

オンライン沖縄語辞典

中村仁宣 (Hisanobu Nakamura)

March 24, 2024

自己紹介

初めていうがなびら。

わんねー...

- 名前のーじや中村なかむらひさのぶ仁宣いんでい言いちょーいびーん。
- 出身すだちや愛知県あいちけん名古屋なごや市しやいびーん。イギリスいぎりすんかい
ん暮くらちょーたくとうる事ことんあいびーん。
- 2021 年ねんに那覇なはうなてい暮くらし始はじみびたん。
- 大学・大学院うとーてい専攻せんこうさる分野ぶんぎや数学・物理すうがく・ぶつりや
いびーたん。
- 前職ぜんしつや機械学習きかくがくしゆエンジニアえんじニアやいびーたん。
- 言語くとうば習ならいしえー好しちやいびーん。

はじめに

今日、話す内容は、元々沖縄に縁がなかった登壇者が、沖縄語を学習するにあたり体験した困難と、持ち合わせていたプログラミングの知識を使ってそのほんの一部を自分なりに解決したストーリーについてです。よって、この話は

1. ^{いち}一沖縄語学習者としての視点
2. ツールの作成者としての視点

の2つの視点から書かれています。このお話が、マイノリティ言語の新規学習者に対してどのように情報発信、または学習環境の整備等をすれば良いかのヒントになれば、と思っています。

今日話す内容

1. オンライン沖縄語辞典開発の経緯
2. オンライン沖縄語辞典の紹介
3. 得られた知見
4. 今後の展望

オンライン沖縄語辞典開発の経緯

オンライン沖縄語辞典開発の経緯

開発者が沖縄語学習を開始

- 2021 年 6 月に那覇に移住。
- 2022 年 6 月に沖縄語学習を始める。
- 教材・辞書を探し始めるが、選択肢は少ない。
- Web 上で使える辞典は、カジュアル・断片的なものが多い。

オンライン沖縄語辞典開発の経緯

『沖縄語辞典』と『うちなーぐち活用辞典』

- 沖縄語辞典、国立国語研究出版、1963 年発行、約 12000 語、¥15,000～
- うちなーぐち活用辞典、編・著者宮良信詳、国立国語研究所言語変異研究領域、2021 年発行

オンライン沖縄語辞典開発の経緯

それぞれの辞書の特徴

- 『沖縄語辞典』
 - 豊富な語彙数、例文、用言活用の解説等を含む。
 - 士族・平民発音の区別。古い言葉・意味の記載。
- 『うちなーぐち活用辞典』
 - 現代の沖縄語の豊富な例文を含む。

オンライン沖縄語辞典開発の経緯

NINJAL データにたどり着く

- 紙の『沖縄語辞典』『うちなーぐち活用辞典』を入手するが、重さのため持ち運べず。
- このような語彙数・例文などが豊富なオンライン辞典があればいいのになと思う。
- NINJAL で『沖縄語辞典』の **Excel データがダウンロード可能**だと知る。
- ないのなら、『沖縄語辞典』のデータをオンライン上で簡単に検索できるようにすれば良い！
- 作るのなら、他の人にも使って貰わないと勿体ない。

オンライン沖縄語辞典開発の経緯

開発を開始

- しかし、Excel のデータをそのまま使ったのでは、一般ユーザーにとっては使いにくい所もある...

沖縄語辞典 ホーム	
<hr/>	
アサ	
発音	/asa/, アクセント◎
品詞	【名】
意味	朝。早朝にはあまり使わない。asa(sakene)〈朝晩〉, asajasa (朝夕) などの複合語に使われる。普通はasumamiti。また,asumamitiの方がasaよりも早い時間をさす感じがする。
<hr/>	
アサ	
発音	/asa/, アクセント◎
品詞	【名】
意味	阿波。+他・900。
<hr/>	

オンライン沖縄語辞典開発の経緯

『沖縄語辞典』を使う上でのハードル

音素記号を知らない利用者、初級学習者、などの一般ユーザーにとって、、、

1. 音素記号を知らないと**使う事ができない**（検索も例文を読むこともできない）
 - ・ 沖縄語は音素記号のみの表記。
2. 付録の「解説篇」を読まない**と動詞活用がわからない**
3. **初見では解読困難な知識が必要。**
4. 語彙の説明文の空白が圧縮されていて**読みづらい事**も

オンライン沖縄語辞典開発の経緯

電子情報の利点を活かす

しかし、これは、『沖縄語辞典』は元々紙の辞書として作成されたので、仕方のない事だったのではないかと思われる。紙の辞書は紙面の節約の必要性があるため、重複を省いて、変化部分のみを記す経済的な表記法にならざるをえないが、一方、電子空間上ではその制限は緩くなり、わかり易さのための冗長性を、許容範囲内で持たせる事ができる。テキスト処理と Web 技術でそれが可能となる。

オンライン沖縄語辞典開発の経緯

『沖縄語辞典』をより多くの人へ

そこで、以下の点を情報技術により克服することで、より多くの人にとって使いやすい（アクセシブルな, accessible）オンライン沖縄語辞典を作る事を目標とした。

1. 音素記号をカナと IPA で表示する。
2. 語彙説明の中の沖縄語もカナ表示する。
3. 動詞の活用表を自動で生成する

オンライン沖縄語辞典

オンライン沖縄語辞典

概要

- 収録辞書：『沖縄語辞典』（沖日、日沖）、『うちなーぐち活用辞典』
- URL：https://okinawago.app/
- 使用料：なし

沖縄語辞典 ホーム 見出しとかな表記 この辞典について

沖縄辞典・日沖辞典・うちなーぐち活用辞典

このページでは、国立国語研究所の『沖縄語辞典』と『うちなーぐち活用辞典』内の単語を、かなの見出し語で検索できます。
かな表記については、こちらのページを参照して下さい。かな表記について

検索語の入力（かな表記）

検索する辞書の選択

☒ 沖一日
☒ うちなーぐち活用辞典
☒ 日一沖

検索語と見出し語の一致条件の選択

☒ 検索語を含む
☐ 前方一致
☐ 後方一致

検索の仕方と結果

検索語の入力（かな表記）

ストゥミティ

検索する辞書の選択

- ☒ 沖→日
- ☒ うちなーぐち活用辞典
- ☒ 日→沖

検索語と見出し語の一致条件の選択

- ☒ 検索語を含む
- ☐ 前方一致
- ☐ 後方一致

検索

「ストゥミティ」の検索結果(「ストゥミティ」を含む語)

[沖]: 沖→日、[う]: うちなーぐち活用辞典、[日]: 日→沖

6件の検索結果が見つかりました。

1. [沖] ストゥミティ
2. [う] ストゥミティ
3. [沖] ストゥミティウキ
4. [沖] ストゥミティムン
5. [沖] ストゥミティウブン
6. [沖] アチャストゥミティ

項目のページ

ストウミティ

発音

音素	/sutumiti/
カナ	ストウミティ
IPA	[sutumiti]
アクセント	⑥

品詞

【名】

意味

朝。「つとめて」に対応する。スティミティ [sutimiti] ともいう。アサ [ʔasa] という語は単独ではあまり用いない。その項参照。

すとうみてい, すていみてい, していみてい, ひていみてい

品詞

【名】

意味

〈早朝〉

例文

- ・ すとうみていぬ ほーちかちとう すすいかちえー わらびぬ めーなち しーびちー くとう やたん。
 - 朝早くのほうき掃除と雑巾掛けは子どもが毎日すべきことだった。
- ・ していみていむのー なんじぐる かむが？。
 - 早朝食は何時頃食べるか。
- ・ ぶかりー なすくとう、していみてい なーらー っちゅぬ やーんかい じん あちみーが いちゅるむのー あらん。
 - 緑起が悪いので、朝早くから人の家に集金しに行くものではない。

全体文章表示比較

ʔaaki=juN⑧ (自 =raN, =ti) ㊦割れる。
 裂けて離れる。また、ひびが入る。裂け目
 ができる。ʔicanu kariti ~. 板が枯れて
 割れる。muraziNmi sjaʃiga, ʔikuçi-
 nuNkai ʔaakitan. 村の協議をしたが、
 意見がいくつにも割れた。㊦言行が食い違
 う。つじつまが合わない。矛盾する。
 taiga ʔjuru kutubanu ʔaakitooN. ふた
 りの言うことばは矛盾している。ʔariga
 ʔjuru kutoo ʔatutu sacitu caa ~. 彼
 の言うことは前とあとといつも食い違
 う。

アーキユン

発音

音素	/ʔaaki=juN, ʔaaki=juN /
カナ	アーキユン
IPA	[ʔa:kiɯɴ]
アクセント	⑧

品詞

〔自〕

意味

1. 割れる。裂けて離れる。また、ひびが入る。裂け目ができる。

イチャヌ カリティアーキユン。
 (ʔiɕanu karitiʔa:kiɯɴ.)

板が枯れて割れる。

ムラジンミ サシガ、イクチヌンカイ アーキタン。(主: ムラジンミ サシガ、イクツィヌンカイ アーキタン。)
 (muradSimmi saʃiga, ʔikutiɲuɴkai ʔa:kitas. (主: muradSimmi saʃiga, ʔikutsiɲuɴkai ʔa:kitas.))

村の協議をしたが、意見がいくつにも割れた。

2. 言行が食い違う。つじつまが合わない。矛盾する。

タイガアフル クトゥバヌ アーキトーン。
 (taiga ʔjuru kutubanu ʔa:kitooɴ.)

ふたりの言うことばは矛盾している。

アリガアフル クトアットゥ サチトゥ チャー アーキヌン。
 (ʔariga ʔjuru kutoo ʔatutu ɕaɕi ʔa:kiɯɴ.)

彼の言うことは前とあとといつも食い違う。

活用

語幹	活用	音素	カナ	IPA
基本語幹 /ʔaaki-/	否定形	ʔaakiraN	アーキラン	[ʔa:kiɾan]
連用語幹 /ʔaaki-/	連用形	ʔaakii	アーキー	[ʔa:ki:]
	過去形	ʔaakitaN	アーキタン	[ʔa:kitan]
音便語幹 /ʔaakit-/	てい形	ʔaakiti	アーキティ	[ʔa:kiti]
	継続形	ʔaakitooN	アーキトーン	[ʔa:kitooɴ]

発音テーブルの例

スン, シュン

発音

音素	/sjuN /
カナ	スン ,(士)シュン
IPA	〔suN〕 ,(士) 〔ʃuN〕
アクセント	①

カナ表記について

開発者が今まで見かけた中で、取り入れやすい表記法をできるだけ多くのバリエーションで検索できるように努力した。「どれか一つでしか検索できない」ではなく、なるべく「どの表記法でも検索できる」を目指したかった。特に、グロツタルストップの有無をどう表記するかについては、現在様々な表記法が使われているので、よく使われているであろう書籍のものを取り入れた。

発音テーブルの例 (声門閉鎖音の有無)

スン, シュン

発音

音素	/sjuN /
カナ	スン ,(土) シュン
IPA	[suN] ,(土) [ɕuN]
アクセント	①

品詞

【他・不規則】

意味

1. する。

'ウン, 'ウン, ムウン, ウウン, ヲウン

発音

音素	/ʔuN /
カナ	'ウン, 'ウン, ムウン, ウウン, ヲウン
IPA	[ʔuN]
アクセント	①

品詞

【自・不規則】

意味

1. いる。おる。(人間・動物などが) 存在する。ンカシ スイナカイ

かな表記の解説ページ

発音とかな表記について

本オンライン辞典では、国立国語研究所発行の *沖縄語辞典* の音素表記とその解説に基づき、IPA音声記号とカタカナを生成し、それを用いて単語の検索をできるようにしています。

カタナの表記は、*沖縄語辞典* が規定した『*沖縄県における「しまくとぅば」の表記について*』の沖縄語の表記や、ある程度一般に広まっていると思われる表記法を参考にしました。沖縄語のかな検査では、それぞれの単語の首文字から生成された幾通りかのかな表記を見出し語として、全ての単語の見出し語一覧表を作成し、そこから検索語とある一定の条件で一致する見出し語を検索結果として表示します。

音素記号と音声記号、カナ表記

この辞典では、音素はスラッシュ記号「/」で括って表記し、実際の音声は、括弧記号「()」の中に、IPAの音声記号を用いて表記します。

原著の『*沖縄語辞典*』では、見出し語は全て「音素記号」で表記されています。「音素(phoneme)」とは、その言語の話者に単語を構成している認識される「音」の最小単位のことです。発声方法が異なり実際の物理的な音が違っても話者には同じ「音素」として認識される場合もあります。例えば、/N/という音素は、その後に続く子音によって、/Npa/→ [mpa]、/kaNgeei/→ [kangei]、/kuNda/→ [kunda] というように音素記号とIPA音声記号が対応します。

声門閉鎖音の表記

母音の前

声門閉鎖音(glottal stop)「ʔ」(または「ʔ」)のある語頭の母音については、かな表記では無標にしています。(例: /ʔIN/ イン、/ʔaN/ アン)。

声門閉鎖音のない語頭の母音のかな表記は、*沖縄語辞典* のガイドラインはあるものの、様々な表記法が使用されているのが現状です。この事を鑑みて、全ての表記を網羅することはありませんが、可能な限り多くの表記法で検索できるようにしました。

以下は、全ての母音に共通する、声門閉鎖音のない場合の表記方法です。

- ・「ア」: 半角、または全角のクォーテーション記号(引用符)を母音の前に置く
- ・「アッア」: 前または後に同じ母音の小文字を置く
- ・「ア」: 語中の場合は無標

/u/と/i/については、「ウ/あ」、「ヨウ/う」の表記も検索使用可能にしました。

子音の前

子音「j」、「w」、「m」の前では、以下の3つの表記法で検索できるようにしています。(ʔwaa/のカナ表記を例とします。)

- ・「ʔ」: 半角の疑問符。例: ʔワア
- ・「ʔ」: IPAの音声記号。例: ʔワア
- ・「ッ」: 全角の小さいツ。例: ヲワア

長母音

長母音は「ー」で表記します。(直前の音節の母音を続ける表記法では、検索にヒットしません。例: xアアケエジュウ、○アアケエジュウ)

土族発音

/h/, /q/, /q/, /g/ (後ろの3つは/S/、/Z/、/C/とも表記します)の音素は、土族階級教育を受けた男性で実際の音声と異なっていました。このアプリでは、『*沖縄語辞典*』の解説に記載されている規則に基づき、土族発音のある場合はそのカナと音声記号を併記しています(一部で省略してあります)。土族発音の詳細については以下の一覧表と、『*沖縄語辞典*』の解説を参照して下さい。

オンライン沖縄語辞典

音素・カナ・IPA 対応表

発音表記一覧表

	a	i	u	e	o	-
	音素 /a,a,'a/	/i,i,i,'i/	/u,'wu,u,wu,'u/	/e,e,'e/	/o,o,'o/	
	IPA [a]	[i]	[u]	[e]	[o]	
	カナ 'ア,'ア,アア,アア	'イ,'イ,イイ,イイ,イイ,キ	'ウ,'ウ,ウウ,ウウ,ヲウ	'エ,'エ,エエ,エエ	'オ,'オ,オオ,オオ,オオ	
?	音素 /ʔa/	/ʔi/	/ʔu/	/ʔe/	/ʔo/	
	IPA [ʔa]	[ʔi]	[ʔu]	[ʔe]	[ʔo]	
	カナ ア	イ	ウ	エ	オ	
k	音素 /ka/	/ki/	/ku/	/ke/	/ko/	
	IPA [ka]	[ki]	[ku]	[ke]	[ko]	
	カナ カ	キ	ク	ケ	コ	
kj	音素 /kja/		/kju/		/kjo/	
	IPA [kja]		[kju]		[kjo]	
	カナ キヤ		キユ		キヨ	
kw	音素 /kwa/	/kwi/		/kwe/		
	IPA [kwa]	[kwi]		[kwe]		
	カナ クワ	クイ		クエ		
g	音素 /ga/	/gi/	/gu/	/ge/	/go/	
	IPA [ga]	[gi]	[gu]	[ge]	[go]	
	カナ ガ	ギ	グ	ゲ	ゴ	
gj	音素 /gja/		/gju/		/gjo/	
	IPA [gja]		[gju]		[gjo]	
	カナ ギヤ		ギユ		ギヨ	

カナ表記で参考にした書籍、文献

- ・ 白水社『沖縄語の入門』西岡敏、仲原穰著
- ・ 研究社『初級沖縄語』花蔭悟著
- ・ 研究社『沖縄語辞典』内間直仁、野原三義編著
- ・ 沖縄県『沖縄県における「しまくとうば」の表記について』
- ・ しまくとうば普及協議会の教材

語彙説明文と沖縄語例文の分離

sjuN④ (他・不規則) ⑤する。'jamatuguci
〜. 日本語を話す。bincoja siibusikoo
neeNkutu saN. 勉強はしたくないからし
ない。siijuusjuN. できる。なしうる。⑤
(強意の補助動詞として) 'jumidu sjuru.
読むのだ。ʔaee sani. ありはしないか。

意味

1. する。

ヤマトグチスン. (土: ヤマトグチシュン.)
[ʔamatugutʃisun. (土: ʔamatugutʃifun.)]

日本語を話す。

ビンチョーヤ シーブシコーネンクトゥ サン.
[bintʃo:ja ʃi:buʃiko:ne:nktu san.]

勉強はしたくないからしない。

シーユースン. (土: シーユーシュン.)
[ʃi:ju:sun. (土: ʃi:ju:fun.)]

できる。なしうる。

2. (強意の補助動詞として)

ユミドゥ スル. (土: ユミドゥ シュル.)
[ʔumidu suru. (土: ʔumidu furu.)]

読むのだ。

アエー サニ.
[ʔae: sani.]

ありはしないか。

分離の処理の難しさ

- 「ほとんど」の例文を分離できているが、一部できていない箇所もあるかもしれない。
- 正規表現による処理を行っているため、正確な処理は対象データに対する網羅的な知識が必要。
- エラーを見つけた場合はご報告を。

オンライン沖縄語辞典

動詞活用表の例

ʔaaki=juN① (自 =ran, =ti)

活用

語幹	活用	音素	カナ	IPA
基本語幹 /ʔaakir-/	否定形	ʔaakiraN	アーキラン	[ʔa:kiɾan]
連用語幹 /ʔaakij-/	連用形	ʔaakii	アーキー	[ʔa:ki:]
音便語幹 /ʔaakit-/	過去形	ʔaakitaN	アーキタン	[ʔa:kiɾan]
	てい形	ʔaakiti	アーキティ	[ʔa:kiti]
	継続形	ʔaakitooN	アーキトーン	[ʔa:kiɾo:n]

sjuN① (他・不規則)

活用

語幹	活用	音素	カナ	IPA
基本語幹 /s Qs-/	否定形	saN	サン	[san]
連用語幹 /s sj-/	連用形	sii	シー	[ʃi:]
音便語幹 /sj-/	過去形	sjaN	サン(土:シャン)	[san] (土: [ʃan])
	てい形	Qsi	ッシ	[ʃʃi]
	継続形	sjooN	ソーン(土:ショーン)	[so:n] (土: [ʃo:n])

うちなーぐち活用辞典

ヤンムチ

発音

音素	/jaNmuci /
カナ	ヤンムチ
IPA	[ˈjamutʃi]
アクセント	①

品詞

【名】

意味

鳥もち。小鳥や昆虫をとらえるもち。ガジマルの木からとる粘液で作る。

やんむち

品詞

【名】

意味

〈鳥モチ〉

例文

- やんむちえー きーぬ かー ひらべったーいしんかい うちきてい、げんの一さーに わっくわさーに どうるどうる ないるあたい ちャー ちちちし ちゅくいん。
 - 鳥モチは木の皮を平たい石に置いて、金槌で解きほぐしてどろどろになるまでずうっと突っついて作る。
- んかしゅー やんばるだきぬ すーらんかい やんむち たっくわさーに、くりし さんさなー、なーびかちかちー、しゅーぐわー、はーべーるー、そーみなー、ある うっさぬ いちむしゅー どうとーたん。
 - ヤンバル竹の先に鳥モチくっ付けて、それでクマゼミ、アブラゼミ、バッタ、蟻、メジロ、すべての昆虫を探っていた。

「この辞書について」のページ

この辞典について

本オンライン辞典のデータは、国立国語研究所が発行している以下の辞書データを元に作られています。

- ・沖縄語辞典
- ・うちなーぐち語彙辞典(宮良信洋著)

かなの表記は、沖網星が策定した

- ・「沖網星における「しまくとぅば」の表記について」

を一部参考にしました。

リンク集

ソースコード

<https://github.com/hissanovao/okinawago-dictionary-app.git>

沖縄語辞典の電子データ

https://github.com/hissanovao/okinawago_dictionary.git

Twitter アカウント

@okinawago_app

このサイトのURLを以下のQR Codeからスキャンできます:



まとめ1

ユーザビリティ向上のための施策

1. シンプルで、一貫性のあるページデザイン
2. カナ表記の解説ページの作成
3. カナで検索可能に。
 - ・ 可能な限り多様な表記法に対応
4. 土族発音も併記
5. 語彙説明文内の沖縄語の例文の分離。
 - ・ 語彙説明文内の沖縄語の例文もカナで表示可能に。
 - ・ 改行を挿入する事で、視認性も向上。
6. 動詞活用表の生成。

到達しやすさへの施策

1. シンプルな URL

- URL を覚えてもらえる。チラシに乘せる時にコンパクト。

2. twitter アカウントとチラシ作り

- SNS での発信と対面での配布、店舗に配置

3. QR-code 掲載

- スマートフォンなどで、隣の人にすぐに URL をシェアなどできる

得られた知見

アクセシビリティ (accessibility) の重要性

アクセシビリティ:= アクセスのしやすさとは...

- 見つけやすさ (searchability)
 - 見つけやすい場所にある。
- 入手しやすさ (reachability)
 - なるべく多くの人が常識的な手段・コストで手に入れられる事。
- 使いやすさ (usability)
 - 使いこなすのに特別な訓練を必要としない事。
- **利用する側からの視点が重要。**

誰にとって利用しやすいデータか

- 一般ユーザーにとって
 - インターネットなら、検索にヒットしやすい、雑多な情報に埋もれていない。
 - 紙の書籍なら、街の本屋で購入可能である、図書館にある、高くない等
 - 利用料が相応、または無料。
 - スマホ・PC 上で、カナで検索できる等
- 開発者にとって
 - データがダウンロード可能で、ある程度構造化 (Excel, CSV, JSON, etc) されている。
 - ソースコードがレポジトリで公開されている。

今後の展望

アプリの改良・充実

- 辞書機能の充実 (活用表の充実、お気に入り語彙、履歴、モバイルアプリ化等)
- 学習サポート機能 (動詞活用練習、語彙暗記テスト等)
- 他辞書の取り込み (さらなる沖縄語、琉球諸語の辞書の追加)
- Wiki 作成
- 共同開発者・支援

う ち くい
終わいまでい聞ち呉みそーち
にふえー
いっぺー御拝でーびる。

参考文献

- ・ 白水社『沖縄語の入門』西岡敏、仲原穰著
- ・ 研究社『初級沖縄語』花蘭悟著
- ・ 研究社『沖縄語辞典』内間直仁、野原三義編著
- ・ 沖縄県『沖縄県における「しまくとぅば」の表記について』
- ・ しまくとぅば普及協議会の教材：学習ツールダウンロードのページ