国立国語研究所学術情報リポジトリ

Possibilities of Meta-research on Japonic Descriptive Linguistics : Evaluation of 40 Years of Lexical Research on Southern Ryukyuan and Its Implications for Future Methodology

メタデータ	言語: jpn 出版者: 公開日: 2022-07-27 キーワード (Ja): キーワード (En): 作成者: 麻生, 玲子, セリック, ケナン, 中澤, 光平, ASO, Reiko, CELIK, Kenan, NAKAZAWA, Kohei メールアドレス: 所属:
URL	https://doi.org/10.15084/00003567

国立国語研究所論集 (NINJAL Research Papers) 23: 75–98 (2022)
ISSN 2186-1358

日琉諸語の記述言語学を対象としたメタ研究の試み
――南琉球諸語の過去 40 年間の語彙研究の評価と課題――

麻生玲子[a]　　セリック・ケナン[b]　　中澤光平[c]

[a] 名桜大学／国立国語研究所 共同研究員
[b] 国立国語研究所 研究系 非常勤研究員
[c] 信州大学／国立国語研究所 共同研究員

要旨

本論文では，過去 40 年間の南琉球の語彙研究を事例に，従来の調査方法や成果に対して詳細な評価を行い，それに基づき，日琉諸語を対象に多地点で詳細な語彙研究を行うという目的を達成するための効果的な研究方法について論じる。過去 40 年間に行われた南琉球の語彙研究成果を収集および評価した結果，研究者と母語話者が協力して行うハイブリッド型研究が，質と量の両面から最も有効な研究手段であることが分かった。このため，今後の語彙研究にとってハイブリッド型の研究形態を積極的に活用していくことが大きな可能性を秘めていると主張する。一方で，南琉球諸語は消滅危機言語であるため研究期間に制限があるにもかかわらず，語彙研究が行われている地点には激しい偏りが存在していることも指摘する。この問題に対して，ハイブリッド型研究であっても面接調査のみに頼ることは非現実的であることを明らかにし，語彙収集に関わる作業を一部遠隔化することで解消できると指摘する。以上の結果を踏まえ，我々が実施している事例を参照しながら，作業を細分化・分担し，各自が居ながらにして作業を効率的に行うハイブリッド遠隔型の語彙研究を提案する *。

キーワード：南琉球，語彙研究，meta-research（メタ研究），フィールド調査法

1. はじめに

我々は，ある問いを解明するために研究を行うわけであるが，その問いが研究方法そのものである場合がある。これはメタ研究（meta-research），すなわち研究そのものの研究と呼ばれ，その分野で採用されている研究手法等を対象とした研究である（Ioannidis 2018）。これまで特に医学や心理学といった研究分野で発展してきた（Ioannidis et al. 2015 など）。メタ研究は，当該研究分野で採用されている研究手法が，問い・目的を解決・達成するために適切かどうかを客観的に判断し，当該手法を改善する目的で行われる。時代や環境によって変化してきた研究手法の全体像を把握することで，問題点の発掘および洗い出しを行い，その結果，明らかになった問題点に

* 本研究は以下の助成を受けている：科研費 18K12390（日本の消滅危機言語を対象とした大量の言語資料収集・蓄積方法に関する基礎研究），19H05353（南琉球八重山諸語における伝播過程の解明と言語系統樹の構築），19K13174（南琉球宮古語の語彙体系の多様性を探る：通方言的な音声付の語彙データベースの構築），21H00353（南琉球諸語を対象とした言語変化モデルの構築と系統樹への応用）。本論文の一部は国立国語研究所の共同研究プロジェクト「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）の研究成果であり，2021 年 12 月 7 日に開催された第 232 回 NINJAL サロンにおける発表を基にしたものである。なお，インフォーマントをご紹介くださった川平村の高嶺善伸様，竹富町教育委員会の皆様，波照間永吉先生（名桜大学），査読者，そして遠隔調査にご協力いただいた各地のインフォーマントの皆様に，心より感謝を申し上げます。

対する対策の検討が初めて可能となるのである。

言語研究の分野におけるメタ研究は，第二言語習得といった応用言語学の分野では 2000 年代から活発に行われ始めている一方（Plonsky and Oswald 2012），基礎研究の分野では少数である。メタ研究はあらゆる研究分野に対して適用できるため，本研究が対象とする記述言語学の分野でもその実施が可能であるが，そのような研究は管見の限り見当たらない。特に日琉諸語を研究対象にする場合，伝統的な方言の多くが消滅危機言語であるため，他の研究分野にはない「時間」という制約が加わる。このような背景から，限られた時間の中で問いの解決や，目的の達成を目指すのであればデータ収集の「効率」という点を考えざるを得ない。従って，従来の研究手法に対して評価を行い，それが目的を達成するために適切な方法かどうかを検討すること，すなわちメタ研究が必要だと言える。

本論文では，南琉球の語彙研究を事例に，これまでとられてきた調査方法や成果に対して詳細な評価を行い，目的を達成するための効果的な研究方法について論じる。結論として，日琉諸語を対象に多地点で詳細な語彙研究を行うという目的であれば，研究者と当該地域の方言の母語話者（以下，話者）が協働で行うハイブリッド型研究が最も有効であると主張する。さらに，地点の偏りを解消し，多くの地点を網羅するためには，ハイブリッド型研究であっても面接調査のみに頼ることは非現実的であるため，作業を細分化および分担し，各自が居ながらにして作業を実施するハイブリッド遠隔型研究を提案する。

本論文は次のように構成されている。まず 2 節では，日琉諸語を対象とする語彙研究の調査方法について概要を述べる。次に 3 節では，南琉球諸語を対象に過去 40 年間で行われてきた語彙研究の評価を行う。続く 4 節では，3 節で得られた結果から導き出される調査・研究手法による語彙研究の実践報告を行い，最後に 5 節でまとめと今後の展望を述べる。

本論文で対象とするのは，南琉球の語彙研究である。語彙研究と一口に言っても，様々な研究が含まれるため，本論文での定義を述べる。本論文が指す語彙研究とは，ある地点で使用されている語彙体系を記述することを目的とした研究である。語彙体系を記述することとは，語彙的な基礎情報を記述することを意味し，具体的には語形，意味，アクセント，品詞，活用クラスの情報，すなわち語彙的に決まっている情報を記述することを指す。この他にも派生関係（自他ペアなど）や語形成（複合語など）の記述も語彙研究に含まれうるが，本論文では上記 5 つの情報に絞って取り扱うこととする。

2. 先行研究

日琉諸語を対象とする語彙研究の代表的な調査方法は，①研究者自ら調査対象地域に足を運び調査を行う面接調査，②研究者が調査対象地域に調査票を送付して回答を待つ通信調査，③話者の内省によるものの 3 つに大きく分けられる。本節では，この 3 つの調査方法の概要と問題点，さらに，その他に，語彙研究以外で報告されている調査方法について述べる。

2.1 面接調査による語彙研究

面接調査とは，「言語調査（linguistic survey）の一つで，相手と直接やり取りをしながら行なう調査をさす」（亀井他（編）1996: 1338）。面接調査は古くから存在し，これまでも多くの研究者が採用してきた。例えば，南琉球に限定すると，古くは田代（編）（明治年間）がある。当該研究は 1880 年代に行った面接調査で収集した語彙資料である（セリック・麻生・中澤 2021）。交通の便が圧倒的に良くなった現在では，多くの研究者が語彙研究および語彙研究に含まれるアクセント研究等を面接調査によって実施している。

面接調査は，話者を目の前にして研究者が確認したいことを自ら確かめながら行うため，確実に回答を得られる調査方法であると言える。しかし一方で，研究者の交通費・滞在費や，スケジュールを調整した上での双方の拘束時間を考えると，コストが高い調査方法である。

2.2 通信調査による語彙研究

通信調査は，語彙研究が主とした目的ではないとはいえ，日本では古くは 20 世紀初頭から実施され（国語調査委員会（編）1905 など），その後，小林（1988，2004）などを中心に発展させられてきた方法である。

通信調査は，調査票を対象地域（個人や教育委員会等の自治体組織）に送付し，回答を待つというものである。一度に多地点で行える上，研究者が現地に足を運ぶ必要がなく，被調査者も時間がある時に回答するため，かかるコストは低い。一方で，小林（2004: 94）によると「通信調査法には面接調査法に比べて，回答の質量共に様々な問題があることは確かである」という。例えば，話者が文字のみで回答するため，実際の音声との異同に関して証拠は得られないという問題や，音声が分からないゆえ表記の統一性が図れないといった問題などが挙げられるだろう。さらに，通信調査を用いて各地点の語彙体系を明らかにしようとした先行研究は管見の限り見当たらない。このような背景から，語彙研究の成果として，一地点の辞書といった成果は生まれにくい可能性がある。南琉球における通信調査の実施報告や成果は見当たらない。

2.3 話者の内省による語彙研究

話者の内省による語彙研究は，話者自身による辞典編集の活動とも言い換えられる。語彙量が豊富かつ説明が詳細である一方，語形の正確さ・アクセント情報や品詞情報の欠如，項目の立て方など，少なからず問題も存在する。ただし，富浜（2013）の『宮古伊良部方言辞典』は上記の想定される問題を全て克服しているものである。しかし，このような質の成果が各地で自然に生まれることは非常に稀だと言えよう。例えば，アクセントの型を正しく記述するためには，ある程度の言語学的な知識と技能が必要であるが，一般的にこれらを自然と備えている者は（研究者を含め）ほぼいないと言える。

2.4 その他の方法

上記 3 つの他に，国内外で報告されている調査方法について，本論文と関連のあるものを大き

く 4 つに分けて紹介する。1 つ目は海外の事例，2 つ目は日本語（いわゆる共通語）の事例，3 つ目はコロナ禍での方言調査事例，4 つ目は琉球地域で調査地の人々と協働して行う活動事例である。

まず，海外の事例として挙げられるのが，スマートフォンアプリケーション LigAikuma（Gauthier et al. 2016）を利用しての言語資料収集である。この手法は，話者自身がスマートフォンのアプリを操作して自らの言葉を録音し，資料を蓄積していくものである。Blachon et al.（2016）では，アフリカのコンゴ共和国の首都で行われた事例が報告されている。このような手法を用いた研究は語彙研究とも言えるが，言語ドキュメンテーションの色がやや強い。

次に，コーパス作成の事例として国立国語研究所機関拠点型基幹研究プロジェクト「大規模日常会話コーパスに基づく話し言葉の多角的研究」がある。当該プロジェクトでは，首都圏在住のいわゆる共通語話者に会話録音を依頼する手法を採用している（田中他 2017，2018）。

上に挙げた 2 つの事例は，いずれも話者自身が言語資料を収集する手法を用いたものである。しかし，伝統的に用いられている日琉諸語にこの手法を適用することは，それが置かれた状況を考慮すれば不可能である。すなわち，消滅危機言語としての日琉諸語を用いる話者はおそらく 80 〜 90 代である一方，上記いずれの事例も参加者の年齢層が若い。Blachon et al.（2016）は，消滅危機言語を対象としているとの記載はあるものの，参加者の年齢層は明らかになっていない。しかし参加者の写真を見る限りかなり若い世代（20 〜 30 代か）である。田中他（2017，2018）の事例も協力者は 20 代〜 60 代以上とあるが，おそらく 80 代の（録音機器を操作する）協力者はいないと予測される。

3 つ目は，コロナ禍で面接調査を遠隔で行っている事例である。小西（2021）は，オンラインで面接調査をしながら，その状況を話者側で録音する手法を紹介している。研究者が調査地に赴けないため，録音やパソコンのセッティングには話者以外の方の協力を得ている。話者と研究者が顔を合わせるという点で，本事例は一種の面接調査に分類されるだろう。

最後に，調査地に住む人々との協働による資料収集の事例である。これは調査地に住む非母語話者が同じ調査地に住む母語話者に調査を行うという試みである。このような事例には重野・白田（2017），林他（2021）などがある [1]。研究者が単独で行う調査とは異なり，これらの事例の多くは，方言の記録者や伝承者の育成が第一義的な目的である。調査現場に研究者が同伴する場合としない場合があるが，いずれにせよ一種の面接調査に分類される。

以上，日琉諸語の語彙研究で報告されている調査方法を見てきた。琉球諸語を対象にした場合，面接調査による研究と話者の内省による研究の 2 つが主となる。3 節では，南琉球地域の語彙研究を事例に研究成果を詳しく見ていく。

[1] この他，ウェブサイトで確認できるものでは，「しまむに教室」（URL: https://researchmap.jp/akikoyokoyama/presentations）（2021 年 12 月 4 日確認）や，「なりきり！方言研究者プロジェクト」（URL: https://sites.google.com/view/narikiri-linguist-pj）（2021 年 12 月 2 日確認）なども調査地に住む非母語話者が同じ調査地に住む母語話者に調査を行う試みにあたる。

3. 南琉球の 40 年間（1980 年〜 2020 年）の語彙研究の評価

2.1 で述べたように，日琉諸語の方言を対象としてきた語彙・アクセント研究では，従来研究者による面接調査が採用されてきた。一方で，調査方法そのものについて，客観的な指標を基に評価されたことは（少なくとも日琉諸語を対象にしたものは）これまでないと言えよう。小林（2004: 94）も「通信調査法が面接調査法に比較してどのような点でどの程度問題なのか，具体的な事例の対比に基づいて検討されたことはあまりなかった」と述べている。従って，本論文の対象とする語彙研究という目的に対して，面接調査がどの程度適切であるかについては客観的な評価が必要であると言える。

そこで本節では，言語研究の客観的評価の試みとして，評価対象を 1980 年から 2020 年までの南琉球の語彙研究に限定し，調査方法およびその成果を評価するための客観的な指標を提示し，分析を行った上で，語彙研究の方法について考察を行う。3.1 で評価対象となる文献と情報の付与方法を述べ，3.2 で語数を中心とした量的評価，3.3 でアクセント情報を中心とした質的評価，3.4 で面接調査の効率の評価を行い，3.5 でまとめを提示する。

3.1 評価対象となる文献と情報の付与方法

1980 年から 2020 年までの南琉球の語彙研究を評価するため，南琉球に関する語彙資料を，種類（論文，辞典など）を問わず網羅的に収集・調査した。ただし，未公刊の原稿やデータ，あるいは中本（1981）のように地図の形でしか語彙が報告されていない文献は調査の対象外とした。収集の結果，193 点の文献が得られた。地元で出回っている私家版など，見逃した文献が幾点もあろうが，南琉球の語彙研究の動向を見るのに十分な資料が収集できたと考える。調査対象となった文献リストは関連データとして提示する。以下では，評価の際に用いる用語および情報付与の方針について述べる。

まず，収集した文献を検討した結果，当該期間における南琉球の語彙研究は，研究主体および採用する調査方法によって 3 つの研究形態に分類できることが分かった。すなわち，193 点の文献は，①研究者による面接調査の成果物（以下「研究者型」，153 点），②話者のみの内省による語彙収集の成果物（以下「話者型」，18 点），③話者の内省による語彙収集の上，研究者との協働によって生まれた成果物（以下「ハイブリッド型」, 22 点），の 3 つのカテゴリーに分けられる。なお，ハイブリッド型には八重山鳩間出身の加治工真市や宮古西原出身の名嘉真三成のように方言母語話者である研究者の成果物も含めた。

次に，情報の付与方針について述べる。書誌情報に加え，各文献に対して「対象地点」,「項目数」,「アクセント情報の有無」の情報を付与した。「項目数」は次のようにカウントした。電子ファイルがある文献は, 電子ファイルから項目数を抽出した。電子ファイルがない文献は, 項目数（または語数）が掲載されていれば，掲載されている数字をそのまま採用した（「約」も含む）。項目数の掲載がない文献は分量によって異なる手法を用いた。分量の少ない文献は項目数を手で数えた。分量の多い文献は頁の一部（10 頁ごとなど）を対象に項目を数え，頁ごとの項目数の平均を基に文献全体の項目数を推測した。また，電子ファイルが部分的に存在する文献は電子ファイ

ルの項目数を基に残りの項目数を推測したこともある。これ以降，本論文では項目数の代理変数として「語数」を用いる。続いて,「アクセント情報の有無」については次のカテゴリーを設けた。すなわち，①アクセント情報が欠落している，②ピッチ情報のみが報告されている，③アクセント型が報告されている，④対象地点はアクセントの対立がない，の 4 つである。

次節以降，上記の用語および情報を用い，南琉球の 40 年間の語彙研究の評価を行う。

3.2 量的評価

本節では，主に語数情報に着目し，語彙研究の量的な評価を行う。3.2.1 で南琉球の語彙研究の動向について，3.2.2 で各変数から見た南琉球の語彙研究の特徴，そして 3.2.3 で語数から見た地点の偏りについて述べる。

3.2.1 南琉球の語彙研究の動向

まず，評価対象期間中の南琉球の語彙研究の動向を見る。年毎に報告されてきた語数の累計を図 1 に示す。

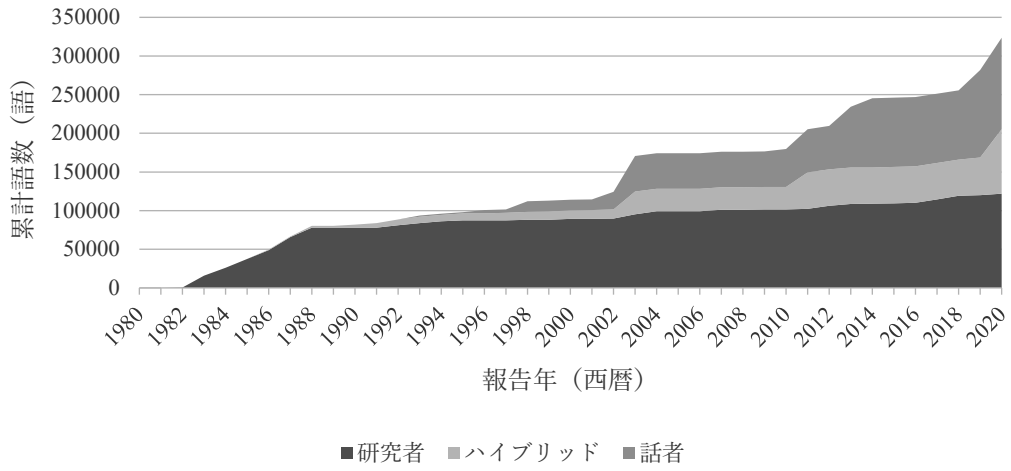

図 1　1980 年～ 2020 年の間に報告された南琉球の語数（累計・研究形態別）

図 1 から分かるように，評価対象期間のはじめの 80 年代には研究者型の研究が非常に活発に行われており，その成果として 1990 年までに 77,825 語が報告された[2]。その背景にはそれぞれ 15,396 語と 12,029 語を含む平山（1983）と平山（1988），約 16,643 語を含む中松（1987）の研究と 27,223 語を含む沖縄言語研究センターの全集落調査（1984 ～ 1986）の一連の報告書がある。しかし，90 年代になると，研究者型の語彙研究の勢いが大きく落ち込み，毎年報告される語数は今日にいたるまでずっと低迷している。その結果，1990 年から 2020 年までの 30 年の間に研

[2] ただし,「語数」には，同一地点の同一単語の重複が相当数含まれていることに注意されたい。

究者型の研究によっては 44,048 語しか報告されてこなかった。1990 年以降の語彙研究の動向の変化には研究者の関心の変化（語彙から文法へ）が関わっている可能性も考えられる。

1990 年以降研究者型の語彙研究が衰退していく代わりに，2000 年代からハイブリッド型とそれに勝って話者型の活動が活発化していく。2000 年から 2020 年までの間に話者型の研究成果は 104,565 語にも達している。これは，南琉球の語彙研究において話者型の研究の貢献度がいかに高いかをよく物語っている。

2000 年代を皮切りに，話者を中心とした語彙研究が活発化していくことには次の要因が考えられる。まず，方言に対する社会的意識が大きく変わった。つまり，伝統的方言の再評価を経て，研究者以外でも方言に対する関心が高まってきた。1994 年の宮古島の鳴りとぅゆんみゃーく方言大会の創設や 2006 年の沖縄県による「しまくとぅばの日」の制定などはこのような社会的な動きをよく反映している。さらに，琉球列島の諸言語を「消滅危機言語」と認定したユネスコの 2009 年の「消滅の危機にある言語」の発表もあり，伝統的な方言が消滅の危機に瀕しているという意識も高まった[3]。その結果，伝統的方言を流暢に話せる最後の世代は危機感を覚え，自分たちの方言を残そうと語彙集・辞典の事業に積極的に取り組むようになったと考えられる。

3.2.2 各変数から見た南琉球の語彙研究の特徴

続いて，南琉球の語彙研究の特徴をもう少し詳しく見る。表 1 に文献数，対象地点数，語数の概要を示す。

表 1　文献数・地点数・語数概要

変数	研究者型	ハイブリッド型	話者型	合計
文献数	153	22	18	193
地点数	56	6	12	57
語数（合計）	122,797	83,292	121,630	326,795
語数（平均）	802	3,786	6,757	1,702
語数（中央値）	280	857	4,834	321

40 年の間に南琉球の 57 地点（うち宮古 38 地点，八重山 19 地点）について延べ 326,795 語が報告された。語数だけで言うと，研究者型と話者型の研究成果は同程度で，研究者型は延べ 122,797 語，話者型は延べ 121,630 語の成果を出している。それに次いで，ハイブリッド型は延べ 83,292 語の成果を出している。これらの数字で分かるように，上で述べたように南琉球の語彙研究においては，話者のみの活動による貢献が非常に高く，研究者に引けを取らない。しかし，研究者型と話者型の研究は特徴が大きく異なる。研究者型の研究は文献数（153 点）が多く，平均語数（802 語）が少ないのに対して，話者型の研究は文献数（18 点）が少なく，平均語数（6,757 語）が多い。平均語数の差は複数の要因によると考えられる。

この差を説明するためにまず発信形態の違いが挙げられよう。話者型の研究発信は全て語彙

[3] 実際，『南琉球宮古語多良間方言辞典』の第一著者はユネスコの発表をきっかけに，辞典の原稿をしたため始めたという（渡久山春英氏，私信）。

集・辞典の形をとっており，1 点の成果物につき多くの語が含まれる。これに対して，研究者型の研究は紙数制限のある論文，あるいはその年度の調査報告書の形をとるものが多く，その必然の結果として 1 点の成果物に報告されている語数は少ない。研究者型の中央値（280 語）の低さは，研究者型のこのような研究発信形態の特徴を特によく反映していると言える。つまり，逆に言えば，平山（1983，1988）や中松（1987）のように数千語以上を報告する 80 年代の成果物は例外である。

しかし，研究者型と話者型の平均語数の差には，発信形態の違いはあれ，調査方法の違いも大きく関わっている可能性があることが指摘できる。研究者型の全ての研究は面接調査を用いているが，面接調査は制約が多いと言わざるを得ない。まず，調査内容（調査票）自体の制約がある。具体的に言うと，面接調査では，特にインフォーマントの参加が消極的であれば，用意された調査内容の分の情報しか得られない。仮に調査票を可能な限り網羅的に作成したとしても，それでも限界がある。なぜならば，その存在が予測できない項目はそもそも調査項目として立てられないからである [4]。実際，研究者が発表してきた語彙のほとんどは基礎語彙に限る。これに対して，話者の内省による語彙収集は内容に関する制約がないため，基礎語彙の範囲を遥かに超える語彙資料を提供できている。そればかりではない。面接調査は，調査内容の制約に加え，調査時間の制約も加わる。つまり，面接調査は研究者と話者の都合が合うという局所的な形で行われていることがほとんどで，原則的に限られた時間でしか実施できない [5]。当然ながら，話者の内省による語彙収集にはそのような制約がない。これは研究者型と話者型との大きな違いである。つまり，研究者による調査は単発的，短期的なものが多いのに対し，話者による語彙収集は長期的に行われる場合が多い。

以上，平均語数に見られる研究者型と話者型との差について述べたが，対象の地点数においても同様に大きな差が観察される。話者型の研究が対象としている地点数（12 地点）に比べ，研究者型が対象としている地点数（56 地点）は圧倒的に多い。この差は，次のように説明することができる。つまり，研究者型の研究では 1 地点における収集語数は相対的に少ないが，興味関心によって，多くの地点を対象としやすい。これに対し，話者型の研究では基本的に自分の母語以外の言語資料を収集しない。この点は話者型研究の最も大きな限界であると言える。

続いて，ハイブリッド型研究の結果について簡単に述べる。ハイブリッド型の研究は話者型の傾向とほとんど同じであると言える。各成果物の平均語数が高いのに対して，対象地点は少ない。なお，話者型に比べ，平均語数と中央値がやや低くなっているのは母語話者兼研究者の名嘉真三成と加治工真市が論文の形（それぞれ 10 点と 7 点）で語彙を報告しているからである。これらの研究を除くと，ハイブリッド型の成果物は話者型と同様に書籍の形になっている。

[4] 例えば，宮古語には「脱皮中の蝉」を指す語が存在する。
[5] 面接調査の効率については 3.4 を参照されたい。

3.2.3 語数から見た地点の偏り

最後に語数から見た地点の偏りを見る。図 2 に宮古語および八重山語の各地点で報告された語数の合計を示す。図 2 で分かるように地点の偏りが激しい（地点別の語数の平均は 5,726 語，中央値は 1,777 語，標準偏差は 8,981 語）。数万語が報告されている地点（与那国，平良，多良間，鳩間，石垣（四箇））もあれば，数百語しか報告されていない地点や，1 語も報告されていない地点も多い。この偏りはどのような要因によるものであろうか。

まずこの偏りは，評価対象期間（1980 年～ 2020 年）以前の研究成果との整合性を図るために生じたものではないことが指摘できる。なぜならば，1980 年以前の語彙研究の最も大きな成果物の対象方言と，評価対象期間中に語数が多く報告された方言とは重なっているからである。例えば，1980 年以前の主要な成果物（宮良（1930），Jarosz（2015）に収録されているネフスキーの戦前の語彙研究や宮良（1975），下地（1979））は宮古平良方言と八重山石垣四箇方言を対象としているが，評価対象期間中にこの 2 つの方言について報告された語数の多さはそれぞれ 2 番目（34,881 語）と 5 番目（22,768 語）である（図 2）。

では，その他に地点の偏りを説明できる要因はあるのだろうか。ここには，各研究形態の構造自体が大きく関わっていると考えられる。本節で明らかにしたように，多くの地点を対象とするのは研究者型の研究であるが，報告する語数は平均的に少ない。これに対して，話者型およびハイブリッド型の研究により数多くの語数が報告されているが，対象地点が少ない。この 2 つの特徴を合わせると，地点の偏りが必然的に出てくる。

さらに，各地点の話者人口の構造も偏りの要因になっていると思われる。つまり研究者型の研究では，当該地点の話者数が多いほど相対的に協力者が見つかりやすいということである。一方で，話者型の研究は，語彙集・辞典を編集できる人材の出現が前提となっており，研究成果を出した著者の多くは教師を含めた公務員等を経験している教養層に属している。しかし，このような人材は，話者数が多い地点だからと言ってどこにでもいるとは限らない。結局のところ，内省による語彙収集を体系的に行う話者の存在が，南琉球の語彙研究の主要なボトルネックとなっていると言える。

なお，もう 1 つの要因として，研究者型の研究の，同じ地点への集中が挙げられる。最も語数が多い与那国（36,984 語），平良（34,881 語），多良間（31,629 語）についてはそれぞれ 11 点もの研究者型の成果物がある。これに関しては様々な要因が考えられる。例えば，協力者が見つかりやすい，先行研究に惹かれその地点の研究を始めようとする者がさらに出てくる，などが挙げられよう。考えうる様々な要因に関して定量的に考察する必要があるが，今回は扱わない。

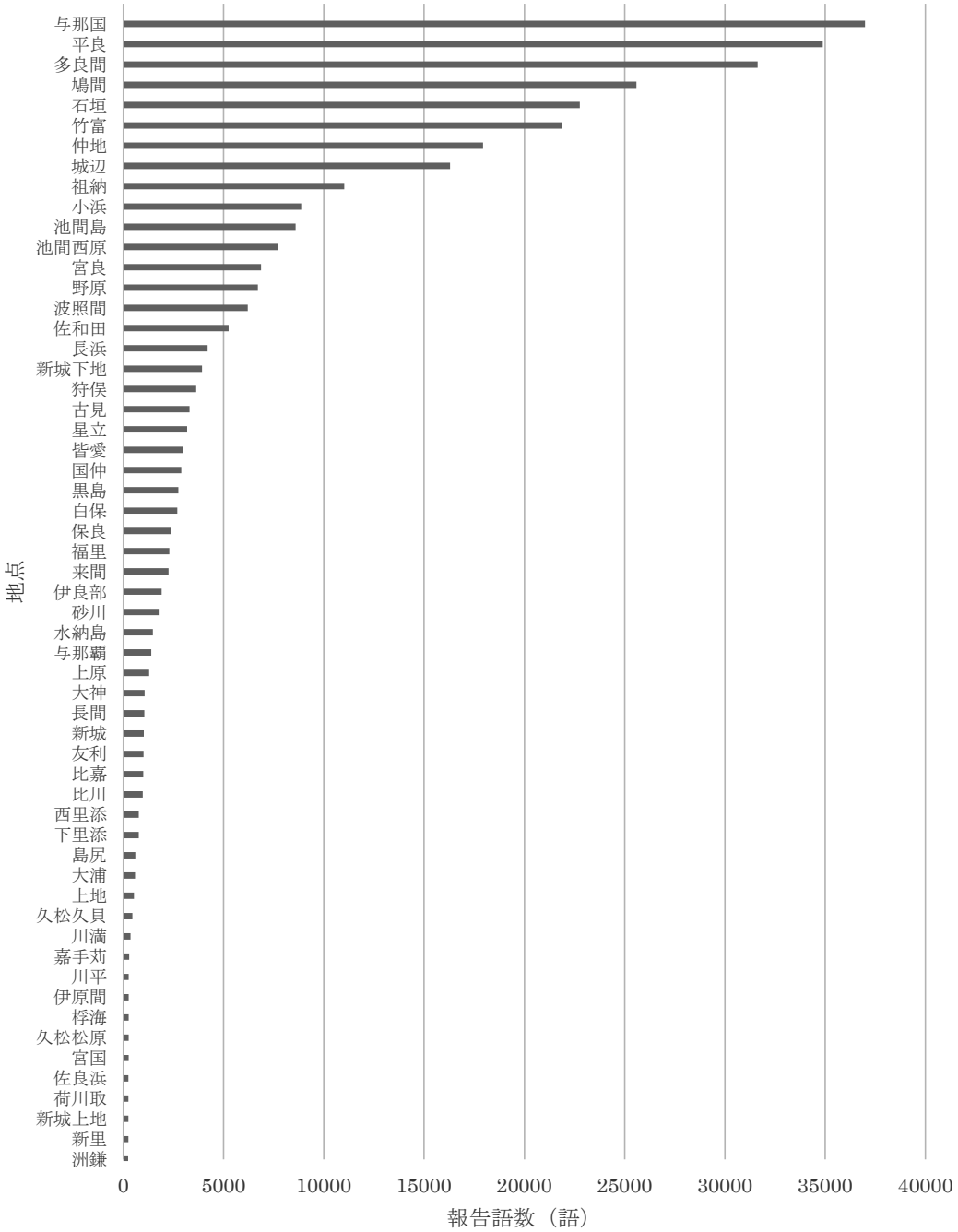
与那国
平良
多良間
鳩間
石垣
竹富
仲地
城辺
祖納
小浜
池間島
池間西原
宮良
野原
波照間
佐和田
長浜
新城下地
狩俣
古見
星立
皆愛
国仲
黒島
白保
保良
福里
来間
伊良部
砂川
水納島
与那覇
上原
大神
長間
新城
友利
比嘉
比川
西里添
下里添
島尻
大浦
上地
久松久貝
川満
嘉手苅
川平
伊原間
桴海
久松松原
宮国
佐良浜
荷川取
新城上地
新里
洲鎌
地点
0
5000
10000
15000
20000
25000
30000
35000
40000
報告語数（語）

図 2　地点別の報告語数

3.3 質的評価

本節では，アクセント情報を中心に各研究形態の質的評価を行う。本論文で定義した語彙研究にとっては「語形」「意味記述」「品詞」「（対立があれば）アクセント」「活用クラス」の情報が必要となる。しかし，文献によっては情報の欠落もあり，また，情報があっても精度が低いことがある。本節では研究形態によってどのような傾向があるかを見る。3.3.1 で結果を概観し，3.3.2 では特にアクセント情報に絞って結果を比較する。

3.3.1 質的評価結果の概要

総合的に見ると，ハイブリッド型の研究の質が最も高いと言える。語彙研究にとって必要となる情報が全て揃うことが多い上に，各情報の精度が高いからである。これに対して，話者型の研究は品詞，活用クラス，アクセントなどの情報が欠ける傾向が目立つ。さらに，複雑な音素体系を持つ言語の場合は適切な表記が確立していないこともあり，「語形」の記述精度が劣っていることがあると言わざるを得ない。

研究者型は精度が高い傾向にあるが，ハイブリッド型ほど情報の種類がカバーされているとは言えない。ただし，これは研究者型の研究成果の発信形態にもよると考えられる。例えば，各方言の音素を対象とした論文（加治工（1982），仲原（2003，2005）等）には品詞やアクセントの情報が掲載されていないが，これらの論文が解き明かそうとしている課題（音素体系）に対してはそれほど支障がない。

3.3.2 アクセント情報を例にした研究形態別質的評価

各研究形態の質をもう少し詳しく比較するために，アクセント情報の掲載の有無と仕方を例に取り質的評価を行う。南琉球の大半の方言にアクセントの対立があることは，昔からよく知られている（秋永 1960，平山他 1967）。そのため，南琉球諸方言の語彙研究では，アクセントの所属が重要な語彙的情報となる。研究形態別にアクセント情報の掲載の有無とその仕方（語形のピッチとアクセント型の明記）を図 3 に示す[6]。文献数で見ると，研究者型で最もアクセント情報の掲載がある（68%）のに対して，話者型はピッチを示す 1 点の文献（池間 2003）を除いて，アクセント情報の掲載が皆無である。ハイブリッド型は研究者型に近い（50%）。

[6] アクセントの対立がないとされる方言を対象とした文献は除外した。

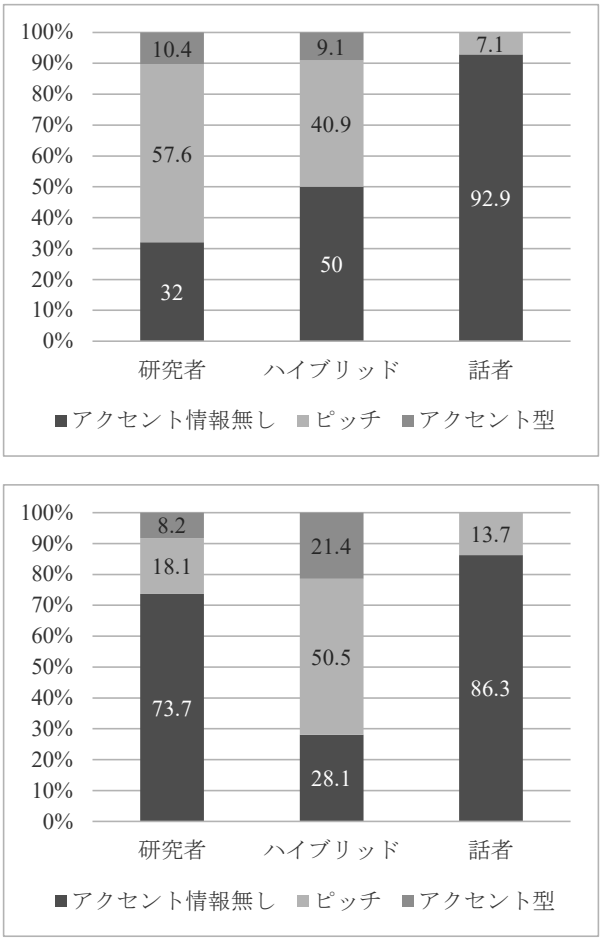

図 3　アクセント情報の掲載（上：文献数，下：語数）

しかし，文献数ではなく，報告語数で見ると状況が大きく変わる。研究者型の研究によって報告された語数の 4 分の 3 にアクセント情報がなく（73.7%），ハイブリッド型の研究はちょうどその逆で, 報告された語数の 4 分の 3 弱にアクセント情報がある（71.9%）。さらに驚くべきことに，表 1 から研究者型と話者型の総語数がそれぞれ 122,797 語と 121,630 語であることを踏まえ，アクセント情報が掲載されていない語数を比較すると，研究者型と話者型ではあまり差が見られないという結果が得られるのである。

言語学の専門家ではない話者がアクセント情報を掲載しないのは説明がつくが，言語学の専門家であるはずの研究者もアクセント情報を掲載しない傾向があるのは何によるものだろうか。そこでいくつかの可能性が挙げられる。まず，上で述べたとおり，文献によっては発信形態（音素体系を対象とした論文など）が関わっていることが指摘できる。しかし，語彙体系を研究対象とし，大量の語数を報告した平山（1983，1988）および中松（1987）などがアクセント情報を掲載

していない点を，発信形態で説明することはできない。発信形態の他には，対象言語のアクセント認定の難易度が関与していることが考えられる。例えば，アクセント型の対立を持つ宮古語の諸方言では単語の単独発話ではアクセント型が認定できないことがよく知られている[7]。そのため，宮古語を対象とした語彙研究では，アクセント型の認定に特化した，場合によっては非常に複雑な調査が別途必要となる。しかし，アクセント情報の欠落傾向は単独の発話でアクセント型が特定できる方言についても認められるため，アクセント体系の難易度だけでもアクセント情報の欠落傾向は説明しきれない。例えば，単独の発話あるいは任意の単純な枠文でアクセント型が特定できる方言には，八重山の諸方言のうち，石垣四箇方言，石垣宮良方言，石垣白保方言，竹富方言，鳩間方言，西表網取方言，波照間方言，与那国方言などがある。これらの方言を対象とした研究者型の研究でもアクセント情報が報告されない傾向が顕著である。例えば，与那国方言を対象とした平山（1988）や法政大学沖縄文化研究所（1987a, b）ではおよそ 10,000 語の資料に対してアクセント情報が含まれていない[8]。

残された可能性として，アクセント情報の欠落傾向に，従来の面接調査方法自体が関係しているということが挙げられる。南琉球の研究者型の研究は，そのほとんどが現地において面接調査を実施している。従来の面接調査では 1 つの調査で「語形」「意味」「品詞」など複数の語彙的情報を調べているわけであるが，よく考えると，語彙研究のために収集しなければならない情報（「語形」「意味」「アクセント」など）はそれぞれ性質が異なっており，当然ながらこれらの情報を収集するための適切な調査手法もそれぞれ異なるはずである[9]。つまり，従来の面接調査では同一の調査セッションで性質の全く異なる情報を収集しようとしてきたわけであるが，異なる手法が求められる調査を同時にかつ十分に遂行することは困難であろう。実際，本論文で対象とする従来の面接調査を採った研究者型の研究の中には，語彙的な基礎情報（語形，意味，アクセント，品詞，活用クラス）を全て報告している研究は 1 つもない。

さらに文献を検討してみると，アクセント情報の有無と，他の語彙的情報の精密さとの間には相関関係が認められる。それは「意味記述が詳細であればあるほどアクセント情報がなく，調査語彙が基礎的なものであればあるほどアクセント情報がある」という関係である。例えば，基礎語彙だけを対象とし，意味記述を簡潔にしている全集落調査はピッチが記されていることが多い。これに対し，調査範囲が基礎語彙を超え，品詞情報も付けながら，例文付きの詳細な意味記述を行った平山（1983，1988）にはアクセント情報が掲載されていない。このことは，1 回の面接調査で全ての情報を揃えることがいかに難しいかを強く示唆している。次節では，面接調査の効率

[7] 対象の語を複数の生産的な複合語の前部要素にしない限りそのアクセント型が特定できない極端な例については，セリック（2021）を参照されたい。
[8] ただし，これらの方言のアクセント体系が，それまで本土で慣習的に用いられてきたアクセント記述のフレームワークにそもそも当てはまらず，研究者がアクセント情報を記述しにくかったことによる可能性も大いに考えられる。
[9] ここで対象とするのはあくまで本論文で定義する語彙研究に対してのものである。言語を包括的に記述する際などには，様々な文法体系を同時に考えながら面接調査を進めることが必要であろう。そのような意味では，包括的な文法書執筆のための面接調査で語彙研究のための資料を同時に集めるというのは当然のことながら難しいとも言える。

について評価を行う。

3.4 面接調査の効率評価

南琉球諸方言を対象とした研究者型の研究は，そのほとんどが面接調査の方法を採ってきた。しかし，これまで面接調査の効率に関する評価がなされていない。従って，本節では南琉球諸方言の語彙集等の成果物のうち，調査参加者（調査者と被調査者）数，および調査時間が明記されているものを対象に，面接調査の効率について予備的な評価を行う。調査時間が分かる研究がほとんどなかったため，本論文が対象とする評価期間（1980 年〜 2020 年）からは外れるが，調査時間が分かるセリック・麻生・中澤・中川（2021）も含めた。

評価対象となるものは，木部（編）（2012），ローレンス（2013，2019），セリック・麻生・中澤・中川（2021）である。このうち，ローレンス（2013，2019）は語形の確認の他，アクセント調査や談話録音を兼ねている。表 2 に一覧を挙げる。語数，調査時間，調査者数，被調査者数は全て延べ数である。

表 2　評価対象となる文献と各種数値

	木部（編）（2012）	ローレンス（2013）	ローレンス（2019）	セリック・麻生・中澤・中川（2021）
調査地点	宮古各地	竹富	竹富	石垣川平
調査者数	28	1	1	4
被調査者数	22	1	1	3
セッション数	20	11	20	5
総調査時間	40h	11h	20h	10h
語数	3390	432	840	881
1 時間当たりの語数	84.8 語	39.2 語	42.0 語	88.1 語

1 時間当たりの収集語数を計算すると，興味深いことに結果がまとまる。すなわち，語形を集中的に確認する場合には 1 時間に 84 〜 88 語，アクセント情報や談話収集も兼ねたものであれば，そのおよそ半分の 39 〜 42 語が収集できるという結果が得られた。おそらく研究者の平均的な数値として見ることができるだろう。

以上の結果に基づいて，次のような簡単なシミュレーションを行う。これまでと同様の面接調査にて，最大級の方言辞典，すなわち約 15,000 語を収録するアクセント付き辞典を作成するためには，375 時間の調査時間が必要ということになる。仮に 1 日に 5 時間調査するとしても 75 日かかる計算になる。しかもこの時間はあくまで資料収集に要する時間であり，面接調査の後に行うデータ処理の時間を含んでいない。研究者の多くは自分の研究の他に，その研究とは無関係な業務（大学教育業務や大学運営業務等）を持っていることを考慮すると，1 年間に行える調査は多くてもせいぜい 10 日〜 14 日程度だろうから，辞書編纂までには少なく見積もっても 5 年〜 7 年かかることになる。これらのことを踏まえると，研究者による面接調査のみで大規模な方言辞典を作成することは現実的であるとは思えない。南琉球の各地点の本格的な辞典を作ろうと思えば，別の方法を探る必要があることは一目瞭然である。

3.5 まとめ

これまでの評価の結果を表 3 にまとめる。

表 3　各研究形態（研究者型・ハイブリッド型・話者型）の比較

		研究者型	ハイブリッド型	話者型
特徴	調査方法	面接	面接・内省	内省
	プロジェクト形態	単発的・短期的	長期的	長期的
	調査範囲制約	有り（基礎語彙）	無し	無し
	調査時間制約	有り	無し	無し
	調査効率	低い	高い	高い
	調査費	有り	有り	無し
	発信形態	論文，報告書	本，論文	本
量的指標	対象地点	多い	少ない	少ない
	（地点別）語数	少ない	多い	多い
質的指標	立項	語彙素	語彙素	語形
	表記	正確	正確	不正確な場合も
	アクセント情報	あまり無い	よく有る	無し
	品詞情報	よく有る	有り	無いことが多い

表 3 を見ると，質的にも量的にも最もバランスの取れた成果が期待できるのはハイブリッド型研究であることがよく分かる。研究者型は面接調査の限界に縛られており，大規模な語彙データを収集することが難しい。話者型は多くの言語的資料を提供してはいるものの，学術的な観点を欠いていることがあるため，質が劣る傾向がある。以上のことから，南琉球の総合的な語彙研究を目指すのであれば，ハイブリッド型の研究が望ましいと結論付けられる。まずは，すでに存在する研究者型あるいは話者型研究のハイブリッド化を積極的に試みることが非常に有効であろう。

一方で，効率化が図れるのはあくまで一地点内の話であり，ハイブリッド型の研究形態を採用したとしても全体を見た際の地点の偏りはこれまでと同様に生じる。

上記の問題点に関しては，大規模実現可能性すなわち「スケーラビリティ（scalability）」という観点から，ハイブリッド遠隔型の研究形態を提案する。ここでいうハイブリッド遠隔型の研究は，これまで面接調査で行っていた作業のうち，遠隔化できるものを抽出し遠隔で実施するものである。遠隔化できる代表的なものとしては，原稿作成作業と録音作業が挙げられる。ハイブリッド遠隔型研究で採用される遠隔調査は，原稿作成を依頼し，郵送等でやり取りする点では 2.2 で述べた通信調査と同じ特徴を持つが，両者は似て非なるものである。本論文が提案するハイブリッド遠隔型研究は，面接調査で得られるであろうデータの内容・質を変えずに遠隔化するため，どちらかと言えば調査者と被調査者の相対する時間を極力減らした面接調査と言えるかもしれない。その意味で，2.4 で述べた面接調査を遠隔（オンライン）で行う手法とも異なる。可能であれば話者自身が録音も行うため，調査者は拘束時間が減る分，空いた時間を用いて別の地点で調査を始めることができるだろう。

以上，効率化の観点およびスケーラビリティの観点から，語彙研究に必要な作業の流れを合理

化するために，ハイブリッド型の研究形態に遠隔調査を取り入れたハイブリッド遠隔型研究を提案する。4 節では，実際にハイブリッド遠隔型で進めた研究事例を紹介する。

4. ハイブリッド遠隔型の語彙研究の試み

本節では，ハイブリッド遠隔型で進めた南琉球の語彙研究について，筆者らの事例を紹介する。4.1 で概要を述べた後，4.2 から 4.3 で個別の事例を紹介し，4.4 で筆者らの研究結果を考察する。紹介する事例の地点を図 4 に示す。北から宮古の狩俣および水納，八重山の石垣宮良，小浜，西表網取，波照間である。

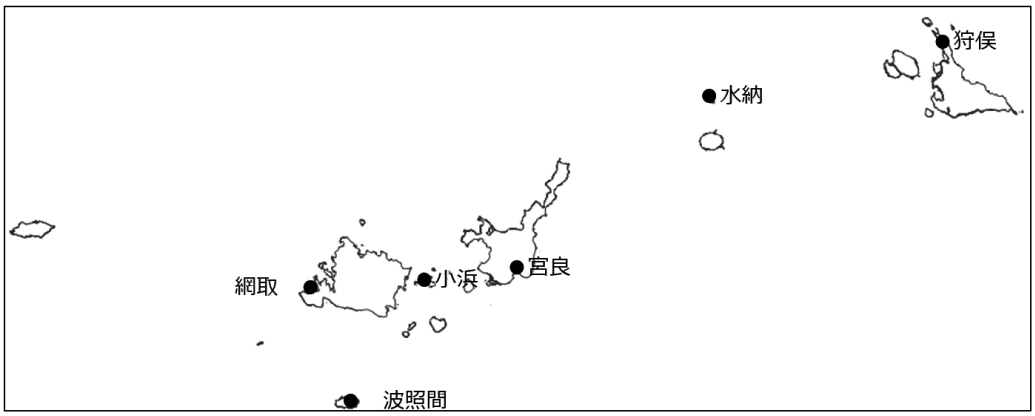

図 4　本節で紹介する事例の対象地（南琉球の 6 地点）[10]

4.1 概要

ハイブリッド遠隔型研究の特徴の 1 つは，作業工程を明確に分け，分担し，各地で実施する点である。例えば，原稿作成作業と録音作業は完全に分離した。我々が試みた基本的な調査の流れは次の 1 から 6 のとおりである。一連の作業は複数回繰り返される。3 の工程では一連の作業を複数回繰り返す過程で，原稿作成作業の回と録音作業の回が交互に繰り返される。

〈調査の流れ〉

1. 研究者：話者に電話や手紙等で連絡し，協力依頼を行う。状況により挨拶にも伺う。
2. 研究者：録音機等の調査セット（図 5）を郵送し，文面や電話で説明を行う。
3. 話者：原稿作成作業あるいは録音作業を行う。
4. 話者：作業完了後，調査セットを返送する。
5. 研究者：受け取ったデータを確認し，整理・処理作業を行う。
6. 研究者：簡易語彙集を作成し，話者に送付する。継続可能な場合は，2 を同時に行う。

[10] 国土地理院（https://maps.gsi.go.jp/development/ichiran.html）の地図を利用し筆者らが改変した。

話者に送付する調査セットには次のものを同梱し，レターパックを利用した（図 5）。

〈調査セット〉

- 録音機
- 録音機の取扱説明書
- 調査指示書
- 調査票
- 宛名等記入済みの返送用レターパック

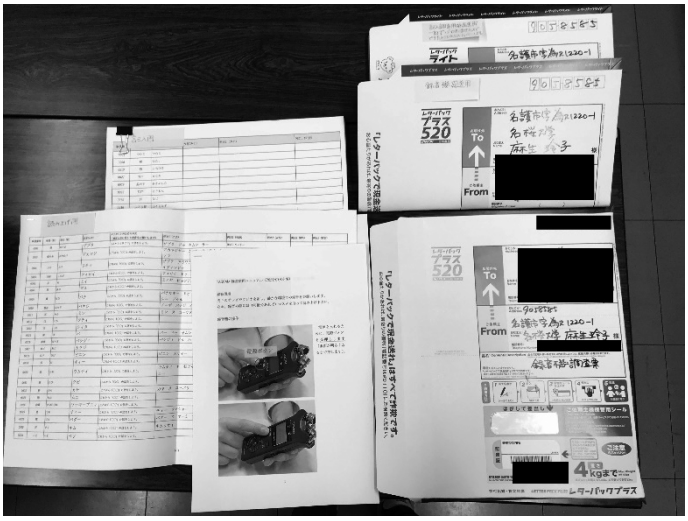

図 5　調査セットの例

音声データの収集には，操作が比較的簡単なパナソニックの RR-XS470 と，多少慣れが必要となる TASCAM の DR-07MKII を，筆者らで選択して送付した。いずれも wav 形式で録音が可能なマイク内蔵型の録音機である。

協力者とは，知り合いからの紹介（さらにその方からの紹介も含む），郷友会への連絡を通した紹介，教育委員会からの紹介等で知り合った。電話や手紙によるコミュニケーションと，これまでの成果等の送付を通して，人間関係を構築し，調査を開始している。成果を見える形にするという点で，ある程度の作業を行った段階で，簡易語彙集にして協力者に送付している。

次節以降，筆者らが現在 6 地点で実施しているハイブリッド遠隔型の語彙研究の事例を述べる（表 4）。その際に，語彙が記された原稿の有無，原稿がある場合にはすでに出版されたものかどうかという観点で分類し述べていく。

表 4　本論文で紹介するハイブリッド遠隔型調査事例一覧

地点	話者の年齢層	原稿	録音[11]	録音者	確認できた語数
八重山・石垣宮良	80 代後半	既存の原稿（出版）	遠隔 P	話者・第三者	778
八重山・小浜 1	80 代後半	既存の原稿（出版）	遠隔 T	話者	407
八重山・小浜 2	80 代前半	既存の原稿（出版）	遠隔 T	話者	未
宮古・狩俣	80 代前半	既存の原稿（出版）	遠隔 T	第三者	約 1,000
八重山・波照間	80 代後半	既存の原稿（私家版）	遠隔 P/T	話者	2,795
八重山・西表網取	70 代前半	原稿作成	遠隔 T	話者	269
宮古・水納	80 代前半	原稿作成	面接	研究者	約 5,000

4.2 原稿がある場合（話者型からハイブリッド遠隔型へ）

4.2.1 出版された語彙研究を出発点に行う研究方法

本方法は，話者がすでに出版した語彙研究を基に，話者型の研究で不足しがちな，立項･表記･アクセント情報･品詞情報に関する問題（表 3）を整理し，再度ハイブリッド遠隔型で研究を進め，内容の充実を図るものである。話者型の研究成果の著者と協力者は，同一人物の場合もあるが，一致するとは限らない。

整理した語彙一覧を調査票とし，話者に送付する。話者には各語形の確認と録音を依頼する。話者は作業後，研究者に一式を返送する。研究者は返送された原稿に対して文字入力や音声の書き起こし等の処理を行う。そして次の調査票とともに処理した原稿を協力者に送付し，同時に確認を依頼する。この一連の流れを繰り返す。

次の 3 地点で調査を実施した。

- 八重山語石垣宮良方言（2020 年 12 月～ 2021 年 2 月）：協力者とは別の話者が出版したものを音声化する試みである（単独発話 778 点，用例 792 点）。引き続き音声化作業を依頼予定である。同居する家族に録音を手伝ってもらうこともあったそうである。
- 八重山語小浜方言（2021 年 7 月～）：2 人に依頼し，出版物の音声化を実施している。1 人は出版物の著者である（用例 407 点）。
- 宮古語狩俣方言（2021 年 9 月～）：出版物の音声化を実施している。本事例では，録音を地元の協力者に委託している（単独発話約 1,000 点）。

4.2.2 未出版の語彙研究に途中から研究者が関与する研究

本方法は，出版予定あるいは個人が取りまとめた語彙研究に，研究者が途中から関与し，話者と協働で研究を進める方法である。言い換えれば，話者型で進んでいたものをハイブリッド遠隔型にした例である。本方法の事例は一例のみであり，すでにある語彙研究の著者と協力者は同一人物の例である。基本的には 4.2.1 と同様の手順を踏む。

[11] 送付した録音機を P（パナソニックの RR-XS470）と，T（TASCAM の DR-07MKII）で記す。

- 八重山語波照間方言（2020 年 1 月～）：個人がまとめた語彙集（2,795 語収録）に対し録音を行い，音声記号とアクセント情報を付与した。現在は動詞の活用形の統合等の立項に関する整備を行っている。今後，例文や語数を拡張した原稿を作成予定である。

4.3 原稿がない場合（新規に始める語彙研究）

本方法は，まとまった語彙資料が見当たらない地点において，新たにハイブリッド遠隔型で語彙研究を始めるものである。この場合，4.2 の方法とは異なり，元となる原稿がないため，まずは基礎となる原稿を作成する必要がある。我々は原稿の作成にあたり，すでにある調査票を活用する方法（4.3.1）と，近隣方言の辞書等を活用する方法（4.3.2）をとった。

4.3.1 調査票を用いた原稿作成

新規に始める語彙研究の 1 つの方法として，原稿作成に調査票を用いるものがある。ここで紹介する事例では，国立国語研究所の共同研究プロジェクト「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（2016 年 4 月～ 2022 年 3 月；代表：木部暢子）」の語彙調査で使用している語彙調査票（未公開）を利用している。現在，西表網取方言で調査を実施している。

- 八重山語西表網取方言（2021 年 7 月～）：単語とその例文を収集し，語彙集を作成している（単独発話 269 点と用例 89 点の原稿と音声）。引き続き原稿作成・録音作業を依頼中である。図 6 は本地点で調査セットが返却された際のものである。

図 6　調査セット返却例（西表網取）

4.3.2 近隣方言の語彙資料を用いた原稿作成

もう 1 つの事例では，原稿作成にあたり，近隣方言の語彙研究資料を利用し，話者に翻訳してもらう方法をとっている。水納方言で調査を実施しており，本事例でのみ，研究者が録音を実施している。

- 宮古語水納方言（2020 年 1 月～）：筆者の一人（セリック ケナン）が収集中の語彙をまとめた宮古諸方言の語彙集を参考に，当該方言での翻訳を依頼した。2 か月で 5,000 語分の原稿が完成した（枠文を利用した発話 15,000 点を含む）。語数を 6,500 語まで増やし，2022 年 3 月に辞書が刊行された（セリック・大浦 2022）。

4.4 考察

以上，4.2 と 4.3 で，ハイブリッド遠隔型の語彙研究事例を見てきた。研究者が直接現地に行かないため，協力者が原稿作成に要した時間を正確に測ることは難しい。そのため効率について厳密に述べることはできないが，例えば波照間方言の事例であれば，200 分の録音に 3,000 語近く収録されている。西表網取方言の事例でも，20 分で 269 語と用例 89 点の録音を終えている。従って，平均して 1 時間でおよそ 800 ～ 900 語の録音を終えられるという結果が得られた。もちろん，録音までの準備（新規調査の場合には自身による原稿の作成）には相当な時間がかかっていると予想できるが，従来の，研究者が協力者を目の前にして行う面接調査の場合には研究者と協力者の双方に同じ時間がかかるため，単純に考えてもかかった時間は半分で済んでいると言える。3.5 で触れたスケーラビリティという観点からは，本事例がどれも 2020 年以降に開始していることからも，同時並行的に多地点の語彙研究を行えていると言えよう。

語彙の収集効率がよいという他に，データ処理作業の効率化という観点からもメリットがある。この観点は，主に研究者側の視点ではあるものの，原稿作成作業と録音作業を分けることによりその後のデータ処理作業の効率化が進んでいる。例えば，データ処理しやすいように録音しているため，研究者でなくても処理が可能となる。水納方言の事例に関して言うと，研究者が行う作業（ピッチの認定など）の前に，作業を平易化するために個々の発話に対して ID の紐付けを行った。この 15,000 点の音声に対する ID の紐付け作業は全て業務委託し，完了させた。効率化によって捻出された時間を使って，論文化や協力者へ語彙集を渡すといった，協力の可視化のスパンを短くすることができるだろう。あるいは，別地点での語彙研究の開始も可能である。これらのことは，協力の継続やひいては潜在的な協力者への働きかけにつながり，よいサイクルが生まれる可能性が高い。

よいサイクルは，協力者に対してだけではなく研究業界全体にも波及しうる。データのノイズ（調査から脱線した会話等）が少なくなり処理がしやすいということは，データベース等で公開しやすいというメリットがある。公開することにより第三者がアクセス可能となり，各研究者の研究結果の信憑性が高まるだろう。例えば，アクセント研究においては一次データを公開する慣習が確立されていないが，その理由として，公開に向けたデータ整理・処理にかける時間が取れ

ず，公開が困難になっていることが推測される。資料を提示することによって，アクセント研究などの研究全体の信憑性が高まると言える。

一方で，調査の完全な遠隔化はこれまでのところ不可能であったことも指摘する必要がある。2020 年から始まった新型コロナウイルス感染症の感染拡大により，意図せず調査が完全に遠隔化した地点があった（石垣宮良・小浜・西表網取）。特に録音作業に関しては，データが録音されていない場合や，筆者らの指示（枠文を用いた発音等）が伝わらない場合もあり，難航する場面も経験した[12]。この経験から言えることは，遠隔調査に限界がある点は認めつつも，ハイブリッド遠隔型の研究では，話者との関係を築いた上で，互いにできる部分はなるべく居ながらにして進めるということである。

5. おわりに

本論文では，南琉球の語彙研究を事例に方言の調査方法に関するメタ研究を試みた。交通の利便性や録音機器の性能，録音データを処理する技術は格段に上がったが，130 年以上前の田代（編）（明治年間）の語彙研究と比べても調査方法自体にこれといった大きな変化はなかったと言って差し支えないだろう。今回，従来の調査・研究手法に関するメタ研究を行うことにより，不足している観点は多くあるとは思うが，一定の客観的な評価指数を提示した。

結論として，過去 40 年間に行われた南琉球の語彙研究成果を調べた結果から，話者の内省による語彙収集の上，研究者と協力して調査を進める「ハイブリッド型研究」が質的にも量的にも最もバランスのとれた成果が期待できるということを述べた。さらに大規模実現可能性すなわち「スケーラビリティ（scalability）」という観点から，ハイブリッド型研究で行う作業のうち，遠隔でできるものをなるべく遠隔で行うという「ハイブリッド遠隔型研究」が効率的かつ合理的であるということを述べ，我々が実施している事例を共有し，その有効性を見てきた。

本研究は，語彙研究で採られるデータ収集に関する調査・研究方法に焦点を当てたわけであるが，一方で，得られた結果は単にハイブリッド型あるいはハイブリッド遠隔型研究の有用性を認識するだけではなく，作業を分担することの重要性をも示唆していると言える。すなわち，語彙収集の研究形態・調査方法という点のみならず，データ処理および研究プロジェクトの運営手法にも当てはまると考える。例えば筆者らは，南琉球の系統関係を紐解くために，南琉球各地の語彙収集を広く行うべくプロジェクトチームを組み，紹介した調査・研究手法で語彙収集を行っている。作業担当については，記述言語学分野で従来採用されている「地域分担」制ではなく「業務分担」制をとり，各自が得意な業務（様々なレベルにおける協力者とのやりとり，データ処理，タグ付け，分析など）を多量にこなすことで，同時並行的に多地点の調査が実現した。効率化というと，ともすると協力者との事務的な対応など，ヒトを対象とした人文研究とは相容れないという印象を持たれかねない。しかし，各人が得意な業務で互いをカバーし，時間を捻出することは，方言研究業界全体で問題として挙げられる「データの死蔵」（小西 2021）というリスクを未

[12] 調査が難航したケースもある一方で，話者が思い思いに原稿を執筆し録音することで，こちらが全く予想しなかった現象が観察されたケースもあった（セリック他 2022）。Serendipity と言えるケースである。

然に防ぎうるだけでなく，協力者に対しては語彙集の返礼対応の迅速化につながり，研究面だけでなく話者との信頼関係の構築にも有用な方法であると考える。

なお，当然のことながら，本論文は決して面接調査を否定しているわけではない。方言研究の基本は研究者型の面接調査であると筆者らも考えている。一方で，いわゆる伝統的な南琉球諸語は，今後 10 年から 20 年で急激に，かつ間違いなく消滅する。琉球列島で初めて経験するこのような状況にどう対応していくかについて，従来の方法に限界があるならば全く新しい方法を模索する必要がある。本論文では実例を提示していないが，研究者型の面接調査から始めて，ハイブリッド型，さらにハイブリッド遠隔型に切り替えていくといった方法も十分に考えられる。本研究の結果がこれからの伝統方言の記述方法について考えるきっかけになれば幸いである。

参照文献

秋永一枝（1960）「八重山方言一・二音節名詞のアクセントの傾向」『国語学』41: 121–125.

Blachon, David, Elodie Gauthier, Laurent Besacier, Guy-Noël Kouarata, Martine Adda-Decker and Annie Rialland (2016) Parallel speech collection for under-resourced language studies using the Lig-Aikuma mobile device app. *Procedia Computer Science* 81: 61–66.

セリック ケナン（2021）「下地皆愛方言のアクセント体系に関する予備的報告」『言語記述論集』13: 215–290.

セリック ケナン・麻生玲子・中澤光平（2021）「明治時代の八重山語語彙資料『海南諸島單語篇』の翻刻」『言語記述論集』13: 139–177.

セリック ケナン・麻生玲子・中澤光平（2022）「南琉球八重山語宮良方言の名詞アクセント資料」『国立国語研究所論集』22: 157–176.

セリック ケナン・麻生玲子・中澤光平・中川奈津子（2021）「川平方言語彙集」川平村の歴史編纂委員会（編）『新版 川平村の歴史』436–410. 沖縄：川平公民館.

セリック ケナン・大浦辰夫（2022）『みんなふつ語彙集』東京：国立国語研究所言語変異研究領域.

Gauthier, Elodie, David Blachon, Laurent Besacier, Guy-Noel Kouarata, Martine Adda-Decker, Annie Rialland, Gilles Adda and Grégoire Bachman (2016) LIG-AIKUMA: A mobile app to collect parallel speech for under-resourced language studies. *Interspeech 2016* (short demo paper), Sep 2016, San-Francisco, France. hal-01350062.

林由華・宮古総合実業高等学校生活福祉科・山本史・上田寛人（2021）「地域コミュニティとの協働による宮古島諸方言記録活動の事例報告」『日本方言研究会 第 112 回研究発表会発表原稿集』41–48.

平山輝男・大島一郎・中本正智（1967）『琉球先島方言の総合的研究』東京：明治書院.

平山輝男（編著）（1983）『琉球宮古諸島方言基礎語彙の総合的研究』東京：桜楓社.

平山輝男（編著）（1988）『南琉球の方言基礎語彙』東京：桜楓社.

法政大学沖縄文化研究所（1987a）「八重山・与那国島」『琉球の方言』11: 1–359.

法政大学沖縄文化研究所（1987b）「八重山・与那国島」『琉球の方言』12: 1–241.

池間苗（著）池間龍一・池間龍三（編）（2003）『与那国語辞典』私家版.

Ioannidis, John P.A. (2018) Meta-research: Why research on research matters. *PLoS Biology* 16(3): e2005468. https://doi.org/10.1371/journal.pbio.2005468.

Ioannidis, John P.A., Daniele Fanelli, Debbie Drake Dunne and Steven N. Goodman (2015) Meta-research: Evaluation and improvement of research methods and practices. *PLoS Biology* 13(10): e1002264. https://doi.org/10.1371/journal.pbio.1002264.

Jarosz, Aleksandra（2015）Nikolay Nevskiy's Miyakoan Dictionary. Reconstruction from the manuscript and its ethnolinguistic analysis. 博士論文 . Poznan: Adam Mickiewicz University Faculty of Modern Languages and Literatures.

加治工真市（1982）「琉球，小浜方言の音韻研究序説」仲宗根政善先生古稀記念論集刊行委員会（編）『琉球の言語と文化：仲宗根政善先生古稀記念』81–106. 沖縄：論集刊行委員会.

亀井孝・河野六郎・千野栄一（編）（1996）『言語学大辞典　第 6 巻（術語編）』東京：三省堂.

木部暢子（編）（2012）「消滅危機方言の調査・保存のための総合的研究　南琉球宮古方言調査報告書」国立国語研究所共同研究報告書. 東京：国立国語研究所.

小林隆（1988）「通信調査法の再評価」『方言研究法の探索』221–340. 東京：国立国語研究所.
小林隆（2004）『方言学的日本語史の方法』東京：ひつじ書房.
国語調査委員会（編）（1905）『音韻分布図』（国立国語研究所蔵）https://dglb01.ninjal.ac.jp/ninjaldl/bunken.php?title=oninbunpuzu（2021 年 12 月 4 日確認）
小西いずみ（2021）「コロナ禍のもとでの方言研究」『方言の研究』7: 37–55.
ローレンス ウエイン（2013）「竹富島方言アクセントと「系列別語彙」: 附 竹富島方言版「北風と太陽」」『琉球の方言』37: 1–24.
ローレンス ウエイン（2019）「竹富島方言アクセント（2）」『琉球の方言』43: 97–129.
宮良当壮（1930）『八重山語彙』東洋文庫叢刊（2）. 東京：東洋文庫.
宮良泰平（1975）『八重山方言の素姓』沖縄：宮良作.
仲原穣（2003）「石垣島宮良方言の音韻研究序説」『琉球の方言』27: 139–157.
仲原穣（2005）「小浜方言と宮良方言の音韻の比較研究」『琉球の方言』29: 107–120.
中松竹雄（1987）『琉球方言辞典』沖縄：那覇出版.
中本正智（1981）『図説琉球語辞典』東京：金鶏社.
Plonsky, Luke and Frederick L. Oswald (2012). How to do a meta-analysis. In: Alison Mackey and Susan M. Gass (eds.), *Research methods in second language acquisition: A practical guide*. 275–295. London: Basil Blackwell.
重野裕美・白田理人（2017）「高校生による奄美大島方言調査実践―鹿児島県立大島北高等学校「聞き書きサークル」活動の一環として―」『日本方言研究会 第 104 回研究発表会発表原稿集』25–28.
下地一秋（1979）『宮古群島語辞典』私家版.
田中弥生・柏野和佳子・角田ゆかり・伝康晴・小磯花絵（2017）「『日本語日常会話コーパス』構築における会話収録方法」『言語処理学会 第 23 回年次大会 発表論文集』481–484.
田中弥生・柏野和佳子・角田ゆかり・伝康晴・小磯花絵（2018）「『日本語日常会話コーパス』の構築―個人密着法に基づく会話の収録―」国立国語研究所『「日常会話コーパス」プロジェクト報告書 2』東京：国立国語研究所.
田代安定（編）（明治年間）『海南諸島單語篇　八重山之部』（東京大学理学図書館所蔵）https://iiif.dl.itc.u-tokyo.ac.jp/repo/s/tashiro/page/home（2021 年 12 月 4 日確認）
富浜定吉（2013）『宮古伊良部方言辞典』沖縄：沖縄タイムス社.

Possibilities of Meta-research on Japonic Descriptive Linguistics: Evaluation of 40 Years of Lexical Research on Southern Ryukyuan and Its Implications for Future Methodology

ASO Reiko[a] CELIK Kenan[b] NAKAZAWA Kohei[c]

[a]Meio University / Project Collaborator, NINJAL
[b]Adjunct Researcher, Research Department, NINJAL
[c]Shinshu University / Project Collaborator, NINJAL

Abstract

In this study, we evaluated the methodology and results of 40 years of lexical research on Southern Ryukyuan, and, based on this evaluation, we explored the feasibility of a more efficient lexical research methodology for the Japonic languages. First, we found that a hybrid approach to lexicography with native speakers and researchers working together was the most efficient, both qualitatively and quantitatively. Therefore, there is a strong possibility of active adoption of this approach in future lexical research. Second, lexical research on Southern Ryukyuan was found to be characterized by an extreme variability in data availability for each doculect. Despite adopting such a hybrid approach, however, it is not possible to resolve this data variability by relying solely on the highly time-consuming methodology of traditional fieldwork, especially given that all the languages under study are endangered. This problem can be resolved by devising novel remote data gathering techniques for lexical research. Therefore, we devised a new methodology that combines both the hybrid and remote approaches and report on its implementation in this article.

Keywords: Southern Ryukyuan, lexical research, meta-research, field method