | ふけ | 名詞 | 人体 | c |
| sa:[gi | 白髪 | 名詞 | 人体 | c |
| aci[ci | 歯茎 | 名詞 | 人体 | c |
| so:kibu[ni | 肋骨 | 名詞 | 人体 | c |

図 2 辞書項目の画面

4. ケーススタディー 伊平屋方言

4.1 伊平屋と伊平屋方言

伊平屋村は沖縄本島北部の北に位置する離島村であり、伊平屋島と野甫島からなる。田名、前泊、我喜屋、島尻、野甫という5つの集落がある。伊平屋方言は沖縄語の地域変種であり、沖縄県伊平屋村で話されている。主な話者は50代後半以上で、若年層への継承が行われていない危機言語である。

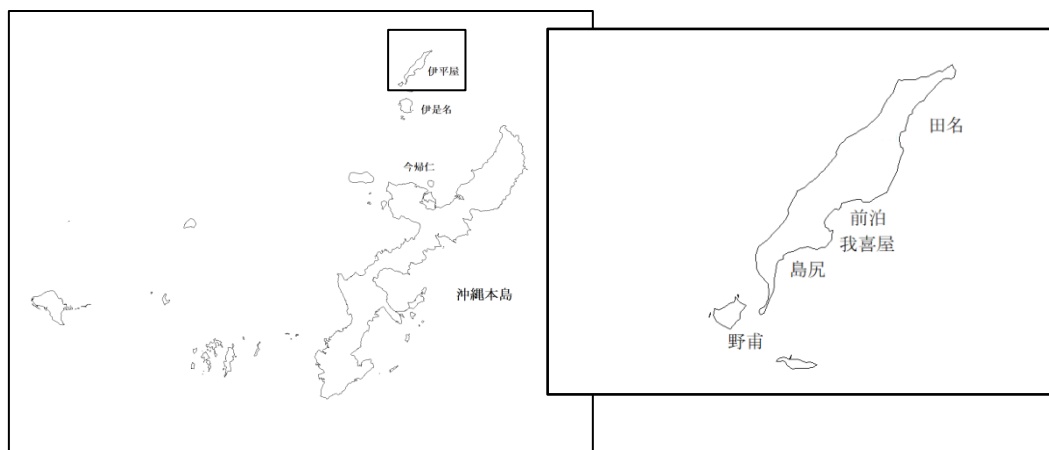


図 4 沖縄諸島と伊平屋村⁴

4.2 オンライン辞書で使用するデータの概要・調査方法

本辞書で使用しているデータは現地調査によって収集した質問調査によるデータ、談話資料から抽出したデータ、オンライン調査のデータ、文献から抽出したデータからなる。現地調査の調査は2017年7月～2019年8月の間に行われ、国立国語研究所の危機言語の共同研究プロジェクトの基礎語彙リスト、五十嵐(2019)の語彙リスト、発表者が作成した語彙リストを使用して実施された。調査協力者は田名から11人、島尻から4人、我喜屋から2人、野甫から1人であるが、この辞書では発表者の伊平屋方言の記述で主な調査の対象だった田名で収集したデータを使用する。談話資料は伊平屋村田名出身の男性3人の90分の会話からなる。オンライン調査は田名の話者1人の協力で、スカイプ及びラインを使用して行われた。文献から抽出したデータは主に地名、屋号、民俗に関する語彙である。

4.3 項目数の現状

現在は、2007項目を収集している。いわゆる基礎語彙だけではなく、地域特有の動植物、民俗、伊平屋村内外の地名を収集している。基礎語彙以外の語彙を収集する必要性はすでにかりまた(2004)に指摘されており、収集することによって言語学者だけではなく、民俗、歴史などの研究者

⁴ 地図は白地図 KenMap Ver9.2 を使用して作成した

にも役立つ情報を含むことになる。

4.4 意味分類、品詞分類と表記の基準

4.4.1 表記

伊平屋方言のカナ表記は當山(2015)及び又吉(2015)を参考にしながら、伊平屋方言の音韻体系に基づいて発表者が考えたものである。音素は筆者の記述(Carlino 2019)をもとにしている。アクセントを示す記号として上昇と下降を[]で示している。この表記法は上野(2012)をもとにしている。問題点として、文献から抽出したデータはカナ表記の問題がある。これらはカナによるものであり、表記にゆれが見られ、音声・音素表記を確かめるには話者と相談する必要がある。

4.4.2 品詞・意味分類

現時点で設定している品詞と意味分類は表 1, 2 に示す通りである。意味分類と品詞情報からも検索できるため、特定の意味分類や品詞のデータを収集することを目的とする利用者にとっても有用である。

表 1 品詞などの分類

| | | |
|-----|-----|------|
| 品詞 | 名詞 | 1297 |
| | 動詞 | 394 |
| | 形容詞 | 117 |
| | 助詞 | 42 |
| | 副詞 | 117 |
| | 連体詞 | 5 |
| | 感嘆詞 | 5 |
| その他 | 接辞 | 30 |

表 2 意味分類

| | | | | | | | | | |
|-----|-----|------|----|----|----|----|-----|----|--------|
| 名詞 | 人体 | 人間 | 貝類 | 魚類 | 昆虫 | 動物 | 時間 | 地名 | 宗教 |
| | 食 | 衣服 | 道具 | 自然 | 植物 | 住居 | 代名詞 | 数詞 | 地名(村内) |
| | 屋号 | その他 | | | | | | | |
| 動詞 | 状態 | 動作 | | | | | | | |
| 形容詞 | 大きさ | 色 | 質 | 他 | | | | | |
| 副詞 | 副詞 | 擬声態語 | | | | | | | |

4.4.3 例文のグロス

例文のグロスが5段表記である。1 段目はカナ表記、2 番目は音素表記、3 段目は形態素、4 番目がグロスで、最後 5 番目は翻訳である。具体的な例は図 3 で示している。グロスの付与は Max Planck Institute(2015)によって開発された Leipzig Glossing Rules や「下地理則の研究室・方言グロスリスト」を参考にしている。

5. 辞書データを増やす

本辞書はまず 2021 年の後半から伊平屋方言の辞書として公開する予定である。その後、発表者の所属大学である九州大学の研究者を中心に地点を増やし、そして 2022 年から外部からのデータの受け入れを開始する予定である。

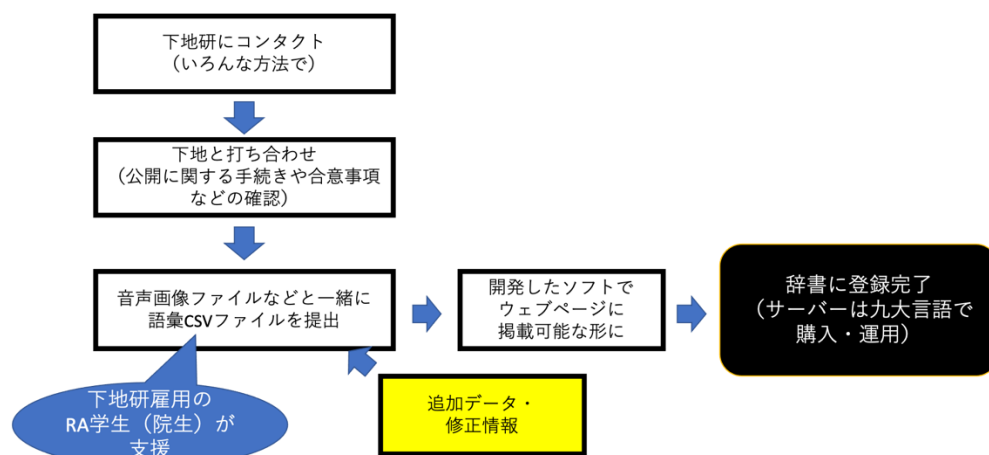


図 5. 辞書データのコントリビューションの大まかな流れ

引用文献

- 上野善道(2012)「N型アクセントとは何か」『音声研究』16(1):44-62. Evans, N., & Dench, A. (2006) Introduction: Catching Language. In F. K. Ameka, A. Dench, & N. Evans (Eds.), *Catching Language: The Standing Challenge of Grammar Writing*, 1-39. De Gruyter Mouton. 小川晋史編(2015)『琉球のことばの書き方 琉球諸語一般表記法』東京:くろしお出版. かりまたしげひさ(2004)「危機言語としての琉球語の文法研究の課題」『日本東洋文化論集』10, 57-77. Carlino Salvatore (2019) 『北琉球沖縄語伊平屋方言の文法』一橋大学, 博士論文. 菊千代・高橋俊三(2005)『与論方言辞典』東京:武蔵野書院. 国立国語研究所(編)(2001)『沖縄語辞典』東京:財務省印刷局. Shimoji, Michinori (2008) *A Grammar of Irabu: A Southern Ryukyuan Language*. Unpublished PhD dissertation, The Australian National University. Shimoji, Michinori (2017) *A Grammar of Irabu: A Southern Ryukyuan Language*. Fukuoka: Kyushu University Press. 前新透・波照間永吉・高嶺方祐・入里輝男(2011)『竹富方言辞典』沖縄:南山舎. 當山奈那(2015)「首里方言」小川晋史編『琉球のことばの書き方』151-168. 東京:くろしお出版. 富浜定『宮古伊良部方言辞典』沖縄タイムス社. Niinaga, Yuto (2014) *A Grammar of Yuwan, A Northern Ryukyuan Language*. Ph.D. Dissertation, The University of Tokyo. 又吉里美(2015)「津堅方言」小川晋史編『琉球のことばの書き方』135-150. 東京:くろしお出版. 原田走一郎(2015)「南琉球八重山黒島方言の文法」博士論文, 大阪大学. ウェブ資料:五十嵐陽介(2019)「日琉語類別語彙(2019年05月17日版)」電子データ <https://researchmap.jp/muuu4evwk-1856949/NINJAL> 危機言語データベース <http://kikigengo.ninjal.ac.jp/> 下地理則の研究室・方言グロスリスト <https://www.mshimoji.com/blank-12> 琉球語音声データベース <http://ryukyu-lang.lib.u-ryukyu.ac.jp/> 白地図 KenMap Ver9.(アクセス 2020/07/03) <http://www5b.biglobe.ne.jp/t-kamada/CBuilder/kenmap.htm> Max Planck Institute for Evolutionary Anthropology (2015) Leipzig Glossing Rules <https://www.eva.mpg.de/lingua/pdf/Glossing-Rules.pdf>