



Title	基礎語彙による方言分類の諸問題 : Asai (1974) のアイヌ語方言データを最大限活用するために
Author(s)	深澤, 美香; Fukazawa, Mika; 小野, 洋平 他
Citation	北方言語研究, 15, 141-163
Issue Date	2025-03-20
DOI	https://doi.org/10.14943/112975
Doc URL	https://hdl.handle.net/2115/94619
Type	departmental bulletin paper
File Information	08_Fukazawa_Ono.pdf



基礎語彙による方言分類の諸問題
—Asai (1974) のアイヌ語方言データを最大限活用するために—

深澤 美香
(国立アイヌ民族博物館)

小野 洋平
(聖路加国際大学大学院公衆衛生学研究科)

キーワード：アイヌ語、文献学、特徴量抽出、文献資料論、基礎語彙統計学再考

1. はじめに

本研究は、基礎語彙によって得られる方言的な情報を余すところなく、最大限に活用することを目的とし、その目的を達成するための仮説的な試みを行うものである。服部・知里 (1960) の 19 方言・基礎語彙 200 項目のデータは、話者が少なくなりフィールド調査が困難になっていく時代のなかで、重要なアイヌ語方言の調査記録となっている。二人の言語学者が調査項目を揃えて、同時期に各地の話者を訪ねて採録した均質なデータは、アイヌ語の方言研究をはじめ歴史言語学的研究にも活用されてきた。とくに、本稿でとりあげる Asai (1974) の研究は、服部・知里 (1960) のデータを用いてクラスター分析を行い、アイヌ語の方言分類を提示したものであり、現在も非常に大きな影響力を持っている。

服部・知里 (1960) の研究は 1955 年 4 月に行われた調査¹ を基にしており、調査者・アイヌ語話者・調査日・調査経緯などが同論文や服部 (1964, 1999) にも詳細に記録されている。調査・研究の背景には Swadesh (1952) などをはじめとする基礎語彙による言語年代学の研究がある。服部は「基礎語彙統計学」の注で次のように述べている。

一般的な基礎語彙を統計的【…中略…】に研究するという意味ではなく、SWADESH の調査表 (本稿では新しい 200 項目のものをを用いる) などによって諸方言 (諸言語) を調査し、その資料を比較言語学的方法によって比較して、できれば同源の形態素を含む形式とそうでない形式とを区別して「残存語」の%を求め、更に可能ならば祖語からの分裂年代を推定しようとする研究方法のことである。服部はこれを「言語年代学」と呼んで来たが、この名称を用いると、祖語からの分裂年代を算定することのみが目的であるように誤解する向きもあるので、「基礎語彙統計学」という名称を用いる。

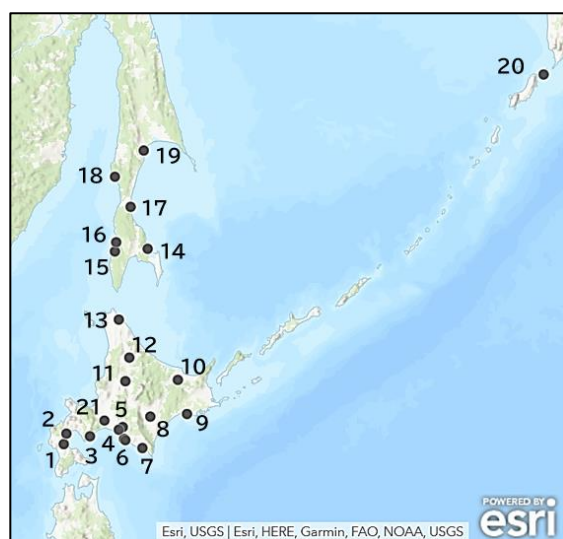
(服部・知里 1960 : 307)

Swadesh が提案する基礎語彙の項目や数が研究の進展とともに移り変わっていくなかで、服部の述べる Swadesh の「新しい 200 項目」がいずれを指しているのか不明であったが、深

¹ 国立アイヌ民族博物館には、1955年4月の服部と知里の調査をまとめた基礎語彙リスト (カーボンコピー版) が保管されている (管理資料名: 「アイヌ語の基礎語彙調査表 (原稿)」、収蔵資料番号: H30H17-44) (Fukazawa 2023、深澤2025)。

澤（2025）によって次のように特定された²。服部（1954）によれば、Swadesh は 1954 年 9 月 18 日に服部へ調査表「新調査語彙」を送付し、Swadesh (1952) で発表された 200 項目に 16 項目を追加、さらに前半の 100 項目がより安定的な残存率を示すという項目分類を行った。さらに、翌年の 1955 年 3 月 3 日付で Swadesh が服部宛に送付した最新の調査表の項目では、一部前半の項目と後半の項目で入れ替えが行われた。この 1955 年 3 月時点での 1～200 番目の 200 項目が服部・知里（1960）の 1955 年 4 月に開始された基礎語彙調査の 200 項目に対応している。なお、1955 年の調査表における前半の 100 項目は、Swadesh (1955) の「新しい 100 項目リスト (The new list of 100 items)」と呼ばれるものとも一致しているが、後半の 100 項目は同論文と完全には対応していない。

その後の Asai (1974) の研究では、服部・知里（1960）の基礎語彙データを基にしつつも、文献や自身の調査データによる追加や修正がみられる。最も大きな変更は、服部・知里（1960）の 19 方言に 2 方言（千歳方言・千島北部方言）が追加されたことと、基礎語彙 200 項目中 2 項目（「雨」(76. rain) と「かぐ」(187. smell)）がそれぞれ二分割されて 202 項目となっていることである。図 1 は Asai (1974) が対象にした 21 方言である。



1. 八雲、2. 長万部、3. 幌別、
4. 平取、5. 貫気別、6. 新冠、
7. 様似、8. 帯広、9. 釧路、
10. 美幌、11. 旭川、12. 名寄、
13. 宗谷、14. 落帆、15. 多蘭泊、
16. 真岡、17. 白浦、
18. ライチシカ、19. 内路、
20. 千島北部、21. 千歳

図 1：方言の地点

ここで改めて述べる必要があるのは、服部・知里（1960）も Asai (1974) も、基礎語彙 200（あるいは 202）項目の全てを分析に採用しているわけではないということである。ある方言において調査に対して疑問がある場合や話者が回答できなかった場合、調査漏れの場合、そして、すべての方言において語形が類似または一致する場合などは、その項目が分析されるデータから除外される傾向にある。とくに、語形の類似判断は、語根レベルで判断するの

² 深澤（2025）でも述べたが、深澤（2017）や Nakagawa & Fukazawa (2022) で解説している Swadesh の調査表と服部の基礎語彙調査表との関係から進展がみられたことになる。Swadesh の 1954 年の調査表において、Swadesh (1952) の 200 項目へ新たに追加された 16 項目は、3a. arm, 17a. breast, 20a. claw, 25b. dark, 30a. down, 59a. full, 59b. fur, 74a. horn, 74b. hot, 82a. knee, 89a. lip, 96a. milk, 96b. moon, 101a. navel, 113a. pus, 118a. ripe。さらに、Swadesh の 1954, 1955 年の調査表で 201～216 番にあたり、服部の基礎語彙調査から除かれた項目は、2. and, 3. animal, 9. because, 18. to breathe, 24. to count, 47. to fight, 51. to float, 76. to hunt (game), 112. to play, 113a. pus, 128. to scratch (as with fingernails to relieve itch), 146. some, 153. stick (of wood), 175. to turn (change one's direction), 193. to wipe, 196. woods。

か、音韻レベルで判断するのかなどによって研究者の方針が分かれるため、その類似判断の違いが項目数を含む分析対象のデータの違いにも反映され、方言分類の結果の違いを引き起こすことがわかっている（詳しくは、小野・深澤（2024）を参照）。

Asai (1974) は、110 項目を抜き出して自らの分析に採用しているが、いずれの項目なのか同論文中には明記されておらず長らく不明とされてきた。小野・深澤（2023）の研究によって Asai (1974) が採用した 110 項目が特定できたため、本研究では、分析に採用されなかった 92 項目に焦点を当てようとするものである。

本稿の構成は、2 節で Asai (1974) のデータをより適切に活用するための試みとして、本研究の仮説と、除外された 92 項目の基礎語彙データから統計分析にかけるデータを作成する方法について述べる。3 節では、2 節で述べた方法に沿って、具体的に Asai (1974) を用いて作成したデータについて紹介する。4 節ではデータ作成に関する課題について論じ、5 節において結果を踏まえた今後の展望について述べる。

2. 仮説とデータ作成の方法

本節では、本研究の仮説を提示したうえで、基礎語彙のデータのうち除外された 92 項目のデータを含めたデータ作成の方法について述べる。

2.1. 仮説

本研究では、既存の基礎語彙のデータを余すところなく最大限に活かすことにより、方言分類などの分析結果にも適切な影響を及ぼすという仮説を立てる。この仮説の検証をするために、本稿では、Asai (1974) で分析に採用されなかった項目があったことに着目した。前節で述べたとおり、Asai (1974) は服部・知里（1960）が扱う 19 方言に 2 方言追加し、基礎語彙 200 項目のうち 2 項目を分割して 202 項目にしている。Asai (1974) は 202 項目全てを計算に含めたわけではなく、そのうちの 110 項目を自らの分析に採用した。しかし、同論文中には、除外された 65 項目の情報のみが記載されており、小野・深澤（2023）において残りの除外された 25 項目を特定したという経緯がある（図 2）。

Asai (1974) が除外した 92 項目のうち、除外が明記されていなかった 25 項目は、その類似判断の規準においてすべての方言で語形が「類似」あるいは完全一致すると判断された項目であった。小野・深澤（2023）は、Asai (1974) の論文で語形同士の「類似」に関する記載を頼りに類似判断の基準を復元し、該当する 25 項目を特定した。Asai (1974) は、ある項目のなかで語形同士が殆ど類似、もしくは明らかに類似している場合、語形同士を括弧でくくって表示している。例えば、「魚」(19. fish) という項目は、Asai (1974) の論文で (1) のように示される。

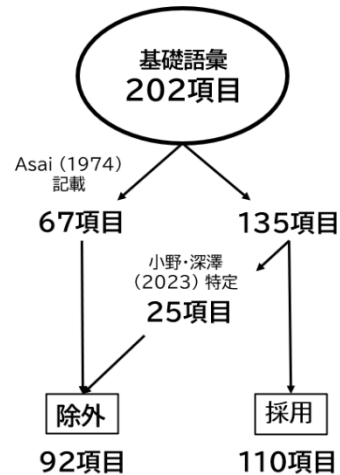


図 2：Asai (1974) において除外・採用された項目の特定の経緯

(1) 19. fish...(ciep (1-3, 21), cep (3-13, 15, 19-21)), ceh (14, 16-18)

(1) からわかることは、①「魚」³ (19. fish) の項目には *ciep*, *cep*, *ceh* という三つの語形があること。また、② *ciep* と *cep* という語形が一つの括弧でくくられており、Asai (1974) の類似判断において *ciep* と *cep* は「類似」とされること。さらに、③ *ceh* は *ciep*, *cep* に対し「類似」とはみなされないということである。なお、語形の後ろについている括弧付きの数字は、図 1 の方言番号に対応し、どの方言がその語形を有しているかを表している。このようにして、Asai (1974) の類似判断を一つずつ確認し、整理することにより、小野・深澤 (2023) では (2) で表される類似判断の基準を得た。

(2) Asai (1974) の類似判断の規準⁴ (小野・深澤 2023)

- a. 母音 1 タイプ : 1-1 ($V_1V_1 : V_1$), 1-2 ($V_1V_1 : V_1y$)
- b. 母音 2 タイプ : 2-1 ($V_1y : V_1$), 2-2 ($V_1y : V_1i$)
- c. 母音 3 タイプ : ($V_1V_2 : V_2$)
- d. 母音 4 タイプ : 4-1 ($C_1e : C_1i$), 4-2 ($C_1V : C_1$) / 4-2' ($CyV_1 : y \sim i$), 4-3 ($C_1 : C_1 + V$)
- e. 半母音タイプ : 1-1 ($y : w$), 1-2 ($iw : uy$)
- f. h- タイプ : 1-1 ($hV_1 : 'V_1$), 1-2 ($'V_1 : h- + V_1$)
- g. 子音タイプ : 1-1 ($C_1C_1 : C_1$), 1-2 ($C_1C_2 : C_2$)
- h. r タイプ
- i. 接辞タイプ
- j. 所属形タイプ
- k. 単複タイプ
- l. 重複タイプ
- m. 位置名詞タイプ

つまり、ある項目において、すべての語形同士の対応関係が上記 (2) の類似判断の規準のいずれかを満たす場合は、「すべての方言で類似」となる。Asai (1974) では、ある項目において「すべての方言で類似もしくは一致」となった場合、その項目は計算に含める必要のないものとして除外されるので、小野・深澤 (2023) は該当する除外対象の 25 項目を明らかにしたということになる。

類似判断の規準のうち a~h のタイプは、音対応に関する方言間の語形の規則性を捉えたものである。Asai (1974) はこれらの規則性をもつ語形を「類似」とみなして分析するデータから除外したが、本研究ではこの規則性こそが方言特徴としての情報となり、基礎語彙 202 項目の情報を最大限に活用することに繋がるのではないかと考えた。

以上のことから、本研究では計算から省かれた基礎語彙 92 項目を中心として、「類似」とみなされて失われた語形の差異の情報、つまり音対応による「規則性」について各方言デー

³ 項目の日本語訳は、服部・知里 (1960) を参考にした。

⁴ 本稿では、Asai (1974) で用いられるアイヌ語の j は y と表記する。

タの情報として取り入れることを試みる。また、これまでの方言分類などの結果にも適切な影響を与えるのではないかという仮説をたて、基礎語彙 92 項目と「類似」とみなされて失われた語形の差異の情報を取り込む方法について言語学的な側面から提案する。

2.2. データ作成の方法

既存の基礎語彙のデータを余すところなく最大限に活かすという目的を達成するために、本研究では (3) のような方法でデータを作成する。

- (3) 「語形」の情報（採用された 110 項目）＋「規則性」の情報（除外された 92 項目を中心とした類似判断の規準による情報）

本研究では、(2) の類似判断の規準のうち音対応による規則性 (a ~ h のタイプ) を、(3) の「規則性」の情報として扱う。類似判断の基準 (i ~ m タイプ) は文法要素に関わるものであり、方言調査段階による問題も多く含まれる可能性があることから除外した⁵。

データ作成にあたり、まず「語形」の情報として扱われていない除外された 92 項目（特に「すべての方言で類似」とされた 25 項目）に注目し、どの類似判断の規準によって「類似」と判断されて除外されたのかを再検証した。基本的な流れは (4) のとおりである。

- (4) 「規則性」をデータ化する方法

- ① 除外された 92（特に 25）項目の類似判断に着目：

83. ashes...(uuna (14-19), una (4-13, 20, 21))

- ② ①から抽出される類似判断のタイプと方言の種類の確認：

母音 1-1 タイプ...(V₁V₁ (14-19) : V₁ (4-13, 20, 21))

- ③ 除外された 92 項目のうち母音 1-1 タイプを含む項目が他にないか確認：

87. red...(huure (14-19), hure (1-13, 20, 21)) など

- ④ ③から抽出される方言情報の追加：

母音 1-1 タイプ...(V₁V₁ (14-19) : V₁ (1-13, 20, 21)) 方言 1, 2, 3 の追加

- ⑤ 採用された 110 項目に母音 1-1 タイプの「類似」と判断された語形がないか確認：

23. tree...(ni (1, 3, 4, 7-13, 20), nii (14-19))

- ⑥ ⑤から抽出される方言情報の追加：

母音 1-1 タイプ...(V₁V₁ (14-19) : V₁ (1-13, 20, 21)) 追加なし

(4) によって得られた「規則性」の情報は、V₁V₁(14-19) : V₁(1-13, 20, 21) となり、これを「語形」の情報のように扱うというのが本研究の試みとなる。また、除外された 92 項目以

⁵ 例えば、「結ぶ」(181. tie) は、sina (1, 3, 9, 15, 16, 20, 21), sinasina (7) という語形を持ち、1. 重複タイプに位置付けられるが、sinasina という語形をもつ類似方言が sina という形式を持たないのか不明である。服部 (1964 : 145) では、八雲方言と旭川方言も重複形の sinasina を持つことが報告されている。服部・知里 (1960 : 337) によれば「この項は、互いに比べることのできない形式が記録されている蓋然性があるので、削除する。「結ぶ」「くる」「しぼる」などの区別を被調査者らにわからせるのが困難だった。北海道方言に、標準語と違う形式が行われているためでもある」と述べている。

外にも、採用された 110 項目のなかに「類似」と判断されて失われてしまった特徴的な規則性の情報があれば「規則性」の情報として追加した。(5) は本研究で統計分析を適用した規則性のベースになっている類似判断のタイプである。(2) よりもタイプ数は少ないが、実際には「規則性」の情報を作成する際、語形の情報によってタイプを組み合わせたり、一つのタイプが複数の規則性にまたがることもあり、規則性と類似判断のタイプが一対一になるとは限らない。具体的には 3 節で見ていくことにする。

(5) 除外された 92 項目を中心として抽出した類似判断の規準

- a. 母音 1 タイプ : 1-1 ($V_1V_1 : V_1$), 1-2 ($V_1V_1 : V_1y$)
- b. 母音 2 タイプ : 2-1 ($V_1y : V_1$)
- c. 母音 3 タイプ : 3 ($V_1V_2 : V_2$)
- d. 母音 4 タイプ : 4-1 ($C_1e : C_1i$), 4-2 ($C_1V : C_1$) / 4-2' ($CyV_1 : y \sim i$)
- e. 半母音タイプ : 1-1 ($y : w$), 1-2 ($iw : uy$)
- f. h- タイプ : 1-1 ($hV_1 : 'V_1$), 1-2 ($'V_1 : h- + V_1$)
- g. 子音タイプ : 1-3 ($C_1C_2 : C_2C_2$)
- h. r タイプ

3. 規則性のデータ : 基礎語彙の項目と類似判断のタイプ

本節では、Asai (1974) のデータから抽出した「規則性」のデータについて解説する。また、その際に用いた基礎語彙の項目と類似判断のタイプについても提示する。

3.1. 規則性 1 : 母音 1-1 タイプ ($V_1V_1 : V_1$)

規則性 1 は、母音 1-1 タイプのうち、 V_1V_1 は樺太方言の長母音を表し、 V_1 は多くの北海道方言が昇り核アクセントを持った母音となる (知里 1973 [1942]; 服部 1967)。

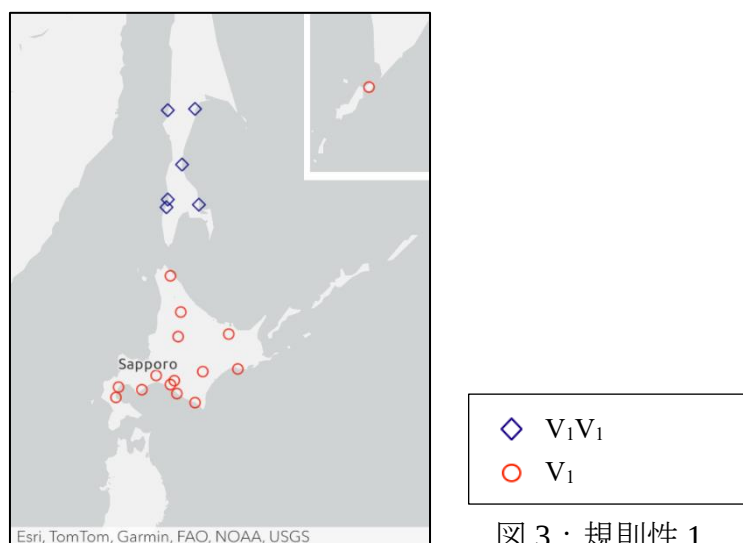


図 3 : 規則性 1

昇り核アクセントを持つ北海道方言では、開音節が二つ続く場合、原則的には第二音節にアクセント (CVCV́) がある。一方、例外的に第一音節にアクセント (ĆVCV) があるような場合、母音の長短アクセントを持つ樺太方言では CVVCV のように第一音節が長母音となり、これに相当するものを規則性 1 とした。一音節語も同様に樺太方言では長母音となる。

規則性 1 に含まれる語彙例は非常に多く、Asai (1974) において、すべての語形が類似または一致する項目として計算から除外されていることも多い。また、項目として除外されていなかったとしても、語形同士が一つの括弧に括られて「類似」とみなされる。

【規則性 1 : V₁V₁ (14-19) : V₁ (1-13, 20, 21)】

語彙例 (○は除外された 92 項目、そのうち●は除外された 25 項目)⁶ :

- 23. tree...**nii** (14-19) : **ni** (1, 3, 4, 7-13, 20)
- 24. seed...**pii** (14-19) : **pi** (1-9, 11-13, 21)
- 54. drink...**kuu** (14-19) : **ku** (1-13, 20, 21)
- 58. hear...**nuu** (14-19) : **nu** (1-13, 20, 21)
- 63. swim...**maa** (14, 15, 17-19) : **ma** (1-12, 20, 21)
- 68. sit...**aa** (14-19) : **a** (1-6, 8-13, 21)
- 71. say...**yee** (14-19) : **ye** (1-13, 21)
- 74. star...**noociw** (19) : **nociw** (1-9, 11, 12, 21)
- 81. smoke...**paa** (14-19) : **pa** (12, 13)
- 83. ashes...**uuna** (14-19) : **una** (4-13, 20, 21)
- 85. path...**ruu** (14-18), **tuu** (19) : **ru** (1-13, 20, 21)
- 87. red...**huure** (14-19) : **hure** (1-13, 20, 21)
- 100. name...**ree** (14-18), **tee** (19) : **re** (1-13, 20, 21)
- 107. here...**teeta** (14-18) : **teta** (1-5, 11-13, 21)
- 115. day...**toono** (15-19) : **tono** (13)
- 120. lake... **too** (14-19) : **to** (1-13, 20, 21)
- 130. lip...**caapus** (14-19) : **capus** (7-13)
- 135. milk...**toope** (14-19) : **tope** (1-13, 20, 21)
- 144. father...**aaca** (18) : **aca** (7, 9, 10, 13)
- 151. child...**poo** (14) : **po** (20)
- 167. smooth... **raarak** (15), **raarah** (14, 16-18), **taarak** (19) : **rarak** (1-7, 11, 13, 20, 21)
- 171. right...**siimon** (14, 15, 17, 18) : **simon** (1-9, 11, 13, 21)
- 183. fall...**haaciri** (14-19) : **hacir** (1-4, 7-13, 20, 21)
- 186. sing...**yuukara** (14-19) : **yukar** (20)
- 200. year...**paa** (14-19) : **pa** (1-14, 20, 21)

⁶ Asai (1974) の「寒い」(94. cold) の項目に *me rayke* (11, 14-16), *meeraiki* (17, 18) とあるのに対し、服部・知里 (1960 : 322) では樺太の落帆、多蘭泊、真岡方言 (14-16) が *me rayke* ではなく *meerayke* という語形で採録されている。よって、この項目は本来 *me rayke* (11), *meerayke* (14-16), *meeraiki* (17, 18) とすべき可能性があり、Asai (1974) に見られる語彙例としては除くことにした。

3.2. 規則性 2 : 母音 1-1 タイプ ($V_1V_1 : V_1$) + 母音 1-2 タイプ ($V_1V_1 : V_{1y}$) + 母音 2-1 タイプ ($V_{1y} : V_1$)

母音 1-1 タイプは、前述のとおり、 V_1V_1 が樺太方言の長母音を表し、 V_1 がそれに対する北海道方言の昇り核アクセントを持った母音を表す。規則性 2 は、ここに母音 1-2 タイプと母音 2-1 タイプに含まれる V_{1y} が追加され、それぞれが類似を表す場合である。

なお、Asai (1974) は「灰」(83. ashes) の項目で、uyna (1-3), (una (4-13, 20, 21), uuna (14-19)) とし、una と uuna を一つの括弧に入れて「類似」と明記しているが、uyna は括弧から外している。しかし、Asai (1974) の表記上には見られないものの、「灰」(83. ashes) の項目は、すべての語形で「類似」と判断され、計算から除外された 25 項目のうちの 1 項目であることが特定されている (小野・深澤 2023)。なお、uyna という語形は北海道の南西部方言に集中しており、地理的な分布状況からも una よりも古い語形であると考えられる。また、uyna は樺太方言の長母音の語形 uuna と北海道方言の昇り核アクセントの語形 una との歴史的な関係を考える上で、uuna と una の間を埋め、通時的な解釈を補完するような語形である可能性がある (Fukazawa 2018: 95–96)。「ここに」(107. here) を表す teyta (19) は樺太方言であり、「灰」(83. ashes) と比べて調査ミスや表記上の問題も検討すべきであるが、これも歴史的な解釈に参考となるデータと言える。

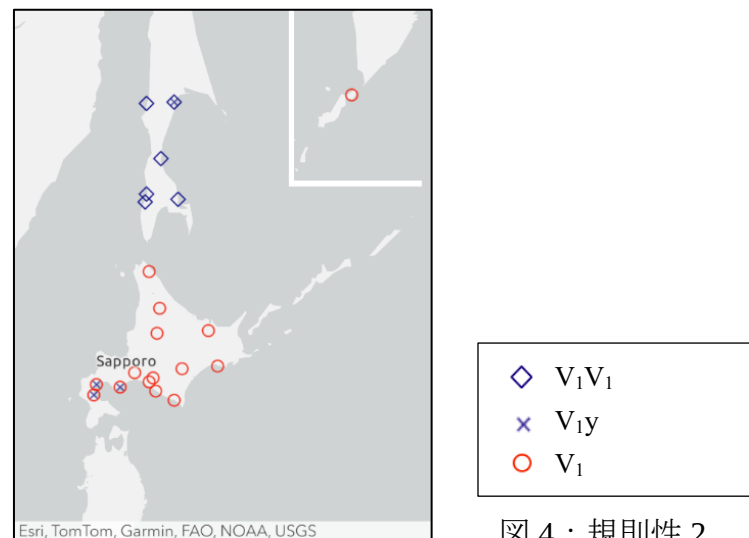


図 4 : 規則性 2

【規則性 2 : V_1V_1 (14-19) : V_{1y} (1-3, 19) : V_1 (1-13, 20, 21)】

語彙例 (●は除外された 25 項目) :

●83. ashes...uuna (14-19) : uyna (1-3) : una (4-13, 20, 21)

107. here...teeta (14-18) : teyta (19) : teta (1-5, 11-13, 21)

3.3. 規則性 3 : 母音 2-1 タイプ ($V_{1y} : V_1$)

ここでは、母音 2-1 タイプに当てはまるものとして、互いに類似する構成要素をもつ人称代名詞のみをとりあげる。人称代名詞は次のように分析的に解釈される。例えば、「私たち」

(3. we) の項目で 1 人称複数除外形の人称代名詞 **ciokay** は、**ci-** が 1 人称複数除外形の人称接辞、**oka(y)** が存在動詞の複数形であると考えられている（アイヌ語の人称代名詞の歴史的変遷は、中川 1988 : 248。本稿の 3.4 節も参照）。

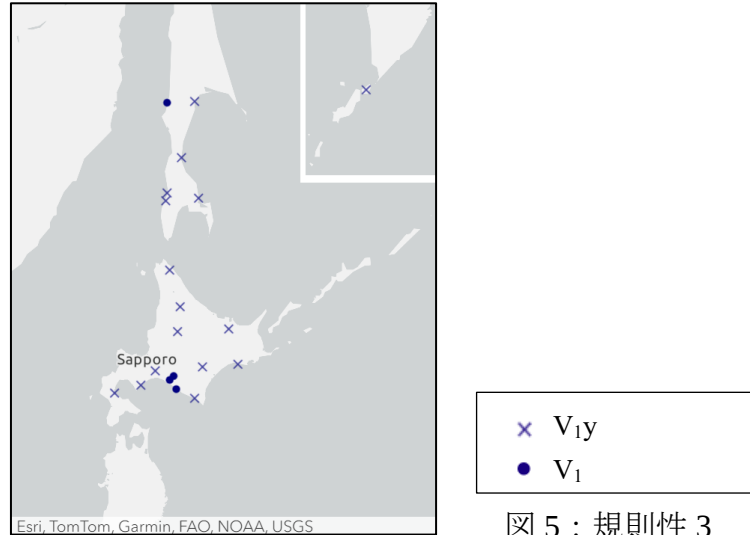


図 5 : 規則性 3

【規則性 3 : V_{1y} (1, 3, 7-17, 19-21) : V_1 (4-6, 18)】

語彙例 (○は除外された 92 項目) :

○3. we...**ciokay** (1, 3, 7-12), **cokay** (20), **cokayutari** (13) : **cioka**(6), **coka** (4, 5)

1. I...**ciokay** (21), **cokay** (13), **cookay** (16, 17, 19)

○3. we...**anokay** (9, 11-13, 17), **anokay utari** (13, 15), **anokay utara** (14, 16, 19), **aokay** (1, 3) :
anoka (18), **aoka** (4)

1. I...**anokay** (14-16, 19) : **anoka** (18)

101. ye...**eciokay** (1, 3, 7-10, 17), **ecokay** (20), **ecookay** (19), **ecookay utari** (15), **ecookay utara** (16, 19), **esokay** (11) : **ecioka** (4, 6)

3.4. 規則性 4 : 母音 3 タイプ ($V_1V_2 : V_2$) + 母音 1-1 タイプ ($V_1V_1 : V_1$)

人称代名詞の項目のなかにはもう一つの規則性が確認できる。母音 1-1 タイプの $V_1V_1 : V_1$ を $V_2V_2 : V_2$ と読み替えると、規則性 4 のような三つの対応関係をつくることができる。なお、「私」(1. I) の項目にある 1 人称単数形の人称代名詞 **kuani** はやはり分析的に考えられ、**ku-** が 1 人称単数形の人称接辞、**an** が存在動詞の単数形と考えられているが、最後の **-i** の解釈について「～こと」を表す形式名詞や名詞化接尾辞であるなど、さまざまに議論がある。また、前述の 1 人称複数除外形の人称代名詞 **ciokay** の **-y** についても、**an** と同様に **-i** として分析的に捉えるかどうかということについては議論があるところである（金田一・知里 1936 : 53 ; 田村 1971 : 3-5 ; 中川 1988 : 250）。

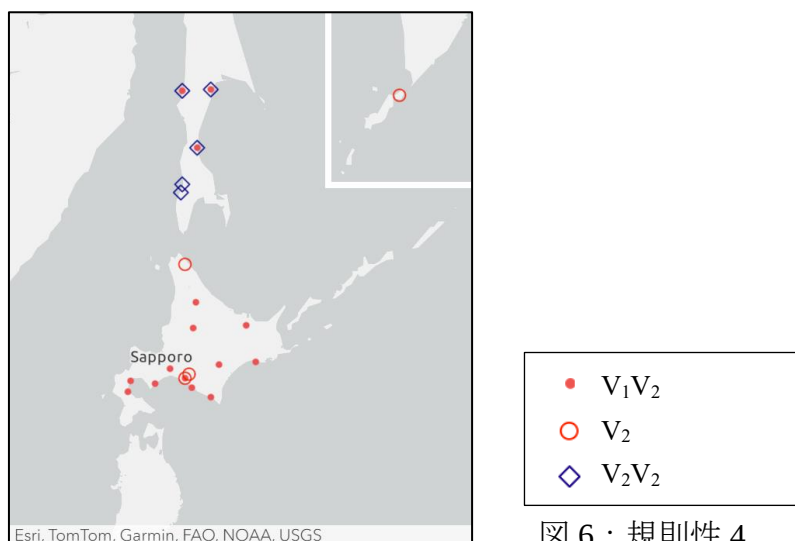


図 6 : 規則性 4

【規則性 4 : V_1V_2 (1-4, 6-12, 17-19, 21) : V_2 (4, 5, 13, 20) : V_2V_2 (15-19)】

語彙例 (○は除外された 92 項目) :

- 3. we...ciokay (1, 3, 7-12), cioka (6) : cokay (20), coka (4, 5), cokayutari(13) : cookayahcin(18)
 1. I...ciokay (21) : cokay(13) : cookay(16, 17, 19)
 1. I...kuani (1-3, 6-12, 17-19, 21) : kani(4, 5, 20),
 101. ye...eciokay (1, 3, 710, 17), ecioka (4, 6), eciokayahcin (18) : ecokay (20) : ecookay (19),
 ecookayutari (15), ecookayutara (16,19)

3.5. 規則性 5 : 母音 3 タイプ (V_1V_2 : V_2)

規則性 5 は、母音 3 タイプに相当する「魚」(19. fish) の項目を一つの規則性と捉えたものである。

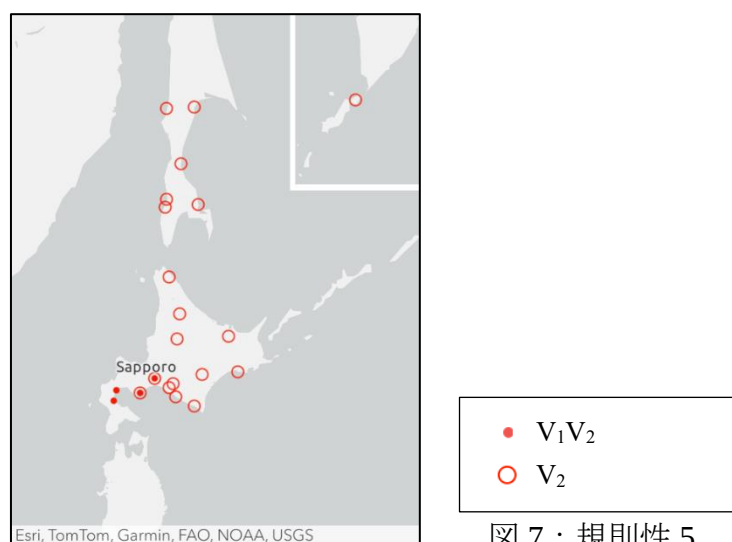


図 7 : 規則性 5

「魚」という語彙は、北海道の南西方言の **ciep** という語形から分析的に考えられてきた。**ci-** は 1 人称複数除外形人称接辞、**e** は「～を食べる」という 2 項動詞、**p** は「もの」を意味する形式名詞と考えられている（知里 1976 [1962] ほか）。

除外された 92 項目には含まれないが、語構成に人称接辞が含まれ、その母音が脱落するという現象自体は規則性 4 に含まれる母音 3 タイプと軌を一にするものである。規則性 5 は母音 3 タイプに含まれる人称代名詞ではない語彙例として別に扱った。規則性 5 については、本稿の 4 節と Ono and Fukazawa (2025) でも詳しく述べる。

【規則性 5 : $V_1V_2(1-3, 21)$: $V_2(3-21)$ 】

語彙例：

19. fish...**ciep** (1-3, 21) : **cep** (3-13, 15, 19-21), **ceh** (14, 16-19)

3.6. 規則性 6 : 母音 4-1 タイプ (C_1e : C_1i)

規則性 6 は、子音の **k** の後ろの母音が **e** と **i** で異なる母音 4-1 のタイプによるものである。

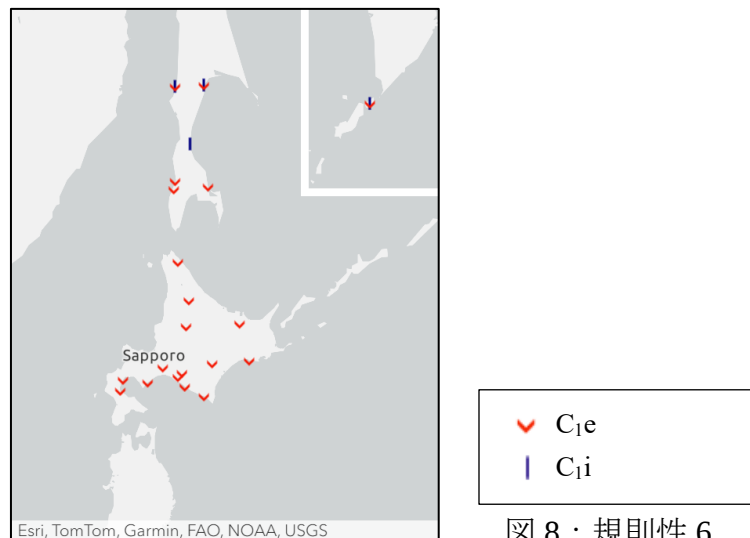


図 8 : 規則性 6

子音 **k** に続く母音が **i** になるものは樺太方言によく見られる。「寒い」(94. cold) は **me-rayke**「寒さ・～を殺す」と分析でき、「～を殺す」(62. kill) が含まれる語構成になっている。「寒い」(94. cold) と「～を殺す」(62. kill) に共通する **ray-ke(/-ki)** は「死ぬ・使役接尾辞」という語構成から成り立っているため、本規則性において対象となる部分は使役接尾辞である。一方、「汚い」(164. dirty) や「突き刺す」**pierce** (179. stab) の **-ke(/-ki)** も擬態を表す語根について動詞化接尾辞として働くものと考えられているが、「殴る」(177. hit) に関しては接尾辞としては分析できない。ここでは方言情報に共通性が見られたため、語彙例の一つとして取り上げた。

【規則性 6 : C₁e (1-16, 18-21) : C₁i (17-20)】

語彙例 (○は除外された 92 項目) :

62. kill...rayke (1-6, 10-16, 21) : rayki (17, 18), tayki (19)

94. cold...me rayke (11, 14-16) : meeraiki (17, 18), imeraykir (20)

164. dirty...icakkere (1-13, 19-21), icakere (14, 15), icahkere (16) : icahkire (17)

○177. hit...sitayke (13-16) : sitayki (19), isitayki (20)

○179. pierce (stab)...otke (1-6, 11, 12, 21), okke (7) : otki (20)

3.7. 規則性 7 : 母音 4-2 タイプ (C₁V : C₁) / 母音 4-2' タイプ (yV₁ : y ~ i)

規則性 7 は千島北部方言がその他の方言とは異なる特徴を示す母音 4-2 および 4-2' タイプを包含する。Asai (1974) においてその他の方言と同一の括弧にくくられて「類似」とされている事例も多く見られる。

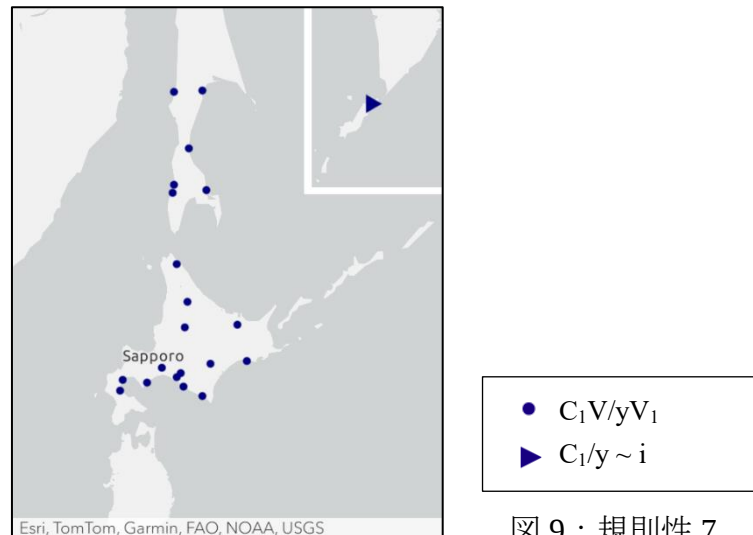


図 9 : 規則性 7

【規則性 7 : C₁V/yV₁ (1-19, 21) : C₁/y ~ i (20)】

語彙例 (●は除外された 25 項目) :

49. belly...pise (9) : pse (20)

53. liver...uraka (15-18) : urak (20)

123. wash...huraye (1-7, 13-19, 21), uraye (8-12)⁷ : uray (20)

●172. old...husiko (1-19, 21) : husko (20)

174. pull...etaye (1-12, 21) : etay (20)

175. push...oputuje (1-6, 8-11, 21), opituye (12, 13), opicuye (7) : optuy (20)

180. dig...poye (14-19) : poi (20)

⁷ Asai (1974) では uraye (6-12) となっているが、服部・知里 (1960) のデータを参照する限り uraye (8-12) となるため、本研究のデータとしては後者を採用した。

3.8. 規則性 8 : 半母音 1-1 タイプ (y : w)

規則性 8 の半母音タイプは語末もしくは語中の環境で y か w として表れるものである。ただし、語中であっても語構成上わたり音ではない場合もある。例えば「はらわた」(133. guts) は、tuy(/tuw)-or-o-p(/-h)「腹腔 (?)・のなか・～にある・もの」と分析できる。「切る」(153. cut) は、1 項動詞の tuy「切れる」に -e という使役接尾辞がついたものと解釈されるが、多蘭泊方言 (15) で「切れる」という 1 項動詞が tuw となるかは一考を要する。

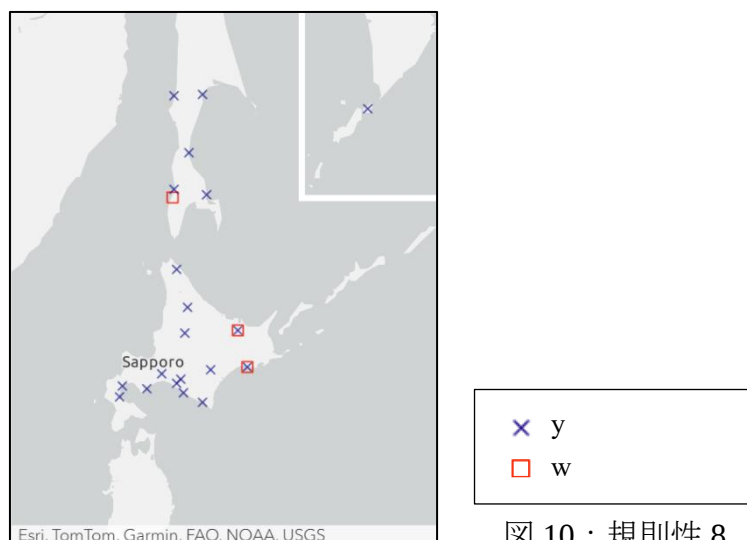


図 10 : 規則性 8

【規則性 8 : y (1-14, 16-21) : w (9, 10, 15)】

語彙例 :

- 17. man...okay (7-11, 20, 21) : okkaw (15)
- 76. rain (the rain)...ruyanpe (7-9, 11-13) : ruwanpe (9, 10)
- 133. guts...tuyorop (6), tuyoroh (16-18) : tuworop (15)
- 153. cut...tuye (1-6, 8-14, 16-19, 21), cuye (7) : tuwe (15)

3.9. 規則性 9 : 半母音 1-2 タイプ (iw : uy)

このタイプは iw と uy が音位転倒を起こしている例である。基礎語彙 200 項目は「星」(74. star) と「突き刺す」(179. pierce (stab)) に関わるが、基礎語彙以外の語彙にも見られる特徴であることが知られている (Kirikae 1988: 29-38)。

【規則性 9 : iw (1-9, 12-17, 19 21) : uy (9, 10)】

語彙例 (○は除外された 92 項目) :

- 74. star...nociw (1-9, 11, 12, 21), noociw (19) : nocuy (9)
- 179. pierce (stab)...ciw (3, 4, 8, 9, 12-17, 19) : cuy (10)

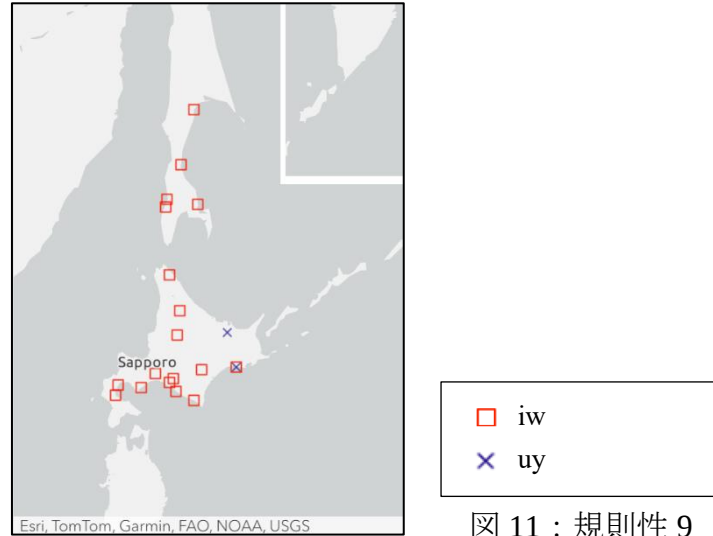


図 11：規則性 9

3.10. 規則性 10：h- 1-1 タイプ (hV_1 : $'V_1$)

h- 1-1 タイプは、基礎語彙 200 項目以外の語彙にも広く見られ、それらは北海道の北東と南西に分かれるような分布を示す。歴史的には語頭の h- が北海道の北東方言で脱落したと考えられ、南西と北東の境界線上にある地点ではどちらか一方ではなく両方の語形を持っていることも多い（方言境界については、深澤・小野（2024）を参照）。

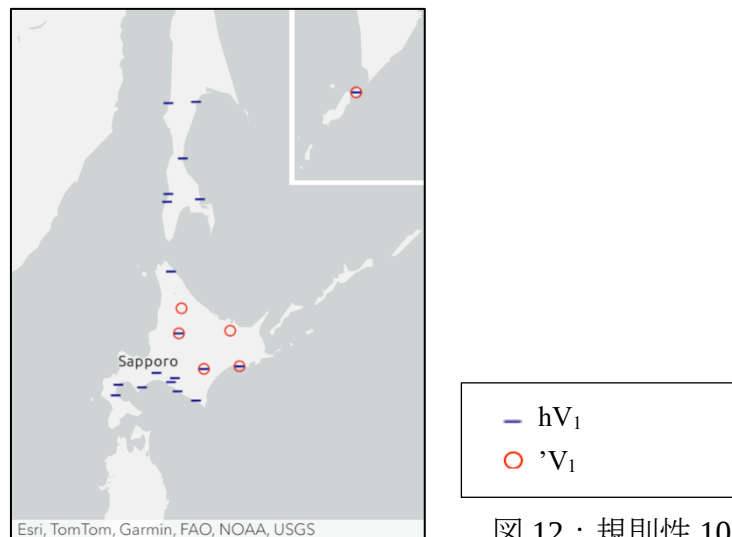


図 12：規則性 10

【規則性 10： hV_1 (1-9, 11, 13-21), $'V_1$ (8-12, 20)】

語彙例（○は除外された 92 項目）：

64. fly...**hopuni** (1-6, 13), **hopunpa** (21) : **opuni** (9, 11, 12)

123. wash...**huraye** (1-7, 13-19, 21) : **uraye** (8-12), **uray** (20)

145. husband...**hoku** (1-9, 11, 13, 15, 16, 18, 20, 21), **hoko** (14, 17, 19) : **oku** (8-12))

151. child...**hekaci** (1, 13, 15-19, 21), **hekattar** (2-7, 11) : **ekaci** (8-11), **ekattar** (11, 12)
 ○194. down...**heras** (1, 2, 21), **herasi** (3-7) : **eras** (11, 12), **erasi** (8, 9)
 ○195. up...**herekas** (1,2), **herikasi**(3-7), **herikas**(21) : **erekas**(9), **erikasi**(8, 10), **erikas** (9, 11, 12)

3.11. 規則性 11 : h- 1-2 タイプ ('V₁ : h-+V₁)

規則性 11 は、同じ h- タイプの規則性 10 とは h- の歴史が異なる。規則性 11 に含まれる h- 1-2 タイプは、北海道の周縁部にある方言で局所的に h- が現れるものを示す (cf. Nakagawa and Fukazawa 2022)。

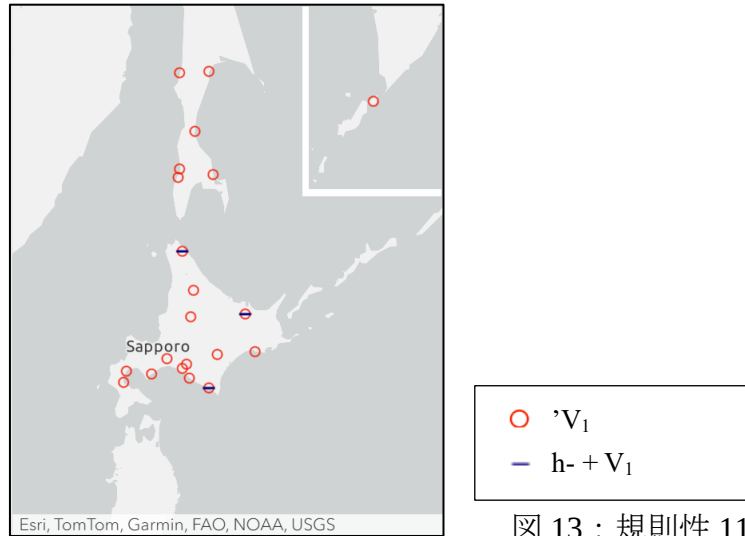


図 13 : 規則性 11

規則性 10 に含まれる h- 1-1 タイプは語頭の h- が歴史的に脱落したものと考えるが、規則性 11 に含まれる h- 1-2 タイプは h- のつく語彙が非常に限定されていることから、歴史的には h- があとから付加されたと考えられる。なお、「霧」(116. fog) は樺太の長母音をもつ語形 *uurara* (14-19) もあるが、厳密には母音 1-1 タイプと組み合わせることになることから、語彙例からは除外した。

【規則性 11 : 'V₁ (1-21) : h- + V₁ (7, 10, 13)】

語彙例 (○は除外された 92 項目) :

45. claw...**am** (1-9, 11-21) : **ham** (10)
 78. sand...**ota** (1-4, 6, 8-21) : **hota** (7)
 116. fog...**urar** (1-9, 11-13, 20, 21) : **hurar** (10)
 ○137. flower...**apappo** (9, 10) : **hapappo** (13)

3.12. 規則性 12 : 子音 1-3 タイプ (C₁C₂ : C₂C₂)

規則性 12 は、子音 1-3 タイプのなかでも、北海道南西方言で C₁C₂ であるものに対し、北海道北東方言で C₂C₂ となるようなものを対象にした。子音 C₂ は /p, t, k/ に相当し、北海

道北東方言では逆行同化が起こった語形を示す。また、樺太方言では C_1C_2 の C_1 が /h/ になる方言があることからそれを hC_2 とし、規則性 12 のデータとして C_1C_2 とは別に表した。

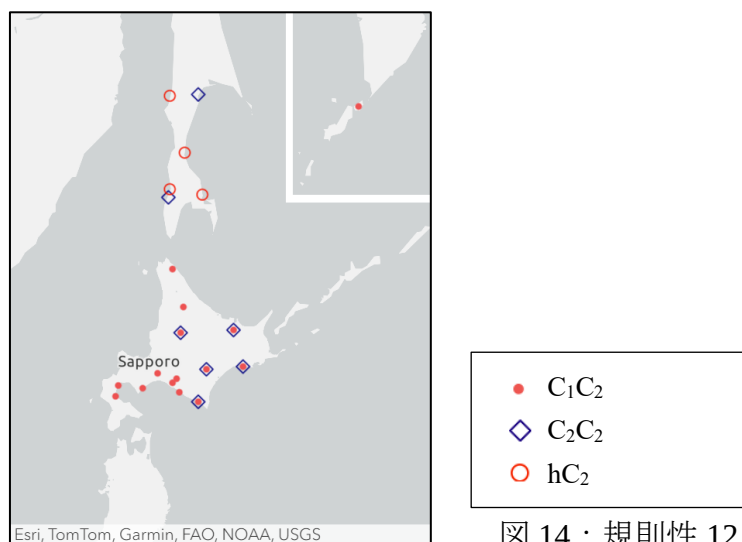


図 14：規則性 12

【規則性 12 : C_1C_2 (1-13, 20, 21) : C_2C_2 (7-11, 15, 19) : hC_2 (14, 16-18)】

語彙例 (○は除外された 92 項目) :

- 65. walk...**apkas** (1-13, 20, 21) : **akkas** (15, 19) : **ahkas** (14, 16-18)
- 67. lie...**hotke** (1, 3-5, 9, 11-13, 21) : **hokke** (8-11)
- 76. rain...**apto** (3-6, 21) : **atto** (15, 19) : **ahto** (14, 16-18)
- 179. pirece (stab)...**otke** (1-6, 11, 12, 21), **otki** (20) : **okke** (7)
- 195. up...**rik ta** (9, 13) : **ritta** (7, 9, 15)

3.13. 規則性 13 : r タイプ

規則性 13 は、r タイプのうち、北海道方言で音節末の r を持つ語形に対し、樺太方言では r の後ろに母音が後続するものを対象にする。北海道と樺太で共通に r の音を持っていれば生じることから、規則性 1 と同様に多くの項目を包含する。なお、千島北部方言に関しては、rV と r の両方が見られることもあるようであるが、使い分けがあるのではなく鳥居 (1903) における表記上の問題ではないかとも考えられる。その他の千島北部資料のなかでは r の後ろの母音が表記されないことも多いことから、語彙例では rV を星印 (*) 付きで表した (Bugaeva and Sato 2021; 村山 1971)。

【規則性 13 : rV (14-19, 20*) : r (1-13, 20, 21)】

語彙例 (○は除外された 92 項目、そのうち●は除外された 25 項目)⁸ :

⁸ 使役接尾辞や所属形によって r に母音が後続するものは、語彙例から省いている。

22. louse...**uriki** (14-16) : **urki** (1-13, 21)
32. grease...**kiripu** (14, 19), **kiripo** (15, 16), **kirupu** (17, 18) : **kirpu** (1-6, 20, 21), **kirpo** (7-13, 21)
- 39.**kisara** (14, 16, 17, 19), **kisaru** (15, 18) : **kisar** (1-13, 20, 21)
- 57. see...**nukara** (14-19) : **nukar** (1-13, 20, 21)
- 60. sleep...**mokoro** (14-19) : **mokor** (1-13, 20, 21)
90. white...**tetara** (14-19) : **retar** (1-13, 20, 21)
- 96. new...**asiri** (14-19, 20*) : **asir** (1-13, 21)
- 97. good...**pirika** (14-19, 20*) : **pirka** (1-13, 20, 21)
101. ye...**eani utara** (14, 16) : **eani utar** (2,12)
116. fog...**uurara** (14-19) : **urar** (1-9, 11-13, 20, 21), **hurar** (10)
- 125. worm...**kikiri** (14-19) : **kikir** (1-13, 20, 21)
126. back...**seturu** (14, 16-19, 20*), **seturi** (15) : **setur** (1-6, 8-13, 21), **secur** (7)
159. thin...**kapara** (14-19) : **kapar** (1-13, 21)
- 162. dull...**enukara** (14, 19) : **enukar** (1-13, 21)
- 183. fall...**haciri** (14-19) : **hacir** (1-4, 7-13, 20, 21), **hacir** (21)
- 186. sing...**yukara** (14-19), **yukari** (20) : **yukar** (20)
- 197. dust...**sirima** (17, 18) : **sirma** (3, 11)

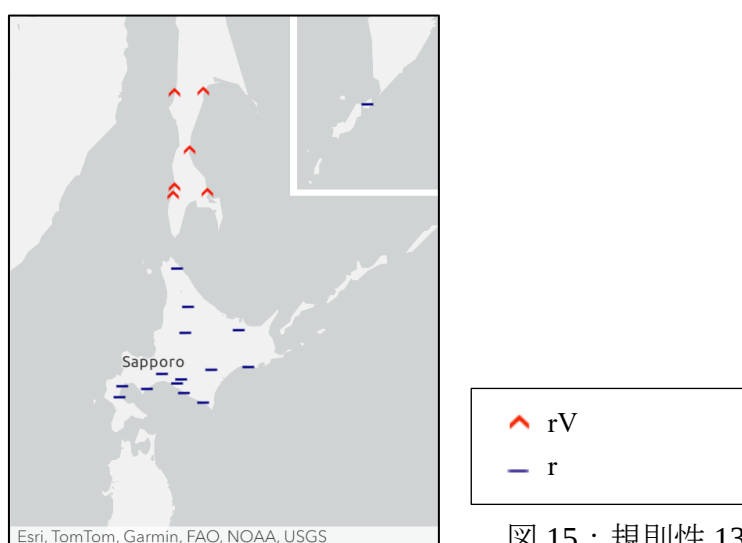


図 15 : 規則性 13

4. 規則性のデータ作成に関わる課題

本研究では、(3) のデータについて統計分析を適用した。

(3 再掲) 「語形」の情報(採用された 110 項目) + 「規則性」の情報(除外された 92 項目を中心とした類似判断の規準による情報)

規則性は除外された 92 項目を中心に作成し、2 節の (4) でも示したとおり複数の語彙例から規則性に対する方言情報を抽出するという方針をとった。しかし一部、例外的に作成した規則がある。その一つが「魚」(19. fish) から作成した規則性 5 である。この項目は、除外された 92 項目には入っておらず、たった一つの項目だけで規則性 5 を構成する。

1 節で述べたとおり、「魚」(19. fish) は、ciep, cep, ceh の三つの語形を持っており、Asai (1974 : 69) では、ciep と cep が「類似」とされ、ceh は（除外された 25 項目にも含まれていないことから）ciep や cep に対して「非類似」であることがわかる。

(1 再掲) 19. fish...(ciep (1-3, 21), cep (3-13, 15, 19-21)), ceh (14, 16-18)

しかし、北海道方言でみられる音節末の -p/-t/-k と樺太の多くの方言で見られる -h は、この他の多くの語彙にも確認でき、一つの規則性にまとめることが可能である。実際、「魚」(19. fish) の項目は服部・知里 (1960) では、「類似」として分析から除外されている。もし (1) の「非類似」の情報から規則性の情報として抽出するとすれば、(5) のようにまとめることができる。

(5) 「非類似」の情報から抽出した規則性のデータ

規則性 14 : V₁C(-p/-t/-k) (1-13, 15, 19-21) : V₁C(-h) (14, 16-18)

語彙例の一部（○は除外された 92 項目）：

19. fish...ciep (1-3, 21), cep (3-13, 15, 19-21) : ceh (14, 16-18)

○20. bird...cikap (1-13, 15, 19, 21) : cikah (14, 16-18)

66. come...ek (1-13, 15, 19-21) : eh (14, 16-18)

155. narrow...ohotne (13), ohupne (19) : ohohne (14, 16), ohuhne (17, 18)

【参考】Asai (1974) が「類似」とした情報から抽出した規則性のデータ
規則性 5 : V₁V₂ (1-3, 21) : V₂ (3-21)

語彙例：

19. fish...ciep (1-3, 21) : cep (3-13, 15, 19-21), ceh (14, 16-18)

本研究では、「類似」の情報から規則性 5 を取り出して、「規則性」のデータとして扱い、110 項目の語形のデータはそのまま「語形」のデータとして扱った。仮に、規則性 14 のように 110 項目に含まれる「非類似」のデータからも規則性を抽出するのであれば、110 項目の「語形」データを「規則性」のデータに変換することも可能になってくる。例えば、本研究で「語形」のデータとして扱われている「来る」(66. come) の項目は規則性 14 の方言情報とも完全に一致しており、「魚」(19. fish) と合わせて「規則性」のデータとしてまとめていくことも可能である。

以上のことを踏まえると、語根レベルで異なる項目を「語形」のデータとして扱い、音韻

レベルで異なる項目を「規則性」のデータとして扱っていくということが、論理的には可能となる。「語形」と「規則性」のデータを同列として扱うことや、複数の語彙にまたがる規則性（例えば、規則性 1 や規則性 13）と目下一つの項目にしか見られない規則性（例えば、規則性 5）を同列に扱うべきかどうかというのが次に考えるべき課題である。本稿ではすべて同列に扱ったことになるが、本研究の目的であった「基礎語彙によって得られる方言的な情報を余すところなく、最大限に活用する」ということは、将来的な課題を明確にしたなかで、概ね達成したと考えられる。

5. 結果と今後の展望

本研究では、Asai(1974) において計算から除外された基礎語彙 92 項目を中心として、「類似」とみなされて失われた語形の差異の情報、つまり音対応による「規則性」について各方言データの情報として取り入れることを試みた。また、これまでの方言分類などの結果にも適切な影響を与えるのではないかと仮説を立て、基礎語彙 92 項目と「類似」とみなされて失われた語形の差異の情報を取り込む方法について提案した。

結果の詳細については、Ono and Fukazawa (2025) にて述べるが、Asai (1974) のアイヌ語諸方言の分類結果について、それぞれの方言グループがより明確に区別された。言い替えると、Asai (1974: 100) では、北海道方言を北部方言、東部方言、西部方言、南部方言⁹ の四つに分類しているが、この分類がより弁別的になったと言える（図 16）。例えば、Asai (1974) によって北部方言に分類される宗谷方言や、南部方言に分類される様似方言は、それぞれ東部方言や西部方言に含める研究者も多い。そんななか、Asai (1974: 100) が宗谷方言と様似方言をそれぞれ北部と南部として別に扱うという判断をしたのは、クラスター分析の結果のみならず Asai 自身の直観にも合っていたのではないかと考える。

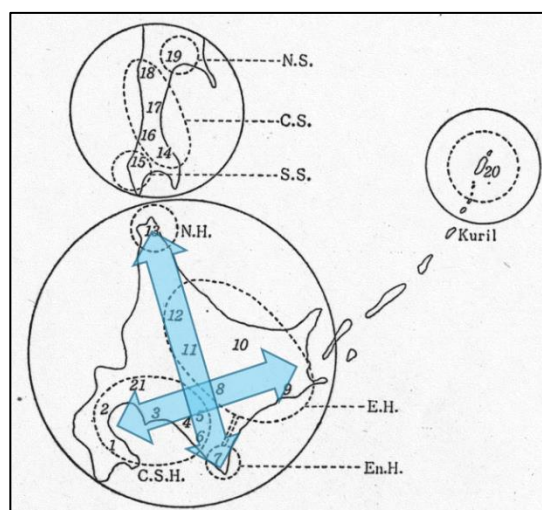


図 16 : Asai (1974: 100) の方言分類と北海道方言の 4 分類

⁹ Asai (1974: 100) の方言分類をわかりやすくするために、直訳を使用しなかった。北部方言 (North Hokkaido)、東部方言 (East Hokkaido)、西部方言 (Central South Hokkaido)、南部方言 (Eastern Hokkaido) となる。

また、さらなる展開として、基礎語彙 200（あるいは 202）項目に対応する語形がほぼ含まれない場合も、規則性の情報を抽出して統計分析を適用することが考えられる。近世アイヌ語資料の一つである『蝦夷生計図説』¹⁰ は、アイヌの文化や風習等を絵図とともに解説をつけたものであり、イナウ（木幣）や衣服の種類、家屋や舟の種類やつくり方などの内容が含まれる。基本的には民具名称などが多く記載されており、文や句レベルやカタカナの濁点や半濁点の有無などの表記違いも見られるが、300～400 ぐらいの語彙を拾うことができる。そのなかから、基礎語彙を合成語からも含めて拾おうとすると 30 弱の語彙数となり、分析に活用できる語彙は全体の 10 分の 1 未満である¹¹。

- 8. not... 「シヨ」 [so(mo)?] (合成語中)
- 16. woman... 「メノコ」 [menoko] (合成語中、分析あり)¹²
- 19. fish... 「チエツプ」 [ciep, cep] (合成語中、分析あり)
- 20. bird... 「チカツ」 [cikap] (合成語中、分析あり)
- 23. tree... 「チクニ」 [cikuni]、「ニ」 [ni] (合成語中、分析あり)
- 24. seed... 「ヒ／ピ」 [pi] (分析あり)
- 27. bark... 「カフ／カプ」 [kap, kapu] (合成語中、分析あり)
- 37. hair... 「ケ（フ）（?）」 [ke, kep?, kepu?, (kepur?)] (合成語中、分析あり)
- 38. head... 「シヤハ」 [sapa, (sapane?)] (合成語中で「人の上に立つ」とある)
- 41. nose... 「イト」 [etu] (分析あり)
- 55. eat... 「イヘ／イベ」 [ipe] (合成語中、分析あり)
- 58. hear... 「ヌウ」 [nu]
- 70. give... 「コレ」 [kore]
- 75. water... 「ワツカ」 [wakka]
- 79. earth... 「トイ」 [toy] (分析あり)
- 82. fire... 「アヘ／アベ」 [ape] (合成語中、分析あり)
- 84. burn... 「ウフイ」 [uhuy] (分析あり)
- 97. good... 「ヒリカ」 [pirka, pirika] (副詞化接尾辞のついた形式中、分析あり)
- 99. dry... 「シヤツハケ」 [satke, sakke] (分析あり)
- 129. wing... 「ラフ、ラブ」 [rap] (合成語中、分析あり)
- 138. grass... 「ムン」 [mun] (分析あり)
- 140. in, 141. at... 「ヲツタ」 [otta (< or ta)]
- 153. cut... 「トイ」 [tuy, tuye?] (分析あり)

¹⁰ 作者は秦檜丸（村上島之丞）、間宮林蔵により増補。『蝦夷島図説』『蝦夷常用集』『蝦夷産業図説』などと呼ばれることもある。多くの写本や類本をもつ。

¹¹ カタカナは『蝦夷生計図説』（東京大学総合研究博物館所蔵資料）の原典ママであり、ローマ字はカタカナから推定される表記である。

¹² 「分析あり」という注記は、「これをムンウフイと称す。ムンは草をいひ、ウフイは焼く事をいひて」のように基礎語彙が語構成要素となり、それぞれの要素に説明がついているという意味である。ただし、名詞と 2 項動詞の組み合わせの場合は、目的語として抱合されて 1 項動詞のようにふるまっているのか、抱合されていないのか判断が分かれるため、「合成語中」とは注記しなかった。

165. bad...「ウエン」[wen]

179. pierce (stab)...「ヂ」[ci, ciw?] (合成語中、「小さなものが高く突出する」とある)

195. up...「リキタ」[rik ta] (句中、分析あり)

上記の基礎語彙によって北海道方言の特徴を持つことは判断できるが、それ以上になると「語形」データとしては数が少なく、量的に検討することが困難になる。しかし、本研究のように「規則性」をデータとして抽出できれば、基礎語彙以外からも方言情報を取得し、服部・知里(1960)やAsai(1974)から抽出した規則性のデータと統合して扱えるようになる可能性がある。さらに、本研究は基礎語彙の語形のみで頼らない方言分類を考えていく出発点として、基礎語彙から規則性の情報を抽出し、データ化したという点が重要である。Swadeshらがこれまで基礎語彙リストを幾度となく更新してきた背景には、統計分析に使用する言語データのうちノイズとなる借用語などの情報を予め取り除いておく必要性があった。要するに、無限に情報を増やしていくよりも、基礎語彙リストのように情報を予め取捨選択することで必要な情報を引き出そうとしてきたわけである。本研究がまず基礎語彙に絞って規則性の情報をデータ化したのは、基礎語彙がもつ「語形」のみをデータとする枠組みから一歩外に出るという重要なステップとなった。将来的には、音韻レベルだけではなく、形態素や統語レベルの文法的特徴も同様にデータ化し、基礎語彙に相当するような一定の規準でもって新たな方言データを作成することを構想している。

謝辞

本研究は、国立アイヌ民族博物館調査研究プロジェクト 共同研究「蝦夷生計図説の研究：ウェブサイト構築と公開に向けて」(2024K01、代表者：深澤美香)の研究成果の一部である。日本北方言語学会第7回大会の口頭発表における質疑および本稿の2名の査読者からは、本研究に対する有益なコメント等を頂いた。ここに記して御礼申し上げる。

参考文献

- 小野洋平・深澤美香(2023)「アイヌ語諸方言の語形の類似に関する基礎データの復元：論文に書ききれなかった研究者の判断・思考に迫る」『北方言語研究』13: 213-246.
- 小野洋平・深澤美香(2024)「比較不可能だったアイヌ語方言分類：統計的方言分類を類似判断の点から再考する」『アイヌ・先住民研究』4: 93-126.
- 金田一京助・知里真志保(1936)『アイヌ語法概説』岩波書店.
- 田村すゞ子(1971)「アイヌ語沙流方言の人称代名詞」『言語研究』59. (再録：(2001)『アイヌ語考④』263-276. ゆまに書房.)
- 知里真志保(1973 [1942])「アイヌ語法研究：樺太方言を中心として」『知里真志保著作集』3: 455-594. 平凡社.
- 知里真志保(1976 [1962])「分類アイヌ語辞典動物編」『知里真志保著作集別巻I：分類アイヌ語辞典植物編・動物編』平凡社.
- 鳥居龍蔵(1903)『千島アイヌ』東京：吉川弘文館.
- 中川裕(1988)「アイヌ語千歳方言の人称代名詞とその歴史的的位置」『東京大学言語学論集'88』

- 239-253. 東京大学文学部言語学研究室.
- 秦憶丸『蝦夷生計図説』(東京大学総合研究博物館所蔵資料)
- 服部四郎(1954)「言語年代學」即ち「語彙統計學」の方法について: 日本祖語の年代『言語研究』26, 27: 29-37.
- 服部四郎(編)(1964)『アイヌ語方言辞典』岩波書店.
- 服部四郎(1967)「アイヌ語の音韻構造とアクセント」『音声の研究』13: 207-223.
- 服部四郎(1999)『日本語の系統』岩波文庫.
- 服部四郎・知里真志保(1960)「アイヌ語諸方言の基礎語彙統計学的研究」『民族學研究』24(4): 307-342.
- 深澤美香(2017)『加賀家文書におけるアイヌ語の研究』千葉: 千葉大学.
- 深澤美香(2025)「国立アイヌ民族博物館所蔵知里真志保記入のアイヌ語基礎語彙調査表」『国立アイヌ民族博物館研究紀要』3. (印刷中)
- 深澤美香・小野洋平(2024)「方言境界再考: アイヌ語の方言境界を例として」『北方言語研究』14: 155-176.
- 村山七郎(1971)『北千島アイヌ語』吉川弘文館.
- Asai, Toru (1974) Classification of dialects: Cluster analysis of Ainu dialects. *Bulletin of the Institute for the Study of North Eurasian Cultures, Hokkaido University* 8: 45-136.
- Bugaeva, Anna and Tomomi Sato (2021) A Kuril Ainu Glossary by Captain V. M. Golovnin (1811). *International Journal of Eurasian Linguistics*, 3(2): 171-216.
- Fukazawa, Mika (2023) Tracking basic Ainu vocabulary: Updates and changes of geographical distributions. In: Lan, Trinh Cẩm, Hạnh, Trần Thị Hồng, Suzuki, Hiroyuki and Endo, Mitsuaki (eds.) *Proceedings of the fifth International Conference on Asian Geolinguistics*, 34-48. Tokyo: Geolinguistic Society of Japan.
- Fukazawa, Mika (2018) Geographical distribution patterns of basic Ainu vocabulary in Hattori and Chiri (1960). In: Suzuki, Hiroyuki and Mitsuaki Endo (eds.) *Studies in Asian Geolinguistics, Monograph Series No. 4: Papers from the Fourth International Conference on Asian Geolinguistics*, 91-103.
- Kirikae, Hideo (1988) Ainu Minzoku Cyoosani cuite 『ウエネウサラ』3: 12-40. (in Japanese)
- Nakagawa, Hiroshi and Mika Fukazawa (2022) Hokkaido Ainu dialects: Towards a classification of Ainu dialects. In: Bugaeva, Anna (ed.) *Handbook of the Ainu Language*, 253-328. Berlin: Mouton De Gruyter.
- Ono Yohei and Mika Fukazawa (2025) Statistical approaches to the quantification of regularities in languages and dialects: An exercise in Ainu dialects of Asai (1974). *Northern language studies*, 15. (in print).
- Swadesh, Morris (1952) Lexico-statistic dating of prehistoric ethnic contacts: With special reference to North American Indians and Eskimos. *Proceedings of the American Philosophical Society*, 96: 452-463.
- Swadesh, Morris (1955) Towards greater accuracy in Lexicostatistic dating. *International Journal of American Linguistics*, 21(2), 121-137.

Challenges of Dialect Classification in Extracting Maximum Information from the Basic Vocabulary of Ainu in Asai (1974)

Mika FUKAZAWA

(National Ainu Museum)

Yohei ONO

(St. Luke's International University)

Keywords: Ainu language, philology, feature engineering, document and bibliographic resource studies,
beyond lexicostatistics

This study attempts to establish a foundation for dialectal classification that does not depend on “basic vocabulary” and Swadesh’s lists. Dialectal information from the basic Ainu vocabulary of Hattori and Chiri (1960) encompasses not only “lexical forms,” but also phonological, morphological, and syntactic “regularities.” We extracted and quantified both sets of information, using the data of Ainu dialects in Asai (1974).

In the context of dialect and historical research of Ainu, the data from Hattori and Chiri (1960), including 19 dialects and 200 items of basic vocabulary, occupy a pivotal position in both the present and the past. The data of Asai (1974) used in this paper are also based on those of Hattori and Chiri (1960). However, it should be noted that both Hattori and Chiri’s (1960) and Asai’s (1974) studies utilized only a portion of the 200 (or Asai’s 202) basic vocabulary items in their analyses. While Asai (1974) did not specify which 110 items out of the 202 he applied for his analysis, Ono and Fukazawa (2023) identified them. Consequently, this study concentrated on the 92 items not applied in the analysis and attempted to quantify the data.

The extraction of regularities may suggest the feasibility of obtaining dialect information from independent sources of the basic vocabulary. Specifically, this approach may facilitate the creation of data on grammatical features beyond the phonological and lexical levels, extending to the morphological and syntactic levels. This study focuses on and delineates the methodology of data preparation, including the 92 items excluded from the basic vocabulary data in Asai’s (1974) analysis. The statistical results of this study, which demonstrated that the dialectal classification of Ainu was more strongly elicited, are detailed in Ono and Fukazawa (2025).

(ふかざわ・みか mk.fukazawa@hotmail.co.jp)

(おの・ようへい mathematical.humanities@hotmail.com)