

ISSN 2436-6471

地理言語学研究
Studies in Geolinguistics

2



日本地理言語学会
Geolinguistic Society of Japan
2022

ISSN 2436-6471

地 理 言 語 学 研 究

第2号

Studies in Geolinguistics

No. 2

2022

日本地理言語学会編集委員会 | Editorial board of Geolinguistic Society of Japan

コーディネーター | Coordinator
中井精一 (Seiichi NAKAI)

委員 | Editors

斎藤純男 (Yoshio SAITÔ)
遠藤光暁 (Mitsuaki ENDO)
福嶋秩子 (Chitsuko FUKUSHIMA)
鈴木博之 (Hiroyuki SUZUKI)

2022年9月30日発行

doi: <https://doi.org/10.5281/zenodo.7122479>

© 2022 Individual contributors

編集・発行 日本地理言語学会
学会ウェブ <https://geolinguistics.sakura.ne.jp/>
事務局 東京都渋谷区渋谷4丁目4-25青山学院大学G1305研究室内

目 次

論文/Research articles

鈴木留美子

日本人の名字の地域性と遺伝学..... 1

SUZUKI, Hiroyuki, Sonam Wangmo, and Tsering Samdrup

Connecting Southern Khams in geolinguistics:

A brief survey on ‘fish’ and ‘pig’ beyond provinces 29

鈴木博之

東部チベット系諸言語における「野うさぎ」の語形：

チベット文語形式との音対応を中心に..... 40

ITO, Reiko, Takamasa SEIMIYA, and Yuji KAWAGUCHI

The word for ‘mirror’ in French: Analysis of *miroir* and *glace*

in literary works, 1650–1799 53

上田博人

スペイン語高頻度語 «así», «mismo», «ahora」の地理・社会的変異と

歴史的変化の要因..... 66

TOKIZAKI, Hisao

Stress and word order in Asian languages with reference

to the *Linguistic Atlas of Asia* 102

SHINAGAWA, Daisuke and Junko KOMORI

‘Pig’ in Bantu 122

福嶋秩子

日本語における「兄弟姉妹」を表す総称の分布..... 126

八木堅二

中国語方言における「兄」「弟」「姉」「妹」を表す語の分類と分布..... 133

HIRANO, Ayaka

Terms for ‘elder brother’ and ‘elder sister’ in Kra-Dai 153

富田愛佳

Kra-Dai における「弟・妹」語彙..... 165

彙報：2021年度..... 176

Contents

Research articles

SUZUKI, Rumiko

Locality of Japanese surnames related with genetics 1

SUZUKI, Hiroyuki, Sonam Wangmo, and Tsering Samdrup

Connecting Southern Khams in geolinguistics:

A brief survey on ‘fish’ and ‘pig’ beyond provinces 29

SUZUKI, Hiroyuki

Lexical forms of ‘hare’ in eastern Tibetic languages:

Focusing on a sound correspondence with Literary Tibetan forms 40

ITO, Reiko, Takamasa SEIMIYA, and Yuji KAWAGUCHI

The word for ‘mirror’ in French: Analysis of *miroir* and *glace*

in literary works, 1650–1799 53

UEDA, Hiroto

Factors of the geographical and social variation and historical change

of Spanish frequent words: «*así*», «*mismo*», «*ahora*» 66

TOKIZAKI, Hisao

Stress and word order in Asian languages with reference

to the *Linguistic Atlas of Asia* 102

SHINAGAWA, Daisuke and Junko KOMORI

‘Pig’ in Bantu 122

FUKUSHIMA, Chitsuko

Distributions of general terms for ‘sibling’ in Japanese 126

YAGI, Kenji

Classifications and distributions of sibling terms in Chinese dialects 133

HIRANO, Ayaka

Terms for ‘elder brother’ and ‘elder sister’ in Kra-Dai 153

TOMITA, Aika

Terms of ‘younger brother’ and ‘younger sister’ in Kra-Dai 165

Memoranda and personalia 2021 176

論文

日本人の名字の地域性と遺伝学

鈴木留美子
国立遺伝学研究所

Locality of Japanese surnames related with genetics

Suzuki, Rumiko
National Institute of Genetics

Abstract: Japanese surnames have a complex origin. Some historical surnames derived from honorable titles given by the emperor of the time, while some surnames were brought by people from China or Korea. Afterwards, warriors dispersed through the war period and started using land names as their surnames. Therefore, the distribution of a surname reflects its origin and history. Even highly frequent surnames showed uneven geographical distribution, which can be explained by historical consequences. Surnames are also expected to work as a marker of clan members who share the similar genetic background. Since the 12th century, the surname of the family was inherited by sons, while daughters bore her husband's surname after marriage. This paternal inheritance of the surname is similar with the genetic inheritance of Y chromosome. This article illustrates frequency maps of various surnames in Japanese prefectures based on telephone number book data, then compares the distribution of surnames with genetic backgrounds of the residents. Principle component analysis of surnames and genetic data showed similarity of the plot patterns when focused on local areas.*

キーワード：日本人の名字；漢字；地理分布；主成分分析；遺伝学

Keywords: Japanese surname; kanji; geographical distribution; principle component analysis; genetics

鈴木留美子 (2022) 「日本人の名字の地域性と遺伝学」 『地理言語学研究』 2: 1–28. doi: <https://doi.org/10.5281/zenodo.7121484>

* 本研究に際しては、2018～2022年度文部科学省新学術領域研究「ゲノム配列を核としたヤポネシア人の起源と成立の解明」（研究代表者：斎藤成也、課題番号18H05505）の援助を受けている。

1. はじめに

本稿では、全国電話帳データに登録された情報を元に、日本人の名字の地理分布を調査した。歴史的には「姓」は氏族が天皇から賜ったもの（藤原、平など）であり、「名字」は、同じ氏の公家が違いの区別のために使った「称号」が元になったもの（九条、堀川など）や、東国の武士が本領の地名を名乗ったもの（三浦、伊東など）である。「苗字」という表記は江戸時代以降に広まった。全国民が名字を持つことになったのは、明治8年（1875）の苗字強制令以降である。江戸時代、農民は名字を名乗ることが許されていなかったが、名字が全くなかったわけではないので、この令が出ると多くは先祖伝来の名字を届けた。しかし代々の名字がわからない都市部の住民は自己設定することも多かったらしい。

平安時代の妻問婚から鎌倉時代の嫁入婚に移行すると、名字は父から息子へ父系で伝わるものとなった。これと同じように、父系で伝わる生物学的要素にY染色体がある。人間の染色体は22対の常染色体（44本）と、性別を決める2本の性染色体（X染色体又はY染色体）から成り、性染色体がXXの組み合わせなら女性、XYなら男性となる。従ってY染色体は父親だけが持ち、父から息子へと受け継がれる。Y染色体はDNAの塩基配列の違いによってタイプ分けされている（ハプログループA～R）。名字は養子のように必ずしも実の息子が受け継ぐとは限らないが、名字とY染色体ハプログループの間にはある程度の並行性があることが知られている。本稿では名字の地域分布を調べると共に、名字とYハプログループ、さらには沖縄地方のゲノム先行研究と照らし合わせて考察する。

2. 名字データの概要

2.1. 都道府県別登録件数

本研究では、CD版全国電話帳データに記録されている1625万315件の名字データを利用した。県別登録件数が最も多かったのは北海道、最も少なかったのは鳥取県であった（図1）。電話帳データにはふりがながないので、名字は「読み」ではなく「表記」によって区別し、例えば「斉藤」「斎藤」「齋藤」などは別々の名字として扱った。名字の種類は68,112、使用されている漢字の種類は3,311であった。

2.2. 名字の種類数と人数

図2に各都道府県別の名字の種類数を、図3に名字あたりの人数（登録人数/名字数）を示す。名字あたりの人数が最も多かったのは新潟県、最も少なかったのは奈良県であった。

図1 都道府県別登録件数

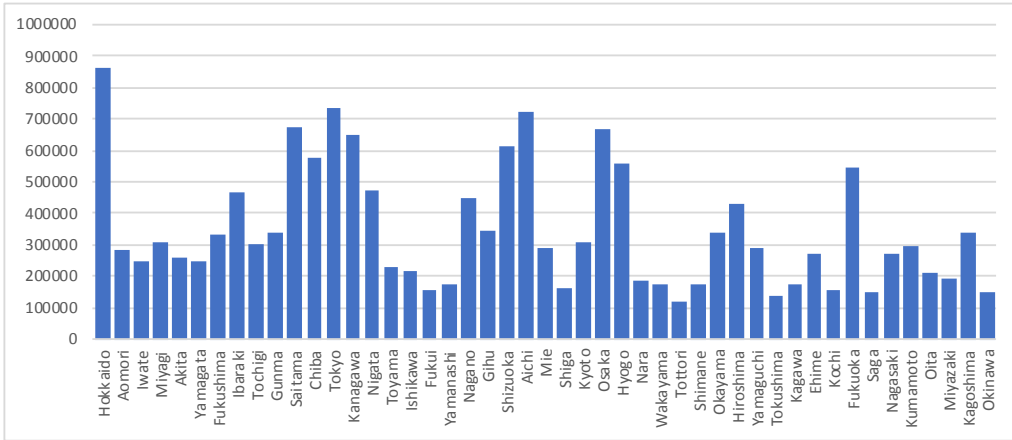


図2 都道府県別名字数

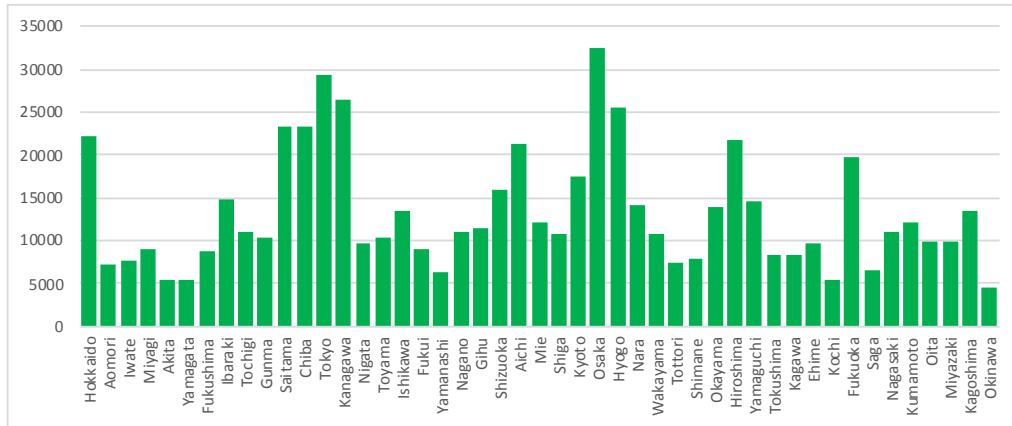
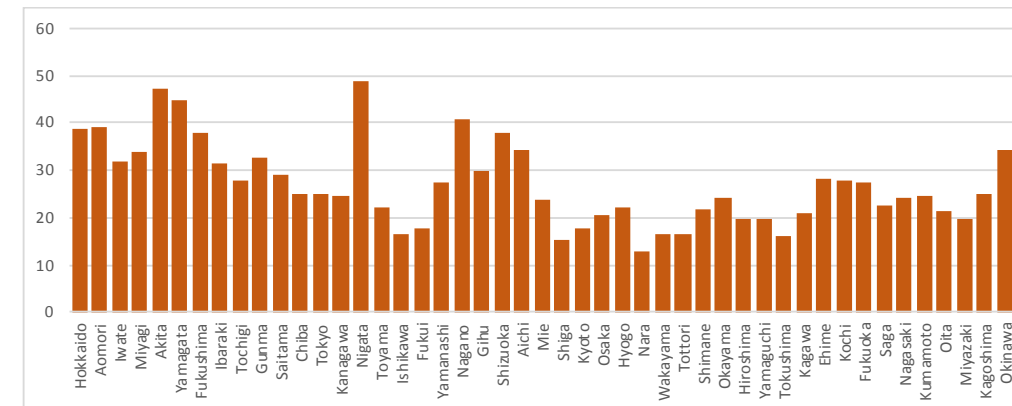


図3 名字あたり人数

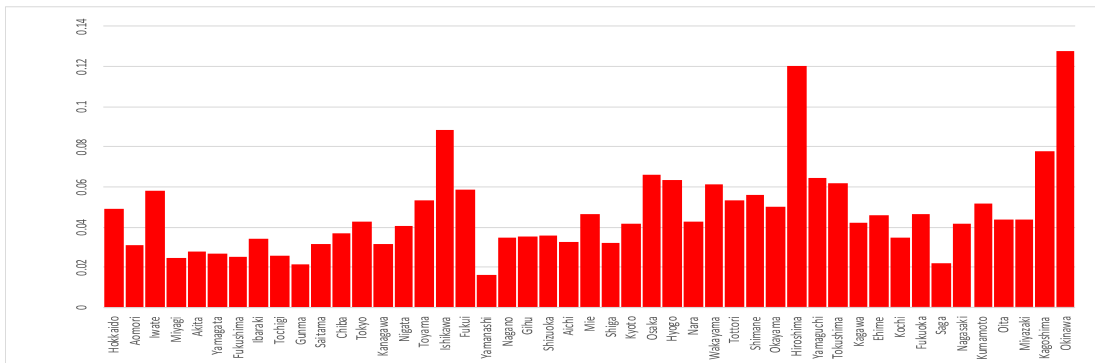


3. 名字の地域特異性

3.1. その都道府県内にしか見られない名字

どの都道府県にも、他の都道府県では見られない特有の名字が存在する。都道府県内の名字数のうち、特有な名字が占める比率を図4に示す。東京、大阪、福岡などは各地から人が集まるので名字の種類数は多いが、特有の名字の比率は高くない。特有な名字の比率が最も高いのは沖縄県で、広島県、石川県、鹿児島県が続く。比率が最も低いのは山梨県で、群馬県、佐賀県、宮城県が続く。

図4 都道府県特有な名字が占める比率

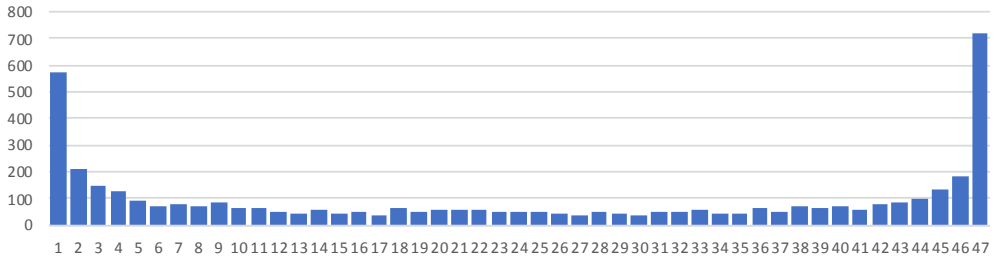


3.2. 名字の共有度別人数

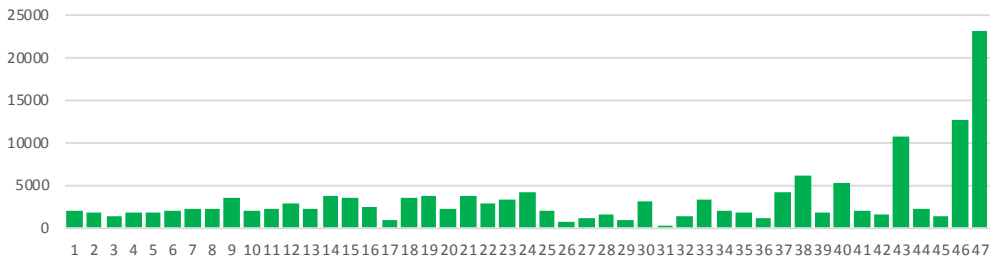
特有の名字の多い沖縄県と広島県であるが、名字の他都道府県との共有度別に人数を調べると差が見られる。図5に、都道府県間共有度（1～47）ごとに広島県と沖縄県の名字数と、人数を示す。左端の共有度1がその都道府特有の名字、右端の47が全都道府県で見られる名字である。広島県は特有の名字の数は多いがその名字を持つ人数は少なく、大多数の人が47都道府県共通の名字を持っていることがわかる。それに対し、沖縄県は都道府県共有度の低い名字を持つ人数が多い。沖縄特有の名字は人数が多い順に喜久村、慶留間、宜志富、多宇、饒辺など、広島県は沖辺、家頭、栗末、巳之口、大高下などであった。

図5 都道府県間共有度と名字数、人数

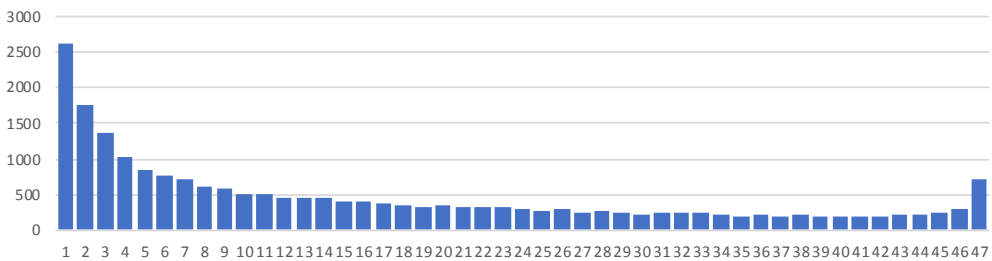
(1) 沖縄県・名字数



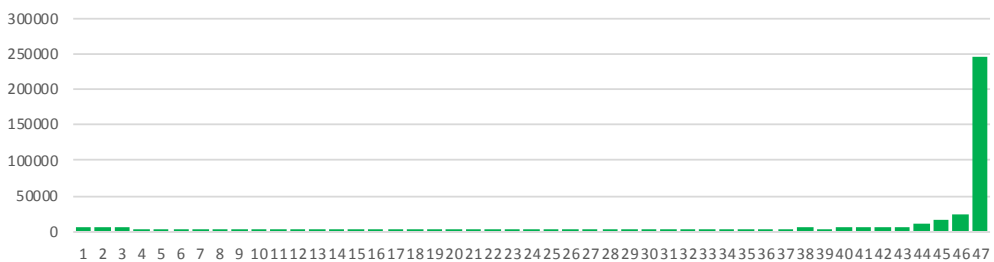
(2) 沖縄県・人数



(3) 広島県・名字数



(4) 広島県・人数



4. 名字の地理分布

名字の都道府県別頻度マップはさまざまなソースから発表されている。大規模なものに矢野桂司によるマップがあるが、ここでは電話帳データを元に、頻度の高い名字の地域差と、歴史的に古い名字の分布を見ていく。

4.1. 頻度の高い名字の分布

今回利用した電話帳データでは、頻度が高い順に、1. 佐藤、2. 鈴木、3. 高橋、4. 田中、5. 伊藤、6. 山本、7. 渡辺、8. 中村、9. 小林、10. 加藤、だった。先行の各種データと同様の順位である。サイトウは異字表記が多いため、単独では10位以内に入らなかったが、これも頻度の高い名字なので異字別に分布を見た。

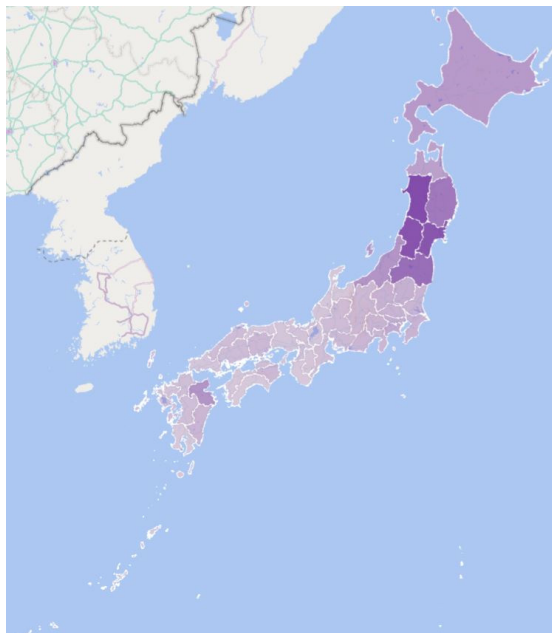
4.1.1. 頻度上位10種の地理分布

頻度上位10種の名字の地理分布を図7に示す。頻度が高い名字でも全国一律に多いわけではなく、地理的な偏りが見られる。名字の分布は、新潟、長野、岐阜、三重以東が東日本型、それより西が西日本型に大きく分けられる。

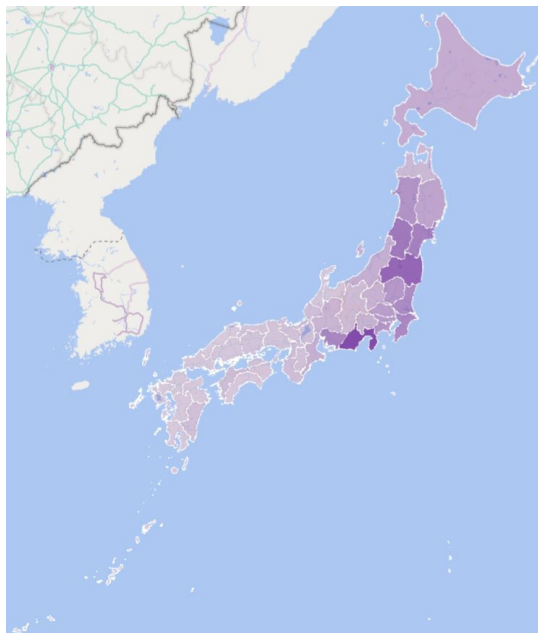
佐藤、高橋は東北に多く、鈴木は東北南部のほか静岡に多い。田中と山本が西日本に多いことも知られているが、田中は青森でもやや多い。伊藤は佐藤と同様東北に多いが、三重県を中心とした中部にも多い。これは伊藤氏が「伊勢の藤原氏」をルーツとするからであろう。渡辺は摂津国西成郡渡辺の地名をとった渡辺綱が起源とされるが、現在の大阪では頻度は高くなく、山梨、福島、新潟で多い。中村は長崎、山口、三重、青森などで頻度が高く、東日本にも西日本にも見られる。中心の村という意味合いで、特定の地名や人物と特に強く結びついていないためであろう。小林は地名に由来する名字とされ、長野では伊那郡小林村が発祥とされる。山梨でも多いのは、上野国緑野郡小林の小林氏が戦国時代に上杉氏に仕えていたという歴史的な経緯によるものと思われる。加藤は藤のつく名字として東北にも見られるが、愛知、岐阜に多い。これは美濃加藤氏の流れをひくものとされる。佐藤、伊藤、加藤で分布が異なるのは藤原氏がそれぞれ各地で分派したたからであろう。

図7 頻度上位10種の地理分布

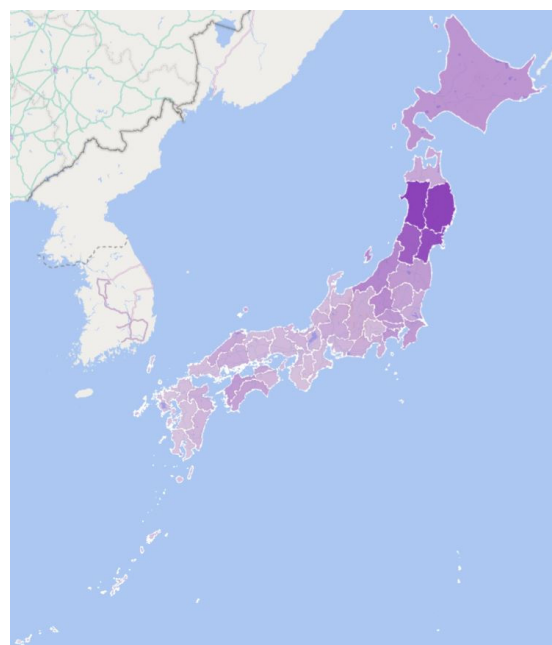
1. 佐藤



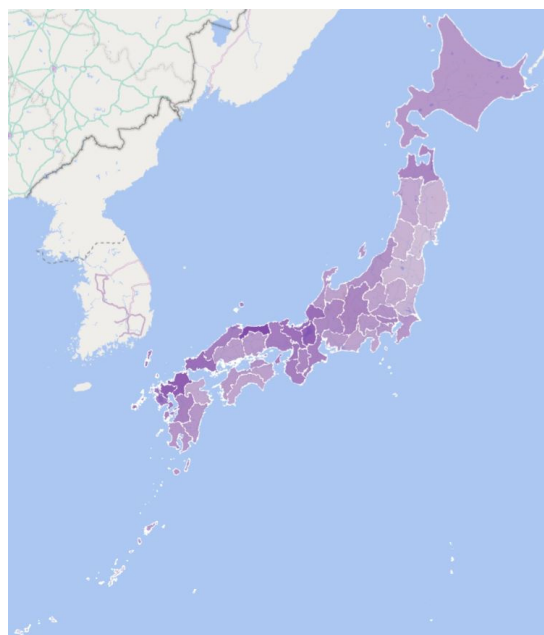
2. 鈴木



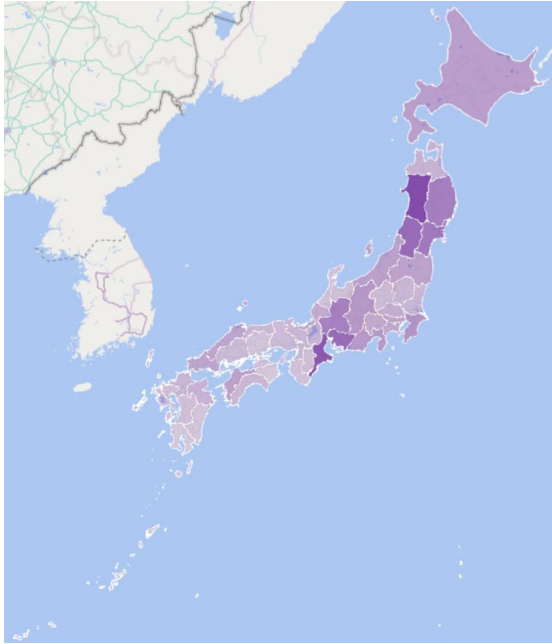
3. 高橋



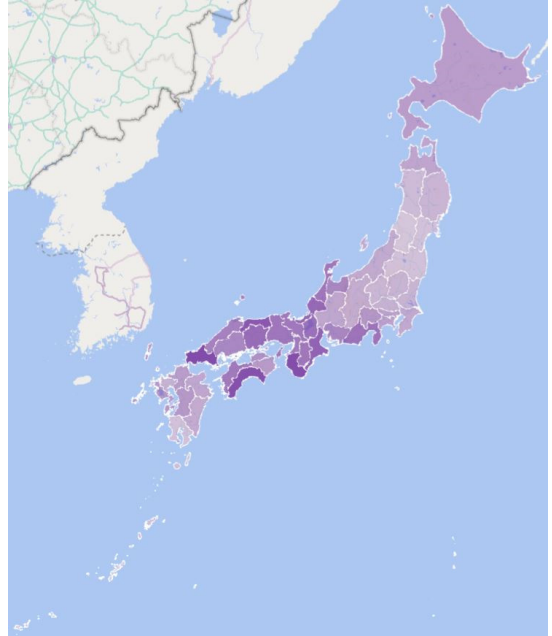
4. 田中



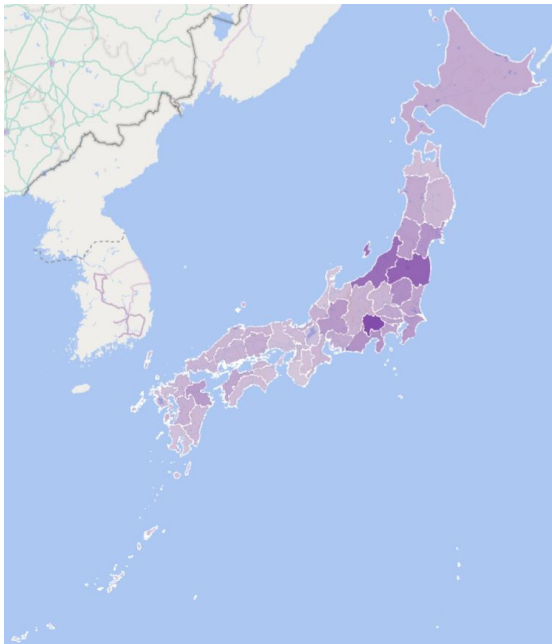
5. 伊藤



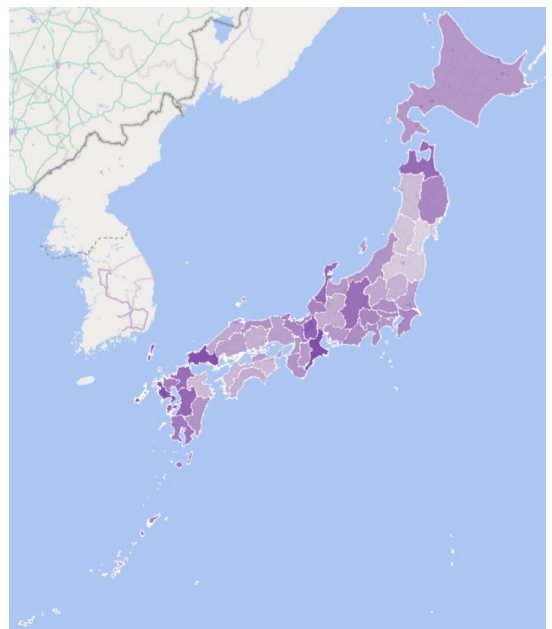
6. 山本



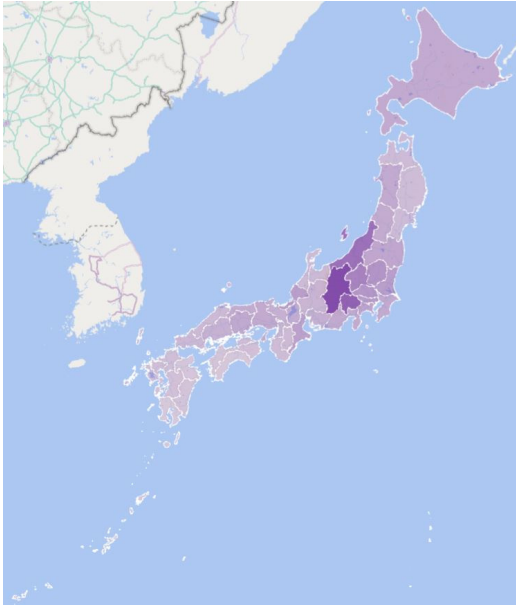
7. 渡辺



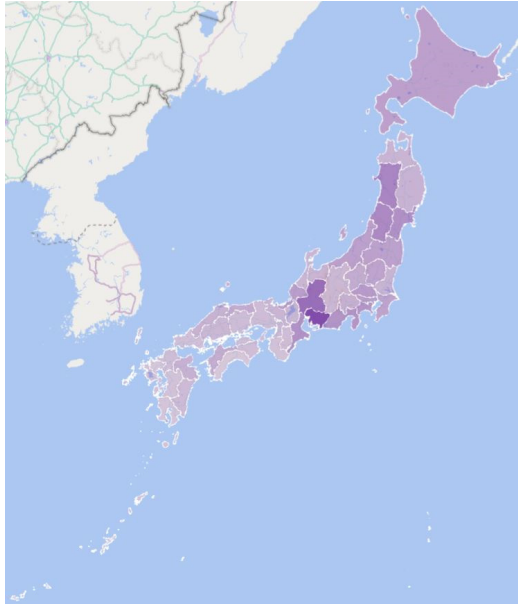
8. 中村



9. 小林



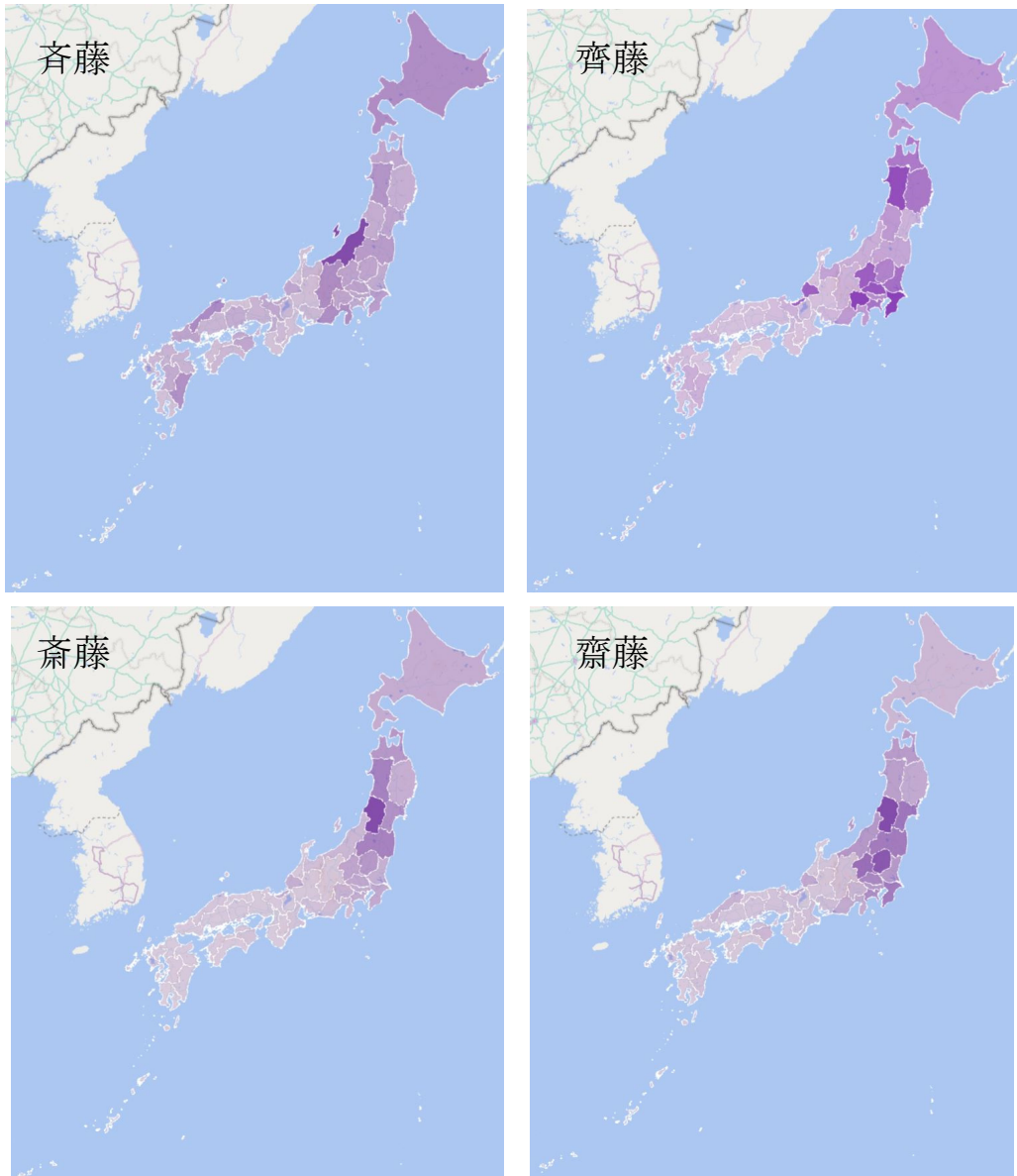
10. 加藤



4.1.2. 斉、齊、斎、齋

サイトウも頻度の高い名字であるが、サイは異字体が多く手書きの時代は85種あったとも言われる。図8に斉藤、齊藤、斎藤、齋藤の分布を示す。齊は斉の、齋は斎の旧字体である。どのサイトウも東日本に多いが、斉藤は新潟県で頻度が高く、旧字体の齊藤とかぶる所は少ない。斎藤と齋藤はどちらも山形で頻度が高いが、齋藤の方がやや南よりに分布する。

図8 斉藤、齊藤、斎藤、齋藤



4.2. 歴史的な名字の分布

現在は名字が新設されることは稀だが、昔は氏族の人数が増えれば互いの識別のために称号を使ったり、武士が地名や主君から賜ったり、明治の「苗字強制令」で新設されたりして増えていった。ここでは歴史を遡り、古い起源を持つ名字の分布を見ていく。

4.2.1. 源、平、藤原、橘

源、平、藤原、橘は四姓とよばれる由来の古い姓である。四つの分類を設ける制度はインドの四種制（カースト）に起源を持ち、中国、匈奴、朝鮮半島などで「四姓ノ制」として広まり、日本にも伝わった。日本の四姓は身分の序列ではなく氏族の出自を表し、近世以前は「姓」と、分家した各家の名称「名字」は別物として併存していた。従って姓をそのまま名字として使うことは少なかったかもしれないが、現代の名字における四姓の分布を図9に示す。

源は他の名字マップでも見られるように石川、富山に多いほか徳島、愛媛でもやや頻度が高い。石川、富山は橘の頻度も高い。平は頻度の高い県が山形、鹿児島、沖縄と南北に別れる。沖縄には平良も多いので、平氏とは異なる起源を持つのかかもしれない。藤原は奥州藤原氏の流れをくむ秋田、岩手に多いほか、兵庫、岡山、島根にも多い。橘は愛媛、兵庫、和歌山、石川、富山、新潟、福島、秋田など、広い範囲で頻度の高い県がみられる。愛媛、和歌山はミカンの産地として知られているが、同じくミカン産地の静岡では橘の頻度は低い。

4.2.2. 渡来系の名字

渡来系の名字としては秦、狛、太秦、丹波などがあるが、ここでは宮廷で雅楽を演奏した渡来系の楽家として知られる、狛系の東、窪、奥、辻、芝、上、西、久保と太秦系の東儀、林、藪、岡の分布を図10に示す。

狛系では、東は南九州、紀伊半島、石川で頻度が高く、窪は鹿児島、奈良、石川、富山に多い。窪と読みが同じ久保はやはり鹿児島で頻度が高く、四国、山陽、近畿などにも頻度の高い県が見られる。奥と西は鹿児島と和歌山で頻度が高く、辻は滋賀を中心に近畿地方で頻度が高い。芝は愛媛、高知、和歌山などに多く、上は新潟で頻度が高いほか、和歌山、奈良にも分布する。

太秦系では、東儀は埼玉、千葉、島根に多いが、埼玉、千葉の分布は近年の移住によるものかもしれない。林は岐阜を中心に北陸・東海から西に分布する。藪は和歌山県で特に多い。岡は長崎、島根、広島、香川、徳島など西南日本で広く見られるほか、山梨、栃木など東日本にも見られる。

県別に見るとこれらの名字は、鹿児島、和歌山、石川で頻度が高い傾向がある。明確な理由はわからないが、島津、紀州、加賀など江戸時代に有力な藩であった地域であることは興味深い。

図9 源、平、藤原、橘

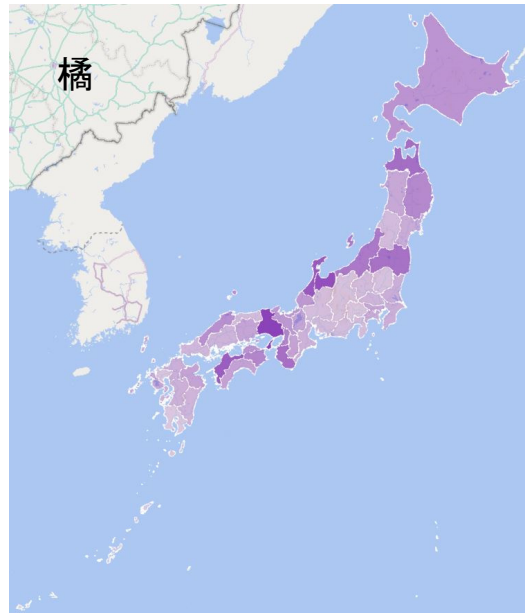
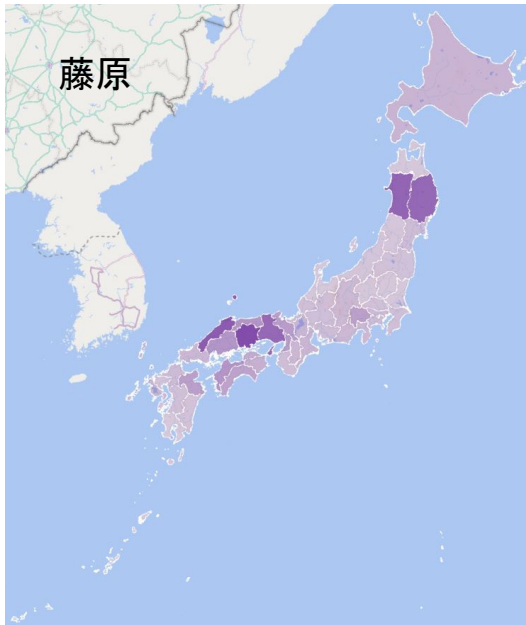
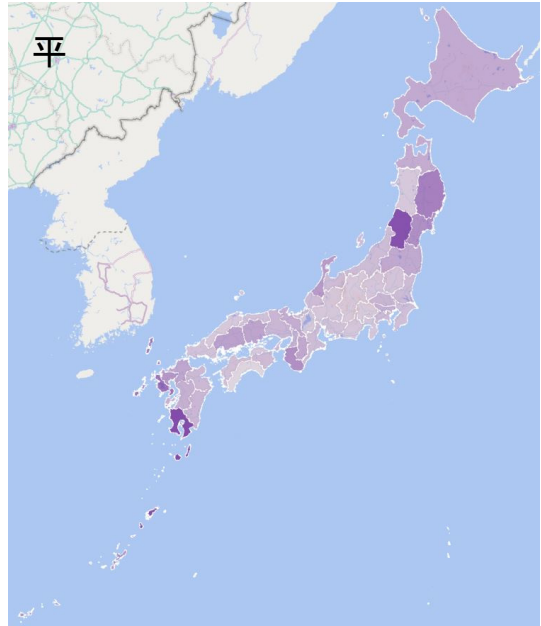
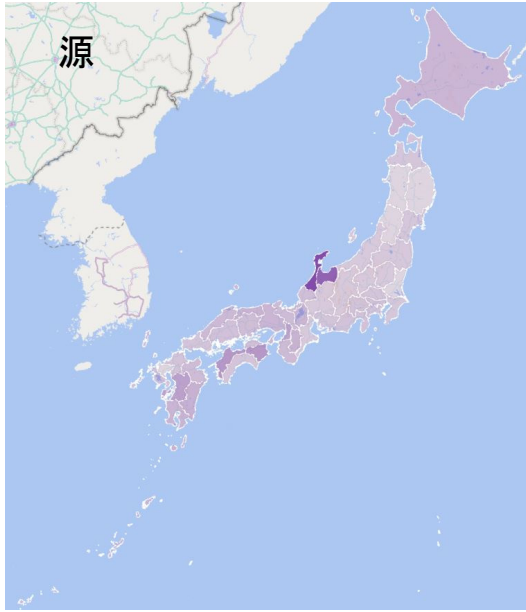
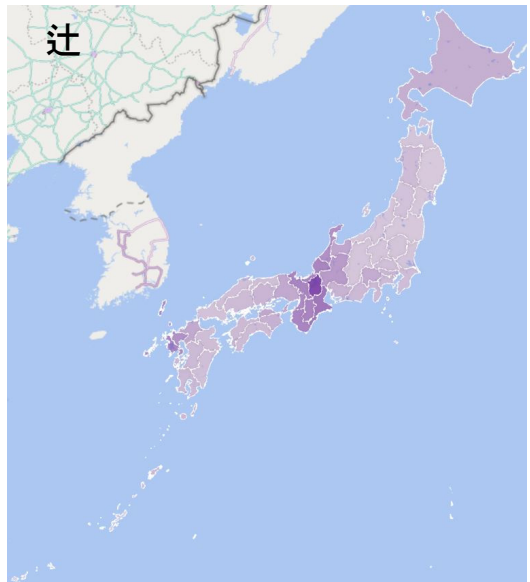
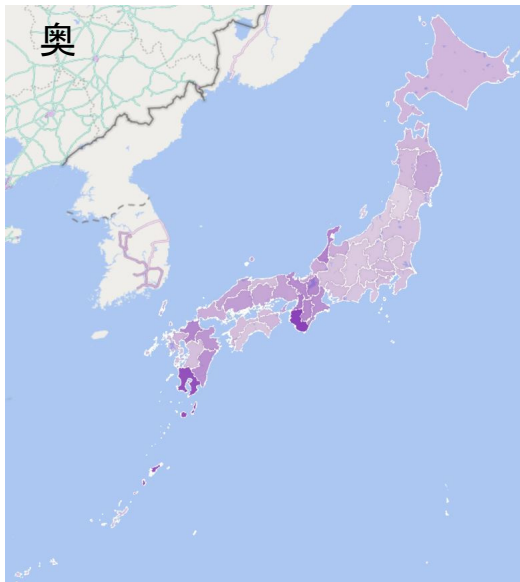
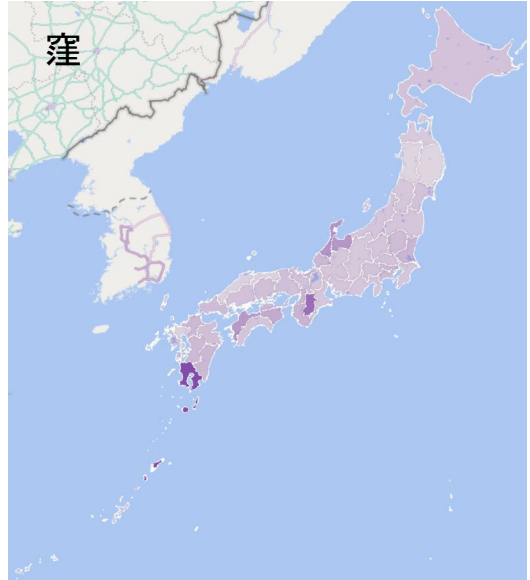
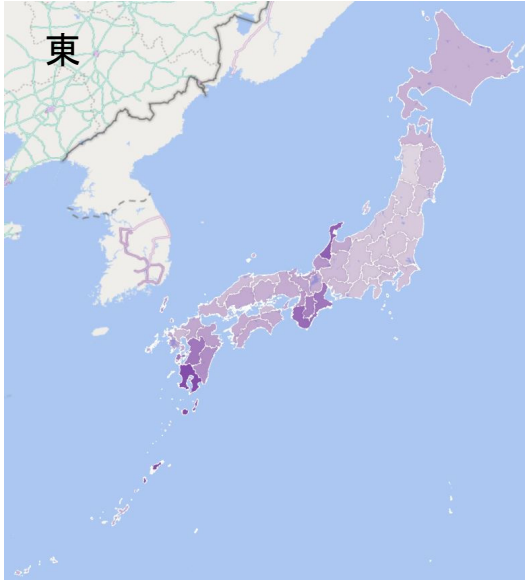
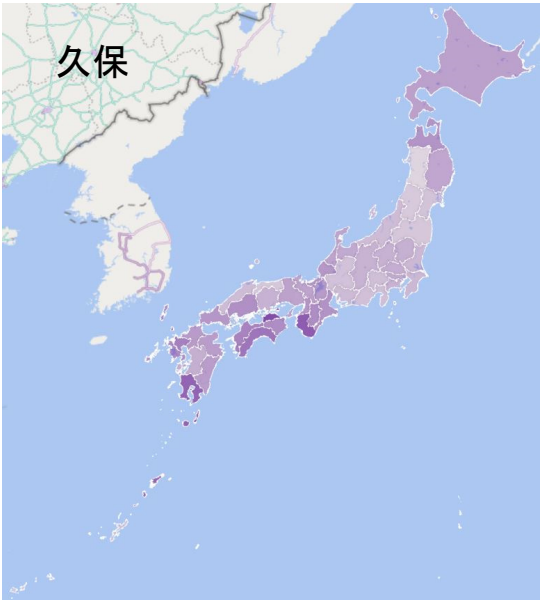
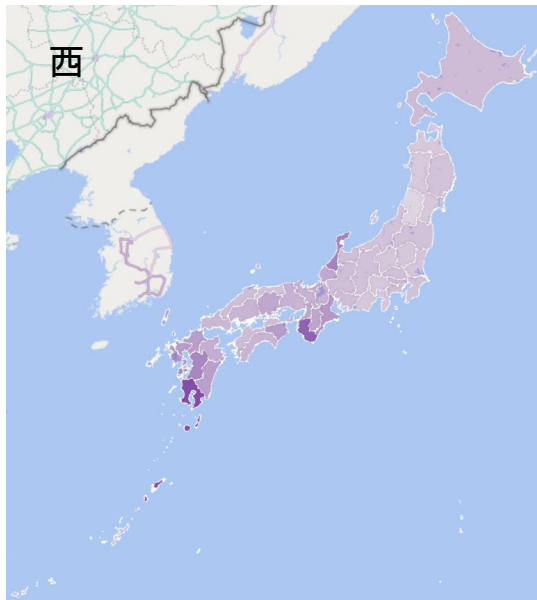
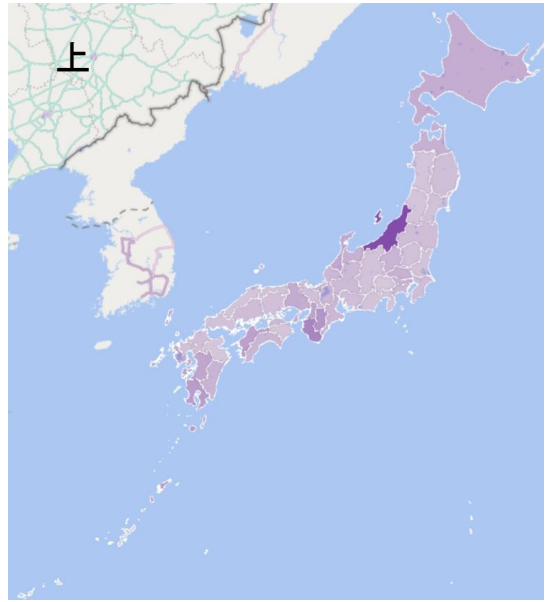
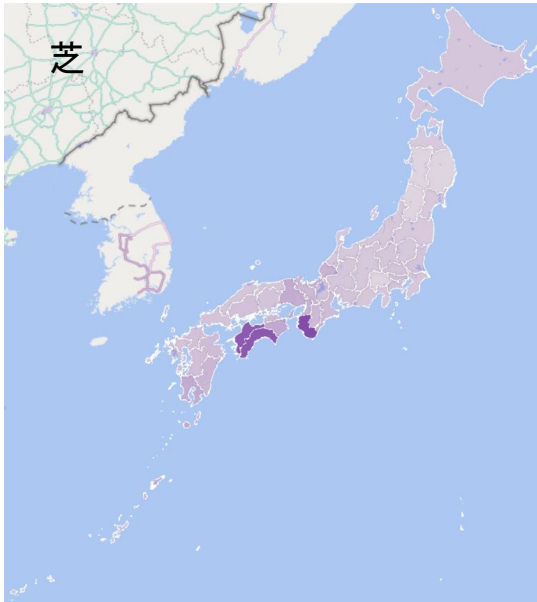
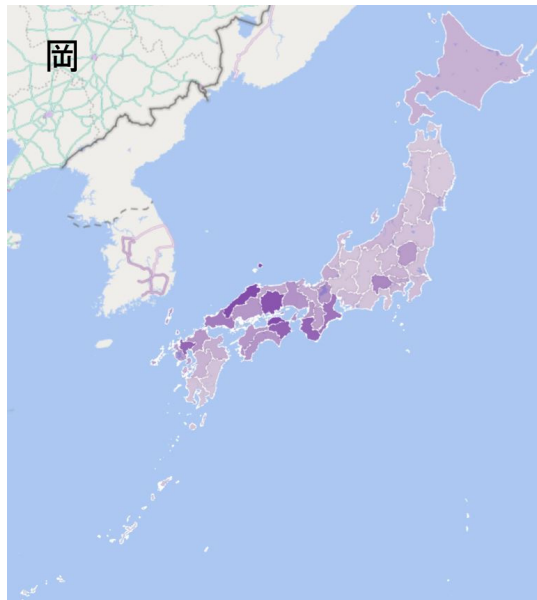
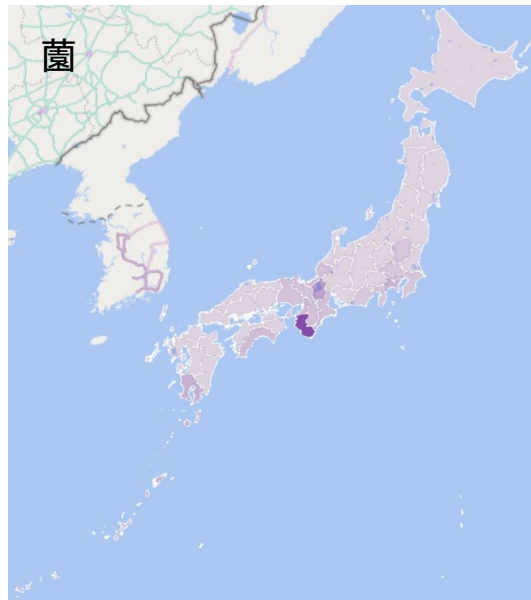
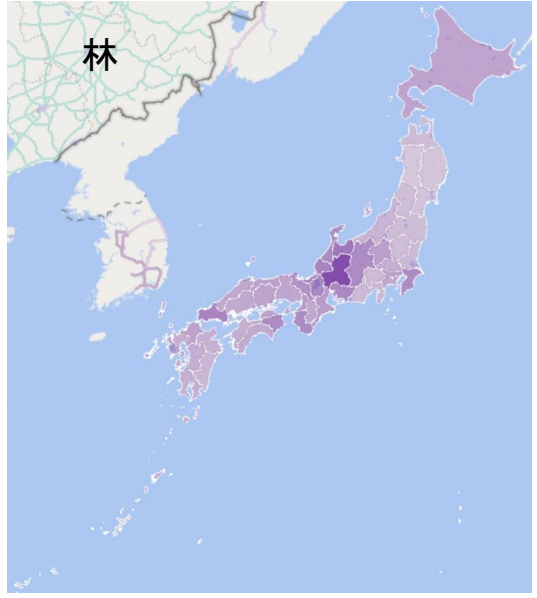
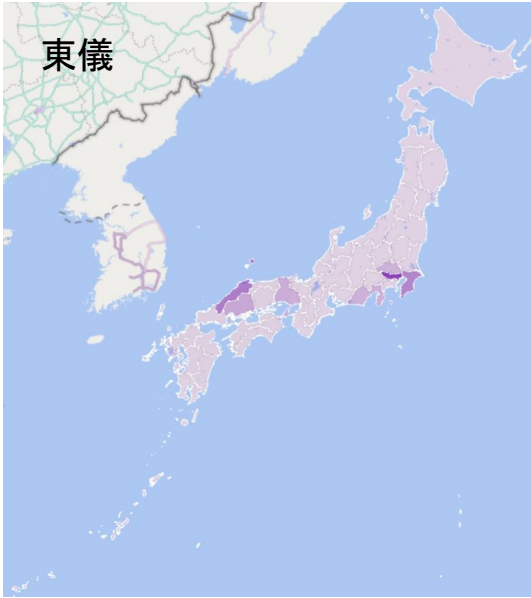


図 10 渡来系楽家
拍系





太秦系



5. 漢字単位での頻度と地理分布

名字の分布マップは書籍やウェブサイトにも多数発表されているので、ここでは漢字単位に分解して分布調査を試みた。

5.1. 頻度の高い漢字

名字に使われている頻度の高い漢字、上位から 30 位までを表 1 に示す。

表 1 使用頻度の高い漢字 1 位～30 位

1.田	2.藤	3.山	4.野	5.川	6.本	7.村	8.木	9.井	10.中
11.小	12.原	13.大	14.佐	15.松	16.谷	17.高	18.上	19.岡	20.島
21.橋	22.口	23.石	24.崎	25.西	26.吉	27.林	28.森	29.沢	30.内

10 位までのうち、野、川、井以外は名字頻度の 10 位に含まれる漢字である。鈴木は名字としては 2 位であるが、「鈴」単独だと 33 位であった。田や藤のように他の漢字との組み合わせが多いものに比べて、鈴は鈴木以外の名字で使われることが少ないからであろう。野、川、井の分布を図 11 に示す。野は満遍なく分布し鳥取、奈良、長野、秋田、青森などでやや少ない。川は近畿、四国、三重、滋賀、福井などから新潟にかけての日本海側に多い。井は瀬戸内、近畿、中部に多く、東北と九州は少ない傾向である。

5.2. 動物を表す漢字

動物を表す漢字で頻度順位の高かった、馬、熊、亀、鳥、鶴、猪、鹿、牛、犬、猿の分布を図 12 に示す。漢字の後のカッコは順位を表す。馬の頻度が高い秋田は相馬のほか神馬がある。熊は岩手、宮城など東北から東日本で多く、熊襲の故郷である熊本では少ない。亀の頻度の高い栃木、香川、愛媛は亀井が共通しているが、栃木では亀田、香川では亀山、愛媛では亀岡が見られる。京都に亀山市が、三重に亀岡市が存在するが、両府県とも亀の頻度は高くない。鳥が中部に多いのは鳥羽の影響だが、秋田の鳥は鳥潟による。鶴は九州で頻度が高く、特に鹿児島では鶴田が多い。猪の多い福島は猪狩と猪俣が見られる。鹿は青森で男鹿、雄鹿が、宮城で鹿野、鹿又が多い。牛の頻度が高い愛知には牛田がある。犬は秋田で犬貝、福岡で犬丸、熊本で犬童が見られる。猿は長野では猿田が、福岡と熊本は猿渡が多い。

動物を表す漢字としては十二支に使われる子、丑、寅などもあるが、子、辰、巳、未以外は通常の動物漢字の頻度の方が高かった。子と未は動物と関係ない意味でも、辰巳は方角由来の名字としても使われるので頻度が高いのだろう。通常の動物漢字では十二支のうち十一が名字に使われていたが、羊を使った名字はなかった。羊は平安時代ごろには大陸から移入されていたようだが、日本

の風土では羊毛の需要が少なく、また肉食の習慣もなかったため家畜として広まらなかったからだろう。

図 11 野、川、井の分布

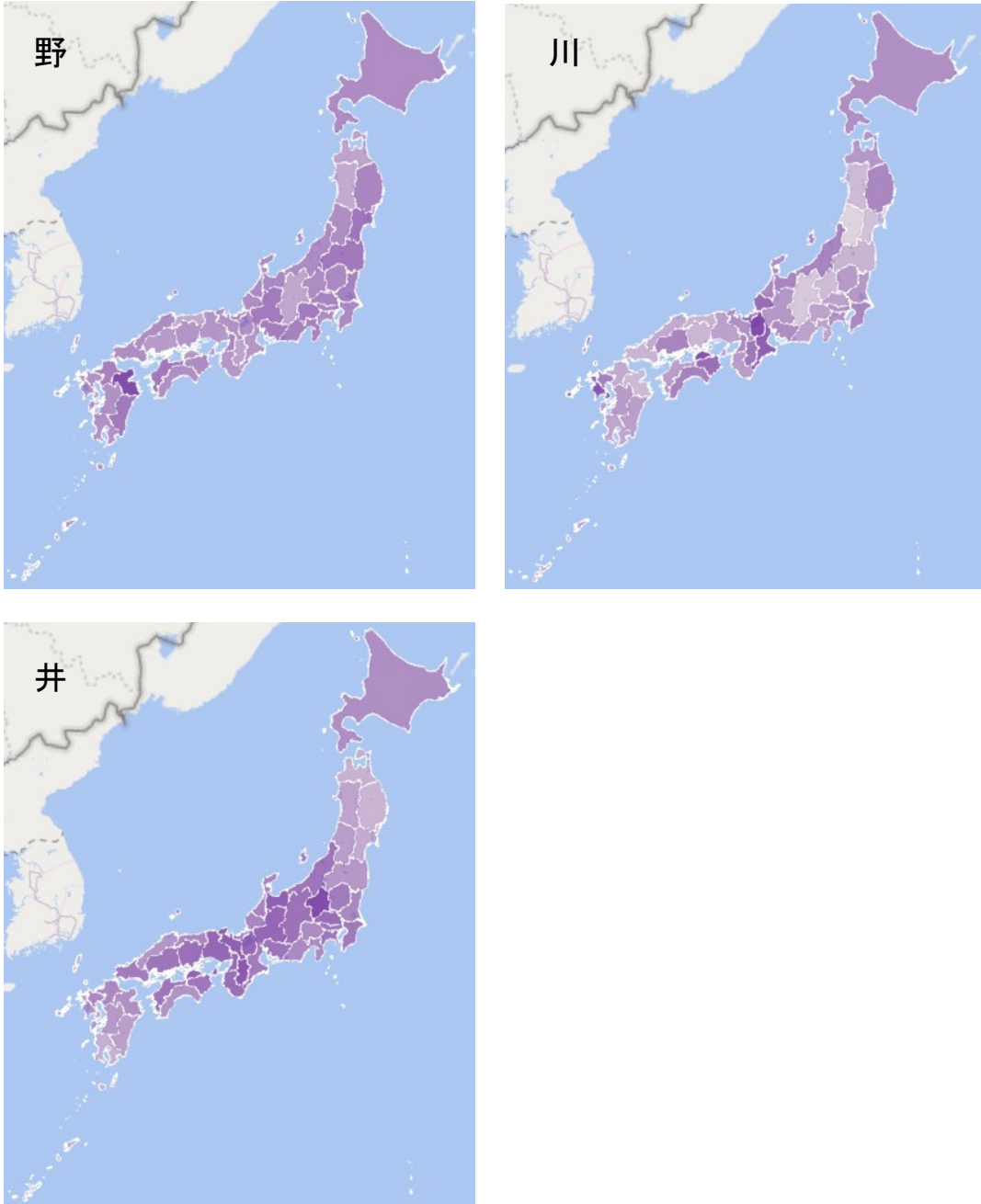
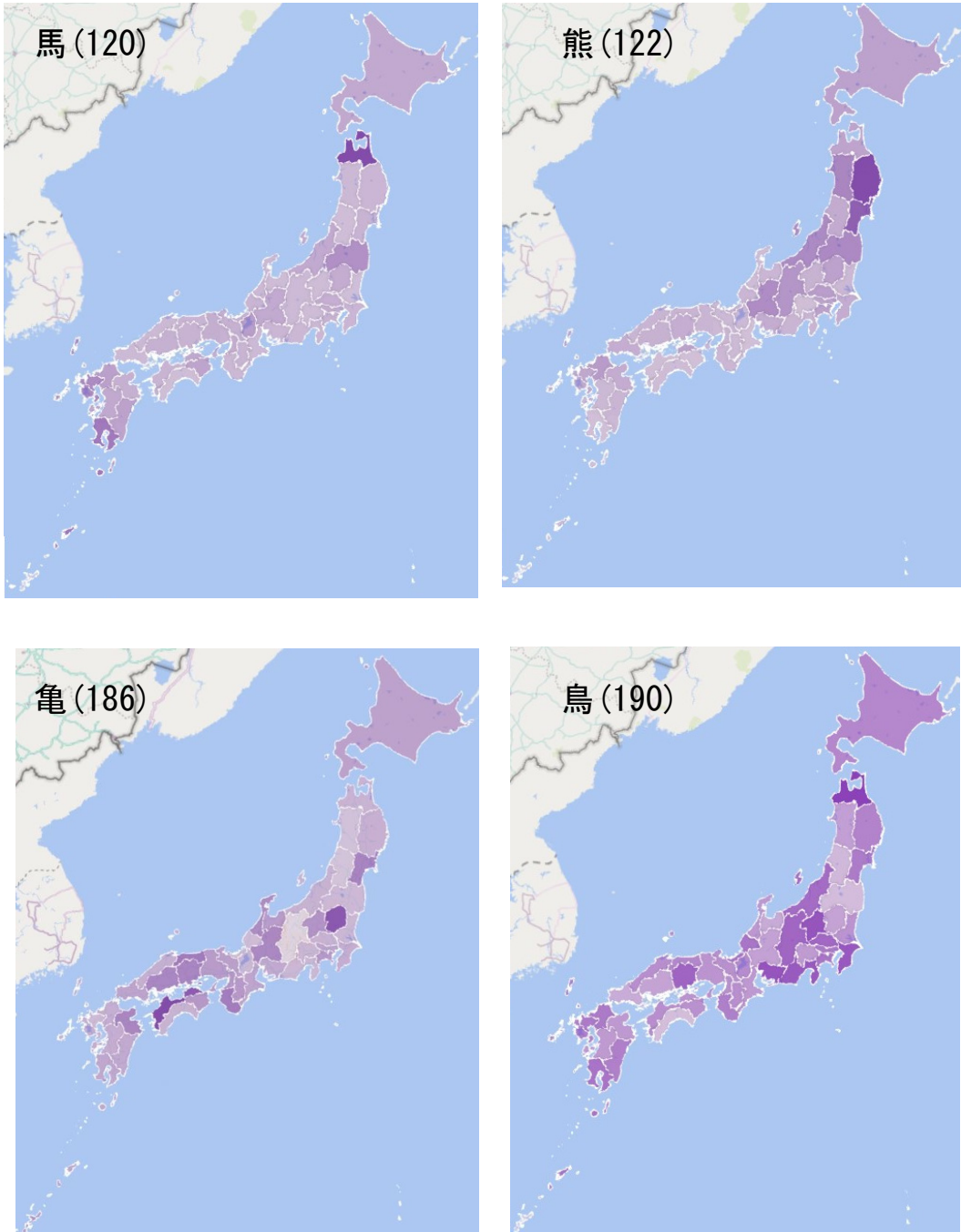
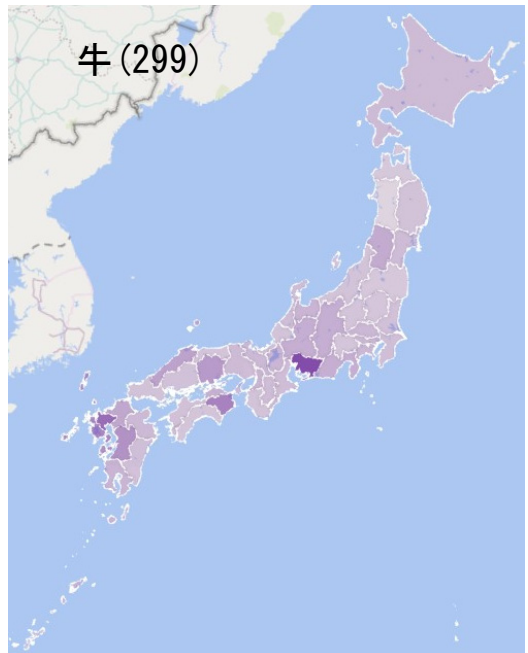
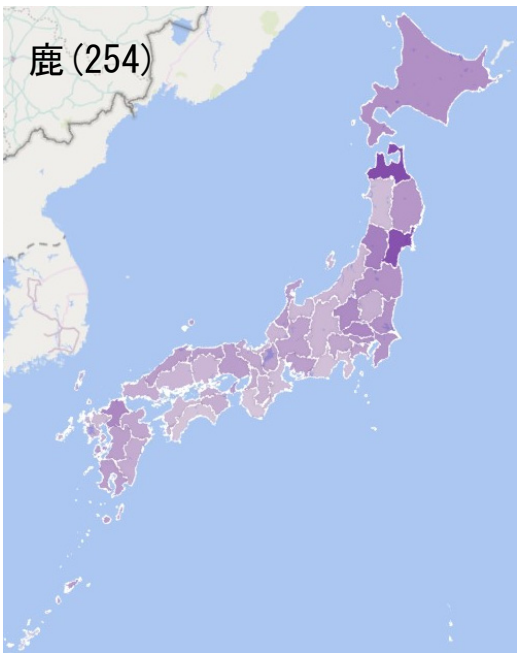
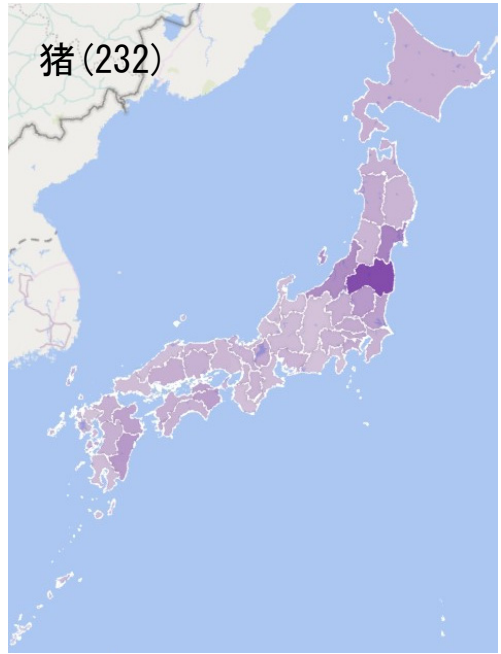
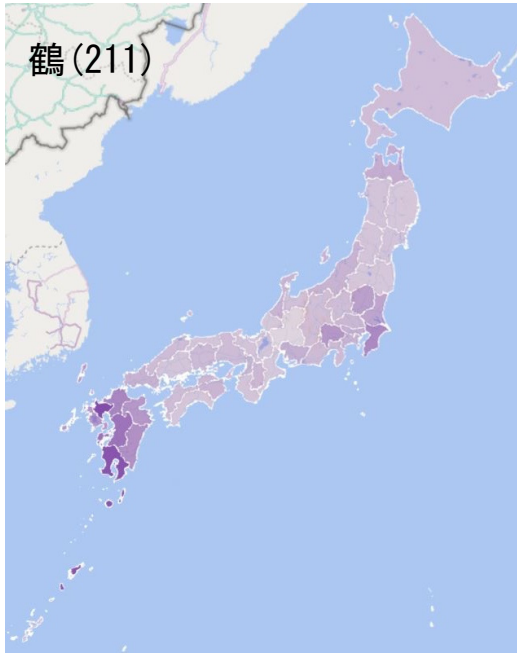
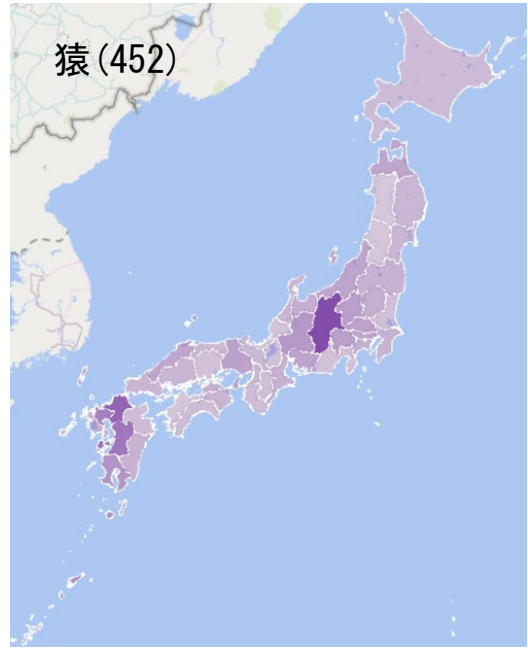
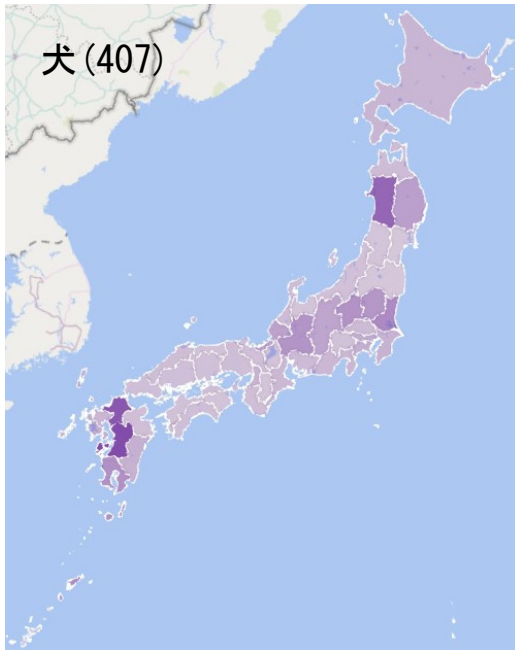


図 12 動物を表す漢字







5. 名字の頻度による主成分分析

全国の半数以上の都道府県 (24 都道府県以上) に存在する名字 6639 種 (1459 万 1476 人) について、県別に頻度をとって行った主成分分析の結果を図 13 に示す。点の色は図の左上に示した日本地図上の色に対応している。横軸の第一主成分 (PC1) の左から右に沿って、都道府県が北から南に並んでおり、沖縄は縦軸の第二主成分 (PC2) で他から大きく離れる。

沖縄県を除いて行なった主成分分析の結果が図 14 である。図 14 も左から東北、関東、中部が並び、右に近畿、中国、四国、九州が固まるが、大分は西日本クラスターから外れて東日本側に位置している。これは源平合戦の時代に源義経について平家追討に加わった豊後の緒方氏が、後に頼朝によって流罪となり、豊後を頼朝の直轄地として関東系の武士が支配したことによるらしい。大分はアクセントも東京型であり、九州の中では東日本に近い。

図 13 名字頻度による主成分分析

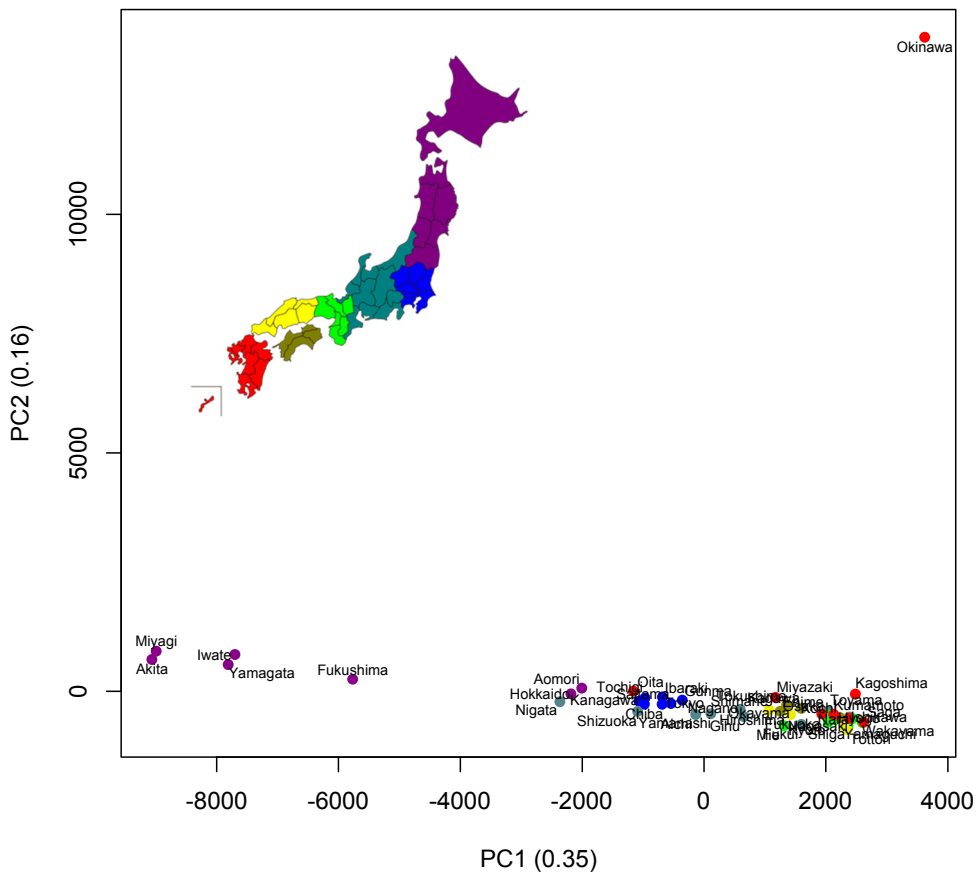
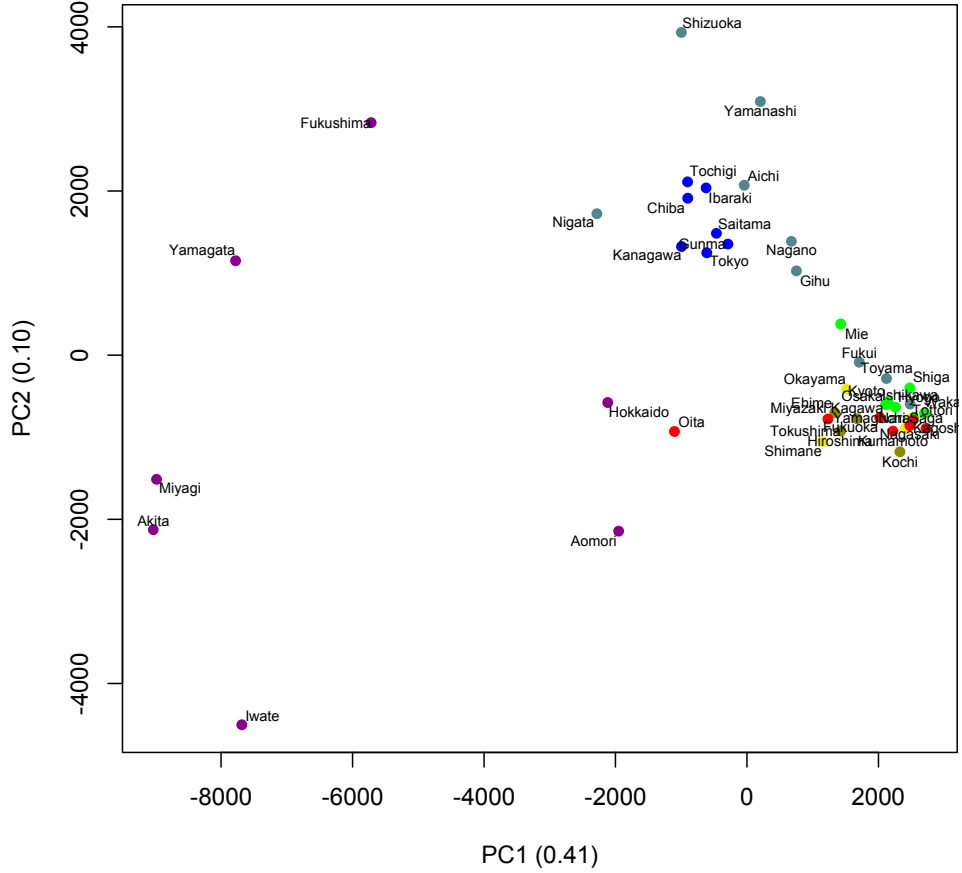


図 14 名字頻度による主成分分析（沖縄県を除く）



6. 名字と遺伝学的データの関連性

ここからは、名字と遺伝学的データの関連性を見ていく。男性が持つ Y 染色体は名字と同様に父から息子へ父系で伝わるが、図 13 の名字頻度による主成分分析のプロットは Timothy Jinam らによる母系のミトコンドリア DNA の解析結果に近い。Y 染色体の 5 大ハプログループ頻度による解析では沖縄が大きく離れることはなかったが（データ未発表）、ハプログループをさらに細かいサブグループに分ければ解像度が上がるかもしれない。都道府県レベルではなく、人の移動の少ない地域の市町村ごとに見ると、名字と遺伝学的データの間に関連性が見られたので以下に述べる。

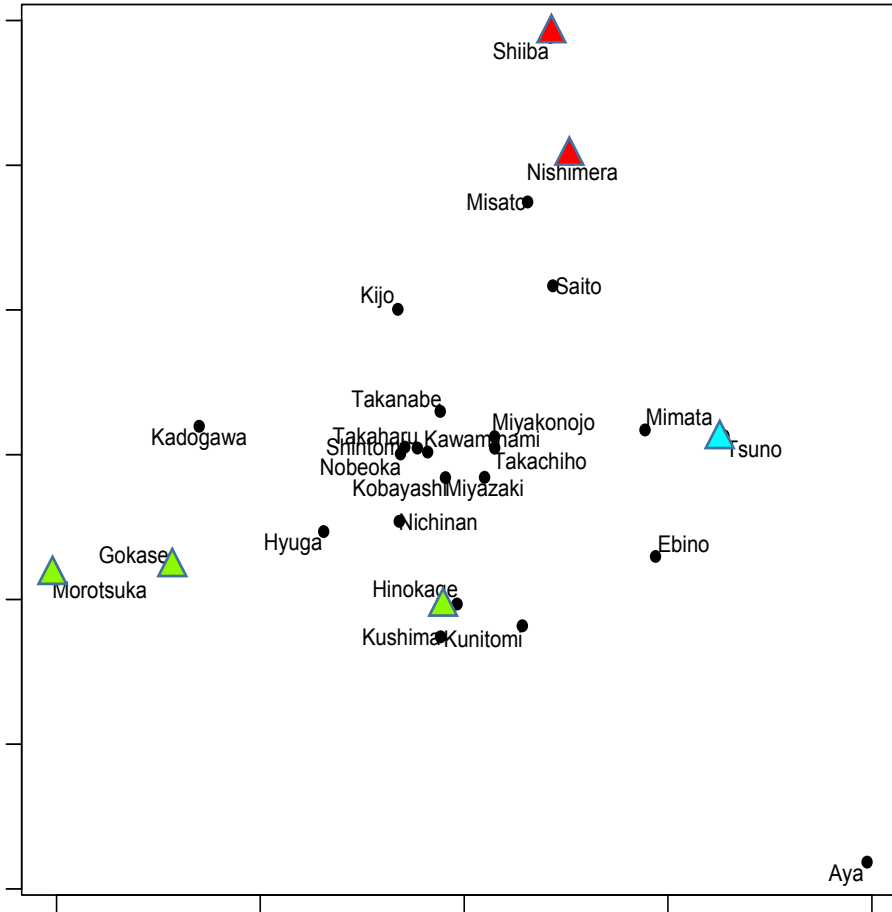
6.1. 宮崎県における名字と Y ハプログループ

野原寛文らによる、宮崎県内の 26 市町村における Y ハプログループ頻度の報告を元にした主成分分析（図 15）と、宮崎県の市町村単位の名字頻度による主成分分析（図 16）を行なった。Y ハプログループによる分析では、D-M125 の多い Shiiba, Nishimera, D-M157 と D-M125 の多い Tsuno, O-M154 の多い Morotsuka, Gokase, Hinokage などが周縁部に位置している。名字による分析ではプロットの位置関係は異なるが、Shiiba, Nishimera, Tsuno, Morotsuka, Gokase, Hinokage がやはり周縁部に位置する。どの名字にどの Y ハプログループが多いかを直接調べることはできないが、上記の地区には特定の名字と Y ハプログループを持つ住民が多いのかもしれない。

6.1. 沖縄における名字とゲノムデータ

沖縄では宮古諸島について、Matsunami らによるヒトゲノムデータを使った詳しい解析が行われている。本稿では、宮古諸島の地区別電話帳データによる名字分析を行い、ヒトゲノムデータによる主成分分析プロットと比較した（図 17）。ヒトゲノムデータの結果では Okinawajima（沖縄本島）は本土に近く、Gusukube, Hirara, Irabu, Ikema の順に遠くなる。同様の傾向は名字データによるプロットでも見られた。これはこの地域の名字が、Y 染色体に限らずゲノム全体とも関連性があることを示唆している。

図 15 宮崎県 26 市町村の Y ハプログループによる主成分分析



右図は、野原らの宮崎県在住者における Y-STR ハプロタイプおよび Y-ハプログループの地理的分布 (2021) による

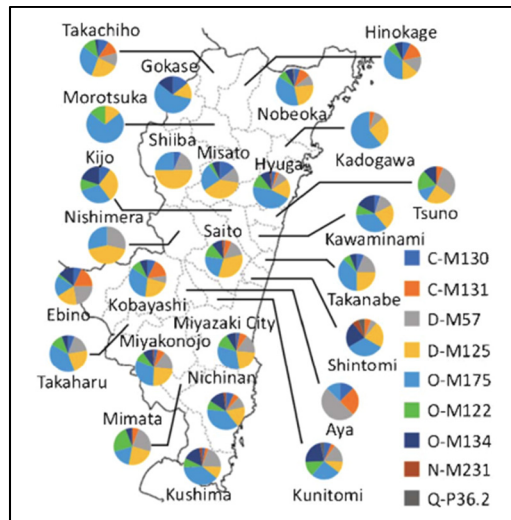


図 16 宮崎県 26 市町村の名字頻度による主成分分析

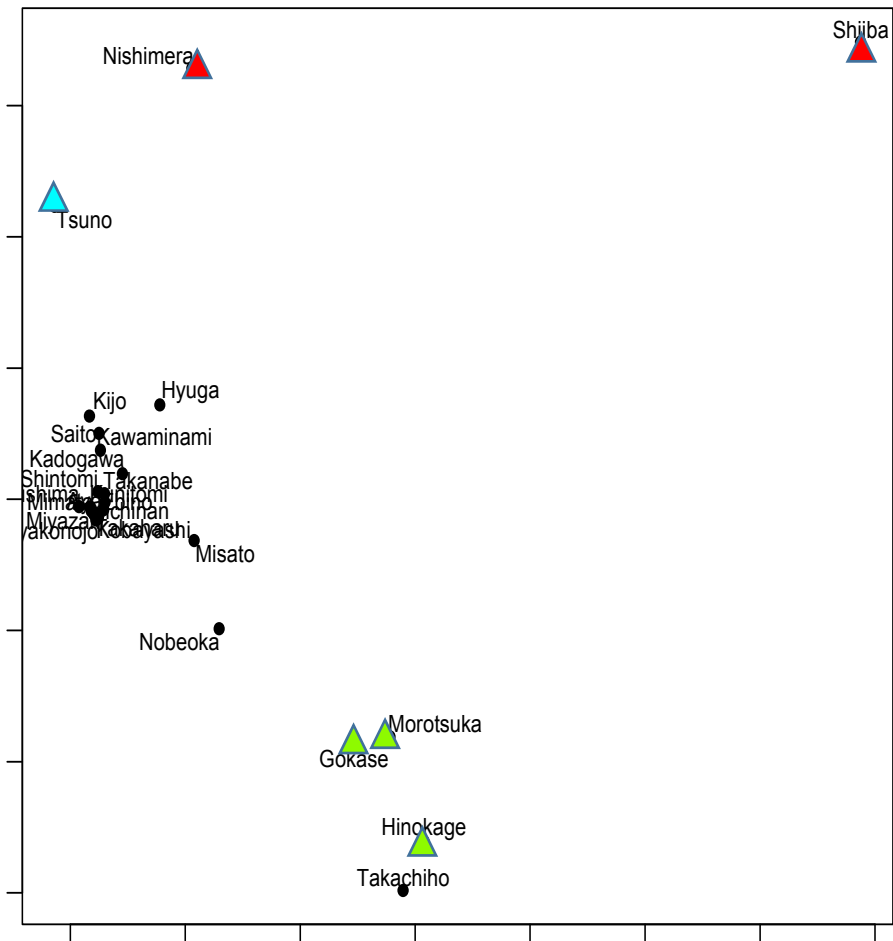
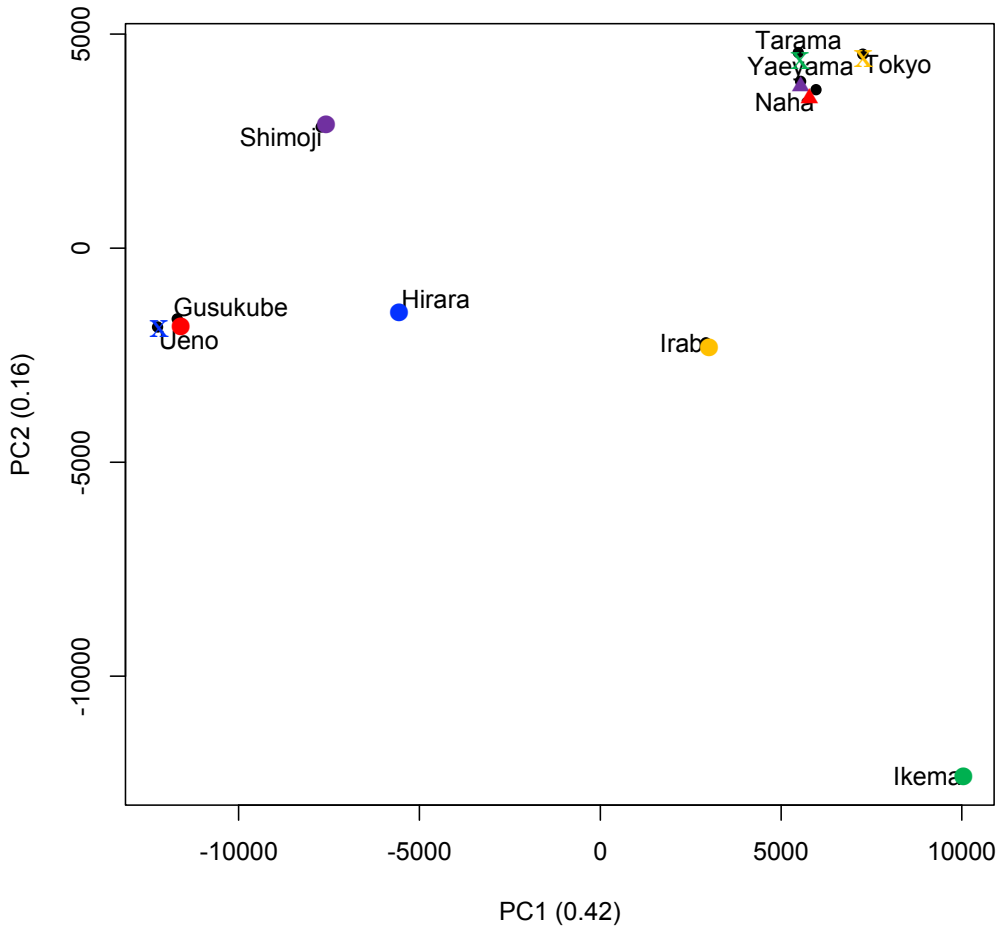
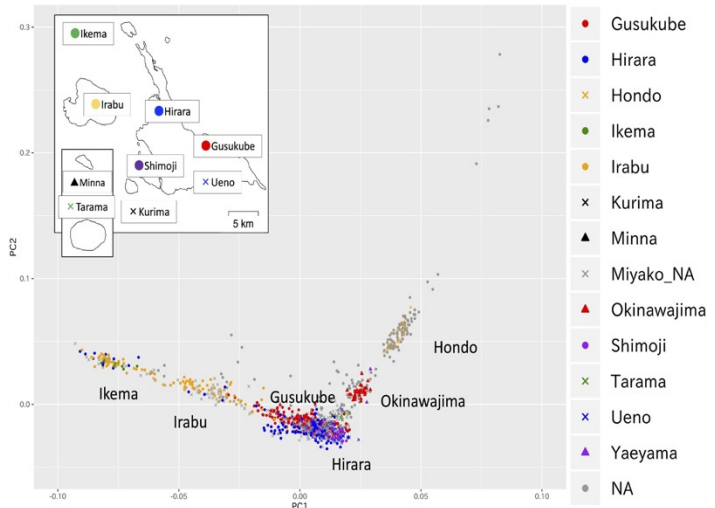


図 17 沖縄県宮古島地区名字による主成分分析



右図は Matsunami らの Fine-Scale Genetic Structure and Demographic History in the Miyako Islands of the Ryukyu Archipelago (2021) による



7. まとめ

名字は頻度の高いものでも地理的なかたよりが見られ、その分布は歴史を反映している。地名からとられた名字も多いが、特定の地名がある県で、必ずしもその名字の頻度が高いとは限らなかった。県単位ではなく市町村単位で調べれば地名と名字の関連が見えてくるのかもしれない。人口流動の少ない地域における名字と遺伝学的データの比較では、父系の Y ハプログループだけではなく、ゲノムデータとの関連性も見られた。

参考文献/References

- 奥富敬之 (2017) 『名字の歴史学』 東京：講談社
- 尾脇秀和 (2021) 『氏名の誕生』 東京：筑摩書房
- 桐村喬 (2015) 「名字からみた日本人の大都市圏における沖縄系住民の分布パターン」 『地理情報システム学会講演論文集』 24:2015.
- 武光誠 (2012) 『知っておきたい日本の名字と家紋』 東京：KADOKAWA
- 豊田武 (2012) 『苗字の歴史』 東京：吉川弘文館
- 野原寛文, 前田育子, 久澄倫之介, 内山岳人, 平島紘子, 中田大仁, 大野里香, 長谷川哲朗, 清水健史 (2021) 「宮崎県在住者におけるY-STRハプロタイプおよびY-ハプログループの地理的分布」 『日本法科学技術学会誌』 26(1): 17-27.
- 森岡浩 (2011) 『名字でわかる日本人の履歴書』 東京：講談社
- 森岡浩 (2017) 『名字でわかるあなたのルーツ』 東京：小学館
- 森岡浩 (2019) 『47都道府県・名字百科』 東京：丸善出版
- 矢野桂司 (2007) 「日本の名字マップとその応用可能性について」 『人文科学とコンピュータシンポジウム論文集』 2007-15, 47-54.
- Matsunami, Masatoshi, Kae Koganebuchi, Minako Imamura, Hajime Ishida, Ryosuke Kimura, Shiro Maeda (2021) Fine-Scale Genetic Structure and Demographic History in the Miyako Islands of the Ryukyu Archipelago, *Mol Biol Evol.* 4;38(5): 2045-2056. doi: <https://doi.org/10.1093/molbev/msab005>
- Jinam, Timothy, Yosuke Kawai, Yoichiro Kamatani, Shunro Sonoda, Kanro Makisumi, Hideya Sameshima, Katsushi Tokunaga, Naruya Saitou (2021) Genome-wide SNP data of Izumo and Makurazaki populations support inner-dual structure model for origin of Yamato people. *J Hum Genet.* 66(7): 681-687. doi: <https://doi.org/10.1038/s10038-020-00898-3>

出版情報

採用決定日：2022年8月31日

Research article

Connecting Southern Khams in geolinguistics: A brief survey on ‘fish’ and ‘pig’ beyond provinces

SUZUKI, Hiroyuki
Kyoto University

Sonam Wangmo
Universität Bonn

Tsering Samdrup
SOAS, University of London

Abstract: This article introduces the geolinguistic methodology for the principal part of the ‘Southern Khams dialects’ in traditional Tibetan linguistics to examine the dialectal variations within them. The target dialects are spoken in the crossing area of Yunnan, Sichuan, and Tibet Autonomous Region, China, connected by two national routes—G318 and G214. The article discusses dialectal variations of two word forms—‘fish’ (*nya*) and ‘pig’ (*phag*)—in Tibetic languages, focusing on phonological features corresponding to three Literary Tibetan (LT) patterns: initial consonant corresponding to LT *ny*, vowel corresponding to LT *a* in an open syllable, and rhyme corresponding to LT *ag*.*

Keywords: Southern Khams region; Southern Route group; sDerong-nJol group; lexical form; dialectology

1. Introduction

Previous geolinguistic research on Khams Tibetan has faced the issue that linguistic data from Tibet Autonomous Region (TAR) are rarely accessible. Hence, the quantity of datasets is unbalanced between TAR and non-TAR regions. This article primarily connects the data of Khams Tibetan dialects spoken in Sichuan and Yunnan with TAR along two main traffic routes—G318 and G214—to overview a geolinguistic situation of ‘Southern Khams’. This includes dialects located in Yunnan and Sichuan as well as Chamdo Municipality and rDzayul County (Nyingkhri Municipality) beyond the Jinshajiang River.

SUZUKI, Hiroyuki, Sonam Wangmo, and Tsering Samdrup. 2022. Connecting Southern Khams in geolinguistics: A brief survey on ‘fish’ and ‘pig’ beyond provinces. *Studies in Geolinguistics* 2: 29–39. doi: <https://doi.org/10.5281/zenodo.7121496>

* The present study is funded by two Grants-in-Aid for Scientific Research from the Japan Society for the Promotion of Science [JSPS]: “Investigation of Undocumented Languages in the Eastern Tibetosphere and their Geolinguistic Research” (headed by Hiroyuki Suzuki, No. 17H04774) and “Geolinguistic Studies of China and Adjacent Multilingual Areas Using High-resolution and Wide-area Maps” (headed by Mitsuaki Endo, No. 18H00670).

The target area includes the following regions:

Yunnan Province

nJol [Deqin]

mBalung [Weixi] (Budy [Badi] Township only)

Sems-kyi Nyila (rGyalthang) [Xianggelila] (except for the southern half area)

Gongshan

Sichuan Province

mBathang [Batang]

sDerong [Deirong]

Chaphreng [Xiangcheng]

Lithang [Litang] (the southwestern part only)

TAR¹

sMarkhams [Mangkang]

mDzogong [Zuogong]

rDzayul [Chayu] (Tshawarong [Chawalong] Township)

The article principally deals with Khams Tibetan,² although other languages are also spoken in the target area, such as Lisu (Mu and Sun 2012, Suzuki 2019a), Naxi (He 2015), Nosu, Selibu (Zhou 2018, Zhou and Suzuki 2020, 2022), Nung, Trung (Qin and Suzuki 2016), Lamo, Larong sMar, and Drag-yab sMar (Suzuki et al. 2018, 2021ac, 2022ab; Tashi Nyima and Suzuki 2019), as well as various regiolects of Southwestern Mandarin. See Roche and Suzuki (2017) for their distribution (except for TAR).

Map 1 shows the target area and location of the varieties by province. Generally, geolinguistic studies do not limit specific administrative units; rather, they consider linguistic and cultural areas. Among the previous studies in the present target area, we find, for example, Suzuki (2018, 2021) for Yunnan Tibetan. However, linguistic classification, as in Map 2, is different from the boundary of administrative units. The target varieties belong to various dialectal groups, including Southern Route (SR), sPomgorgang (PG), Chaphreng (CP), sDerong-nJol (DJ), rDzayul (DZ), and Sems-kyi-nyila (SN) from Khams Tibetan, as well as Amdo Tibetan (AM).³

¹ See, for example, Suzuki et al. (2021b) for a geolinguistic approach to Tibetic varieties in the TAR's target region.

² A similar approach is taken by Suzuki (2022b) on the word form for 'dog'.

³ Based on the description of Suzuki (2018abc) with updates of Suzuki (2020, 2022). The classification of SR is preliminary and warrants a rigorous examination. See also Tournadre and Suzuki (2022).

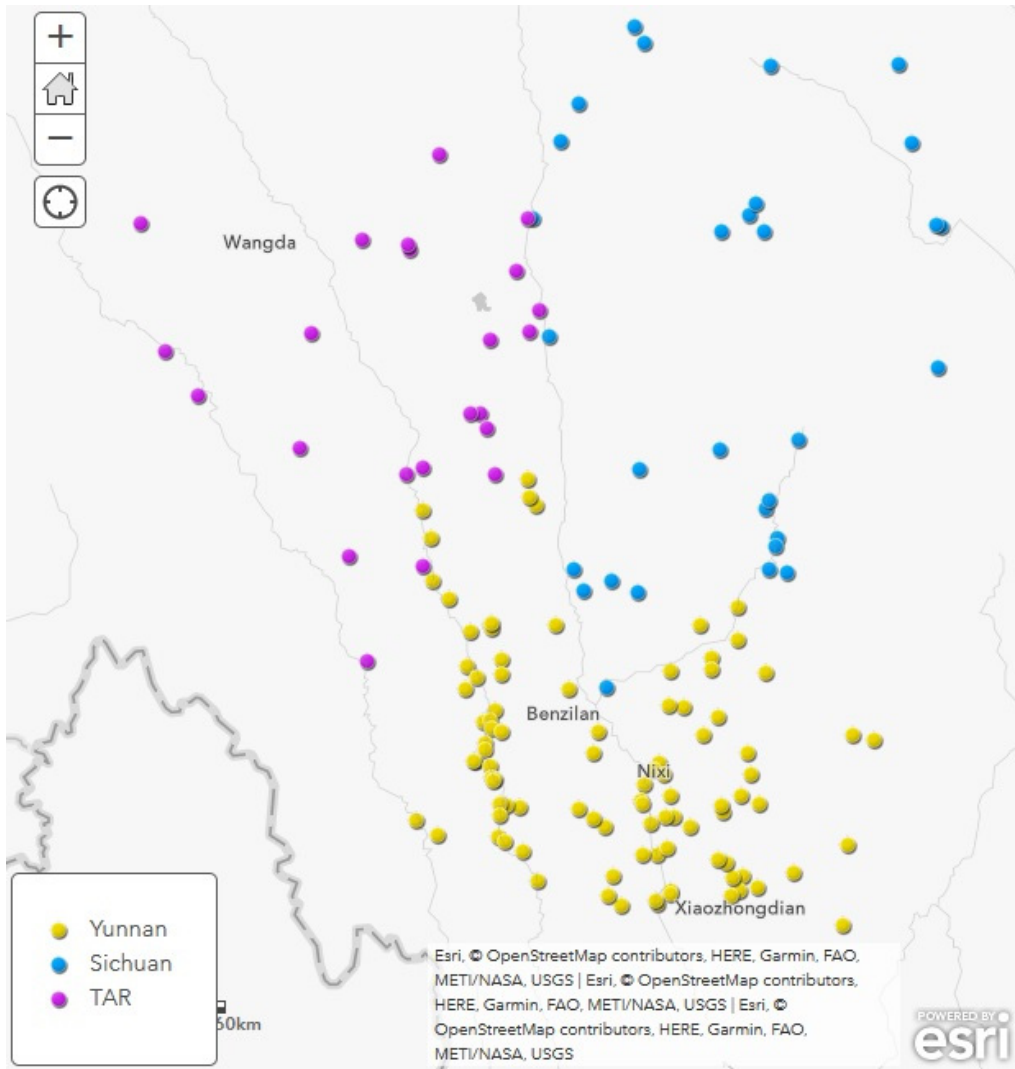


Fig. 1 Target area and data points for mapping.

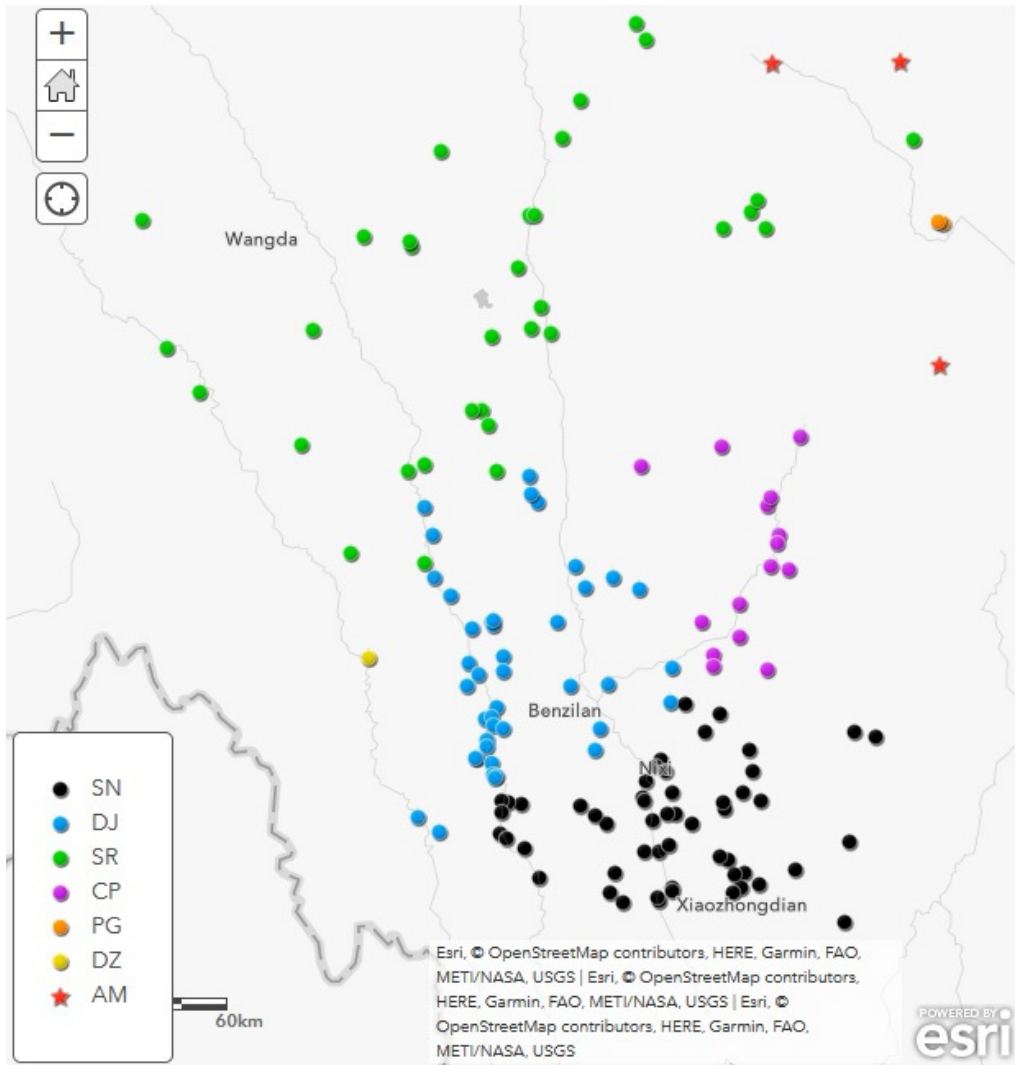


Fig. 2 Language classification in the target area.

This article examines sounds of two monosyllabic words—‘fish’ and ‘pig’. Their Literary Tibetan (LT) forms are *nya* and *phag*, respectively. It focuses on three of their features:

- (1) Head of ‘fish’: initial consonant corresponding to LT *ny*;
- (2) Tail of ‘fish’: vowel corresponding to LT *a* in an open syllable; and
- (3) Tail of ‘pig’: rhyme corresponding to LT *ag*.

We do not discuss the initial consonant of the word form for ‘pig’, as it is characterised as /p^h/; however, in Amdo Tibetan and a few other varieties, it corresponds to /h/. See also Suzuki (2007, 2019b). For the phonetic description in the data, we follow Suzuki’s (2016) method.

2. Head of ‘fish’

This section examines the initial consonant corresponding to LT *ny*. As Map 3 displays, two sound correspondences are attested in the target area.

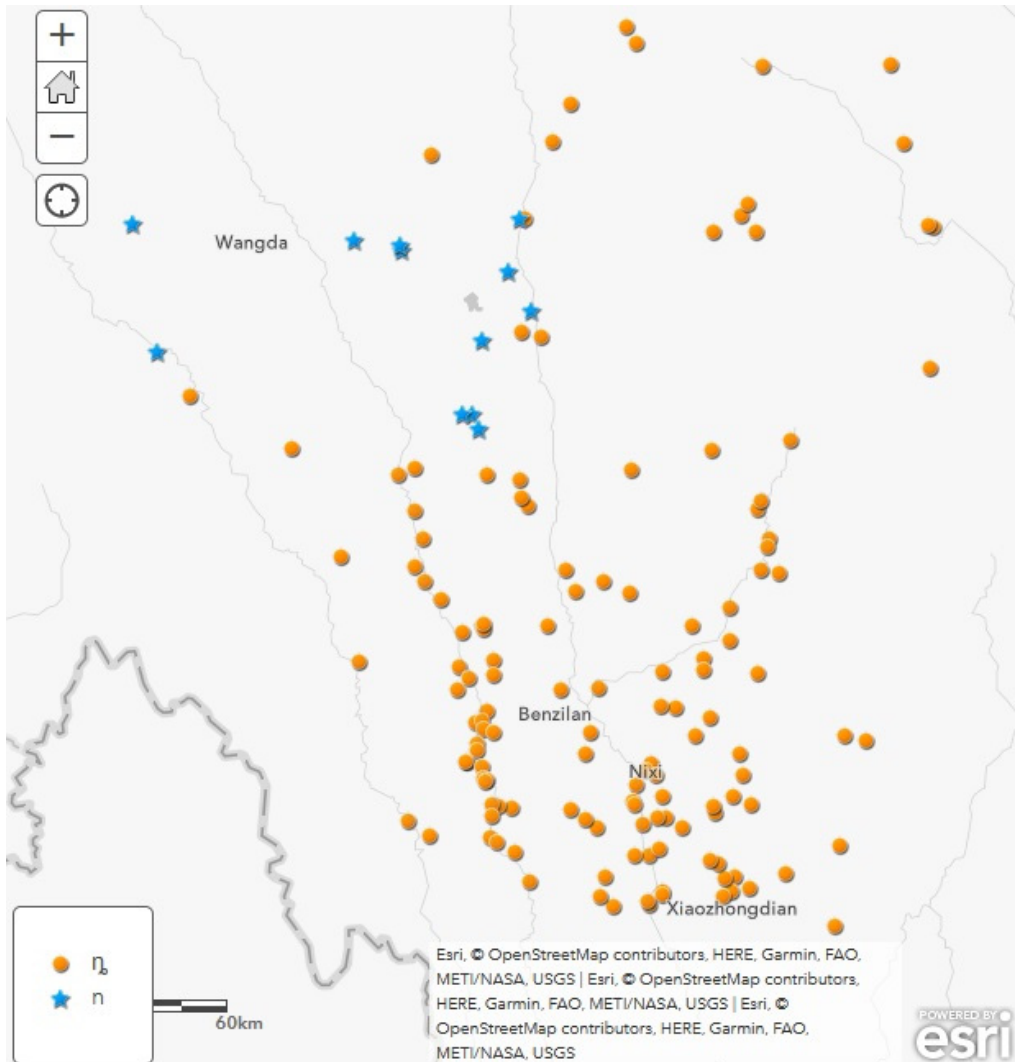


Fig. 3 Initial consonant of the word form for ‘fish’.

Type /ŋ/ is ubiquitous, and Type /n/ is marked in the Tibetic languages. As Map 3 shows, the latter appears principally in sMarkhams as well as mDzogong counties. The distribution range is characterised by Route G318 between the Jinshajiang and Nujiang rivers. Notably, along the Jinshajiang River, some varieties use Type /ŋ/, while some others use Type /n/. The river has probably been a boundary of the given linguistic feature; however, we must investigate the mutual relationship, including intermarriage and business traffic, between the riversides.

This classification influences the word forms for ‘fish’ in non-Tibetic languages. For example, the Lamei dialect of Lamo uses /nɛ/ ‘fish’, which can be borrowed from the donor language (Tibetic) using Type /n/. Map 3 shows that Tibetic varieties used in the Lamo-speaking area belong to Type /n/. Therefore, it is reasonable that Lamo has borrowed the word from a variety distributed in its vicinity.

3. Tail of ‘fish’

This section examines the vowel corresponding to LT *a* in an open syllable. As Map 4 displays, six sound correspondences are attested in the target area.

The unmarked sound is /a/, which is assumed to be a straightforward sound correspondence with LT *a*. Type /a/ is the most widely distributed. Of the remaining sounds, /e/ is considered exceptional as a sound corresponding to LT *a* in Tibetic languages. A similar sound is known as ‘brightening’ (Matisoff 2004) in languages of the Qiangic and rGyalrongic branches, including Lamo, Larong sMar, and Drag-yab sMar; see some lexical forms cited in Suzuki et al. (2018, 2021ab, 2022a) and Tashi Nyima and Suzuki (2019).

We analyse the other sounds as follows: The sound /ɐ/ is considerably close to the unmarked type /a/, which is only attested in the Myigzur dialect. The remaining three sounds /ɑ, ɔ, o/ form a category that can be labelled as the ‘back’ vowel. We can hypothesise a sound change process /ɑ/ > /ɔ/ > /o/ based on the synchronic distribution. Note that the sound /o/ is found more widely in the east of the target area (dialects of the sPomborgang group); see Suzuki (2018ac), Li and Suzuki (2020), and Suzuki and Li (2022).

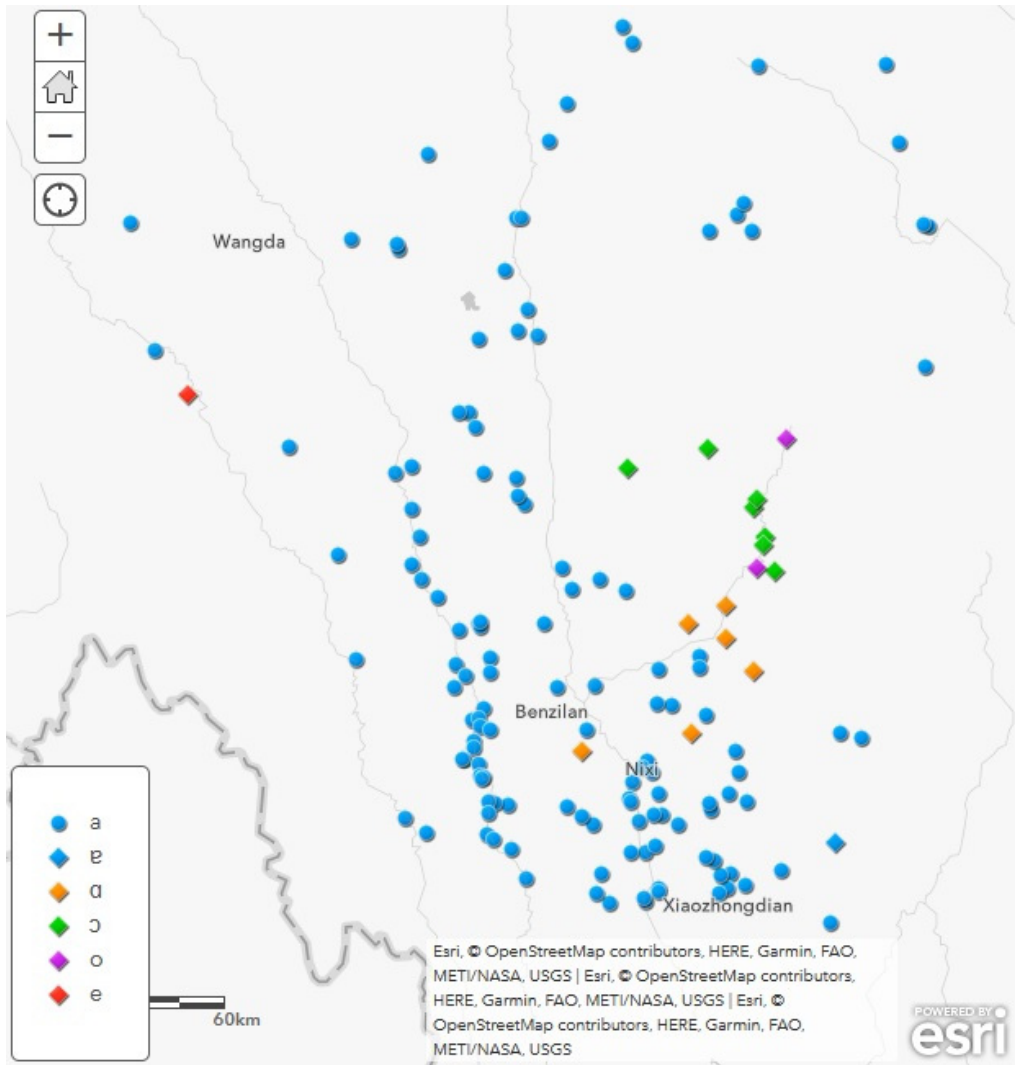


Fig. 4 Vowel of the word form for 'fish'.

4. Tail of 'pig'

This section examines the rhyme corresponding to LT *ag* in an open syllable. As the sound correspondence of LT *ag* is concerned, a glide is also included despite its phonological status as part of initial consonant clusters. As Map 5 displays, eight sound correspondences are attested in the target area.

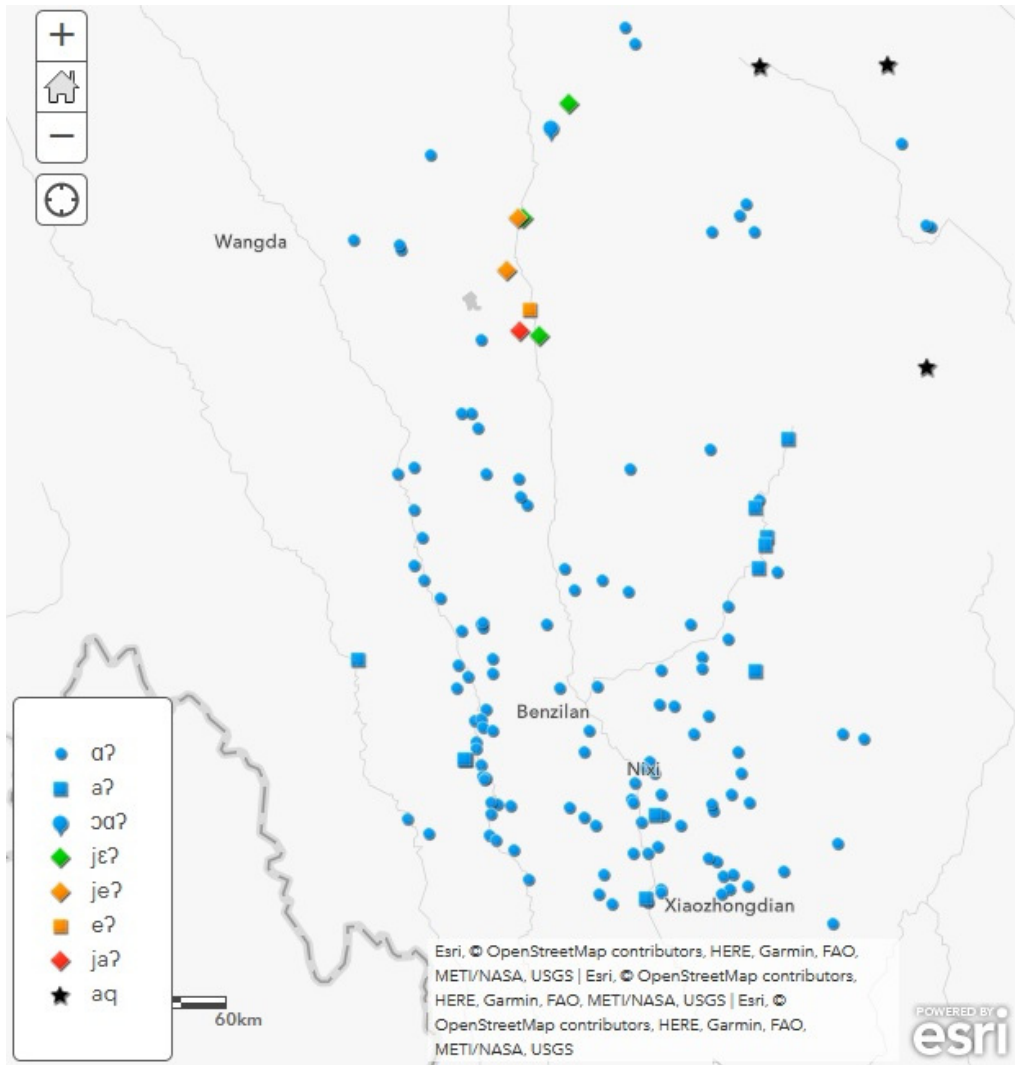


Fig. 5 Rhyme of the word form for 'pig'.

The unmarked sound is /aʔ/; it is the most widely distributed. A similar form /aʔ/ is found around Type /aʔ/. Type /ɔaʔ/ is only found in the mBathang dialect; owing to its phonetic feature, we believe that it is derived from the back vowel form /aʔ/. Type /aq/ is found only in Amdo Tibetan. The remaining form is connected to front mid vowels /e/ and /ɛ/, as well as a palatal glide /j/, which we call Type E for simplicity. These types are principally found in varieties along the Jinshajiang River. From the geolinguistic viewpoint, the form of the mBathang dialect (Type /ɔaʔ/) is peculiar, as it is surrounded by varieties using Type E. Type E was once discussed in varieties distributed in a wider region (Suzuki 2007:37-38), displaying a scattered distribution

(see also Suzuki 2010, 2011). Map 5 shows that the distribution of Type E expands to a certain geographical range; hence, we can collect more data to examine whether these varieties can form an independent dialect (sub-)group.

5. Conclusion

The article examined three linguistic phenomena to overview the geolinguistic situation of ‘Southern Khams’, including dialects located in Yunnan and Sichuan as well as TAR. They are (1) initial consonant corresponding to LT *ny*, (2) vowel corresponding to LT *a* in an open syllable, and (3) rhyme corresponding to LT *ag*. Each linguistic map shows different distribution patterns of the given features in the target area. The results suggest a possibility of further division into (sub-)groups of Khams Tibetan, including varieties spoken in TAR.

References

- He, Jichuan [和继全] (2015) *Baidi Bowancun Naxi Dongbawen diaocha yanjiu* 《白地波湾村纳西东巴文调查研究》[*Research in Naxi Dongba scripts in Bowan Village, Baidi*]. Beijing: Minzu Chubanshe.
- Li, Chunmei and Hiroyuki Suzuki [李春梅、铃木博之] (2020) Kangdingshi Jiju Zangyu suoshu wenti 康定市吉居藏语所属问题 [Issues in the dialectal affiliation of sKyidyul Tibetan in Darmdo Municipality]. *Minzu Xuekan* 5: 102–109+154–156.
- Matisoff, James A. (2004) ‘Brightening’ and the place of Xixia (Tangut) in the Qiangic subgroup of Tibeto-Burman. In: Dah-an Ho (ed) *Studies on Sino-Tibetan languages: Papers in honor of Professor Hwang-Cherng Gong in his 70th birthday*, 327–352. Taipei: Academia Sinica.
- Mu, Yuzhang and Hongkai Sun [木玉璋、孙宏开] (2012) *Lisuyu fangyan yanjiu* 《傈僳语方言研究》[*Study on Lisu dialects*]. Beijing: Minzu Chubanshe.
- Qin, Liying and Hiroyuki Suzuki (2016) Chasing a cat from the Mekong to the Salween: A geolinguistic description of ‘cat’ in Trung and Khams Tibetan in North-western Yunnan. *Studies in Asian Geolinguistics I —Sun—*, 61–71. [Revised edition available in Suzuki (2022a: 165–178)] URI: https://publication.aa-ken.jp/sag1_sun_2016.pdf
- Roche, Gerald and Hiroyuki Suzuki (2017) Mapping the linguistic minorities of the eastern Tibetosphere. *Studies in Asian Geolinguistics VI —“Means to Count Nouns” in Asian Languages—*, 28–42. URI: https://publication.aa-ken.jp/sag6_count_2017.pdf
- Suzuki, Hiroyuki [鈴木博之] (2007) Sensei minzoku sooroo Tibettogo hoogen ni okeru ‘buta’ wo arawasu go 川西民族走廊・チベット語方言における「ぶた」を表す語 [Words for ‘pig’ in Tibetan dialects spoken in the Ethnic Corridor of West Sichuan]. *Kyoto University Linguistic Research* 26: 31–57 doi: <https://doi.org/10.14989/57308>
- Suzuki, Hiroyuki [鈴木博之] (2010) Kamutibettogo Qiongbo/Chongcang [Khyungpo/Khromtshang] hoogen no onsei bunseki to sono hoogen tokutyoo カムチベット語瓊波/冲倉 [Khyungpo/Khromtshang] 方言の音声分析と方言特徴 [Khams Tibetan Khyungpo/Khromtshang dialect: Phonetic analysis and dialectal characteristics]. *Journal of Asian and African Studies* 79: 95–120. URI: <http://hdl.handle.net/10108/57437>

- Suzuki, Hiroyuki (2011) Phonetic analysis of dGudzong Tibetan: The vernacular of Khams Tibetan spoken in the rGyalrong area. *Bulletin of National Museum of Ethnology* 35.4: 617–653. doi: <https://doi.org/10.15021/00003880>
- Suzuki, Hiroyuki (2016) In defense of prepalatal non-fricative sounds and symbols: towards the Tibetan dialectology. *Researches in Asian Languages* 10: 99–125. URI: <http://id.nii.ac.jp/1085/00002195/>
- Suzuki, Hiroyuki [鈴木博之] (2018a) Litang xian ji qi zhoubian Zangzu yuyan xianzhuang diaocha yu fenxi 理塘县及其周边藏族语言现状调查与分析 [Current situation of Tibetans' languages in Lithang County: Research and analysis]. *Minzu Xuekan* 2: 35–44+106–109.
- Suzuki, Hiroyuki (2018b) *100 linguistic maps of the Swadesh word list of Tibetic languages from Yunnan*. Fuchu: Research Institute for Languages and Cultures of Asia and Africa. URI: https://publication.aa-ken.jp/sag_mono3_tibet_yunnan_2018.pdf
- Suzuki, Hiroyuki [鈴木博之] (2018c) Kangba Zangyu Bengbogang fangyanqun ji qi yuyin tezheng 康巴藏语崩波岗方言群及其语音特征 [Khams Tibetan sPomborgang dialect and its phonetic features]. *Sichuan Minzu Xueyuan Xuebao* 5: 66–72.
- Suzuki, Hiroyuki [鈴木博之] (2019a) Tibetto bunkaken ni kurasu Risuzoku no hanasu Kamutibetogo hoogen: Deqin-ken Xiarmo-kyoo Buyapei-mura no gengo zizyoo チベット文化圏に暮らすリス族の話すカムチベット語方言：徳欽県霞若郷布亞培村の言語事情 [Khams Tibetan vernacular spoken by Lisu in the Tibetosphere: Language situation in Buyapei Hamlet, Xiarmo Township, Deqin County]. *Journal of Kijutsuken* 11: 1–16. URI: <http://id.nii.ac.jp/1422/00003017/>
- Suzuki, Hiroyuki (2019b) How Tibetans classify *pigs* in their languages in the eastern Tibetosphere: Revisiting the *pig* issue through geolinguistics. In: Hiroyuki Suzuki, Keita Kurabe, and Mitsuki Endo (eds) *Papers from the Workshop “Phylogeny, Migration, and Contact of East and Southeast Asian Languages and Human Groups”*, 40–53. Fuchu: Research Institute for Languages and Cultures of Asia and Africa. [Revised edition available at Suzuki (2022a: 315–331)] URI: https://publication.aa-ken.jp/sag_mono7_phylogeny_dispersion_contact_2019.pdf
- Suzuki, Hiroyuki (2020) Geolinguistic significance of the Phongpa dialect in the history of Yunnan Tibetan. *Proceedings of the Second Annual Meeting of the Geolinguistic Society of Japan*, 25–29. [Revised edition available in Suzuki (2022a: 141–148)]
- Suzuki, Hiroyuki (2021) Cong diliyuyanxue de jiaodu kan Yunnan Zangyu /l/ j/ de lishi fazhan 从地理语言学的角度看云南藏语/l/及j/的历史发展 [Historical development of the /l/ and /j/ sounds in Yunnan Tibetan from a geolinguistic context]. In: Hiroyuki Suzuki and Mitsuki Endo (eds) *Zhongguo yuyan dili yanjiu lunwenji*, 21–38. Fuchu: Research Institute for Languages and Cultures of Asia and Africa. doi: <https://doi.org/10.15026/116769>
- Suzuki, Hiroyuki (2022a) *Geolinguistics in the eastern Tibetosphere: An introduction*. Tokyo: Geolinguistic Society of Japan. doi: <https://doi.org/10.5281/zenodo.5989176>
- Suzuki, Hiroyuki (2022b) Dian-Chuan-Zang jiaojiechu Zangyuzhi yuyan zhong de “gou”: Xiangguan yuyin xingshi de diliyuyanxue fenxi 滇川藏交界处藏语支语言中的“狗”——相关语音形式的地理语言学分析 [‘Dog’ in Tibetic languages of the boundary area of Sichuan, Yunnan, and Tibet: A geolinguistic analysis of its word form]. In: *Iwata Ray jiaoshou rongxiu jinian lunwenji bianjizu* (ed.) *Iwata Ray jiaoshou rongxiu jinian lunwenji (shangce)*, 253–263. Tokyo: Geolinguistic Society of Japan. doi: <https://doi.org/10.5281/zenodo.6342364>
- Suzuki, Hiroyuki and Chunmei Li (2022) Vocabulary of the Chahualin dialect of sPomborgang Khams (Jiulong County). *Asian and African Languages and Linguistics (AALL)* 16: 261–289. doi: <https://doi.org/10.15026/117165>
- Suzuki, Hiroyuki, Sonam Wangmo, and Tsering Samdrup (2021a) Lamei, another dialect of Lamo (mDzong, TAR): Vocabulary and sentence structure. In: Yasuhiko Nagano and Takumi Ikeda (eds) *Grammatical phenomena of Sino-Tibetan languages 5: Link languages and archetypes in Tibeto-Burman*, 25–69. Kyoto: Institute for Research in Humanities, Kyoto University. URI: <http://hdl.handle.net/2433/263977>

- Suzuki, Hiroyuki, Sonam Wangmo, and Tsering Samdrup (2021b) Geolinguistic analysis of ‘hand’, ‘wind’, and ‘moon’ in Tibetic languages in sMarkhams, mDzogong, and rDzayul counties. *Studies in Asian and African Geolinguistics II —Grammatical relations—*, 48–56. URI: https://publication.aa-ken.jp/saag2_grammatical_relations_2021.pdf
- Suzuki, Hiroyuki, Sonam Wangmo, and Tsering Samdrup [鈴木博之、四郎翁姆、才讓三周] (2022) Raron-Ma [Larong sMar] go Rumei [Rongsmad] hoogen no goi siryo (Nitiei taisyoo) ラロン・マ [Larong sMar]語如美[Rongsmad]方言の語彙資料 (日英対照) [Wordlist of the Rongsmad dialect of Larong sMar (Japanese-English)]. *Journal of Kijutsuken* 14: 27–63. URI: <http://id.nii.ac.jp/1422/00004410/> [Suzuki et al. (2022a)]
- Suzuki, Hiroyuki, Tsering Samdrup, and Sonam Wangmo (2018) Contrastive word list of three non-Tibetic languages of Chamdo——Lamo, Larong sMar, and Drag-yab sMar——. *Kyoto University Linguistic Research* 37: 79–104. doi: <https://doi.org/10.14989/240980>
- Suzuki, Hiroyuki, Tsering Samdrup, and Sonam Wangmo [鈴木博之、才讓三周、四郎翁姆] (2021) Taya-Ma (Drag-yab sMar) go Bae (mBengo) hoogen no goi siryo (Nitiei taisyoo) タヤ・マ(Drag-yab sMar)語巴俄(mBengo)方言の語彙資料 (日英対照) [Wordlist of the mBengo dialect of Drag-yab sMar (Japanese-English)]. *Journal of Kijutsuken* 13: 189–213. URI: <http://id.nii.ac.jp/1422/00004151/> [Suzuki et al. (2021c)]
- Suzuki, Hiroyuki, Tashi Nyima, Tsering Samdrup, and Sonam Wangmo [鈴木博之、扎西尼瑪、才讓三周、四郎翁姆] (2022) Changdu shinei xin renzhi yuyan de shuci jieou 昌都市内新认知语言的数词结构 [Numeral construction of the newly-recognised languages in Chamdo Municipality]. *Nankai Yuyan Xuekan* 1: 148–158. [Suzuki et al. (2022b)]
- Tashi Nyima and Hiroyuki Suzuki (2019) Newly recognised languages in Chamdo: Geography, culture, history, and language. *Linguistics of the Tibeto-Burman Area* 42.1: 38–82. doi: <http://doi.org/10.1075/ltba.18004.nyi>
- Tournadre, Nicolas and Hiroyuki Suzuki (2022) *The Tibetic languages: An introduction to the family of languages derived from Old Tibetan* (with the collaboration of Xavier Becker and Alain Brucelle for the cartography). Villejuif: LACITO Publications.
- Zhou, Yang [周洋] (2018) Yunnan Shuimofanghua de gebiaoji ji qi lai yuan 云南水磨房话的格标记及其来源 [Case markers in Selibu in Yunnan and their origin]. *Fangyan* 3: 357–369.
- Zhou, Yang and Hiroyuki Suzuki [周洋、鈴木博之] (2020) Shuimofanghua ti fanchou de hunhe tezheng 水磨房话体范畴的混合特征 [Mixed features of the aspect category in Selibu]. *Minzu Yuwen* 4: 43–56. URI: <http://www.mzyw.net.cn/Magazine/Show?id=75666>
- Zhou, Yang and Hiroyuki Suzuki (2022) Evidentiality in Selibu: A contact-induced emergence. *Diachronica* 39.2: 268–309. doi: <https://doi.org/10.1075/dia.19055.zho> ; With an online appendix. doi: <https://doi.org/10.1075/dia.19055.zho.additional>

Publication history

Date received: 18 April 2022

Date accepted: 31 August 2022

論文

東部チベット系諸言語における「野うさぎ」の語形 チベット文語形式との音対応を中心に

鈴木博之
京都大学

Lexical forms of ‘hare’ in eastern Tibetic languages: Focusing on a sound correspondence with Literary Tibetan forms

SUZUKI, Hiroyuki
Kyoto University

Abstract: This article discusses word forms for ‘hare’ in Tibetic languages spoken in the eastern Tibetosphere (principally Gansu, Sichuan, and Yunnan, as well as a part of Chamdo and Myanmar). There are two principal word forms derived from Literary Tibetan *ri bong* and *spang g.yag* (literally meaning ‘grassland yak’). The former is attested widely, and the latter is attested in the south-eastern edge of the Tibetosphere. There are various phonetic realisations for the word forms derived from *ri bong*. Of them, the article focuses on the variation of sound correspondences of the second syllable, which is classified in several types: labial, velar, uvular, and glottal. By drawing linguistic maps, the article first describes details of the distribution of the form derived from *spang g.yag*: its essential distribution area is the south-eastern part of the Tibetosphere in Yunnan, of which varieties spoken in the rGyalthag area use the form derived from *ri bong*. Second, it describes the sound differences in the *ri bong*-type, exhibiting a relationship between the velar and uvular types as well as between the glottal and monosyllabic types.*

キーワード: チベット系諸言語; 東チベット文化圏; 語形式; 音対応

Keywords: Tibetic languages; eastern Tibetosphere; word form; sound correspondence

1. はじめに

チベット系諸言語では、一般的に「うさぎ」の語形について、「野うさぎ」と「家うさぎ」が語彙的に区別される。チベット文語（蔵文）形式¹では、前者

鈴木博之(2022)「東部チベット系諸言語における「野うさぎ」の語形：チベット文語形式との音対応を中心に」『地理言語学研究』2: 40–52. doi: <https://doi.org/10.5281/zenodo.7121520>

* 本稿にかかる現地調査は、科研費（課題番号16102001, 07J00250, 21251007, 25770167, 16H02722, 17H04774, 18H00670）の援助を受けている。

¹ 蔵文形式は、de Nebesky-Wojkowitz (1956)によるローマ字転写で示す。

が *ri bong*、後者が *yos* として知られる^{2,3}。この中で、特に前者については語形・音形ともにさまざまな方言的・地理的差異が現れる。筆者は東部チベット文化圏 (eastern Tibetsphere) のチベット系諸言語について、300 地点前後の資料を収集した。その中で、「野うさぎ」の語形は次の2点に基づいて分類することが可能である。

- ・ *ri bong* に語彙的な対応関係が認められるか否か
- ・ *ri bong* に語彙的な対応関係がある場合、その音対応がどうであるか

特に、*ri bong* に対応する形式で、その第2音節に対応する形式が多様であることから、これについて、地理言語学的に意味のある解説を与えられるかは、興味のある問題となる。本稿では、東チベットのチベット系諸言語を対象に、以上の2点に基づいて語形を整理したのち、ArcGIS online を用いて言語地図を作成し、地理言語学的に有意な特徴が現れるかを検討する。語形の表記には音標文字を用いて記す⁴。地点・方言名について、中国国内の地点は漢字を用い、それ以外はローマ字表記とする。

2. 語形

本節では、「野うさぎ」の具体的な語形を整理し、具体的な形式を音標文字を用いて記述する。

2.1. *ri bong*に語彙形式上の対応関係が認められるか否か

東部チベット文化圏の多くのチベット系諸言語における「野うさぎ」の形式は、文語でもある *ri bong* に対応関係をもつ (以下、「*ri bong* 類」)。これについての詳細は 2.2 に整理し、先にそれ以外の形式について見ていく。

² 日本語ではいずれも「うさぎ」として問題なくとも、たとえば英語では *hare* と *rabbit* の区別があり、それぞれ蔵文形式の *ri bong* と *yos* に対応させることができる。ただし、各種英訳の環境では、*ri bong* を *rabbit* とすることもあり、混乱が認められる (Shelton 1925、華侃 主編2002、Kajihama 2004、Suzuki and Sonam Wangmo 2018など)。筆者のチベット系諸言語の現地調査で、「*ri bong* という動物は野生であり家で飼われることはない」という調査協力者による解説が得られた (Suzuki and Sonam Wangmo 2021、Zou and Suzuki 2022参照)。このようなことから、日本語であっても「野うさぎ」と訳すよう心がけることが有益であると考えられる。

³ 東チベットのチベット系諸言語の形式を収録語を含む『チベット語・ラテン語・フランス語辞典』DTLF (1899:936)では、見出し語に *ri gong* も掲げている。語義としては、*lièvre/lepus* 「野うさぎ」を当てているが、*lapin* 「家うさぎ」は *khang pa'i ri gong* (直訳は「家の野うさぎ」という形式を与えている。筆者はこの形式については調査記録には認められない。一方、蔵文 *yos* にも *lièvre/lepus* 「野うさぎ」と訳されている (1899:919)。Girardeau et Goré (1956) の『フランス語・東チベット語辞典』は、南部カム地域、特に雲南で話される口語形式を反映している (鈴木2021参照) が、*lièvre* 「野うさぎ」にも *lapin* 「家うさぎ」にも *ri gong*、*ri bong* を掲げ、十二支の「う」に *yos* を掲げている (1956: 166, 168)。

⁴ 分節音については、Suzuki (2016) を参照。超分節音の表記については、Suzuki (2022) を参照。

ri bong 類以外で複数の地点で記録されたものとしては、spang g.yag 類と呼んでまとめられる一連の形式が認められる。この語形式は蔵文には認められないが、直訳すれば「草原のヤク」となる。たとえば、^hpō^hjaʔ, ^hpū zaʔ, ^hpu jā, ^hpō zaʔ/のような形式があり、これらをそれぞれの言語における蔵文との音対応を踏まえて考えれば、蔵文で spang g.yag のような形式が想定される。

これ以外には、次のような形式を認めることができる。

- ・蔵文 a khu ri bong (直訳「野うさぎおじさん」) の対応形式
- ・蔵文 a myes ri bong (直訳「野うさぎおじいさん」) の対応形式
- ・蔵文 yos (直訳「家うさぎ」) の対応形式
- ・/tə^hdə/ (蔵文との対応関係は不明)

以上のうち、最初の2つは、語形式の後半に ri bong 対応形式を含むことから、広義の ri bong 類に含めることができる。ただし、親族名称を伴う点については注目に値する⁵。

2.2. ri bong類の音対応

ri bong 類は多数派であると言えるが、その音対応は多様である。また、音対応の中の多数派は、必ずしも蔵文のつづり字と一対一の対応を示していない。以下に整理する。方言名は省略する。

表 1 : ri bong類の音対応

分類	説明	語形
Ra1	2音節形式かつ第2音節が唇音	ʼrə be, ʼrə bō, ʼrə ^h bō, ʼrə bu: ; ^rə βu: ; ^rə mō, ^fiə- mū ; ʼri wō, ^rə wō, rə wu:
Ra2	2音節形式かつ第2音節が軟口蓋音	ʼri ko:, ʼrə kō, ʼrə kū ; ^rə gō, ʼrə gō, ʼrə ^h gō, ʼrə gū ; rə ʃoŋ, rə ʃō, rə ʃu, rə ʃo, ʼro ʃā, ʼru ʃu: ; rə ʃwu:
Ra3	2音節形式かつ第2音節が口蓋垂音	^rə qō ; rə ʋoŋ
Ra4	2音節形式かつ第2音節が声門音	ʼrə hō ; ʼre: fiu, rə fiu, rə fiə
Rb	1音節形式	ru:, rō, ʼrō:

⁵ いくつかの動物名が、普通名詞としても親族名称を伴う現象は広く認められる。特に、「虎」や「熊」の事例が多い (Tournadre and Suzuki 2022参照)。動物が登場する昔話などの語りにおいては、Suzuki and Sonam Wangmo (2021)、Zou and Suzuki (2022) などにあるように、呼びかけ語として頻繁に当該現象が認められる。ただし、それらはあくまでも呼びかけであり、普通名詞として親族名称を伴うかどうかは別問題であるが、当該語形が現れる背景として考慮すべき事例である。

以上のうち、Ra1 に分類した/w/を含む形式は、実際には両唇と軟口蓋における同時調音であるから、Ra2 との調音位置上での連携もあるといえる。Ra2 には、ちょうどこの両者をつなぐように見える/rə ywu:/という形式も含めた⁶。

「野うさぎ」の蔵文 ri bong という語形について、Hill (2011:116-118)がチベット系諸言語の資料を引用して議論している。蔵文 ri bong という形式は、「山」と「ロバ」という形態素からなる複合語であると述べるが、第2音節初頭子音の音形式について、b は蔵文の指小辞'u の音対応と共通する部分があることを指摘している。指小辞'u は蔵文 bu に由来し、語中にあるときに[b > w > y]と弱化したとする。この音変化の過程が「野うさぎ」にも生じていると理解できる。

Hill (2011)は蔵文形式 ri bong の第2音節が b と対応する諸方言を主にチベット文化圏の西部に位置する地域（ラダック、阿里など）の変種に見出しているが、表1の Ra1 類に示すように、同文化圏の東部にも認められることに注意できる。また、Ra1, Ra2, Ra3, Ra4, Rb の順で第2音節の初頭子音が弱化していると見ることは可能である。一方で、Ra2 に含めた/y, g, k/は弱化と認めるには困難がある。つまり、内部において異なる音変化をたどったか、そもそも異なる語源があるのか、慎重に検討する必要がある。

3. 言語地図

本節では、2節で記述した事柄を言語地図上に表し、その地理的分布について考察を加える。各地図は2枚で1組（東チベット全体⁷及び雲南地区の拡大図）であるが、各図に通し番号を振る。

3.1. 語彙形式に関する地図

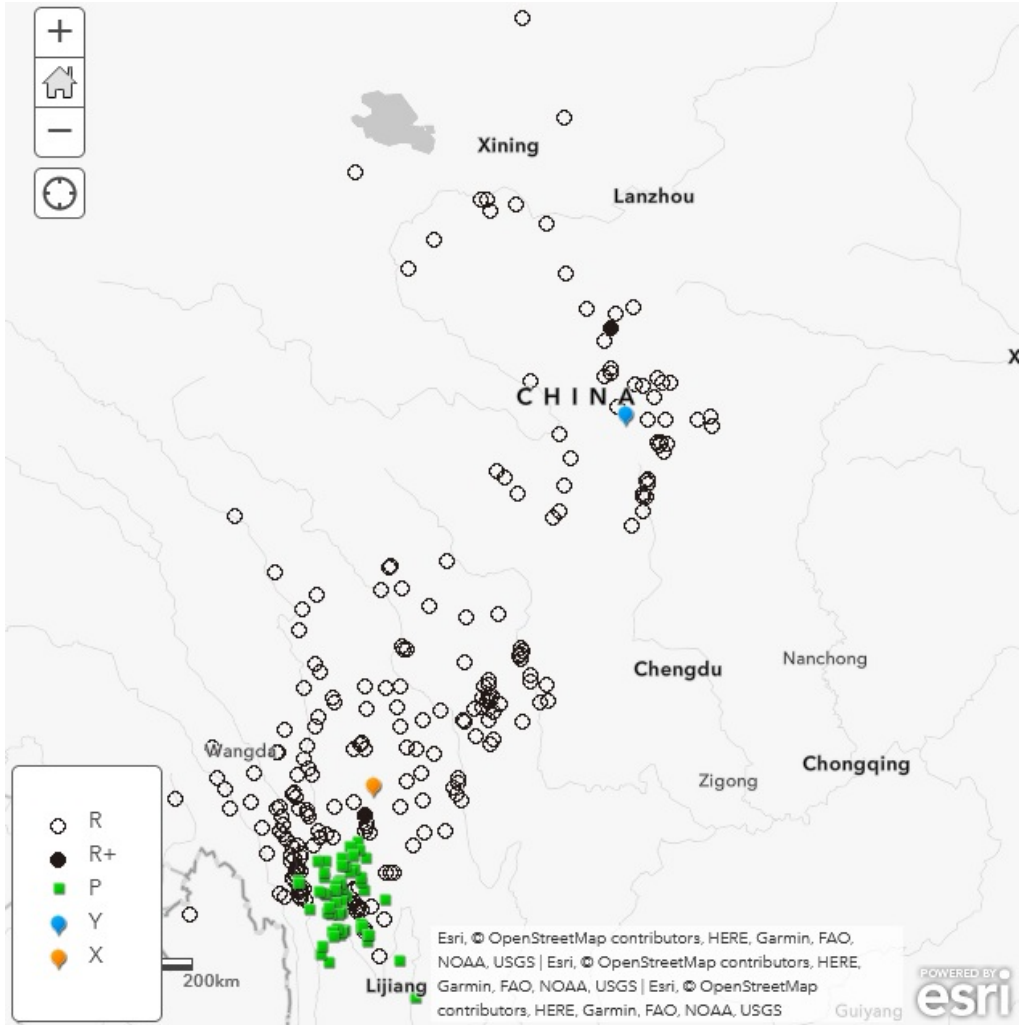
ここでは、2.1節に掲げた語形の分布を地図上に示す。表1は東チベット全体の言語地図、表2は表1の中のカム地域南部（雲南のチベット文化圏を中心とする）の拡大図である。

図1を見ると分かるように、R類（蔵文 ri bong 対応形式）が広範囲にわたり多くの地点を占めている。R+類（2地点のみ）は、R類に相当する形式が語の内部に含まれ、R類に囲まれるように分布している。Y類、X類はそれぞれ1点のみ認められ、いずれもR類に囲まれる形で分布している。一方、P類は雲南を中心に認められ、その分布地域はR類とも重なりながら、一定の連続

⁶ わたり音として現れる/w/や/j/について、蔵文との対応関係を調べると、初頭子音の性質による場合のほかにも母音+末子音との対応関係に由来するものも少なくない。後者については、鈴木(2012)などを参照。

⁷ 本稿で扱う「東チベット」の範囲は、中国青海省、甘肅省、四川省、雲南省のチベット系諸言語の分布地域で、玉樹市を除く地域、およびチベット自治区昌都市の南東部、林芝市の東端部、ならびにミャンマーのカチン州に限定する。

した範囲を形成している。R 類と P 類の重なりについては、図 2 を参照するとよくわかる。



凡例：R: 蔵文ri bong ; R+: 蔵文ri bongを含む形式 ; P: 蔵文spang g.yag ;
Y: 蔵文yos ; X: その他

図1：「野うさぎ」の語形式

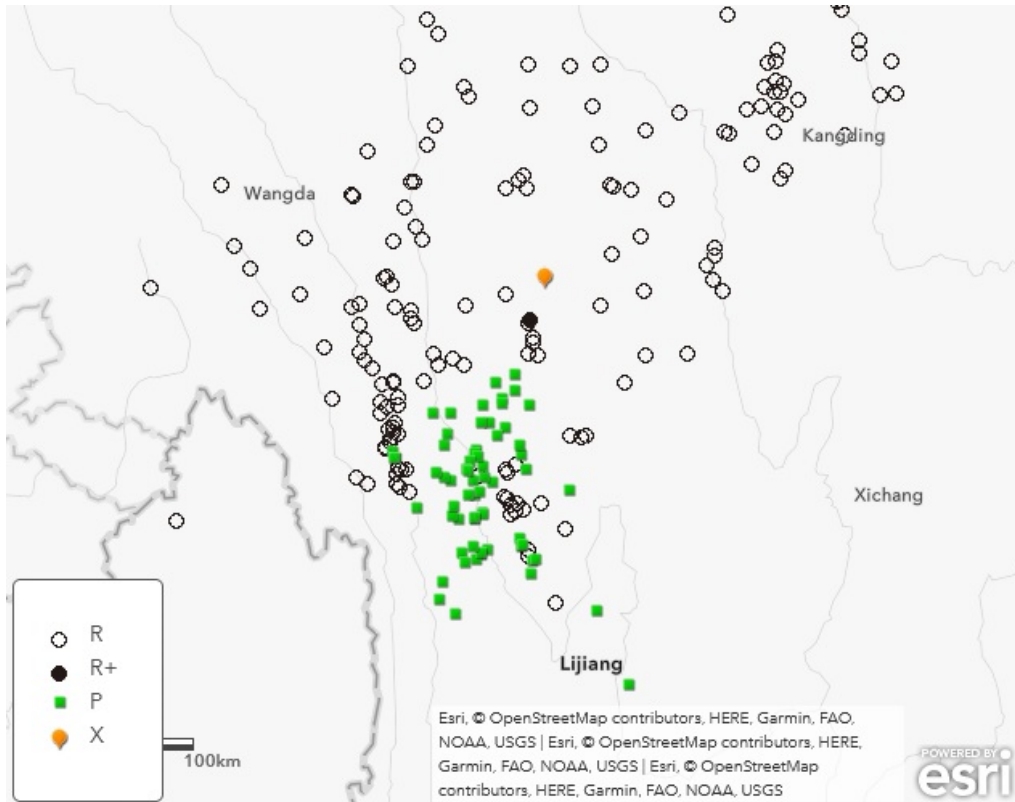


図2：「野うさぎ」の語形式（カム地域南部の拡大版）

図2を見ると、P類の分布する中に、複数のR類を用いる地点が囲まれていることがわかる。これは建塘鎮を中心とする地域である。鈴木(2017)、Suzuki (2018, 2022)などの地理言語学的研究が示すように、建塘鎮は雲南のチベット文化圏における政治・経済・宗教の中心地であり、音韻変化の特徴でABA分布を見せる事例の中心地、すなわちより新しい形式を生み出す地域にも対応することが分かっている。つまり、建塘鎮がR類を見せているとすれば、それが新しく用いられるようになった形式である可能性をまず検討する必要がある。

P類とR類はおそらく一定程度の地点で共存していたと考えることができる。なぜなら、チベット文語においてP類の形式が書かれることがなく、「野うさぎ」の形式は文脈によることなく常に *ri bong* とつづられ、かつこのつづり字が読書音として発音されるからである。すなわち、P類は完全な口語形式である一方、R類は文語の語彙であるが口語にも現れていた、という関係が存在するといえる。

すると、R類がP類の分布域に紛れ込んでいる状態は不自然であるとはいえない。しかし、そのR類の分布が建塘鎮周辺にかたまっているのは、R類の使用を支持するコミュニティーが、ABA分布でいうところのBすなわち新しい

形式として、R 類を周辺へと拡大させつつある状態を示していると考えられる。すなわち、文語に属する形式を口語形式へと拡張させるという改新が起きたと考える。

それでは、蔵文には語彙形式としては認められない P 類の起源はどこにあったと言えるだろうか。語形の分布を見る限り、雲南のチベット文化圏の中で、瀾滄江流域を除く地域に主たる連続する分布範囲があり、瀾滄江流域の数地点に P 類の分布域が拡張したと考えてよいだろう。この点的な分布を見せる地点は、他の語彙形式についても Suzuki (2018)でその関連性が指摘され、Suzuki (2022)では、方言区分が建塘鎮に分布する諸方言と同じ区分（香格里拉方言群）に入ることが提起された。このことも踏まえると、「野うさぎ」の語形の分布は、香格里拉方言群が東から西へ向かって拡大していった可能性を示唆する。

このように考えると、P 類は香格里拉方言群の諸方言について、本来の語形であった可能性が高いと言える。それが、建塘鎮を中心に文語形式に一致する R 類に置き換わってきたと見ることができる。

次に、建塘鎮付近において、移民の歴史が比較的明確に知られる地点の語形式に注目してみたい。すべて図 2 の南部に見える Lijiang の文字の周辺に分布する地点である。

まず、Lijiang の文字の北側に R 類が認められるが、これは虎跳峡鎮魯堆村である。このあたりは主に彝族やリス族が居住している（呉光範 2009:270-285）が、魯堆村にはチベット族がまとまって暮らしている。魯堆村のチベット族の祖先は、建塘鎮、小中甸郷から主に 2 世代前に移民してきたという口述史がある⁸。つまり、魯堆村の言語は建塘鎮付近の変種を反映しているといえ、2 世代前の建塘鎮の言語では、すでに R 類が用いられていたものと推定できる。

次に、Lijiang の文字の北側に見える P 類について考える。この地点は雪花村であり、香格里拉市格咱郷、東旺郷、および四川省甘孜州郷城県などの地域から主に 3～4 代前に移民してきたという口述史がある⁹。調査協力者によれば、雪花村の老年層の発音には個人間に大きな異なりが認められるとされるが、現在の中年層以下の世代では、発音は一定しつつあるという。格咱郷の多くは香格里拉方言群に属する変種を用いる地点であるが、東旺郷と郷城県は、郷城方言群という香格里拉方言群と意思疎通に困難を生じる程度の異なりをもつ別個の方言群に属する変種を用いる地域にある。調査協力者の観察は、異なる方言群を話す世代から、時代が下るにつれて共通語のような変種を生み出したことを意味し、その「共通語」の音声に関わる諸特徴は、香格里拉方言群建塘下位方言群の特徴を強く示している。「野うさぎ」の語形が P 類である点については、建塘方言群の中でも格咱郷では P 類を用い、東旺郷でも P 類を用い

⁸ 筆者の現地調査による(2015年)。

⁹ 筆者の現地調査による(2014年)。

る。このような地域の出身者が「野うさぎ」の P 類の語形を反映させたと考えられることができる。このことは、格咱郷や東旺郷でも現在から 3～4 代前の言語で「野うさぎ」の語形が P 類であったことも示唆する。

最後に、Lijiang の文字の南東側に見える P 類について考える。この地点は大安納西族郷であり、和旭東主編(2001:247)に記録され、また鈴木(2009)などの言語記述がある。この地点は民族分布の点でもチベット族居住地が飛び地を形成しており、彼らの来歴は 16 世紀にさかのぼる移住である。筆者の現地調査(2014 年)によると、原住地は現在のチベット自治区察隅県察瓦龍郷である説と芒康県鹽井郷である説の 2 通りがある。この 2 地点は図 2 の北西部にあたる。その言語特徴については、それぞれ鈴木(2020)と鈴木(2019)を参照すると、香格里拉地域の言語とは関連がないと考えられる。一方、鈴木(2009)や Suzuki(2018, 2022)では、大安方言は香格里拉方言群維西塔城下位方言群に属すると結論する。「野うさぎ」の語形式を見る限り、維西塔城下位方言群の特徴を反映していると言え、移民という史実が確かであるとしても、現在のチベット族の話す言語は途中で維西塔城下位方言群に置き換わったか、16 世紀では図 2 の北西部においても香格里拉方言群に近い言語が話されていたか、といった仮説を検証しなければならないと言えるだろう。ただし、この問題は本稿で扱う範囲を超えるものである。

3.2. ri bongの音対応に関する地図

ここでは、2.2 節に掲げた蔵文 ri bong の対応形式の分布を地図上に示す。2.1 節で言及した形式の中で、蔵文 ri bong 対応形式を含む形式も含む。表 3 は東チベット全体の言語地図、表 4 は表 3 の中のカム地域南部(雲南のチベット文化圏を中心とする)の拡大図である。なお、凡例中の分類は表 1 にあるものと同じである。

図 3、4 を見ると、蔵文のつづりに近い Ra1 類は地図南部に散発的に、かつ広範囲に認められる。なお、表 1 に示すように、Ra1 類は語形の第 2 音節初頭子音に /b, m, w/ など複数の形式をもつ。特に鼻音 /m/ が現れる点は、単なる特異な音変化として考えるほかに、そもそも蔵文 ri bong という形式と対応するものではないという見方もできる。一方、2.2 節にも述べたように、/w/ は両唇と軟口蓋における同時調音であるから、Ra2 類と調音位置上で連携していると考えられる。ただし、Ra1 類の /w/ から Ra2 類へ移行したと理解するのは、チベット系諸言語全体で見たとき、Ra1 類の分布の地理的偏りとデータの少なさ(Hill 2011)を説明できるものではないため、慎重であるべきである。

Ra2 類と Ra3 類は、図 3、4 にも反映させたように、互いに関連するものとして解釈した。Ra3 類にある口蓋垂音は、チベット系諸言語での位置づけに諸説ある(黄布凡 2012、鈴木 2014、鈴木、四郎翁姆 2021 など)。Ra3 類の分布を見ると、口蓋垂音を音素として体系に持つすべての変種が「野うさぎ」の

語形で口蓋垂音が現れているとは言えない。すなわち、音体系に口蓋垂音があるものが Ra3 類となっているわけではなく、そのうちの一部の変種に認められる。口蓋垂音と軟口蓋音の合流する過程については鈴木、四郎翁姆(2021)に例示されている。これを踏まえると、Ra2 類と Ra3 類の関係については、Ra3 類 > Ra2 類の順序を示唆するが、逆の順序も可能で、Ra2 類が先に存在し、各変種で並行例などからの類推により Ra3 類に変化した可能性もある。現状では、どちらの過程がより蓋然性が高いかを示すことはできない。

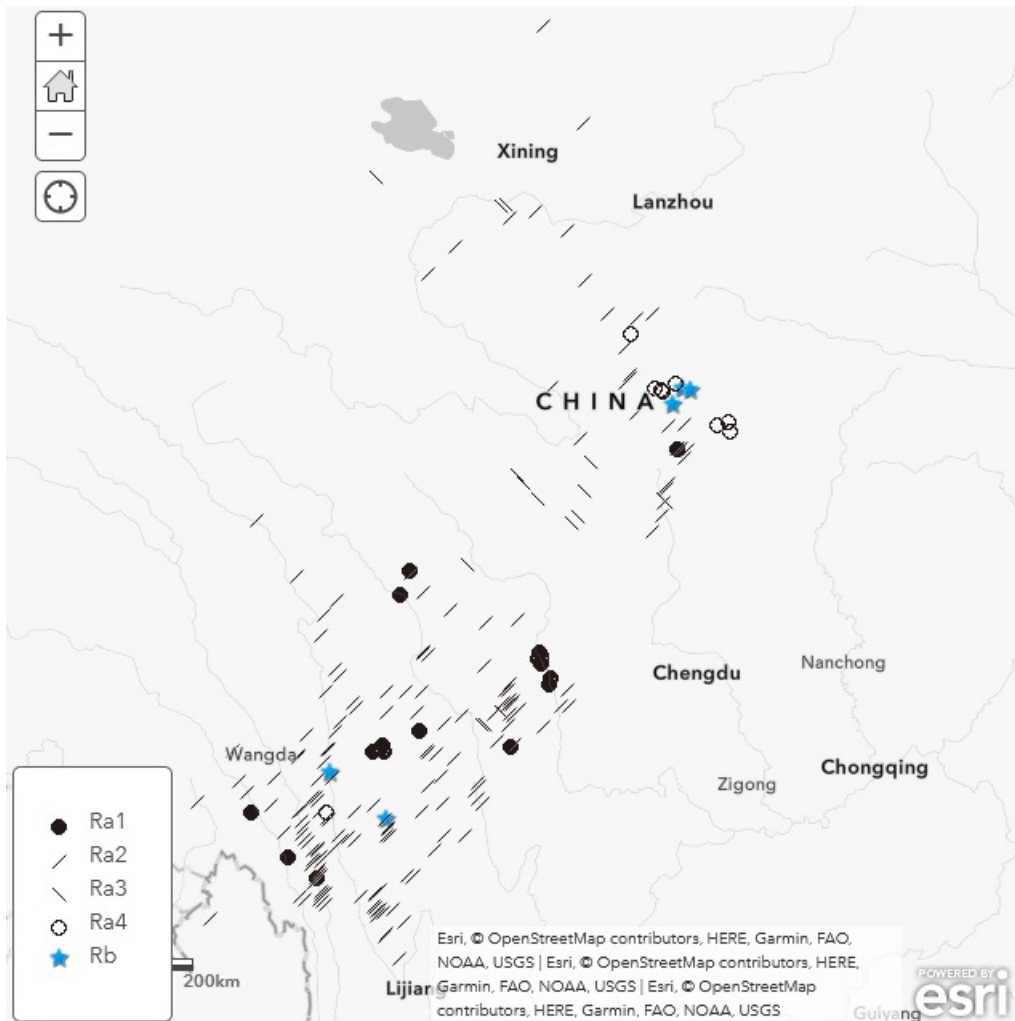


図3 : ri bongの音形式

Rb 類は1音節形式を用いる変種であるが、図3が示すように、地図の北東と南西に認められ、またその周辺に Ra4 類が分布する点に注目できる。Ra4 類

は、語形の第2音節初頭子音が声門摩擦音になる点に特徴づけられる。声門摩擦音が前後の母音に影響されて単音節化する過程は容易に理解できる。加えて、Rb類が現れる変種やその周辺に分布する変種には、音節縮約の実例も認められる(鈴木 2022, forthcoming)。このことから、Ra4類とRb類は互いにつながりがあると言え、語形の変遷として、Ra4類 > Rb類という順序とすることに問題はないだろう。

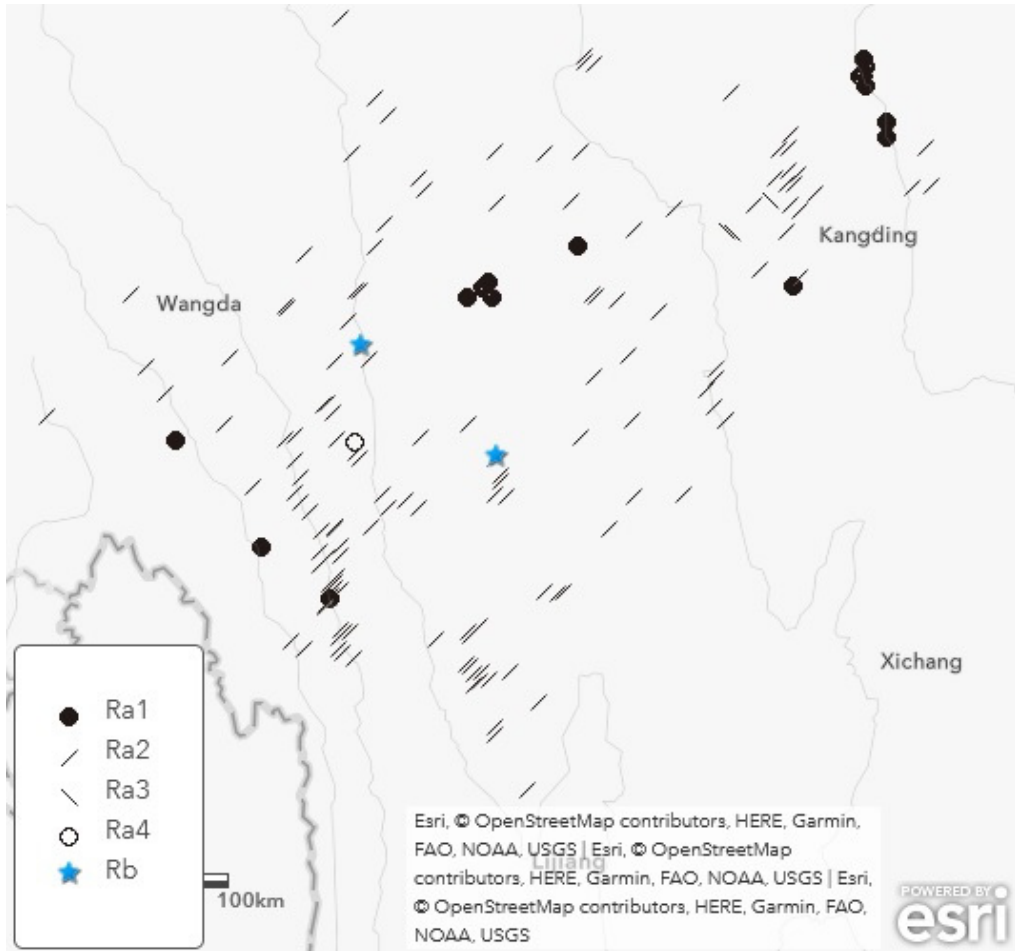


図4：ri bongの音形式（カム地域南部の拡大版）

以上、藏文 ri bong 対応形式を分類し、その分布について検討した。それを踏まえて、「野うさぎ」の語形式について疑問が生まれる。もし当該語が一般に説明されるように「山」と「ロバ」という形態素からなる複合語であるならば、Ra1類以外の音形式は、複合語であることを話者が意識しなくなって以降に音変化を起こしたと考えなければならない。bong は語幹であり、指小辞 bu

とは異なり、より独立性の高い、派生について生産的でない形態素である¹⁰。これが第2音節に現れ、かつ第1音節が開音節であったとしても、東部チベット系諸言語の多くの口語形式では、このbは閉鎖音を保ちうるといえる。加えて、対応形式に **bong** が含まれていると考えるのであれば、Ra2類で軟口蓋閉鎖音が現れる形式に対し、別途解説を与えなければならない。軟口蓋摩擦音であれば、指小辞と同様の初頭子音部の弱化として説明できるが、先行音節が開音節という条件で閉鎖音はならないからである。そこで1つ検討の可能性がある仮説は、そもそも軟口蓋摩擦音であったものが、過剰修正の結果、閉鎖音として認識されるようになったという変化である。

以上のような変化を仮説としてのみ提示するのは、「野うさぎ」の語形の形態素分析が母語話者にとって自明のことであるとはいえないからである。筆者の調査によれば、多くの市井の調査協力者は「野うさぎ」の語形について解説できず、蔵文の知識がある一部の協力者から「山のロバ」という含意があることの説明を得られた。これは、文字形式 **ri** と **bong** に基づいて「山」と「ロバ」という語義にその場で分析した可能性がある。同様のことは蔵文 **spang g.yag** の対応形式についても見られ、一般の話し手は「草原のヤク」と理解しているのではなく、やはり蔵文の知識のある人が分析して提示した程度である。それゆえ、蔵文を知らない話し手の間で独自の音変化が起こり、^hpu jā/といった蔵文との対応関係から若干不規則的な音形式が用いられるようになってきたとも考えることもできる。

本稿の言語地図を見るだけで「野うさぎ」の語源についてさらに考察を進めることは難しいが、もし音形それ自体を言語地図化した場合、語源についての考察を深めることができる可能性もある。これについては別稿にゆずる。

4. まとめ

本稿では、チベット文化圏東部におけるチベット系諸言語の「野うさぎ」の語形について分析した。まず、語形式としては、主に蔵文 **ri bong** 対応形式、**spang g.yag** 対応形式の2種類が用いられることを記述した。次に、前者について、音形式に注目し、より詳細な分類を行い、語形を特徴づけた。

以上の2点について言語地図を作成することで、特に蔵文 **spang g.yag** 対応形式の分布地域について詳細に検討し、その分布域に現れる蔵文 **ri bong** 対応形式について解釈を与えた。音形式の言語地図では、Ra2類とRa3類の関連及びRa4類とRb類の関連について解釈を与えた。

¹⁰ 蔵文で「ロバ」は通例 **bong bu** となる。

参考文献/References

- 鈴木博之 (2009) 「納西文化圏のチベット語・永勝県大安[Daan]方言の方言所属」『国立民族学博物館研究報告』34.1: 167–189. doi: <https://doi.org/10.15021/00003919>
- 鈴木博之 (2012) 「カムチベット語燕門・斯嘎[Sakar]方言の文法スケッチ」『地球研言語記述論集』4: 123–158. URI: <http://id.nii.ac.jp/1422/00000849/>
- 鈴木博之 (2014) 〈尼汝藏語的小舌輔音与其藏文对应規律〉《東方語言學》14: 1–12.
- 鈴木博之 (2017) 「音韻現象のABA分布をめぐる解釈の方法とその実際---チベット文化圏南東端のカムチベット語を例に」『言語記述論集』9: 43–64. URI: <http://id.nii.ac.jp/1422/00000911/>
- 鈴木博之 (2019) 〈瀾滄江流域盐井至巴迪段藏語土話中の音韻及词汇异同概況〉鈴木博之、倉部庆太、远藤光暁編《東部亞洲地理語言學論文集》, 14–42. 府中: 東京外國語大學アジア・アフリカ言語文化研究所. URI: https://publication.aa-ken.jp/sag_mono6_eastern_asian_2019.pdf
- 鈴木博之 (2020) 「カムチベット語察瓦龍[Tshawarong]方言の音声記述と語彙」『アジア・アフリカの言語と言語学』15: 105–137. doi: <https://doi.org/10.15026/99899>
- 鈴木博之 (2021) 「『フランス語・チベット語辞典』(1956)に記録される雲南チベット語の特定に向けて」『日本地理言語学会第3回大会予稿集』2–6. doi: <https://doi.org/10.5281/zenodo.5546119>
- 鈴木博之 (2022) 「下迭部チベット語阿夏[Azha]方言のチベット文語形式との音対応と語彙: 迭部県のチベット系諸言語の概観とともに」『言語記述論集』14: 65–114. doi: <http://id.nii.ac.jp/1422/00004411/>
- 鈴木博之 (forthcoming) 「カムチベット語索多西/改那西[Gadnagshod]方言の方言特徴と語彙」『アジア・アフリカの言語と言語学』17.
- 鈴木博之、四郎翁姆 (2021) 「カムチベット語塔公[Lhagang]方言における口蓋垂音」『言語記述論集』13: 1–12. URI: <http://id.nii.ac.jp/1422/00004143/>
- 華侃 主編 (2002) 《藏語安多方言词汇》兰州: 甘肃民族出版社.
- 黄布凡 (2012) 〈藏緬語的小舌音〉《語言學論叢》45: 157–174.
- 和旭東 主編 (2001) 《麗江地區民族志》昆明: 雲南民族出版社.
- 吳光範 (2009) 《迪慶・香格里拉旅遊風物志——沿着地名的線索》昆明: 雲南人民出版社.
- DTLF = Les Missionnaires Catholiques du Thibet (1899) *Dictionnaire tibétain-latin-français*. Hong Kong: Imprimerie de la Société des Missions Étrangères.
- Giraudeau, Pierre-Philippe et François L. N. Goré (1956) *Dictionnaire français-tibétain (Tibet oriental)*. Paris: Adrien-Maisonneuve.
- Hill, Nathan W. (2011) Alternances entre ɣ et b en tibétain ancien et dans les langues tibétaines modernes. *Revue d'Études Tibétaines* 20: 115–122. Online: http://himalaya.socanth.cam.ac.uk/collections/journals/ret/pdf/ret_20_04.pdf
- Kajihama, Ryoshun (2004) *Folk tales from Eastern Tibet*. Dharamsala: Library of Tibetan Works & Archives.
- de Nebesky-Wojkowitz, René (1956) *Oracles and demons of Tibet: The cult and iconography of the Tibetan protective deities*. 's-Gravenhage: Mouton.
- Shelton, A. L. (1925) *Tibetan folk tales*. New York: George H. Doran Company.
- Suzuki, Hiroyuki (2016) In defense of prepalatal non-fricative sounds and symbols: towards the Tibetan dialectology. *Researches in Asian Languages* 10: 99–125. URI: <http://id.nii.ac.jp/1085/00002195/>

- Suzuki, Hiroyuki (2018) *100 linguistic maps of the Swadesh word list of Tibetic languages from Yunnan*. Fuchu: Research Institute for Languages and Cultures of Asia and Africa. URI: https://publication.aa-ken.jp/sag_mono3_tibet_yunnan_2018.pdf
- Suzuki, Hiroyuki (2022) *Geolinguistics in the eastern Tibetosphere: An introduction*. Tokyo: Geolinguistic Society of Japan. doi: <https://doi.org/10.5281/zenodo.5989176>
- Suzuki, Hiroyuki and Sonam Wangmo (2018) Geolinguistic approach to the route of Tibetic loanwords in Lhagang Choyu. In Hiroyuki Suzuki and Mitsuaki Endo (eds.) *Papers from the Fourth International Conference of Asian Geolinguistics*, 115–126. Fuchu: Research Institute for Languages and Cultures of Asia and Africa. URI: https://publication.aa-ken.jp/papers_4IC_Asian_geolinguistics_2018.pdf
- Suzuki, Hiroyuki and Sonam Wangmo (2021) Two folktales in Lhagang Tibetan of Minyag Rabgang Khams: *The Sheep and the Wolf* and *The Hare and the Tiger*. *Tokyo University Linguistic Papers (eTULIP)* 43: e114–e142. doi: <https://doi.org/10.15083/0002002787>
- Tournadre, Nicolas and Hiroyuki Suzuki (2022) *The Tibetic languages: An introduction to the family of languages derived from Old Tibetan* (with the collaboration of Xavier Becker and Alain Brucelle for the cartography). Villejuif: LACITO Publications.
- Zou, Yuxia and Hiroyuki Suzuki (2022) Five folktales in Braghoglung Tibetan of Cone. *Himalayan Linguistics Archive* 11: 1–85.

出版情報

投稿受理日：2022年4月30日

採用決定日：2022年8月31日

Research article

The word for ‘mirror’ in French Analysis of *miroir* and *glace* in literary works, 1650–1799

ITO, Reiko

Tokyo University of Foreign Studies, Doctoral Program

SEIMIYA, Takamasa

Tokyo University of Foreign Studies, Doctoral Program

KAWAGUCHI, Yuji

Tokyo University of Foreign Studies

Abstract: Romance languages such as Portuguese, Galician, Spanish, and Italian inherited words descended from Latin SPECULUM ‘mirror’. However, in the Gallo-Romance languages, Occitan, and Catalan, words derived from the Latin deponent verb MIRARI are used: *miroir* in French, *miralh* in Occitan, and *mirall* in Catalan. In French, *glace*, descendant of Latin GLACIA, is also utilized. The words *miroir* ‘mirror’ and *glace* ‘ice’ appeared in literature from the 12th century. The meanings of *glace* were extended to include ‘mirror’ in 1825, although it is possible that *glace* was so used earlier. In this paper, we analyze the occurrences of *miroir* and *glace* meaning ‘mirror’ in literary works in the period 1650–1799. We clarify that *glace* was used to mean ‘mirror’ before 1825. We also point out that the usage frequency of *glace* meaning ‘mirror’ gradually increased between 1650 and 1799. Furthermore, in the first half of the 18th century, the use of *glace* to indicate the glass part of a mirror began to decline, while that of *glace* to mean ‘mirror’ started increasing. It was around this time that the metonymy of *glace* for *miroir* began in earnest.*

Keywords: miroir (mirror); glace (glass); diachronic linguistics; dialectology; geolinguistics

1. Introduction

Given that Narcissus was looking at his reflection in a body of water in Greek mythology, surfaces of water were probably the first mirrors. Following that, stones and metals whose surface had been polished were used as mirrors. In the 14th century, the technique of processing glass to reflect images was developed.

First, the present article will give an overview of the words for ‘mirror’ in the Romance languages. Second, we will focus on the situation in French. French has two words for ‘mirror’: *miroir* and *glace*. These two words have different etymons.

ITO, Reiko, Takamasa, SEIMIYA, and Yuji, KAWAGUCHI. 2022. The word for ‘mirror’ in French: Analysis of *miroir* and *glace* in literary works, 1650–1799. *Studies in Geolinguistics* 2: 53–65. doi: <https://doi.org/10.5281/zenodo.7121541>

* This research was supported by Research Center for Science Systems (RCSS) in Japan Society for the Promotion of Science (JSPS).

2. Overview of the situations in the Romance languages

2.1. ‘Mirror’ in the Romance languages

Based on the term for ‘mirror’, the Romance languages can be divided into three differentiated areas (see the linguistic landscape presented in Figure 1). Portuguese *espelho*, Galician *espello*, Spanish *espejo*, and Italian *specchio* all come from the Latin etymon SPECULUM ‘mirror’. On the contrary, French *miroir*, Occitan *miralh*, and Catalan *mirall* are nouns derived from the Latin deponent verb MIRARI ‘to watch’. The isolated Romanian form *ogîndă* probably comes from the Old Church Slavic *se oglindati* ‘to look around’ (cf. *DLR* under *ogîndă*).

Traditionally, geolinguistic theory teaches us that the distribution of the term ‘mirror’ in the Romance languages seems to bear witness to the distribution often referred to as “center versus periphery.” According to this theory, the reflexes of Latin SPECULUM, i.e., those in Portuguese, Spanish, Galician, and Italian, show the older layer, while those descending from the Latin MIRARI in French, Occitan, and Catalan are innovations that occurred in an ancient area of SPECULUM. However, this geolinguistic interpretation misses evidence of the pre-existence of forms derived from SPECULUM in French, Occitan, and Catalan during more or less ancient times. The form *espelh* in old Provençal would only be a loan from Italian, and the Piedmont dialect actually knows the loan *spetš* from the Italian *specchio* (cf. *FEW* XII, p. 162). In our current state of knowledge, we do not know if the innovation of MIRARI really occurred in an old field dominated by SPECULUM.

In summary, the greatest question regarding the name for ‘mirror’ in the Romance languages is the absence of forms derived from the etymon SPECULUM in the Gallo-Romance domain and that of Catalan. Let us now review the names for ‘mirror’ in *Atlas Linguistique Roman (ALiR)*.¹

¹ In the following lines, we will not analyze in detail all the variants found in *ALiR*.



Fig. 1 'Mirror' in the Romance languages

The forms [ʃp'aʎu] in Portuguese, [esp'eʎo] in Portuguese and Galician, and [esp'exo] in Castilian all presuppose as the etymon the Latin SPECULUM, although the fate of the consonantal sequence -CL- is not identical: On the western outskirts of the Iberian Peninsula, it is a palatal consonant /ʎ/, and at the center of the peninsula, -CL- changed to a fricative consonant /x/. The borrowing [isp'iʎua] in Basque teaches us that the mirror was not an autochthonous object in the Basque-speaking region but rather a cultural object introduced from the Romance languages.

Despite the voicing of -C- to [g] and the vowel change to [a], we can say that the Rhaeto-Romance form [ʃp'iagal]² still retains the skeleton of SPECULUM. Various forms such as one finds in and around Italy are all derived from SPECULUM: [spj'etʃ] in northern Italy, [sp'ekkjo] in Central Italy, [spicc'ali] in Southern Italy, [sp'idʒu] in Sicily and Sardinia, and [sp'eccu] in Corsica. A single exception would be the form [sp'era] in Central Italy, which comes from the etymon SPĒRA 'ball'.

The situation of Romanian is quite different. First, it can be confirmed that almost everywhere in the Romanian-speaking region, 'mirror' did not belong to the linguistic layer of Romance languages but rather to that of other languages. The most extended form [ogl'ində] in Dacoromanian is borrowed from Old Slavic. Aromanian and Megleno-Romanian, respectively, present the borrowings [ɣil'ie] from Greek *ύαλί*³ (Segura Da Cruz & Augusto 1993: 118) and [vid'elə] from Bulgarian *vidjalo*. Only one

² It is strictly a phonetic form of Romansh, one of the Rhaeto-Romance languages. The influence of Swiss German on the realization of this form cannot be ruled out.

³ In modern Greek, it is spelled γυαλί 'glass'.

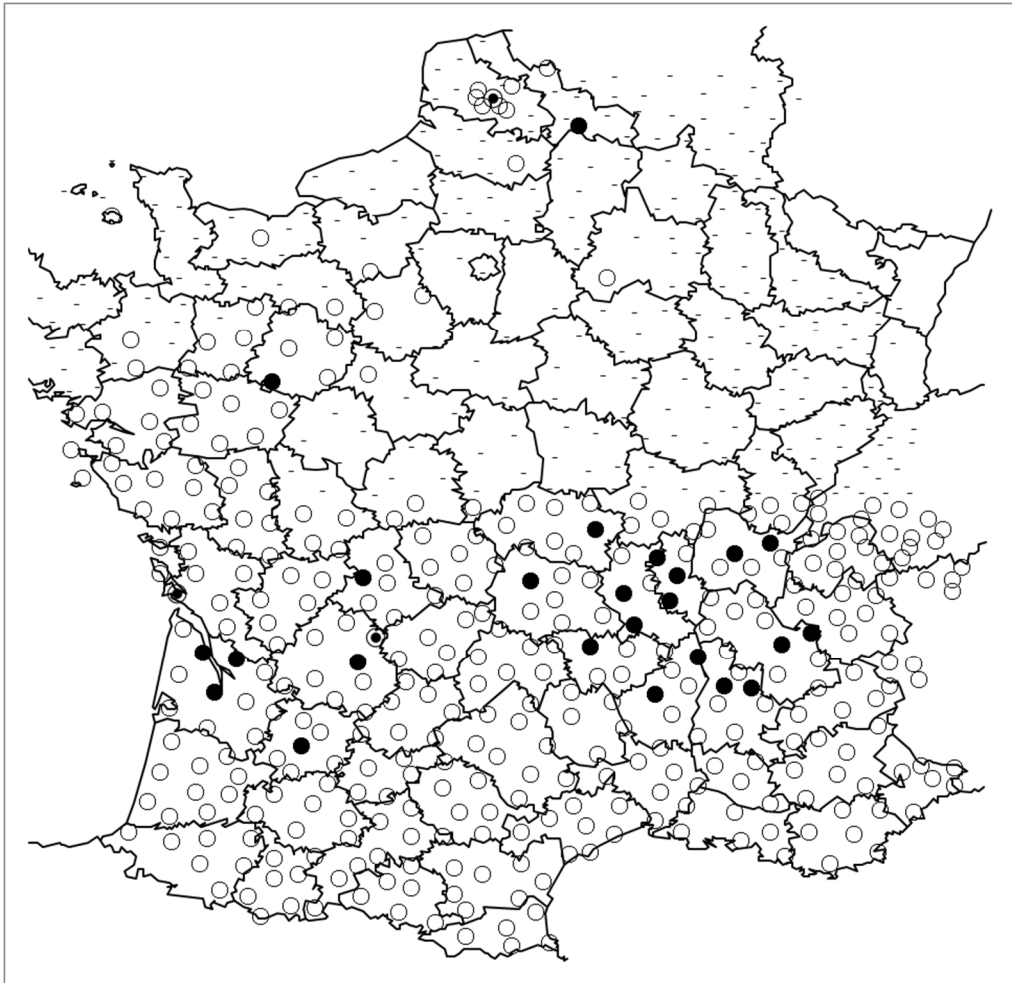
presumably autochthonous form [kətətʷ'are] is attested in the Maramureş region of Transylvania. This form probably comes from the Romanian *căuta-oare* ‘to capture the image’. In the following sections, we will tackle the difficult question of the word for ‘mirror’ in the Gallo-Romance and Catalan areas.

2.2. ‘Mirror’ in French

Figure 2 shows the situation of *miroir* ‘mirror’ in the Gallo-Romance, Occitan, and Catalan dialectal areas in the early 20th century. As mentioned above, descendants of Latin MIRARI, including *miroir* and *miralh*, spread in the southern area of ALF, where descendants of Latin GLACIA ‘ice’ such as *glace* are also present. As Edmont’s interviews did not cover the northern part of ALF, the situation of the Gallo-Romance area is unclear.

There are two possible reasons for this data loss. The first possibility is that even though Gilliéron had included ‘mirror’ in the questionnaire, Edmont forgot to ask about it. The second possibility is that Gilliéron excluded it, but Edmont added it during his inquiry. Gilliéron (1902: 4) describes the word choice of ALF as follows: isolated words, chosen from the popular repertoire, grouped together by similarity of meaning, and more particularly, designated to establish the phonetic laws of the languages.⁴ Edmont started the inquiry in Pas-de-Calais, the area where he was born, situated in the north of France, in 1897. Figure 2 depicts the presence of dialectal forms in this area. However, in *Matériaux de l’Atlas linguistique de la France: N°s 276–286*, one of his original notebooks, the forms for *miroir* are not registered, even though they are listed in the map. This fact suggests that Edmont had not elicited the forms for *miroir* in the first survey, but after having finished the survey, he returned to Pas-de-Calais and surveyed some additional words, including *miroir*. Thus, the first possibility is more realistic.

⁴ “... mots isolés, choisis dans le répertoire populaire, groupés par similitude de sens, et plus particulièrement désignés pour établir les lois phonétiques des parlers.”



○ < Lat. MIRARI ● < Lat. GLACIA □ = both ○ & ●

Fig. 2 ALF1648: *miroir*

According to Rey et al. (1992: 891 & 1251), the word *miroir* was first attested in the 12th century, and it acquired the meaning of “which offers the image of the ideal representation of things and people”⁵ in the 13th century, and to the present day it carries the meaning ‘mirror’. The word *glace* also first appeared in the 12th century. In about 1165, it already had the meanings ‘indifference’ and ‘non-tinned glass plate’. As time passed, the meaning of this word was extended to include ‘mirror’ in 1825. Mirrors are made by silvering or applying a film of silver to plate glass. This indicates that *glace*

⁵ “ce qui offre l’image des choses et des gens, représentation idéale de”

and *miroir* are inseparably linked, which explains why *glace* means ‘mirror’ in modern French.

Table 1 Meanings of *miroir* and *glace* in Rey et al. (1992: 891 & 1251)

century	<i>miroir</i>	<i>glace</i>
12 th	mirror	ice, indifference, non-tinned glass plate
13 th	+ ideal representation	
15 th		+ ice block
17 th		+ icing, window (of coach), ice cream
19 th		+ window, mirror, congealed food

3. Objective

Rey et al. (1992) indicated the first use of *glace* as ‘mirror’ in 1825, but such a use might have occurred much earlier. To clarify this, we analyzed the changes in the number of occurrences of *glace* as ‘mirror’ and *miroir* in literary works before 1800.

4. Methodology

4.1. Corpus

The French written corpus *Frantext 22.2 – Classique* was used as the corpus for this study. *Frantext Classique* includes works from 1650 to 1799, with a total of 1,106 works and 44,746,239 tokens. The words analyzed are *glace*⁶ and *miroir*⁷ in this corpus.

4.2. Method

For this study, we gathered all occurrences of *miroir* and *glace* that appear in the corpus.

We classified *glace* into four types by meaning:

- ice (and derived meanings: cold-hearted, indifference, ice cream)
- plate glass (and derived meaning: window of coach)
- mirror
- unidentifiable

⁶ Including the following spelling variations and plural forms: *glaces*, *glasse*, *glasses*, and *glacies*.

⁷ Including the following spelling variations and plural forms: *miroirs*, *mireur*, and *mireurs*.

A questionnaire was then administered regarding the classifications of *glace*. Four participants (three doctoral students in French studies and one former doctoral student, including Seimiya and Ito) identified the meaning of *glace* on a scale of 0 to 3 (0: *unidentifiable*, 1: *ice*, 2: *plate glass*, 3: *mirror*) for 20 examples of *glace*. Cronbach’s alpha for the questionnaire was $\alpha = .87$. We therefore deemed the questionnaire appropriate.

Examples (1) to (3) show the co-occurrence of *glace* and *miroir*. In these sentences, *glace* indicates ‘glass’, not ‘mirror’. For example, ‘*glace de miroir*’ in (3) is not ‘mirror of mirror’ but rather ‘glass of mirror’. In other words, the meaning of *glaces* here is ‘the material comprising the mirror’ or ‘a part of the mirror’.

(1)	dans	vn	brilliant	miroir
	PREP	INDEF.ART.M	shiny-ADJ.M	mirror-NOM.M
	dont	la	glace	fidelle
	REL.PRON	DEF.ART.F	glass-NOM.F	faithful-ADJ.F

‘in a shiny mirror whose faithful glass’
 (ASSOUCY Charles COYPEAU d’ 1653, “Poësies et lettres ... contenant
 diverses pièces héroïques, satiriques et burlesques,” p. 41, Frantext Q975)

(2)	du	grand	miroir	son
	of. DEF.ART.M	big-ADJ.M	mirror-NOM.M	POSS.ADJ.M
	poing	brise		la
	fist-NOM.M	break-IND.PRES.3SG		DEF.ART.F

glace
 glass-NOM.F

‘of the big mirror his fist breaks the glass’
 (DEMARETS DE SAINT-SORLIN Jean 1657, “Clovis ou la France
 chrestienne,” p. 124, Frantext Q748)

(3)	il	ressemble	à
	SUBJ.PERS.PRON.3SG	look like-IND.PRES.3SG	PREP

une	glace	de	miroir
INDEF.ART.F	glass-NOM.F	PREP	mirror-NOM.M

‘it looks like a glass of mirror’

(DU HALDE Jean-Baptiste 1735, “Description géographique, historique, chronologique, politique et physique de l’Empire de la Chine et de la Tartarie chinoise : t. 2,” p. 176, Frantext Q494)

Example (4) shows that *cette glace* refers to the front part of *un miroir*. In this case, *glace* means ‘mirror’ as a kind of metonymy.

(4) on	nomme	cela	
INDEF.PRON.3SG	call-IND.PRES.3SG	DEM.PRON.N	
un	...	<i>miroir</i>	ce
INDEF.ART.M		mirror-NOM.M	DEM.PRON.N
que	tu	vois	
REL.PRON	SUBJ.PERS.PRON.2SG	see- IND.PRES.2SG	
n’	est	que	
ADV.NEG	be-IND.PRES.3SG	CONJ	
ton	image	que	cette
POSS.ADJ.M	image-NOM.M	REL.PRON	ADJ.DEM.F
<i>glace</i>	réfléchit		
mirror-NOM.F	reflect-IND.PRES.3SG		

‘it is called a mirror ... what you see is only your image that this mirror reflects’
(DELISLE DE LA DREVETIÈRE Louis-François 1737, “Arlequin sauvage, Théâtre du XVIIIe siècle 1,” p. 485, Frantext S274)

In (5), *glace* appears without *miroir*. The use of the verb *se mirer* clearly indicates that *glace* here means ‘mirror’.

(5) je	la
SUBJ.PERS.PRON.1SG	PERS.PRON.F

meaning ‘mirror’ gradually increased from 3% in 1650–1674 to 33% in 1775–1799. On the other hand, that of *miroir* declined. Evidently, the use of *glace* for ‘mirror’ expanded during this 150-year period.

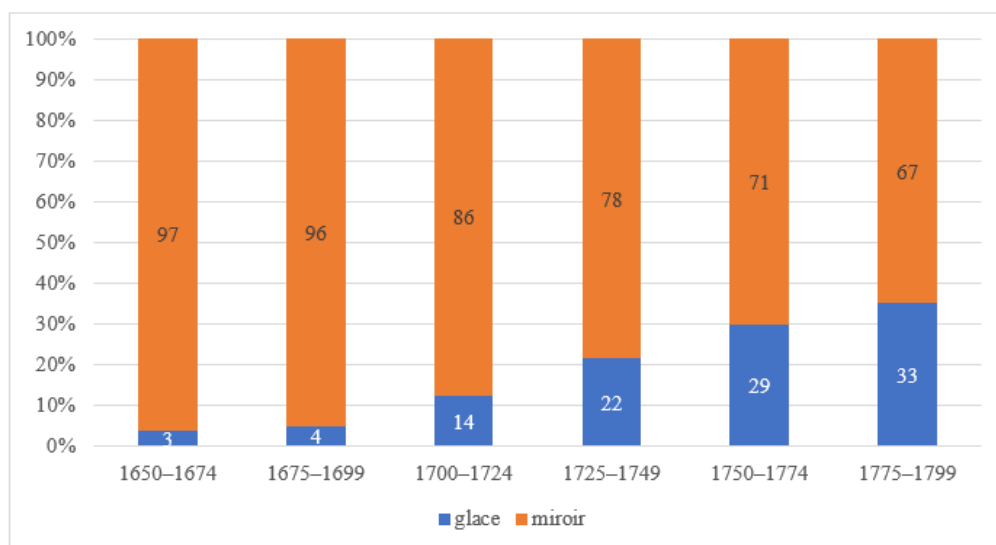


Fig. 3 Proportion of the relative frequencies (per 1 million words) for *miroir* and *glace* meaning ‘mirror’

5.2. *Glace* as part of a mirror and *glace* meaning ‘mirror’

We also compared the relative frequencies (per 1 million words) of *glace* as part of a mirror (i.e., *glace du miroir*, *miroir dont la glace*, etc.) and *glace* meaning ‘mirror’. Decimal places are rounded off (see Figure 4.)

The relative frequencies (per 1 million words) of both uses of *glace* were the same from 1650 to 1699. The incidence of *glace* as part of a mirror was maintained at the same level until 1724. Thereafter, such usage was very rare. On the other hand, the relative frequency of *glace* meaning ‘mirror’ has increased since 1700. In other words, *glace* could have two aspects - ‘part of a mirror’ and ‘mirror’ - until the end of the 17th century, but since then, *glace* on its own has increasingly meant ‘mirror’.

Examples such as *glace du miroir* imply that *glace* in reference to the constituent material might have been distinguished from *miroir*, a product with the function of reflecting images. Alternatively, *miroir*, referring to the entire object, could have been distinguished from a reference to the glass part, *glace*, of an object that reflects images. It would appear that the distinction between function (*miroir*) and material (*glace*) or the whole object (*miroir*) versus the glass part (*glace*) fell out of use, which made possible the metonymy of *glace* for *miroir*. It is unclear why this happened in the early

18th century. It could relate to the total amount of glass mirrors produced and their prevalence.

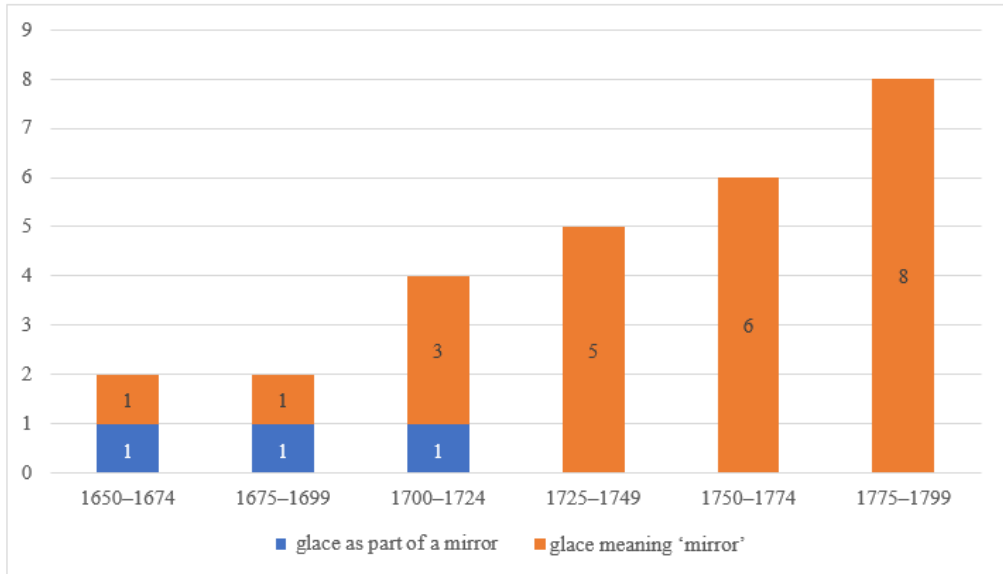


Fig. 4 Relative frequencies (per 1 million words) for *glace* as part of a mirror and *glace* meaning ‘mirror’

6. Discussion and conclusion

The purpose of this study was to determine whether *glace* was used in the sense of ‘mirror’ before 1825.

We have confirmed that the word *glace* was already being used to mean ‘mirror’ during the period 1650–1799 (see Table 3 and Figure 3). *Miroir* was always the predominant word for ‘mirror’, while usage of *glace* to mean ‘mirror’ gradually increased during the examined period.

Furthermore, until the early 18th century, *glace* was used to describe both ‘part of a mirror’ and ‘mirror’. Thereafter, the use of *glace* solely to indicate ‘mirror’ became widespread. This use constitutes a kind of metonymy wherein *glace* refers to *miroir*.

Abbreviations

<i>ALF</i>	<i>Atlas linguistique de la France</i>
<i>ALiR</i>	<i>Atlas Linguistique Roman</i>
<i>ATILF</i>	<i>Analyse et Traitement Informatique de la Langue Française</i>
<i>DLR</i>	<i>Dicționarul limbii române</i>

FEW *Französisches etymologisches Wörterbuch. Eine Darstellung des galloromanischen Sprachschatzes*

References

- Contini, Michel (1992) *L'Atlas Linguistique Roman* : état d'avancement du projet. In: Real Academia de la Lengua Vasca (ed.). *Actes du Congrès International de Dialectologie, Iker* 7: 339–356. Bilbo: Euskaltzaindia.
- Contini, Michel & Gaston Tuaillon (1993) Introduction. In: Michel Contini et Gaston Tuaillon (eds.). *Atlas Linguistique Roman (ALiR), Volume 1, Présentation*, 1–18. Roma: Istituto poligrafico e zecca dello stato/ Libreria dello stato.
- Gilliéron, Jules & Edmond Edmont (1897) *Matériaux de l'Atlas linguistique de la France*. <https://gallica.bnf.fr/ark:/12148/btv1b525056205.r=Mat%C3%A9riaux%20de%20l%27Atlas%20linguistique%20de%20la%20France%E3%80%80276?rk=21459:2> [accessed July 2022].
- Gilliéron, Jules & Edmond Edmont (1902) *Atlas linguistique de la France : Notice servant à l'intelligence des cartes*. Paris: Honoré Champion.
- Gilliéron, Jules & Edmond Edmont (1902–1910) *Atlas linguistique de la France*. Paris: Centre national de la recherche scientifique.
- Iordan, Iorgu, Alexandru Graur & Ion Conteanu (1965–) *Dicționarul limbii române*. București: Editura Academiei.
- Kawaguchi, Yuji [川口裕司] (2022) Romansushogo no kagami「ロマンス諸語の鏡」[Mirror in Romance languages]. *Renzokukōen Romansushogo no rekishi to gogen Romansushogo no goi* 『連続講演 ロマンス諸語の歴史と語源 ロマンス諸語の語彙』 [Series of lectures The History and Etymology of Romance Languages, The Vocabulary of Romance Languages]. Asahi Karuchā Sentā Shinjuku 2022/3/5 朝日カルチャーセンター新宿 2022/3/5[at Asahi Culture Center Shinjuku, on March 5, 2022].
- Kawaguchi, Yuji, Reiko Ito, Takamasa Seimiya and Kaho Okawara [川口裕司, 伊藤玲子, 清宮貴雅, 大河原香穂] (2021) *Furansu chirigengogaku to gengochizu tanjō kara genjō made* 「フランス地理言語学と言語地図 誕生から現状まで」 [French geolinguistics and linguistic atlases: From nascency to the status quo]. *Studies in Geolinguistics* 1: 151–168. <https://doi.org/10.5281/zenodo.5529208> [accessed August 2022].
- Le Dù, Jean, Yves Le Berre et Guylaine Brun-Trigaud (2005) *Lectures de l'Atlas linguistique de la France de Gilliéron et Edmont : du temps dans l'espace : essai d'interprétation des cartes de l'Atlas linguistique de la France de Jules Gilliéron et Edmond Edmont*. Paris: Editions du Comité des travaux historiques et scientifiques (CTHS).
- Matsubara, Hideichi [松原秀一] (1996) Kagami Le miroir 「鏡 Le miroir」 [The mirror]. *Furansu kotoba jiten* 『フランスことば事典』 [Encyclopedia of French language]: 90–94. Tokyo: Kōdansha.
- Rey, Alain, Marianne Tomi, Tristan Horde et Chantal Tanet (1992) *Dictionnaire historique de la langue française*. Paris: Le Robert.
- Segura Da Cruz, Maria Luísa & Celeste Augusto (1993) Les désignations romanes de MIROIR. In: Tuaillon Gaston & Michel Contini (eds.) *L'Atlas Linguistique Roman, Volume 1, Commentaires*. 115–126. Roma: Istituto poligrafico e zecca dello stato/ Libreria dello stato.
- von Wartburg, Walther (1922–2002) *Französisches etymologisches Wörterbuch. Eine Darstellung des galloromanischen Sprachschatzes*, 25 Bde.

URL

ATILF (1998-2022) *Frantext*. Nancy. URL: <https://www.frantext.fr/> (accessed 2022-07-25).

Publication history

Date accepted: 19 August 2022

論文

スペイン語高頻度語 «así», «mismo», «ahora」の 地理・社会的変異と歴史的变化の要因

上田博人
東京大学・名誉教授

Factors of the geographical and social variation and historical change of Spanish frequent words: «así», «mismo», «ahora»

UEDA, Hiroto
University of Tokyo, professor emeritus

Abstract: On this occasion provided by the Geolinguistic Society of Japan, I investigate the factors of the historical, geographical and socio-stylistic variation of Spanish words using the methods of Spanish philology and corpus linguistics. I analyze three words used with high frequency throughout history: *así* [adverb] ‘so’, *mismo* [adjective] ‘same’ and *ahora* [adverb] ‘now’. These words competed with the old variants, *ansí*, *mesmo* and *agora*, to arrive at the current unified forms.

First I read *Ingenioso hidalgo Don Quijote de la Mancha* (Miguel de Cervantes) published at the beginning of the 17th century for treating the social variation of these words. Second, I observe the chronological vicissitudes of these words from the 1200s to the 1700s in the large corpus and dictionaries. Next, I observe dimensions of the historical, geographical and socio-stylistic variables with the corpus of historical notarial documents and another of current conversations. In analyses, I use R and its packages together with the system developed by us that works on the internet: LYNEAL.

キーワード: スペイン語; 高頻度語; 地理的変異; 社会文体的変異; 歴史的变化

Keywords: Spanish; frequent words; geographical variation; socio-stylistic variation; historical change

1. はじめに

スペインのスペイン語研究は伝統的に文献学と言語学の分野で行われてきた。文献学は文学と言語の関係を重視し、文学作品の読解・研究を主な目的に掲げている。一方、言語学はとくに文学との関係を見ず、独立して言語の構造・機能・体系を追究する。両方の分野は共通して言語を対象とするので、学会は両方の研究発表を含み、研究者はどちらの接近法も可能とすることが多い。日本地理言語学会から投稿の依頼をいただいたこの機会に、スペイン文献学とコーパス言語学の方法を使って語形の地理・社会的変異と歴史的变化の要因を探ろうと思う。取り上げる語は歴史上高頻度で使用されてきた *así* [副]「そのように」、*mismo* [形]「同じ」、*ahora* [副]「今」である。これらは標準化された現代語形に至るまで、それぞれの古形 *ansí*, *mesmo*, *agora* と競合した。高頻度語を対象とする理由は統計的有意性を重視するためである。

出発点として 17 世紀初頭に発表された『才知ある郷士ドン・キホーテ・デ・ラ・マンチャ』(Miguel de Cervantes) を読み、そこに見られる語形の社会的変異を扱う。次に、大規模コーパスと各種辞書を使って 1200 年代から 1700 年代までの年代的推移を直線的に観察する。続けて歴史的な公証文書と現代の話し言葉のコーパスを使って年代・地域・文書の種類を多元的に分析し、現代語形に至った要因を考察する。分析手段として R (R Core Team 2021) と R の各種パッケージ (該当部分で引用)、および私たちが開発しウェブで機能する LYNEAL (*Letras y Números en Análisis Lingüísticos*) を使用する¹。

2. 資料

スペイン語の歴史は概略して初期 (700-1200)、中世 (1200-1500)、古典期 (1500-1700)、近代 (1700-) に区分される (Lapesa, 1981; Cano Aguilar, 1988; 山田, 1995, 1996; 寺崎, 2011)。スペイン文学史上最大の出版数を重ねた Miguel de Cervantes (1547-1616) の *Ingenioso hidalgo Don Quijote de la Mancha* 『才知ある郷士ドン・キホーテ・デ・ラ・マンチャ』(I. 1605; II. 1615: 以下では DQ とする) はスペイン語史の中で古典期中期に位置する。この時期のスペイン語では中世の言語的大変異・変化の余波がほぼ収束し、その大多数の語彙は、一部の表記法を改めれば現代スペイン語と変わらない。よって、現代スペイン語の中級コースを終えた日本の大学生でもほとんど読解できる (清水 1986: 17)。しかし、標準化された現代語形と異なり、当時の音韻・形態・語形のバリエーションが各所に見られることも事実である (Rico 2010: 1176-1177)。

¹ <https://lecture.ecc.u-tokyo.ac.jp/~cueda/lyneal>

語形のバリエーションとして特に *ahora - agora* (Frago 2015: 144), *así - ansí* (同: 142, 144), *mismo - mesmo* (同: 158) の頻度が高い²。

DQ の語彙を網羅的に扱った先行研究として、Cejador / Frauca (1906), Echeverría / Reyes (1931), Fernández Gómez (1962), Caballero (1970) などが挙げられるが、その中ですべての単語を独力で数え上げた Fernández Gómez (1962: 1099-1136) は *ahora* (頻度: 512), *así* (1004), *mismo* (626) を載せている。しかし、上に挙げた諸研究には語形の変異の情報はない。筆者が参加している共同研究では、全語彙の代表形 (レンマ)、品詞、変化形、異形の頻度を文脈の中で記録し、登場人物との相関を追究している。以下ではその研究資料を利用する。

17世紀初頭に出版された DQ は当時の印刷の規範に従っており、中世の手稿文献に多く見られる表記の揺れは少ない³。しかし、その比較的安定した表記の中で異形が見られるものがある。異形が出現する条件として、登場人物の属性 (性・年齢・社会層など) が考えられる。DQ の登場人物が使用する語には、人物が属する社会層・発話場面の状況による文体・出身地の地方差が生き生きと描かれている (Hatzfeld 1972: 9194; 牛島1989: 246)。しかし、その語形がそのまま同じ条件で当時使用されていた、と考えることはできない。よって、語彙の通時的变化を俯瞰するためには、作品をその歴史的な文脈の中で位置づけなければならない (Medina Morales, 2005)。本研究では、DQ で観察した3語の語形の歴史的推移をラテン語・中世・古典期・近代・現代の各種資料を通じて観察する。

語彙の歴史的変遷を辿る資料として各時代に発行された辞書が利用できる。その中でも Real Academia Española (1726-1739) は必須であり、さらに、それ以前の多くの辞書 (総数145冊) に掲載された語を収録した Nieto Jiménez / Alvar Ezquerro (2007) が優れている。

辞書資料は中世の末期 (1490) まで存在しなかったため、中世を含めた語形の変化を知るためには、幅広く収集された文献資料を使用しなければならない。そのためには Real Academia Española が公開している『スペイン語歴

² Cervantes と同時代に発行された辞書 Cobarrubias (1611 / 1977) は *agora, assí, mesmo* を代表形として掲げ、*mesmo* の項に *mismo* の用例を載せている。

³ このような変異 (頻度分布) は作者 Cervantes 自身に帰属するものなのか、または一定の規範を示した印刷業者の手が入ったものなのか判断できない。Rico (2010: 1176) によれば、印刷業者は作者の独自の書き方よりも当時の一般的な表記法を採用したようである。DQ (II, cap. 62) には、Quijote がバルセロナの印刷所を訪れ、校正作業を見る場面が描かれている。その場面には著者 (イタリア語の翻訳者) がいたが、とくに校正には関わっていた様子は見えない。一般に著者は表記には拘らず、一切を印刷業者に任せたようである。Rico (ibid.) によれば Cervantes の自筆では 'b' を使って *Cerbantes* と署名したが、印刷物ではすべて 'v' (*Cervantes*) で記された、と述べている。言語研究資料として文学作品を扱うとき、それが作者の特徴を示すものであれ、それとも当時の規範を示すものであれ、どちらにしても当時の言語状態の一端を示している、と考えてもよいだろう。このように文学作品の中で見た語形のバリエーションは17世紀初頭の言語状態の一端を示すものであるが、当時の共時態を代表するものである、とは言い難い。同時代の辞書の記述の代表性も保証されていない。

史辞典コーパス』(*Corpus del Diccionario Histórico de la Lengua Española: CDH*, 2013)が有用である⁴。本研究ではこれを利用する。

言語の歴史をコーパスを使って観察するとき、歴史的な変化だけでなく地理的な変異にも留意しなければならない。そのために、各地に保存された公証文書(王室発行文書・行政文書・裁判記録・売買契約書・私文書など)の多くには日付と発行地が記載されているので、それらを変数としてデジタル化した CODEA (*Corpus de documentos Españoles Anteriores a 1800*) と、現在の半島各地の老齢者との会話を転写した COSER (*Corpus Oral y Sonoro del Español Rural*) を使用する⁵。

3. 分析

3.1. «*así*」〔副〕「そのように」

3.1.1. 社会

はじめに DQ のテキストを分析する⁶。ここでは変異語形と登場人物をクロス集計し、その絶対頻度と正規化頻度⁷と共起係数(→【補説】)を計算した。絶対頻度は使用頻度を示すが、それぞれの登場人物の総使用語数が異なるので比較できない。正規化頻度は頻度の大きさを比較できるが、[0, 1]の範囲で標準化されていないので同じスケールで数値を評価することができない。そこで、共起係数を用いて正規化頻度を標準化した。

DQ における *ansí* と *así* の出現数を登場人物の性(男性/女性)・年齢(若者/中高年)・社会層(上流/平民)でクロス集計すると次の結果となった⁸。

表1 : *ansí* / *así*: 性 (絶対頻度/正規化頻度/共起係数)

語形	男性	女性	男性	女性	男性	女性
<i>ansí</i>	56	4	12.6	9.9	0.108	0.101
<i>así</i>	876	66	196.4	163.8	0.69	0.614

⁴ <https://apps.rae.es/CNDHE/view/inicioExterno.view>

⁵ <https://lecture.ecc.u-tokyo.ac.jp/~cueda/lyneal/codea.htm>

<https://lecture.ecc.u-tokyo.ac.jp/~cueda/lyneal/coser.htm>

⁶ テキストとして原典の語形を尊重した Sevilla Arroyo (2005) を使用する。

⁷ 絶対頻度を登場人物の総使用語数で割り適当な乗数を掛けて求める。本研究では乗数を絶対頻度と桁数が一致するように設定した。この場合は 10^5 である。「 5 」は冪数を示し 10^5 は100,000を表す。それぞれの登場人物のグループの総使用語数は次のとおりである：男性: 445,992, 女性: 40,301; 若者: 5,772, 中高年: 480,521; 上流: 377,602, 平民: 108,691。

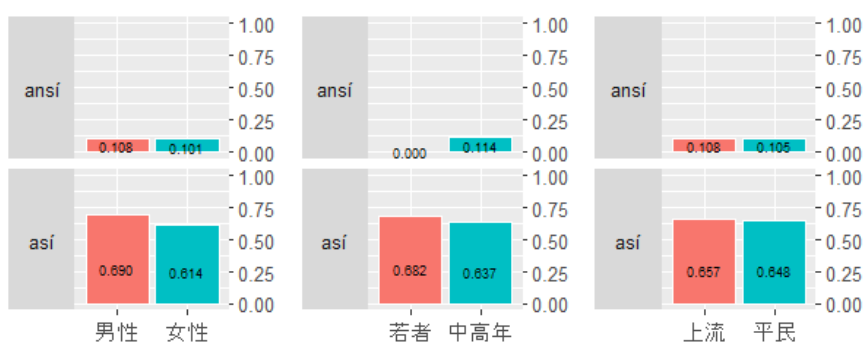
⁸ R のパッケージである ggplot2 を使って以下の棒グラフと折れ線グラフを作成した。

表2 : *ansí* / *así*: 年齢 (絶対頻度/正規化頻度/共起係数)

語形	若者	中高年	若者	中高年	若者	中高年
<i>ansí</i>	0	60	0.0	12.5	0.000	0.114
<i>así</i>	12	930	207.9	193.5	0.682	0.637

表3 : *ansí* / *así*: 社会層 (絶対頻度/正規化頻度/共起係数)

語形	上流	平民	上流	平民	上流	平民
<i>ansí</i>	47	13	12.4	12	0.108	0.105
<i>así</i>	735	207	194.6	190.4	0.657	0.648

図1 : *ansí* / *así* (共起係数)

このように、ほとんど全体的に性・年齢・社会層の区別なく圧倒的多数が *así* を使用している。一方、*ansí* の頻度は低いが、絶対頻度を見ると無視することはできない。以下では、この作品で *ansí* の頻度が小さかった理由を探る。

3.1.2. 年代

次のリストは Nieto Jiménez / Alvar Ezquerro (2007) と Real Academia Española (1726-1739) に収録された *así* / *ansí* の語形の歴史的分布を示す。

así : (発行年) 1490, 1492, 1495, 1505, 1520, 1526, 1534, 1551, 1554, 1562, 1568, 1570, 1580, 1582, 1587, 1591, 1599, 1601, 1604, 1611, 1617, 1636, 1639, 1642, 1645, 1660, 1661, 1666, 1670, 1679, 1693, 1721, 1723, 1726.

ansí : 1562, 1587, 1599, 1601, 1604, 1607, 1609, 1617, 1620, 1621, 1636, 1639, 1670, 1693, 1705, 1706, 1726.

このように辞書資料では *ansí* の初出は1562年である。Corominas / Pascual (1980: s.v. *así*) は中世・古典期に *ansí* が高頻度で用いられた、と記述している。一方、次の『スペイン語歴史辞典コーパス』(CDH) の検索結果を見ると、中世 (1064-1500) よりも古典期 (1501-1700) で *ansí* の出現が際立っている。

しかし、古典期においても *así* の恒常的な優勢が観察され、大きな言語変化は見られない。その中で *ansi* はたしかに一時期使用されたが、*así* (*assi*) を凌駕することはなかった。

表4 : *ansi* / *así*: CDH (RAE)

絶対頻度	<i>ansi</i> / <i>así</i>		正規化頻度	<i>ansi</i> / <i>así</i>	
	<i>ansi</i>	<i>así</i>		<i>ansi</i>	<i>así</i>
1064-1500	1,687	90,926	1064-1500	43.09	2322.49
1501-1700	29,307	161,277	1501-1700	339.03	1865.71
1701-1800	10	9,409	1701-1800	0.62	1512.15
1801-1900	474	49,703	1801-1900	10.7	1122.26
1901-2005	375	162,066	1901-2005	2.12	917.28

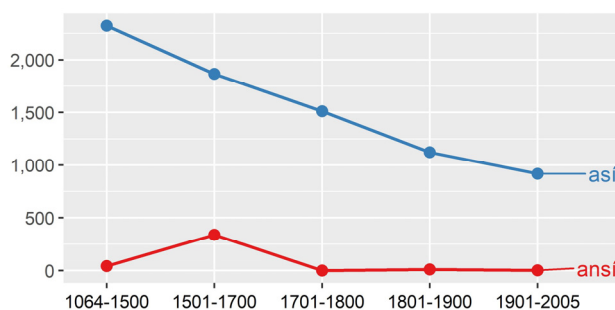


図2 : *ansi* / *así*: 正規化頻度 CDH (RAE)

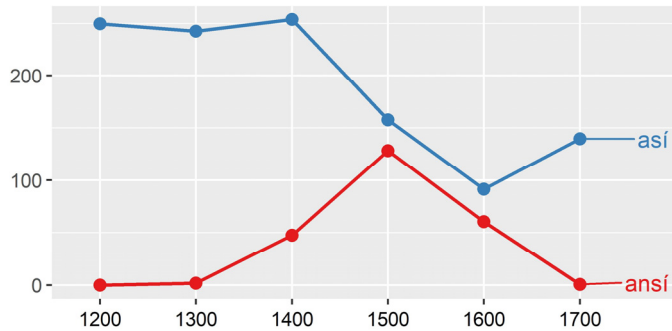
3.1.3. 歴史・地理・社会

以下では歴史・地理・社会の観点から *ansi* / *así* の変異を観察するために CODEA と COSER のコーパス資料を使用する。はじめに CODEA の資料を使って歴史的变化を見る。次の表の左側はそれぞれの絶対頻度を示す。右側は乗数を 10^5 にしたときの正規化頻度である。グラフは正規化頻度で作成した⁹。

表5 : *ansi* - *así*. 絶対頻度・正規化頻度. CODEA.

絶対頻度	<i>ansi</i> / <i>así</i>		正規化頻度	<i>ansi</i> / <i>así</i>	
	<i>ansi</i>	<i>así</i>		<i>ansi</i>	<i>así</i>
1200	0	334	1200	0	249.8
1300	4	509	1300	1.9	242.6
1400	131	700	1400	47.5	253.9
1500	372	457	1500	128.3	157.6
1600	108	163	1600	60.6	91.4
1700	1	190	1700	0.7	139.5

⁹ 年代: 総語数 = 1200: 133,693, 1300: 209,804, 1400: 275,709, 1500: 289,897, 1600: 178,337, 1700: 136,183.

図3 : *ansi* - *asi*. 年代. CODEA (正規化頻度, 10⁵)

先の CDH の検索結果と比べると、どちらも *ansi* が1500年代に上昇していることが確認できる。

次の表と地図は *ansi* - *asi* の年代と地域のクロス集計の対照値 (→【補説】) を示している¹⁰。対照値の負値は *ansi* の勢力を示し、正値は *asi* の勢力を示す。

表6 : *ansi* - *asi*. CODEA. 対照値

N.	1200	1300	1400	1500	1600	1700
1.AS	0.000	0.000	1.000	0.000	0.000	0.000
2.PV	0.000	0.000	1.000	0.895	0.000	1.000
3.GA	0.000	0.000	0.000	0.000	0.000	1.000
4.NA	1.000	1.000	0.000	0.000	0.000	0.000
5.LE	1.000	0.971	0.585	0.051	0.500	0.000
6.CV	1.000	0.962	0.361	0.162	0.158	1.000
7.CA	0.000	0.000	0.000	0.000	0.000	1.000
8.AR	1.000	1.000	1.000	0.650	0.000	0.000
9.CN	1.000	1.000	0.701	-0.281	0.036	0.966
10.VA	0.000	0.000	0.000	0.000	1.000	1.000
11.EX	1.000	1.000	0.319	-1.000	0.167	0.000
12.AN	0.000	1.000	1.000	0.500	-0.217	1.000

¹⁰ 地方名 = AN: Andalucía, AR: Aragón, AS: Asturias, CA: Cataluña, CN: Castilla la Nueva, CV: Castilla la Vieja, EX: Extremadura, GA: Galicia, LE: León, MU: Murcia, NA: Navarra, PV: País Vasco, VA: Valencia. 地図は Leaflet を導入したLYNEALの出力である。

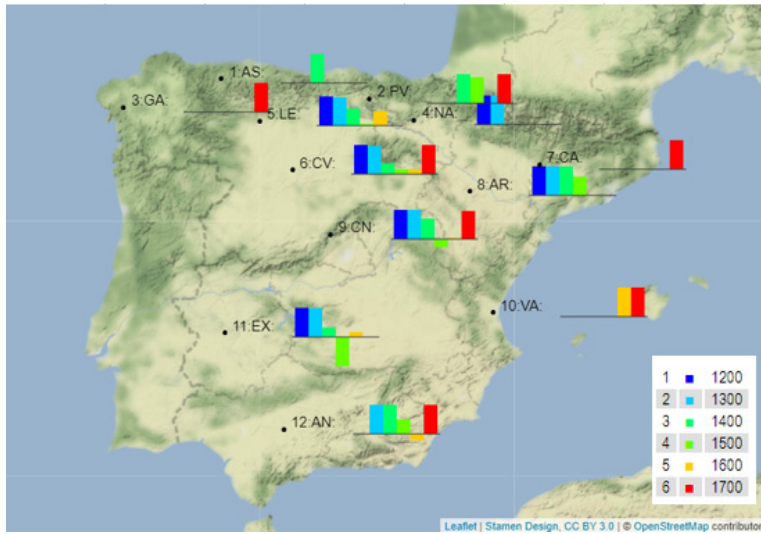


図4: *ansí* - *así*. CODEA. 対照値

これを見ると *ansí* は1500年代の 9.CN と 11.EX, および1600年代の 12.AN に限られていることがわかる。

次に王室文書と私文書を区別し、両者の社会・文体差を歴史を通して観察する。王室文書は王室の言語の規範を示し、私文書は民衆の言語の実態を示している。

表7: *ansí* - *así*. 王室文書・私文書. 正規化頻度. CODEA.

<i>ansí</i>	王室文書	私文書	<i>así</i>	王室文書	私文書
1200	0.0	0.0	1200	24.2	26.8
1300	0.2	0.3	1300	25.8	21.3
1400	2.3	5.2	1400	30.6	19.5
1500	9.2	11.5	1500	19.1	20.3
1600	2.4	5.0	1600	10.6	7.9
1700	0.0	0.1	1700	0.0	13.4

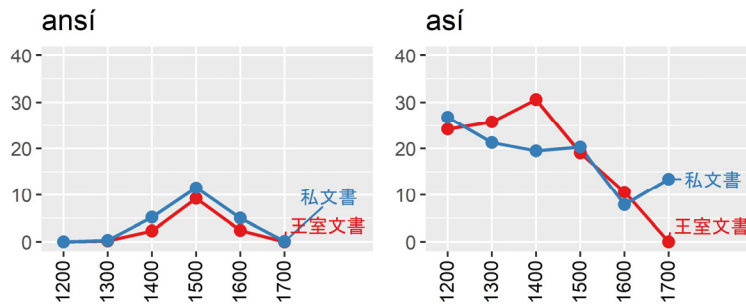


図5: *ansí* - *así*. 王室文書/ 私文書 CODEA. (正規化頻度, 10^5).

このように特定の年代・地域に現れる *ansí* は社会・文体の特徴はなく、王室文書と私文書の大きな違いは検出されなかった。

現代のスペイン語の状況を見るために COSER コーパスを検索すると、半島中北部の Burgos に1例だけ見つかった。よって現代語で *ansí* は皆無に近いと言える。

3.1.4. 考察

Rodríguez Molina (2015: 1053) によれば *así* の語源はラテン語形 AD SIC 「このように」である。一方、Corominas / Pascual (1980: s.v. *así*) は、*así* の語頭の *a-* は副詞に見られる単なる延長であった、という。それに加えて、SIC > [副詞] *sí* 「はい (肯定の返事)」や接続詞 SI > *si* 「もし～ならば」と語形が同じになるので *a-* によって同音衝突を避けた可能性がある。また、一般に語形が短いときは補強のために延長されることがあったことも要因の一つと見てよいだろう。この *así* は当初から全時代を通じて圧倒的多数であった。

一方、*ansí* については、筆者は1300年代に無意味な *-n-* が音節末に生じた結果と考える。類例として、*mucho* > *muncho* [形] 「多くの」、*estonces* > *entonces* [副] 「そのとき」、*desde* > *dende* [前] 「…から」などがある¹¹。他にも LOCUSTA > *langosta* [名] 「ロブスター」、MACULA > *mancha* [名] 「汚れ」、HIBERNU(S) > *invierno* [名] 「冬」、*maçana* > *manzana* [名] 「リンゴ」などの変化に「理由不明の *-n-*」(Menéndez Pidal, 1968: 189)、「偶然の *-n-*」(Rodríguez Molina, 2015: 1055) が見られる¹²。このような *-n-* は *ansí* だけでなく *asín*, *asina*, *ansín*, *ansina* のように位置が変わったり繰り返されたりすることもある。音節末・語末の *-n-* は頻度が高いので¹³、意味はなく偶然に付加された要素であったと思われる。また、音節末子音の情報量は小さいので音声変化が起こりやすい(上田, 1975)。一方、Corominas / Pascual (同) はこの *-n-* に前置詞 *en* 「～の中に」の影響を認めるが、両者の意味関係が不明である。

Alvar / Pottier (1983: 339) は *ansí* が *asín*, *ansina*, *asina* などと共に俗語 (vulgar) であった、と述べている。また、1500, 1600年代の書簡や日記などの私文書を調べた Blas (2021) は *ansí* が低社会層で使われていた、と述べている。逆に、Del Barrio de la Rosa (2017: 254) は 1581-1620 の間 *ansí* は *así* よりも社会的に上位の語であった、という。古典期の文学作品を分析した Medina Morales (2005: 252-253) も、該当する時期で *ansí* が標準形であった、

¹¹ *estonces* > *entonces*, *desde* > *dende* は16世紀の文人 Juan de Valdés (1535; ed. Lope Blanch 1969: 101) が挙げている例である。彼は *assi* > *ansi* も含めて、この *-n-* よりも *-s-* のほうがよい、と述べている。

¹² 中尾(1996: 150)は〔英〕*messenger* (< *message*), *nightingale* (< ME. *nihtegale*) の例を載せている。

¹³ 1400年代の Castilla la Nueva 地方で検索された語形の語中音節末音素の頻度は次のようであった：n : 4632 (43.3%), s : 2424 (22.7%), r : 2047 (19.1%), l : 865 (8.1%), b : 273 (2.6%), k : 187 (1.7%), g : 142 (1.3%), z : 35 (0.3%), m : 28 (0.3%), p : 25 (0.2%), d : 22 (0.2%), t : 6 (0.1%), x : 4 (0.0%), ll : 2 (0.0%), rr : 1 (0.0%)。このように語中の音節末子音の中で 'n' の頻度が最大であった。

と述べている。筆者は、CODEA の資料によれば 1400, 1500, 1600年代で私文書とともに王室文書でも使われていたので中世・近代において両者に社会・文体差はなかった、と考える (図5)。Frago (2015: 142, 145) は DQ の *ansi* は古形であり農民の言葉であった、と述べているが、登場人物の会話には *ansi* - *asi* の間に社会層の区別はない (図1)。社会・文体差が生まれたのは1700年代以降であり、その理由は *ansi* が低頻度で地域的に限定されていたため低く評価され、広域・高頻度の標準形 *asi* の威信が高められたことによる、と思われる (Rodríguez Molina 2015: 1051)。

よって、スペイン語史上 *asi* > *ansi* > *asi* という変化はなく¹⁴、常に *asi* が優勢であった。図式的に示せば両者の関係は次のようになる (A: *asi*, B: *ansi*, A:B は A が B よりも優勢であったことを示す)。

A → A:B → A

ansi は、-n- の挿入が有標であり、その地理的分布が限定的であり、使用頻度が比較的低かったため、1500年代後半に移動した新首都 Madrid で標準形に成り得なかったのであろう。

3.2. «mismo» [形] 「同じ」

3.2.1. 社会

mesmo - *mismo* の変異と DQ に登場する人物の属性 (性・年齢・社会層) の関係を見る。

表8: *mesmo* / *mismo*: 性 (絶対頻度/正規化頻度/共起係数)

語形	男性	女性	男性	女性	男性	女性
<i>mesmo</i>	318	24	71.3	59.6	0.549	0.467
<i>mismo</i>	256	26	57.4	64.5	0.458	0.525

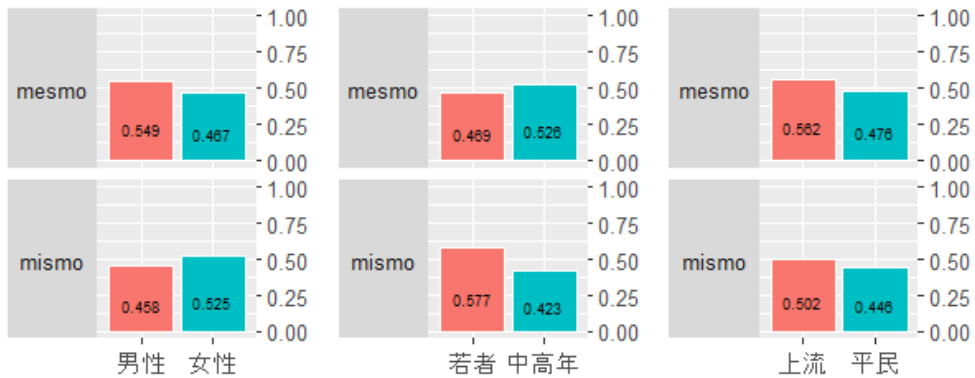
表9: *mesmo* / *mismo*: 年齢 (絶対頻度/正規化頻度/共起係数)

語形	若者	中高年	若者	中高年	若者	中高年
<i>mesmo</i>	4	338	69.3	70.3	0.469	0.526
<i>mismo</i>	5	277	86.6	57.6	0.577	0.423

表10: *mesmo* / *mismo*: 社会層 (絶対頻度/正規化頻度/共起係数)

語形	上流	平民	上流	平民	上流	平民
<i>mesmo</i>	282	60	74.7	55.2	0.562	0.476
<i>mismo</i>	231	51	61.2	46.9	0.502	0.446

¹⁴ Medina Morales (同) は *ansi* > *asi* の過程で語中音 ('n) の脱落を想定しているが、筆者は逆に *asi* > *ansi* の過程で語中音 ('n) の挿入を想定する。

図6: *mesmo* / *mismo*: 性・年齢・社会層 (共起係数)

このように、*mesmo* は男性・中高年・上流に多く、*mismo* は女性・若者に多いが、差はあまり大きくはない。よって、DQにおける両語形はほとんど自由変異に近い状態であった可能性がある。

3.2.2. 年代

Nieto Jiménez / Alvar Ezquerro (2007) と Real Academia Española (1726-1739) に収録された *mesmo* / *mismo* の語形の歴史的分布を見よう。

mesmo: 1490, 1492, 1495, 1499, 1505, 1551, 1562, 1565, 1570, 1591, 1599, 1601, 1609, 1611, 1617, 1621, 1634, 1636, 1639, 1642, 1661, 1670, 1705, 1706, 1721, 1726.

mismo: 1587, 1590, 1599, 1601, 1604, 1609, 1617, 1620, 1621, 1636, 1639, 1661, 1670, 1705, 1706, 1721, 1723, 1725, 1726.

この年代の分布を見ると、たしかに *mesmo* のほうが古く、*mismo* が後発であるように見える。しかし、17世紀にどちらが優勢であったのかは判然としない。また、中世には辞書資料がなかったために1490年以前の状況は不明である。

一方、Real Academia Española の『スペイン語歴史辞典コーパス』(CDH) を検索すると、次に示すように *mesmo* > *mismo* の変化はすでに1500年代にその趨勢が決定していたことがわかる。

表11: *mesmo* / *mismo*: 絶対頻度・正規化頻度 (RAE: CDH)

絶対頻度	mesmo	mismo	正規化頻度	mesmo	mismo
1064-1500	12,861	1,453	1064-1500	328.50	371.13
1501-1700	18,547	6,437	1501-1700	214.55	744.65
1701-1800	181	24,134	1701-1800	11.34	1512.21
1801-1900	678	51,747	1801-1900	15.3	1168.42
1901-2005	863	151,074	1901-2005	4.88	855.07

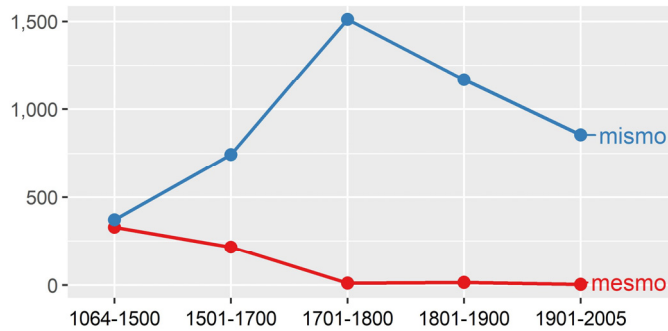


図7 : *mesmo* / *mismo*: 正規化頻度 (RAE: CDH)

3.2.3. 歴史・地理・社会

CODEA en LYNEALを用いて *mesmo* > *mismo* の年代推移を見る。

表12 : *mismo* - *mesmo*. 絶対頻度・正規化頻度. CODEA.

絶対頻度	絶対頻度		正規化頻度	正規化頻度	
	<i>mismo</i>	<i>mesmo</i>		<i>mismo</i>	<i>mesmo</i>
1200	51	6	1200	83.6	9.8
1300	98	39	1300	61.4	24.4
1400	59	79	1400	28.4	38.1
1500	181	35	1500	84.2	16.3
1600	125	38	1600	81.8	24.9
1700	286	6	1700	211.5	4.4

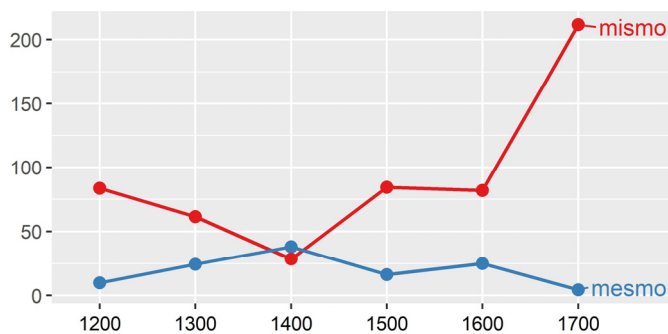


図8 : *mismo* - *mesmo*. 年代. CODEA (正規化頻度, 10⁵)

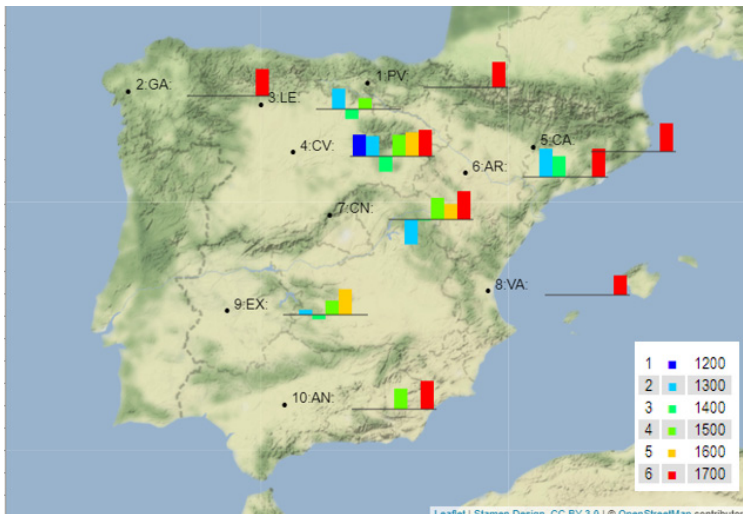
このように *mesmo* と *mismo* は1400年代に拮抗したが、先の CDH と同様に1500年代に *mismo* の優勢が決定している。

次の表は各年代と地域(地方)のクロス集計から導出した *agora* と *ahora* の対照値とその地図プロットである¹⁵。

¹⁵ 地方名 = AN: Andalucía, AR: Aragón, AS: Asturias, CA: Cataluña, CN: Castilla la Nueva, CV: Castilla la Vieja, EX: Extremadura, GA: Galicia, LE: León, MU: Murcia, NA: Navarra, PV: País Vasco, VA: Valencia.

表13 : *mismo* - *mesmo*. 対照値. CODEA.

年代	1.PV	2.GA	3.LE	4.CV	5.CA	6.AR	7.CN	8.VA	9.EX	10.AN
1200	0.000	0.000	0.000	0.765	0.000	0.000	0.000	0.000	0.000	0.000
1300	0.000	0.000	0.722	0.714	0.000	1.000	-0.882	0.000	0.176	0.000
1400	0.000	0.000	-0.351	-0.545	0.000	0.739	-0.043	0.000	-0.158	0.000
1500	0.000	0.000	0.385	0.756	0.000	0.000	0.766	0.000	0.500	0.727
1600	0.000	0.000	0.000	0.846	0.000	0.000	0.547	0.000	0.905	0.000
1700	0.900	0.947	0.000	0.935	1.000	1.000	1.000	0.692	0.000	1.000

図9 : *mismo* - *mesmo*. CODEA. 対照値

この地図によれば、各地でほとんど全年代を通じて現代語形 *mismo* が多数を占めている。とくに1700年代にほとんど統一して *mismo* となった。古形 *mesmo* は1300年代の CN と 1400年代の CN, CV, EX, LE に限定され、それ以上の年代・地域の広がりはない。

次に社会・文体差を示す王室文書と私文書の中で使われる *mismo* と *mesmo* の社会的文体特徴の年代推移を見る。

表14 : *mismo* - *mesmo*. CODEA. 王室文書・私文書, 正規化頻度 (10⁵)

<i>mismo</i>	王室文書	私文書	<i>mesmo</i>	王室文書	私文書
1200	77.1	64.3	1200	1.4	0.0
1300	31.8	76.1	1300	35.8	13.2
1400	32.2	36.8	1400	30.8	38.1
1500	144.3	76.6	1500	9.4	16.4
1600	117.6	73.4	1600	0	21.9
1700	105.6	138.4	1700	0	5.6

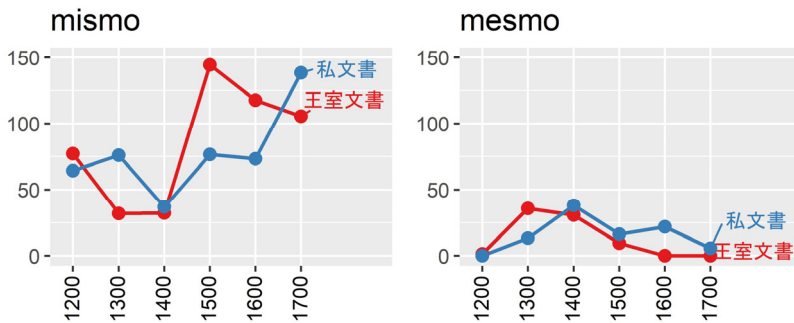


図10 : mismo - mesmo. CODEA. 王室文書/ 私文書 (正規化頻度, 10⁵)

このように1500年代に王室文書・私文書共に *mesmo* が減少したが、私文書では1600年代で少し回復している。一方、*mismo* は王室文書で1500年代に急上昇し私文書が後続した。このことは *mesmo* > *mismo* の変化は「上から普及」した可能性を示しているが、1300年代の私文書での *mismo* の使用は無視できない。

次は現代スペイン語を収集したコーパス COSERで古形 *mesmo* を検索すると、Zaragoza, Albacete, Badajoz の3地点で見つかった。いずれも半島周辺部に位置する。



図11 : mesmo. COSER. 現代スペイン語

3.2.4. 考察

Corominas / Pascual (1980: s.v. *mismo*) は、*mesmo* の語源を〔想定ラテン語形〕 *MED+IPSIMUS (IPSIMUS: IPSE「自ら」の最上級) の強調形 > *meismo* とし、中世では一般に *mesmo* が使われ、1600年代の初めに *mismo* が優勢になった、と述べている。Penny (2006: 171) によれば、スペイン語の初期には *meismo* と *me(esmo)* が見られ、14世紀から黄金世紀 (1550-1650) は *mesmo* が *mismo* より多用され、17世紀に *mismo* が標準形になり、18世紀以降は *mesmo* は農村部に残存した。このように、どちらも一致して1600年代に

mismo が優勢になった、と述べている。しかし、本節での観察によれば *mesmo* > *mismo* の時期は100年早く、1500年代である。

Hanssen (1913: 85) と Alvar / Pottier (1983: 109) は *mismo* と *mesmo* がどちらも *meismo* に遡る、という。

- (a) *meismo* > *mismo*
- (b) *meismo* > *mesmo*

筆者は、次のようにラテン語形から派生した *meismo* の強勢はラテン語の強勢配置に従って *meísmo* とし、次のようにそれぞれの音韻変化に強勢の条件を設定する¹⁶。

- (a') *meísmo* > *mísimo*
- (b') *meísmo* > *méismo* > *mésmo*

スペイン語は一般に「開母音+強勢母音」という「母音分立」(hiatus)を嫌うので (Torrens Álvarez, 2018: 52-53), (a') では単母音化して母音分立を回避し(*ei* > *i*), (b') では強勢の位置を先行させて下降二重母音を作り母音分立を回避している (*ei* > *éi*)¹⁷。どちらもスペイン語で自然な音韻変化である。その結果生まれた *mismo* も *mesmo* も安定した音韻形式であった。筆者はこの理由から *meísmo* > *méismo* > *mísimo* という連鎖的な変化を考えず、当初から *mísimo* ~ *méismo* という語形の変異があり、中世 (1200-1400年代) では常に競合していた、と考える (図8)。

1500年代に *mismo* が *mesmo* を凌いで圧倒的に優位になった理由は、中世の競合状態で *mismo* の方が優勢であったためであろう。語形の競合状態は言語の通信機能に支障を起こすので、王室文書などでは統一されることが好ましい。そこで、当時まで優勢であった *mismo* が採用され、この「上からの変化」(Labov, 1994: 78; Trask 2000: 56; Labov 2001: 196) が完了したのが1500年代であった。よって、*mismo* と *mesmo* の歴史的関係を図式で示すと次のようになる(A: *mismo*, B: *mesmo* ; A=Bは拮抗状態を示す)。

A → A=B → A

Medina Morales (2005: 97-98) は *mesmo* > *mismo* の過程で、*e* > *i* という音韻変化を認めているが、筆者は *mismo* が当初から優勢であったことを鑑み (図7, 8), そのような音韻変化を認めない。また強勢がある 'e' が 'i' に変化することはスペイン語史上一般的ではない。よって、*mismo* ~ *mesmo* は音韻の変異ではなく語形の変異であった、と考える。歴史上 *mismo* が先行した理由は、(a') の変化を見るとわかるように、強勢移動のステップを必要とし

¹⁶ この段落ではすべての語形の強勢母音にアクセント記号をつける。

¹⁷ Torrens Álvarez (2018: 52) は *veinte* > *véinte*, *reina* > *réina* の例を挙げている。

なかったためである。一方、(b') の *mesmo* は強勢の移動を待たなければならなかった。

DQ で観察される *mesmo* の多用は幾分不自然である。Cervantes は独特の擬古文で当時すでに古形となっていた語形を使ったようだ。

3.3. «ahora» [副] 「今」

3.3.1. 社会

次に、DQ の中で *agora* ~ *ahora* の語形のバリエーションの社会差を観察する。

表15 : *agora* / *ahora*: 性 (絶対頻度/正規化頻度/共起係数)

語形	男性	女性	男性	女性	男性	女性
<i>agora</i>	124	12	27.8	29.8	0.368	0.374
<i>ahora</i>	294	29	65.9	72	0.569	0.601

表16 : *agora* / *ahora*: 年齢 (絶対頻度/正規化頻度/共起係数)

語形	若者	中高年	若者	中高年	若者	中高年
<i>agora</i>	2	134	34.7	27.9	0.317	0.357
<i>ahora</i>	7	316	121.3	65.8	0.707	0.469

表17 : *agora* / *ahora*: 社会層 (絶対頻度/正規化頻度/共起係数)

語形	上流	平民	上流	平民	上流	平民
<i>agora</i>	86	50	22.8	46	0.308	0.424
<i>ahora</i>	212	111	56.1	102.1	0.473	0.667

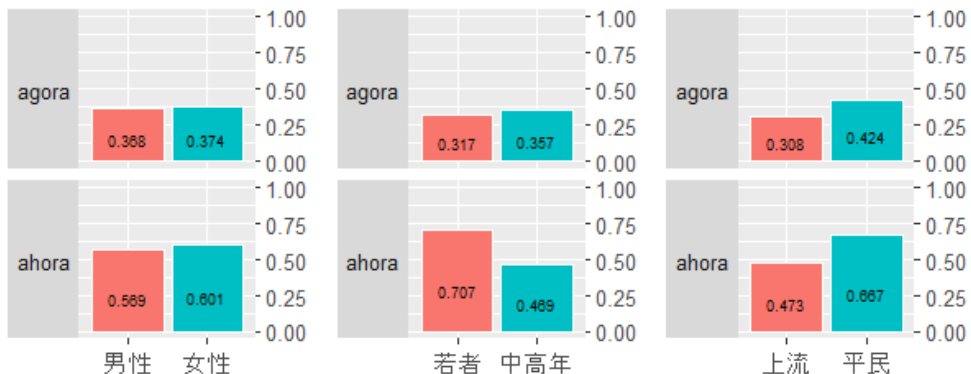


図12 : *agora* / *ahora*: 性・年齢・社会層 (共起係数)

このように、DQ において *ahora* はとくに若者・平民に好まれていたので、*agora* > *ahora* の推移は「下からの変化」であった可能性がある。

3.3.2. 年代

Nieto Jiménez / Alvar Ezquerro (2007) を参照すると *agora* と *ahora* の語形を載せている辞書の発行年は次のようになる。

agora: 1490, 1492, 1495, 1499, 1505, 1520, 1526, 1535, 1553, 1554, 1565, 1568, 1570, 1582, 1587, 1599, 1601, 1604, 1607, 1609, 1614, 1620, 1621, 1634, 1639, 1645, 1660, 1666, 1670, 1676, 1693, 1700, 1705, 1706, 1721, 1723.

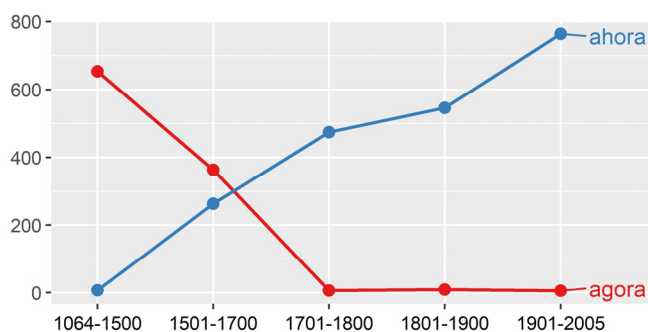
ahora: 1551, 1570, 1587, 1599, 1607, 1609, 1617, 1620, 1636, 1639, 1661, 1666, 1670, 1726.

これを見ると、確かに *agora* が古形であり、*ahora* は比較的新しく、その初出が1551年であることがわかる。Cervantes (1547-1616) の時代ではすでに *ahora* が一般に使用され、*agora* は古形であることから高雅な文体であった。また、逆に地方に残存する語形であることから揶揄されることもあったようだ。いずれにしても特殊な語形であった。

次は『スペイン語歴史辞典コーパス』(CDH)の *agora/ahora* の検索結果である。

表18 : *agora / ahora* (RAE: CDH)

絶対頻度	<i>agora / ahora</i>		正規化頻度	<i>agora / ahora</i>	
	<i>agora</i>	<i>ahora</i>		<i>agora</i>	<i>ahora</i>
1064-1500	25,593	273	1064-1500	653.7	6.96
1501-1700	31,375	22,757	1501-1700	363.0	263.26
1701-1800	107	7,574	1701-1800	6.7	474.57
1801-1900	398	24,197	1801-1900	9.0	546.35
1901-2005	1,029	13,573	1901-2005	5.8	764.04

図13 : *agora / ahora* 年代. 正規化頻度 (RAE: CDH)

この表とグラフを見ると、*agora > ahora* の変化は古典期 (1500-1700) と近代 (1701-) の間にあったことがわかる。これは

A → A:B → B:A → B

という図式に従う典型的な言語変化である。しかし、比較的わずかであるが、中世 (1064-1500) にも *ahora* が出現していたことは、先述の Alvar / Pottier (1983) の説明を参照すれば無視できない。RAE のコーパス資料 (CDH) の頻度情報の年代区分はとくに初期 (700-1200)・中世 (1200-1500) の下位区分がないため、すべてのコンコーダンスラインを調べると *a(h)ora* は次のような頻度 (絶対頻度) であった。

1200 (頻度: 21), 1300 (22), 1400 (201).

このように1200, 1300年代における *a(h)ora* はたしかに存在したが頻度は低かった。

3.3.3. 歴史・地理・社会

CODEAとCOSERのコーパス資料を使って *agora* と *ahora* の歴史・地理・社会の状態を観察する。はじめに年代推移を見る。

表19 : *agora - ahora*. 絶対頻度・正規化頻度. CODEA.

絶対頻度	agora	ahora	正規化頻度	agora	ahora
1200	75	0	1200	56.1	0
1300	236	0	1300	112.5	0
1400	315	1	1400	114.3	0.4
1500	231	19	1500	79.7	6.6
1600	38	68	1600	21.3	38.1
1700	5	63	1700	3.7	46.3

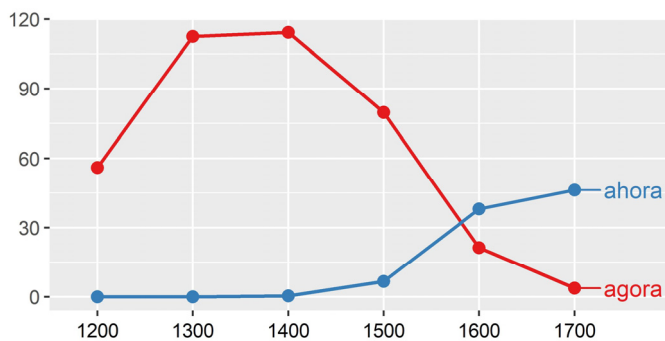


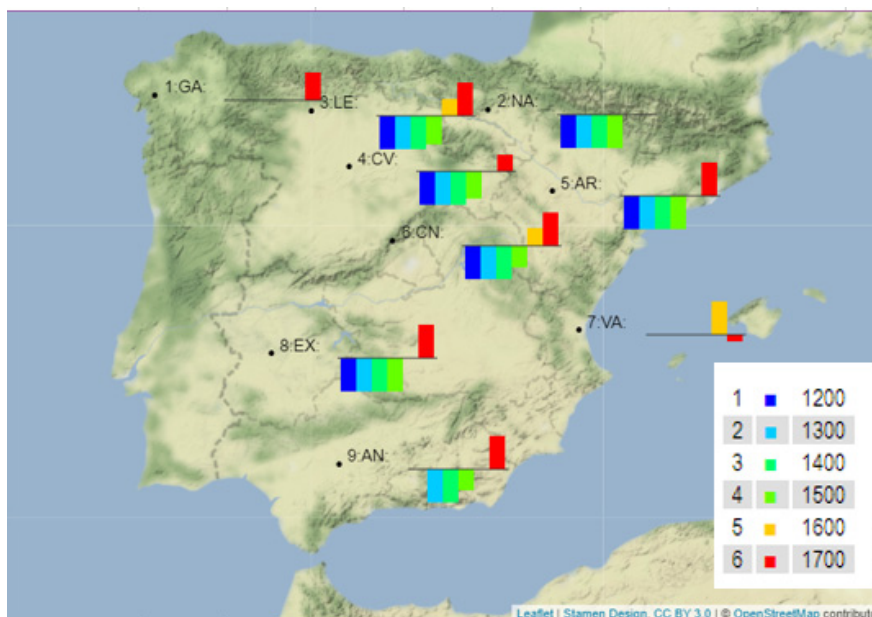
図14 : *agora / ahora* 年代. CODEA. (正規化頻度, 10^5)

上のグラフを見ると *agora* > *ahora* の変化は1600年代であったことがわかる。この観察は先の CDH の検索結果 (図13) と整合する。

次の表は各年代と地域 (地方) のクロス集計から導出した *agora* と *ahora* の対照値とその地図プロットである (*ahora* は *aora*, *a ora* を含む)。

表20 : *ahora* - *agora*. CODEA. 対照値

*	1.GA	2.NA	3.LE	4.CV	5.AR	6.CN	7.VA	8.EX	9.AN
1200	0.000	-1.000	-1.000	-1.000	-1.000	-1.000	0.000	-1.000	0.000
1300	0.000	-1.000	-1.000	-1.000	-1.000	-1.000	0.000	-1.000	-1.000
1400	0.000	-1.000	-1.000	-1.000	-1.000	-1.000	0.000	-1.000	-1.000
1500	0.000	-1.000	-0.857	-0.818	-1.000	-0.647	0.000	-1.000	0.185
1600	0.000	0.000	0.500	0.000	0.000	0.532	1.000	0.000	0.000
1700	0.833	0.000	1.000	0.500	1.000	1.000	-0.200	1.000	1.000

図15: *agora* - *ahora*. CODEA. 対照値

上の表と地図を見ると、*ahora* が1600年代に半島中央部の LE, CN に出現し、1700年代にほぼ半島全域を占めたことがわかる。中央の CN に位置する Madrid が首都に定められたのは1561年であり、その後各地から首都に移住した人々の方言差は首都で標準化され、その標準化された言語が地方にも強い影響を与えた (Del Barrio de la Rosa 2018: 233)。 *agora* > *ahora* の変化はその典型的なケースである。

資料 CODEA の中で文体差の大きな王室文書と私文書を選択すると *agora* と *ahora* の正規化頻度はそれぞれの文書で次のようになる。

表21: *agora - ahora*. 王室文書・私文書. 正規化頻度. CODEA.

agora	王室文書	私文書	ahora	王室文書	私文書
1200	43.4	2.2	1200	0.0	0.0
1300	50.5	17.6	1300	0.0	0.0
1400	40.6	25.0	1400	0.4	0.0
1500	33.5	17.2	1500	0.3	3.8
1600	11.8	3.4	1600	7.3	19.1
1700	0.0	3.7	1700	0.0	32.3

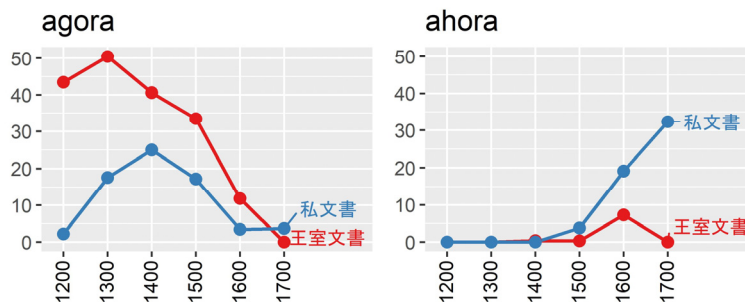


図16: *agora - ahora*. 王室文書・私文書. CODEA (正規化頻度, 10^5)

上のグラフは、古形 *agora* が1300年代に王室文書で多く使用され、私文書は100年遅れて後続していること、そして新形 *ahora* が1600年代でとくに私文書で大きく上昇していることを示している。王室文書は「上からの普及」を示し、私文書が「下からの普及」を示していると考えられるので (Del Barrio de la Rosa 2018: 144-145), *agora* が上位 (王室) の語形であり、*ahora* は下位 (民衆) の語形であったようだ。

最後に、古形 *agora* の現状 (21世紀) を確認するために COSER で検索すると、次の結果となった (正規化頻度, 10^5)。

表22: *agora*. 現代. 正規化頻度. COSER

*	Albacete	Ávila	Cantabria	León	Teruel	Zaragoza
agora	9.4	2.6	1.7	2.4	183.5	1.0

これら各地を地図にプロットすると次のようになる。

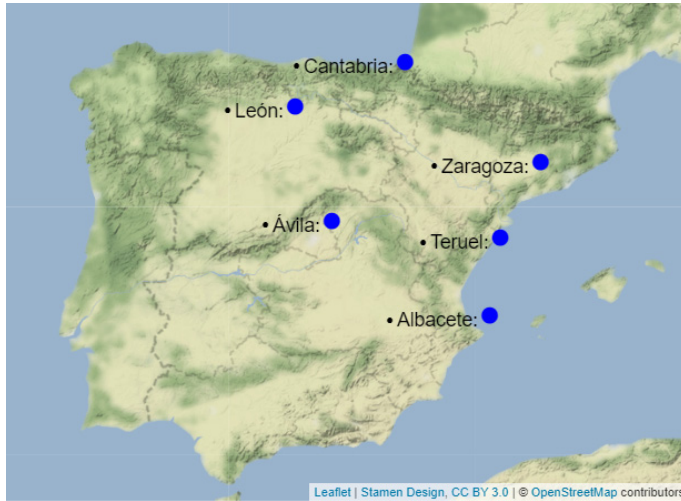


図17: agora. 現代. COSER.

とくに頻度が高い Teruel とそれに近い Albacete は半島の周辺部に位置する。周辺分布は多くの古形の特徴である。

3.3.4. 考察

以上の観察をまとめると、1500年代まで *agora* は古形・教養層・正規の文体・首都・王室文書の語形であり、*ahora* は新形・非教養層・民衆の文体・私文書の語形であった。現代スペイン語では新形 *ahora* が標準形であり、古形の *agora* は半島の周辺部にわずかに残存するだけである。

Corominas/Pacual (1980. s.v. *hora*) や Penny (2006: 158) は〔ラテン語〕HAC HORA〔奪格〕「この時に」>*agora*>*ahora*〔副〕「今」という通時的変化を想定する¹⁸。よって DQ における *agora* ~ *ahora* の社会・文体的変異は、*agora* > *ahora* という通時的変化の移行期の状態を示していることになる可能性がある。しかし、ここで、

$g \rightarrow$ [無音]/a_o

という音韻変化を想定しているが、このような変化はスペイン語で一般的ではない (*agosto*〔名〕「8月」、*aguja*〔名〕「針」、*agudo*〔形〕「鋭い」などは -g- を保持している)。

一方、Alvar/Pottier (1983: 334-335) は、*ahora* の初出が *agora* と並んで早かったため (12世紀)、*agora* > *ahora* という通時的変化を考えず、*ahora* には *agora* とは異なる語源 AD HORA を設定した。Alonso (1986: s.v. *ahora*) も同

¹⁸ Real Academia Española の *Diccionario de la lengua española* も、*agora* > *ahora* を認めている。

<https://dle.rae.es/ahora?m=form>

じ語源を載せている。しかし、ラテン語形 AD HORA(M)〔対格〕は「定刻に」という意味で使われたので (Segura Mungía, 2007, s.v. *hora*), *ahora*「今」にこの語源を認めることは難しい。また, Alvar / Pottier と Alonso による AD HORA > *ahora* というラテン語源は *ahora* の早期の出現を前提とするが, 筆者が扱った資料では僅少であった。

よってAD HORA > *ahora* というラテン語源を支持することは難しく, *agora* → *ahora* の年代的推移を認めざるを得ない。しかし筆者は, 次の理由で, これが音韻変化ではなく形態変化であった, と考える。

先に見たように (図15), 1500年代に新形が出現したのは半島中央部 (LE, CN) であったが, これは王室の規範によるものではなく, 民衆の「下からの変化」であった (図16)。それが1600年代に一举に全国に普及した理由を筆者は次のように考察する。*agora* → *ahora* の変化は高頻度で用いられる *hora*〔名詞〕「時間」からの派生という言語内の要因によって生じた, と考えられる。すなわち, *ahora* は *hora* に「様態・方向」を示す接頭辞 *a-* が付加されて規則的に派生した語形である, と見なす。「*a*+名詞」→ 副詞という語形成の類例として *a* + *caso*〔名〕「場合」→ *acaso*〔副〕「たぶん」, *a* + *prisa*〔名〕「急ぎ」→ *aprisa*〔副〕「速く」などが挙げられる。そして, 「接頭辞 + *hora*」の例として *a deshora*「突然」(→「折り悪く」)がある。接頭辞・接尾辞は語基に隣接し, 意味と語形の関係が明示されるため形態的に自然である (Dressler, 2003: 464)。*ahora* の綴りで無音の *-h-* が用いられるのは規範的にラテン語 HORA に合わせたためではなく, 民衆がスペイン語 *hora* と同調したためであろう。また, CODEA コーパスで *a ora* という分離された表記例が1400, 1500, 1600年代に9例見つかることも「*a*+名詞」→副詞という語形成に根拠を与える。このような自然な(無標の)語形成は言語使用の負担が小さい。一方, *agora* は *hora* から独立した特殊形であったので形態的に有標であり, 言語使用の負担が大きかったはずである。

言語が接触して生まれるクレオール言語では元の言語の構造が規則化・単純化することが知られている (細川, 1992; 田中, 1999)。また, 同一言語の方言間の接触により生成された新しい標準形 (コイナー) でも不規則形が淘汰される (Tuten, 2003; Del Barrio de la Rosa, 2018)。*ahora* は1500年代まで私文書で僅かに使われていたが, 1600年代になるとコイナー化が進んだ半島中央部, とくに新首都 Madrid が位置する CN で使用が一気に加速し, さらに標準形として全国に普及したようだ。

agora > *ahora*の歴史は, 次のような典型的な言語変化の図式で示される (A: *agora*, B: *ahora*)。

A → A : B → B : A → B

4. 結論

以上で、各種の資料を使って、「*así*», «*mismo*», «*ahora*» という高頻度で使われる語の異語形 (*así* ~ *ansí*, *mismo* ~ *mesmo*, *ahora* ~ *agora*) の年代的变化と地理・社会的変異を観察し、変化と変異の要因を考察した。

«*así*»: *así* の語源はラテン語形 AD SIC または副詞に見られる単なる *a-* による延長であった、という説に加えて、筆者は SIC > *sí* [副詞] 「はい」(肯定の返事) や S I > *si* [接続詞] 「もし~ならば」との同音衝突を避けた可能性と、一般に語形が短いときは補強のために延長されることがあったことも原因の一つとして考えた。この *así* は当初から連綿と全時代を通じて圧倒的多数であった。一方、*ansí* は無意味な *-n-* が音節末に生じた結果である。この *ansí* はCODEAの資料によれば1400, 1500, 1600年代で私文書とともに王室文書でも使われていたので中世・近代において両者に社会・文体差はなかった、と考えられる。事実、DQ の登場人物の会話には *ansí-así* の間にそのような区別はない。図式的に示せば両者の関係は次のようになる (A: *así*, B: *ansí*; A:BはAの優位を示す)。

(1) A → A:B → A

«*mismo*»: *mismo* ~ *mesmo* の変異に強勢の条件を設定して、次のような変化を想定した。

(a) *meísmo* > *mismo*

(b) *meísmo* > *méismo* > *mésmo*

両者は中世 (1200-1400年代) では常に競合していたが、その中で優勢であった *mismo* が1500年代に圧倒的に優位を占めた。語形の競合状態は言語の通信機能に支障を起こすので、とくに王室文書で当時まで優勢であった *mismo* が採用された。よって、*mismo* と *mesmo* の歴史的関係を図式で示すと次のようになる (A: *mismo*, B: *mesmo*; A=Bは拮抗状態を示す)。これは音韻変化ではなく、語形の交替であった。

(2) A → A=B → A

«*ahora*»: *agora* は古形・教養層・正規の文体・首都・王室文書の語形であり、*ahora* は新形・非教養層・民衆の文体・私文書の語形であった。1500年代に新形が出現したのは半島中央部であった。それが1600年代に一挙に全国に普及した。これは王室の規範による変化ではなく、民衆の「下からの変化」であった。よって、*agora* → *ahora* の歴史は次の図式 (3) で示される (A: *agora*, B: *ahora*)。この図式は典型的な言語変化を示している。これも音韻変化ではなく、語形の交替であった。

(3) A → A : B → B: A → B

図式 (1) では、歴史的に B (*ansí*) の出現が認められるが、それが古形 A (*así*) と交替することはなかった。*ansí* は、*-n-* の挿入が有標であり、その地理的分布が限定的であり、使用頻度が比較的lowだったため、1500年代後半に移動した新首都 Madrid で標準形に成り得なかったのである。

図式 (2) は A (*mismo*) と B (*mesmo*) の拮抗状態が生じたことを示しているが、守旧的な王室文書が伝統的に優勢であった A を採用した。たしかに A と B は拮抗していたが、B が A を凌駕する積極的な理由はなかった、と考えられる。このように「上からの変化」によって普及した語形は意識して学ばれる規範形である (Labov, 1994: 78; Trask 2000: 56; Labov 2001: 196; エルスパス, 2015: 61; 家入・内田, 2015: 125)。そして、Labov (1994: 453) によれば「上からの変化」は一般に語の借用 (lexical borrowing) を含む。ここで確認した *mismo* は借用語ではないが、語の古形が採用されたので、やはり語彙的性格を持つ。規範の個別的学习には負担・コストがかかるが、この場合 *mismo* は伝統的に頻用されていたので負担は少なく、よって全国的に普及したようだ。

図式 (3) の B (*ahora*) が1600年代に「下からの変化」の結果として一挙に普及した理由として、高頻度で用いられていた *hora* [名詞] 「時間」との派生関係が考えられる。1600年代になると、コイネー化が進んだ半島中央部で *ahora* の使用が加速し、さらに標準形として全国に普及した。一般に「下からの変化形」は話者が意識的に学ぶ規範形ではなく、音韻・形態変化のように無意識で自然に獲得するものである。無意識の自然獲得には負担・コストがかからないので、獲得された音韻・形態は一般的な標準形となった。

言語が歴史・地理・社会の範囲の中で変化する「要因」(factor) を考察することは可能であるが、それは物質が自然の中で変化する「原因」(cause) とは異なる。たとえば水が温度と気圧の条件によって氷・水蒸気に変化することは普遍的な法則なので、温度・気圧を「原因」と呼ぶことができる。一方、言語形式は派生・類推・同化・異化などの言語内的条件と、年代 (歴史)、地域 (地理)、人口移動・標準化・規範、対話者・場面などの言語外的条件が作用する。しかし、温度・気圧の条件の作用によって水の状態が変化することと異なり、条件が変化しても言語形式が変化するとは限らない。Weinreich 他 (1968: 188) や Medina Morales (2005: 17-18) の言うように、通時的変化は共時的変異を前提とするが (図式 (3)), 共時的変異が常に通時的変化を引き起こすわけではない (図式-(1), (2))。よって A → B という単純で完全な変化は現実的ではない。つまり、共時的変異は通時的変化の必要条件であるが、十分条件ではない (Guy, 2003: 371)。従って本研究は、自然現象の変化の普遍的・先験的予測を可能にするような「原因」ではなく、言語現象の一般的・後験的解釈を可能にする「要因」を探る試みである。考察した「要因」の一般性を求めるためにさらに多くの分析を積み上げていかなければならない。

【補説】

(1) 相対頻度

たとえば、次の頻度は DQ (Miguel Cervantes: 『才知ある郷士ドン・キホーテ・デ・ラ・マンチャ』) における *agora* と *ahora* が二人の主要登場人物 (Q: Quijote, S: Sancho) と作者のナレーション(N) に出現した絶対頻度を示している。

データ行列：周辺度数 (sum)

	Q	S	N	sum
agora	50	40	5	95
ahora	109	70	18	197
sum	159	110	23	292

次の表では、各頻度に加えて、その順位 (降順) を括弧内 (..) に示した。

データ (順位)

	Q	S	N
agora	50 (3)	40 (4)	5 (6)
ahora	109 (1)	70 (2)	18 (5)

このデータ行列の数値は絶対頻度で計算しているので比較できない。たとえば、*agora:Q=50* は *agora:S=40* よりも大きいですが、そもそも Qの縦和 = 159 がSの縦和 = 110よりもかなり大きいので本来比較できない数値である。そこで、一般に頻度行列を扱う研究では、絶対頻度を縦和で割った相対頻度やパーセントを使って比較することが多い¹⁹。次のアステリスク (*) より右の行列は、それぞれの絶対頻度を縦和で割って求めた相対頻度 (「行による相対頻度」) とその順位を示している。

相対頻度-列 (順位)

	Q	S	N	Q	S	N
agora	50 (3)	40 (4)	5 (6)	* 0.314 (5)	0.364 (4)	0.217 (6)
ahora	109 (1)	70 (2)	18 (5)	* 0.686 (2)	0.636 (3)	0.783 (1)

このように、データ行列 (絶対頻度) と縦和による相対頻度(列)を比較すると、順位が大きく異なる(3,4,6,1,2,5 : 5,4,6,2,3,1)。

次に「行による相対頻度」を見る。

¹⁹ または絶対頻度をそれぞれのグループの総語数で割って、適当な数値 ($10^3=1,000$, $10^6=1,000,000$ など) を掛けた正規化頻度が使われる。

相対頻度-行 (順位)

	Q	S	N		Q	S	N
agora	50 (3)	40 (4)	5 (6)	*	0.526 (2)	0.421 (3)	0.053 (6)
ahora	109 (1)	70 (2)	18 (5)	*	0.553 (1)	0.355 (4)	0.091 (5)

このように「列による相対頻度」と「行による相対頻度」の順位はまったく異なる (5,4,6,2,3,1 : 2,3,6,1,4,5)。「列による相対頻度」(縦方向で見る)は「各登場人物が使用している語形の割合」を示し、「行による相対頻度」(横方向で見る)は「各語形を使用している登場人物の割合」を示している。どちらも同じことを述べているようだが、実際に計算された数値と順位が大きく異なるので注意しなければならない。たとえば agora:Q は「列による相対頻度」では0.314 (5)で第5位となっているが、「行による相対頻度」の agora:Q は 0.526 (2) で第2位を占めている。よって、同じ現象を扱っているのに、視点が変わると大きく異なる結論になる。頻度行列を扱った研究では、行・列の一方だけの相対行列を作成し分析していることが多い。その結論は選択された視点に立てば有効であるが、別の視点に立てば大きく異なることが多い。よって、理想的には横行と縦列の視点によって変わるような指標ではなく、コンスタントな指標が望ましい。

池田 (1976: 122) はデータ行列全体を総和(=292)で割って、全体を相対化する方法を扱っている。これは「全体のなかでそれぞれの条件を同時に満たすものがどのくらいの割合であるかをみようとするもので、総和が一定(=1.00)になるように相対化されたものである。」

相対頻度-全体 (順位)

	Q	S	N		Q	S	N
agora	50 (3)	40 (4)	5 (6)	*	0.171 (3)	0.137 (4)	0.017 (6)
ahora	109 (1)	70 (2)	18 (5)	*	0.373 (1)	0.240 (2)	0.062 (5)

たしかに、表全体の相対頻度はそれぞれのセルが全体の中で占める割合がわかる。しかし、大小関係はデータ行列そのものと同じになる。そして、データ行列と同様に、横和(行和)と縦和(列和)の情報を生かしていないので、たとえば agora:Q と agora:S を直接比較できない。Q と S のそれぞれの縦和が勘案されていないためである。この問題は次の「共起回数」で解決される。

(2) 共起係数

石川 (2021: 124-133) は2語 (X, Y) の連鎖の重要度を共起頻度を使って計算する方法として類似度を取り上げ、Dice係数, Jaccard係数, Cosine係数などを紹介している。他にも Romesburg (1989, 1992: 187) は多くの指標を挙げている。次の一覧は、それぞれの指標の数式と範囲を示す。

$(a+d) / (a+b+c+d)$, [0, 1] # Simple matching
 $2*a / (2*a+b+c)$, [0, 1] # Dice
 $a / (a+b+c)$, [0, 1] # Jaccard
 $a / \sqrt{(a+b)*(a+c)}$, [0, 1] # Cosine
 $(a+d-b-c) / (a+b+c+d)$, [-1, 1] # Hamann
 $(a*d-b*c) / (a*d+b*c)$, [-1, 1] # Yule
 $(a*d-b*c) / \sqrt{(a+b)*(c+d)*(a+c)*(b+d)}$, [-1, 1] # Phi

上の a, b, c, d はそれぞれ次の表のX:Y, X:!Y, !X:Y, !X:!Y (強調記号 ! は「以外」) を意味する。n は 総和 (a+b+c+d) を示す。

*	Y	!Y (Y以外)	和
X	a	b	a + b
!X (X以外)	c	d	c + d
和	a + c	b + d	n = a+b+c+d

ここで、次のようなクロス集計表に共起の指標を適用する方法を考える²⁰。

共起回数 (順位)			
	Q	S	N
agora	50 (3)	40 (4)	5 (6)
ahora	109 (1)	70 (2)	18 (5)

たとえば, agora:Q = 50 は「agora (X)とQ(Y) が共起する度数(a)」を示している。一方, agora:S = 40と agora:N = 5の和 (40+5 = 45) は「agora (X) と Q 以外 (!Y) が共起する度数 (b)」を示している (プログラムでは横和から a を引いて求める)。また, ahora:Q = 109は「agora 以外 (!X)と Q (Y) が共起する度数 (c)」を示す (プログラムでは縦和から a を引いて求める)。最後に ahora:S = 70 と ahora:N = 18 の和 (70+18=88) は「agora 以外 (!X) と Q以外 (!Y) が共起する度数 (d)」を示す, と考えられる(プログラムでは総和から a+b+cを引いて求める)。よって, たとえば agora:Q の単純一致係数 (Simple matching: Sm) は, 次のように計算される。

$$Sm(\text{agora:Q}) = (a + d) / n = (50 + 88) / (50 + 45 + 109 + 88) = 0.473$$

同様にして, agora:S, agora:N, ahora:Q, ahora:S, ahora:N の単純一致係数を計算すると, 次の結果になる。

²⁰ ここでは簡単化して絶対頻度を使って共起係数を計算した。実際には, 正規化頻度を使うべきである。

Simple matching (順位)

	Q	S	N	*	Q	S	N
agora	50 (3)	40 (4)	5 (6)	*	0.473 (4)	0.572 (2)	0.630 (1)
ahora	109 (1)	70 (2)	18 (5)	*	0.527 (3)	0.428 (5)	0.370 (6)

単純一致係数 (simple matching), Dice, Jaccard, Cosine は範囲が [0,1] となり, それぞれのセルの数値の比較ができる。しかし, 後で見るように, 乱数を使って実験すると, その平均値は 0.5 にはならない。

Dice (順位)

	Q	S	N	*	Q	S	N
agora	50 (3)	40 (4)	5 (6)	*	0.394 (3)	0.390 (4)	0.085 (6)
ahora	109 (1)	70 (2)	18 (5)	*	0.612 (1)	0.456 (2)	0.164 (5)

Jaccard (順位)

	Q	S	N	*	Q	S	N
agora	50 (3)	40 (4)	5 (6)	*	0.245 (3)	0.242 (4)	0.044 (6)
ahora	109 (1)	70 (2)	18 (5)	*	0.441 (1)	0.295 (2)	0.089 (5)

Cosine (順位)

	Q	S	N	*	Q	S	N
agora	50 (3)	40 (4)	5 (6)	*	0.407 (3)	0.391 (4)	0.107 (6)
ahora	109 (1)	70 (2)	18 (5)	*	0.616 (1)	0.476 (2)	0.267 (5)

Hamann, Yule, Phi では分子で対称的な数値 (a+d, b+c; a*d, b*c) の差が使われている。このために次のような2行のデータ(または2列のデータ)では, かならず同じ数値の正と負の値となる。

Hamann (順位)

	Q	S	N	*	Q	S	N
agora	50 (3)	40 (4)	5 (6)	*	-0.055 (4)	0.144 (2)	0.260 (1)
ahora	109 (1)	70 (2)	18 (5)	*	0.055 (3)	-0.144 (5)	-0.260 (6)

Yule (順位)

	Q	S	N	*	Q	S	N
agora	50 (3)	40 (4)	5 (6)	*	-0.054 (4)	0.138 (2)	-0.288 (6)
ahora	109 (1)	70 (2)	18 (5)	*	0.054 (3)	-0.138 (5)	0.288 (1)

Phi (順位)

	Q	S	N	*	Q	S	N
agora	50 (3)	40 (4)	5 (6)	*	-0.025 (4)	0.064 (2)	-0.067 (6)
ahora	109 (1)	70 (2)	18 (5)	*	0.025 (3)	-0.064 (5)	0.067 (1)

このことは具体例で見るとわかりやすい。たとえば次は Yule の式である。

$$(a*d - b*c) / (a*d + b*c) \quad \# \text{ Yule}$$

この分子の $a*d$ と $b*c$ は、次の agora:Q の場合では

	Q	S	N
agora	50 (3)	40 (4)	5 (6)
ahora	109 (1)	70 (2)	18 (5)

$a*d = 50*(70+18)$, $b*c = (40+5)*109$ となる。そして、ahora:Q の場合は、それらが交替して、 $a*d = (40+5)*109$, $b*c = 50*(70+18)$ となる。その結果、 $Yule(agora:Q) = -1*Yule(ahora:Q)$ となる。Hamann, Phi も同様である。よって、この3係数 (Hamann, Yule, Phi) は2語形の頻度を比較するような本研究のように2行のデータを扱うときは使用できない。

Dice係数の乗数2は2行2列の行列には有効であるが、一般に $n * p$ の行列 (クロス集計表) では、後述するように乱数の多数回の適用の平均値が範囲 [0,1] の中間の 0.5 にはならないので不都合である。その平均値を 0.5 とするために、乗数 (m) を $n+p-2$ にする必要がある。その理由は b と c のセル数が、それぞれ $p-1$, $n-1$ となるためである。Dice が0.5となるのは $2a = b+c$ のときであるが、このとき2行2列の行列では $2*a$ は2個分のセルに対応するが ($2*a$), $b+c$ も2個分のセルになる (bとc)。しかし、 n 行 p 列の行列では $b+c$ のセルの個数は $(p-1) + (n-1) = n+p-2$ 個のセルになるので、a をそれと同じ個数にするために乗数を2ではなく、次のように $m = n+p-2$ としなければならない。

$$\frac{2*a}{2*a+b+c} \# \text{ Dice}$$

$$\frac{m*a}{m*a+b+c} \# \text{ Dice-m, } m = n+p-2$$

次が乗数 (m) を補正した Dice係数(Dice.m) の結果である。先のDice係数よりも全体的に数値が上昇していることがわかる。後で見るように、乱数を使って発生させた n 行 p 列の行列に適用すると平均値が0.5になる。よって、Dice.m係数が0.5以上ならば、平均よりも大きな数値と言える。

Dice.m (順位)	Q	S	N	Q	S	N
agora	50 (3)	40 (4)	5 (6)	* 0.493 (3)	0.490 (4)	0.122 (6)
ahora	109 (1)	70 (2)	18 (5)	* 0.703 (1)	0.557 (2)	0.227 (5)

一般にクロス集計表を類似度を使って分析するときは、Dice.m ([0,1]) を使用するとよい。Dice.m は2行2列の場合は $m = n+p-2 = 2$ となり、Dice と一致するので、すべてのクロス集計表に適用できる (Diceは2行2列の行列だけに適用できる)。

次は Data と Dice-m の数値を視覚化したグラフである。このように両者はかなり異なる。

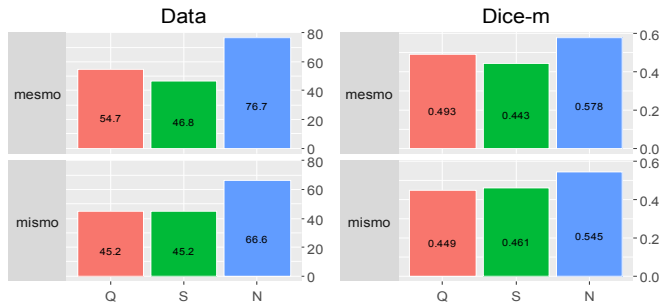


図18: データ行列とDice-m

(3) 共起係数の比較

これまでに扱ってきた分布の特徴を示す各指標の間には、どのような関係があるのかを調べたい。そのために、正規分布乱数を12個発生させ、これを3行4列にした行列を使って各指標を計算し、行列の最初のセル(1行:1列)の指標を取り出す、という操作を100回行った。こうして作成した100行のデータフレームの相関行列と相関グラフを出力した (Rのパッケージ psych を使用した)。

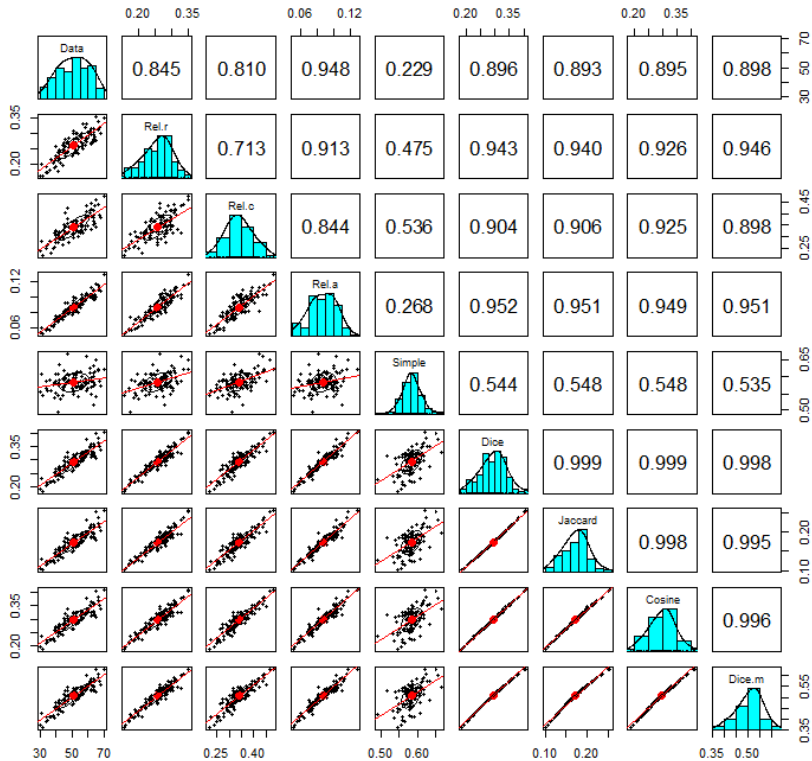


図19: 指標間の相関

(Data: データ行列, Rel.r: 相対頻度(行), Rel.c: 相対頻度(列),
Rel.a: 相対頻度(全体), Simple: Simple matching)

この図を見ると次のことがわかる。(1) データ行列 (Data) と相対頻度 (Rr, Rc, Ra) は相関しているが、相関係数はあまり高くなく、散布図は大きなブレを見せている。とくに行の相対頻度 (Rr) と列の相対頻度 (Rc) の相関が弱い。よって行と列の視点は大きく異なる様相を示す、と言える。(2) 類似度係数の中で Simple matching が他の類似度係数と相関しない。(3) Dice, Jaccard, Cosine, Dice.m の相関が非常に強い。(4) すべての指標はその平均値に集中し、正規分布に近似する。

このように相対頻度は他の指標と整合しない。また、行の相対頻度 (Rr) と列の相対頻度 (Rc) の相関が弱いので注意が必要である。逆に、行の視点または列の視点だけを意識した一方向的分析には有用である。しかし、本研究では語形と登場人物・年代・地域・社会の関係を双方向的に分析するので、相対頻度を採用しない。

次にそれぞれの指標の分布をボックスチャート(箱ひげ図)で観察する。一般にボックスチャートは中央値と四分位 (25%点, 50%点, 75%点) を使って作成されるが、ここでは平均値の位置に関心があるので平均±標準偏差を用いたボックスチャートを作成した (Rのパッケージ ggplot を使用した)。

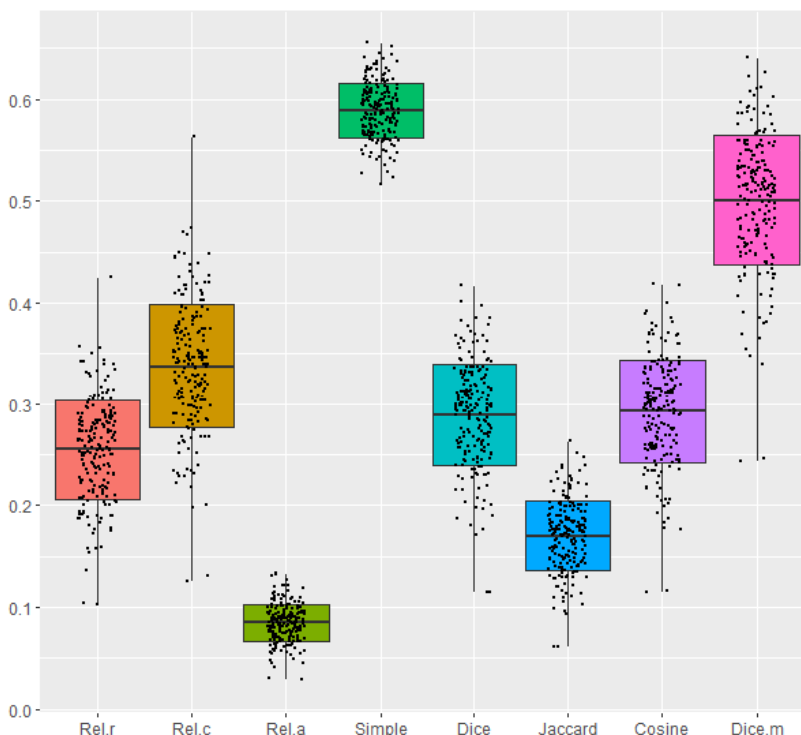


図20: 各指標の分布 (平均±標準偏差)

扱った指標の理論的範囲はたしかに [0, 1] であるが、平均値は必ずしも 0.5 にはならず、Dice.m を除いて多くの指標の平均値は 0.5 以下になっている (Rel.r, Rel.c, Rel.a, Dice, Jaccard, Cosine)。一方、Simple matching の平均値は 0.5以上になる。

行の相対頻度 (Rel.r = $a/(a+b)$) の平均値が 0.5 以下になる理由は下の表を見るとわかる。a は1つのセルであるが、b は3個のセルなので、 $a/(a+b)$ は $1/(1+3) = 1/4$ になる。確かに先の図の Rel.r の平均値の位置は 0.25 の近傍になっている。同様にして、Rel.c の平均値は $a/(a+c)=1/(1+2)=1/3$ になる。Rel.a の平均値は $a/(a+b+c+d) = 1/12 = 0.083$ になる。Dice, Jaccard, Cosineの平均値も同様にして計算できる。

*	1	2	3	4
1	a	b	b	b
2	c	d	d	d
3	c	d	d	d

Simple matching の平均値が 0.5 以上になる理由は次の通りである。a(X:Y) のセルは1個であるが、d (!X:!Y) のセルは、上の表のように6個である。そして、b (X: !Y)のセルは3個、c (!X)のセルは2個である。よって、Simple matching の値は $a+d = 1+6 = 7$ 個のセルの和を12個のセルの和で割って求めるので、平均値が0.5を超えることになる ($7/12 = 0.583$)。

Dice-m の平均値は $m = 3 + 4 - 2 = 5$ となるので、

$$m * a / (m*a+b+c) = 5 / (5+3+2) = 0.5$$

となる。このように、乗数 m を使う Dice.m は、どのような行列でもその平均値は 0.5 になる。

範囲が [-1, 1] である Hamann, Yule, Phi は、先に述べたように、2行(または2列)のデータには使えないので除外する。範囲が[0,1]の係数の中で、Dice-m だけが平均値 = 0.5 となるので、Dice-m を使えば同じ尺度で数値を比較することができる。そして、Dice-m の実際の分布図 (図-20) を見ると大きな振幅が観察されるので、その感度は良いと言える。よって、本研究では Dice.m を採用した。

(4) 対照値

$(x-y) / (x+y)$ で導出される標準的な数値を「対照値」(contrast)と名付ける。対照値の最小値は $x=0$ のときの -1 であり、最大値は $y=0$ のときの 1 である。 $x = y$ のときに中間の 0 となる。 $x=y=0$ のときの対照値は0とする。対照値は僅かの絶対頻度でも大きな数値になることがあるので本研究では $x+y$ が10未満の場合の対照値を0とした。

次の表では各年代と地域(地方)のクロス集計である。最初の表 (表23) が *ahora* の分布であり、次が *ahora* の分布である(表24: *ahora* は *aora*, *a ora* を含む)。

表 23 : *ahora*. CODEA. 絶対頻度

<i>ahora</i>	1.GA	2.NA	3.LE	4.CV	5.AR	6.CN	7.VA	8.EX	9.AN
1200	0	0	0	0	0	0	0	0	0
1300	0	0	0	0	0	0	0	0	0
1400	0	0	0	0	0	0	0	0	0
1500	0	0	1	3	0	9	0	0	5
1600	0	0	3	1	0	36	10	2	2
1700	11	0	1	3	10	16	2	5	1

表 24 : *agora*. CODEA. 絶対頻度

<i>agora</i>	1.GA	2.NA	3.LE	4.CV	5.AR	6.CN	7.VA	8.EX	9.AN
1200	0	3	2	9	1	1	0	1	0
1300	0	11	13	20	35	38	0	11	1
1400	0	6	58	51	46	27	0	3	8
1500	0	1	13	30	21	42	0	3	22
1600	0	0	1	1	0	11	0	2	2
1700	1	0	0	1	0	0	3	0	0

これらの2つの表から次の式で *agora* と *ahora* の対照値 (contrast) を求める。

$$\text{contrast}(\textit{ahora}, \textit{agora}) = (\textit{ahora} - \textit{agora}) / (\textit{ahora} + \textit{agora})$$

表 25 : *ahora - agora*. CODEA. 対照値

N.	1.GA	2.NA	3.LE	4.CV	5.AR	6.CN	7.VA	8.EX	9.AN
1200	0.000	-1.000	-1.000	-1.000	-1.000	-1.000	0.000	-1.000	0.000
1300	0.000	-1.000	-1.000	-1.000	-1.000	-1.000	0.000	-1.000	-1.000
1400	0.000	-1.000	-1.000	-1.000	-1.000	-1.000	0.000	-1.000	-1.000
1500	0.000	-1.000	-0.857	-0.818	-1.000	-0.647	0.000	-1.000	0.185
1600	0.000	0.000	0.500	0.000	0.000	0.532	1.000	0.000	0.000
1700	0.833	0.000	1.000	0.500	1.000	1.000	-0.200	1.000	1.000

対照値は統計分析でしばしば使われている。たとえば、先述の Hamann 係数と Yule 係数はそれぞれ $a+d = x$, $b+c = y$; $a*d = x$, $b*c = y$ とすれば、対照値になる。

$$\begin{aligned} & (a+d-b-c) / (a+b+c+d), [-1, 1] \text{ \# Hamann} \\ & (a*d-b*c) / (a*d+b*c), [-1, 1] \text{ \# Yule} \end{aligned}$$

略号

CDH: *Corpus del Diccionario Histórico de la Lengua Española* (Real Academia Española).

CODEA: *Corpus de documentos Españoles Anteriores a 1800*.

COSER: *Corpus Oral y Sonoro del Español*.

DQ: *Ingenioso hidalgo Don Quijote de la Mancha* (Miguel de Cervantes) 『才知ある郷士ドン・キホーテ・デ・ラ・マンチャ』

LYNEAL: *Letras y Números en Análisis Lingüísticos*.

参考文献

家入葉子・内田充美 (2015) 「15世紀の英語とフランス語の接触：キャクストンの翻訳を通して」
高田博行・渋谷勝己・家入葉子(編)『歴史社会言語学入門』116–133. 東京：大修館.

池田央 (1976) 『統計的方法 I 基礎』東京：新曜社.

石川慎一郎 (2021) 『ベーシックコーパス言語学』東京：ひつじ書房.

エルスパス, シュテファン (2015) 「下からの言語史：19世紀ドイツの『庶民』のことばを中心にして」(佐藤恵訳), 高田博行・渋谷勝己・家入葉子 (編)『歴史社会言語学入門』55–69.
東京：大修館.

上田博人 (1975) 『スペイン語音節の音韻論的考察』東京外国語大学卒業論文.

<https://lecture.ecc.u-tokyo.ac.jp/~cueda/kenkyu/onsei/onsetu.pdf>

上田博人 (2016) 『言語研究のための数値データ分析法』「5.得点」

<https://lecture.ecc.u-tokyo.ac.jp/~cueda/gengo/4-numeros/doc/n5-tokuten.pdf>

牛島信明 (1989) 『反=ドン・キホーテ論』東京：弘文堂.

清水憲男 (1986) 『「ドン・キホーテ」をスペイン語で読む』東京：PHP研究所.

田中克彦 (1999) 『クレオール語と日本語』東京：岩波書店.

寺崎英樹 (2011) 『スペイン語史』東京：大学書林.

中尾俊夫 (1996) 『音韻における通時的普遍—最小変化の原理』東京：リーベル出版.

細川弘明 (1992) 「ピジン・クレオール語」 亀井孝・河野六郎・千野栄一編『言語学大辞典』3,
437–466.

山田善郎〔監修〕(1995) 『中級スペイン文法』東京：白水社.

山田善郎〔監修〕(1996) 『スペインの言語』京都：同朋舎出版.

Alonso, Martín (1986) *Diccionario medieval español*. 2 tomos. Salamanca: Universidad Pontificia de Salamanca.

Alvar, Manuel y Bernard Pottier (1983) *Morfología histórica del español*. Madrid: Gredos.

- Blas Arroyo, José Luis (2021) Apogeo y declive de *ansí* en los Siglos de Oro: nuevos datos desde la sociolingüística histórica, *Boletín de Filología*. vol. 56 no.1 versión On-line
<http://dx.doi.org/10.4067/S0718-93032021000100263>
- Caballero, Justo (1970) *Guía - diccionario del Quijote*. Ciudad de México: Editorial España Errante.
- Cano Aguilar, Rafael (1988) *El español a través de los tiempos*. Madrid: Arco / Libros.
- Cejador Frauca, Julio (1906) *La lengua de Cervantes. Gramática y diccionario de la lengua castellana en El Ingenioso hidalgo don Quijote de la Mancha*. Madrid: Establecimiento Tipográfico de Jaime Ratés.
- Cobarrubias, Sebastián de (1611 / 1977) *Tesoro de la lengua castellana o española*. Madrid: Ediciones Turner.
- Corominas, Joan y José A. Pascual (1980) *Diccionario crítico etimológico castellano e hispánico*. 6 vols. Madrid: Gredos.
- Del Barrio de la Rosa, Florencio (2017) Piezas léxicas y variación morfosintáctica en la historia del español. Tres casos en el español de los Siglos de Oro (1581-1620). En Del Barrio (ed.) *Palabras vocabulario léxico. La lexicología aplicada a la didáctica y a la diacronía*, 251–266. Venezia: Edizioni Ca Foscari.
- Del Barrio de la Rosa, Florencio (2018) *Espacio variacional y cambio lingüístico en español*. Madrid: Visor Libros.
- Dressler, Wolfgang U. (2003) Naturalness and morphological change. In Joseph, B. D. and R. D. Janda (eds.) *The handbook of historical linguistics*, 461–471. Oxford: Blackwell.
- Echeverría Reyes, Aníbal. (1931) Vocabulario de *El Ingenioso Hidalgo don Quijote de la Mancha*, compuesto por Miguel Cervantes y Saavedra. *Anales de la Universidad de Chile*, núm. 2, 409–491.
- Fernández Gómez, Carlos (1962) *Vocabulario de Cervantes*. Madrid: Real Academia Española.
- Frago, Juan Antonio (2015) *Don Quijote, lengua y sociedad*. Madrid: Arco / Libros.
- Gui, G. R. (2003) Variationist approaches to phonological change. In Joseph, B. D. and R. D. Janda (eds.) *The handbook of historical linguistics*, 369–400. Oxford: Blackwell.
- Hanssen, Federico (1916) *Gramática histórica de la lengua castellana*. París: Ediciones Hispano-Americanas.
- Harzfeld, Helmut (1972) *El «Quijote» como obra de arte del lenguaje*. 2a ed. Madrid: C.S.I.C.
- Labov, William (1994) *Principles of linguistic change. Vol. 1. Internal factors*. Oxford: Blackwell.
- Labov, William (2001) *Principles of linguistic change. Vol. 2. Social factors*. Oxford: Blackwell.
- Lapesa, Rafael (1981) *Historia de la lengua española*. 9a ed. Madrid: Gredos.
- Lope Blanch, Juan M. (ed.) (1969) *Juan de Valdés, Diálogo de la lengua*. Madrid: Castalia.
- Medina Morales, F. (2005) *La lengua del Siglo de Oro. Un estudio de variación lingüística*. Granada: Editorial Universidad de Granada.
- Menéndez Pidal, Ramón (1968) *Manual de gramática histórica española*. 13 ed. Madrid: Espasa-Calpe.
- Penny, Ralph (2006) *Gramática histórica del español*. Barcelona: Ariel.
- R Core Team (2021) *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.
<https://www.R-project.org/>
- Real Academia Española (1726-1739) *Diccionario de autoridades*. Madrid: Real Academia Española.
<https://apps2.rae.es/DA.html>
- Real Academia Española (2013) *Corpus del Diccionario histórico de la lengua española (CDH)* [en línea].
<https://apps.rae.es/CNDHE>
- Rico, Francisco (2010) *Miguel de Cervantes, Don Quijote de la Mancha, edición, notas y anexos*. Madrid: Punto de lectura.

- Rodríguez Molina, J. (2015) El adverbio *así* en español medieval: variantes morfonéticas. En García Martín, J. *Actas del IX Congreso Internacional de Historia de la Lengua Española* (Cádiz, 2012), vol. 1, 1049–1064. Madrid / Frankfurt am Main: Iberoamericana / Vervuert.
- Romesburg, H. Charles (1989, 1992) *Cluster analysis for researchers*. Florida: E. Krieger Publishing Company. 西田英郎・佐藤嗣二訳『実例クラスター分析』東京：内田老鶴傳圃.
- Sánchez-Prieto Borja, Pedro (2012) Desarrollo y explotación del Corpus de documentos españoles anteriores a 1700 (CODEA). *Scriptum Digital* 1: 5–35.
- Segura Mungia, Santiago (2007) *Diccionario por raíces del latín y de las voces derivadas*. Bilbao: Universidad de Deusto.
- Sevilla Arroyo, Florencio (2005) (ed.) *Miguel de Cervantes Saavedra, Don Quijote de la Mancha*. Madrid: SIAL Ediciones.
- Torrens Álvarez, María Jesús (2018) *Evolución e historia de la lengua española*. 2a. ed. Madrid: Arco / Libros.
- Trask, R., L. (2000) *The dictionary of historical and comparative linguistics*. Edinburgh: Edinburgh University Press.
- Tuten, Donald N. (2003) *Koineization in medieval Spanish*. Berlin: Mouton de Gruyter.
- Ueda, Hiroto (2022) Cambios históricos frecuentes de la ortografía léxica. Análisis de los documentos castellanos del corpus CODEA. *VII Congreso Internacional CHARTA. Universidad de Granada, 8-10, junio de 2022*.
<https://lecture.ecc.u-tokyo.ac.jp/~cueda/kenkyu/rekisi/periodiza/periodiza.pdf>
- Weinreich, U., W. Labov, and M. Herzog (1968) A theory of language change. In Lehmann, W. and Y. Malkiel (eds.) *Directions for historical linguistics*, 97–195. Austin: University of Texas Press.

出版情報

採用決定日：2022年8月31日

Research article

Stress and word order in Asian languages with reference to the *Linguistic Atlas of Asia*

TOKIZAKI, Hisao
Sapporo University

Abstract: This paper gives an overview of the *Linguistic Atlas of Asia* (LAA) (Endo et al. 2021) and examines the hypothesis of stress-order correlation (Tokizaki 2011 et seq.) based on its descriptions of Asian languages. Generalizing the idea of holistic typology by Bally (1944) (German vs. French) and Donegan and Stampe (1983) (Munda vs. Mon-Khmer), Tokizaki (2011 et seq.) proposes the hypothesis that word-stress location correlates with word order (head-initial/head-final) in the world's languages: languages with word-initial stress have head-final order (OV, postposition, suffixing, modifier-noun) while languages with word-final order have head-initial order (e.g. VO, preposition, prefixing, noun-modifier). In this paper, I examine the word-stress location in each group of Asian languages described in Chapter 7 on tone and accent in LAA. Analyzing more descriptions of stress in literature other than LAA and the word order data in Dryer (2013a, b, c, d, e, f, g) in the *World Atlas of Linguistic Structure Online* (WALS), I argue that the stress-order correlation hypothesis generally holds in Asian languages. It is also argued that word-stress location correlates with word order because the word-stress pattern is projected onto the phrase-stress pattern where the main stress universally falls on the complement rather than the head (e.g. on O rather than on V in a verb phrase) (cf. Cinque 1993).*

Keywords: *Linguistic Atlas of Asia*; word stress; word order; WALS; syntax-phonology interface

1. Introduction

The purpose of this paper is to show that word-stress location (initial/final) correlates with word order (head-final/head-initial) in Asian languages as well as other languages of the world, with reference to the *Linguistic Atlas of Asia* (LAA) (Endo et al. 2021) and the *World Atlas of Linguistic Structure Online* (WALS) (Haspelmath and Dryer 2013).

TOKIZAKI, Hisao. 2022. Stress and word order in Asian languages with reference to the *Linguistic Atlas of Asia*. *Studies in Geolinguistics* 2: 102–121. doi: <https://doi.org/10.5281/zenodo.7121553>

* This paper was presented at the 5th meeting of a ILCAA Joint Research Project “Exploration into the Mechanism of Language Change and Variation through the Dialogue between Theoretical Linguistics, Linguistic Typology, and Quantitative Linguistics (Online, March 8, 2022). I would like to thank Yoshiki Ogawa for giving me the chance to present my research. Thanks also go to Hituzi Syobo for giving me the digital data of the *Linguistic Atlas of Asia*. I am grateful to Mitsuaki Endo and Satoko Shirai for the original data in the *Linguistic Atlas of Asia*. This work was supported by KAKENHI 20H01269.

In section 2, I briefly review two linguistic atlases WALS and LAA. Section 3 illustrates Chapter 7 on tone and accent in LAA and discusses the stress-order correlation hypothesis. Section 4 investigates the stress system of each Asian language group and examine the correlation between word-stress location and word order. In section 5, I briefly explain why word-stress location correlates with word order based on the minimalist program of linguistic theory (Chomsky 1995, 2012). Section 6 concludes the discussion.

2. Linguistic Atlases: WALS and LAA

2.1. *World Atlas of Linguistic Structure*

Publication of the *World Atlas of Linguistic Structure* (WALS) (Haspelmath et al. 2005) made much impact on the study of linguistic typology. The maps in WALS provide an excellent overview of the features of the world's languages. The accompanying CD-ROM, which contains data and interactive software, was also useful. In addition to this book, the online version of WALS was published in 2008 (Dryer and Haspelmath 2008), which greatly enhanced the usability of the maps and the data. Revised versions of WALS with minor corrections have been published periodically in 2011, 2013, 2014 and 2020.

2.2. *Linguistic Atlas of Asia*

A recently published atlas, the *Linguistic Atlas of Asia* (Endo et al. 2021, henceforth LAA) focuses on the languages in Asia. LAA is a box set consisting of a book and four sheets of A1-size maps (double-sided and folded to A5 size), which provide detailed descriptions of linguistic features. The book contains eight chapters, the first five of which discuss the lexical items 'sun,' 'rice plant,' 'milk,' 'wind,' and 'iron.' The other three chapters discuss word order (Ch. 6), prosodic features (Ch. 7) and sentence structure (Ch. 8). Each chapter starts with an overview of the item in Asia written by the author shown in (1).

- (1) Introduction (ENDO Mitsuaki)
 - Chapter 1: sun (FUKUSHIMA Chitsuko)
 - Chapter 2: rice plant (SUZUKI Hiroyuki)
 - Chapter 3: milk (EBIHARA Shiho)
 - Chapter 4: wind (SAITÔ Yoshio)
 - Chapter 5: iron (TAGUCHI Yoshihisa)
 - Chapter 6: numeric quantification (KURABE Keita)
 - Chapter 7: tone and accent (IWATA Ray)
 - Chapter 8: it rains (SHIRAI Satoko)

Each chapter also contains 15 articles written by leading researchers on each language or language group, as shown in (2) where the number following the name shows the chapter.

- (2) i. Nivkh (SHIRAIISHI Hidetoshi)
- ii. Ainu (FUKAZAWA Mika)
- iii. Japanese (KISHIE Shinsuke)
- iv. Korean (FUKUI Rei)
- v. Sinitic (UEYA Takashi [6], YAGI Kenji [7], SUZUKI Fumiki 4, 5, [8])
- vi. Hmong-Mien (TAGUCHI Yoshihisa)
- vii. Kra-Dai (ENDO Mitsuaki)
- viii. Tibeto-Burman (SHIRAI Satoko, KURABE Keita, IWASA Kazue, SUZUKI Hiroyuki, EBIHARA Shiho)
- ix. Austroasiatic (KONDO Mika 1, 4, 8, SHIMIZU Masaaki 2, 5, 7, MINEGISHI Makoto 3, 6)
- x. Austronesian (UTSUMI Atsuko)
- xi. Tungusic (MATSUMOTO Ryo) [with Uralic in Ch. 8]
- xii. Uralic (MATSUMOTO Ryo) [with Tungusic in Ch. 8]
- xiii. Mongolic and Turkic (SAITÔ Yoshio)
- xiv. South Asia (YOSHIOKA Noboru)
- xv. Arabic (NAGATO Youichi) [Semitic instead of Arabic in Ch. 4 and 5]

These articles contain a large number of linguistic maps in black and white.¹ The eight color maps in the LAA box correspond to Chapters 1 to 8 in (1). LAA has fifty pages of bibliography, classified by language (group), which give detailed information on the literature cited in the chapters and the primary sources for the maps.

This book is a great contribution to the study of geolinguistics in Asia. Although the range of topics discussed in LAA is not as wide as WALS, the number of data points in Asia described in LAA is far more than that in WALS. For example, Ainu is described as having a simple tone system as a language in WALS (feature 13) while it is reported to have three accent patterns according to dialects in LAA (pitch accent, quantitative accent and none, p. 239).

Below, I will focus on Chapter 7 of LAA, on tone and accent. I will discuss whether the data and descriptions in LAA support the hypothesis in Tokizaki (2019), which

¹ Most of the articles in LAA are based on those in the *Studies in Asian Geolinguistics I to VIII* published electronically by ILCAA (<https://publication.aa-ken.jp>). These original papers include color versions of the maps in LAA.

claims that the stress location in words correlates with the word order of head and complement in the languages of the world.

3. Stress and word order in Asia

3.1. Correlation between word stress and word order

I have been working on the hypothesis that prosody and word order correlate in the languages of the world. This is an extension of holistic typology by Bally (1944) and Donegan and Stampe (1983). Bally compared German and French, and Donegan and Stampe compared Munda and Mon-Khmer in Austroasiatic; they argue that the former has initial stress and head-final order while the latter has final stress and head-initial order. Applying this stress-order correlation to the world's languages in general, Tokizaki (2011) and Tokizaki and Kuwana (2013) argue that languages with word-initial stress tend to have head-final order such as OV while languages with (pen)ultimate stress tend to have head-initial order such as VO.

Tokizaki (2019) investigates Altaic languages (Turkic, Mongolic and Tungusic), which seem to be counterexamples to the hypothesis: most of them have head-final order while a large number of them have been reported to have (pen)ultimate stress (Goedemans and van der Hulst (2005a, b), Schiering and van der Hulst (2010) and Goedemans, Heinz and van der Hulst (2014)). Tokizaki (2019) argues that Altaic languages in fact have word-initial stress and optional word-final pitch accent, the latter of which is sometimes mistaken for word stress/accent. It is argued that Altaic languages have double accent (initial stress and final pitch accent). The idea of word-initial stress is supported by a number of facts: vowel reduction in non-initial syllables, vowel harmony with the initial syllable, alliteration, more variety of vowels in the word-initial position, and emphatic stress/lengthening, emphatic reduplication and gemination of the initial syllable. Thus, Altaic languages are not counterexamples to the generalization that languages with head-final order have word-initial stress.

Tokizaki (2019) focuses on Altaic languages, and does not discuss other languages in Asia in detail. In the next section, I examine the data in Chapter 7 of LAA on tone and accent, and argue that the analysis in Tokizaki (2019) holds true in Asian languages in general.

3.2. Tone and accent in Asia

LAA discusses one phonological aspect of languages, tone and pitch accent, in Chapter 7. The overview is written by IWATA Ray, a leading researcher of Chinese geolinguistics (pp. 232–236). He proposes a division between an inner group and an outer group in order to describe the distribution of suprasegmental features in Asia.

Basically, the inner group is tonal while the outer group is accentual. The map is shown in Fig. 1 (Figure 7.1.1, p. 236).

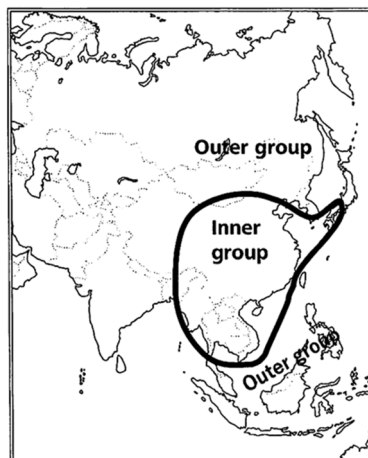


Fig. 1 Dichotomy of Prosodic Features in Asia (Iwata 2019: 236)

One of the difficulties in the study of accent is the use of the terms. Iwata points out that “the phonetic substance of this term [accent (HT)] varies depending on the authors, probably reflecting the research tradition of each language area” (p. 232).

Prosody including tone and accent shows a wide variation in Asian languages. It is understandable that each paper in this chapter describes the language (group) according to the tradition of the previous studies. If we look at the supplementary map of tone and accent in LAA, we recognize how different the prosody of languages in Asia is.

The eight maps in LAA use different colors to show language groups, e.g. red for Sinitic and dark blue for Uralic. In Map 7, the shape of marks distinguishes prosodic properties such as stress/accent location and the number of phonological tones. However, interpreting what these marks are intended to show is not always straightforward. For example, accent on the penultimate syllable is expressed by a semicircle (upper half) in Austronesian and by an outlined circle in Uralic. The semicircle (upper half) mark used for penult in Austronesian is also used for a variety of simple accent-patterns of Keihan type in Japanese and for 2 to 4 tones in Korean. It would be more user-friendly if all the authors followed the same guideline for using shapes.

Chapter 7 in LAA discusses tone and accent in a wide area of Asia where the languages shown in (2) are spoken. The accompanying map (Map 7) consists of the wide area map, the enlarged map and their legends, which are shown in Figures 2 to 5.

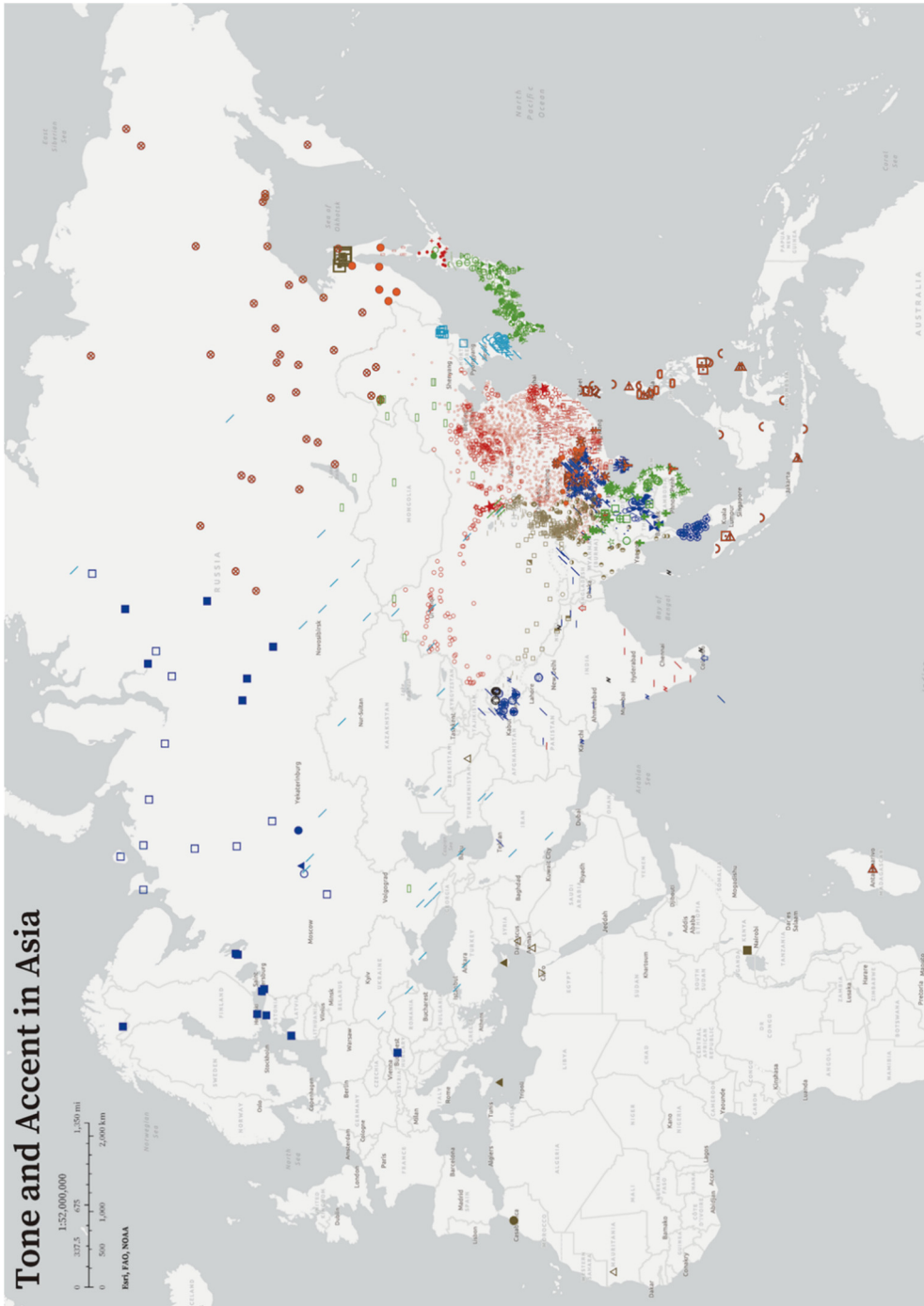


Fig. 2 Tone and Accent in Asia

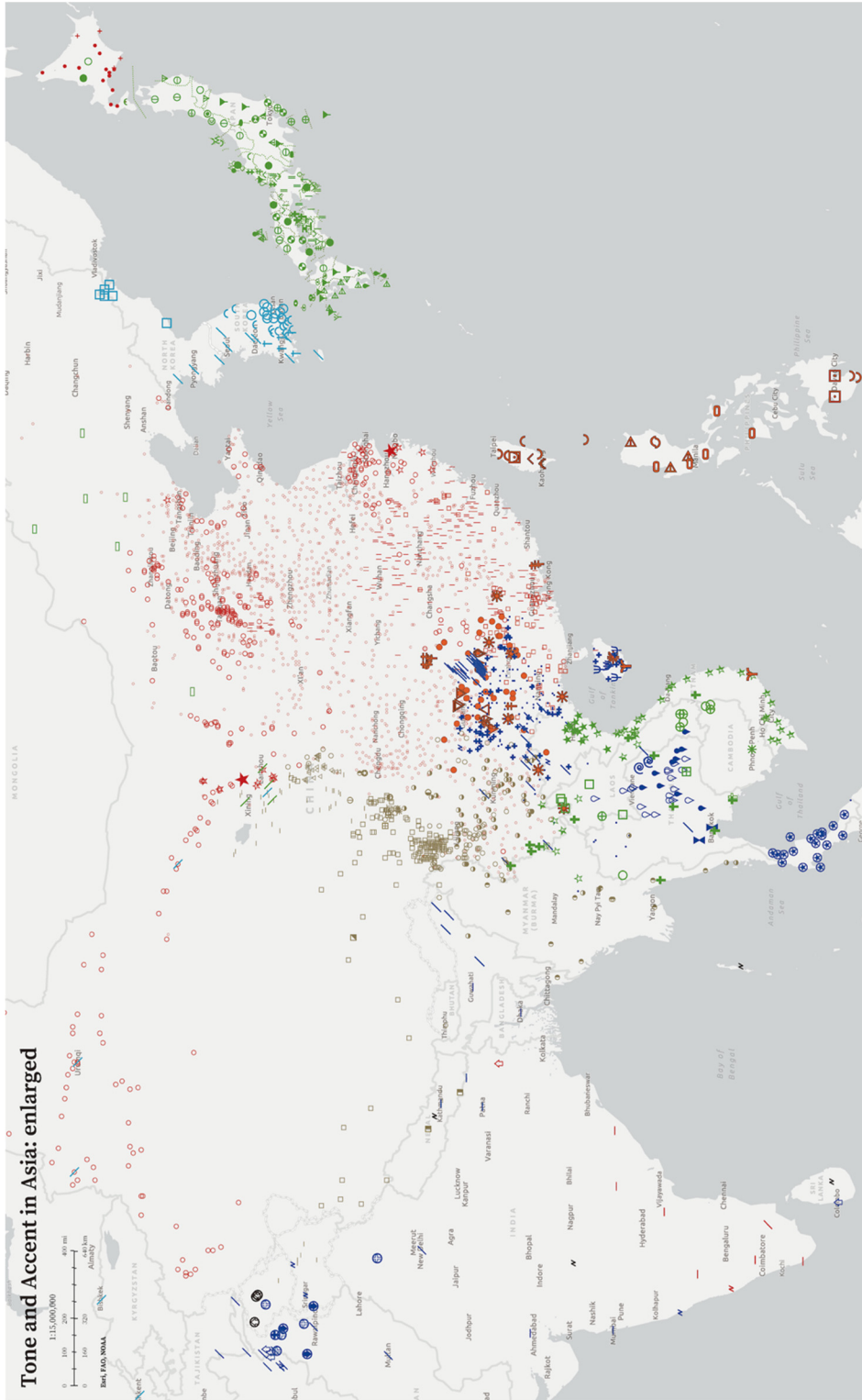


Fig. 3 Tone and Accent in Asia: enlarged

Nivkh	●	A variety of simple accent-patterns of Tokyo Type (Variation III)	⊞	Accent-pattern which has a vestige of Ancient Keihan Type	▶	Accent-pattern similar to Keihan Type (Variation V)
54%	◊	Ambiguous accent-pattern similar to Tokyo Type (Variation I)	>	Simple accent-patterns of Keihan Type (Variation I)	◀	Ambiguous accent-pattern similar to Keihan Type (Variation V)
67-73%	◊	Accent-pattern similar to Tokyo Type (Variation II)	<	Simple accent-patterns of Keihan Type (Variation II)		
90%	●	Accent-pattern similar to Tokyo Type (Variation III)	<	A variety of Simple accent-patterns of Keihan Type (Variation II)		
Ainu	●	A pitch accent				
	◻	B. quantitative accent (vowel length)				
	+	C. none				
Japanese	●	Inner Tokyo Type	▶	One pattern accent (the first mora high and the following ones low)	⊞	A variety of Simple accent-patterns of Keihan Type (Variation II)
	◊	Center Tokyo Type	▶	One pattern accent (the first mora high and the succeeding particle also high)	⊞	Simple accent-patterns of Keihan Type (Variation III)
	⊕	Center Tokyo Type (Variation I)	▶	One pattern accent (the last mora high and the succeeding particle low)	⊞	A variety of Simple accent-patterns of Keihan Type (Variation II)
	◊	Center Tokyo Type (Ambiguous version)	▶	Medival-Modern Keihan Type	⊞	Accent-pattern similar to Keihan Type (Variation II)
	◊	Outer Tokyo Type	▶	Keihan Type	⊞	A variety of accent-pattern similar to Keihan Type (Variation II)
	◊	A variety of Outer Tokyo Type (Variation I)	▶	Keihan Type (partially Tokyo Type)	⊞	Simple accent-pattern similar to Keihan Type
	◊	A variety of Outer Tokyo Type (Variation II)	▶	A variety of Keihan Type (partially Tokyo Type)	⊞	Simple accent-pattern similar to Keihan Type
	◊	A variety of Outer Tokyo Type (Variation I, Ambiguous version)	▶	Keihan Type (partially Tokyo Type)	⊞	Simple accent-pattern similar to Keihan Type
	◊	A variety of Unique Tokyo Type	▶	Keihan Type	⊞	Simple accent-pattern similar to Keihan Type
	◊	Simple accent-patterns of Tokyo Type (Variation I)	▶	Keihan Type (partially Tokyo Type)	⊞	Simple accent-pattern similar to Keihan Type
	●	Simple accent-patterns of Tokyo Type (Variation II)	▶	Keihan Type	⊞	Simple accent-pattern similar to Keihan Type
	◊	Simple accent-patterns of Tokyo Type (Variation III)	▶	Keihan Type	⊞	Simple accent-pattern similar to Keihan Type

Fig. 4 Tone and Accent in Asia, legend I

	Austronesian	Mongolic	Others
△ NT/NA+YC	phonemic length contrast, stress/accnt falling on the last long vowel	Initially-accented	E. Not written
* PA+NP	stress/accnt falling regularly on ultima	Finally-accented	A-1. 2-way Pitch
▢ WT+SA+NP	stress/accnt falling regularly on penult		
▣ WT+YT	stress/accnt falling regularly on antepenult		
● NT/NA+ST+YL	stress/accnt falling on penult, but influence by PAN *e (present day predominantly schewa)	Turkic	Arabic
○ PA+ST+YT	phonomic accent (secondarily introduced)	Finally-accented	A-1. madrasa type
○ ST+SA+NP	phonomic tone		A-2. madrasa type
○ ST+SA+YL			A-3. partial lexicalized
▣ WT+RG+YC			B. Maghreb type
(NP: no phonation pattern, NT/NA: no tone or accent, PA: pitch accent, RG: register, SA: stress accent, ST: syllabic tone, WT: word tone, YC: phonation influenced by initial consonants, YL: distinctive phonation, YT: phonetic phonation)			C. tone type
Austroasiatic			
○ A. Voiced/voiceless contrast in initial consonants			
□ B. Long/short contrast in vowels			
⊕ C. Initial contrast and registers			
⊕ D. Vowel contrast and registers			
⊕ E. 2-way contrast in register			
⊕ F. 4-way contrast in register			
* G. Restructuring of vowels			
☆ H. 2-way contrast in tone			
☆ I. 3-way contrast in tone			
☆ J. 4-way contrast in tone			
☆ K. 5-way contrast in tone			
☆ L. 6-way contrast in tone			
Tibeto-Burman			
□ WT+NP			
○ ST+NP			
○ ST+YL			
- NT/NA+NP			
△ RG+YC			
○ ST+YT			
○ Type 2-d. Tone *A, *C, and *D split			
● Type 2-e. All the proto-tones split through devoicing			
* Type 3. Tone-split 1 and Tone-split 2 are observed			
Kra-Dai			
ψ A. Proto type			
+ B. Split caused by consonants			
· C. Split of tone D			
/ D. Split conditioned by aspirated initial consonants			
⊕ E. Merger between A ^v and B ^v , A ^l and B ^l			
○ F. C ^l and C ^v are identical			
⊕ G. B shows no distinction and C ^v =C ^l			
⊕ H. Bangkok dialect			
⊕ I. Sakon Nakhon and Wanomiwat dialects			
⊕ J. others			
Indo-European			
/ C. Stress			
- D. Not identified			
⊕ E. Not written			
○ A-1. 2-way Pitch			
⊕ B-1. 3-way Tone			
⊕ A-2. 3-way Pitch			
⊕ B-2. 5-way Tone			
Dravidian			
- D. Not identified			
⊕ E. Not written			
⊕ A-1. 2-way Pitch			
/ C. Stress			

Fig. 5 Tone and Accent in Asia, legend 2

As can be seen in Figs. 2 to 5, the prosody of Asian languages is described in detail. A problem for a typology of prosody is that the prosodic patterns and their descriptions are quite different from language to language. For example, Ainu is described as having pitch accent, quantitative accent or none, while the description of Sinitic is about the number of phonological tones (1 to 6).

In this paper, I focus on the location of word-stress/accent. However, the maps and legends in Figs. 2 to 5 show information on word-stress/accent only in some language groups such as Austronesian, Tungusic, Uralic, Mongolic, Turkic and Arabic. I will refer to other studies than LAA for more data and descriptions of the languages in LAA that lack descriptions about stress.

4. Stress and word order in Asian languages

In this section, I investigate the correlation between word-stress location and word order in each language (groups) shown in (2).

4.1. Nivkh

Nivkh is described as having initial stress and head-final order in WALS. In LAA, Shiraishi (pp. 237-238) shows the durational ratio of the word-initial vowel (stressed) and the second vowel (unstressed) in Amur and Sakhalin dialects (54% to 90 %) on the map. Thus, Nivkh data support the hypothesis that languages with initial stress have head-final order (Tokizaki 2011, 2019).

4.2. Ainu

Ainu has head-final word order. WALS and its source StressTyp2 (Goedemans et al. 2014) make no mention about stress in Ainu. Fukazawa in LAA (pp. 239-240) reports that the majority of Hokkaido dialects have pitch accent on the first or the second syllable while Sakhalin dialects have quantitative accent on a long vowel. She points out that “a long vowel on the open first syllable in Sakhalin often corresponds to a high pitch on the open first syllable in Hokkaido” (e.g. *mīina* (Sakhalin) vs. *mīna* (Hokkaido)). Assuming that vowel length is generally related to stress, I claim that Ainu dialects have word-initial stress/accent rather than word-final stress/accent. Then, Ainu also supports the hypothesis that languages with initial stress have head-final order.

4.3. Japanese

Japanese is a typical head-final language. Its prosody is problematic for the idea that head-final languages have word-initial stress: it has been argued that (Tokyo) Japanese

has no stress but pitch accent on the antepenult mora. In Tokizaki (2022), I argue that a content word in Japanese has the stress on its first mora and an optional pitch-fall accent near its end. If this analysis is on the right track, Japanese also confirms the hypothesis that head-final languages have word-initial stress.

4.4. Korean

Korean is also problematic for the hypothesis that head-final languages have word-initial stress. The nature of stress or accent in Korean has been controversial. Lee, H.-B. (1989), Lee, H.-Y. (1990) and Song (2005) argue that Korean words have stress on the first or the second syllable depending on syllable weight, while de Jong (1994), Jun (1995, 2005) and Lim (2001) argue that Korean stress is not a word-level but a phrase-level stress assigned to the initial syllable in a prosodic phrase. Fukui in LAA (pp. 246–247) does not mention stress in Korean. In Tokizaki (2018), I argue that Korean has both word-level stress and phrase-level stress. I argue that word-level stress is projected to phrase-level stress, which is also in the initial position in the phrase. If this analysis is on the right track, Korean also confirms the hypothesis that head-final languages have word-initial stress.

4.5. Sinitic

Sinitic languages and dialects show interesting geographical distribution. Yagi in LAA (pp. 249–253) discusses the number of tones in Sinitic dialects but not stress and accent. Iwata (2017: 67) draws a map of the tone sandhi types in Chinese dialects, which indicates the location of accent (initial/final), as shown in Fig. 6.

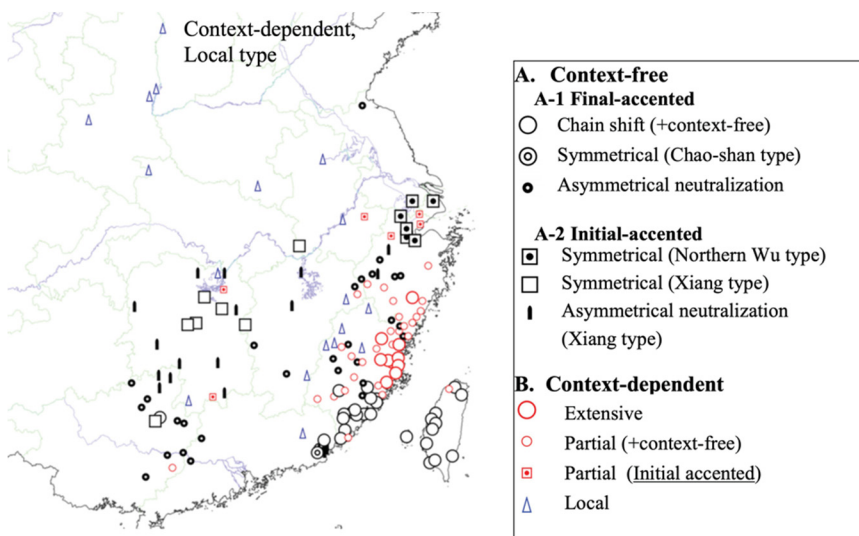


Fig. 6 The tone sandhi types in China (Map 1 in Iwata 2017: 67)

This map shows that the tone sandhi in China can be grouped into initial-accented tone sandhi (marked with a square) in the north-west and final-accented tone sandhi (marked with a circle) in the south-east. If we assume that the accented position is the location of stress, we predict that north-west dialects with initial-accent have head-final word order while south-east dialects with final-accent have head-initial order. This prediction is partially borne out by the data from Hashimoto (1978), which show that the northern dialects have some head-final orders (e.g. modifier-head) while the southern dialects have head-initial orders (head-modifier).

4.6. Hmong-Mien

Hmong-Mien languages have head-initial order (e.g. VO & preposition in Hmong Njua) and head-final order (e.g. genitive-noun) according to Dryer (2013a, b, c, d, e, f, g). In WALS, Maddieson (2013) reports that Mien and Hmong Njua have a complex tone system, while there is no information about stress in Goedemans and van der Hulst (2013a, b). Taguchi in LAA (pp. 254–255) discusses tonal types only. Ratliff (2017) points out that the sesquisyllables in Hmong-Mien, as an areal feature relating to Mon-Khmer and Tibeto-Burman, appear in the word-initial syllable. A sesquisyllable makes an iambic foot with the second, full syllable. Then, the head-initial order in Hmong-Mien also supports the idea that head-initial order matches word-final stress. Of course we need to explain why some Hmong-Mien languages also have head-final order in the genitive-noun and numeral-noun constructions. It is interesting that some Hmong-Mien languages also have head-initial order in NP (noun-adjective and noun-demonstrative) while Sinitic dialects typically have head-final NPs (adjective-noun and demonstrative-noun). It seems that mixed word order in NPs in Hmong-Mien shows a gradation from head-final order in the north to head-initial order in the south.

4.7. Kra-Dai

The prosody of Kra-Dai languages is described by Endo in LAA (pp. 256–259). He discusses the distribution of tonal patterns, but not stress in the languages. Peyasantiwong (1986) argues that Tai has stress on the word-final syllable. If we assume that Kra-Dai languages in general have word-final stress, we can explain the fact that the word orders in Kra-Dai are head-initial in most noun phrases as well as in verb phrases. It is also interesting to note that northern dialects (Dong (Southern), Zhuang (Northern), Gelao, Yay, Nung (in Vietnam) and Hlai (Baoding)) have numeral-noun order while southern dialects (Lao and Thai) have noun-numeral order.

4.8. Tibeto-Burman

Tibeto-Burman prosody is described in LAA by Shirai et al. (pp. 260–266). Their description shows a wide variety of tone, accent and phonation in the languages, but

does not show much about stress location. Caplow (2016a) argues that Balti has stress on the second syllable of nouns and stress on the first syllable of verbs. Caplow (2016b) also argues, based on the data of Balti and Rebkong Amdo, that Proto-Tibetan had the same stress pattern. According to the data in Dryer (2013a, b, c, d, e, f, g) most Tibeto-Burman languages have head-final order (OV, postpositional phrase and genitive-noun, suffixing) while many of them also have head-initial order (noun-adjective, noun-demonstrative and noun-numeral).² It might be the case that head-final orders in verb phrases and postpositional phrases relate to the stress on the first syllable in verbs while head-initial orders in noun phrases relate to the stress on the second syllable in nouns. Of course, this is a speculation, which needs to be proved with more research.

It is also promising to investigate the position of neutral tone (or light tone) or sesquisyllable (minor syllable) in a word. Satoko Shirai (personal communication) points out the fact that central Tibetic languages tend to have neutral tone in the word-final position as Mandarin Chinese does, while Burmese has developed sesquisyllables in the word-initial position due to language contact with the Mon language (cf. Matisoff 1973).

4.9. Austroasiatic

Austroasiatic prosody is described by Shimizu in LAA (pp. 267–268), which discusses voiced/voiceless contrast in initial consonants, long/short contrast in vowels, register contrasts, restructuring of vowels and tonal contrasts, but not stress. He does not discuss Munda in this chapter on tone and accent. Donegan and Stampe (1983) is one of the pioneering works investigating the correlation between linguistic rhythm and word order (cf. Bally 1944 for German and French). They compare Munda and Mon-Khmer and argue that Munda has head-final order and initial stress while Mon-Khmer has head-initial order and final stress (cf. Anderson 2015a, b for the idea that Munda has word-final stress and phrase-initial stress). As Dryer (2013a, b, c, d, e, f, g) shows, Munda languages have OV, postposition and noun-final order in NPs (genitive-noun, adjective-noun, demonstrative-noun and numeral-noun) while other genera in Austroasiatic have VO, preposition and noun-initial order in most NPs (noun-genitive, noun-adjective, noun-demonstrative³) except for numeral-noun.⁴ Then, Austroasiatic shows a clear contrast between head-final order & initial-stress and head-initial order & final-stress.

² Karenic languages in the south of the region have preposition rather than postposition.

³ Nicobarese and Khasian have demonstrative-noun order. Mlabri (Minor) has genitive-noun order, and Nicobarese (Car) has adjective-noun order.

⁴ The order of numeral and noun might be problematic because numerals must occur with classifiers in many languages (Dryer 2013g). We need to investigate the structure of noun phrases with classifiers in more detail. Chapter 6 in LAA on numeric quantification gives us much information about the word-order typology.

4.10. Austronesian

The stress/accent of Austronesian is described in detail by Utsumi in LAA (pp. 269–271). She discusses stress/accent in Proto-Austronesian, Formosan and Malayo-Polynesian, Philippine and Indonesian, and concludes that “original Proto-Austronesian stress pattern is retained resulting in the placement of either on the final or the penult syllable” (p. 270). According to Goedemans and van der Hulst (2013a, b) in WALS, Austronesian languages have the following stress locations (the numbers show the number of languages): initial 4, second 1, antepenult 4, penult 62, ultimate 7; left-edge 1, right-oriented 3, right-edge 29. This shows that most Austronesian languages have final stress rather than initial stress.

The word order in most Austronesian languages is head-initial: VO, preposition, noun-genitive 112 (GN 40), noun-adjective and noun-demonstrative. Thus, Austronesian languages support the hypothesis that final-stress languages have head-initial order.

4.11. Tungusic

Matsumoto in LAA (pp. 272–273) reports that “[t]he accent types of all Tungusic languages are classified as type B, in which accent is basically on the last syllable of the word.” One might think that this final-accent is a kind of final-stress. However, he also mentions that “[i]t is often mentioned that the first syllable is pronounced with a relatively strong aspiration and that the last syllable is accompanied with a musical intonation (pitch accent).” These observations match the idea of double accent (initial stress plus final pitch accent) in Altaic proposed by Tokizaki (2019). If we admit word-initial stress in Tungusic (and Altaic in general), we predict that the languages have head-final word order. This prediction is borne out: word order in Tungusic (and Altaic) is consistently head-final: OV, postposition, strongly suffixing, genitive-noun, adjective-noun, demonstrative-noun and numeral-noun. Matsumoto claims that “[g]rammatical features such as vowel harmony, word-formation and inflection based on suffixing, should be considered together for classifying the accent type and the historical changes.” This claim seems to be based on the same idea as Tokizaki (2019), which argues that Altaic initial-stress correlates with vowel harmony and head-final orders including suffixing.

4.12. Uralic

Matsumoto in LAA (p. 274) observes that accent basically falls on the first syllable in Uralic. He classifies Uralic accent patterns into five types according to the location of accent: the first syllable (west area), the first or the second (if the first syllable contains *i/u*) syllable (Mordvin (Moksha), unbounded in WALS), the last syllable

(Udmurt, n.d. in WALS), penultimate (Hill Mari, combined in WALS) and any syllable (Meadow Mari, unbounded in WALS). Matsumoto also points out the language contact between Uralic and Turkic/Tungusic, which changed the prosody of Uralic languages.

The word order of Uralic is head-final (OV, postposition, strongly suffixing, GN, AN, DN and NumN) except for VO in Finnish, Estonian, Hungarian, Saami (Northern), Mordvin (Erzya) and Komi-Zyrian (Dryer 2013a, b, c, d, e, f, g).⁵ Thus, if we assume that Uralic languages have (or originally had) initial stress, they support the hypothesis that languages with initial-stress have head-final order.

4.13. Mongolic and Turkic

Saitô in LAA (pp. 275–276) reports that “which syllable is phonetically prominent in Mongolic languages has been a matter of controversy.” Using vowel reduction and deletion as a criterion for an accent, he claims that the majority of the languages (Mongol, Oirad, and so on) have an accent on the initial syllable while some languages (Monguor, Dongxiang, Kangjia, Bao’an and Shira Yughur in the Gansu and Qinghai provinces) have an accent on the final syllable.

As for Turkic accent, Saitô argues that higher pitch is normally associated with the last syllable in native words. He also points out that “[s]tress, although not well predictable, often falls on the first syllable, but can be placed on other syllables.” This is what I call “double accent” in Tokizaki (2019): Altaic languages have stress in the word-initial position and pitch accent in the word-final position.

Then, we can generalize that Mongolic and Turkic as well as Tungusic typically have word-initial stress. Altaic languages (Mongolic and Turkic and Tungusic) share the property of initial stress and head-final order (OV, postposition, strongly suffixing and modifier-N), supporting the stress-order correlation hypothesis.

4.14. South Asia

Yoshioka in LAA (pp. 277–279) describes tone/accent in the languages in South Asia, which seems to correspond to Indic and Dravidian languages in WALS. He classifies the languages into seven types: 2-way pitch (H/not H), 3-way pitch (H/M/L), 3-way tone (Rise/Fall/Even), 5-way tone, stress, not identified well, and accent not written. However, he does not mention the stress location of stress languages in South Asia. Goedemans and van der Hulst (2013a, b) in WALS give us some information about stress location in Indic: initial (Bengali), ultimate (Kalami, Romani (North Russian)), left-edge (Nepali), right-oriented (Bhojpuri, Maithili, Hindi), right-edge (Awadhi, Gujarati), combined (Sindhi). They also describe the stress location of two Dravidian

⁵ Most VO languages in Uralic are located in the west.

language: initial (Koya), left-edge (Malayalam). We cannot draw any firm conclusions about the stress location in these languages.

The data in Dryer (2013a, b, c, d, e, f, g) show that Indic languages are head-final (OV, postposition, strongly suffixing and modifier-N). We need more data about the stress location to confirm the stress-order correlation hypothesis.

4.15. Arabic

Nagato in LAA (pp. 280–281) observes that “[i]n Arabic, in general, stress falls on the penultimate syllable.” He also points out Maghreb type (with stress shift to the ultimate) and tone type (in Kenya and Uganda).

Word order in Arabic is mostly head-initial: VO, preposition, noun-genitive, noun-adjective (but demonstrative-noun and numeral-noun). Then, Arabic also supports the stress-order correlation hypothesis: languages with final (including penultimate) stress have head-initial order.

4.16. Summary

In this section, I have investigated the languages in Asia in order to confirm the stress-order correlation hypothesis. The summary of the discussion is shown in (3).

(3)	<u>languages</u>	<u>stress</u>	<u>word order</u>
i.	Nivkh	initial	head-final
ii.	Ainu	initial	head-final
iii.	Japanese	initial (+final pitch accent)	head-final
iv.	Korean	initial or second	head-final
v.	Sinitic (north-west)	initial	head-final
	Sinitic (south-east)	final	head-initial
vi.	Hmong-Mien	final	head-initial (final in NP)
vii.	Kra-Dai	final?	head-initial (final in Num)
viii.	Tibeto-Burman	initial in V, second in N	head-final (initial in NP)
ix.	Austroasiatic	final	head-initial
	AA (Munda)	initial	head-final
x.	Austronesian	final	head-initial
xi.	Tungusic	initial	head-final
xii.	Uralic	initial	head-final
xiii.	Mongolic & Turkic	initial	head-final
xiv.	South Asia (Indic)	initial/final	head-final
	SA (Dravidian)	initial	head-final
xv.	Arabic	penultimate	head-initial

The result of this survey show the stress-order correlation generally holds in Asian languages: languages with initial stress have head-final order while languages with final stress have head-initial order. Of course we need to investigate the nature of prosody in some language groups (3iii, iv, vii, viii, xiv).

5. Why does stress location correlate with word order?

Finally, let us consider why stress location correlates with word order. Tokizaki (2011 et seq.) argues that a language “chooses” either head-initial or head-final order according to its stress pattern. In the minimalist program of generative grammar, Chomsky (1995, 2012 et seq.) assumes that linear order of constituents is determined at the interface between syntax and phonology. For example, syntax makes a structure (a set) without linear order by merging a head X and its complement YP (a phrase headed by Y, e.g. object, modifier) as in (4).

$$(4) \quad \{X \ YP\}$$

This structure is linearized either head-initial (5a) or (5b).

$$(5) \quad \begin{array}{ll} \text{a.} & [X \ \mathbf{Y}P] \\ \text{b.} & [\mathbf{Y}P \ X] \end{array}$$

In (5), the complement YP receives the main stress in the constituent because YP consists of a head Y and its complement: Cinque (1993) claims that stress falls on the most deeply embedded element in a structure, which is in YP. I assume that the word stress pattern is projected onto the phrasal stress pattern in a language: languages with word-initial stress have phrase-initial stress while languages with word-final stress have phrase-final stress. Then, the phrase-final stress in (5a) is chosen by languages with word-final stress (6a) while the phrase-initial stress in (5b) is chosen by languages with word-initial stress (6b).

$$(6) \quad \begin{array}{ll} \text{a.} & [\text{Word } \sigma \ (\cdot) \ \sigma] \\ \text{b.} & [\text{Word } \sigma \ (\cdot) \ \sigma] \end{array}$$

Thus, languages with word-final stress (6a) have head-initial order with phrase-final stress (5a) while languages with word-initial stress (6b) have head-final order with

phrase-initial stress (5b). I argue that this is why word-stress correlates with word order in languages of the world. For the detail of this idea, see Tokizaki (2011 et seq.) and Tokizaki and Okazaki (2022).

6. Conclusion

In this paper, I have investigated the correlation between word-stress location and word order in Asian languages referring to the *Linguistic Atlas of Asia* (LAA). In section 2, I briefly reviewed WALS and LAA. In section 3, I illustrated Chapter 7 on tone and accent in Asia in LAA. In Section 4, I discussed the stress location and word order in each language group in Asia based on the descriptions in LAA and other data in literature including WALS. In Section 5, I explained the reason why word-stress location is correlated with word order.

This paper is an attempt to prove the hypothesis that languages with initial stress have head-final order (e.g. OV) while languages with final stress have head-initial order (e.g. VO). We have seen that the hypothesis is supported by most language groups in Asia. Of course we need to examine the prosodic system of each language (group), which has not been investigated well. I hope that this study will make a contribution to the study of geolinguistics and linguistic theory.

References

- Anderson, Gregory D.S. (2015a) Overview of the Munda languages. In: Mathias Jenny and Paul Sidwell (eds.) *The Handbook of Austroasiatic languages*, 364–414. Leiden: Brill.
- Anderson, Gregory D.S. (2015b) Prosody, phonological domains and the structure of roots, stems and words in the Munda languages in a comparative/historical light. *Journal of South Asian Languages and Linguistics* 2(2): 163–183. doi: <https://doi.org/10.1515/jsall-2015-0009>
- Bally, Charles (1944) *Linguistique générale et linguistique Française*. 2nd edition. Berne: Francke.
- Caplow, Nancy (2016a) Stress patterns and acoustic correlates of stress in Balti Tibetan. *Himalayan Linguistics* 15(2): 1–49. URI: <https://escholarship.org/uc/item/0k86k32k>
- Caplow, Nancy (2016b) Reconstructing stress in Proto-Tibetan: Evidence from Balti and Rebkong Amdo. *Linguistics of the Tibeto-Burman Area* 39(2): 180–221. doi: <https://doi.org/10.1075/ltba.39.2.01cap>
- Chomsky, Noam (1995) *The minimalist program*. Cambridge, MA: MIT Press.
- Chomsky, Noam (2012) *The science of language: Interviews with James McGilvray*. Cambridge: Cambridge University Press.
- Cinque, Guglielmo (1993) A null theory of phrase and compound stress. *Linguistic Inquiry* 24(2): 239–297.

- de Jong, Ken (1994) Initial tones and prominence in Seoul Korean. In: Sook-hyang Lee and Sun-Ah Jun (eds.) *Ohio State University Working Papers in Linguistics No. 43: Papers from the Linguistics Laboratory*, 1–14.
- Donegan, Patricia Jane and David Stampe (1983) Rhythm and the holistic organization of language structure. *Papers from the parasession on the interplay of phonology, morphology and syntax*, 337–353. Chicago: Chicago Linguistic Society.
- Dryer, Matthew S. (2013a) Prefixing vs. suffixing in inflectional morphology. In: Matthew S. Dryer and Martin Haspelmath (eds.) *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. URI: <http://wals.info/chapter/26>
- Dryer, Matthew S. (2013b) Order of object and verb. In: Matthew S. Dryer and Martin Haspelmath (eds.) *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. URI: <http://wals.info/chapter/83>
- Dryer, Matthew S. (2013c) Order of adposition and noun phrase. In: Matthew S. Dryer and Martin Haspelmath (eds.) *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. URI: <http://wals.info/chapter/85>
- Dryer, Matthew S. (2013d) Order of genitive and noun. In: Matthew S. Dryer and Martin Haspelmath (eds.) *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. URI: <http://wals.info/chapter/86>
- Dryer, Matthew S. (2013e) Order of adjective and noun. In: Matthew S. Dryer and Martin Haspelmath (eds.) *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. URI: <http://wals.info/chapter/87>
- Dryer, Matthew S. (2013f) Order of demonstrative and noun. In: Matthew S. Dryer and Martin Haspelmath (eds.) *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. URI: <http://wals.info/chapter/88>
- Dryer, Matthew S. (2013g) Order of numeral and noun. In: Matthew S. Dryer and Martin Haspelmath (eds.) *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. URI: <http://wals.info/chapter/89>
- Dryer, Matthew S. and Martin Haspelmath (eds.) (2008, 2013) *World Atlas of Linguistic Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. URI: <http://wals.info>
- Endo, Mitsuaki, Makoto Minegishi, Satoko Shirai, Hiroyuki Suzuki and Keita Kurabe (eds.) (2021) *Linguistic atlas of Asia*. Tokyo: Hituzi Syobo.
- Goedemans, Rob. and Harry van der Hulst (2013a). Fixed stress locations. In: Matthew S. Dryer and Martin Haspelmath (eds.) *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. URI: <http://wals.info/chapter/14>
- Goedemans, Rob and Harry van der Hulst (2013b) Weight-sensitive stress. In: Matthew S. Dryer and Martin Haspelmath (eds.) *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. URI: <http://wals.info/chapter/15>
- Goedemans, Rob, Jeffrey Heinz and Harry van der Hulst (2014) StressTyp2. URI: <http://st2.ullet.net>
- Hashimoto, Mantaro [橋本萬太郎] (1978) *Gengo ruikei chiriron* 『言語類型地理論』 [*Linguistic typogeography*]. Tokyo: Kobundo.
- Haspelmath, Martin, Matthew S. Dryer, David Gil and Bernard Comrie (2005) *The World Atlas of Linguistic Structures*. Oxford: Oxford University Press.
- Iwata, Ray (2017) Typology of Chinese tone sandhi and its historical implication. In: Mitsuaki Endo (ed.) *Studies in Asian geolinguistics VII “tone and accent,”* 66–79. Fuchu: Research Institute for Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies. URI: https://publication.aa-ken.jp/sag7_tone_2017.pdf
- Jun, Sun-Ah (1995) A phonetic study of stress in Korean. *Journal of Acoustical Society of America* 98(5-2): 2893.

- Jun, Sun-Ah (2005) Korean intonational phonology and prosodic transcription. In: Sun-Ah Jun (ed.) *Prosodic typology: The phonology of intonation and phrasing*, 201–229. Oxford: Oxford University Press.
- Lee, Hansol H.B. 1989. *Korean grammar*. Oxford: Oxford University Press.
- Lee, Ho-Young. 1990. The structure of Korean prosody. Unpublished doctoral dissertation, University College London.
- Lim, Byung-jin (2000) The role of syllable weight and position on prominence in Korean. *Japanese/Korean Linguistics* 9: 139–150.
- Maddieson, Ian (2013) Tone. In: Matthew S. Dryer and Martin Haspelmath (eds.) *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. URI: <http://wals.info/chapter/13>
- Matisoff, James A. (1973) Tonogenesis in Southeast Asia. In: Larry M. Hyman (ed.) *Consonant types & tone: Southern California occasional papers in linguistics no. 1*, 71–95. Los Angeles, CA: University of Southern California. URI: <https://dornsife.usc.edu/assets/sites/56/docs/SCOPIL1-consonant types and tone.pdf>
- Peyasantiwong, Patcharin (1986) Stress in Thai. In: Robert J. Bieclner, Thomas J. Hudak and Patcharin Peyasantiwong (eds.) *Papers from a conference on Thai studies in honor of William J. Gedney*, 211–230. Ann Arbor, MI: The University of Michigan.
- Ratliff, Martha (2017) Structure of Hmong-Mien Languages, Session #8 Language contact. Slides presented at LSA Institute, University of Kentucky. URI: <https://www.linguisticsociety.org/e-learning/structure-hmong-mien-languages-session-8-language-contact>
- Schiering, René and Harry van der Hulst (2010) Word accent systems in the languages of Asia. In: Harry van der Hulst, Rob Goedemans and Ellen van Zanten (eds.) *A survey of word accentual patterns in the languages of the world*, 509–613. Berlin: Mouton De Gruyter.
- Song, Jae Jung (2005) *The Korean language: Structure, use and context*. London: Routledge.
- Tokizaki, Hisao (2011) The nature of linear information in the morphosyntax-PF interface. *English Linguistics* 28(2): 227–257. URI: https://www.jstage.jst.go.jp/article/elsj/28/2/28_227/_pdf
- Tokizaki, Hisao (2018) Is Korean stress word-level or phrase-level? Poster presented at NINJAL ICPP 2018 (5th NINJAL International Conference on Phonetics and Phonology), Tokyo. October 27, 2018.
- Tokizaki, Hisao (2019) Word stress, pitch accent, and word order typology with special reference to Altaic. In: Rob Goedemans, Jeff Heinz and Harry van der Hulst (eds.) *The study of word stress and accent: Theories, methods and data*, 187–223. Cambridge: Cambridge University Press. doi: <http://doi.org/10.1017/9781316683101.007>
- Tokizaki, Hisao (2022) Stress and pitch accent in Japanese word prosody. In: Hisao Tokizaki (ed.) *Phonological Externalization volume 7*, 27–35. Sapporo: Sapporo University. URI: <http://id.nii.ac.jp/1067/00007905/>
- Tokizaki, Hisao and Yasutomo Kuwana (2013) A stress-based theory of disharmonic word orders. In: Theresa Biberauer and Michelle Sheehan (eds.) *Theoretical approaches to disharmonic word orders*, 190–215. Oxford: Oxford University Press. doi: <https://doi.org/10.1093/acprof:oso/9780199684359.003.0007>
- Tokizaki, Hisao and Masao Okazaki [時崎久夫・岡崎正男] (2022) *On'inron-to tano bun'ya-tono intāfeisu* 『音韻論と他の分野とのインターフェース』 [The interface between phonology and other fields] Tokyo: Kaitakusha.

Publication history

Date accepted: 31 August 2022

Research article

‘Pig’ in Bantu

SHINAGAWA, Daisuke
ILCAA, Tokyo University of Foreign Studies

KOMORI, Junko
Osaka University

Abstract: This brief report provides an overview of geographic distribution of words denoting ‘pig’ in Bantu languages. While reflexes of the Proto-Bantu reconstructed form **-gòdòbè* are widely spread in Eastern Bantu zones, forms traced back to the (shortened) PB form **-gòdó* are dominantly distributed in Western Bantu zones. This article also provides information on several forms that are only locally observed especially in North Eastern languages.

Keywords: ‘pig’, geographical distribution, Bantu languages, Proto-Bantu, Niger-Congo

1. Introduction

The present paper aims to present a brief overview of geographical distribution of words denoting ‘pig’ in present Bantu languages. The data are mostly based on the two online lexical databases, namely the ‘Bantu Lexical Reconstructions 3’ (Bastin *et al.* 2002) and the ‘Tanzanian Language survey’, which is a digital compilation of vocabulary lists of Eastern Bantu languages originally collected by Derek Nurse and Gérard Philippson in the early 70’s (Nurse and Philippson 1975).

2. General distribution of the reconstructed forms

In the list of Proto-Bantu lexicon by Meeussen (1969), a lexical form for ‘pig’ is reconstructed as **-gòdòbè* [BLR-MAIN-1494] (as registered in Bastin *et al.* (2002)), which has two shortened variants with a slight difference of vowel height, namely **-gòdó* [BLR-MAIN-1493] and **-gùdú* [BLR-MAIN-1536].

As a general distributional tendency, these two types show the ‘East vs. West’ distribution pattern. The longer form **-gòdòbè* and its descendant forms are widely spread in the Eastern Bantu area including zones D, E, F, G, J, L, M, N, P, R, and S (for the Bantu zones, see Guthrie 1967–71; Maho 2009; Hammarström 2019), e.g., in Holoholo [D28b] *ηoloβε*, Bungu [JD53] *inguluwe*, Embu [E52] *ngorwe*, Bende [F12] *ngulubhe*, Luguru [G35] *nguluwe*, Luba [L31a] *ηulube*, Ila [M63] *tfulube*, Ngindo [P14] *ligolobe*, Ngandjera [R24] *siηguruβε*, Sethoto [S33] *k’olobe*.

On the other hand, the shortened forms are exclusively distributed in the Western zones including zones A, B, C, H, K, L, and R, e.g., in Bulu-Bene [A74] *ηgoe*, Nzadi [B865] *ηgûl*, Lingala [C30B] *ngûlu*, Congo-Yoombi [H16c] *yingulu*, Luwena-Lubale [K14] *ηgulu*, Pende [L11] *ηgulu*, Khumbi [R14] *ocingulu*.

This ‘East vs. West’ distribution is quite suggestive about the process of historical development considering the distribution pattern attested e.g., in the case of ‘chicken’ (Shinagawa and Komori 2022b), which is divided into ‘North-West vs. the rest’, as it may suggest these patterns reflect different stages in the process of ‘Bantu expansion’ (for a current discussion of the migration route, see de Filippo *et al.* 2012, Grollemund *et al.* 2015, among others).

3. Local variation

A closer look at the Eastern belt reveals more variation. For example, a number of forms descended from **-gòdòbè* take another shortened form that drops the syllable in the middle. The forms, which we tentatively label as *<gubi>*, are rather limitedly distributed in zones E, G, and P such as in Kilimanjaro Bantu (Chaga) languages [E60] *nguwe*, in Hehe [G62] *ingubi*, and in Matumbi [P13] *ngobe*.

There are also many different forms which are only shared in a very closely related group of languages, which might be seen as a local lexical innovation. An interesting distribution is observed around the Lake Victoria, where there are at least three clusters of languages that share a common form denoting ‘pig’. In the northern area, *<mbizi>* and its related forms are observed, e.g., in Lusoga [JE16] *embidhi*, Gwere [JE17] *mbizi*, Masaaba [JE31] *imbizi* etc. In the eastern shore of the lake, *<beche>* is the common form, e.g., Saamia [JE34] *embeche*, Suba [JE403] *embichi*, and Kuria [JE43] *embeshe*. Finally in the western part of the lake, *<punu>* and the related forms are shared among Nyoro-Ganda group [JE10] in Uganda as well as in Haya-Jita group [JE20] in Tanzania. Like in the case of ‘rat’ (cf. Shinagawa and Komori 2022a), we can also find out a lexical unity among the languages spoken in the corridor between Lake Tanganyika and Lake Rukwa [M10–20], where *<kapoli>* and the related forms are commonly used.

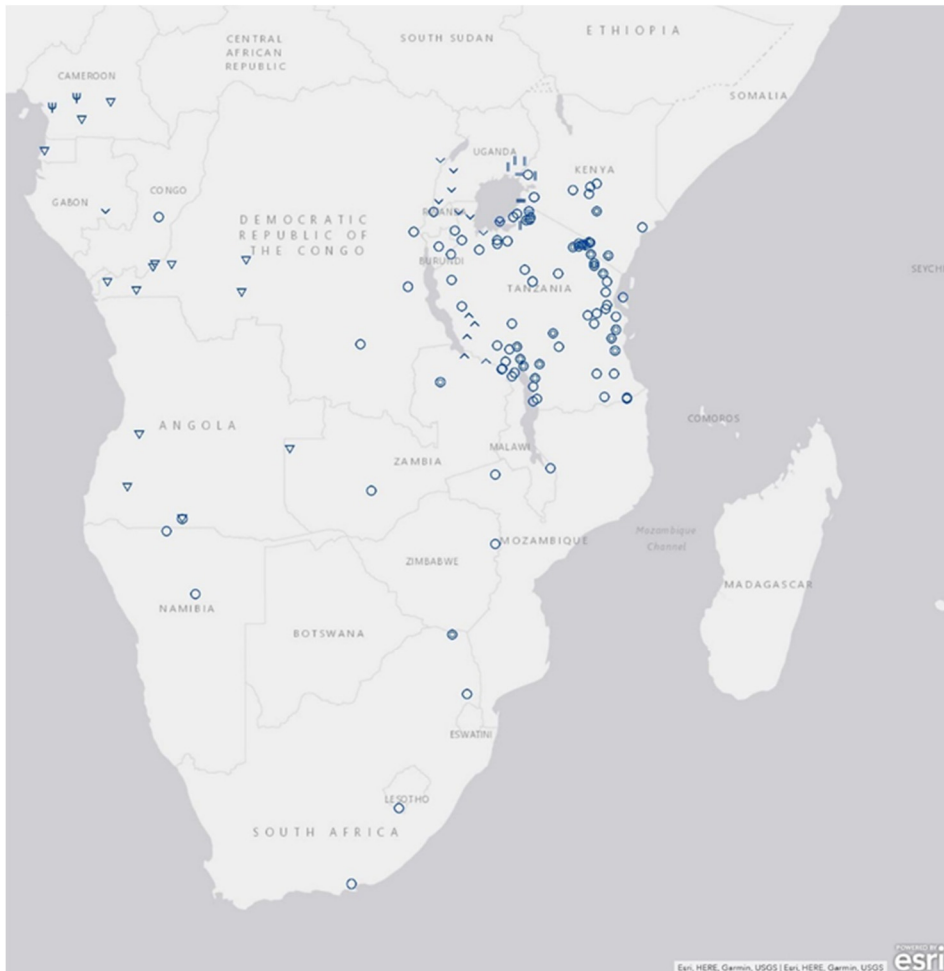


Fig. 1 Geographical distribution of the word ‘pig’

Forms traced back to reconstructed PB forms	Other common forms
○ *-gòdòbè; *-gòdòbì	∨ <punu>
⊙ *-gòdòbè; *-gòdòbì <gubi>	∧ <kapoli>
▽ *-gòdò	= <beche>
	<mbizi>
	Ψ <goy>

References

- Bastin, Yvonne, André Coupez, Evariste Mumba, and Thilo C. Schadeberg (eds.) (2002) Bantu lexical reconstructions 3 / Reconstructions lexicales bantoues 3. <http://linguistics.africamuseum.be/BLR3.html>
- de Filippo, Cesare, Koen Bostoen, Mark Stoneking, and Brigitte Pakendorf (2012) Bringing together linguistic and genetic evidence to test the Bantu expansion. *Proceedings of the Royal Society B*. doi: <https://doi.org/10.1098/rspb.2012.0318>
- Grollemund, Rebecca, Simon Branford, Koen Bostoen, Andrew Meade, Chris Venditti, and Mark Pagel (2015) Bantu expansion shows that habitat alters the route and pace of human dispersals. *PNAS* 112(43): 13296–13301. doi: <https://doi.org/10.1073/pnas.1503793112>
- Guthrie, Malcolm (1967–71) *Comparative Bantu: An introduction to the comparative linguistics and prehistory of the Bantu languages*. 4 Vols. London: Gregg Press.
- Hammarström, Harald (2019) An inventory of Bantu languages. In Van de Velde, Mark, Koen Bostoen, Derek Nurse, and Gérard Philippson (eds.) *The Bantu languages. Second edition*, 17–78. London and New York: Routledge.
- Maho, Jouni Filip (2003) Classification of the Bantu languages: An update of Guthrie's referential system. In Nurse, Derek and Gérard Philippson (eds.) *The Bantu languages*, 639–651. London and New York: Routledge.
- Meeussen, Achille Emile (1969) *Bantu lexical reconstructions*. Ms. Tervuren: Royal Museum for Central Africa
- Nurse, Derek and Gérard Philippson (1975) The Tanzanian language survey. Dar es Salaam: Department of Foreign Languages and Linguistics of the University of Dar es Salaam. doi: <https://doi.org/10.5281/zenodo.5121819>
- Shinagawa, Daisuke and Junko Komori (2022a) 'Rat' in Bantu. In Hiroyuki Suzuki et al. (eds.), 68–70. doi: <https://doi.org/10.5281/zenodo.7117887>
- Shinagawa, Daisuke and Junko Komori (2022b) 'Chicken' in Bantu. In Hiroyuki Suzuki et al. (eds.), 114–116. doi: <https://doi.org/10.5281/zenodo.7117922>
- Suzuki, Hiroyuki, Mika Fukazawa, Akiko Yokoyama, and Mitsuaki Endo (eds.) (2022) *Linguistic atlas of Asia and Africa I*. Tokyo: Geolinguistic Society of Japan. doi: <https://doi.org/10.5281/zenodo.7118188>

Publication history

Date accepted: 15 September 2022

論文

日本語における「兄弟姉妹」を表す総称の分布

福嶋秩子
新潟県立大学

Distributions of general terms for ‘sibling’ in Japanese

FUKUSHIMA, Chitsuko
University of Niigata Prefecture

Abstract: The author is involved in the making of *the Linguistic Atlas of Asia and Africa* concerning the system of ‘sibling’ terms. In the process, it has been noticed that general terms for ‘sibling’ often reflect the system which might be alive or outdated (obsolete). In this paper, general terms for ‘sibling’ in Japanese dialects and Ryukyuan dialects are examined and made into linguistic maps using ArcGIS Online. The changes shown in the maps are interpreted referring to the variation of systems. A map of a different lexical item ‘the day before yesterday’ which has forms similar to general terms for ‘sibling’ is compared to prove the existence of homonymic collision.*

キーワード：キョウダイ名；総称；言語地図；同音衝突

Keywords: Sibling terms; General terms; Linguistic maps; Homonymic collision

1. はじめに

『地理言語学研究』1に「日本語における「キョウダイ名」と題する小論を投稿した（福嶋 2021）。東京外国語大学アジア・アフリカ言語文化研究所の共同利用・共同研究課題「アジア・アフリカ地理言語学研究」における『アジア・アフリカ言語地図』作成のプロジェクトで「キョウダイ名」の体系をとりあげることが決まり、筆者がそのコーディネーターを引き受けたため、その準備のつもりであった。体系を考えるにあたっては、以下の三つの観点から整理できる（①本人から見た年齢差：年上~年下、②相手の性別：男性~女性、③本人から見た性差：同性~異性）（松本 2006、Murdock 1968）。日本語のキョウダイ名の地図化を行っているときに、体系図ではなく、「兄弟姉妹」を表す総称が、キョウダイ名の体系を示しているケースがままあることに気づいた。漢

福嶋秩子 (2022) 「日本語における「兄弟姉妹」を表す総称の分布」 『地理言語学研究』 2: 126–132. doi: <https://doi.org/10.5281/zenodo.7121605>

* 本研究は、JSPS科研費 JP 19K00555 の助成をうけたものです。また、東京外国語大学アジア・アフリカ言語文化研究所の共同利用・共同研究課題「アジア・アフリカ地理言語学研究」の成果の一部です。

語の「兄弟」自体が、「兄」＋「弟」という構造をもっている。したがって、総称に注目することで、隠れた体系が見えてくる。ただし、当代においては、すでに元の語源意識はなくなっていて、過去の体系の痕跡にすぎないこともあるので、注意が必要である。

2. 日本語方言資料における「兄弟姉妹」を表す総称

図1は平山輝男他編『現代日本語方言大辞典』の「きょうだい」のデータを地図化¹したものである。漢語のキョウダイとその変種が本土方言域のみならず琉球方言域にもある。これは借用語であり、「兄弟」という語源意識は必ずしもない。一方、西日本に分布するオトドイは、古代日本語にあったとされる「本人から見た年齢差」が問題とされる体系（日本国語大辞典）を背景とする造語で、妹・弟を示すオト（ト）と兄・姉を示すエとの複合語（オトト＋エ）に由来する。語源意識はなくなっているかもしれないが、古代語の痕跡と考えれば興味深い²。八丈方言のオトーネは、「オトは年下のきょうだい、アネは年上のきょうだいの意。両者をあわせてきょうだいの意になる」と記述されているので、これも同じく古代語の痕跡といえよう。なお、琉球方言は、「本人から見た性差」が問題となる体系で、異性のきょうだいについての名称をもっている。古典沖縄語では、男性から見た女きょうだいはヲナリ、女性から見た男きょうだいはウエケレとなる。図1では宮古方言に bikii、bunai があるが、これはウエケレの変種で男きょうだいを意味しているはずである。厳密にいうと総称ではないが、「きょうだい」の項目で回答されている。図1では重なってみえにくくなっているが、宮古平良方言にウトウザ、多良間方言にウトウダラがある。これは別の方言でウトウジャとなる語の変種で、本来（性別にこだわらず）年下のきょうだいを表す名称であるが、方言によっては「兄弟姉妹」を表す総称として使われているようだ。

3. 琉球語方言資料における「兄弟姉妹」を表す総称

キョウダイ名を表す体系の言語地図を描くために『図説琉球語辞典』を使ったが、「兄弟姉妹」を表す総称の項目は含まれていなかった。そのため、琉球語の様々な方言辞典や方言調査資料³を参照して総称をぬきだし、地図化した（図2参照）。キョウダイとその変種が全域にあるが、「本人から見た年齢差」が背景にある妹・弟を示すウトウトウあるいはウットウと兄・姉を示すセダとの複合語、男性から見た女きょうだいヲナリと女性から見た男きょうだいうエケレとの複合語が、与論、首里および八重山にある。一方、（性別にこだわ

¹ 本論における言語地図はいずれもArcGIS Onlineにより描画した。

² なお、オトドイの分布の不思議については4.でとりあげる。

³ 図2作成のために参照した資料名は文末にあげる。



図1：日本語における「兄弟姉妹」を表す総称（『現代日本語方言大辞典』により地図化）

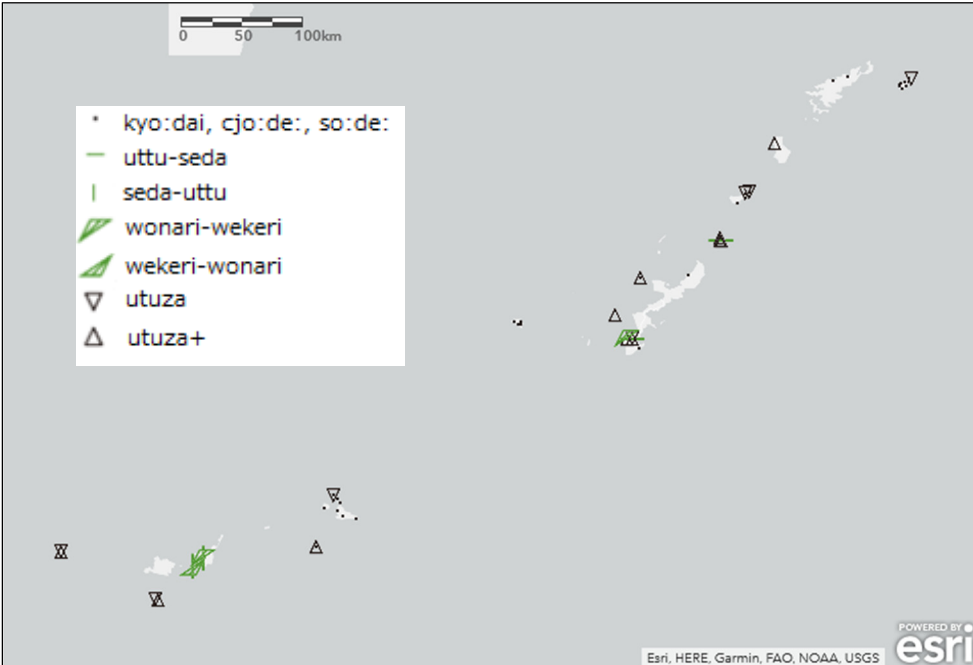


図2：琉球語における「兄弟姉妹」を表す総称（多種方言辞書・方言調査資料により地図化）

らず) 年下のきょうだいを表す名称ウトウザとその変種(複数形か)が奄美から沖縄、宮古、八重山に広く分布する。後者のウトウザの類が意外に古いのではないかと思う。セダウットウやヲナリウエケレの類は一時的に用いられる語で、結合度はあまり高くないのかもしれないが、「本人から見た年齢差」や「本人から見た性差」を組み込んだ体系の存在はうかがえる。

4. オトドイの分布

図1でとりあげたオトドイは、ジャパンナレッジ『日本国語大辞典』で調べると、西日本にさらに広く分布することがわかったので、地図化した(図3参照)。近畿地方から中国・四国地方にかけてのかなり広い地域にある。このオトドイには、清音のオトトイという形式もあり、こちらは主に近畿地方の周辺に分布しているように見える。兵庫県にあるオトニーは、兄をあらわすニーを使っているので、新しい語だろう。

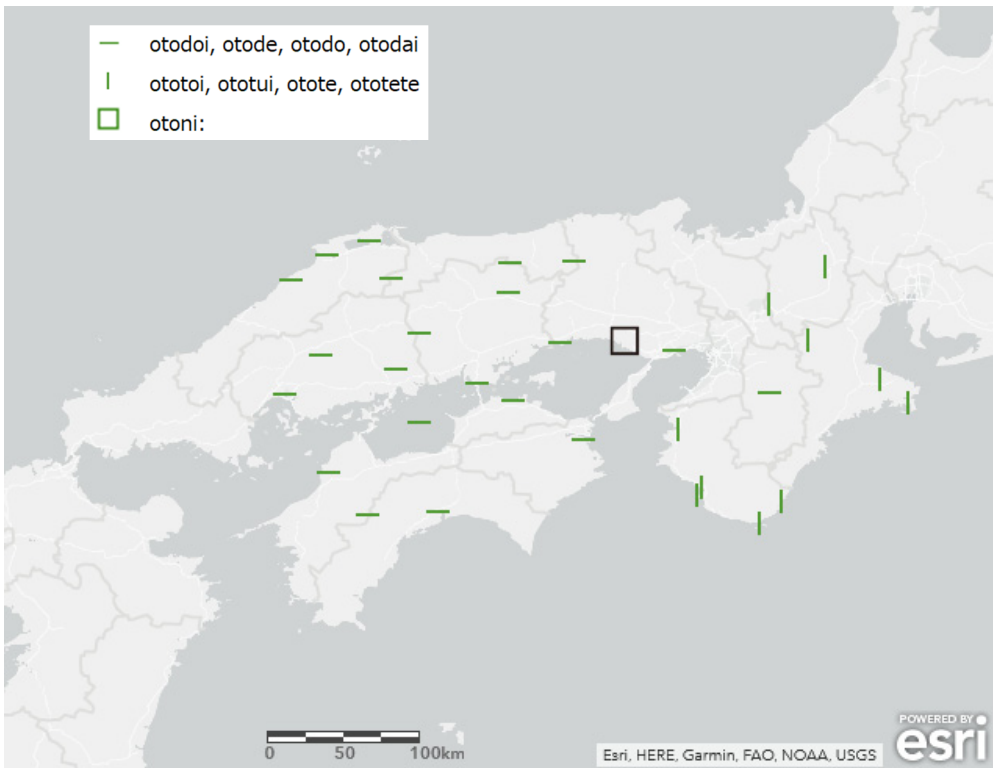


図3：「兄弟姉妹」を表す総称オトドイの分布(ジャパンナレッジ『日本国語大辞典』)

古い語とも考えられるオトドイ・オトトイが西日本だけに分布するのはやや不思議である。その理由を考えるために、類似の語形をもつ「一昨日」の方言

分布と比較する（佐藤 2002 : 33）（図 4 参照）。中央にオトツイ、周辺にオトトイ（東北地方を中心にオトトイナ）が分布し、ABA 分布であるので、「一昨日」については、オトトイ > オトツイ という変化が起きたと考えられる。

I - 2 月日・時間

33

おととい（一昨日）

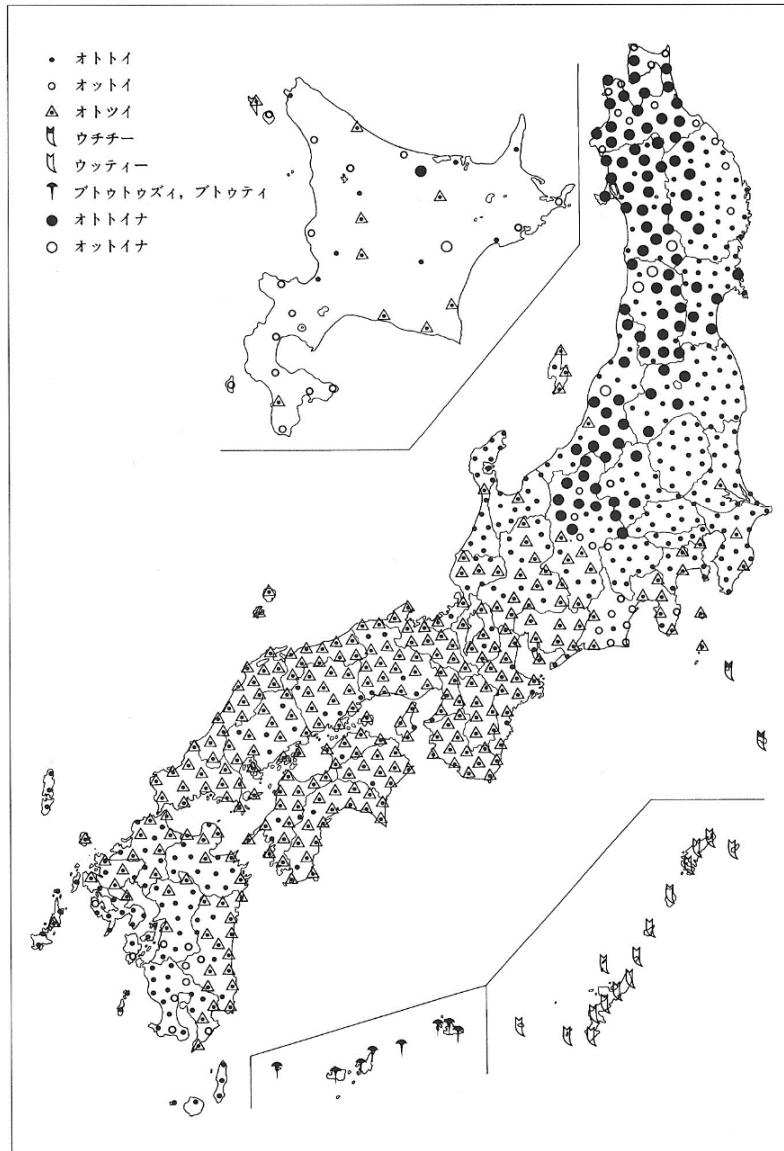


図 4 : 「一昨日」を表すオトトイの分布 出典 : 佐藤 (2002: 33)

二つの地図の類似語形の分布の対応関係は表 1 のようになる。

表 1：オトドイ・オトトイをめぐる分布の比較

	琉球	九州	中国・四国	近畿	東日本
一昨日	オトトイ	オトトイ オトツイ	(オトトイ) オトツイ	(オトトイ) オトツイ	オトトイ オトトイナ
兄弟姉妹			オトドイ	オトトイ オトツイ	

「一昨日」でオトトイが優勢な地域には、「兄弟姉妹」のオトドイ・オトトイが分布していないことが見て取れる。「兄弟姉妹」のオトドイ・オトトイは、「一昨日」のオトトイに比べて新しく、伝播にあたって同音衝突が起きた結果、それ以上広がらず、すみ分けることになったと考えられる。「兄弟姉妹」の総称としては、オトトイが古く、異化としてオトドイが生まれ、中国・四国地方に広がったのではないだろうか。

5. まとめ

東京外国語大学アジア・アフリカ研究所の共同利用・共同研究課題「アジア・アフリカ地理言語学研究」の研究会で、アジア・アフリカの言語・言語群における「キョウダイ名」の体系について地図化を進めている。「キョウダイ名」の体系に属する各語や総称の分布図は興味深いのが、言語・言語群を越えての比較はしにくくなる。小論では、日本語における「キョウダイ名」の総称について地図化を試み、体系との関係性を探った。また、類似語形がある言語地図と比較し、分布の要因について考察した。

参考：図 2 を作成するために使用した琉球語方言辞書・方言調査資料

- 内間直仁・野原三義編著 (2006) 『沖縄語辞典 那覇方言を中心に』 研究社
 岡村隆博他編 (2014) 『徳之島方言辞典 基礎データ版』 徳之島方言の会
 沖縄古語大辞典編集委員会編 (1995) 『沖縄古語大辞典』 角川書店
 生塩睦子 (1999) 『沖縄伊江島方言辞典』 伊江島教育委員会
 菊千代 (2005) 『与論方言辞典』 武蔵野書院
 国立国語研究所編 (1975) 『沖縄語辞典』 大蔵省印刷局
 尚学図書編 (1989) 『日本方言大辞典』 小学館
 富浜定吉 (2013) 『宮古伊良部方言辞典』 沖縄タイムス社
 長田須磨・須山名保子・藤井美佐子共編 (1980) 『奄美方言分類辞典』 笠間書院
 半田一郎編著 (1999) 『琉球語辞典』 大学書林

- 平山輝男 (1966) 『琉球方言の総合的研究』 明治書院
平山輝男他編著 (1992) 『現代日本語方言大辞典』 明治書院
前新透 (2011) 『竹富方言辞典』 南山舎
宮城信勇 (2003) 『石垣方言辞典』 沖縄タイムス社
国立国語研究所 危機言語調査報告書 (喜界島、与論・沖永良部、久米島、南琉球宮古)
ジャパンナレッジ 『日本国語大辞典』

参考文献/References

- 佐藤亮一 (2002) 『お国ことばを知る方言の地図帳』 東京：小学館.
中本正智 (1981) 『図説琉球語辞典』 東京：力富書房金鶏社.
平山輝男他編 (1992) 『現代日本語方言大辞典』 東京：明治書院.
福嶋秩子 (2021) 「日本語における「キョウダイ名」」『地理言語学研究』 *Studies in Geolinguistics* 1: 115-122. doi: <https://doi.org/10.5281/zenodo.5529312>
松本克己 (2006) 「第18章 世界諸言語のきょうだい名」『世界言語への視座 歴史言語学と言語類型論』 東京：三省堂.
Murdock, George Peter (1968) Patterns of sibling terminology. *Ethnology* 7(1): 1-24. URI: <https://www.jstor.org/stable/3772805>
ジャパンナレッジ <https://japanknowledge.com/> 日本国語大辞典

出版情報

採用決定日：2022年8月31日

論文

中国語方言における「兄」「弟」「姉」「妹」を表す語の分類と分布

八木堅二
国士舘大学

Classifications and Distributions of Sibling Terms in Chinese Dialects

YAGI, Kenji
Kokushikan University

Abstract: The purpose of this paper is to classify and draw maps of distributions of words representing “elder brother”, “younger brother”, “elder sister”, “younger sister”, “brothers”, “sisters”, and “brothers and sisters” in Chinese dialects. Chinese dialects basically possess a system that divides all of “elder brother”, “younger brother”, “elder sister”, and “younger sister”, but when looked the dialects as a whole, there are word forms or phenomena that indicate commonality between each sibling term. According to Matsumoto’s (2006) classification, all sibling term systems of Chinese dialects are identical, but in the future, it is necessary to consider the formation of distributions and relationships with surrounding languages, while paying attention to such commonalities.*

キーワード: 中国語方言 ; 兄弟姉妹を表す語 ; アジア・アフリカ地理言語学
Keywords: Chinese dialects; Sibling terms; Studies in Asian and African Geolinguistics

1. はじめに

松本 (2006) において全世界 1479 言語を対象とした sibling term のタイプ分けと大まかな分布状況が示されたが、2022 年のアジア・アフリカ地理言語学研究 (遠藤光暁教授主催) の例会において、アジア・アフリカにおける sibling term 体系のさらなる詳細な分布状況が各言語担当者から示された。

ユーラシア大陸において「兄」「弟」「姉」「妹」の四種を分ける体系 (松本(2006)の D 型) は特に東アジアにおける分布の集中が顕著に見られ、なかで

八木堅二 (2022) 「中国語方言における「兄」「弟」「姉」「妹」を表す語の分類と分布」『地理言語学研究』2: 133–152. doi: <https://doi.org/10.5281/zenodo.7121644>

* 本研究の一部はJSPS科研費JP20K13032の助成を得て行った。

もほぼ全ての方言で四種を区別する体系を保持する中国語諸方言はその分布の中核をなしていると言えよう。

これらの分布の形成過程や通言語的な関連性についてはこれからさらなる検討が加えられることになるが、それに先立ち中国語方言の詳細な分布状況を示しておくことは他言語の研究者にとっても有益と考え、中国語諸方言における「兄」「弟」「姉」「妹」「兄弟」「姉妹」および「兄弟姉妹」を表す語の分類と分布状況について概括しておきたい。

2. 「兄」

2.1. 「兄」を表す語の分類¹

- A1 兄/xiong/ : [eian²¹], [fie⁴⁴], [fæ⁵⁵], [hin³³], etc.
- A2 兄兄/xiong xiong/ : [hiã¹¹hiã³⁵]
- A3 阿兄/a xiong/ : [a¹¹(h-)jiã⁵⁴], [a¹³ia⁵³³], [a¹¹(h-)ie⁵³³]
- A4 老兄/lao xiong/ : [lau⁴¹ eion³³], [lau²⁴çian³³]
- A5 兄佬/xiong lao/・兄郎/xiong lang/ : [eiã³³lɔ⁴¹], [ei⁵²4lɔ⁴⁴], [ei⁵²lɔ⁵⁰]
- A6 兄哥/xiong ge/ : [hiã³³kx⁵⁵], [hiã³³ko⁵⁵]
- A0 兄（その他）：兄长[eyəŋtɕaŋ], 尼兄[ni¹¹hia²⁴]
- B1 哥/ge/ : [kə⁵³], [ko⁵⁵], [ke³⁴], [kə⁴⁴3], [kou⁵¹]
- B2 哥哥/ge ge/ : [kə⁴⁴kə⁴], [kou⁴⁴kou⁴⁴], [ko²²ko⁵⁵], [kuə³¹kuə⁰]
- B3 阿哥/a ge/, 依哥, 阿哥儿 : [a⁰kə⁴⁴3], [a¹¹ko¹¹], [a[?]5ko³³], [ʔə[?]3ku⁵¹]
- B4 哥子/ge zi/, 哥仔, 哥崽, 哥嘢 : [ko⁵⁵tsɿ⁰], [ko¹³tsɿ³³], [kuo³³tsai³³]
- B5 哥儿/ger/ : [kən¹¹], [ko⁵⁵ə⁵⁵]
- B6 哥郎/ge lang/ : [ko³³lan⁰], [kuo¹¹lən⁰]
- B7 哥老信/ge lao guan/ : [ko⁵⁵nau⁴²kuan⁵⁵], [ko⁵⁵lau⁴²kuan⁵⁵]
- B8 老哥/lao ge/ : [lau⁵⁴ko⁴⁵]
- C1 大/da/ : [ta²¹]
- C2 大大/da da/ : [da³⁴da⁴⁴], [ta³⁵ta³⁵], [da²²da²⁴]
- C3 阿大/a da/ : [a⁴⁴dəu³⁴]
- C4 大佬/da lao/・大郎/da lang/ : [tho¹¹lau¹¹], [tai²²lu³⁵]
- C5 老大/lao da/ : [lɑ²³tha⁵³], [lei⁵⁵da²⁴]
- C6 大哥/da ge/ : [tɑ⁴²3ko⁵⁵], [thai²¹ko¹³]
- D1 pa : [pa²⁴]

¹ これらのデータは既刊の方言報告類から収集したもので、分類は主に漢字語形に基く。以下アルファベットで大分類を示し、数字は下位分類を示す。次に代表的な漢字語形を示し、//内に標準語の音声（ピンイン）を例示する。最後に[]内に実際の音声形式（音声記号は基本的にはIPAだが一部中国方言学慣用のものが含まれる。数字は5度制調値で、0は轻声を示す）の例を挙げる。なお本稿では漢字語形に基き議論を進め音声記号はあくまで参考として付し、音声記号を一部ないし全て省略する場合がある。

D2 老伯/lao bo/, 老伯哩 : [lɔpa], [laupak], [laupa, lɔupeu?lei]

D3 阿伯/a bo/ : [a33pek5], [ua213pek5]

D0 p- (その他) : [p^hiau44pa41tsɿ0]

0 その他 : 爰[ei51], 拐子[kuai52tsɿ0], [mai42lian35], [tsa21tsa33], 阿吾, 阿卧, 相汉

2.2. 「兄」を表す語の分布

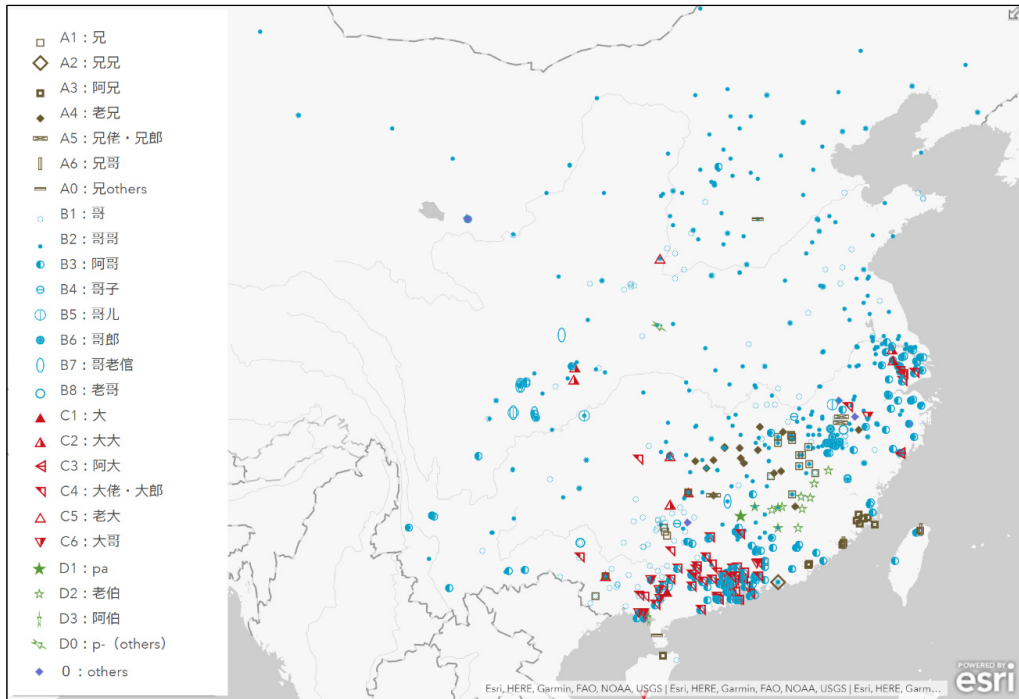


図1a : 「兄」を表す語の分布 (全国)

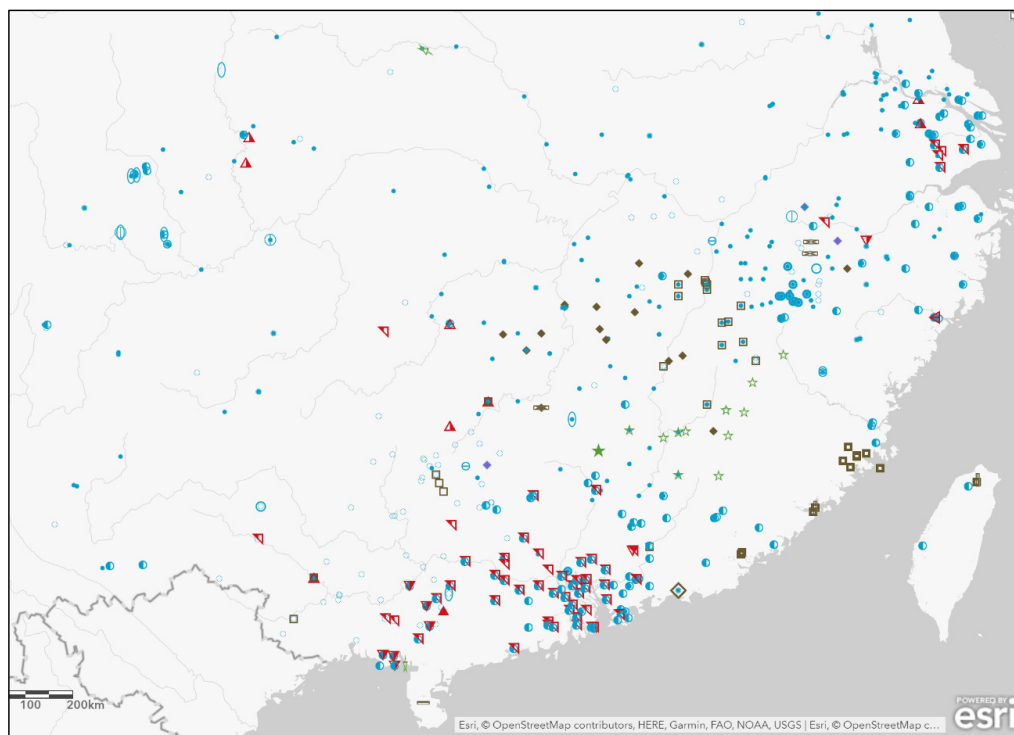


図1b: 「兄」を表す語の分布 (南方拡大)

まず「兄」を表す語の語形を主に四種: A「兄」を有するもの(兄タイプ)、B「哥」を有するもの(哥タイプ)、C「大」を有するもの(大タイプ)、D「pから始まる音節」を有するもの(Pタイプ)に分類する。さらにそれぞれのタイプの中で語構成などを勘案して下位分類を行った。

「兄」を表す語のタイプは比較的明瞭な分布傾向を示しており、大局的にはB「哥」が全国的に分布し、その他のタイプが主に南方に分布する形になっている。A・C・Dタイプは南方において周圈的な分布を見せており、D(Pタイプ)が主に江西南部の客家語を中心にまとまって分布し、A(兄タイプ)が主に江西・湖南と福建に分布しDの分布を北西方面と南東方面から取り囲んでいる。さらにC(大)タイプが広東周辺と長江下流域に遠隔的に分布し、四川や安徽その他の北方地域にも散発的に分布が見られる。

八木「中国語方言における「兄」「弟」「姉」「妹」を表す語の分類と分布」

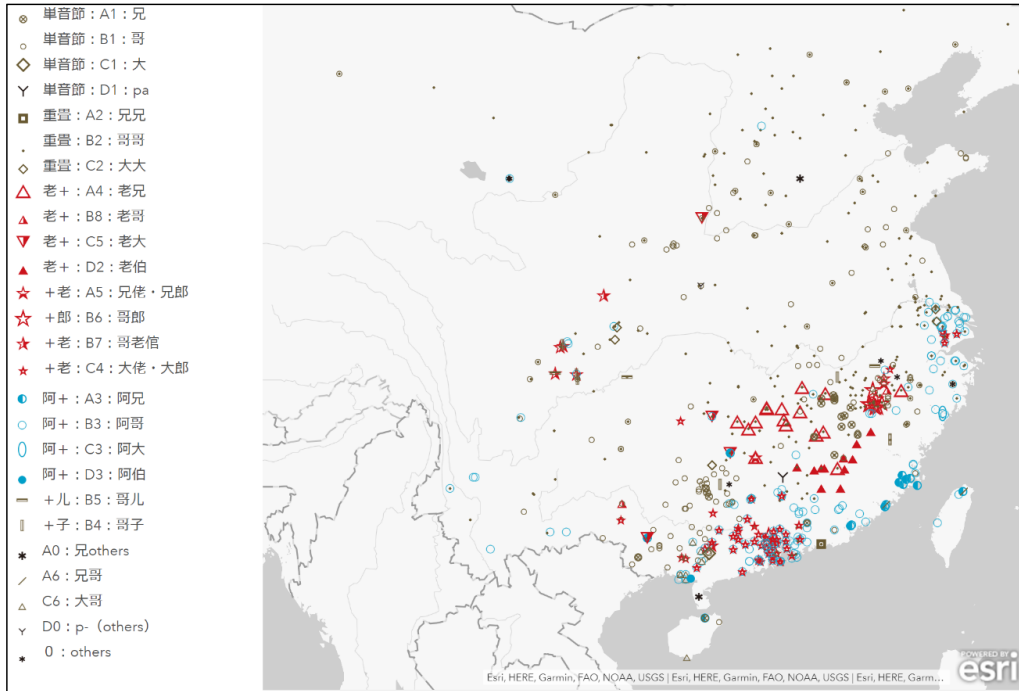


図2a : 「兄」を表す語の修飾成分の分布 (全国)

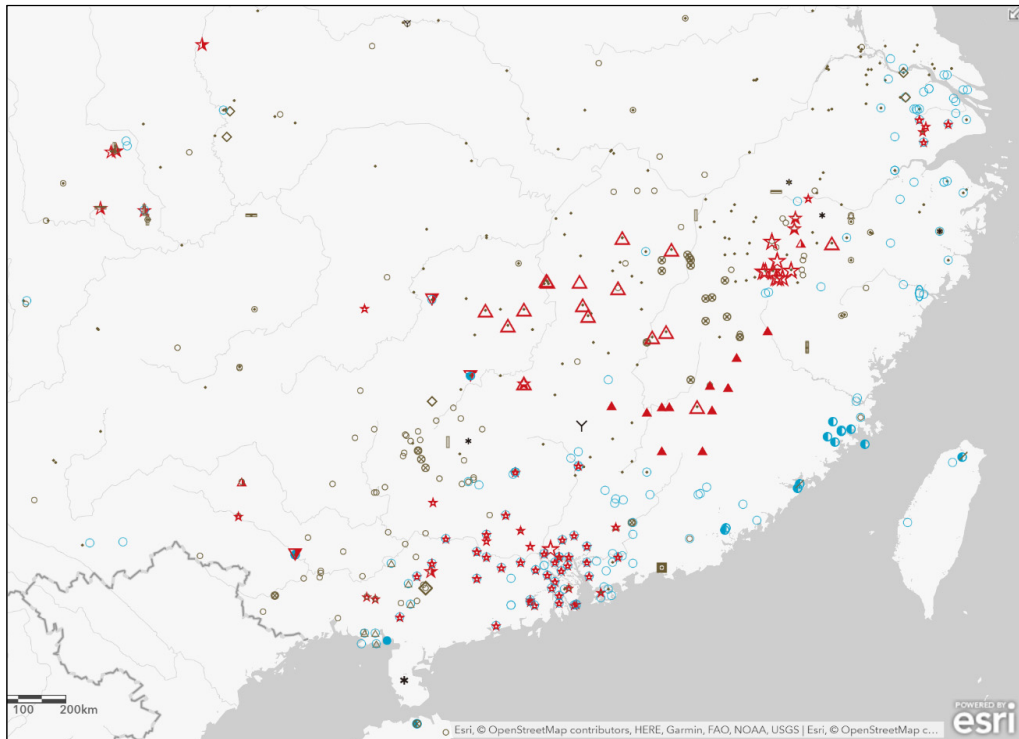


図2b : 「兄」を表す語の修飾成分の分布 (南方拡大)

次に「兄」の語形の修飾成分に着目してみたい（図2：分類は図1と同じだが、修飾成分を基に並べ替え記号を変更した）。

「兄」の語形を語構成を基準に分類すると、単音節のもの、重ね型のもの、前置成分として「老」を有するもの、後置成分として「老」を有するもの、接頭辞の「阿」を有するもの、接尾辞（子や儿など）を有するもの、およびその他と分類される。

このうち接頭辞の「阿」を有する語形は南部の沿岸部を中心に、広く分布し、内陸部や一部北方地域でも分散的な分布が見られる。「老」は語の前部に出現する場合と後部に出現する場合がある。本論では語構造に関する判断を避けてそれぞれ前置成分・後置成分と呼ぶことにする。「老」が前置成分となる語形は江西・湖南を中心にまとまった分布を示し、後置成分となる語形は広東を中心に多く分布し、その他江西東北部から長江下流域の地域と、四川などにも分布し、前置成分となる語形の分布を取り囲んでいる。

修飾成分と語幹の分布状況を比べると、それぞれにおいてタイプの異なる地域的連続性が看取される点が興味深い。

3. 「弟」

3.1. 「弟」を表す語の分類

- A1 弟/di/ : [di31], [dai213], [thai11], [ti52], [tuæ3], [tie41]
- A2 弟弟/di di/ : [ti24ti0], [ti42ti0], [thi55thi22]
- A3 弟子(仔)/di zi/ : [di6tsɿ], [thi1tsɿ], [thai53teie0]
- A4 弟兄/dir/ : 弟兄[ti213ə21], 弟兄弟兄[dɪŋ21dɪŋ33], [dɛŋ23dɛŋ44]
- A5 弟娃/di wa/ · 弟娃儿/di war/ : [ti35ua21], [ti13uaə21]
- A0 弟（その他） : 幺弟
- B1 兄/xiong/ · 兄兄/xiong xiong/ · 兄儿/xiongr/ : [ɕior31]
- B2 兄弟/xiong di/ : [ɕyŋ33ti0], [ɕiʌŋ1thi], [huai43ti21], [ɕyŋ214-31 li0], [haŋ13thoi21]
- B3 大哥/di ge/ : [die22ko45]
- C1 老老(佬佬)/lao lao/ : [lɔ44lɔ44], [lɔ55lɔ0]
- C2 細老(佬)/xi lao/ · 小老/xiao lao/ : [sɕi33lou35], [ɬai33lou55], [sɕi33lɔ35]
- C3 弟老(佬)/di lao/ · 弟郎/di lang/ : [thi lɔ], [thi33laŋ0], [thi33lan0]
- C4 弟老倌/lao di guan/ : [thi12lɔ3kuə43]
- C5 老弟/lao di/ : [lau213thi0], [lɔ44thi24], [lu33thie33], [lau352li31]
- C6 佬細/lao xi/ : [lɔ22sɕiə433]
- D1 阿弟/a di/ : [a13(t-)li11], [aʔ5di13], [a0tie223]

- D2 小弟/xiao di/・細弟/xi di/ : [sɪx55ti33], [sio55ti33]
- D3 舍弟/she di/ : 舍弟[sA42thi213]
- D4 毛弟/mao di/ : 毛弟,毛弟兄[mau21tir213]
- E1 kua・kuai : 瓜弟[kua55ti13], 弟乖[kua55ti13]
- 0 その他 : [mo33mo33], [miemie], 吊汉, 阿宝, 溺浊公本文

3.2. 「弟」を表す語の分布

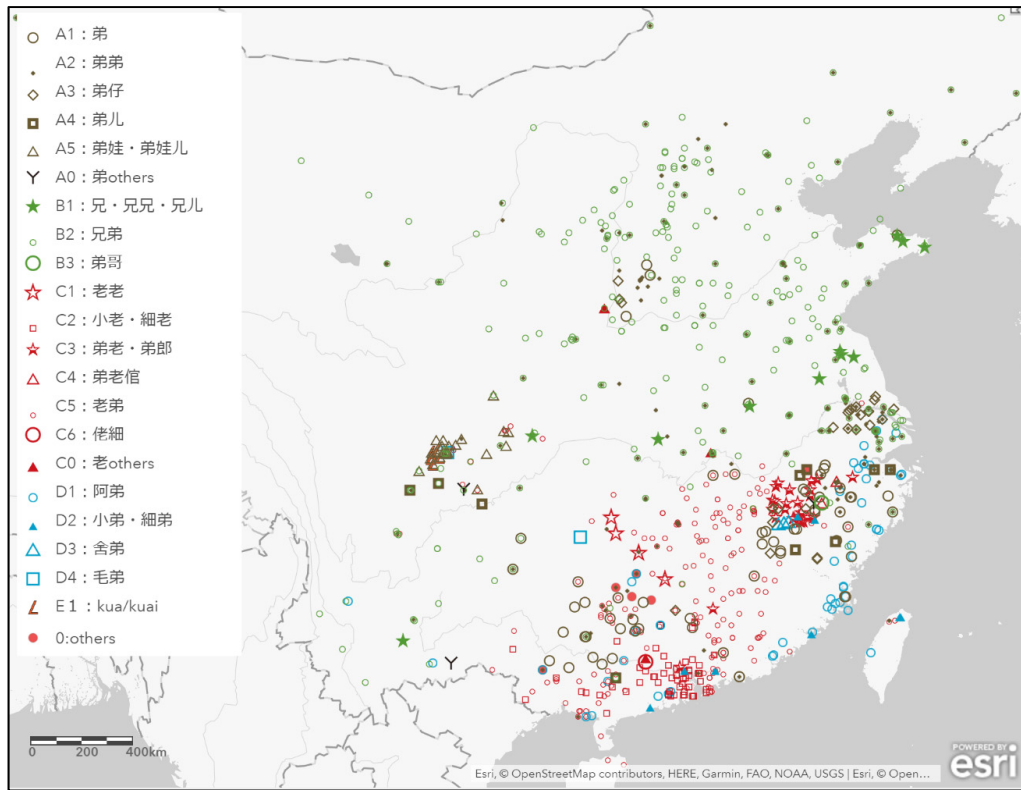


図3a : 「弟」を表す語の分布 (全国)

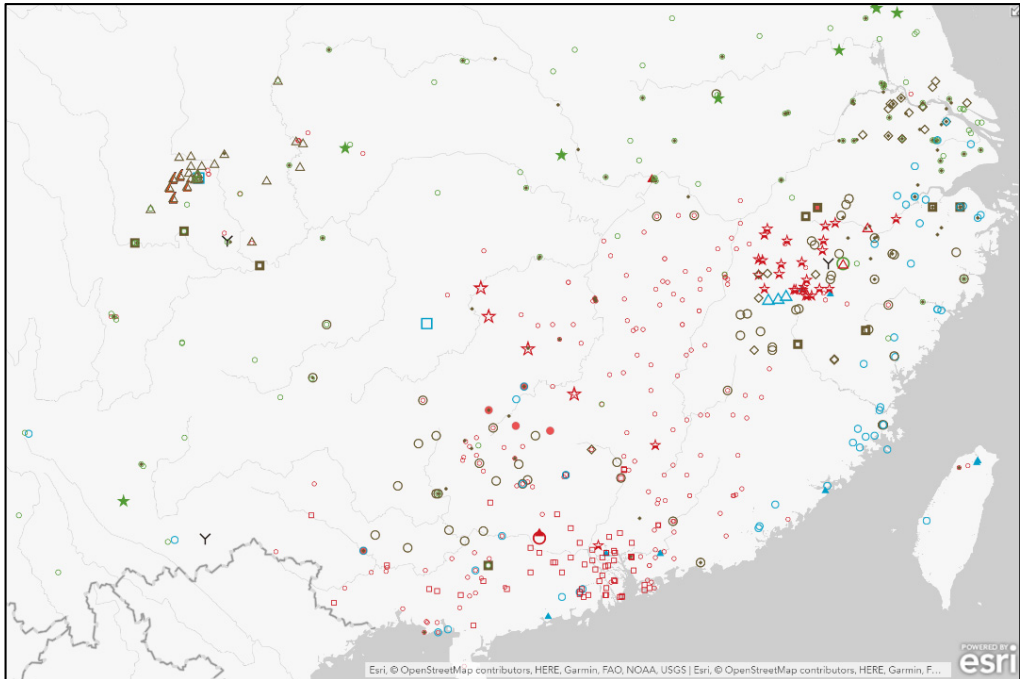


図3b：「弟」を表す語の分布（南方拡大）

弟の語形は基本的に「弟」を含んでいるが、「兄」と「老」のみで構成される語形もある。本論の分類では「弟」単独で構成される語および後置成分を含むものをA、「弟」に前置成分を含むものをD、語中に「兄」を含むものをB、語中に「老」を含むものをCとする。

B「兄」を語中に含むタイプのうち、B2「兄弟」となるタイプが北方を中心に広く分布している。B1は「兄」や「兄兄」のように語幹が「兄」で、「弟」を含まない形式となっている。B1はまとまった分布を見せないが、山東・江蘇・河南から長江流域にかけて北方方言南部の境界付近に分布する傾向が見られる。これらの地域で「兄」を示す語として用いられるのは「哥」である点にも留意しておく必要がある。「兄弟」を示す語に関する分布との関連については後述する。

(B2 兄弟タイプ以外の)「弟」に前置成分を含む語形は主に長江以南に分布している。「老」を有する語形の分布領域は、「兄」の地図における分布状況と類似しているが、わずかに「兄」における「老」よりも分布が多い。また湖南を中心にC1「老老」のような「老」のみで構成される語形がいくつか見られる。

常宁：佬佬[lɔ44lɔ44] (吴启主 1998:149)

沅陵：[lɔ55lɔ0] (杨蔚 1999:138)

このような語形は「兄」の地図には見られない。広東・広西、江西東北部などでは「老」が後置成分となるものが多く、江西・湖南などでは「老」が前置成分となるものが多く見られる。

4. 「姉」

4.1. 「姉」を表す語の分類

- A1 姊/zi/: [tsəi33], [tsi33], [tei54]
- A2 姊姊/zi zi/: [tsa33tsa35], [tse55tse53], [tsiʔ5tsiʔ3], [tei33teie0]
- A3 姊子(仔)/zi zi/: [tsi33tsie21], [tsi51a0]
- A4 姊儿/zir/: [teii], [teien]
- A5 阿姊/a zi/: [aʔ21tsi453], [a55tsi33], [aʔ5teie44], [a33ti55], 阿姊仔
- A6 大姊/da zi/: [thai21tsi33],
- A7 老姊/lao zi/: [lau tsu]
- A8 姊佬/zi lao/: [tei lo]
- A9 姊哥/zi ge/: [ʧi34ɔ443]
- A0 姊others: 姊阿, 尼姊, □姊, [len55teie42]
- B1 姐/jie/: [teie214], [teia35], [teie13], [tse53], [tsie53], [ti54]
- B2 姐姐/jie jie/: [teie55teie0], [tei22teie33], [teia31teia31], [tsie324-24tsie0]
- B3 姐子/jie zi/: [tsia352te0], [teia41ti0], [tsi33tsai33]
- B4 姐儿/jier/: [teier53]
- B5 阿姐/a jie/: 依姐, [aʔ5teie43], [ai44tsi35]
- B6 大姐/da jie/: [ta213teie44], [tæ31ti42]
- B7 家姐/jia jie/: [ka55tse55], [ka55tsie55], [ka55tsø55]
- C1 大/da/: [da51], [tai35]
- C2 大大/da da/: [ta44ta44], [tai44tai44], [təʔ3tə0]
- C3 阿大/a da/: [a33thai33], [a33thei]
- D1 家/jia/: [ka55]
- D2 阿家/a jia/: [a33ka55], [a33ka33]
- D3 姪老倌/cha lao guan/: [tsa55lɿ32kuŋ54]
- D4 霞~/xia~/: [xa24teie42]
- E1 阿保/a bao/: [a33pau24]
- E2 阿哈/a hai/: [aʔ2hai54]
- E3 阿姪/a cha/: [a3tsa33]
- F1 娘娘/niang niang/: [lianlian]
- 0 その他: 亚甫, [tei53ia21], 姑妈[ku1ma1], [vi35]

4.2. 「姉」を表す語の分布

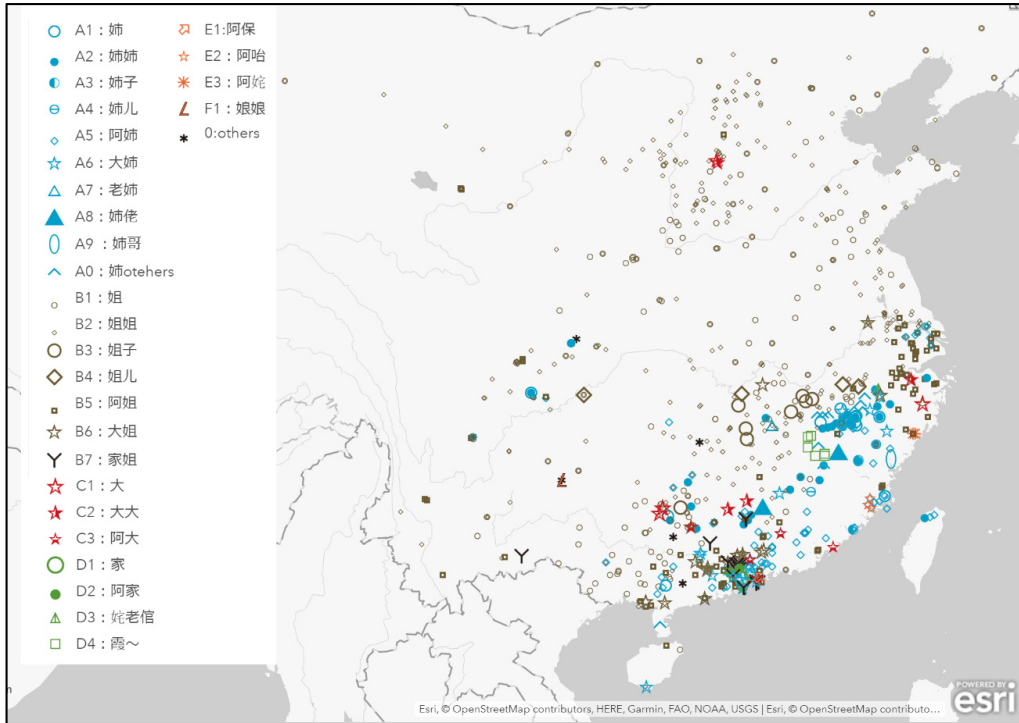


図4a: 「姉」を表す語の分布 (全国)

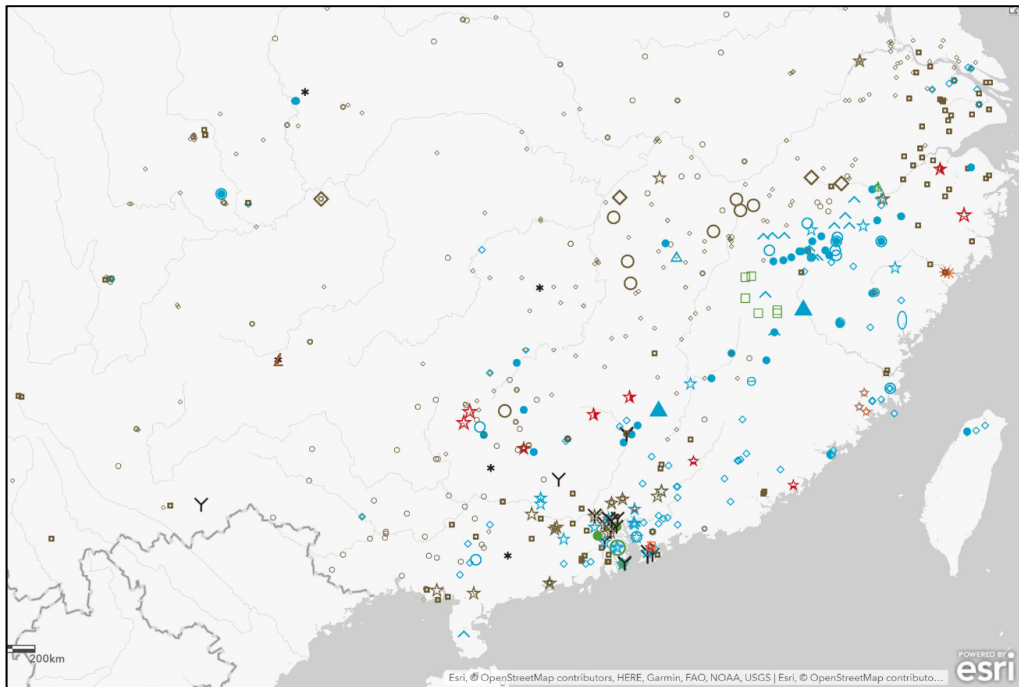


図4b: 「姉」を表す語の分布 (南方拡大)

姉を表す語の分類は漢字語形に着目し「姊」を語幹として有するものをA、「姐」を語幹として有するものをB、「大」を語幹として有するものをCとし、その他の形式については主に音声的・形態的な特徴を勘案してD・E・Fおよびその他として分類した。なお「姊」と「姐」の音形は方言によっては非常に近くその関係については今後音声表記を基に再検討する必要があるが、本稿では全て出典元の漢字表記に従って分類を行った。

A「姊」を有する語形は南方の特に東南部の浙江・福建・広東を中心に分布が広がっている。B「姐」を有する語形は、全国的に分布が広がる。北方方言ではB1の姐単音節型やB2の姐姐のように重ね型となる形式が主となる。長江中流域を中心にB3姐子・B4姐儿のような接尾辞を有する形式が見られるほか、B5阿姐やB6大姐・B7家姐のような前置成分を有する形式は南方に多く分布が見られる。C大を語幹として持つ形式は広東・広西や浙江などに分散的に分布が見られる。A6大姊やB6大姐のような前置成分として大を有する形式南方を中心に分布が見られ、分散的ではあるが分布自体はCよりも広範囲に見られる。なお、「佬（老・郎）」を語中に有する形式は、A7老姊、A8姊佬、D3姥老信などがあるが、分布は少数にとどまる。

5. 「妹」

5.1. 「妹」を表す語の分類

- A1 妹/mei/ : [mai31], [mei53], [mi21], [moi21], [muẽ], [mɛn33]
 A2 妹妹/mei mei/ : [mei42mei0], [me55me53], [mi24mi3], [muẽ11muẽ35]
 A3 妹子/mei zi/ : [me53tsɿ0], [mei24tsɿ0], [min31tsɿ1], [mi31tɛ0]
 A4 妹儿[meir/ : [mər213], [mior55]
 A5 妹娃娃/mei war/ : [mei213uar21]
 A6 妹姆/mei mu/ : [me33m55]
 A0 妹～（その他） : [mæ sə], [muẽi24ku21]
 B1 細妹/xi mei/ : [sɛi33mui35], [siə52mui21], [ɦɛi33mɔi24]
 B2 阿妹/a mei/ : [aʔ21muai42], [ap21mue11], [aʔ2mui52], [a0muai335], 阿妹唧,
 阿妹儿, 依妹
 B3 小妹/xiao mei/ : [sə443mə21], [ɛiə55bə21]
 B4 女妹/nü mei/ : [ny33mɛ0]
 B5 囡妹/nan mei/ : [na22mue212], [nɔ21me212]
 B0 ～妹（その他） : 辣妹, 尼妹, [iəu21məŋ33], 耶妹, 奶妹儿
 C1 姊妹/zi mei/ : [tsɿ42mei0], [tsɿ41 mei213], [tɕɿ212mEi53]
 C2 姐妹/jie mei/ : [tsi33 mE33]
 D1 老妹/lao mei/ : [nau54moi55], [lo53moi21], [lau54mɔi213], [la41me11]
 D2 妹老(佬)/mei lao/ : 妹老仔[mœ44lau53tsi], [mui33la55], 妹郎
 D3 妹老信/mei lao guan/ : [me33lɿ2kua54]

- D4 令妹/ling mei/ : [la33mɿ33], [lɛ11mɿn24]
- E1 姑娘/gu niang/ : [kɔ45nɛŋ31], [kɔ52nɛŋ52]
- E2 娘娘/niang niang/ : [liɑŋ liɑŋ]
- E3 姑xiang/gu ~/ : [kɔu24xiɑŋ22]
- E4 姨xiang/yi ~/ : [i21xiɑŋ21]
- E5 细姨/xi yi/ : [ɬɛi44ji44]
- E6 哥hian/ge ~/ : [ko24hian22]

5.2. 「妹」を表す語の分布

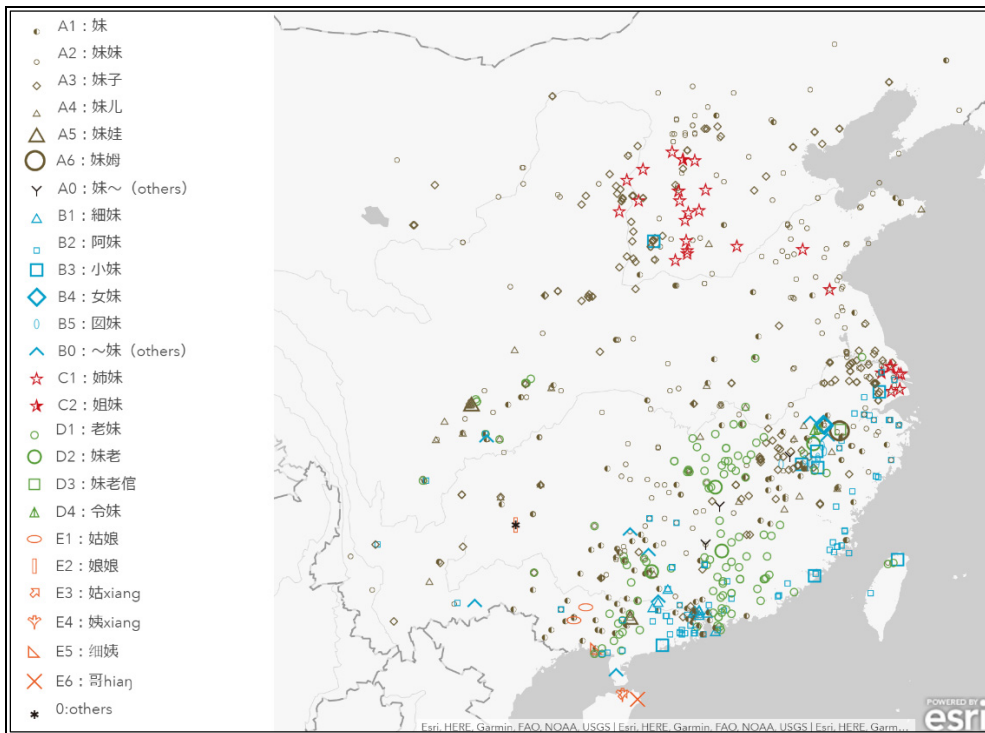


図5a : 「妹」を表す語の分布 (全国)

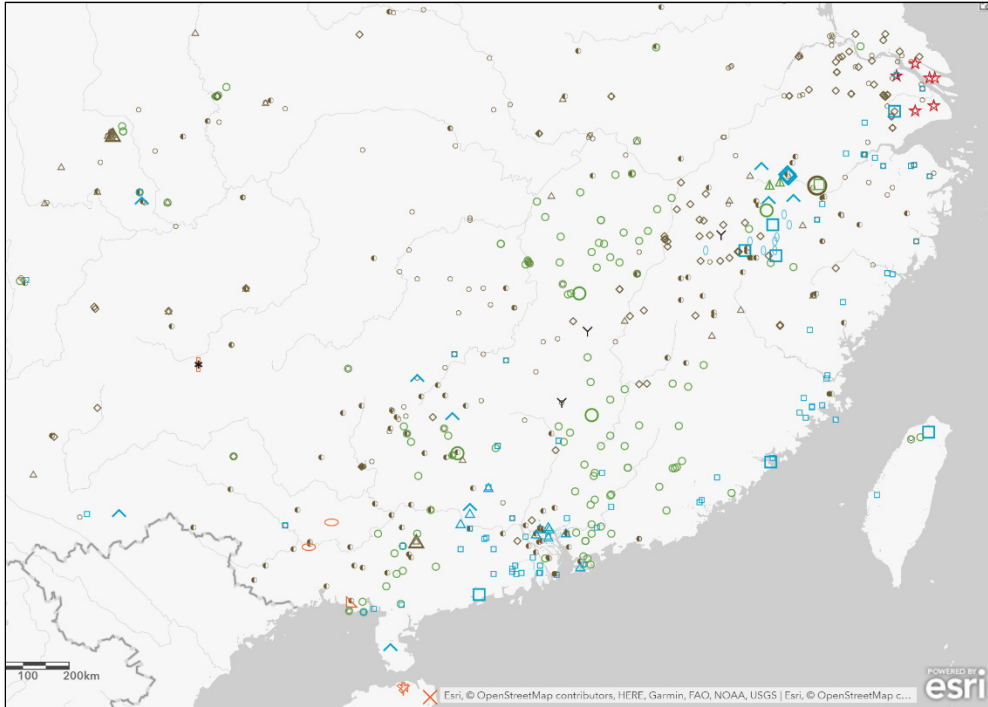


図5b：「妹」を表す語の分布（南方拡大）

妹を表す語は基本的に「妹」を語中に有している。単音節形式（A1）・重ね型形式（A2）・および後置成分を有する語形（A3～A6）をAとし、前置成分を有する語形をBと分類する。姉を示す語（姉・姐）を妹の前置成分として有するタイプ（C1・C2）が山西や江蘇などにややまとまった分布が見られる。山西では「姉」は姉を示す語の語中には用いられないが（「姐」が用いられる）、江蘇では「姉」が姉を示す語の中で用いられる場合がある。

Dは老・佬などを語中に有するタイプである。D1のように老が前置するタイプは江西から広東・広西にいたるまで広範囲に多く分布が見られる。D2のように老が後置されるタイプは数か所にのみ見られるのみだが、分布はD1に隣接しつつ広範に点在する。「妹」の前に前置成分を有する語形は、C以外は基本的に南方に分布している。

「妹」を語中に持たない形式（E）は広西や海南など、南方に分布が少数見られ、その多くは他の親族呼称の形式を語中に有している（E1～E6）。

6. 「兄弟」

6.1. 「兄弟」を表す語の分類

A 兄：兄 xo、兄佬、哥、哥儿、哥儿几个、哥儿们、哥儿们儿、哥几个、

哥们儿

B 弟：阿弟、弟弟

C 兄弟：兄弟、兄弟家、兄弟些、兄弟俩、兄弟仔、兄弟儿、兄弟弭、兄弟们、兄弟伙、兄弟伙的、兄弟伙里、兄弟和里、兄弟伙罗、兄弟俩儿、哥兄老弟、哥弟、亲兄弟、兄弟团、胞兄弟、兄弟哥、哥弟弟

D 弟兄：我家弟兄、弟兄、弟兄家、弟兄子、弟兄淘里、弟兄几个、弟兄弭、弟兄捏、弟兄们、弟兄们子、弟兄伙子、弟兄伙的、弟兄伙里、弟兄伙儿、弟兄伙儿的、弟兄俩、弟兄俩儿

*その他（非表示）

6.2. 「兄弟」を表す語の分布

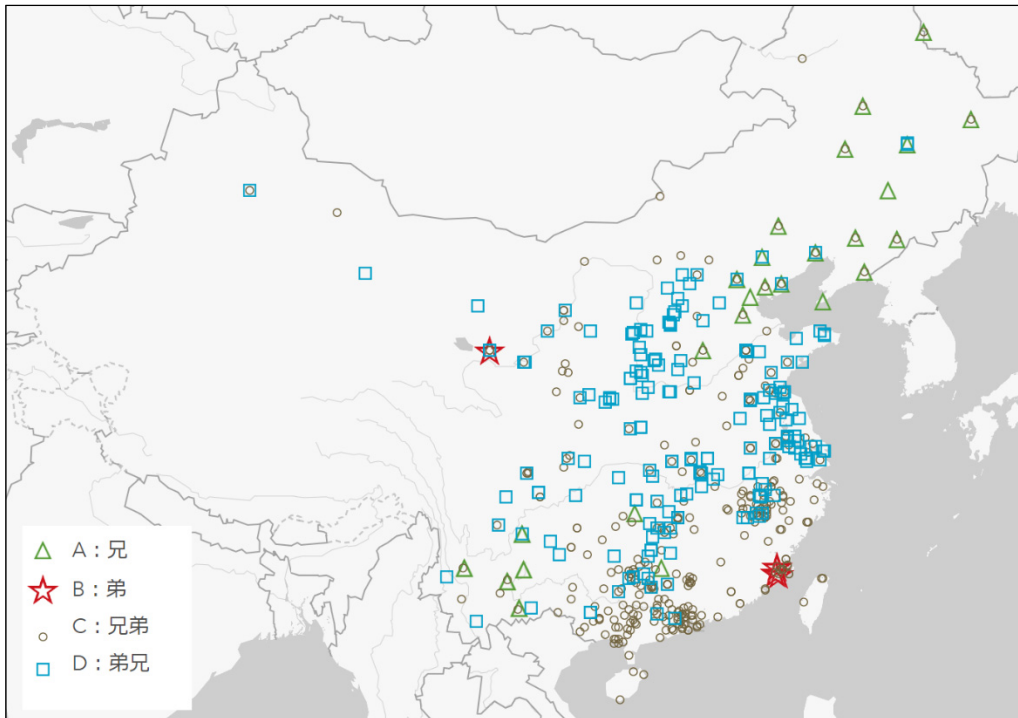


図6：「兄弟」を表す語の分布

「兄弟」を表す語については「兄」を表す語（兄・哥）のみを語中に有し「弟」を表す語（弟）を語中に持たないものをA、「弟」を表す語のみを語中に有し「兄」を表す語を持たないものをB、「兄弟」のように「兄」を表す語が「弟」に前置されるものをC（兄弟哥・哥弟弟のような「兄」を表す語が繰り返し出現するものも暫定的にこの分類に含めている）、「弟兄」のように「弟」を表す語が「兄」に前置されるものをDと分類する。

分布領域を最も広いのはC「兄弟」タイプだが、D「弟兄」タイプが北方を中心に広く分布している。D「弟兄」タイプの成立は弟を表す語の中で「兄弟」の語形が広く用いられている事と関係があるだろう。Aの「兄」を表す語のみで「兄弟」を表す地域が遼寧・吉林などの東北部と雲南を中心とする西南部に遠隔的に分布している点は比較的地点数も多く興味深い。一方B「弟」を表す語のみで「兄弟」を表す地点は少ない。

7. 「姉妹」

7.1. 「姉妹」を表す語の分類

A1：姉・姐：姐 xo、姐儿、姉姐

A2：姐们・姉们：姐儿们、姐们儿、女姊妹们、姊妹们

B1：姐妹、姐妹家、姐妹儿、姐儿妹子、两姊妹、亲姊妹、姊妹、姊妹家、姊妹仔、姊妹子、姊妹淘里、姊妹妹、姊妹里、姊妹和里、姊妹儿、姊妹几个、姊妹弭、姊妹伙子、姊妹伙的、姊妹伙里、姊妹伙儿、姊妹伙罗、姊妹俩、姊妹俩儿、姊妹妹子、姊姊妹妹

B2：姐姐姊妹

B3：女姊妹

B4：姊妹们、姊妹每

C：弟妹得、哥姊妹、姊妹哥

0：その他

7.2. 「姉妹」を表す語の分布

「姉妹」を表す語は「姉」を表す語（姉・姐）のみを語中に有し、「妹」を持たないAタイプ、「姉妹」のように「姉」と「妹」の双方を有するBタイプ、「兄」ないし「弟」を示す語を語中に有するCタイプに大分される。Aタイプの中で語中に複数を示す们/men/を持つものをA2とした（们は音的に妹/mei/と近く注意が必要なため）。

姉妹を表す語は基本的にB1「姊妹・姐妹」が圧倒的に多く分布するが、Aタイプのように「姐」のみで「妹」を語中に有さない語形が北方の沿岸部や長江流域などに分散的に分布している。「哥」「弟」のように「兄・弟」を示す語形を語中に含むCタイプは山西や福建などごく一部で見られる。

忻州：弟妹子[ti mei tə] (温端政1985:93)

寿宁：哥姊妹[kə44 tsi55 muoi335] (秋谷2020:655)

屏南：姊妹哥[tsi11 muai55 ə443] (秋谷2020:655)

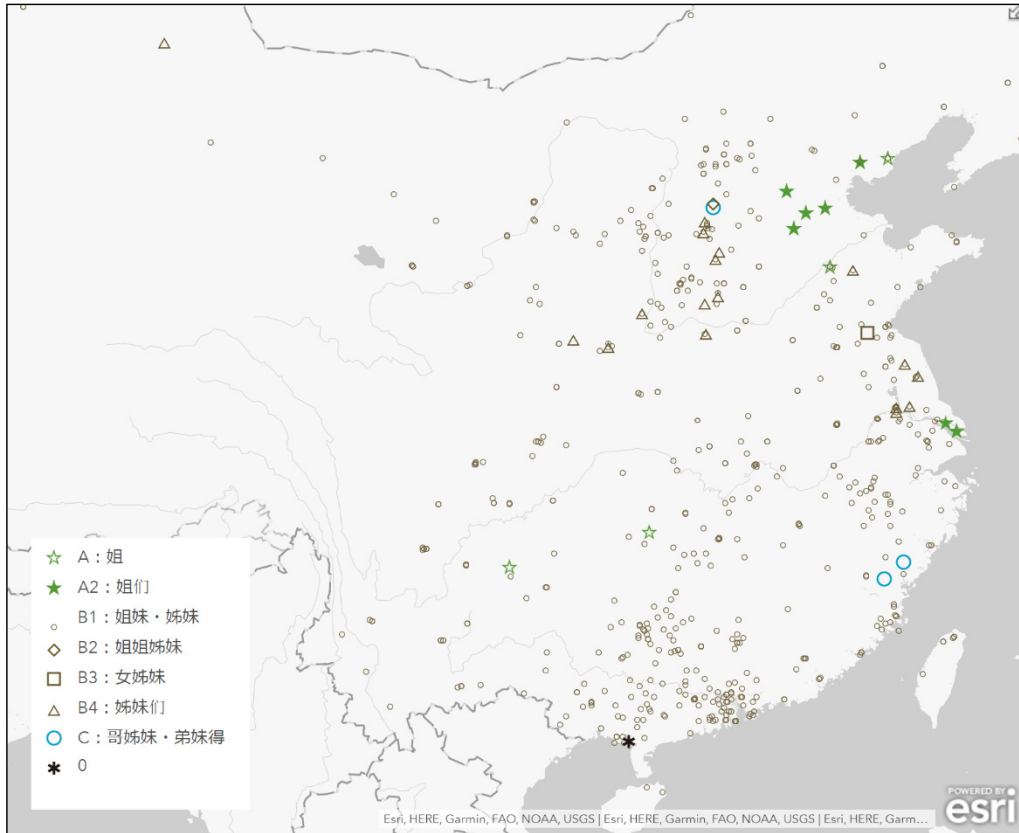


図7: 「姉妹」を表す語の分布

8. 「兄弟姉妹」

8.1. 「兄弟姉妹」を表す語の分類

A1: 姊妹、姊妹们、姊妹子、姊妹妹

A2: 姐妹

A3: 姊妹淘里、姊妹和里

A4: 姊们

A5: 姊妹家

B: 兄弟、弟兄

C: 弟妹

0: その他（非表示）: 兄弟姊妹、弟兄姊妹

8.2. 「兄弟姉妹」を表す語の分布

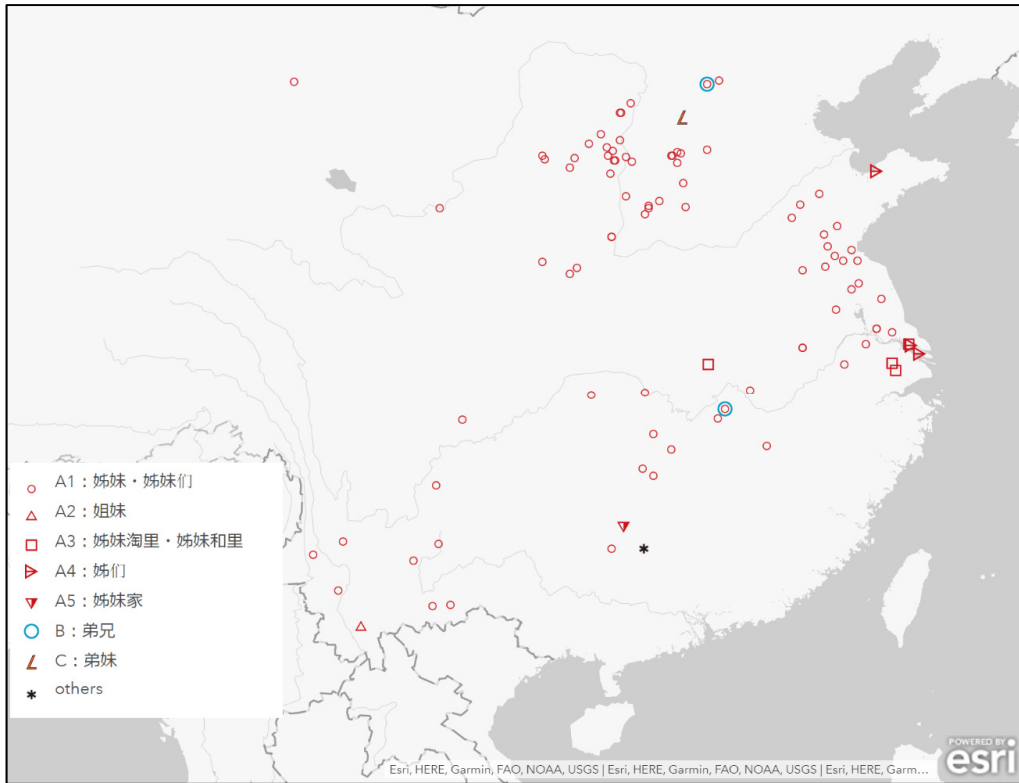


図8a: 「兄弟姉妹」を表す語の分布 (語形が「兄弟姉妹」のもの非表示)

兄弟姉妹を表す語は「兄弟姉妹」ないし「弟兄姊妹」などとして表され、「兄」「弟」「姉」「妹」のそれぞれを表す語を語中に有するのが一般的である (このタイプは地図中では非表示にしている)。

北方を中心に語中に「兄」「弟」を示す語がなく「姊妹」のみとなる形式が多く見られる (Aタイプ)。またごく一部であるがB「兄弟」「弟兄」やC「弟妹」のように、「姊」「妹」を持たない語形も見られる。C「弟妹」は山西忻州に見られる語形で、同一の語形を「兄弟」「姊妹」「兄弟姊妹」のいずれを表すにも用いられるという。

忻州: 弟妹[ti mei] (温端政 1985:93)

9. 中国語方言における「兄」「弟」「姉」「妹」の語構成について

以下、中国語方言を全体として見た場合の「兄」「弟」「姉」「妹」を表す語の構成に関する特徴について述べておきたい。

中国語方言において「兄」「弟」「姉」「妹」を表す語の語幹はそれぞれ「兄・哥」「弟」「姉・姐」「妹」が主となる。

それぞれの語幹は単独で用いられる場合もあるが、語幹に対して（Ⅰ）前置成分（阿など）を加える（Ⅱ）後置成分（子・儿など）を加える（Ⅲ）重ね型にする、といった処理を加えることが「兄」「弟」「姉」「妹」全てに共通して見られる。

主に南方において、老（佬）/lao/が「兄」「弟」「姉」「妹」全てにおいて前置成分としても後置成分としても出現している。

例（括弧内にそれぞれの地図における分類を示す）

○前置成分の老

兄：老兄（A4）	姉：老姊（A7）
弟：老弟（C5）	妹：老妹（D1）

○後置成分の老（佬）

兄：兄佬（A5）	姉：姊佬（A8）
弟：弟佬（C3）	妹：妹佬（D2）

年齢を軸とした共通性（「兄」・「姉」と「弟」・「妹」の対立関係）を示す語としては「大」と「小（細）」が見られる。ただし「大」と「小」は対称的に用いられているわけではなく、「大」は単独で用いられる場合や後置成分（語幹）となる場合もあるのに対し、「小」は基本的に前置成分（修飾成分）として用いられ単独の形式や語幹としては用いられない点が留意される。

○年齢の対立に関する「大」と「小（細）」：

兄：大哥（C6）・大（C1）・阿大（C3）	姉：大姊（A6）・大（C1）・阿大（C3）
↑↓	
弟：小弟・細弟（D2）	妹：小妹・細妹（B3）

性別を軸とした共通性には、兄弟・姉妹ともに年長者を示す語形が年少者を示す語形においても用いられるという点がある。すなわち、兄と姉を示す「兄」「姉」がそれぞれ弟と妹の語形中に見られる。その逆の年少者の語形が年長者の語形に用いられる事がない点も留意される。

○性別の対立に関する「兄」と「姉」

兄：兄（A1）	⇔	姉：姊（A1）
弟：兄（B1）・兄弟（B2）		妹：姊妹（C1）

また、「兄弟」を表すのに「兄」「弟」のいずれかの要素だけを用いる場合や、「姉妹」を表すのに「姉」を表す語のみを用いる場合なども年齢の枠組を超えたつながりを示していると考えることができる。

当然「兄弟姉妹」を「姉妹」のみ表すという場合は性別の枠組を超えたつながりを示す事になる。

以上のような例から、一口に「中国語方言の sibling term 体系は単一である」と言っても方言全体から見た場合や個々の方言内部の語構成を詳細に見た場合には多様な状況があることがわかる。

10. おわりに

以上中国語方言の中に見られる「兄」「弟」「姉」「妹」の分類と分布状況を概観し、それらの共通性を示す文法現象の主なものについて見た。最初に述べたように中国語方言は個別的に見ると全て「兄」「弟」「姉」「妹」を峻別するタイプとなっているが、全体を俯瞰すると共通の語幹や語形が存在している事がわかる。この事は中国語内部における「兄」「弟」「姉」「妹」の成立過程や他言語との関係を考察する上で示唆的であり、さらに検討を加える必要がある。

本稿で見た分類は主に漢字語形に着目した分類であり、一つの観点を提示したに過ぎない。今後は音声形式なども考慮し様々な異なる観点から分類・分析を進めていく必要がある。その上で方言間の変化の過程を考察し、文献資料などと照合しながら、周辺言語との関連や他の親族呼称も視野に入れた sibling term 体系の発展過程およびその類型・普遍性などについても考察していきたい。

参考文献/References

- 秋谷裕幸 (2020)《闽东四县市方言调查研究》上海：上海教育出版社。
福嶋秩子 (2021)「日本語におけるキョウダイ名」『地理言語学研究』1:115-122。
岩田礼 (1989)《中国江蘇・安徽・上海両省一市境内親族称谓詞の地理分布》《開篇》 単刊 No.1。
岩田礼 (2012)《漢語方言解釈地図（続集）》東京：好文出版。
郭明昆 (1962)『中國の家族制及び言語の研究』東方學會。
松本克己 (2006).『世界言語への視座—歴史言語学と言語類型論—』東京：三省堂。
王琪 (2008)《上古汉语称谓研究》北京：中华书局。
温端政 (1985)《忻州方言志》北京：语文出版社。
吴启主 (1998)《常宁方言研究》长沙：湖南教育出版社。

杨蔚 (1999) 《沅陵乡话研究》长沙：湖南教育出版社.

出版情報

採用決定日：2022年8月31日

Research article

Terms for ‘elder brother’ and ‘elder sister’ in Kra-DaiHIRANO, Ayaka
Osaka University

Abstract: In this paper, I aim to analyze terms that refer to elder brothers and elder sisters in Kra-Dai languages geolinguistically. There are two types of elder-sibling term systems in Kra-Dai languages: the ‘elder sibling’ type and the ‘elder brother/elder sister’ type. The ‘elder sibling’ distinctions are found around Guiyang in China and Southeast Asia. The most widespread form for ‘elder sibling’ is derived from Proto-Tai *bi^{B2}. The ‘elder brother/elder sister’ distinctions are found mainly in southern China and Hainan Island. All distinctions on Hainan Island are of the ‘elder brother/elder sister’ type. Although ‘elder sibling’ distinctions are more widely distributed than ‘elder brother/elder sister’ ones, words that mean ‘elder sibling’ have fewer variations. Some languages, in particular Central Tai branch languages, use both systems together. The most widespread forms for ‘elder brother’ and ‘elder sister’ might be derived from the Chinese 哥 (ge) ‘elder brother’ and 姊 (zi) ‘elder sister’, respectively. Kra-Dai has many elements borrowed from Chinese for elder siblings. Not only 哥 (ge) and 姊 (zi) have been borrowed; elements similar to 阿 (a), which is a prefix for elder people in Chinese, are also found in some languages. Although Chinese has a strong influence on elder-sibling terms in Kra-Dai languages, Hlai languages have fewer Chinese-derived elements than others. For this reason, the forms found in Hlai languages might be closer to the ancient form.

Keywords: Kra-Dai; sibling terms; loan words**1. Introduction**

This paper presents a geolinguistic analysis of words referring to elder siblings, such as ‘elder brother’ and ‘elder sister’, in Kra-Dai languages. In some Kra-Dai languages, there is no distinction between ‘elder brother’ and ‘elder sister’, and there is a word for ‘elder sibling’. On the other hand, other Kra-Dai languages have separate words for ‘elder brother’ and ‘elder sister’. In this paper, I first create linguistic maps of the distribution of the elder sibling term system. After looking at the distribution of systems according to whether or not there is a gender distinction, I examine the distribution of each form of ‘elder sibling’, ‘elder brother’, and ‘elder sister’ found in Kra-Dai languages.

HIRANO, Ayaka. 2022. Terms for ‘elder brother’ and ‘elder sister’ in Kra-Dai. *Studies in Geolinguistics* 2: 153–164. doi: <https://doi.org/10.5281/zenodo.7121716>

2. System of Elder Sibling Terms

Figure 1 shows the distribution of the ‘elder brother/elder sister’ type languages and the ‘elder sibling’ type languages. The ‘elder brother/elder sister’ type languages have gender-differentiated words for ‘elder brother’ and ‘elder sister’, whereas the ‘elder sibling’ type languages do not distinguish between the two genders. As shown in Figure 1, the geographical distribution of the ‘elder sibling’ type is wider in Southeast Asia and around Guiyang Province, China. The ‘elder brother/elder sister’ type is more common in southern China and Hainan Island.

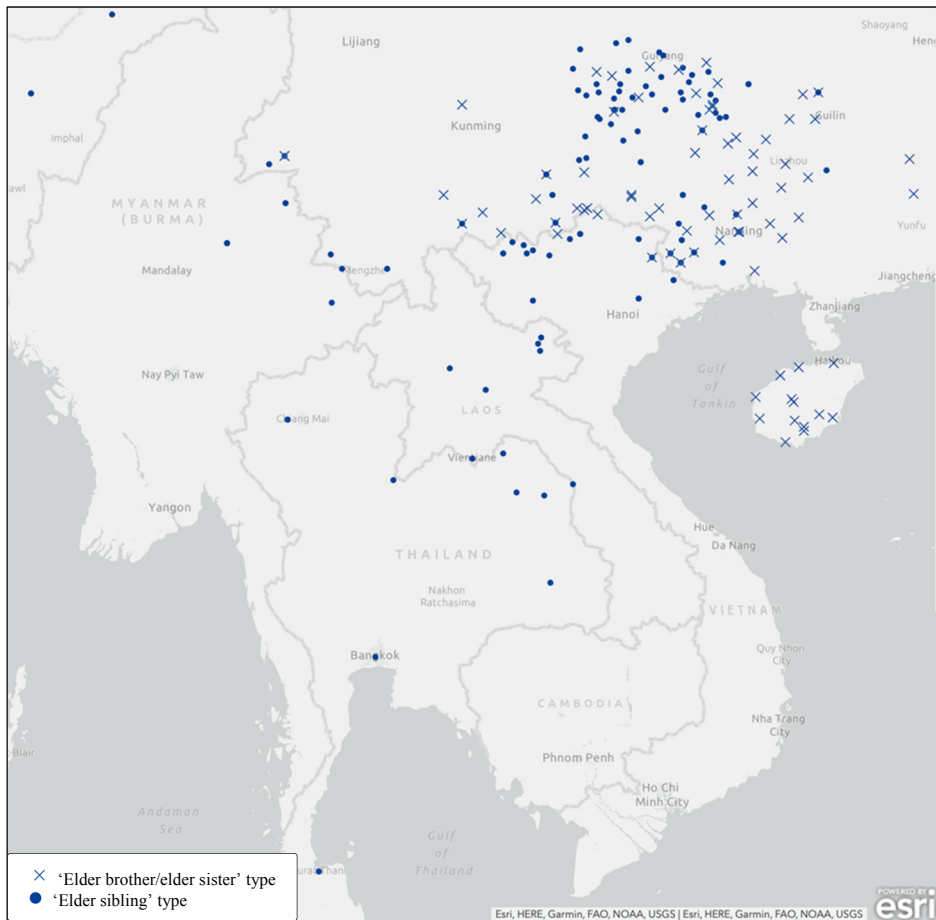


Fig. 1 Distribution of the elder sibling term system in Kra-Dai languages

Each language subclass in the distribution illustrated in Figure 1 shows a different situation. As shown in Figure 2, many languages in the Southwestern Tai languages have a system that does not distinguish elder siblings by gender, while some languages

distributed in China distinguish elder siblings by gender. Some languages use the systems of ‘elder sibling’ and ‘elder brother/elder sister’ together.

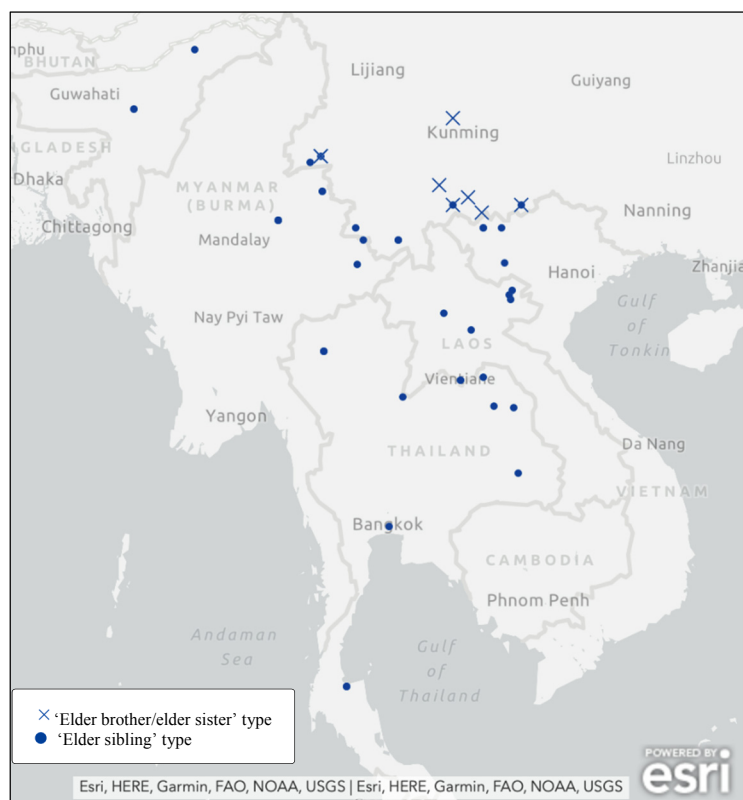


Fig. 2 Distribution of the elder sibling term system in Southwestern Tai languages

Central Tai languages also have ‘elder brother/elder sister’ type languages, which are distributed in China, as shown in Figure 3. However, the boundary between the ‘elder brother/elder sister’ and ‘elder sibling’ types is more ambiguous than in the Southwest Tai languages, and some languages use the two systems together.

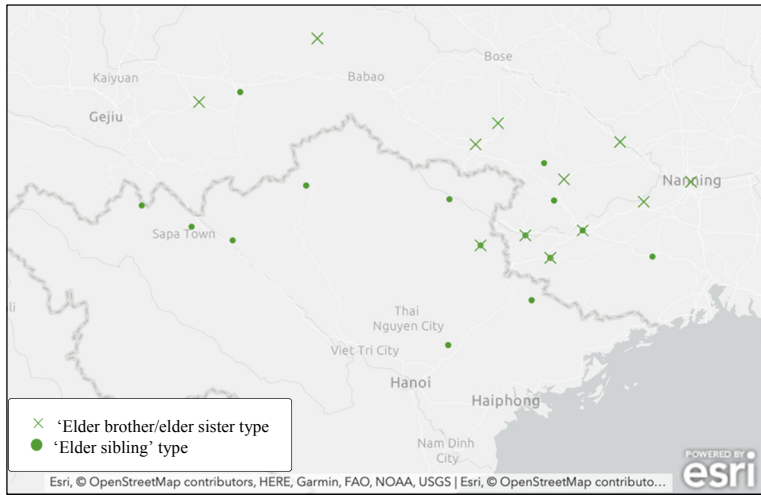


Fig. 3 Distribution of the elder sibling term system in Central Tai languages

In Northern Tai languages, the distribution area of the ‘elder brother/elder sister’ type and the ‘elder sibling’ type is divided. In Bouyei languages, the ‘elder sibling’ type is dominant, while in the northern Zhuang dialects, the ‘elder brother/elder sister’ type is dominant, resulting in the distribution shown in Figure 4. In some northern Zhuang dialects, the ‘elder brother/elder sister’ type system and the ‘elder sibling’ type system are used together.

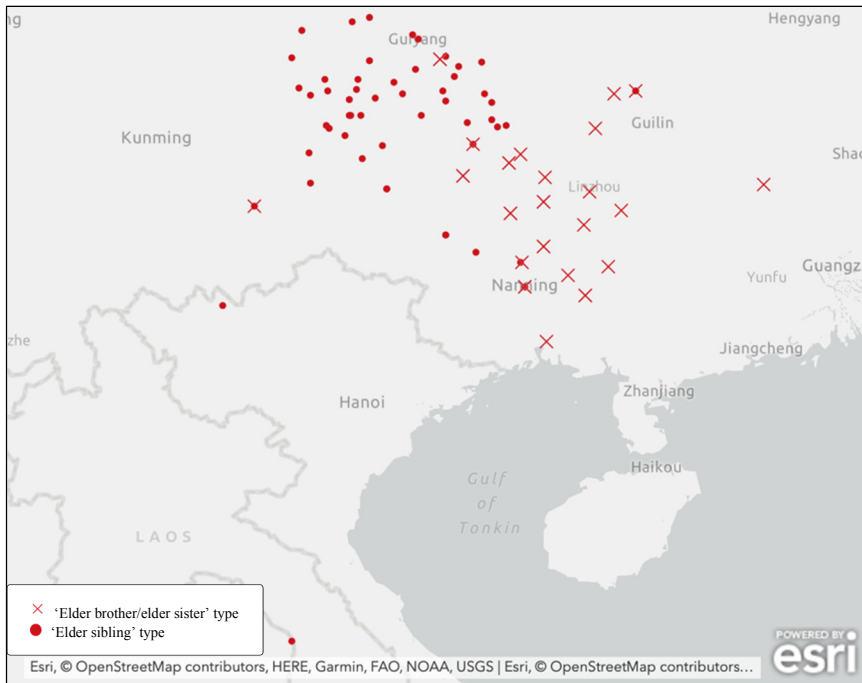


Fig. 4 Distribution of the elder sibling term system in Northern Tai languages

In Kam-Sui and Kra languages, the 'elder brother/elder sister' type predominates, as shown in Figure 5.

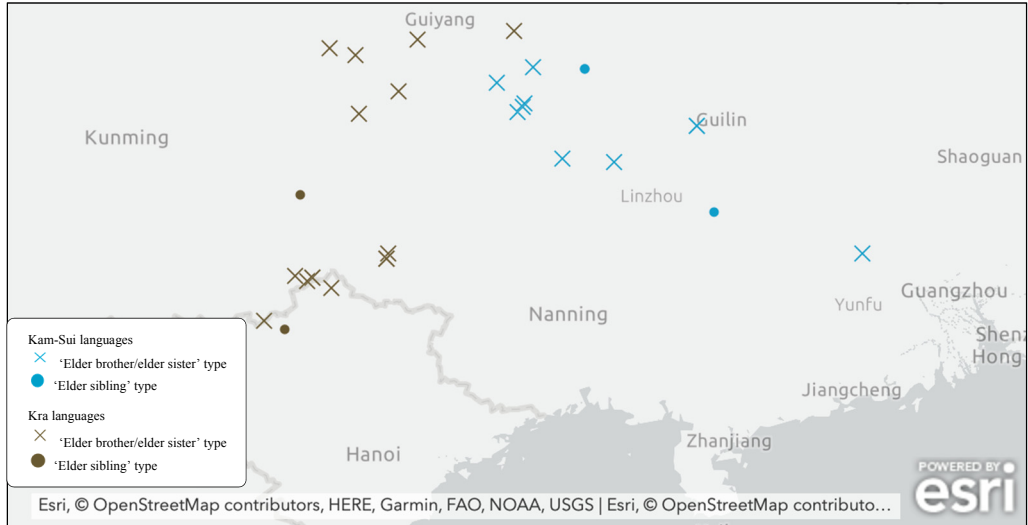


Fig. 5 Distribution of the elder sibling term system in Kam-Sui and Kra languages

Hlai and Ong Be languages spoken on Hainan Island are all of the 'elder brother/elder sister' type system, as shown in Figure 6, as far as I was able to ascertain at this time.

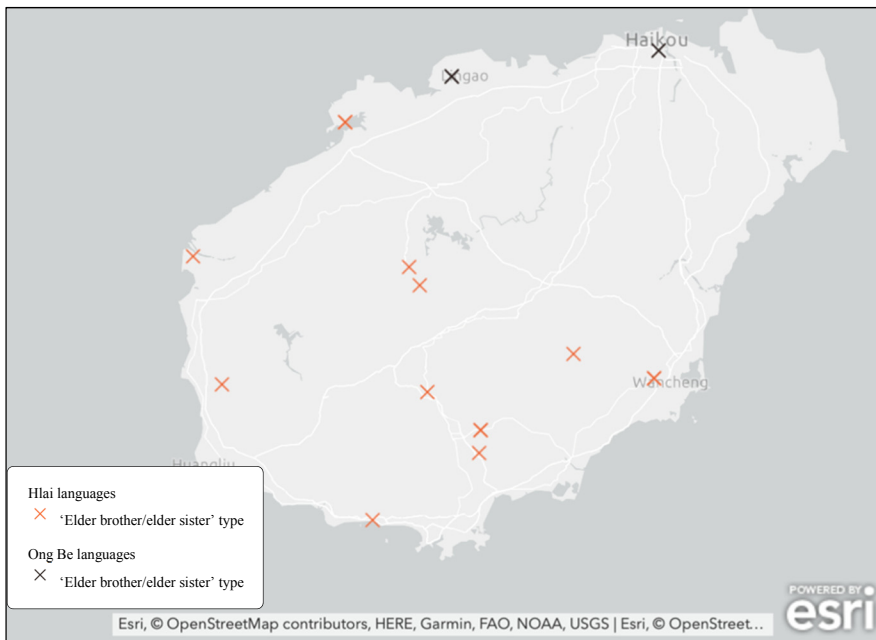


Fig. 6 Distribution of the elder sibling term system in Hlai and Ong Be languages

So far, I have looked at the distribution of languages with regard to whether they distinguish elder siblings by gender or not. In languages distributed in the current territory of China, the ‘elder brother/elder sister’ type is predominant, and the influence of Chinese cannot be denied. However, as seen in the Northern Tai language group, there are cases where the ‘elder sibling’ type is dominant even within Chinese territory. There are also several places, mainly in the Sino-Vietnamese border region, where the two systems are used together.

As seen in the next section, the ‘elder sibling’ type has a wide distribution area, but the number of forms is limited, while the ‘elder brother’ and ‘elder sister’ forms have many types and complex distribution.

3. ‘Elder Sibling’

The ‘elder sibling’ type, which makes no distinction based on gender, is widely distributed, but the forms are more uniform than ‘elder brother’ and ‘elder sister’, which I discuss in the next section and beyond. As shown in Figure 7, the majority are PI/PHI types that use a form similar to *pi* or *phi*. This form can be derived from Proto-Tai **bi*^{B2} (Li 1977). In addition, TA types are found in two languages, *ta:i*⁴ is found in one language, and *tja*⁴ *nji*⁴ is found in one language.

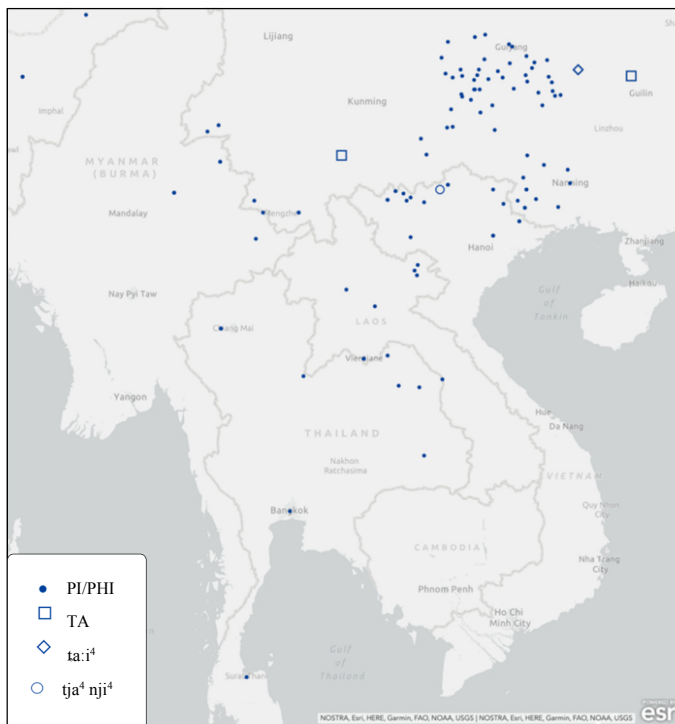


Fig. 7 Distribution of ‘elder sibling’

In some languages, it is reported that the meaning corresponding to ‘elder brother’ or ‘elder sister’ can be expressed by modifying the ‘elder sibling’ element with elements expressing gender. Figure 8 shows the distribution of modifiers in the languages of the ‘elder sibling’ type that can be used to express ‘male elder sibling’. The most widespread type is the CAAI type.

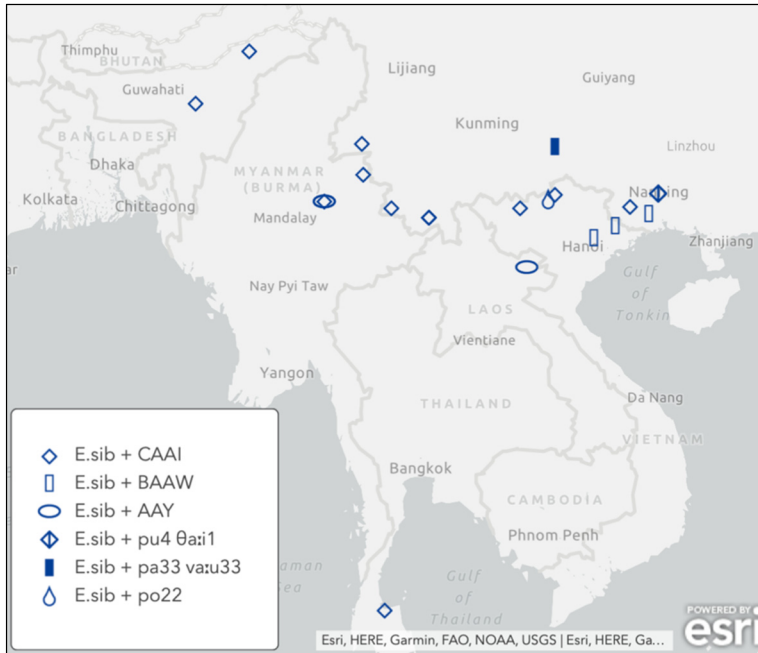


Fig. 8 Distribution of male modifiers for ‘elder sibling’

The distribution of modifiers used to express ‘female elder sibling’ is shown in Figure 9. Although the number of languages I was able to examine is limited, compared to the distribution of modifiers for ‘male’ shown in Figure 8, none of the forms can be observed as widely as the CAAI form for ‘male’.



Fig. 9 Distribution of female modifiers for ‘elder sibling’

4. ‘Elder Brother’

As shown in Figure 10, the most widespread form is similar to *ko*. This KO type is thought to be a borrowing of the Chinese word 哥 (*ge*) ‘elder brother’. The TA type is similar to one of the forms of ‘elder sibling’ observed in section 3, and the PI/PHI type is the most widespread form of ‘elder sibling’ seen in section 3, but it is used in the sense of ‘elder brother’ in some languages. The PI/PHI type form as ‘elder brother’ is sometimes used with the KO type form in one language. The (W)ENG/WING type is predominant in the Hainan Island Hlai languages.

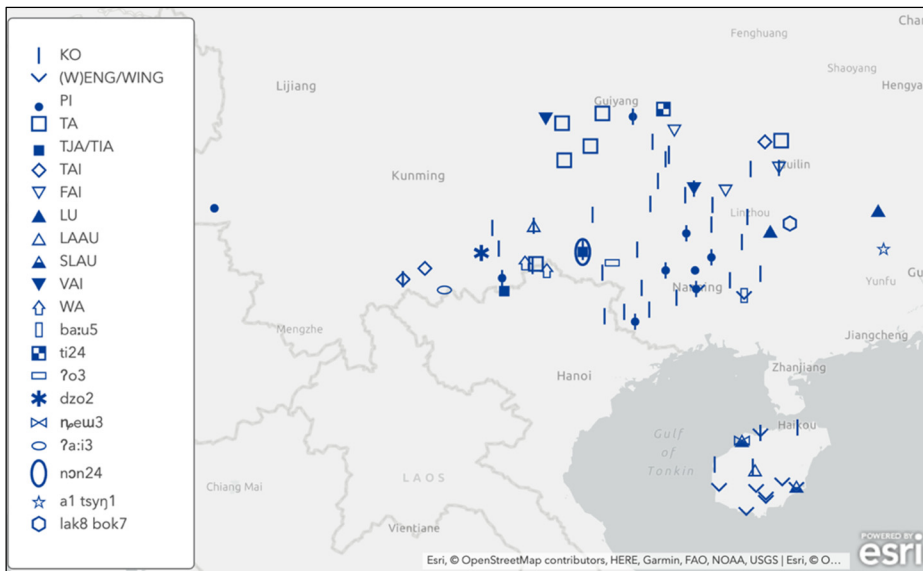


Fig. 10 Distribution of ‘elder brother’

A form prefixed with a kind of prefix similar to the Chinese 阿 (a) is found at 10 sites. As Figure 11 shows, none are found on Hainan Island; all are found on the mainland. 阿 (a) in Chinese is a prefix for elder people. Thus, it may be due to language contact with Chinese.



Fig. 11 Distribution of the ‘A+elder brother’ form

5. ‘Elder Sister’

As shown in Figure 12, the SE/CHE type is the most widespread. The SE/CHE type may be a borrowing from 姊 (zi) since 姊 (zi) is pronounced like *tse* in the Chaozhou dialect of Chinese¹. On Hainan Island, the TA and PI type are also observed as forms of ‘elder sister’. The EI/AU and KHAU types are predominant on Hainan Island.

¹ In southern Vietnamese dialects belonging to the Austro-Asiatic language family, ‘elder sister’ is *ché*, similar to the SE/CHE type of Kra-Dai; Lê Ngọc Trữ (1993) suggests that it may be a loan word from Chaozhou or Cantonese dialects.

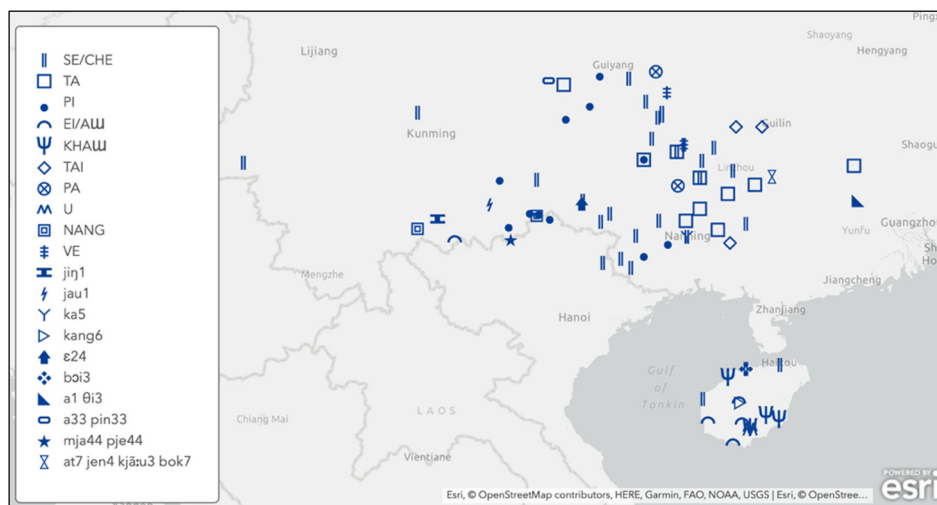


Fig. 12 Distribution of 'elder sister'

Like 'elder brother' in section 4, 'elder sister' is often observed in a form with an attached prefix similar to the Chinese 阿 (a), which is found at 7 sites. Figure 13 shows its distribution; like the A+ 'elder brother' form, it is not observed on Hainan Island.



Fig. 13 Distribution of the 'A+elder sister' form

6. Conclusion

There are two systems in Kra-Dai; the 'elder sibling' system and the 'elder brother/elder sister' system. Some languages use both of these systems together. The most widespread forms of 'elder brother' and 'elder sister' appear to be loan words from Chinese, and the influence of Chinese can be seen strongly in these forms.

However, in the languages of Hainan Island, especially in the Hlai languages, the use of Chinese-derived forms is limited. In other words, the forms found in the Hlai languages may be closer to the original Kra-Dai lexical form.

Data sources

- 平野綾香(2021)『ベトナムランソン省チャンディン県ヌン語の記述研究』大阪大学, 博士論文.
北京大学中国语言文学系语言学教研室(1989)《汉语方音字汇》北京:文字改革出版社出版.
薄文泽(1997)《佯僳语研究》上海:上海远东出版社.
薄文泽(2003)《木佬语研究》北京:中央民族大学出版社.
广西壮族自治区少数民族语言文字工作委员会(2008)『广西民族语言方音词汇』北京:民族出版社.
贺嘉善(1983)《仡佬语简志》北京:民族出版社.
小坂隆一等(1998)《仡央语言词汇集》贵阳:贵州民族出版社.
李锦芳(1999)《布央语研究》北京:中央民族大学出版社.
李锦芳(2006)《西南地区濒危语言调查研究》北京:中央民族大学出版社.
李如龙(2012)《茶洞语研究》北京:民族出版社.
李云兵(2000)《拉基语研究》北京:中央民族大学出版社.
梁敏(1980)《毛难语简志》北京:民族出版社.
梁敏(1980)《侗语简志》北京:民族出版社.
梁敏等(1997)《临高语研究》上海:上海远东出版社.
梁敏等(2002)《普标语研究》北京:民族出版社.
毛宗武等(1982)《瑶族语言简志》北京:民族出版社.
欧阳觉亚(1983)《黎語调查研究》北京:中国社会科学出版社.
欧阳觉亚(1998)《村語研究》北京:民族出版社.
王均(1980)《仡佬语简志》上海:上海远东出版社.
杨通银(2000)《莫语研究》北京:中央民族大学出版社.
喻翠容(1980)《布依语简志》北京:民族出版社.
喻翠容、罗美珍编著(1980)《傣语简志》北京:民族出版社.
张均如(1980)《水语简志》北京:民族出版社.
张均如等(1999)《壮语方言研究》成都:四川民族出版社.
周耀文等(2001)《傣语方言研究》北京:民族出版社.
Hudak, Thomas John (1994) *William J. Gedney's Southwestern Tai dialects: Glossaries, texts, and translations*. Ann Arbor, Michigan: Center for South and Southeast Asian Studies, University of Michigan.
Hudak, Thomas John (1995) *William J. Gedney's Central Tai dialects: Glossaries, texts, and translations*. Ann Arbor, Michigan: Center for South and Southeast Asian Studies, University of Michigan.

- Hudak, Thomas John (1997) *William J. Gedney's Tai dialect studies: Glossaries, texts, and translations*. Ann Arbor, Michigan: Center for South and Southeast Asian Studies, University of Michigan.
- Hudak, Thomas John (2008) *William J. Gedney's Comparative Tai source book*. Honolulu: University of Hawai'i Press. URI: <https://www.jstor.org/stable/20532978>
- Morey, Stephen (2005) *The Tai languages of Assam: A grammar and texts*. Canberra: Pacific Linguistics, The Australian National University. doi: <https://doi.org/10.15144/PL-565>
- Pittayaporn, Pittayawat (2009) *The phonology of Proto-Tai*. Doctoral dissertation, Cornell University.
- Самарина, Ирина Владимировна (2011) *Языки гэлао: материалы к сопоставительному словарю кадайских языков*. Москва: Academia.

References

- Lê Ngọc Trụ (1993) *Tâm-Nguyên Tự-Điển Việt-Nam [Vietnamese etymological dictionary]*. Hồ Chí Minh: Nxb. Thành Phố Hồ Chí Minh.
- Li, Fang-Kuei (1977) *A handbook of comparative Tai*. Honolulu: University of Hawai'i Press. URI: <https://www.jstor.org/stable/20006684>

Publication history

Date accepted: 15 September 2022

論文

Kra-Daiにおける「弟・妹」語彙

富田愛佳
大阪樟蔭女子大学

Terms of ‘younger brother’ and ‘younger sister’ in Kra-Dai

TOMITA, Aika
Osaka Shoin Women’s University

Abstract: The Kra-Dai languages can be divided into two types: ‘younger sibling’ type and ‘younger brother/younger sister’ type languages. The ‘younger sibling’ type is most widely distributed in present, while the ‘younger brother/younger sister’ type is common in the languages that retain old Kra-Dai features (e.g. the Kra languages). The common and widely distributed form of ‘younger brother/sister’ is the NONG form, which means ‘younger sibling’. In the both case of ‘younger brother’ and ‘younger sister,’ an additional gender-indicative form often follows the NONG form in the Tai languages: TSAJ ‘man, male person’ (or BAAW ‘young unmarried man’) and JING ‘woman’ (or SAAW ‘young unmarried woman’, BWK ‘girl’). The Central Tai languages are characterized by the use of BAAW and SAAW as additional forms. The Northern Tai languages is characterized by the use of BWK as an additional form of ‘younger sister’. In Li (Hlai), the RUNG form, which come from Proto-Hlai, was found. The RUNG form means ‘younger sibling’, and the form is the characteristic of Li (Hlai) language. In Buyang, the ZO (‘younger brother’) and ON (‘younger sister’) forms, which trace back to Proto-Kra, were found. Thus, the word forms other than the widely distributed NONG form often have different origin from the Tai languages. In addition, it was pointed out that in Buyang, Proto-Kra ‘younger sister’ is thought to have changed to mean ‘younger sibling’, and that the MOI form and the additional MUI form of ‘younger sister’ may have been influenced by Austro-Asiatic languages, while these forms may also have been borrowed from Chinese.

キーワード : Kra-Dai ; キョウダイ語彙 ; Proto-Kra
Keywords: Kra-Dai; sibling terms; Proto-Kra

1. はじめに

本稿は、Kra-Dai 諸言語における「弟・妹」語彙の地理的分布の特徴を明らかにすることを目的とする。

「弟・妹」言語地図作成にあたっては、Bibliography に記載の言語資料から当該語彙の形式を遠藤光暁氏、平野綾香氏、筆者の3名で Excel に入力し、esri の ArcGIS を使用して地図描画した。「弟・妹」の語彙項目の分類やその考察については筆者がその責を負う。

2. 「年下キョウダイ」型と「弟／妹」型の分布

Kra-Dai において「弟・妹」を指示する語彙は、大きく「年下キョウダイ」型と「弟／妹」型に分けることができる。

「年下キョウダイ」型とは、そのキョウダイの性別を問わず、年下のキョウダイであることを意味する形式を持つタイプである。「弟」「妹」の語彙形式が完全に一致していたり、あるいは「年下キョウダイ」を意味する形式に性別を表す形式を付加して「弟」「妹」を表したりする。

「弟／妹」型とは、「弟」「妹」の語彙に、それぞれ別の形式を使用するタイプである。

Kra-Dai における「年下キョウダイ」型と「弟／妹」型の分布は、図1の通りである。



図1: 「年下キョウダイ」型と「弟／妹」型

「年下キョウダイ」(ySib) 型は、海南島を含む広い範囲に分布する。「弟／妹」(yBro/ ySis) 型は、Zhuang のごく一部、Kam-Sui 諸語、Kra 諸語、Li、Lingao に見られる。これらとは別に、「弟」語彙のデータだけで、「妹」はデータが

ない地点 (no_ySis) もある。そのような地点は「年下キョウダイ」にあたる一つの形式しかないと考えれば「年下キョウダイ」(ySib) 型と見なすことができるかもしれない。

3. 弟

3.1. 「弟」を示す形式の分布

第2節で確認したように、Kra-Daiの「弟・妹」語彙は「年下キョウダイ」型と「弟／妹」型に分けられるが、ここからは「弟・妹」を示す具体的な形式の分布について見ていく。

第3節では「弟」の語彙形式の分布を見る。形式を包括的に見るために、ここでは「年下キョウダイ」型と「弟／妹」型に分割することはしない。「弟」を表す形式をすべて扱ってまとめたのが図2である。各分類の具体的な語彙形式については、付録1を参照されたい。

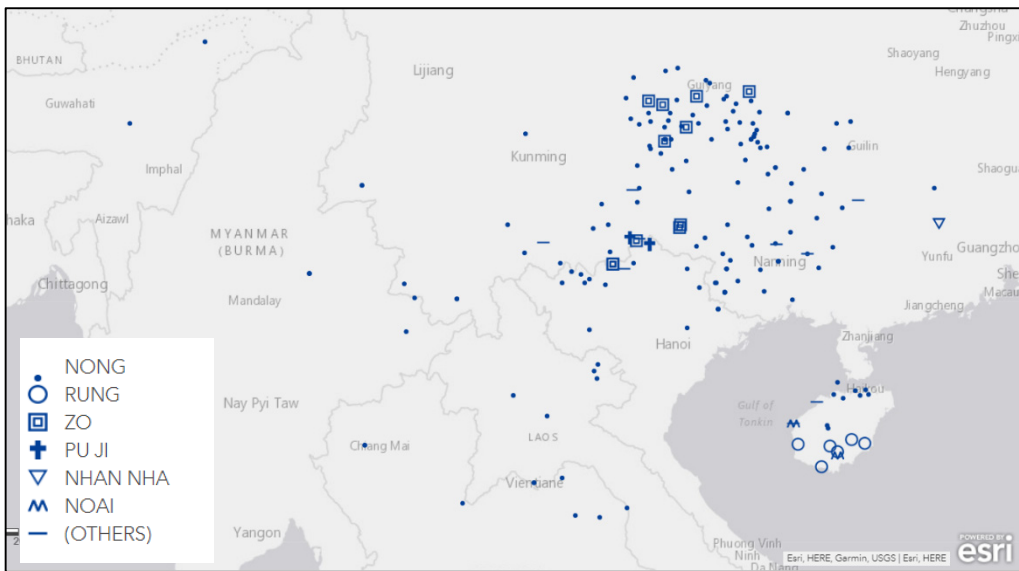


図2：「弟」を示す形式の分布

NONG型は、Proto-Tai *nɔŋ (Li 1977) ‘younger sibling’ に遡る形式で、もともと広汎に見られるが、海南島では海口周辺など一部に留まる。なお、付録1で nu4 (Sui 語・三都) を NONG 型の語形として分類しているのは、Sui 語で末子音に -ŋ を持つ語が少ない (Castro and Pan (eds.) 2015: 104) とされていることから、音韻論上の制約によって語末の -ŋ が脱落したと想定したためである。実

際、Burusphat, Wei and Edmondson (2003) においても漢語借用語 (农业, 农场) を除き、nun という音配列の見出し語は挙がっていない。

RUNG 型はすべて Li (Hlai) 語であり、海南島南部に分布する。初頭子音が付録 1 の語彙形式のように Proto-Hlai *r > r, g, h (Norquest 2007) の規則にしたがうことから、Proto-Hlai から継承された形式であると考えられる。

ZO 型、PU JI 型は両方ともすべて Kra 諸語であり、貴陽周辺とベトナム国境付近に分布する。ZO 型は Buyang と Gelao、PU JI 型は Gelao の地点である。Ostapirat (2000) では、Proto-Kra *ɟauB ‘brother (younger)’ という形式が示されており、ZO 型はこれに遡るものと推測される。PU JI 型においては、語頭の ʔa³³ や ʔɒ³³ などの接頭辞と思われる形式のみ「妹」と共通している。

NHAN NHA 型は Kam-Sui、NOAI 型は Li (Hlai) で、それぞれ孤立的に分布する。

その他 (OTHERS) に分類した形式のうち、on³³ pa³³ va:u³³ (Buyang, Kra) に見られる on³³ は Proto-Kra *ʔonC ‘sister (younger)’ (Ostapirat 2000) に遡る語形であると考えられる。同地点の「妹」は on³³ pa³³ mai¹¹ であるため、Proto-Kra で「妹」を表していた on は、共時的には「年下キョウダイ」を表す語形となり、キョウダイの性別は表さなくなっている¹。また、同じくその他の形式に分類した、tak⁸ tai (Zhuang, 北方タイ諸語) は、Proto-Kra *taiA ‘brother (elder)’ (Ostapirat 2000) に遡る tai に接頭辞 tak⁸ がついた形式であろう²。

3.2. 「弟」において2音節以上の形式を構成する要素

付録 1 からわかるとおり、「弟」を示す形式の中には、2 音節以上であるものが多い。これらは、基本的に「年下キョウダイ」型であると考えられ、3.1 の図 2 で分類のラベルにした NONG、RUNG、ZO 等に対応する「年下キョウダイ」の形式に、性別「男」を意味する要素が付加されている場合が多くある。また、性別を示す要素以外でも、何らかの形式が付加されている場合がある。これらの付加形式の分布をまとめたのが図 3 である。

もっとも広く分布するのは TSAJ 型で、北方タイ諸語、南西タイ諸語を中心に雲南省から広西チワン族自治区の南部に分布し、図 2 の NONG の後部要素として出現する。BAAW 型も NONG の後部要素となるが、こちらは広西チワン族自治区のベトナム国境付近の中央タイ諸語に分布する。Hudak (2008) では、TSAJ にあたる語形は ‘man, male person’、BAAW にあたる語形は ‘young unmarried man’ の項目として北方・中央・南西タイ諸語全般で広く見られる。

¹ 当該地点の「兄」「姉」は共通して pi⁴⁵ という形式を持ち、これが共時的には「年上キョウダイ」を表すと考えられる。一方、Ostapirat (2000) は Proto-Kra の再建形として *taiA ‘brother (elder)’ / *ɟauB ‘brother (younger)’ / *piC ‘sister (elder)’ / *ʔonC ‘sister (younger)’ を挙げており、pi⁴⁵ は歴史的には「姉」の形式である。このように、当該地点では歴史的に「姉」の形式が「年上キョウダイ」の意味に、「妹」の形式が「年下キョウダイ」の意味に変化している。

² Burusphat and Qin (2006) によると、接頭辞 tak⁸ は同年代男性につける接頭辞であるという。

ほかに NONG の後部要素となるのは PAAN 型で、貴州と広西の境界付近の Kam-Sui 諸語で見られる。一方、NONG の前部要素になるのが TAK 型で、南寧市の東方に分布する Zhuang (北方タイ諸語) で見られる。TAK は注 2 で述べたように同年代男性につける接頭辞であろう。

LA 型、?A 型はそれぞれ、図 2 の ZO 型、PU JI 型の前部要素としていずれも Kra 諸語の Gelao で見られる。この要素は、同地点の「妹」語彙にも共通する形式があることから、性別を示す形式ではなく、接頭辞ではないかと考えられる。

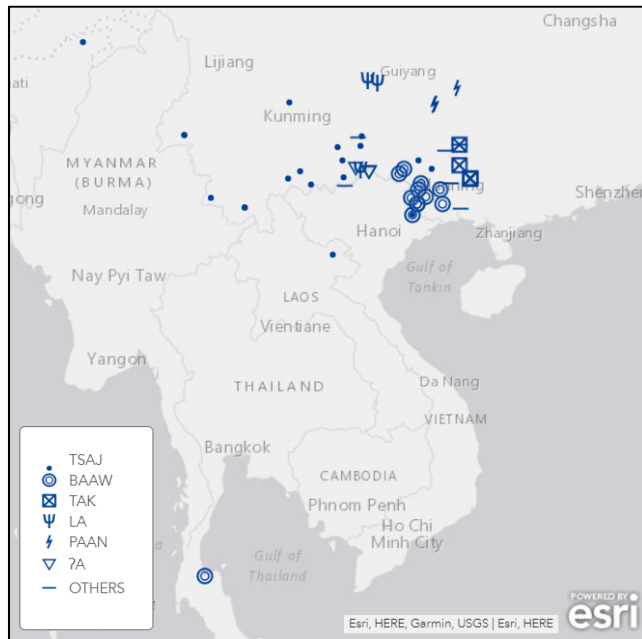


図 3 : 「弟」に見られる付加形式

4. 妹

4.1. 「妹」を示す形式の分布

「妹」を表す形式の分布をまとめると、図 4 のようになる。各分類の具体的な語彙形式については、付録 2 を参照されたい。

NONG、RUNG 型は「弟」の分布でも見られた形式で、「年下キョウダイ」を表す形式だと考えられる。NONG 型はタイ諸語、Kam-Sui 諸語を中心に広く見られ、Li (Hlai) や Lingao などにもごく一部が分布する。RUNG 型はすべて Li (Hlai) である。

ON 型は、3.1 節の最後で言及したように Proto-Kra *ʔonC ‘sister (younger)’ (Ostapirat 2000) に遡る語形である。すべて Kra 諸語である Buyang のデータで

あるが、3地点のうち2地点では「弟」としてやはり Proto-Kra に遡る ZO 型の形式を持っており、Proto-Kra の「弟」「妹」の形式を継承している。

MOI 型は、「弟」では出てこなかった形式であるため、「弟／妹」型の語彙であると見なせる。分布は、貴陽周辺では Gelao、桂林・南寧周辺では Zhuang (中央タイ諸語) と Kam-Sui 諸語、そして海南島では Lingao、のようにばらつきがある。語形も *mui*~*moi* (Zhuang)、*mei* (Kam-Sui)、*məi* (Lingao)、*məu* (Gelao) のように母音に違いがある。SEAlang Mon-Khmer etymological dictionary によると、Headley (1997) では Khmer ‘mei’ に対して意味的に「妹」と重なるような ‘term of address to a younger woman or to a woman of lower status’ という語釈が与えられており、MOI 型はオーストロアジアの語形と関係する可能性もあるが、一方で漢語「妹」の借用とも考えられる。

NI、EN、MIA、VAI 型も「弟」では出てこなかった形式であり、「弟／妹」型の語彙であると考えられる。なお、MIA にあたる形式は、Hudak (2008) では南西タイ諸語で ‘wife’ の語彙形式としてよく現れる。

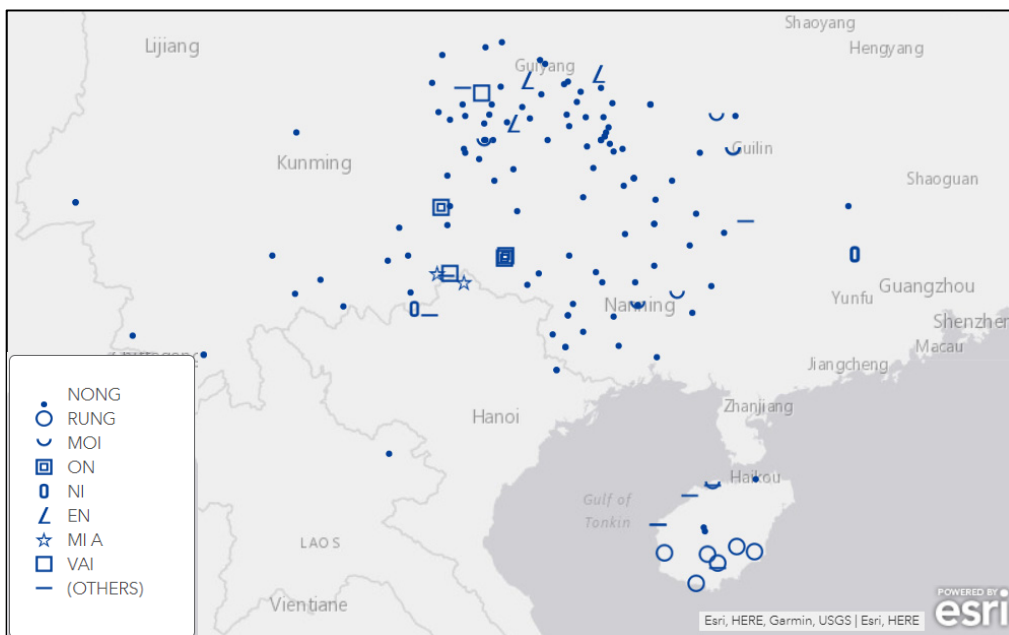


図4：「妹」を示す形式の分布

4.2. 「妹」において2音節以上の形式を構成する要素

2音節以上の「妹」語彙形式のうち、図4の分類の前部または後部についての付加形式によって分類し、分布をまとめたのが図5である。

JING 型、SAAW 型は雲南から広西の南部にかけて同程度に広く分布し、どちらのタイプも図4の NONG 型の後部要素となって、年下キョウダイの性別

を表す。JING 型が南西タイ諸語の Dai を中心とした分布であるのに対して、SAAW 型は、中央タイ諸語の Zhuang を中心とした分布である点が異なる。Hudak (2008) では、JING 型に相当する形式は ‘woman’ を意味し、南西および中央タイ諸語で見られる。また、SAAW 型の形式は ‘young unmarried woman’ を意味し、北方・中央・南西タイ諸語全般に見られる。SAAW 型の語彙形式は広く分布しているが、これを「妹」の性別を示すのに使うのが中央タイ諸語の特徴と言えそうである。

BWK 型は北方タイ諸語の Zhuang を中心に分布する。多くの場合 buk の音形をとり、NONG 型の後部要素となるが、中央タイ諸語の Zhuang の 1 地点 (隆安) のみ、NONG の前部要素となる語形があった。Hudak (2008) によれば、この BWK 型の形式は北方タイ諸語で ‘girl’ の意味であり、南西タイ諸語には見られない形式であるという。

LAAK 型はすべて Kam-Sui 諸語に分布し、ほとんどの場合、やはり NONG 型の後部要素になるが、1 地点のみ Biao で図 4 の NI 型の前部要素として現れる例があった。また、LAAK の後に bjik、ʔbi:k、bi:k、ʔja:k などを伴う地点も多く見られる。

ここまで見てきた JING、SAAW、BWK、LAAK 型とは逆に、図 4 の NONG 型の前部要素となるのが TA 型である。TA 型は BWK 型同様に北方タイ諸語の Zhuang に分布するが、その分布域は BWK 型より東側に寄っている。

MUI 型は Zhuang (北方タイ諸語)、Pubiao の互いに離れた 3 地点で見られ、図 4 の NONG 型あるいは VAI 型の前部または後部要素になっている。音形から、図 4 の MOI 型と関連する形式だと考えられる。MOI 型は「弟」には見られない形式のため、「弟／妹」型の「妹」を表す語彙だと考えられるが、MOI 単独ではなく NONG 型や VAI 型との組み合わせで「妹」を表す地点もあるということである。

ʔA 型は図 4 の MI A 型の前部要素として現れる。このタイプが観察される地点は、「弟」の図 3 でも ʔA 型であった Gelao の地点である。これらの地点では「弟／妹」型で、「弟」PUJI 型と「妹」MIA 型にそれぞれ接頭辞として ʔA 型が付いていると考えられる。

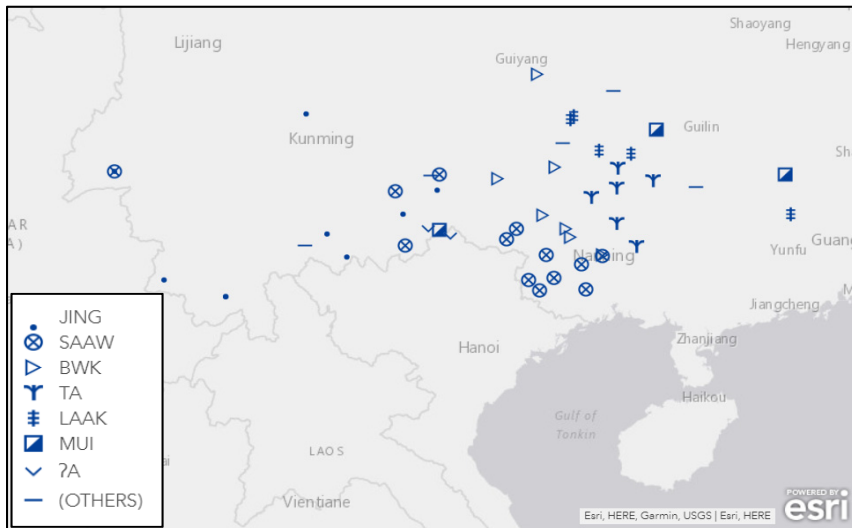


図5：「妹」に見られる付加形式

5. おわりに

本稿では、Kra-Dai 諸語の「弟・妹」を表す形式について、まず「年下キョウダイ」型と「弟／妹」型の言語に大きく分けられること、現在もっとも広汎に分布するのは「年下キョウダイ」型であるが、「弟／妹」型は Kra 諸語など Kra-Dai の古い特徴を残す言語に多いことを確認した。

「弟・妹」の各語形で共通して広く分布するのは、「年下キョウダイ」型の語形である NONG 型であった。「弟・妹」とも、この NONG 型の後部要素として性別を示す形式が付加されることも多くあり、この付加形式はタイ諸語においてごく一般的な語彙である TSAJ「男性」（あるいは BAAW「未婚男性」）、JING「女性」（あるいは SAAW「未婚女性」、BWK「少女」）だが、中央タイ諸語では付加形式として BAAW や SAAW を使う特徴が、北方タイ語では「妹」の付加形式として BWK を使用する特徴がある。

このほか、Proto-Hlai に遡る Li (Hlai) に特徴的な「年下キョウダイ」語形の RUNG 型や、Buyang における Proto-Kra の語形に遡る「弟」の ZO 型および「妹」の ON 型など、タイ諸語とは異なる語形も見られた。

さらに、Buyang で Proto-Kra の「妹」が「年下キョウダイ」の意味へ変化したと考えられること、「妹」の MOI 型および付加形式の MUI 型は漢語の借用とも考えられる一方で、オーストロアジア諸語の影響の可能性があると指摘した。

付録1: 「弟」を示す形式の分類

<NONG 型>

no:ŋ⁶, nɔ:ŋC2, nɔɔŋ⁴, nɔɔŋ⁵, nɔɔŋ⁶, nɔŋ³, nɔŋ⁴, nɔŋ⁵, nɔŋ⁶, nɔŋ⁸, nɔŋC2, nu:ŋ⁴, no:ŋ⁴, noŋ², noŋ⁴, noŋ⁶, noŋC2, nooŋ³, nooŋ⁶, nu:ŋ⁴, nu:ŋ⁴, nu⁴, nuaj³, nuaj⁴, nuaj⁴, nuaj⁶, nuaj⁴, nuaj²³, nuaj⁴, loŋ⁴ tsa:i², nauŋ⁴ caai², no:ŋ⁴ tea:i², nɔɔŋ⁶ baaw², nɔŋ⁴ tea:i², nɔŋ⁴ teai², nɔŋ⁴ tsa:i¹, nɔŋ⁴ tsa:i², nɔŋ⁴ tsa:i², nɔŋ⁵ caay⁴, nɔŋ⁵ saay⁵, nɔŋ⁶ baaw², no:ŋ³ ba:u⁶, no:ŋ³ ma:u⁵, no:ŋ⁴ ba:u⁵, nɔ:ŋ⁴ ba:u⁵, no:ŋ⁴ ba:u⁶, noŋ⁴ pa:n¹, noŋ⁴ tsoi², nooŋ³ baaw⁵, nooŋ⁶ baaw², nooŋ⁶ lay², nu:ŋ⁴ pu⁴ θa:i¹, nuaj⁴ t̄:ai¹, nuaj⁴ θa:i¹, nuaj⁴ θ:ai¹, nuaj⁴ θa:i¹, nuaj⁴ lə² ba:n¹, pi³ naaj⁴ loŋ¹, ro:i⁶ nu:ŋ⁴, tak⁸ nu:ŋ⁴, waj² nuaj⁴, ʔa:u¹ nu:ŋ⁴, ʔoo⁵ nooŋ⁶

<RUNG 型>

ru:ŋ¹, gu:ŋ¹, gu:ŋ⁴, huan⁴, hu:ŋ⁴

<ZO 型>

zo³⁵ (姉による呼称), zo³⁵ tja⁴⁴ (兄による呼称), zu¹¹, zu⁵³, a⁴⁴ zu⁴², la³¹ zu³¹, la³⁵ jau³³, tshə³³, tsə³¹, tsəu³¹, tɛu¹³

<PU JI 型>

ʔa³³ pu³³ zi³¹, ʔo³³ pu³³ ji³¹ nda³³, ʔo³³ pu³³ ji³¹

<NHAN NHA 型>

ŋan³ ŋa⁵, ŋan³ nia⁵

<NOAI 型>

nuai⁵, nɔ³ (姉による呼称)

<その他 (OTHERS)>

muaj⁴, mɔŋ⁴ tsa:i², ʔe:ŋ³, va:i³³, on³³ pa³³ va:u³³, jo² nji² po²², lak⁸ kjai³, tak⁸ tai, sok²

付録2: 「妹」を示す形式の分類

<NONG 型>

nɔŋ⁴, noŋ⁴, noŋ⁴ mjek⁹, nu:ŋ⁴, nu⁴, nuŋ⁴, nuŋ⁴ la:k⁸ bi:k⁸, nuŋ⁴ la:k⁸ ʔja:k⁷, nuŋ⁴ mu:i⁴, ta⁶ nu:ŋ⁴

<RUNG 型>

ru:ŋ¹, gu:ŋ¹, gu:ŋ⁴, hu:ŋ⁴, huan⁴

<MOI 型>

mei³⁵, məi⁴, məu¹³ tjan³¹ tjan³¹, mo:i⁶, mui³⁵, ta⁶ mu:i⁶

<ON 型>

un²⁴, ʔun³³, on³³ pa³³ mai¹¹

<NI 型>

ŋi⁴⁴, ni⁴, lak⁸ ni⁴

<EN 型>

e³³, en⁵⁵

<MI A 型>

ʔo³² mi³³ ʔã³³, ʔo³² mi³³ ʔɿ³³

<VAI 型>

va:i³³ mǝi⁴⁵, uai⁵⁵ pia⁵⁵

<その他 (OTHERS)>

a³³ an³³, a⁴⁴ tɛ⁴² (~tei⁴²), at⁷ jen⁴ kjā:u³-kjai³, ko¹, nuəi⁵, ŋwɛ² nji³, mja²², tset², zɛi² (姉による呼称),
ʔe:ŋ³

データ用文献一覧/Bibliography

- 平野綾香 (2021) 『ベトナムランソン省チャンディン県ヌン語の記述研究』大阪大学, 博士論文.
- 薄文泽 (1997) 《佯僜语研究》上海: 上海远东出版社.
- 薄文泽 (2003) 《木佬语研究》北京: 中央民族大学出版社.
- 广西壮族自治区少数民族语言文字工作委员会 (2008) 《广西民族语言方音词汇》北京: 民族出版社.
- 小坂隆一、周国炎、李錦芳編 (1998) 《仡央语言词汇集》贵阳: 贵州民族出版社.
- 李錦芳 (1999) 《布央语研究》北京: 中央民族大学出版社.
- 李錦芳、吴雅萍、姜莉芳、熊南京、李炜、徐晓丽 (2006) 《西南地区瀕危语言调查研究》北京: 中央民族大学出版社.
- 李如龙、侯小英、林天送、秦恺 (2012) 《茶洞语研究》北京: 民族出版社.
- 李云兵 (2000) 《拉基语研究》北京: 中央民族大学出版社.
- 梁敏編著 (1980) 《侗语简志》北京: 民族出版社.
- 梁敏、张均如 (1997) 《临高话研究》上海: 上海远东出版社.
- 梁敏、张均如 (2002) 《标话研究》北京: 中央民族大学出版社.
- 梁敏、张均如、李云兵 (2007) 《普标语研究》北京: 民族出版社.
- 毛宗武、蒙朝吉、郑宗泽編著 (1982) 《瑶族语言简志》北京: 民族出版社.
- 欧阳觉亚 (1998) 《村语研究》上海: 上海远东出版社.
- 欧阳觉亚、郑贻青 (1983) 《黎语调查研究》北京: 中国社会科学出版社.
- 王均、郑国乔編著 (1980) 《仡佬语简志》北京: 民族出版社.
- 杨通银 (2000) 《莫语研究》北京: 中央民族大学出版社.
- 喻翠容編著 (1980) 《布依语简志》北京: 民族出版社.
- 喻翠容、罗美珍編著 (1980) 《傣语简志》北京: 民族出版社.
- 张均如編著 (1980) 《水语简志》北京: 民族出版社.
- 张均如、梁敏、欧阳觉亚、郑贻青、李旭练、谢建猷 (1999) 《壮语方言研究》成都: 四川民族出版社.
- 周耀文、罗美珍 (2001) 《傣语方言研究》北京: 民族出版社.
- Chen, Yen-ling (2018) *Proto-Ong-Be*, Ph.D. dissertation, University of Hawai'i at Manoa. URI: <http://hdl.handle.net/10125/62511>
- Hudak, Thomas John (1994) *William J. Gedney's southwestern Tai dialects*. Ann Arbor: Center for South and Southeast Asian Studies, The University of Michigan.

- Hudak, Thomas John (1995) *William J. Gedney's central Tai dialects: glossaries, texts, and translations*. Ann Arbor: Center for South and Southeast Asian Studies, The University of Michigan.
- Hudak, Thomas John (1997) *William J. Gedney's Tai dialect studies*. Ann Arbor: Center for South and Southeast Asian Studies, The University of Michigan.
- Hudak, Thomas John (2008) *William J. Gedney's comparative Tai source book*. Honolulu: University of Hawai'i Press. URI: <https://www.jstor.org/stable/20532978>
- Morey, Stephen (2005) *The Tai languages of Assam: A grammar and texts*. Canberra: Pacific Linguistics, The Australian National University. doi: <https://doi.org/10.15144/PL-565>
- Pittayaporn, Pittayawat (2009) *Phonology of Proto-Tai*. Ph.D.dissertation, Cornell University, Ithaca, New York.
- Самарина, Ирина Владимировна (2011) *Языки гэлао: материалы к сопоставительному словарю кдайских языков*. Москва: Academia.

参考文献/References

- Burusphat, Somsong and Xiaohang Qin (eds.) (2006) *Northern Zhuang Chinese-Thai-English dictionary*. Bangkok: Institute of Language and Culture for Rural Development, Mahidol University.
- Burusphat, Somsong, Xuecun Wei and Jerold A. Edmondson (eds.) (2003) *Sui (Shui) Chinese-Thai-English dictionary*. Bangkok: Institute of Language and Culture for Rural Development, Mahidol University.
- Castro, Andy and Xingwen Pan (eds.) (2015) *Sui dialect research*. SIL e-Books 66. SIL International. URI: <https://www.sil.org/resources/archives/64425> (Originally published in Chinese as 《水语调查研究》贵阳: 贵州人民出版社, 2014.)
- Headley, Robert K. (1997) *Modern Cambodian-English dictionary*. Maryland: Dunwoody Press.
- Hudak, Thomas John (2008) *William J. Gedney's comparative Tai source book*. Honolulu: University of Hawai'i Press. URI: <https://www.jstor.org/stable/20532978>
- Li, Fang-Kuei (1977) *A handbook of comparative Tai*. Oceanic linguistics special publication no.15. Honolulu: University of Hawai'i Press. URI: <https://www.jstor.org/stable/20006684>
- Norquest, Peter K. (2007) *A phonological reconstruction of Proto-Hlai*. Ph.D. dissertation. Tucson: Department of Anthropology, University of Arizona. URI: <http://hdl.handle.net/10150/194203>
- Ostapirat, Weera (2000) Proto-Kra. *Linguistics of the Tibeto-Burman Area* 23(1): 1-251. doi: <http://doi.org/10.15144/LTBA-23.1.1>
- SEAlang Mon-Khmer etymological dictionary. URL: <http://sealang.net/monkhmer/dictionary/>

出版情報

採用決定日：2022年9月16日

彙報 2021年度

Memoranda and personalia 2021

第3回大会は岡山大学で開催準備していたが、引き続きコロナのためオンライン開催となった。7件応募があり、すべて採択、うち2件についてはやむを得ない事情により辞退となり、最終的には一般発表は5件、そして国立遺伝学研究所の鈴木留美子先生に「日本人の名字の地域性と遺伝学」と題する招待講演をしていただいた。土曜の半日開催ではあったが、充実した内容であった。

学会誌『地理言語学研究』第1号を2021年9月27日にpdf版公開・翌月に紙版もAmazonより販売開始した。また2022年2月にMonograph No.1としてHiroyuki Suzuki, *Geolinguistics in the eastern Tibetosphere: An introduction*; 2022年3月31日にはNo.2として《岩田礼教授荣休纪念论文集》上下冊も同様にしてpdf版・紙版を公開した。pdf版のほうがデフォルトで、短い間に多くの方々がダウンロードし、学会側・研究者側ともメリットが大きい方法であると実感された。

大会での研究発表・学会誌・モノグラフのいずれも募集要項に記されている趣旨と様式にかなっていれば、無料ないし非常に低廉なコストでどなたでも参画できるので、ぜひ積極的に活用いただくと幸いである。

2021年度の会計報告は以下の通り（会計監査・世話人会確認済）：

収入の部	61,454 円	2020 年度繰越金
支出の部	5,568 円	学会 HP サーバー使用料
	20,000 円	大会招待講演謝金
	5,478 円	学会誌紙版公開手数料（MyISBN/デザインエッグ）
残高	30,408 円	