ISSN 2436-6471

地 理 言 語 学 研 究

Studies in Geolinguistics

4

日本地理言語学会
Geolinguistic Society of Japan
2024

ISSN 2436-6471

地 理 言 語 学 研 究

第4号

Studies in Geolinguistics

No. 4

2024

日本地理言語学会 発行 | Published by Geolinguistic Society of Japan

2024年10月18日発行
doi: https://doi.org/10.5281/zenodo.13948834

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

特集「アジア・アフリカにおける言語構造の地理言語学的アプローチ」

日本地理言語学会　発行 | Published by Geolinguistic Society of Japan

翻訳/Translation

Contents

Translation

研究論文

中世韓国語の母音jeの諸方言における反映と母音体系の変遷について

福井　玲
東京大学名誉教授

Reflexes of Middle Korean vowel *je* in modern dialects and the development of vowel system in Korean

FUKUI, Rei
Professor Emeritus, University of Tokyo

Abstract: The purpose of this paper is to examine reflexes of Middle Korean vowel *je* in modern dialects and to discuss the development of the vowel system in this language. In Middle Korean as well as in Modern Korean, the vowel *je* is peculiar in many ways. First, the diphthong *je* is different from other diphthongs such as *ja, jo* and *ju* in that the latter ones are confined to such word classes as words consisting of Sino-Korean morphemes, onomatopoeia and so on, while the former is not confined to such word classes, and the number of words containing it is much larger than the latter ones. This imbalance shows that the vowel *je* has a kind of special status that can be comparable to ordinally simple vowels such as *a, e, o, u* and *i*. Also, this vowel is peculiar in historical sense. Although this vowel is generally considered to be a central vowel in Middle Korean, there are evidence that it was a front vowel in an earlier stage of this language. This paper tries to reinterpret geolinguistic maps for words containing this vowel and to show that this was indeed the case, thereby proposing possible vowel changes that could have occurred in the vowel system of this language.*

キーワード：韓国語；中世韓国語；母音体系；二重母音；言語地図
Keywords: Korean; Middle Korean; vowel system; diphthong; geolinguistic map

1. はじめに

韓国語の母音je (ㅖ)は，共時的にも歴史的にもさまざまな点で特異である。この母音は半母音j＋母音eという二重母音であるが，jで始まる他の二重母音と

福井玲(2024)「中世韓国語の母音jeの諸方言における反映と母音体系の変遷について」『地理言語学研究』4: 1–19. doi: https://doi.org/10.5281/zenodo.13948538
* 本稿は2024年6月1日に大阪大学で開かれた「地理言語学会」第６回大会で行った講演の原稿を修正・補完したものである。この大会に招待してくださった清水政明先生と講演に際してコメントを下さった先生方に感謝いたします。

はいろいろな点で異なり，あたかも単独の母音であるかのごとく振る舞う。この母音に関する問題点を整理すると次のようになる。

(1) 半母音jで始まる他の母音，すなわちja (ㅑ), jo (ㅛ), ju (ㅠ)の場合，これを第１音節に含む語は，漢字形態素，擬声語擬態語などを除く通常の語彙には非常に少ない。それに対して，第１音節にje (ㅖ)を含む語はごく普通に数多く見られる。
(2) 現代の多くの方言において母音je (ㅖ)はej [e] (ㅔ)と交替する現象が見られる。標準語としてはje (ㅖ)と発音されるべきものが，実際の口語ではej (ㅔ)と発音されることが多い。韓国における標準語であるソウル方言でもこの現象はよく見られる。
(3) 中世語（15～16世紀）の用言語幹においては，母音je (ㅖ)の出現頻度は他の母音に比べて顕著に多い（伊藤智ゆき2019: 124）。
(4) 中世において日本人が朝鮮語を聞き取って仮名で表記する場合に，この母音には「え段」の仮名を用いることが非常に多い（志部昭平1988: 73，辻星児2000）。

本稿では, 小倉進平の資料をもとにした言語地図（福井玲編2017, 2018, 2022）の中から第1音節に母音jeを含む語を選んで，その発音の方言上の分布を改めて再解釈し, 韓国語の母音体系の歴史的な変遷を考えるうえで何らかの示唆が得られるのではないかという立場から吟味していく。

一例をあげれば，「嫁」を意味する語は標準語では [mjɔniɾi] と発音されるが，言語地図上では，そう発音される地域は狭く，京畿道を中心としてその南北の黄海道と忠清道などに見られるのみである。他の多くの地点は [meniɾi] あるいはそのバリエーションで発音されている。そして，ここに見られる母音 [e] が過去の発音を何らかの形で反映しているのではないかと見ることも可能化もしれない。本稿では，こうした例を，語頭子音のあるなしなどいくつかの類型に分けて提示し, さらに韓国語における母音体系の変化の問題ともからめて，歴史的な観点から考察していく。なお，その際，現代語および中世語の母音体系が問題となるので，下の図１に現代語ソウル方言の母音体系，図２に中世語の母音体系を提示しておく。

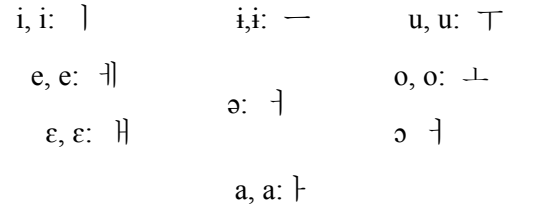


図1：現代韓国語（ソウル方言）の古い世代の母音体系（梅田博之(1994)に基づく）

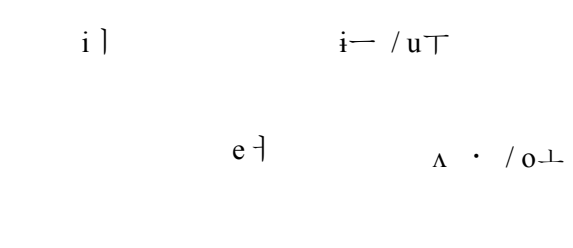


図2：中世韓国語の母音体系（福井玲(2013)に基づく）

以下，本論文ではまず第２節で半母音jで始まる母音の頻度の偏りを論じ，次いで，第３節で現代語の方言における母音je(ㅖ)とej(ㅐ)の交替現象を紹介，第４節でこの母音が過去の文献でどのように記録されているかに基づく史的考察を行う。続いて，第５節では言語地図上に反映された現代語における母音je (ㅖ)の現れを観察し，最後の第６節で，これらすべてを総合して母音je (ㅖ)がどのような変遷をたどったのかを推定する。

2. 第一音節での半母音jで始まる母音の種類の偏りについて

半母音jで始まる二重母音ja (ㅑ), je (ㅖ), jo (ㅛ), ju (ㅠ)が語の第一音節に現れる場合，漢字形態素，擬声語擬態語などを除くと，je (ㅖ)以外のものは，通常の語彙では非常に少ない。それに対してje (ㅖ)を含む語はごく普通に数多く見られる。以下にそれを例示する。扱う語例は，中世語あるいは現代語どちらを選ぶこともできるが，ここでは現代語を対象とすることにする。その理由は，中世語はそれなりに資料の量が多いとはいえ，やはり限りがあって，現代語にしか語例が見られないものもあるあるからである。ここでは便宜上小学館『朝鮮語辞典』において最重要語(1650語)と基本語(4400語)の表示のあるものを中心に考察することにする（以下ではこの２つを合わせて基礎語彙と呼ぶ）。これは，微細なニュアンスの違いを表すような語の変種をすべて扱うと重要語が見えにくくなるからである。当然のことながら，漢字語，外来語，擬声語・擬態語は除き，複合語も原則として除く。ただし，判断に迷うものについては「備考」に追加説明を記した。また，中世語には存在して現代語には存在しない重要な語例についても「備考」欄に記載する。

2.1. 第一音節にja (ㅑ)を含むもの

まず，第一音節に母音ja (ㅑ)を含む語例を，語頭子音の有無で分類してあげる。

語頭子音なし：jaritha- 야릇하- (風変りだ), jak- 약- (賢い), jarmip- 얄밉- (小憎らしい), jarp- 얇-(薄い), jamcenha얌전하- (まじめだ), jath- 얕- (浅い)
語頭子音あり：kjarikha- 갸륵하- (殊勝だ), kjarimha- 갸름하- (長めだ),……

（備考：基礎語彙ではないが，他に，kjauttuŋ 갸우뚱(首をかしげるようす), kjatha- 갸웃하-(傾いている)などもあるが，これらは擬態語的である。そのほか，뺨 (頬)は現代語ではこの母音を含むが，中世語ではspamであって，母音ja (ㅑ)は含まれない。）

2.2. 第一音節にjo (ㅛ)を含むもの

次に，第一音節に母音jo (ㅛ)を含む語例をあげる。

語頭子音なし：なし
語頭子音あり：なし

（備考：jo요 (敷布団)は基礎語彙だが，漢字語「褥」に由来するので省く。また，指示詞jo요 (この)は，同じ意味の指示詞i이の変異形であり，これを含む語はすべて除外する。この他，뾰족하다(尖った)という形容詞は擬態語的である。また，용하다(偉い)に含まれる용は漢字語扱いしない辞書が多いが，漢字語と考える。なお，中世語にはtjoh- (良い), hjok- (小さい)という重要な形容詞が存在する。前者は現代語では音変化によってこの項目に該当しなくなっているが，少ない母音jo (ㅛ)の用例の中では重要なものである。後者は現代語では使われなくなったものである。）

2.3. 第一音節にju (ㅠ)を含むもの

次に，第一音節に母音ju (ㅠ)を含む語例をあげる。

語頭子音なし：junanha- 유난하다 (普通と異なる), juc^{h}윷 (樗蒲(かりうち), juc^{h}nori윷놀이で使われる棒)
語頭子音あり：なし

（備考 junanha- 유난하- のju유は漢字語「類」の可能性がある。なお，現代語には存在しないが，中世語～近代語にはjumu 유무 (書信)という語が存在した。しかしこれも漢字語「有無」に由来する。）

2.4. 第一音節にje (ㅖ)を含むもの

最後に，第一音節に母音je (ㅖ)を含む語例をあげる。これは以下の通り基礎的な語彙を多く含み，多数存在する

語頭子音なし：jeki 여기(ここ), jeki- 여기- (感じる), jeni 여느 (普通の), jeterp 여덟 (八), jetin 여든 (八十), jere 여러 (いろいろな), jerim 여름 (夏), jeses 여섯 (六), jeu 여우 (狐), jewi- 여위- (やつれる), jei- 여의- (死に別れる), jeccu- 여쭈- (覗う), jekk- 엮- (編む), jer 열 (十), jer- 열- (開ける), jerp- 엷- (薄い), jes 엿 (飴), jeŋmun영문 (わけ), jep^{h} 옆 (横)

語頭子音あり：kje 겨 (糠), kjenu- 겨누- (狙う), kjetiraŋi 겨드랑이(脇), kjerej 겨레(はらから), kjeru 겨루- (競う), kjeu 겨우(やっと), kjeur 겨울(冬), kjekk- 겪- (経る), kjenti- 견디- (耐える), kjencu- 견주- (比べる), kjet- 겯- (編む), kjer 결(木目), kjep- 겹- (手に負えない), kjepc^{h}i- 겹치- (重なる), kjet^{h} 곁(横), mjeniri 며느리 (嫁), mjec^{h} 몇 (幾), pje 벼 (稲), pjerak 벼락 (雷), pjeruk 벼룩 (蚤), pjesir 벼슬 (官職), pjer 별 (星), pjeŋari 병아리 (雛), 볕 pjet^{h} (日差し), ppje 뼈 (骨), k^{h}je- (火をつける), p^{h}je- 펴- (開く), p^{h}jerc^{h}i- 펼치- (広げる), hje 혀 (舌)

（備考 jeŋmun영문 (わけ) は，音形からすると漢字語起源のように見えるが，一応あげておく。）

なお，子音で始まる語のうち，t, s, c chで始まるものは中世語から近代語にかけて，tje > cje > ce; sje > se, cje > ce, c^{h}je > c^{h}e という変化が起きたため，上記のデータには含まれないことをお断りしておく。

さて，以上をまとめると次のようなことがいえる。まず，ja (ㅑ), jo (ㅛ), ju (ㅠ)で始まるものは非常に少ない。わずかに見られるものについても，形容詞が多いといった偏りが見られる。ja (ㅑ)については母音調和の上でje (ㅖ)と対になるものがいくつか含まれる。

これらに比べるとje (ㅖ)は語数が多いのみならず，品詞も多様であり，典型的な基礎語彙に含まれるものも多い。このことから，あたかもje (ㅖ)のみが，単母音a, e, u, o, iなどと同じように振る舞うように見える。

3. 母音je (ㅖ)とej (ㅞ)の交替

多くの方言において母音je (ㅖ)はej (ㅞ)と交替する現象が見られ，このことはソウル方言においても同様である。ソウル方言についての記述の一例として，梅田博之(1971)より「嫁」と「稲」の例をあげる。

159-1. 嫁　daughter-in-law
menɯri 며느리 (mjɔnɯriともいうが標準語的……（後略）

165-4. 稲　rice-plant
bjɔ 벼（老年層はbeということもある）……（後略）

この現象は，他方言では一層頻繁に見られるが，詳細は第５節の言語地図の実例において詳しく見ていく。

なお，ソウル方言には，母音je (ㅖ)がej (ㅔ)ではなくi (ㅣ)と交替する例も見られる。

101-1. 火をつける　light，kindle

kinda 켜다 (ki-, N를 V. 標準語ではkjɔ-だがふつうki-という. ……（後略）

4. 史的考察

母音je (ㅖ)は，現代のソウル方言では [jɔ] と発音されるが，中世語では中舌母音 [jə] であった。そして，さまざまな理由によって，母音je (ㅖ)は，さらには，半母音を伴わない母音e (ㅔ)そのものが，中世語以前には前舌母音だったのではないかとする推定も行われてきた。

例えば，モンゴル語からの借用語における音韻の対応を検討した李基文(1972)によれば，中世語では，母音e (ㅔ)が中舌母音であったと考えられるのに対し，12, 13世紀ごろの古代語では，母音e (ㅔ)は前舌母音であったとしている。母音e (ㅔ)が前舌であった時代には，je (ㅖ)も当然，前舌であったことになる。

また，河野六郎 (1979: 510)は朝鮮漢字音の研究から古代の母音体系について「b層では母音 어 ə は ä に対応する。とすればこの母音ももとは前舌的な母音ではなかったかと思われる」として，漢字音導入時においてこの母音が前舌母音であったと推定している。

さらに，je (ㅖ)に関する話に戻ると，中世語では一般的にはje (ㅖ)も，母音e (ㅔ)と同様に中舌であったと考えられているが，その時代には，前舌母音がi (ㅣ)しかないので，je (ㅖ)は異音として前舌的であってもおかしくないと考えることもできる。例えば，志部昭平 (1988: 73) ではje (ㅖ)について次のように述べている。

> 「この朝鮮語の母音E [ə] は，I音 [i, j] の次でより狭く，かつ前舌寄りとなり，母音音節では [je], 子音音節では単母音化して [e] に近く発音されたことになる。この現象が, 二重母音I+Eが母音音節では[je], 子音音節では[e], あるいは[i]になる，という慶尚道方言の特徴（小倉, 1944, pp.28～33, 金周弼, 1986）ともよく一致するのは興味深いことであるただ, D+I+Eの結合のみは，エ段ではなくチヨ(181, 258)と表記されていることから見て未だ[-jə]が保たれていたようである。」

また，別の観点からより古い時代のje (ㅖ)の発音について考察した研究もある(伊藤智ゆき2019)。それによれば中世語（15～16世紀）においては，用言語

幹に現れる母音je (ㅖ)の出現頻度は他の母音に比べて顕著に多い(伊藤智ゆき 2019: 124)。これに基づいて，伊藤智ゆき(2013, 2019)では，中世語における母音je (ㅖ)が朝鮮語祖語では，母音調和上iと対をなすïであったとし，さらにそれは音声的には[e]であったかもしれないとしている。

4.1. 日本語に転写された古代～中世の母音je (ㅖ)の特徴

中世において日本人が朝鮮語を聞き取って仮名で表記する場合に，この母音には「え段」の仮名を用いることが非常に多かった。まず，16 世紀末の文禄慶長の役（壬辰・丁酉倭乱）に際して，日本側が実用的な目的[1] のために簡単な朝鮮語の単語や文を記録した『陰徳記』所載の「高麗詞之事」においては，次のように多くの語がエ段の日本語で転写されている（志部昭平 1988: 73）。

> レンカム rjeŋkam (令監), セギ hjeŋ (兄), セイ hje (舌), キレイ kirjeki (雁), ゾクセン sʌjŋsjen (生魚), ベンボク p^{h}jenp^{h}uŋ (屏風), エ jej (此), チヨ tjeh (笛), チヨロ tjere (アソコ), エクイ jekii (其所), ヘギ bjeŋ (煩), セキ sjeŋ (腹立), エトロフ jetirp (八), エル jer (十), エス jes (六), メツチゴ mjec^{h}inko (イクツカ)　【ローマ字表記は本稿での方式に改める。】

また，これよりも古い時代の資料も存在する。辻星児(2000)によると，鎌倉時代初期成立の『二中歴』と，院政期後半に成立したと考えられる『世俗字類抄』に朝鮮語の数詞と思われるものがカナで表記されている。前者は新村出による紹介をはじめとして古くから知られていたが，後者は辻星児(2000)が初めて紹介したものである。この中で「六」「八」「十」の発音に相当すると思われるものは次のとおりである。ただし，実際には『二中歴』には錯誤があり，「三」と「四」，「五」と「六」が入れ替わり，さらに「八」にあたるものが「九」にあたるものとして書かれているが，それらを正したうえで示すと次のようになる（□の部分は不明の字）。

	六（jesis）	八(jetirp/jetʌrp)	十(jerh)
二中歴　高麗語	エス↘	エタリ	エツ□
二中歴　貴賀国語	エソ↘	エタリ	エ□
世俗字類抄（右）	エスス	エトロフ	セツカ
（左）	エスソ		エル

[1] これは，「高麗詞之事」の冒頭に「日本の諸将通事ナクテハ叶ハマシトテ高麗ノ詞ヲ習セラレケリ　其詞ヲ伝聞所ヲ聊記付ス」と注記されていることから，日本側が作戦行動上必要な朝鮮語を習ったものを記録したものであり，実用的な目的で記録したものと考えることができる。文の内容についても，日本側の軍事行動とかかわっていたことが分かる。

いずれも，je に該当する部分は「エ」と表記されていることが分かる。この「エ」は当時の日本語の発音に照らして言えば [je] に相当するものであったと考えられる。

5. 言語地図上での音je (ㅖ)の反映

ここでは，小倉進平の朝鮮語方言調査結果に基づく『小倉進平『朝鮮語方言の研究』所載資料による言語地図とその解釈』（第 1 集～第 3 集，福井玲編）の中から次の項目を選んで言語地図を提示する。（項目中のかっこ内は小倉進平による項目名。それぞれの項目には担当者による解説が書かれているので，そちらも参照されたい。）

	語形 (現代)	語形 (中世)	項目	所収（担当者）
(1)	mjeniri	mjeniri (LH-)	嫁（子の妻）	第 2 集（徐旼廷）
(2)	pje	pje (H)	稲	第 2 集（福井玲）
(3)	pje:r	pjer (R)	星	第 1 集（李美姫）
(4)	pjeru	pjero (LH)	硯	第 1 集（林茶英）
(5)	hje	hje (H)	舌	第 2 集（李美姫）
(6)	kjeur	kjezirh (LH)	冬	第 1 集（李美姫）
(7)	kjej:cip	kjecip (R-)	女	第 3 集（福井玲）

（中世語の語形のあとの括弧内に入れた L, H, R は声の高さを表す。L は低いピッチ，H は高いピッチ，R は長い上昇調で発音された）

なお，本稿末尾に掲載した地図は，この報告書作成時のデータに基づいて，今回の主題である母音 [jɔ] : [e]の対立を見やすくするように作り直したものである。記号の基本的な形と色使いは次のようになっている。

赤丸：[jɔ]　青三角：[e]　紺横長三角：[ɛ]　　緑縦長三角：[i]

次に，これらの項目をこの順に見ていく。

(1) mjeniri: mjeniri (LH-)「嫁」: 地図 1 参照。[mjɔnuri]ないし[mjɔniri]で現れるのは黄海道の西側と，忠清道の一部である。他の地域は第 1 音節の母音は[e], [ɛ], [i]で現れる。このうち，小倉進平が[ɛ]で表記しているのは，慶尚道で /e/ と /ɛ/ の対立がない地点である場合が多い。[i] は [e]が狭母音化を起こしたものである。よって，これらの地域では[e]で現れるのが基本であると見ることができる。

(2) pje: pje (H)「稲」: 地図 2 参照。この項目では，半島の南半分では narak という，pje とは語源的に異なる語形が使われている。それを除いて考えると，[pjɔ]と発音されるのは黄海道西側の数地点のみで，他の多くの地点では [pe] となっている。

(3) pje:r: pjer (R)「星」: 地図 3 参照。これは上の(2)とは若干異なる。現代語の [pjɔ:l] に相当すると思われる[2]発音かなり広範囲に見られる。また，[pel] のように母音が [e], [ɛ], [i] で現れるものもかなり広範囲に見られ，両者を併用する地点も各地に混在する。両者の違いとしては，黄海道と京畿道では前者のみが見られるという点が注目される。(2)との条件の違いとしては，この単語が中世語では上声を持ち，現代語ではそれに対応して長母音を持っている点があげられる。

(4) pjeru: pjero (LH)「硯」: 地図 4 参照。これは分布の傾向としては(2)と似ている。即ち第一音節が [pjɔ] となる地域は黄海道，京畿道とその周辺が中心で，他の多くの地点では母音が [e], [ɛ], [i] となっている。ただし，咸鏡北道の所謂六鎭方言では，数地点 [pjɔ] となる地域が存在する。

(5) hje: hje (H)「舌」: 地図 5 参照。これは上のいずれとも傾向が若干異なる。もともと /h/ の次に [i] か [j] が後続する場合に/h/は硬口蓋摩擦音 [ç] となり，それがさらに地域によっては [s] や [ʃ] になることを考慮に入れる必要がある。さらに，これらの子音のあとでは，半母音 [j] の有無が対立に関与しなくなることも併せて考えると，母音が [jɔ] であるか [e] であるかに関して，この項目の場合に問題になるのは主母音が [ɔ] であるか [e] であるかということになる。この点で地図を見ると，[ɔ] となるのは黄海道，京畿道周辺，忠清南道から全羅道の西側に帯状に分布し，それらを取り囲むようにして北側および東側には [e] の地域が広く分布していることが分かる。

(6) kjeur: kjezɨrh (LH)「冬」: 地図 6 参照。この項目はまた上のいずれとも異なる独自の特徴を持っている。小倉進平のデータでは，主母音は [ɔ] となる地域が多く，逆に [e] となる地点は見られない。しかし，ここで空白となっている平安北道については『平北方言辞典』(p. 33) によると全域で게울 [keul] となるという記録が見られる。

게울 (名) 〔全域〕겨울. 冬節.

[2] ただし，小倉進平のデータでは，[pjɔ:l] にあたると思われる部分が，なぜか [piɔl] と書かれている。ここでは，[pjɔ:l] と同じものとして扱う。

また，また，平安道と同じく空白となっている咸鏡北道については，『咸北方言辞典』(pp. 56-57)によると，kjeur (겨울)について次のような変種が記録されている。

게울，겨슬，겨울, 겨을，결，동삼，저울，저을

このうち，平安北道と同じく게울 [keul]と発音される地点は鶴城，鏡城，穏城の三地点である。

なお，こうした点に注意すると，これ以外の方言についても[e]が現れないのは小倉進平の調査漏れだった可能性もある。

また，もう１つこの項目が他の項目と異なる要因として，中世語の語形に含まれるｚの消失後は，２音節が母音連続になるという点をあげることができるかもしれない。

(7) kjejcip: kjecip (R-)「女」：地図７参照。これもまた，これまでのすべての項目とは異なる独自の特徴を持っている。それは中世語では [kjə-] だったものが，現代語におけるその反映と考えられる [kjɔ-] で現れるものが全く見られず，第一音節の母音はすべて [e], [ɛ], [i] になっているという点である。また，現代のソウル方言では見られない [kje] という音節が平安道，京城，咸鏡北道に見られる点も注目される。これに関しても中世語で第一音節に上声(R)を持っていることとかかわりがあるのかもしれない。

6. 考察

ここでは第５節で見た地図上での分布と，第４節で見た歴史的な事柄を総合的に考察する。まず，地図に関しては，問題の母音が [jɔ] で現れるか [e] で現れるかに関して顕著な地域差が見られることが大きな特徴である。ただし，実際には語毎に微妙な違いがあり，まさに “chaque mot a son histoire” という状態を呈している。しかしながら，あえて地域的な特徴をまとめると次のようになる。

[jɔ]が優勢：黄海道，京畿道，(項目によっては全羅道西部，咸鏡道六鎭方言)
[e]が優勢：平安道，咸鏡道，江原道，慶尚道，全羅道，済州島

ここで，注目すべきは子音の口蓋音化の傾向との関連である。通常，平安道方言は口蓋音化が最も起こりにくい方言であるのに対して，咸鏡道（の大部分），江原道，慶尚道，全羅道，済州島は口蓋音化が起こりやすい地域である。つま

り，子音の保守性に関しては正反対の地域が，この母音の現れに関しては共通の性質を持っていると言える点には着目しておきたい。

さて，ここで第４節で見た歴史的な資料，およびそれを受けた志部昭平の考察（第１節）を合わせて考えてみよう。

まず，[jɔ]で現れるか[e]に関して，「通説」では，次のような変化が想定される。

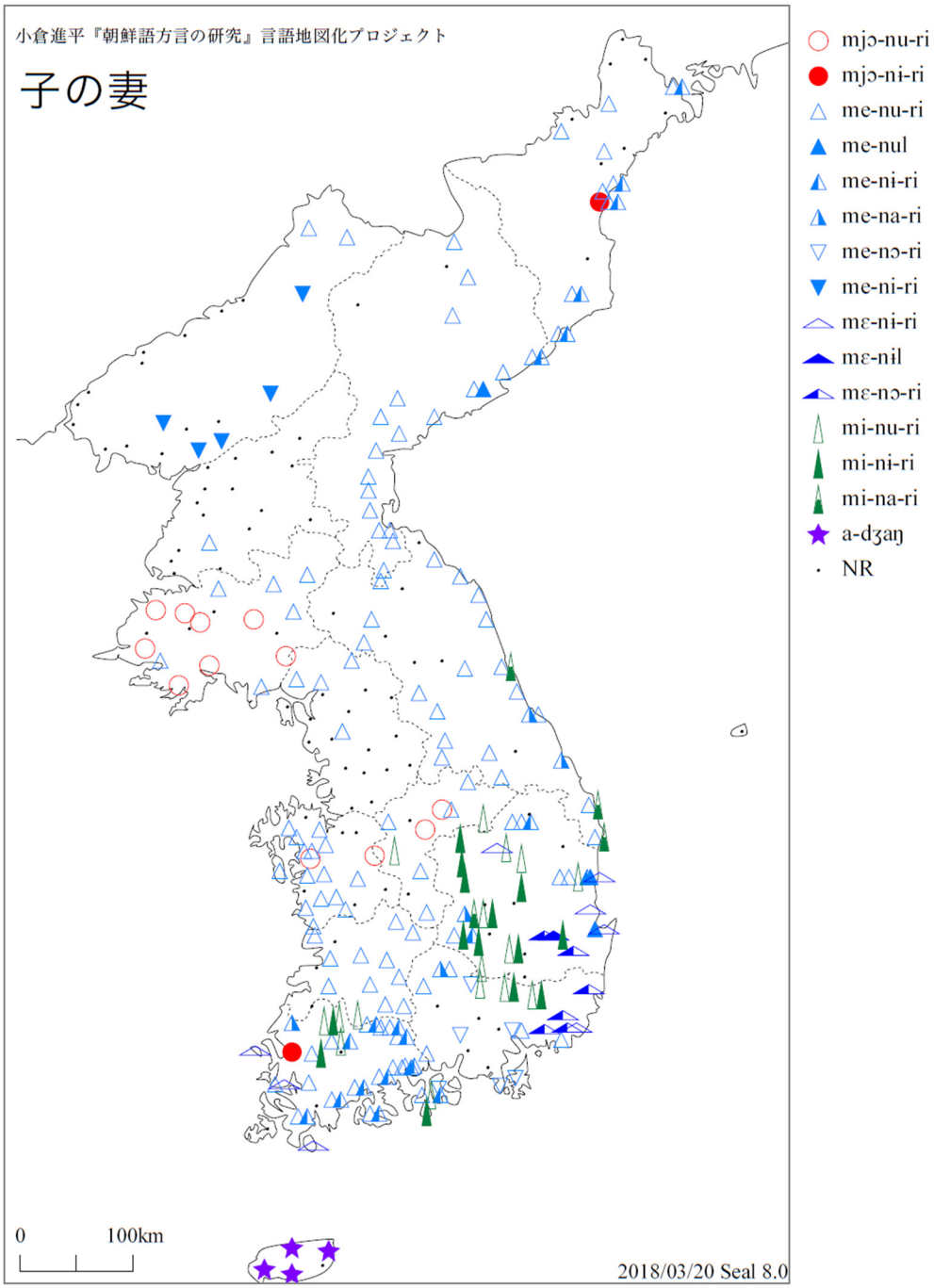
[jə] > [jɔ]　ソウルなど
[jə] > [e]　慶尚道，平安道など

しかし，これでは，多くの相互に地理的に離れた地点で [e] が現れることが説明しにくい。また，かつては je (ㅖ)が前舌母音であった可能性を考慮に入れると，次のような仮説も考えられる。

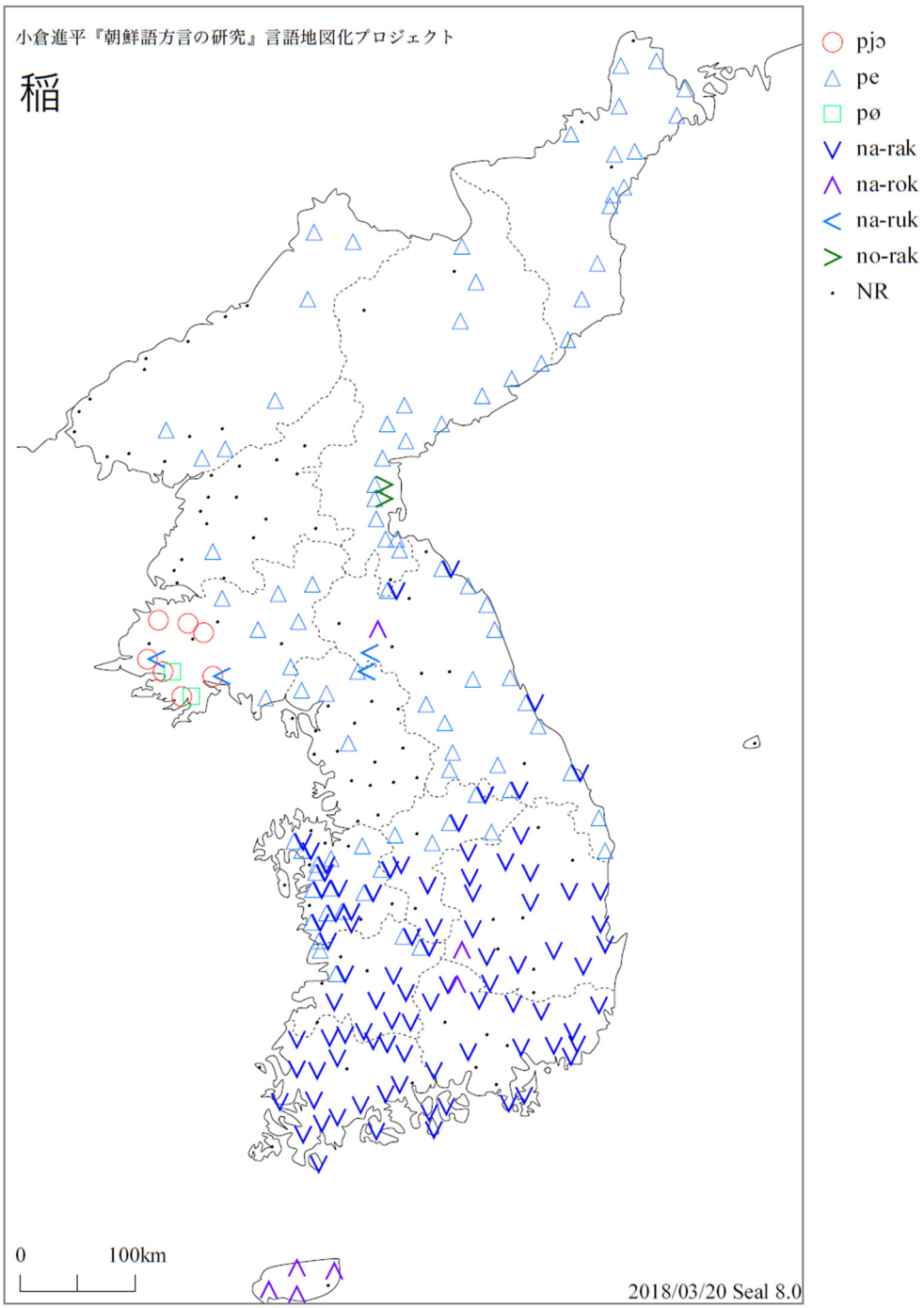
[je] > [jə] > [jɔ]　ソウルなど
[je] > [e]　慶尚道，平安道など

その際，中世語で二重母音 [əj] だったものが現代語では前舌母音[e]となる変化の時期との関りも考慮に入れる必要が出てくる。さらに，上で見たように，項目ごとに [ɔ] が出やすいか [e] が出やすいかが異なる点も問題点として残される。しかし，本稿ではひとまず，現代語で [e] で現れる地域の少なくとも一部は，より古い時代の [je] から来ているという可能性を提示してみようと思う。

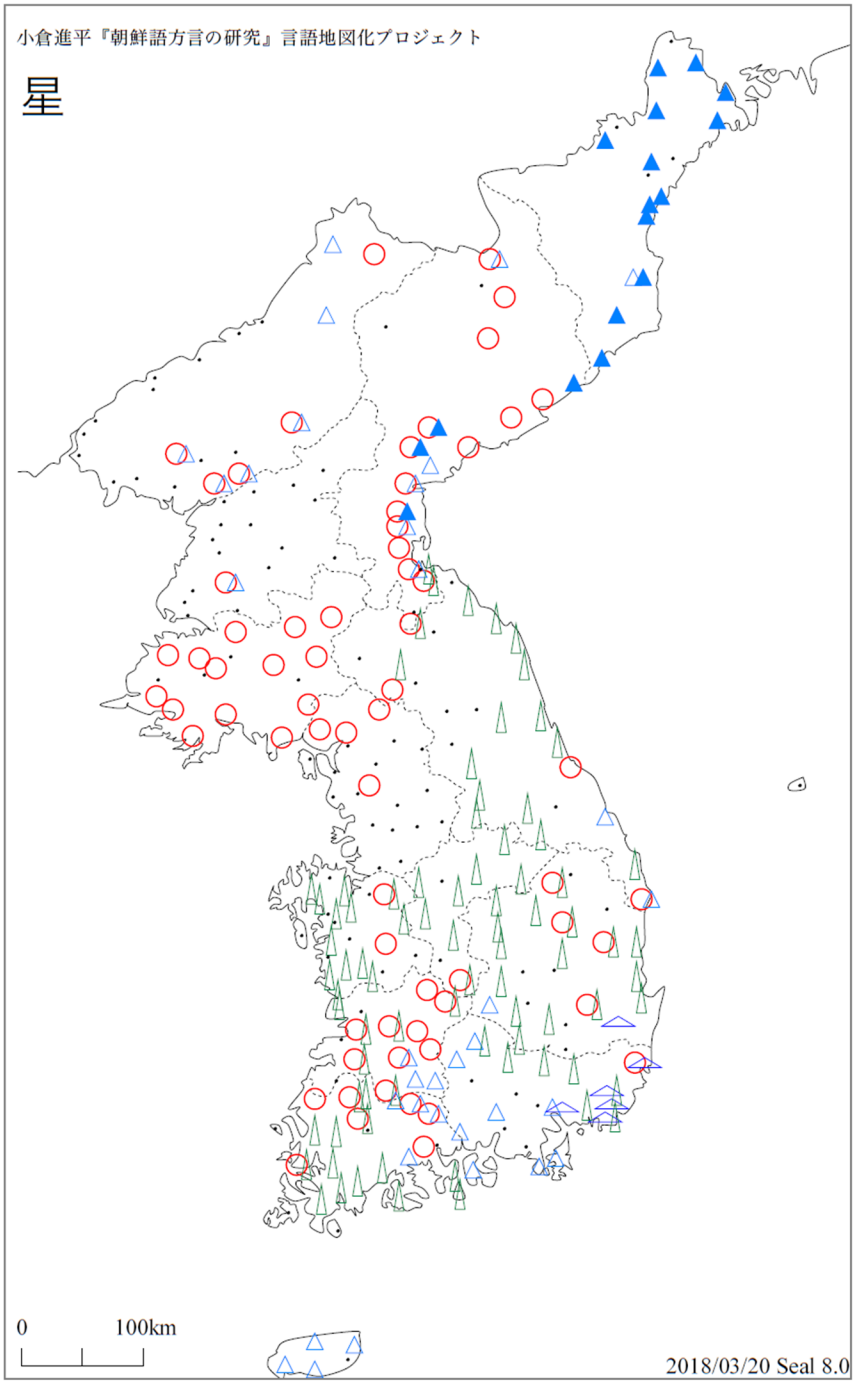
なお，仮にそうだっとすると，もう１つ別の問題点が生じる。母音ㅓとㅖがともに前舌母音であったならば，その２つの区別はどうなっていたかという問題である。上で見た河野六郎説では，古代語でㅓを [ä] としているわけであるが，それならば，ㅖは [je] としても，あるいは伊藤智ゆき(2019)のように [e] としても，区別が保たれる点では問題ない。


小倉進平『朝鮮語方言の研究』言語地図化プロジェクト
子の妻
mjɔ-nu-ri
mjɔ-ni-ri
me-nu-ri
me-nul
me-ni-ri
me-na-ri
me-nɔ-ri
me-ni-ri
mɛ-ni-ri
mɛ-nil
mɛ-nɔ-ri
mi-nu-ri
mi-ni-ri
mi-na-ri
a-dʒaŋ
NR
0
100km
2018/03/20 Seal 8.0

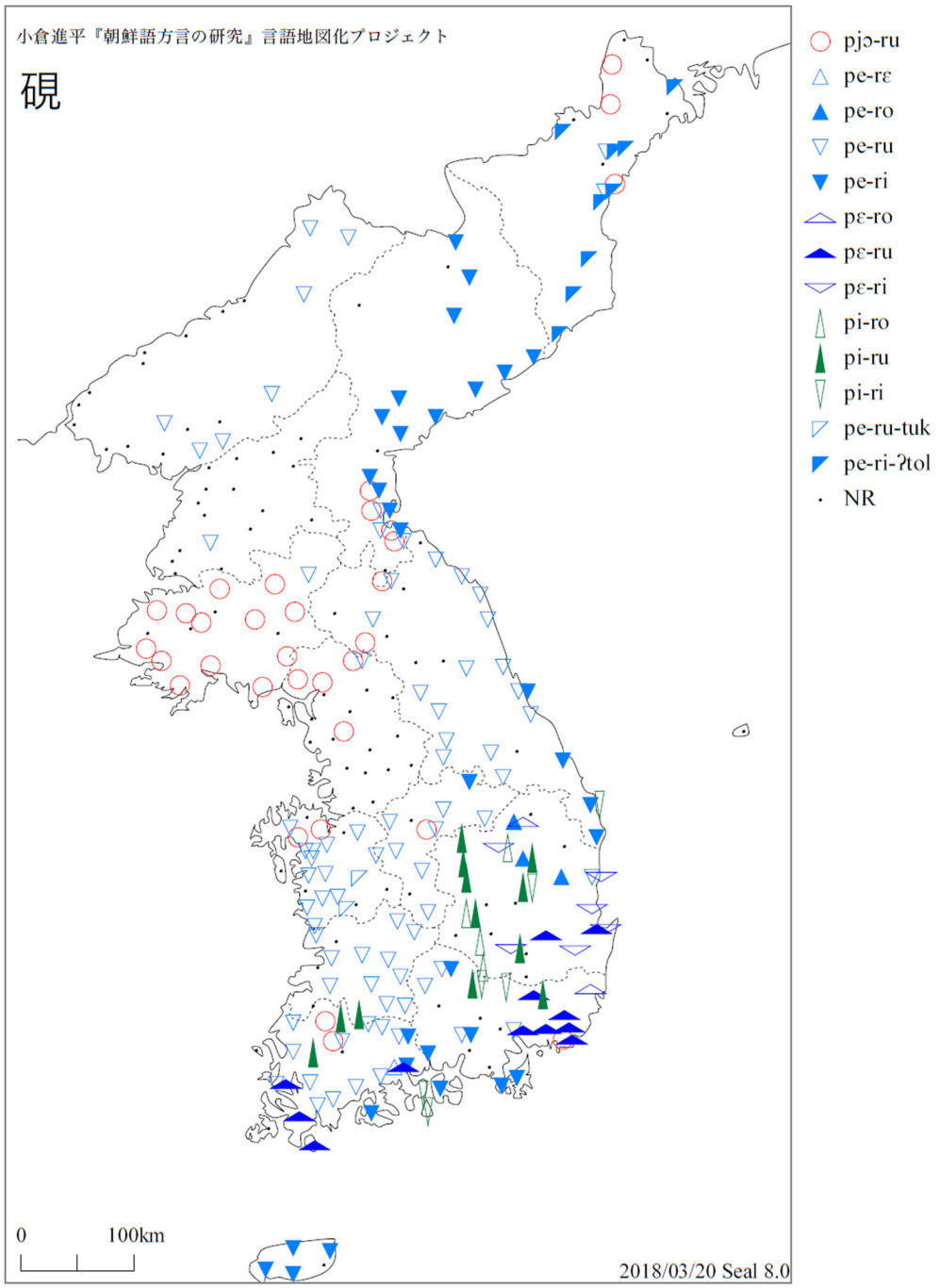
地図 1


小倉進平『朝鮮語方言の研究』言語地図化プロジェクト
稲
pjɔ
pe
pø
na-rak
na-rok
na-ruk
no-rak
NR
0
100km
2018/03/20 Seal 8.0

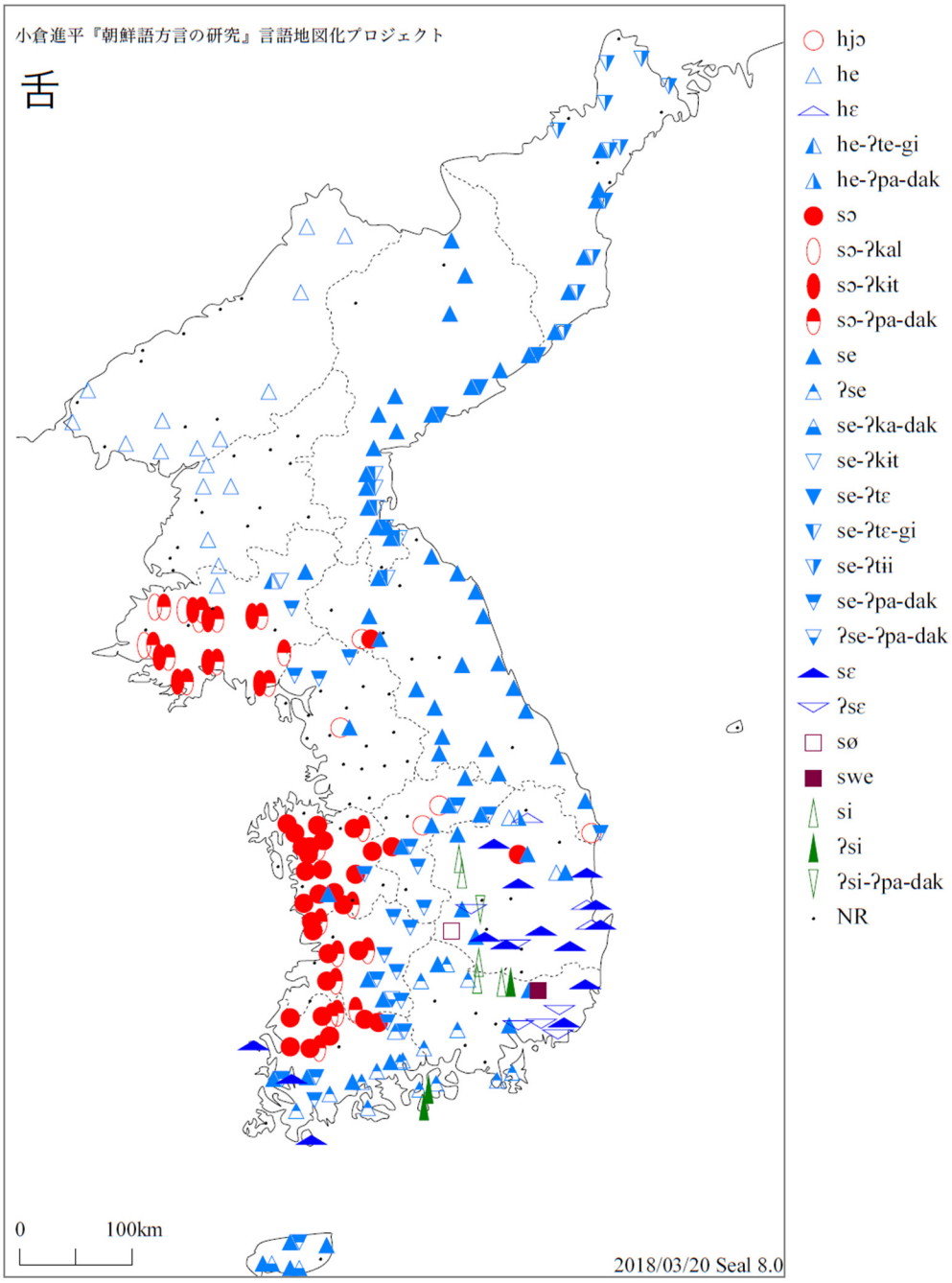
地図２


小倉進平『朝鮮語方言の研究』言語地図化プロジェクト
星
piɔl
pel
pe-ri
pɛ:l
pi:l
NR
0
100km
2018/03/20 Seal 8.0

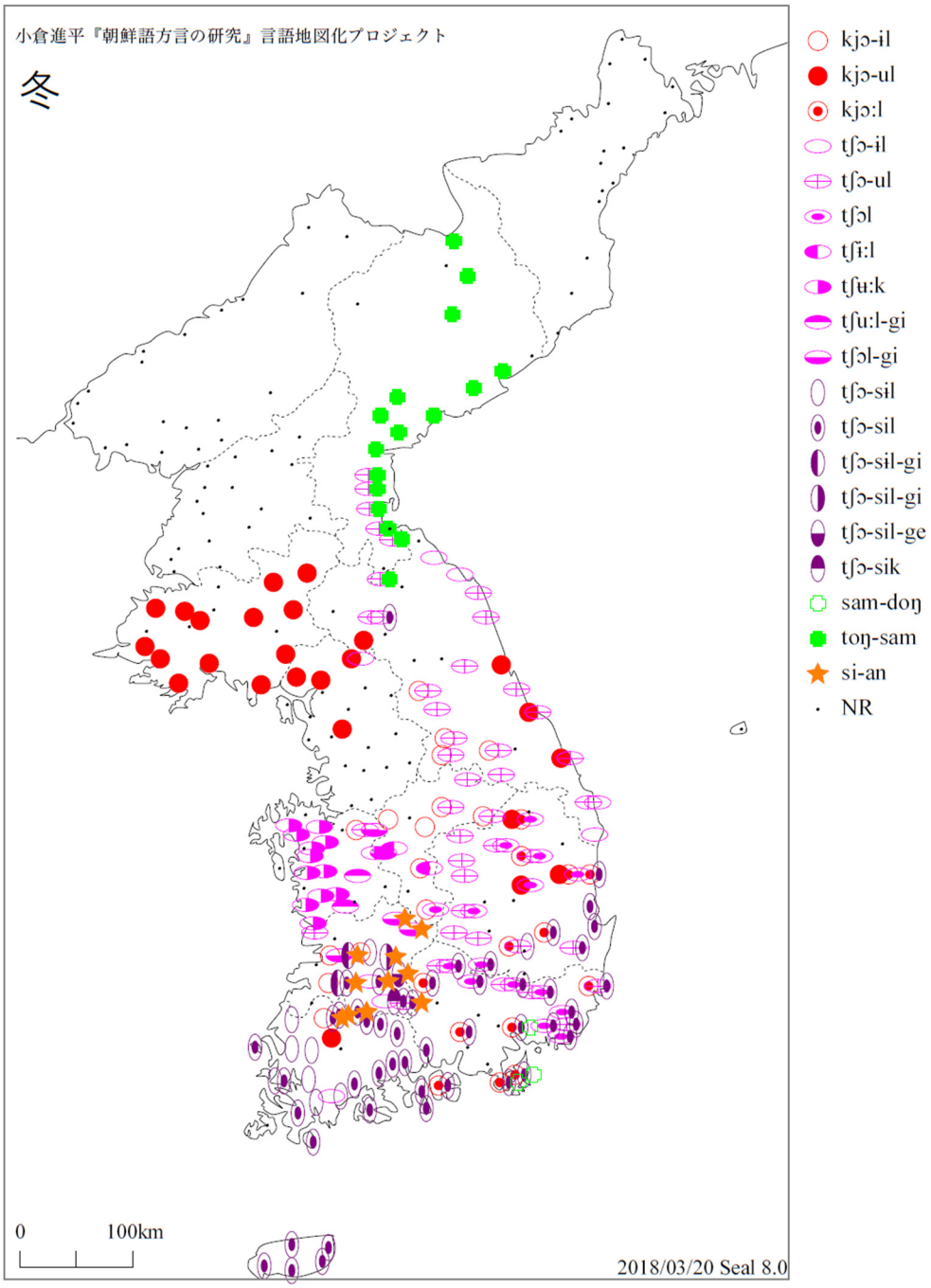
地図３


小倉進平『朝鮮語方言の研究』言語地図化プロジェクト
硯
pjɔ-ru
pe-rɛ
pe-ro
pe-ru
pe-ri
pɛ-ro
pɛ-ru
pɛ-ri
pi-ro
pi-ru
pi-ri
pe-ru-tuk
pe-ri-ʔtol
NR
0
100km
2018/03/20 Seal 8.0

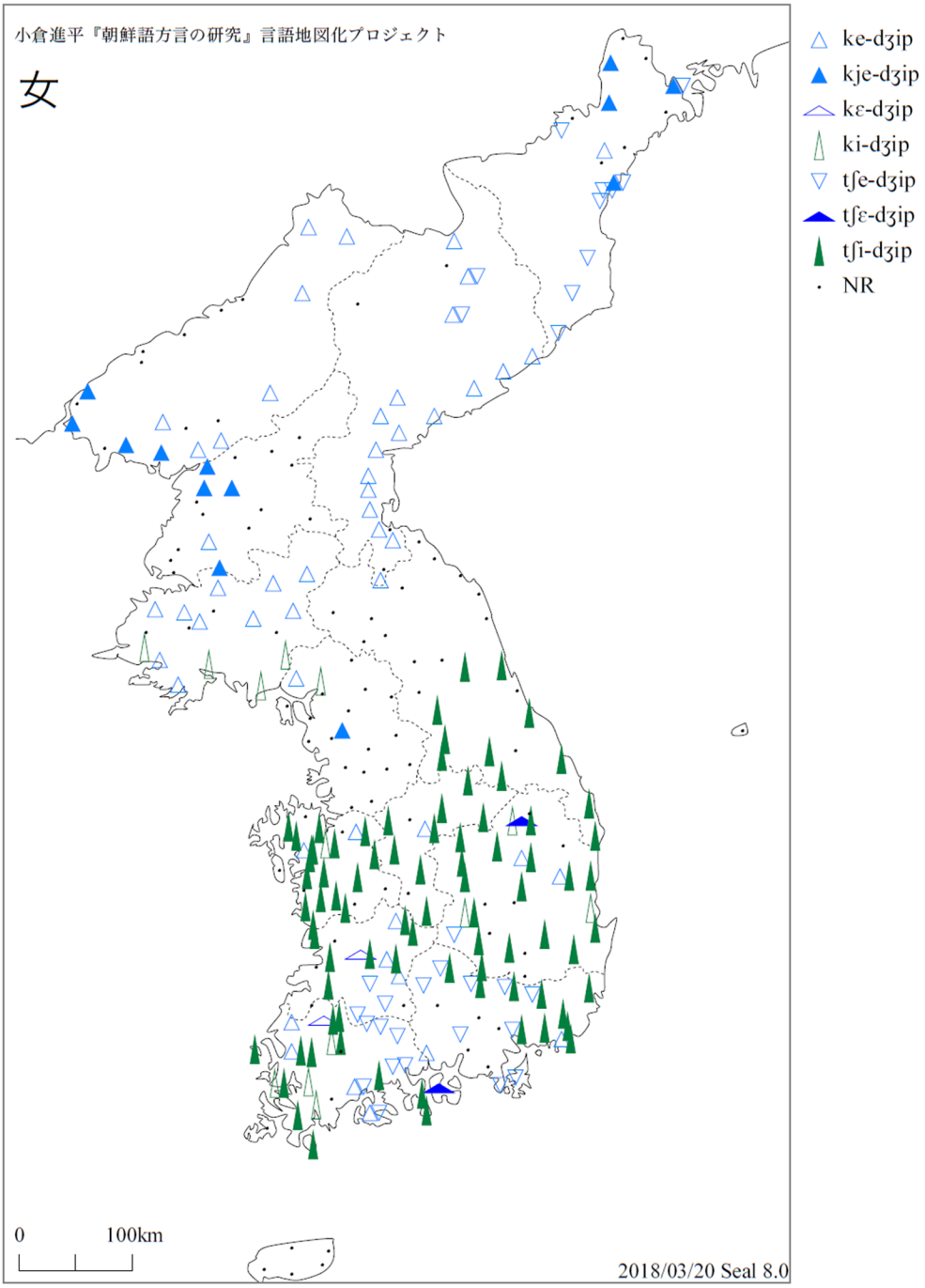
地図４


小倉進平『朝鮮語方言の研究』言語地図化プロジェクト
舌
hjɔ
he
hɛ
he-ʔte-gi
he-ʔpa-dak
sɔ
sɔ-ʔkal
sɔ-ʔkit
sɔ-ʔpa-dak
se
ʔse
se-ʔka-dak
se-ʔkit
se-ʔtɛ
se-ʔtɛ-gi
se-ʔtii
se-ʔpa-dak
ʔse-ʔpa-dak
sɛ
ʔsɛ
sø
swe
si
ʔsi
ʔsi-ʔpa-dak
NR
0
100km
2018/03/20 Seal 8.0

地図５


小倉進平『朝鮮語方言の研究』言語地図化プロジェクト
冬
kjɔ-il
kjɔ-ul
kjɔ:l
tʃɔ-il
tʃɔ-ul
tʃɔl
tʃi:l
tʃu:k
tʃu:l-gi
tʃɔl-gi
tʃɔ-sil
tʃɔ-sil
tʃɔ-sil-gi
tʃɔ-sil-gi
tʃɔ-sil-ge
tʃɔ-sik
sam-doŋ
toŋ-sam
si-an
NR
0
100km
2018/03/20 Seal 8.0

地図６


小倉進平『朝鮮語方言の研究』言語地図化プロジェクト
女
ke-dʒip
kje-dʒip
kɛ-dʒip
ki-dʒip
tʃe-dʒip
tʃɛ-dʒip
tʃi-dʒip
NR
0
100km
2018/03/20 Seal 8.0

地図 7

参考文献/References

梅田博之 (1971)『現代朝鮮語基礎語彙集』東京外国語大学アジア・アフリカ言語文化研究所.
梅田博之 (1994) 韓国語の母音. 『言語研究』106: 1-21. https://doi.org/10.11435/gengo1939.1994.106_1
小倉進平 (1944) 『朝鮮語方言の研究』東京：岩波書店.
河野六郎 (1964-67, 68/1979) 朝鮮漢字音の研究. 『朝鮮学報』31-33, 35, 41-44. 朝鮮学会. （のちに『朝鮮漢字音の研究』(1968)として天理時報社より刊行，河野六郎(1979)所収）
河野六郎 (1979)『河野六郎著作集 第2巻 中国音韻学論文集』東京：平凡社.
志部昭平 (1988) 陰徳記 高麗詞之事について：文禄慶長の役における仮名書き朝鮮語資料. 『朝鮮学報』128: 1-102.
辻星児 (2000)「二中歴」「世俗字類抄」所引の朝鮮語数詞について. 『岡山大学言語学論叢』8: 1-18.
福井玲 (2013) 『韓国語音韻史の探求』東京：三省堂.
福井玲編 (2017)『小倉進平『朝鮮語方言の研究』所載資料による言語地図とその解釈第１集』(PDF 版). 東京大学大学院人文社会系研究科 韓国朝鮮文化研究室.
福井玲編 (2018)『小倉進平『朝鮮語方言の研究』所載資料による言語地図とその解釈第2集』(PDF 版, 冊子版). 東京大学大学院人文社会系研究科 韓国朝鮮文化研究室.
福井玲編 (2022)『小倉進平『朝鮮語方言の研究』所載資料による言語地図とその解釈第3集』(PDF版, 冊子版). 東京大学大学院人文社会系研究科 韓国朝鮮文化研究室.
金履浹編著 (1981)『平北方言辞典』城南：韓国精神文化研究院.
金泰均編著 (1986)『咸北方言辞典』ソウル：京畿大学校出版局.
李基文 (1972) 『国語音韻史研究』ソウル：塔出版社.
Ito, Chiyuki [伊藤智ゆき] (2013) Korean accent: Internal reconstruction and historical development. *Korean Linguistics* 15(2):129-198.
Ito, Chiyuki [伊藤智ゆき] (2019) The accent and phonotactics of Middle Korean verbal stems. 『朝鮮学報』253: 115-169.

出版情報

投稿受理日：2024年9月3日
採用決定日：2024年9月13日

Research article

Linguistic motives common to Japanese and Romance dialects Two examples with maps

DEL GIUDICE, Philippe
Université Côte d'Azur, CNRS, BCL, France

BRUN-TRIGAUD, Guylaine
Université Côte d'Azur, CNRS, BCL, France

Abstract: Linguistic data treatment and mapping can vary greatly depending on the choices made by dialectologists. After recalling the fundamental principles of traditional geolinguistics, this article introduces the motivational point of view and the possibilities it presents in terms of mapping dialectal data. Concepts such as *désignant* and *motive* help focus on the ideas, interpretations and worldviews that are so important to lexical creation and change. By using them to classify the dialectal lexicon, we have found great similarities between the data of the *Linguistic Atlas of Japan* (LAJ) and the *Atlas Linguistique Roman* (ALiR). Those similarities transcend cultural differences. The two main examples that we map and analyze are 'praying mantis' and 'tadpole'. Both in Japan and in the Romance world, the former is generally named after its posture that looks like it is praying or after its bladelike limbs, when it is not seen as some kind of a (holy) horse. The latter gets its names from the fact that it is a baby frog, that it has a big head or that its shape is reminiscent of kitchen utensils. Our motivational maps show that such motives are massively shared by Japanese and Romance dialects as they cover a significant part of both areas. This rules out the possibility that they are due to chance. Etymologists already take advantage of this kind of discovery, that opens up new fields of research at the intersection of linguistics and cognitive sciences.*

Keywords: Linguistic motivation; Romance dialects; Japanese dialects; Praying mantis; Tadpole

1. Traditional dialectology and geolinguistics

1.1. A study of forms

The first major geolinguistic projects date back to the turn of the 20th century, with the monumental *Atlas Linguistique de la France* (ALF) in the West and the smaller *Phonetic Dialect Atlas* (PDA) / *Grammatical Dialect Atlas* (GDA) in Japan. Whether dealing with the lexicon, grammatical morphemes or purely phonetic aspects, the

DEL GIUDICE, Philippe and Guylaine BRUN-TRIGAUD. 2024. Linguistic motives common to Japanese and Romance dialects: Two examples with maps. *Studies in Geolinguistics* 4: 20–39. doi: https://doi.org/10.5281/zenodo.13948553

* We would like to thank our dear colleague Philippe Garat (Université Grenoble Alpes) for allowing us to use the latest version of the ShinyDialect tool, which he is still developing. We would also like to thank Professor Takuichiro Onishi for pointing us to the online resources that enabled us to consult and analyze the Japanese data in detail.

general principle of this kind of undertaking is to present linguistic forms (i.e. phonetic types, figurative pronunciations, signifiers) on a map.

This form-focused approach, referred to in these pages as "traditional", proved to be extremely productive. Its general mapping model, continually reproduced to this day, demonstrates the durability of this tried-and-tested method. More than a century later, the fourth and final volume of the *Atlas Linguistique et ethnographique de la Provence* (ALP), follows the exact same pattern as the ALF, with lexical maps showing the raw dialectal phonetic form of all the localities surveyed for a given meaning (Fig. 1).

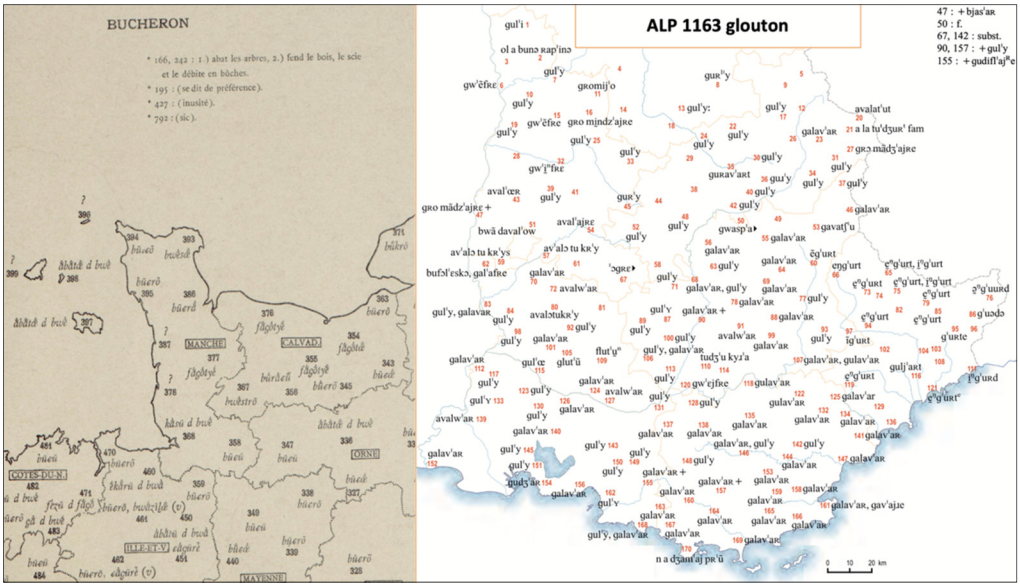


Fig. 1 ALF (1902-1910) (left) et ALP IV (2016) (right): an established model

1.2. A certain amount of cartographic diversity

This durability does not, however, rule out any freedom in the presentation of data. Alongside the raw data atlases characteristic of the old French tradition, we find atlases of interpreted data, in which data deemed to be of a similar nature are grouped together and represented by signs, colors, or a combination of the two. The Japanese tradition provides excellent examples of this. The geolinguist may also be tempted to represent homogeneous areas by drawing isoglosses (Fig. 2) or producing colored area maps.

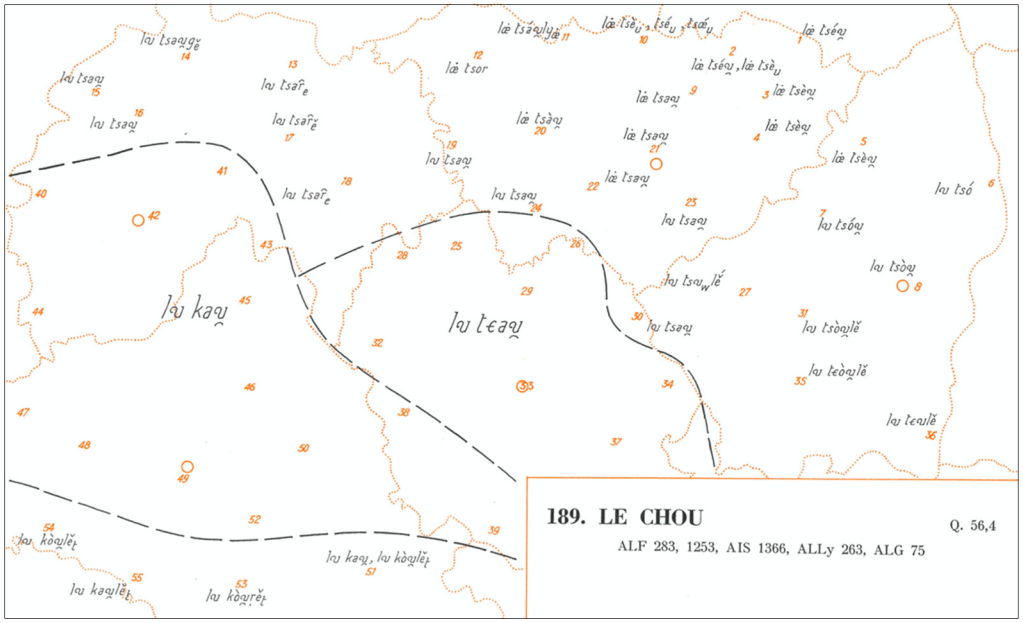


Fig. 2 The isoglosses in the *Atlas Linguistique et ethnographique du Massif Central* (ALMC)

1.3. Results of all kinds

There have been countless advances in geolinguistics, from elucidating the history and spread of individual items (from this point of view, some analyses have become famous, such as Gilliéron's study of rooster names in Gascony), to identifying the external factors – particularly territorial ones (Fig. 3) – that structure linguistic diffusion in a particular dialect area, or to developing terminology and proposing theoretical models of word circulation.

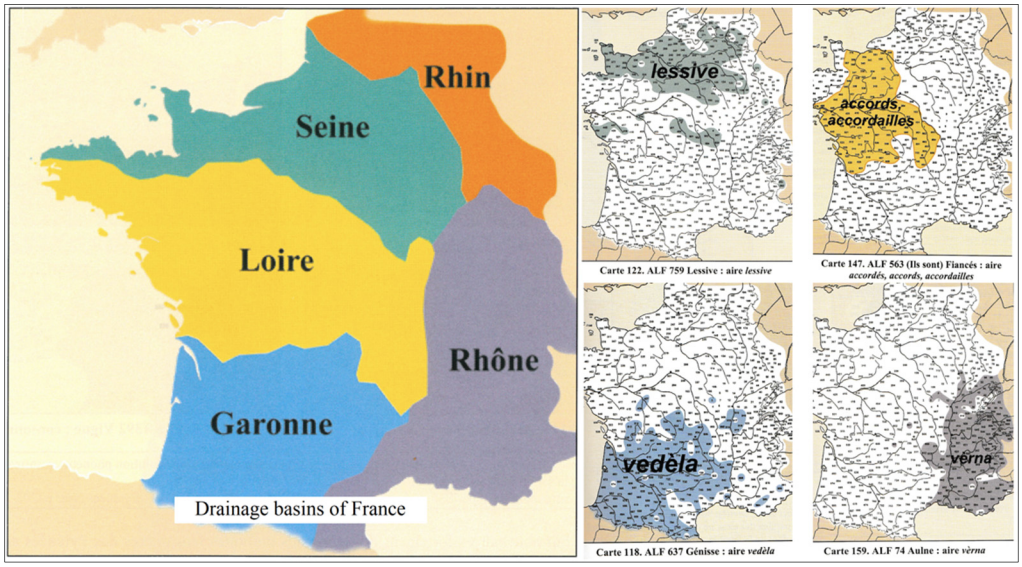


Fig. 3 Drainage basins of France and 4 examples of lexical distribution (Brun-Trigaud & Le Berre & Le Dû, 2005:134)

1.4. The arbitrariness of the sign as axiom

The basic theoretical foundation of this long tradition can be found in Saussure's arbitrary sign (2005 [1916]: 100–103). From arbitrariness flows the principle of the continuity of areas, as described by Dauzat (1922: 42):

> An area that today appears dislocated and fragmented once formed a continuous whole. [...] Since ideas have only factual relationships with expressions, the identity of a root or a form in two different languages to designate the same thing or the same function cannot be the effect of chance, but proves the kinship of the two types, i.e. their prior common existence.

This principle enables Dauzat to assert, for example, that the dialectal distributions of *pott-* 'lips' and its derivative *poton* 'kiss', now scattered across southwestern France, Savoy and the Vosges (Fig. 4), are the remnants of an ancient homogeneous area. These lexical types necessarily predate the Roman conquest, but they declined with the introduction of Latin.

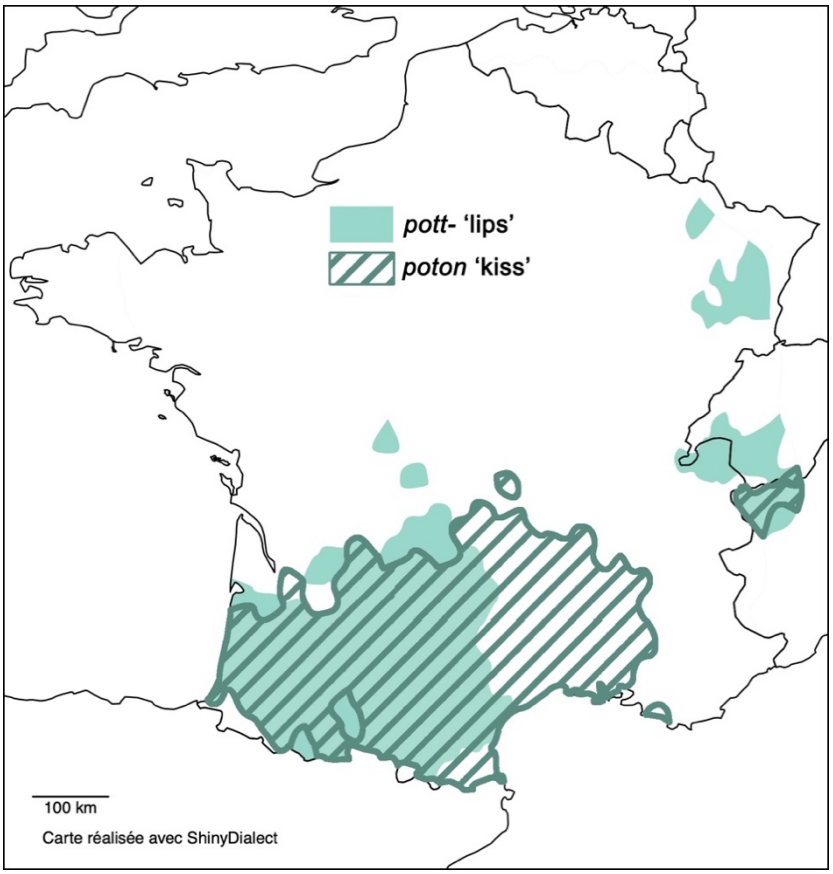


Fig. 4 The distribution of *pott-* 'lips' and *poton* 'kiss' according to the ALF

2. A complementary perspective: motivation

2.1. From arbitrariness to motivation

Once the advantages of the traditional approach were well known, and the methodological apparatus was firmly established within the community of dialectologists, recent decades have seen the development of a new point of view, which is intended to complement standard practices. With this approach, the researcher's gaze can shift from the form to the idea, or representation of the world that underlies that form. The focus then shifts away from the arbitrary nature of the sign to the sign's motivation. What are the causes that explain that a given word is the way it is? For what non-random reasons have such a signifier and such a signified been associated? This type of research constitutes the motivational approach[1]. On this basis,

[1] For a general presentation of motivation, its applications, its usefulness in etymological research, etc., see Del Giudice (2022).

etymologically and phonetically unrelated forms can be grouped together in the same category.

This is why, in the case of dialectal names for the robin in France, *roupie* (< *roux* 'ginger' + *pit* 'chest'), *pierrot* (*pit* + *roux*), *gavach roge* (literally 'ginger throat') and *pitre jaune* (literally 'yellow chest'), which all refer to the fact that the bird's plumage features a reddish medallion, need to be classified in the same motivational category. All these linguistic forms are generated by the same motive. An inventory of data larger than the handful of examples taken here would show that, for the robin, this colored medallion motive is overwhelmingly dominant throughout France. In contrast, the name *bovet* (literally 'little ox'), inspired by the animal's compact shape, is another type of designation that is not very widespread. In fact, the latter term is found very close to the heart of the French bullfighting tradition, where bovids are of particular cultural importance. Fig. 5 contrasts the treatment of the above five names according to a traditional map (left) and according to a motivational map (right), which highlights the originality of the second motive and the fact that it is found in a well-defined area.

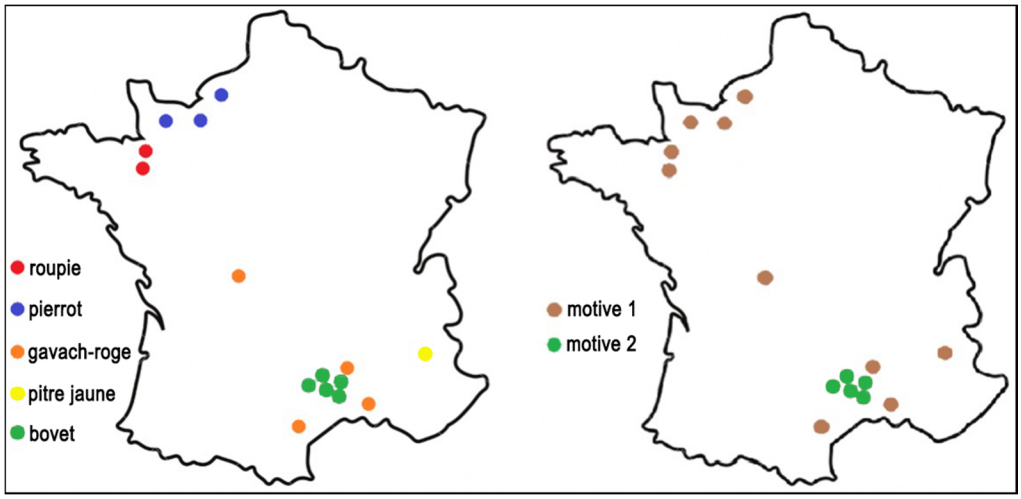


Fig. 5 Traditional and motivation treatments of an identical set of data

2.2. Examples of motivational studies

While the motivational point of view can sometimes call into question the achievements of the traditional approach (e.g. does the principle of the continuity of areas have the value of an absolute law in a motivated conception of language?), in practice it is a means of extending and exploiting differently the results obtained from classical methods, often without opposing them.

For example, having collected hundreds of dialectal designations of ‘mountain’ and ‘mound’ in the Gallo-Romance area in the ALF and ALFR[2], Del Giudice (2020) demonstrates that, beyond their formal differences, they are generated by 6 major fundamental representations – motives – which tend to be distributed differently depending on whether populations are located at altitude or at the foot of mountains, thus resolving the etymological problem of certain forms that had remained obscure until then. To achieve this, it was necessary to carry out a cartography of motives, complementary to the cartography of linguistic forms found in classic atlases.

In another vein, Dalbera (2001; 2006: 219–220) provides an interesting case of motivational research based on formal observations and phenomena familiar to the traditional approach. Based on the recurring motives that he identifies in the Romance-speaking world, Dalbera explains why the slow worm (*anguis fragilis*) may bear names derived from the Arabic *al-ʿaqrab*, originally meaning ‘scorpion’. In the Romance world, it is common for the slow worm to bear the name of the snake, but very unexpected that it should borrow that of the scorpion. The source of this name was therefore to be found elsewhere, and it all began with outcomes of lat. CURTIONE ‘snake’ of the *escurzón* type which, undergoing a paronymic attraction from *escorpión* ‘scorpion’, took on the form of the latter word, following a certain geographical progression, so that certain dialects (in particular in an intermediate zone between the territory where *escurzón* ‘slow worm’ is dominant and the area where *escorpión* ‘slow worm’ is) present mixed forms, *escurzión* being the best example. The massive presence of words evolved from *al-ʿaqrab* in areas in contact with *escorpión*, would then only be the result of a particular case of synonymic substitution[3], from the Romance word to the word of Arabic origin meaning ‘scorpion’. Fig. 6 shows a schematic representation of this change mechanism, based on ALEANR data and the “Orvet – 1” map of the *Atlas Linguistique Roman* (ALiR). Together with the *Atlas Linguarum europae* (ALE), this motivational atlas forms a reference set for large-scale motivational studies.

[2] *Atlas Linguistiques et ethnographiques de la France par Régions*, produced from the 1960s onwards. These are classic geolinguistic tools in the ALF tradition. The ALP, mentioned above, is part of this collection.
[3] See Iwata (2021).

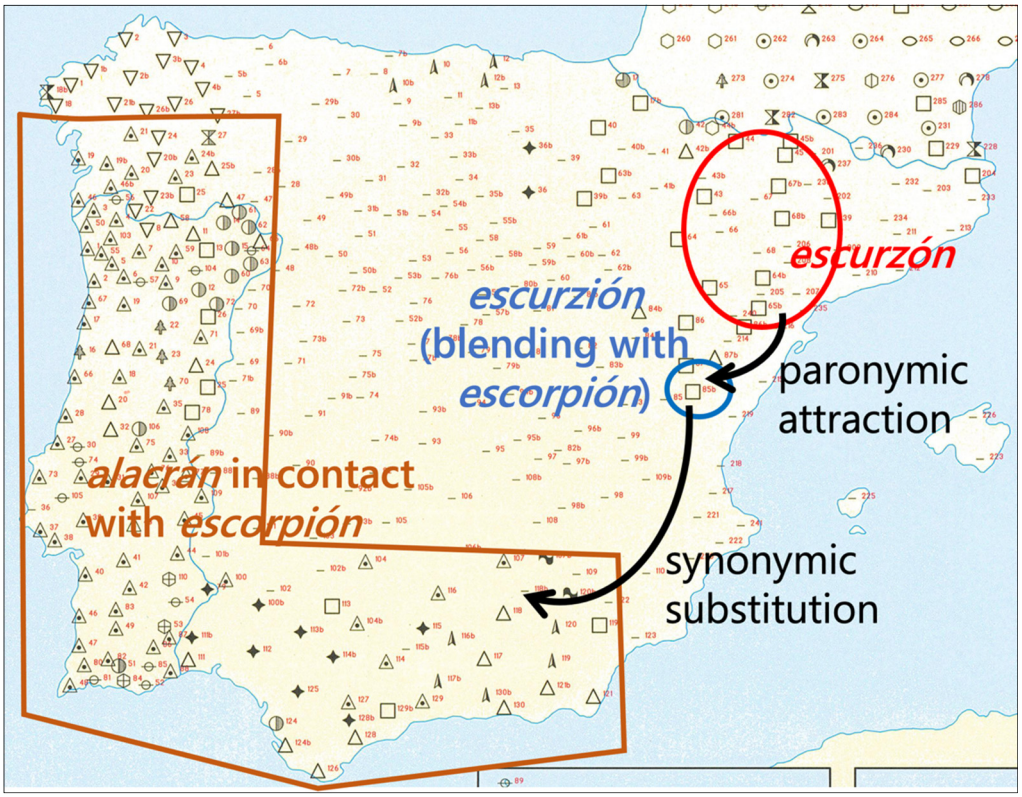


Fig. 6 Iberian names for 'slow worm', from *escurzón* to the words developed from *al-'aqrab*

3. Comparing Japanese dialects and the Romance world

3.1. A unique scale of analysis for dialect data

The studies cited so far might suggest that motivational analysis is confined to explaining phenomena at a more or less local level. However, one of the strengths of this approach dedicated to representations is that it touches on both cultural and cognitive aspects, and thus enables us to observe the potency of certain designation reflexes across borders and continents.

To assess this, we compared Japanese and Romance dialectal items. In other words, we brought together cultures and territories that are extremely far apart. We selected 2 examples ('praying mantis' and 'tadpole') from the onomasiological maps common to the *Linguistic Atlas of Japan* (LAJ) and the ALiR. Everything that follows will show that analyses of such geographical scope and depth in terms of the functioning of the naming act would have been difficult without taking motivational aspects into account. The examples given below therefore draw the reader's attention to the similarities

between Japanese and Romance naming. But one should not lose sight of the fact that there are also important differences between these two dialectal areas.

3.2. The praying mantis

Table 1 shows – purely by way of illustration – a very small sample of dialectal forms meaning 'praying mantis', among all those found in Romance languages and on the Japanese archipelago.

Table 1 A few dialectal forms meaning 'praying mantis'

Romance forms (ALiR 2.a)	Japanese forms (LAJ maps 229, 230)
prˌɛgadˈiw	ogamimusi
rrˈɛzɐ rrˈɛzɐ	ogamu
mɐrˈiɐ lˈovɐ	ogametaroo
kˈorta	itokirimusi
sɛɣamˈanu	kusakarimusi
sɛrramˈanos	nokogirimusi
kavˈallo	syorouma
kˈaḍḍu e dˈeu̯s	hotokeuma
kɐvɐlˈiɲu di nˈɔsɐ siɲˈorɐ	dondomusi

The first observation to be made from this initial comparison is that a dialectology that focuses all its attention on forms is completely at a loss when it comes to comparing languages that have no genetic relationship with each other. In its original state, the LAJ map legend is perfectly inapplicable to a map representing the Romance territory. On this basis, the only comment that can be made is: "the forms meaning 'praying mantis' in Japan and the Romance-speaking area have nothing to do with each other". The impossibility of going any further makes the comparison meaningless.

But this all changes with motivational analysis. It is now a matter of questioning the fundamental representations (i.e. motives) that gave rise to the various names for the mantis. To access these representations, we need to look at how they were "translated" into words during the lexical creation process, since these expressions always have their own literal meaning, which can be called *désignant*, and which, when naming an object in reality, makes it possible to associate linguistic material with a referent in an easily intelligible way. In what follows, *désignants* are presented between curved angle brackets (‹›).

Table 2 repeats the forms in Table 1, adding the *désignants*. Thus prˌɛgadˈiw 'praying mantis' breaks down into *prèga* 'pray' + *Dieu* 'God' and we understand that what motivates this name is the shape and position of the insect that seems to join its limbs to pray. This motive appears in the right-hand column of the table, grouping together a set of *désignants* inspired by the same guiding representation. The whole table functions according to this principle.

Table 2 Same forms as in Table 1, with *désignants* and motives

Romance (ALiR 2.a[4])		Japanese (LAJ maps 229, 230)		Motives
Forms	*Désignants*	Forms	*Désignants*	
prˌɛgadˈiw	‹God-praying (insect)›	ogamimusi	‹praying insect›	1. The posture of the mantis that looks like it is praying
rrˈɛzɐ rrˈɛzɐ	‹pray pray›	ogamu	‹pray›	
mɐrˈiɐ lˈovɐ	‹Mary-praying (insect)›	ogametaroo	‹praying Jack›	
kˈorta	‹cut!›	itokirimusi	‹thread-cutting insect›	2. The mantis's limbs resembling cutting tools
sɛɣamˈanu	‹hand-cutting (insect)›	kusakarimusi	‹grass-cutting insect›	
sɛrramˈanos	‹hand-sawing (insect)›	nokogirimusi	‹saw-insect›	
kavˈallo	‹horse›	syorouma	‹horse (of the soul of the departed?)›	3. A general comparison with the horse, sometimes with a religious reference
kˈadd̬u e dˈeu̯s	‹horse of God›	hotokeuma	‹horse of Buddha›	
kɐvɐlˈiɲu di nˈɔsɐ siɲˈorɐ	‹horse of Saint Mary›	dondomusi	‹horse insect›	

It is striking that dialectal terms that have absolutely nothing in common in terms of form, reveal representations and naming procedures that are so common from the Mediterranean basin to Japan. This method seems to reveal, thanks to dialectal data, certain mechanisms of human cognition, which etymologists are increasingly trying to take advantage of.

To convince ourselves that these are not isolated facts due to mere chance, we can project onto a map the very numerous data relating to the three motives presented above[5]. In the classification of Japanese data for this section and the following one, we used exclusively the resources of the *Linguistic Atlas of Japan Database* (LAJDB)[6] that were a decisive help. The map that we propose here (Fig. 7) was produced using ShinyDialect[7], an online cartographic tool capable of automatically producing areal maps extrapolated from point data, but which does not yet allow the display of multiple responses for the same survey point. When several motives were attested in the same Japanese locality, we chose to systematically favor the least common motive on a national scale. We sometimes tried to apply the same principle to single responses that seemed to condense two different motives.

Some forms were difficult to classify. We chose to classify *ogama-* type forms in the prayer motive, bearing in mind that they may not be unrelated to *kama* 'sickle'. On the

[4] See all the data in García Mouton (2001).
[5] For reasons of size and balance, we have decided not to include Romania on our maps. This does not mean, however, that this language area is at odds with our observations. The patterns described in this article are also found in Romania.
[6] https://www.lajdb.org/. More information about the LAJDB is to be found in Kumagai (2016).
[7] https://ljk-codes.imag.fr/shinydialect-v2/.

other hand, it seemed natural to classify compound forms containing *kama* as ‘bladelike limbs’, even when the other part of the word was not clearly identified. To limit the risk of error, we avoided classifying some uncertain data, unless the LAJ comments invited us to do so. For example, *-kiri*, on its own and as a possible avatar of *kirigirisu* ‘kind of grasshopper’ (LAJ comments, vol. 5, p. 91), did not justify inclusion in the ‘bladelike limbs’ category, even though there should be no shortage of speakers who see it as a form related to *kiru* ‘to cut’. We have therefore decided not to classify forms such as *ibokiri(musi)* or *kubokiri*.

For the rest, we think we have followed common sense. In the motivational category of prayer we have included both explicit designations of prayer and forms such as *negisama* ⟨shinto priest⟩, which are not unlike Romance names for the mantis (one thinks of *monaca* ⟨nun⟩ or *cura* ⟨priest⟩).

The “other” category brings together all the other forms: those that represent other motives, and those we couldn’t classify with any certainty.

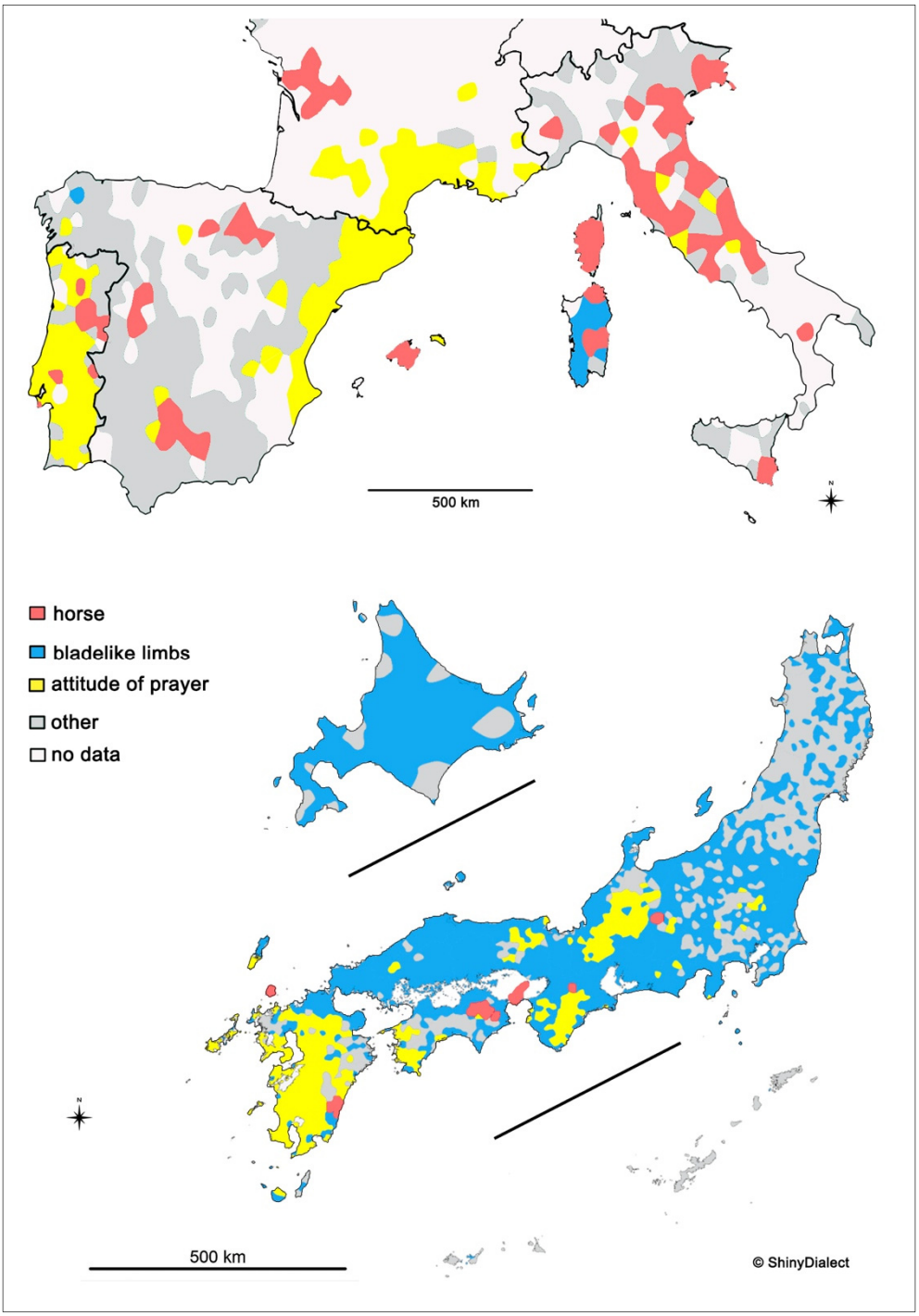


Fig. 7 Principal motives for 'praying mantis'

Despite the limited amount of data available for Spain, Italy and, above all, France, it is very clear that commonalities between the Romance world and Japan can be observed in various parts of the two large areas under consideration. Sardinia and a locality in Galicia thus echo standard Japanese, while western Japan responds to Portugal and the Catalan-Occitan area (plus a few Italian localities) for the image of the praying insect. The image of the horse, dominant in Corsica and scattered throughout the rest of the Romance area, is present very sporadically, but throughout the entire length of Western Japan.

The commonalities between Japan and the Romance domain are therefore undeniable. And yet, within these common motives, an element of cultural specificity emerges, as the ‹horse of God› (or of the Virgin, or of a Saint) becomes notably ‹horse of Buddha› in Tokushima Prefecture and on the island of Awaji[8].

Other phenomena common to both linguistic groups have not been included in this article. One thinks in particular of the many names shared between mantis, grasshopper, cricket or other more unexpected species (‹snake› in Italo-Romanic, ‹lizard› in Japanese).

Finally, some Japanese motives seem absent in the West, such as that manifested by the name *ibomushi* ‹wart insect›, based on the archipelago's popular belief that warts disappear if they are presented to mantises to gnaw on. Conversely, there are romance motives not observed in Japan.

3.3. The tadpole

In a similar vein, Del Giudice (2022) draws an equivalent parallel between Japanese dialects and the languages spoken in the vicinity of the city of Nice, regarding tadpole designations. As with the praying mantis, the motives common to Japan and this tiny Romance area are obvious. Some of these data are recalled as a sample in Table 3 (where Occitan forms from other regions have been added as a complement, indicated in italics).

[8] The fact that the mantis is called ‹horse› may simply be due to the fact that this insect is perceived as a kind of grasshopper, itself sometimes called a horse – at least in Western languages. The explanation given by the ALE commentary (1.1., p. 147), according to which the horse is a jumping animal in the same way as the grasshopper, would deserve to be put in competition with other leads.

Table 3 'tadpole': dialectal forms, *désignants* and motives

Occitan (THESOC, ALF…) Forms	*Désignants*	Japanese (LAJ maps 221, 222, 223) Forms	*Désignants*	Motives
babiòt	‹toad (+ diminutive suffix)›	kaerunoko	‹frog baby›	1. The tadpole is just seen as a frog, a toad (often as a baby frog, a baby toad)
babion	‹toad (+ diminutive suffix)›	hikinoko	‹toad baby›	
granhotil	‹frog (+ diminutive suffix)›	wakudonko	‹frog baby›	
testassa	‹head (+ augmentative suffix)›	atamadekka	‹head (+ big?)›	2. The tadpole has a remarkably big head
testard	‹head (+ suffix)›	atamabuto	‹head (+ thick?)›	
capgròs	‹big head›	atamadaezin	‹head (+ ?)›	
caceta	‹ladle (+ diminutive suffix)›	otamazyakusi	‹ladle›	3. The shape of the tadpole is reminiscent of kitchen utensils
paeleta	‹frying pan (+ diminutive suffix)›	syamozi	‹rice scoop›	
culhièra	‹spoon›	nabehuta	‹pot lid›	

Once again, the three main motives of the Nice region, well shared by Japanese dialects, appear in the right-hand column of the table. In addition to the Occitan-speaking regions, they are also abundantly represented throughout the Romance world.

Extending the study area to a wider Romance scale[9], then projecting onto a map all the ALiR[10] and LAJDB data corresponding to one of these three motives[11], we obtain Fig. 8.

As in Fig. 7, we had to eliminate multiple responses and find a solution for hybrid forms. To this end, we gave priority to non-standard forms. We also favored the representation of the "head" motive, which is quite rare in the Japanese archipelago (it is represented on the map by round dots that make it stand out more clearly). This is why an interesting case such as *atamazyakusi* (where *atama* 'head' appears) has been

[9] And even beyond, since we have added non-Romance data from Brittany and the Basque Country (for the latter, the data is even non-Indo-European). To map these two areas, we have used the *Nouvel Atlas Linguistique de la Basse-Bretagne* (NALBB, map 235) and the *Euskararen Herri Hizkeren Atlasa* (EHHA, map 118).

[10] Completed – and where necessary slightly revised – using information from the *Atlas Linguistique et ethnographique de la Picardie* (ALPic, map 601), the *Atlas Linguistique et ethnographique de la Champagne et de la Brie* (ALCB, map 1115) and Tuaillon & Stiers (1986).

[11] As with mantis, we have taken into account the LAJ's comments for Japanese dialectal terms that we were unable to interpret ourselves. Thus, all the forms *atamabuto*, *atamabukko*, *atamahucyage*, *atamadekka*, *atamadaezin*, *suburubuuta*, *taaciburaa* were linked to the motivation of the head, following the remark "all of these forms are thought to focus on the big head of the tadpole" (「いずれも『おたまじゃくし』の頭の大きいことに注目した語形と考えられる」) (vol. 5, p. 67).

placed in the “head” category and not in “kitchen utensil”, even though the proximity of this form to *otamazyakusi* is obvious.

Alongside the three motives shown in Table 3, we wanted to highlight the distribution of a particular lexical type. This is *otama* when not followed by *-zyakusi*. This type includes forms generally in geographical contact with *otamazyakusi*, such as *otama* and its avatars *otamakko*, *otamaccyo(o)*, *kaerottama*, etc. The LAJ comments themselves hesitate to see *otama* as a simple truncation of *otamazyakusi* or as a base *tama* ‘ball’ (p. 66), itself possibly induced by *atama* ‘head’ (p. 67). On the other hand, our map incorporates into the pattern “frog, toad (baby frog, baby toad)” Romance designations such as ⟨frog egg, toad egg⟩, which the Japanese form *otamago* also seems to vaguely echo (*tamago* meaning ‘egg’). Forms with *otama* are therefore likely to condense several motives. We have preferred to project the *otama* type itself without deciding on its fascinating ambiguity.

In other cases where we had doubts, prudence led us to abstain, which deprives some data of a projection on the map. For example, in *gorokko* we thought we recognized *gaerokko* (= *gaeruko* ⟨baby frog⟩). We were also tempted to see in *genguruko*, *ge(e)kuro* distorted forms – clearly metathetical for the second – of *gaeru(no)ko*. But in all these cases, the comments (pp. 68–69) seemed to interpret things differently or remained rather unexplicit, ultimately convincing us not to map the forms concerned. In this case, the loss is not noticeable, as the “frog, toad” motive is already abundantly represented on the map of Japan. It could even have been more abundant, but we didn’t keep forms where the modifier meant ‘frog’, whenever the identity of the modified word was unclear. This was the case, for example, with *gaerukuta* and *gaeruma(n)cyo*. All this being said, the presence of gaps and probable errors in our map concerns only a very small part of the data.

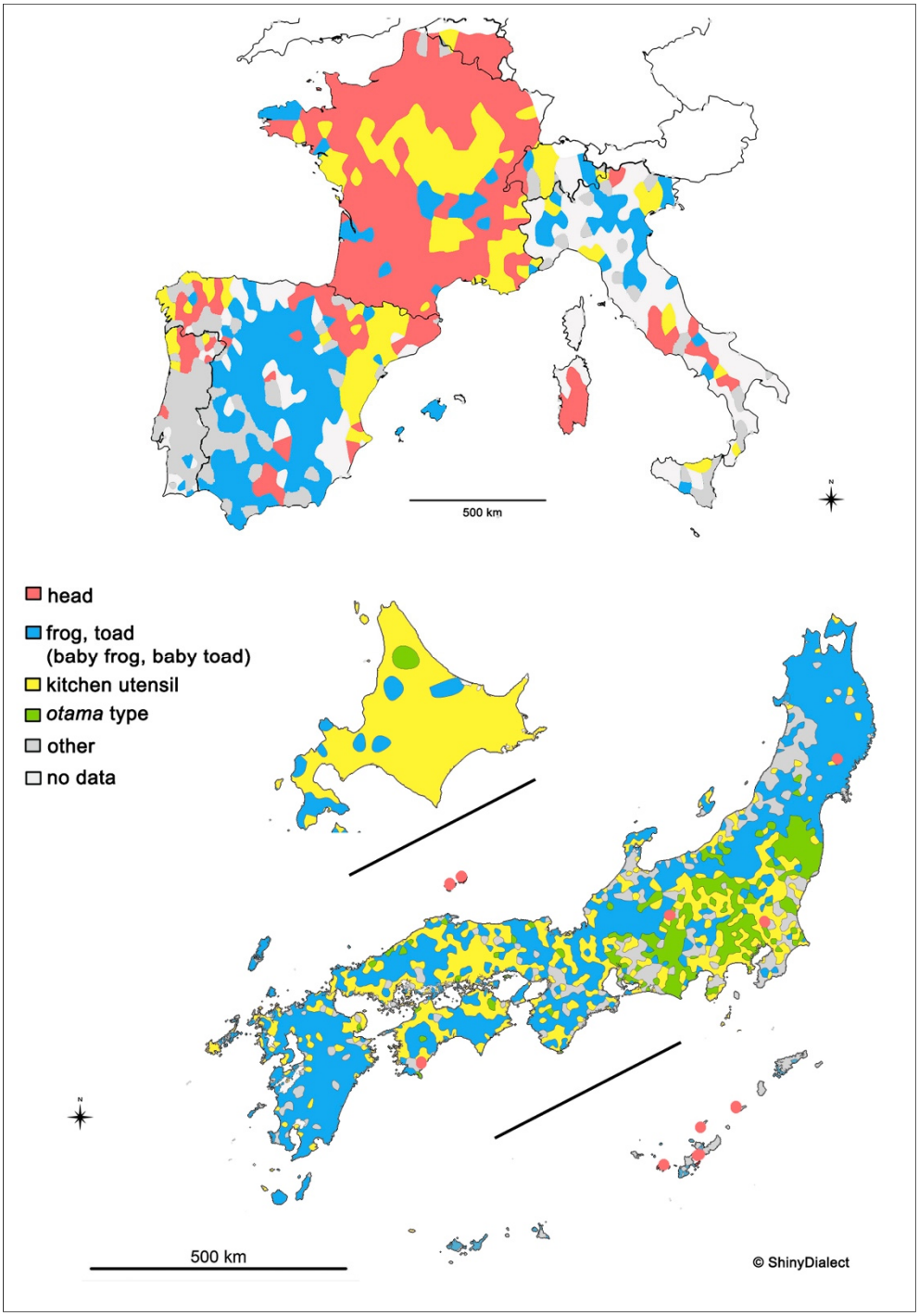


Fig. 8 Principal motives for 'tadpole'

The main impression to be gained from Fig. 8 is that, as expected, motives are common to both Western Europe and Japan, but in very contrasting proportions. The idea of the head, particularly prevalent in France and Sardinia, carries infinitely less weight in Japan, where the corresponding forms are marginal. Paradoxically, the strength of this representation is indisputable, for in this country where *otamazyakusi* and the frog motive reign supreme, it can be found, in a fragmented way and in very different forms, from the far south (Ryukyu Islands) to Iwate Prefecture in northern Honshu. Are these ancient forms in retreat because of the dominant types, or more recent creations/reconstructions, possibly accidental, driven by the obvious anatomical specificity of the tadpole?

The other main motives are fairly evenly distributed between the two main areas studied, although it is worth noting that the comparison between tadpoles and kitchen utensils relies almost exclusively, in Japan, on the *otamazyakusi* type, whose generous distribution may well have been favored by the weight of the standard. Nonetheless, this map allows us to verify once again that Japan and the Romance world do indeed use common representations to name similar referents.

To keep it as simple and evocative as possible, Fig. 8 (like Fig. 7) is limited to 3 outstanding motives, although others could have been mentioned. The motivational classification proposed by Médélice (2009) is much more complete and precise than the overview given here. In particular, she mentions the names of various tools (hammer, fork, nail...) and comparisons with fish, which are not considered in our map[12]. Readers interested in finding out more are invited to take a look at ALIR's comments.

4. Beyond praying mantises and tadpoles

4.1. Complementary data: "easy motives" and more surprising representations

The designations of the praying mantis and the tadpole, in Japanese and Romance dialects, were good examples for outlining a number of naming reflexes widely shared across these two territories. Apart from 'mantis' and 'tadpole', not all concepts allow us to draw such massive parallels from one linguistic space to another, but cross-border motives are nonetheless observed with fairly significant regularity. We might point out that a scarecrow is literally the object that frightens crows, and that the same *désignant* is found in the Japanese word *karasuodosu* (*karasu* 'crow' + *odosu* 'scare', LAJ map 190), alongside other similar names such as *suzumeodosi* (*suzume* 'sparrow' + *odosi* 'what scares'), which is reminiscent of the Nicoise *espavèntapàsseras* (*espavènta*

[12] A number of forms such as *otamazyakko*, *otamadozyoo*, *otamakkozyoro*, whose head refers to the notion of fish, are classified in our map as "*otama* type".

'scares' + *pàsseras* 'sparrows'). Sometimes the idea of scaring is enough in itself, and *odosi*, *odosu* is rendered in Occitan as *espavental*, *espaurugal* (derivatives of *espaventar*, *espaurugar* 'scare'). Another name, *waraningyoo* ('straw' + 'puppet'), is not discordant with the form *palhasso* (traditional carnival puppet made of straw) used in the region of Nice to designate the scarecrow.

These fundamental denominations often seem to be generated by fairly obvious representations (the tadpole is a frog, the scarecrow is a man of straw), whether original forms or forms that were subsequently remotivated (in this respect, the tadpole offers several examples of remotivation-condensations such as *atamajyakusi* and *otama*). But sometimes, shared representations are a little more surprising. It is the case for the mantis, when it is seen as a holy horse, or for the owl (LAJ, map 212) when it is compared with a cat. And once again, forms like *nekodori* ⟨cat-bird⟩ or *mya(a)cuko* ⟨meow + ?⟩ look particularly similar to occitan *catamiaula* ('female cat' + 'to mew 3sg'), *miaula*, *miauneta* ⟨mew (+ diminutive suffix)⟩ or french *chat-huant* ⟨hooting cat⟩. It is also a safe bet that *yo-* in a number of Japanese designations (*yotori*, *yotaka*, *yohukuro*, *yocuku*...) means 'night', just as this bird is called *nuechola* ⟨night (+ diminutive suffix)⟩ in several Occitan dialects. It is on correspondences of this kind that research is now trying to break through the most obscure etymologies.

4.2. Further perspectives

In any case, the common motives evoked in this article speak volumes about the way human beings interpret the world, and about the fundamental workings of language (what is it about a referent that strikes the imagination, and how do we translate this representation into words?). New words are not invented by chance, and the changes they undergo over time are often – probably more than is generally believed – linked to deep-seated representations. We therefore believe that semantic motivation is a possible gateway to more in-depth studies involving cognitivists. The aim would then be to move on from the lexical case study stage to that of the general functioning of our brain and the interface between thought and expression. A fine program for data derived from dialectology!

From a strictly geolinguistic point of view, the main purpose of Figs. 7 and 8 was to show that the phenomena mentioned are significant, since they cover considerable areas within the dialectal spaces considered and given that even the non-dominant motives ("horse" and "head", in Japan) extend in dotted lines over hundreds of kilometers. The precise distribution of motives, however, was not decisive for our purpose. It would still be interesting to see whether words whose motive depends on the place of residence (along the lines of what was attempted in Del Giudice (2020)), give rise to similar distributions in Japan and other parts of the world. The identification of

comparable motivational areas based on geographical or social criteria (types of community) could lead to new discoveries about certain constants in the act of naming things.

References

Atlases

ALCB: Bourcelot, Henri (2011) *Atlas Linguistique et ethnographique de la Champagne et de la Brie*. Paris: Editions du CTHS.

ALE: Alinei, Mario *et al.* (1983-) *Atlas Linguarum Europae*. Assen / Maastricht: Van Gorcum, Rome: Istituto Poligrafico e Zecca dello Stato, Bucarest: Academia Română.

ALEANR: Alvar, Manuel, Antonio Llorente and Tomás Buesa (1979-1983) *Atlas Lingüistico y Etnográfico de Aragón, Navarra y Rioja*. Zaragoza: Departamento de Geografía Lingüistica – Institución Fernando El Católico de la Exma.

ALF: Gilliéron, Jules and Edmond Edmont (1902-1910), *Atlas Linguistique de la France*. Paris: Champion.

ALiR: AA.VV. (1996) *Atlas Linguistique Roman*. Rome: Istituto Poligrafico e Zecca dello Stato, Alessandria: Edizioni dell'Orso.

ALMC: Nauton, Pierre (1955-1961) *Atlas Linguistique et ethnographique du Massif Central*. Paris: Editions du CNRS.

ALP (Vol. 4): Bouvier, Jean-Claude, Claude Martel and Guylaine Brun-Trigaud (2016), *La langue d'oc telle qu'on la parle. Atlas Linguistique et ethnographique de la Provence*. Forcalquier: Alpes de Lumière.

ALPIc: Carton, Fernand and Maurice Lebègue (1998) *Atlas Linguistique et ethnographique de la Picardie*. Paris: Editions du CNRS.

EHHA: Videgain, Charles (dir.) (2008-2020) *Euskararen Herri Hizkeren Atlasa.* Bilbao: Euskaltzaindia. (Map number 118, https://www.euskaltzaindia.eus/dok/iker_jagon_tegiak/ehha/118.pdf)

GDA: Language Research commission (1986 [1906]) *Kogoho Bunpuzu* 『口語法分布図』[Grammatical Dialect Atlas]. Tokyo: Kokushokankokai [国書刊行会].

LAJ: National Language Research Institute (Kokuritsu Kokugo Kenkyujo [国立国語研究所]) (1966-1974) *Nihon Gengo Chizu*『日本言語地図』[Linguistic Atlas of Japan], 6 vols. Tokyo: National Printing Bureau, Ministry of Finance (Okurasyo insatsukyoku [大蔵省印刷局]).

NALBB: Le Dû, Jean (2001) *Nouvel Atlas Linguistique de la Basse-Bretagne*. Brest: CRBC. (Map number 235, https://nakala.fr/10.34847/nkl.7d61xeok).

PDA: Language Research commission (Kokugo Chousa Iinkai [国語調査委員会]) (1986 [1905]) *On'in Bunpuzu*『音韻分布図』[Phonetic Dialect Atlas]. Tokyo: Kokushokankokai [国書刊行会].

THESOC: Dalbera, Jean-Philippe *et al.* (1992-) *Thesaurus occitan*. Université de Nice / Université Côte d'Azur and CNRS UMR 7320. http://thesaurus.unice.fr/

Other references

Brun-Trigaud, Guylaine, Yves Le Berre and Jean Le Dû (2005) *Lectures de l'Atlas Linguistique de la France de J. Gilléron et E. Edmont. Du temps dans l'espace*. Paris: Editions du CTHS.

Dalbera, Jean-Philippe (2001) Les désignations romanes de l'orvet (carte et commentaire). *Atlas Linguistique Roman* vol. 2.a, 391–404. Roma: Istituto Poligrafico e Zecca dello Stato.

Dalbera, Jean-Philippe (2006) *Des dialectes au langage. Une archéologie du sens*. Paris: Champion.

Dauzat, Albert (1922) *La Géographie linguistique*. Paris: Flammarion.

Del Giudice, Philippe (2020) Désignation et représentation des éléments topographiques dans les dialectes de France: l'élévation de terrain. *Dialectologia et Geolinguistica* 28: 55–80. DOI: https://doi.org/10.1515/dialect-2020-0002

Del Giudice, Philippe (2022) Furansu Yoroppano gengogakuni okeru mochibeshon. Sono riron, hohoronto seika『フランス・ヨーロッパの言語学におけるモチベーション。その理論、方法論と成果』[Linguistic motivation in France and Europe, Theory, method and results]. *Flambeau* 48: 10–31. DOI: https://doi.org/10.15026/124877

García Mouton, Pilar (2001) Les désignations romanes de la mante religieuse (carte et commentaire). *Atlas Linguistique Roman vol.* 2.a, 239–280. Roma: Istituto Poligrafico e Zecca dello Stato.

Iwata, Ray [岩田礼] (2021) Parallelisms in lexical changes across languages: Analogical changes in Chinese and French time words. *Studies in geolinguistics*『地理言語学研究』1: 123–141. DOI: https://doi.org/10.5281/zenodo.5529335

Kumagai, Yasuo [熊谷康雄] (2016) Developing the Linguistic Atlas of Japan Database and advancing analysis of geographical distributions of dialects. In: Marie-Hélène Côté, Remco Knooihuizen and John Nerbonne (eds.) *The future of dialects*, 333–362. Berlin: Language Science Press. DOI: https://doi.org/10.17169/langsci.b81.159

Médélice, Jeanine Élisa (2009) Les désignations romanes du têtard (carte et commentaire). *Atlas Linguistique Roman* vol. 2.b, 215–253. Roma: Istituto Poligrafico e Zecca dello Stato.

Saussure (de), Ferdinand (2005 [1916]) *Cours de Linguistique générale*. Paris: Payot.

Tuaillon, Gaston and Roger Stiers (1986), Les désignations gallo-romanes du têtard. *Géolinguistique* 2: 97–118.

Publication history

Date received: 28 June 2024
Date accepted: 28 July 2024

Research article

Geolinguistic analysis of the spelling in the 1923 Burgundian text *L'âme du Morvan*

ITO, Reiko
Japan Women's University

Abstract: The first edition of the text, *L'âme du Morvan*, in the Morvan dialect was published in 1923. Morvan is in central Burgundy. Historically, it belonged to neighboring regions and was never united (Richard 1973: 153, 291). According to Ito (2020, 2021, 2022), no unified language was spoken in Morvan during the mid-20th century. In the 39 stories in this book, we find variations in spelling. Because some words are spelled differently in different stories, we hypothesize that the 39 stories could have been written in different spoken language and seek to determine which regional spoken language the stories were written in. First, we presume the pronunciations of forms representing 261 words across 39 stories. Second, we compare them with the pronunciations of 18 Morvan locations in the *Atlas Linguistique et Ethnographique de la Bourgogne* (ALB), the Linguistic Atlas of Burgundy. Based on the coincidence ratio of the 18 locations, we classify the 39 stories using cluster analysis. Next, we visualize the group that is the most consistent with the locations. Finally, we consider whether the 39 stories were written in the same spoken language or several spoken languages. The spelling analysis suggests that the 39 stories may have been written slightly differently; however, they were written in the dialect of the Saulieu region—the Bas-Morvan dialect.*

Keywords: *L'âme du Morvan*; Morvan; Burgundy; dialectology; geolinguistics

1. Introduction

Morvan is a region in the center of Burgundy, France, covering an area of 60 km from east to west and 100 km from north to south. It is a naturally beautiful area, enriched with forests and pools. Its central high-altitude area reaches approximately 900 m above sea level. The region was designated as *le Parc Naturel Régional du Morvan* (the Morvan Regional Natural Park) in 1973. The spoken language of Morvan is written as *morvandeau*, *morvandiau*, or *morvandelle*.[1] The dialect ceased

ITO, Reiko. 2024. Geolinguistic analysis of the spelling in the 1923 Burgundian dialect text *L'âme du Morvan*. *Studies in Geolinguistics* 4: 40–58. doi: https://doi.org/10.5281/zenodo.13948557

* This article is a further analysis based on the presentation at the 6th Annual Conference of the Japanese Association for Geographic Linguistics (Osaka University; June 1, 2024). I would like to thank everyone who provided comments.

[1] Le Trésor de la langue française informatisé (TLFi) http://stella.atilf.fr/

to be spoken in most parts of Burgundy by the 1960s, although there were still residents in Bresse and Morvan who spoke both the dialect and French (Taverdet 1973: 320–325) (see Figure 2).

Fig. 1 France

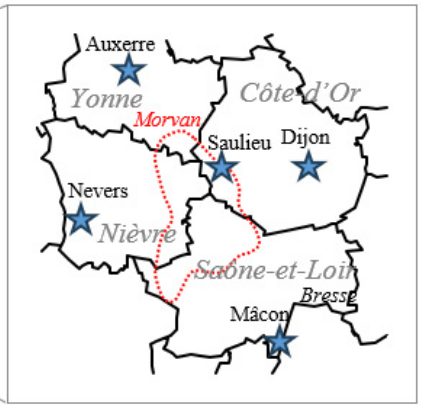


Fig. 2 Burgundy

2. Previous research on the Morvan dialect

2.1. Régnier (1979a, b), Bertrant (1979)

The three books by Régnier (1979a, b) and Bertrant (1979) constitute a series called *Les Parlers du Morvan*. Régnier (1979b) is a linguistic map of his survey in northern and central Morvan in the 1940s–1960s. He analyzed Morvan's spoken language phonetically, morphologically, and lexically (Régnier 1979a). Bertrant (1979) transcribed pronunciations found in the four provinces straddled by Morvan (Côte-d'Or, Saône-et-Loire, Yonne, and Nièvre) into spellings (see Figure 2 for provinces). Régnier (1979a: 186) posits that the spoken language of Morvan was the last barrier to the Burgundian dialect against the pressure of the Nivernais dialect from the West.

2.2. Ito (2020, 2021)

Ito (2020, 2021) geolinguistically analyzed the pronunciations of words with the same etymon, such as C(A)-, -CC(A), PL-, and BL-. The results confirm that in Morvan in the mid-20th century, words with the same etymon were capable of being pronounced differently, even at the same location, and no phonetically unified dialect was spoken. C(A)- and -CC(A) are widely distributed as [ʃ], while [s] forms a language island around the plateau of Morvan—or is continuous with [s] or [ts] in

southeastern Burgundy. This supports Dauzat's theory that [ts] was once widespread throughout Burgundy (Dauzat 1944: 224–231) and Régnier's theory that Morvan's spoken language has something in common with Francoprovençal[2] (Régnier 1979a: 133–134). We also found that the boundaries—between [pl]/[bl], continuous from western Burgundy, and [pj]/[bj] continuous from eastern Burgundy—are complex in Morvan.

2.3. Ito (2022)

Ito (2022) analyzed linguistic-geographical subject personal pronouns and the verb « être » (present indicative) in Morvan in the 1960s. The results confirm the lack of phonetic unity with respect to subject personal pronouns and the forms of conjugation of this verb. In addition, the phonetic variations in the Burgundian dialect remained at relatively high elevations in central Morvan.

3. Dialect text *L'âme du Morvan*

L'âme du Morvan, in the 1923 edition, was published by Mrs. Gervais, who was an editor. It consists of 39 short stories, two theatrical scenarios, and descriptions of beliefs and customs. The author's name is not printed, but "author A. Guillaume" is handwritten on the facing page of the digitized book available on Gallica.[3]

The book was reprinted in 1971 with Guillaume's name as the author by the organization *les Amis du Vieux Saulieu* (the Friends of Old Saulieu). In addition to the contents of the 1923 edition, one scenario has been added. According to the preface of the 1971 edition, Alfred Guillaume was born in Saulieu (Figure 2). He worked as a veterinarian and was elected as the city's first magistrate. The reprint was authorized by Miss Guillaume, Guillaume's daughter, and Mrs. Guerrier, a descendant of Mrs. Gervais. The preface reads: "*L'âme du Morvan* was first published in 1923 … Only a few copies bear the author's name: Mr. A. Guillaume." Based on this description, the 1923 edition is noted as Guillaume (1923). A part of the preface of the 1923 edition reads as follows:

> We have brought together the stories that have already appeared in the published almanacs and those that have not yet been published. Finally, we have added the beliefs, customs, superstitions, proverbs, and sayings peculiar to the region … All of these are written in the dialect of the Saulieu

[2] Francoprovençal is a language originally spoken in east-central France. Southeastern Burgundy belongs to this region.
[3] Gallica is the Digital Library of *la Bibliothèque Nationale de France* (BnF) and its partners. https://gallica.bnf.fr/ark:/12148/bpt6k98015409?rk=21459;2

region, but with some differences; this is the dialect of the entire Bas-Morvan.[4]

3.1. Taverdet (2001)

According to Taverdet (2001: 210–211), Guillaume was a veterinarian and mayor of Saulieu, the capital of the Côte-d'Or canton. Taverdet states that traditional dialect authors did not write for the edification of patois speakers but for readers who were generally educated, wanted not to mock but to appreciate the peasantry, and liked to find old words (2001: 215).

3.2. Variations in spelling in the 39 stories of the 1923 edition

Among the 39 stories, we found variations in spelling, which can be classified into the following types.[5]

- Vowels
  « place » (place): *place*, *plaice*
  « roi » (king): *roi*, *roué*
- Consonants
  « chambre » (room): *chambre*, *chambe*
  « entendre » (hear): *entendre*, *entend'e*
- *l* or *i* after *p*/*b*
  « blé » (wheat): *blé*, *bié*
- With or without a hyphen
  « année » (year): *année*, *an-née*
- Compound variation (missing or replacing letters, inserting apostrophes).
  « horloge » (clock): *heurloge*, *r'loge*
  « regarder » (watch): *eurgairder*, *regairder*, *r'gairder*
- Synonyms
  « brebis » (ewe): *borbis*, *oueille*

In terms of the relationship between stories and their variations, words can be grouped into three categories:

1) Words spelled uniformly in the 39 stories.
2) Words not spelled uniformly in the same story.
3) Words spelled differently in different stories.

We note that some words were spelled differently in different stories. As already mentioned, the preface to the 1923 edition states that the 39 stories are not the creation of a single author but a collection of existing stories and unpublished works.

[4] Bas-Morvan is a plateau in the northern part of Morvan, connected to the Paris basin to the northwest.
[5] Forms written in the corpus are italicized. « » denotes a word in standard French.

It also states that the entire book is "written in the dialect of the Saulieu region," with "slight differences" … nevertheless, "in the spoken language of Bas-Morvan." Therefore, by analyzing the spellings, we can prove that the 39 stories were written in more than one spoken language.

4. Research question

The research question for this study is:

> Which regional spoken language of Morvan were the 39 stories in *L'âme du Morvan* (1923 edition) written in?

The study is original because no research to date has analyzed the spellings of these stories in Guillaume (1923) from a geolinguistic perspective.

5. Methodology

5.1. Corpus

The 39 stories from the 1923 edition were used as the corpus in this study. They comprise utterances (in dialect or in SF[6]) and narration (in dialect). The utterances of certain persons, for example *monsieu* « monsieur, » are always written in SF. Phrases written in SF were excluded from the corpus. The size of the corpus is as follows:

- Word types: 8,308
- Word tokens: 56,673

As apostrophes sometimes appear inside a word, they were counted as letters. In other words, tab'e « table » (table) was recognized as a single word.

5.2. Analyzed words

Headwords of the *Atlas Linguistique et Ethnographique de la Bourgogne* (ALB), the Linguistic Atlas of Burgundy (261 words)[7], were analyzed. They were selected based on the following criteria.

1) The meaning of form, representing 261 words, can be found in Morvan's *Dictionary of Dialects* (Chambure 1878) and the *Dictionary of Burgundian French* (Taverdet and Navette-Taverdet 1990). However, in the first note to

[6] SF means standard French.
[7] The 261 words include subject-verb conjugation combinations. For the sake of convenience, we use "word" in this study.

the first story, *Ein tor de couéchon*, the author writes: "To make patois easier to read and understand, we have adopted spellings that are as close as possible to French and to the pronunciations of words. We do not, therefore, follow the spelling indicated by Chambure in his Glossary" (Guillaume 1923: 7). When referring to Chambure (1878), it was necessary to investigate possible spellings rather than retain his spelling.

2) Substantive (nouns, adjectives, adverbs, infinitives of verbs, and conjugated forms of subjects and verbs).
3) Etymology of the FEW[8].

Figure 3 shows that the forms representing the 261 words appeared 3,874 times in the corpus.

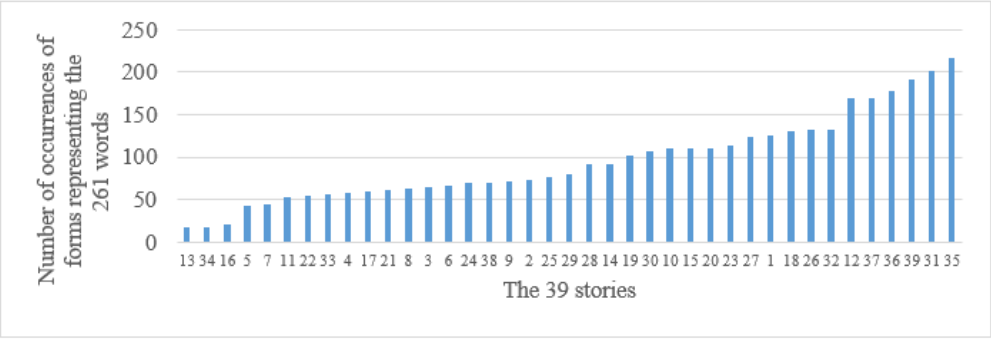


Fig. 3 The number of occurrences of forms representing the 261 words

Figure 4 shows the tokens and the number of analyzed words in each story. As expected, the greater the number of tokens, the greater the number of words analyzed.

Of the 261 words, 187 appeared in unified form in 39 stories. In contrast, 74 words had multiple forms: 32 were spelled differently in different stories, and 42 had multiple spellings in the same story.

[8] von Waltburg (1922-2002) *Französisches etymologisches Wörterbuch. Eine Darstellung des galloromanischen Sprachschatzes.* (French Etymological Dictionary. A presentation of the Gallo-Roman vocabulary).

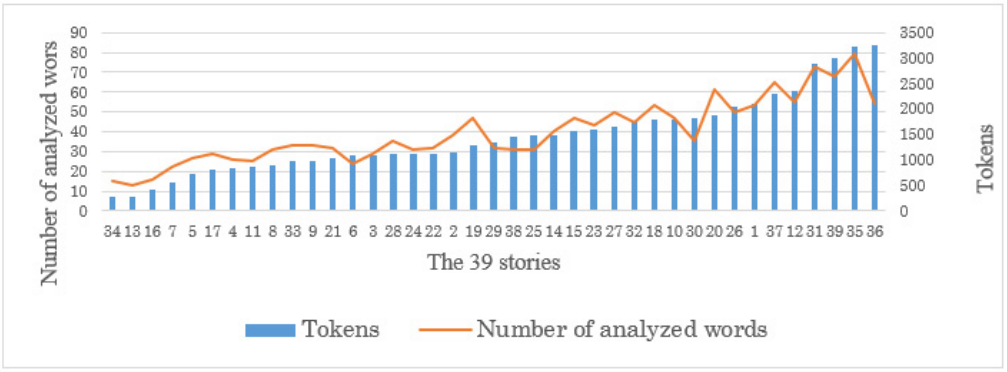


Fig. 4 The tokens and the number of the analyzed words in each story

5.3. 18 locations of Morvan in the *Atlas Linguistique et Ethnographique de la Bourgogne* (ALB)

Figure 5 shows The 18 locations of Morvan in the ALB: L27, L30, L52, L54, L62, L65, L70, L75, L77, L79, L80, L83, L85, L86, L89, L92, L93, and L102. The locations indicated by the dots were not used in this study.

Fig. 5 The 18 locations of Morvan in the ALB

5.4. Procedure

The relationship between spelling and pronunciation in the corpus was estimated by referring to the relationship between spelling and pronunciation in SF, based on Guillaume's approach (1923: 7): "To make patois easier to read and understand, we have adopted spellings that are as close as possible to French and to the pronunciations of words." The pronunciations were inferred from the forms representing the 261 words. Then, they were compared with the pronunciations of the 18 locations in the ALB for each story (classified "1" if they matched, and "0" if they

did not). The 39 stories were analyzed and classified into several clusters. The locations at which each cluster is most consistent was then determined.

6. Analysis

6.1. Pronunciation inference

The following validations of the relationship between spellings and pronunciations in the corpus were conducted to infer possible pronunciations from the forms representing the 261 words. Since the spoken Morvan language that this corpus was written in had not yet been determined, the pronunciations of all locations of Morvan in the ALB were used. Hence, the range of possible pronunciations was expanded.

Since Guillaume was a speaker of both dialect and SF, we assumed that the same spelling forms in SF were pronounced the same way as in SF.

6.1.1. Validation of the relation between vowel letters and pronunciations

When we refer to Catach's (1980: 62–75, 82–142) relationship between vowel graphemes and pronunciations in SF, some spellings are slightly inconsistent with pronunciations in the ALB. Generally, « é » is used for [e] and « è » for [ɛ] in SF. In the corpus, however, *é* is not always used to represent [e] and *è* is not always used to represent [ɛ]. The added graphemes and pronunciations are shown in bold (Table 1).

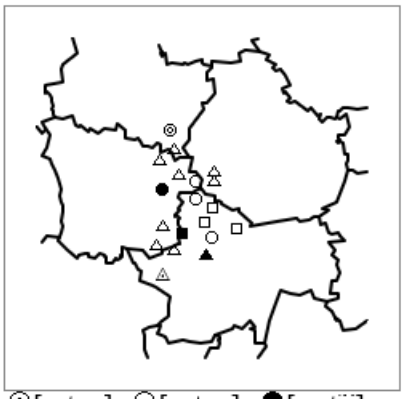
Table 1 Vowels graphemes and pronunciations in the corpus

Grapheme	Pronunciation	Grapheme	Pronunciation
a, à, â	[a], [ɑ]	an, am, em	[ɑ̃], [ã]
é, è, ê, ai, aî , ei, **eai**	[e], [ɛ]	en	[ã], [ɛ̃]
e	[e], [ɛ], **[aᵋ]**, [ə], ∅	in, ain	[ɛ̃], [ɛ̃]
i	[i]	oin, ouin	[wɛ̃], [wɛ̃]
o, ô, au	[o], [ɔ]	on, ôn	[õ]
ou, oû	[u]	oi	[wa], **[wɑ]**
u, û	[y]	**oé**	**[we]**, **[wɛ]**
ü	[y], **[ɥ]**	ou+vowel	[w]
u (+i)	[ɥ], **[y]**	i, y, yi, il, ill	[j]
eu, œu, **eû**	[ø], [œ]		

The letter « â » could represent an open back vowel in SF. The pronunciation of *pâteure* seems to be [pɑtø:r], indicated by the legend ○ in ALB 339: pâturage (pasture) (see figure 6). In the following, symbols represent legends in figures. In this

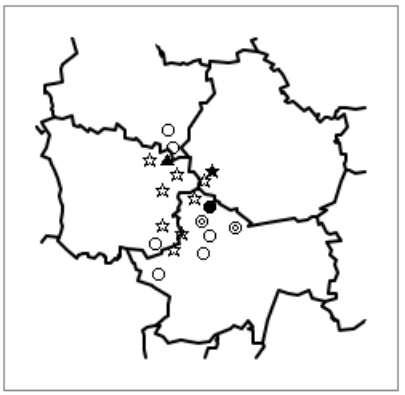
case, *â* represents an open back-vowel [ɑ]. In contrast, *crâcher* could be pronounced as ○[kraʃe] in ALB1328: cracher (to spit), with *â* indicating an open front vowel [a] (see Figure 7). These two maps show that *â* could be either an open front vowel [a] or an open back vowel [ɑ] in the corpus.

In SF, the vowel letters with circumflex such as « â, ê, î, ô, û » have some functions: to represent long vowels, to distinguish homophones, and to indicate traces of lost letters. Here, they were examined to determine whether they always represent long vowels in the corpus. The letter *î* was not used alone in the corpus.



◎[paty:r], ○[pɑtø:r], ●[pɑ:tʲji]
□[pɑ:ke], ■[pɑ:kʲjɛ], △[pɑ:ki],
▲[pɑki], ◬[pɑ:k]

Fig. 6 ALB339: pâturage

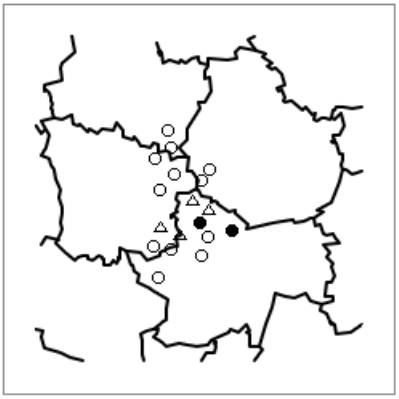


○[kraʃe], ●[krwɛʃe], ◎[krɔʃe],
▲[krɛse], ☆[køpe], ★[kœpre]

Fig. 7 ALB1328: cracher

As already mentioned, *pâteure* could be ○[pɑtø:r] in ALB339: pâturage and *crâcher* seems to be ○[kraʃe] in ALB1328: cracher (see Figures 6 and 7). In ALB111: saison, *sâyon* « saison » (season) seems to be ▲[sɑ:jõ] (see Figure 8). In the corpus, *â* could represent [ɑ], [ɑ:], and [a].

The forms *nouêr* and *nouair* for « noir » (black) seem to correspond to □[nwɛ:r] or to ■[nwɛr] in ALB1032: noir (see Figure 9). It is not possible to determine whether *ê* represents [ɛ:] or [ɛ].



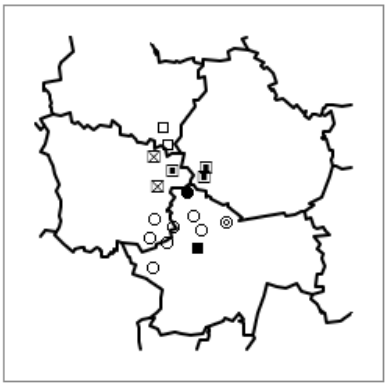
○[sɛ:zõ], ●[sɑ:zõ], ▲[sɑ:jõ]

Fig. 8 ALB111: saison

Côyier for « collier » (necklace) could correspond to ◎[ko:je] or to ○[koje] in ALB1254: collier (see Figure 10). The letter *ô* seems to represent [o:] or [o].

The form *tûyon* for « tison » (firebrand) could correspond to ○[ty:jõ] in ALB1450: tison (see Figure 11). However, *trûyot* and *truyot* for « trèfle » (clover) could correspond to △[tryjo] in ALB330: trèfle (see Figure 12). It means that *û* could be [y:] ou [y].

Thus, the vowel letters with circumflex, *â*, *ê*, *ô*, *û*, could be long or short vowels, depending on forms.

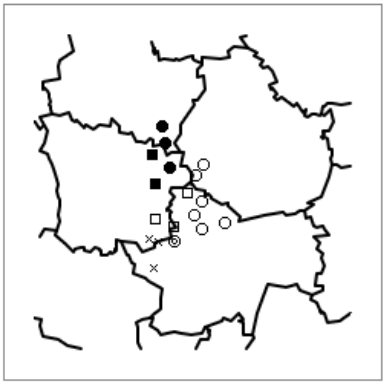


□[nwɛ:r], ■[nwɛr], ⊠[nwe:],
▣[nwɛ], ○[nwar], ●[nar], ◎[nwa]

Fig. 9 ALB1032: noir

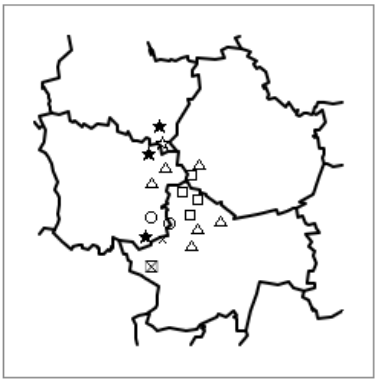
○[koje], ◎[ko:je], ●[kuje]

Fig. 10 ALB1254: collier



○[ty:jõ], ●[ty:ʒõ], ◎[ty:zõ]
□[tø:jõ], ■[tø:ʒõ], ×[brezjo]

Fig. 11 ALB1450: tison



○[trɛf], △[tryjo], □[tru:jo], ⊠[trɔlo]
☆[trylo], ★[try:lo], ×[tru:lo]

Fig. 12 ALB330: trèfle

The vowel letters without the circumflex cannot specifically represent short or long vowels in SF. We assumed that they were both represented in the corpus.

6.1.2. Validation of the relation between consonantal letters and pronunciations

Table 2 shows the relationship between consonant graphemes and pronunciations in the corpus. The grapheme « r » is pronounced as a uvular fricative [ʁ] in SF, but the ALB shows that it is an alveolar trill [r] in Morvan. For others, we referred to the relationship between graphemes and pronunciations in SF, following Catach (1980: 143–194).

Table 2 Consonant graphemes and pronunciations in the corpus

Grapheme	Pronunciation	Grapheme	Pronunciation
p	[p]	x, cc, xc	[ks], [gz]
b	[b]	ch	[ʃ]
t	[t]	j, g	[ʒ]
d	[d]	l	[l]
c, qu, k	[k]	**r**	**[r]**
g, gu	[g]	m	[m]
f, ph	[f]	n	[n]
v	[v]	gn	[ɲ]
s,ss,c,ç	[s]	ng	[ŋ]
vowel+s+vowel, z	[z]		

In SF, word-final consonants are generally voiceless. Conversely, « -l » and « -r » are, in general, voiced at the end of a word. However, « -r » is voiceless in infinitives with the ending « -er » and in some nouns with « -er ». In dialectal texts, because spellings are normally written to represent sounds, it is common to consider a consonant letter at the end of a word as voiced, if it exists. In the following section, we examine whether word-final consonant letters in the corpus could be pronounced. Of the 261 words, *sel* (salt) is spelled in the same manner as SF. However, there is no form with *-l* that differs from that of SF. Thus, *-l* is not verified.

The word « droit » (straight) has a « -t » that is voiceless in SF. In the corpus, it was spelled *drouet*. In ALB1301: droit, [t] was not pronounced in Morvan. The « -s » that indicates plurality is not pronounced in SF. In the corpus, the plural word « plumes » (feathers) is spelled *pleumes*. In ALB1187: les plumes, the [s] are not in Morvan. The unpronounced *-t* and *-s* are written in the corpus as well as in the SF. This clearly indicates the influence of SF on the spelling in the corpus.

The letter « -r » of « -er » in infinitives and some nouns is not pronounced in SF. The letter « -r » of « pêcher » (to fish) is not pronounced in SF. It is spelled *pouâcher* in the corpus, however, [r] does not appear in Morvan in ALB942: pêcher. The letter « -r » of « plancher » (floor) is also not pronounced in SF. It is a spelled *plincher* and [r] is not pronounced in ALB1397: plancher. In Morvan, as in SF, *-r* can be voiceless in infinitives and nouns with *-er*. This clearly indicates the influence of SF.

Perhaps the author wrote the unpronounced word-final consonant letters as they are in SF so that readers would immediately know what the word was. Because the pronunciation of word-final consonants is strongly influenced by SF, the treatment of *-r* is generally considered to be the same as in SF.

6.2. Geolinguistic analysis

6.2.1. Coincident ratio

We compared the pronunciations inferred from the spellings with the pronunciations at the 18 locations. Our analysis of Story 6 is presented in Table 3. The leftmost column contains the 261 words. Of the 261 words, 24 appear in Story 6. The second column on the left represents the form of the story. A hyphen indicates that a word does not appear. The third column lists all pronunciations that can be inferred from these forms. All vowels are indicated as short; however, they may also be long. The right side of the table shows the pronunciations of the 18 ALB locations. « Vider » (to empty) is written as *queurer* in the corpus. From the spelling, pronunciations [køre], [kœre], [kørɛ], or [kœrɛ] can be inferred. At L27, « vider » is pronounced [køre]. Since it matches one of the guessed pronunciations, it counts as "1". Because the inferred pronunciations do not match L30, L93, or L102, these locations are counted as "0." Next, the coincident ratio for Story 6 was determined.

Table 4 shows the coincident ratio for all stories.

Table 3　The coincident ratio of Story 6

24/ 261 words	Corpus		ALB					
	Forms	Inferred pronunciations	Maps no.	L27	L30	…	L93	L102
« aider »	-	-	ALB1785	[ɛːdʲje]	[ɛːdʲje]		[ɛde]	[eːde]
				0	0		0	0
…	…	…	…	…	…	…	…	…
« vider »	*queurer*	[køre][kørɛ] [kœre][kœrɛ]	ALB730	[køre]	[vide]		[vide]	[vide]
				1	0		0	0
« cercueil »	*cerqueu*	[sɛrkø][serkø] [sɛrkœ][serkœ]	ALB1701	[sarkœj]	[sɛrkœj]		[sɛrkœj]	[bjer]
				0	0		0	0
Match score				18	19		9	9
Coincident ratio (= match score / 24 words)				0.75	0.79		0.38	0.38

Table 4　The coincident ratio of the 39 stories

Coincident ratio	ALB				
	27	30	…	93	102
Story 1	0.60	0.66		0.21	0.21
Story 2	0.55	0.63		0.39	0.42
…	…	…		…	…
Story 38	0.68	0.74		0.35	0.32
Story 39	0.60	0.60		0.32	0.31

6.2.2. Cluster analysis

Ward's square Euclidean distance method and the coincident ratio of the 18 locations[9] were used for the cluster analysis (see Table 4). Based on the resulting dendrogram, the 39 stories were classified into four clusters (see Figure 13). Cluster 1 consists of Story 34 only; Cluster 2 consists of stories 1, 10, 12, 14, 15, 18, 20, 21, 22, 24, 25, 27, 28, 30, 31, 33, 35, 36, 37, and 39; Cluster 3 consists of 2, 3, 5, 6, 7, 17, 23, and 26; and Cluster 4 consists of 4, 8, 9, 11, 13, 16, 19, 29, 32, and 38.

[9] We used R that is a language and environment for statistical computing and graphics similar to S. http://cran.r-project.org/

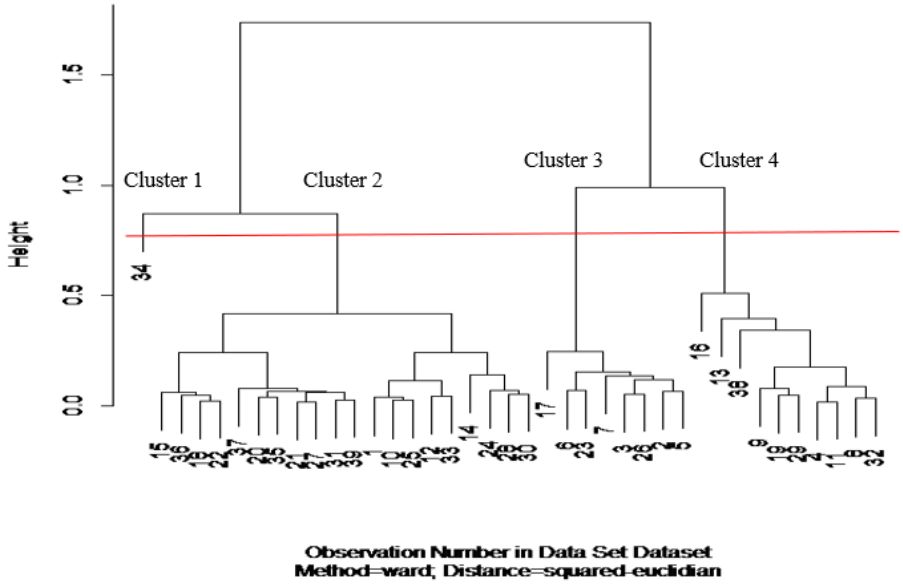


Fig. 13 Dendrogram

The average matching rate was then calculated (Table 7). This was determined by dividing the total matching scores by the sample size (see Tables 5 and 6). The total match score was the sum of the match scores for the stories in each cluster.

Table 5 The total match score

Locations	Cluster 1	Cluster 2	Cluster 3	Cluster 4
L27	10	576	160	211
L30	10	657	187	231
L52	6	490	132	170
L54	10	463	121	172
L62	5	347	115	130
L65	11	486	147	186
L70	3	313	111	101
L75	5	393	134	134
L77	3	277	94	87
L79	5	383	122	148
L80	6	483	150	170
L83	5	393	136	136
L85	1	396	128	131
L86	2	339	115	104
L89	2	390	127	137
L92	2	331	106	96
L93	2	302	104	90
L102	3	297	102	88

The sample size was the total number of words analyzed in the stories.

Table 6 The sample size

	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Sample size	15	974	262	297

Table 7 The average matching rate

Locations	Cluster 1	Cluster 2	Cluster 3	Cluster 4
L27	67%	59%	61%	71%
L30	67%	67%	71%	77%
L52	40%	50%	50%	57%
L54	67%	48%	46%	58%
L62	33%	36%	44%	43%
L65	73%	50%	56%	62%
L70	20%	32%	42%	34%
L75	33%	40%	51%	45%
L77	20%	28%	36%	29%
L79	33%	39%	47%	49%
L80	40%	50%	57%	57%
L83	33%	40%	52%	45%
L85	7%	41%	49%	44%
L86	13%	35%	44%	35%
L89	13%	40%	48%	46%
L92	13%	34%	40%	32%
L93	13%	31%	40%	30%
L102	20%	30%	39%	29%

6.2.3. Mapping

The average matching rate was placed on a map using the cartographic tool—Shiny Dialect.[10] The numbers in the figure represent the 18 locations.

[10] It was developed in the framework of ECLATS (*Extraction de contenus géolinguistiques d'atlas et analyse spatiale*), a project supported by ANR (*Agence Nationale de la Recherche*).
http://lig-tdcge.imag.fr/shiny/ShinyDialectV1_2_10/

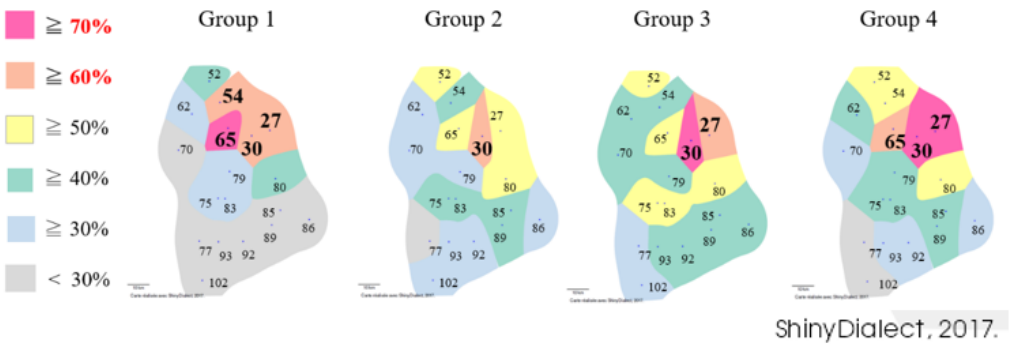


Fig. 14 Mapped average matching rate

Figure 14 shows that Cluster 1 has an average matching rate of at least 70% with L65 and at least 60% with L27, L30, and L54. Cluster 2 has at least 60% with L30. In Cluster 3, L30 and L27 had a match of over 70%, and over 60%, respectively. In Cluster 4, L27 and L30 match by more than 70%, and L65 by more than 60%. None of the clusters have a high average matching rate with the southwestern Morvan.

6.3. Discussion

Clusters 2, 3, and 4 are most consistent with L30. In contrast, Cluster 1 was the most consistent with L65, followed by L27, L30, and L54. Story 34, the only story comprising Cluster 1, has the lowest number of tokens and words analyzed (see Figure 4). While this may have influenced the analysis, Stories 13 and 16—which have similarly low numbers of tokens and analyzed words—are classified in Cluster 4. After all, Story 34 could be unique.

Regardless, the four clusters are close to the spoken language at around L30, L27, and L65. Figure 15 shows that L30 is the closest to Saulieu, with a linear distance of approximately 4 km. L27 and L65 are approximately 9 and 11 km from Saulieu, respectively.

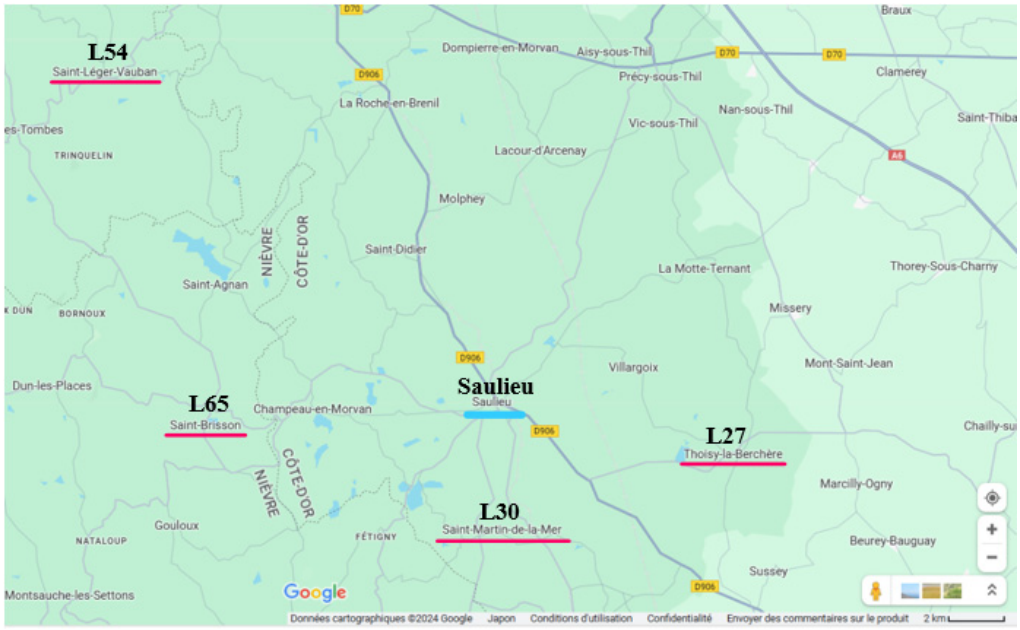


Fig. 15 Saulieu and nearby locations

7. Conclusion

In answer to the research question, "Which regional spoken language of Morvan were the 39 stories in *L'âme du Morvan* (1923 edition) written in?" we can say that they were written in the spoken language of the Saulieu region. Spelling analysis suggests that the 39 stories may have been written slightly differently; however, they were written in the dialect of the Saulieu region—the Bas-Morvan dialect. This confirms the statement in the preface to the 1923 edition.

By incorporating the relationship between the spelling and pronunciation of standard French, the author wrote forms that readers could easily associate with the words in standard French. He also introduced his own spelling to represent pronunciation. On the other hand, he did not insist on faithfully reproducing the pronunciation of word-final consonants but wrote down the unpronounced word-final consonant according to standard French spelling.

A partial summary of the preface to the 1923 edition states:

> Regions of France are losing their originality and individuality. This is inevitable, even in Morvan, which has long remained faithful to its traditions without accepting anything new. We have dreamed of making a small contribution to saving the language of our ancestors from total

extinction or oblivion, in the hope of preserving the memory of this dying life.

This study has shown that the spoken language of the author's hometown has been preserved, mainly from a phonetic point of view, and that his goal has been achieved with the publication of this book.

In the future, we will analyze whether each story's spelling is unique. We will pay attention to Story 34, as we found unique phonetic features in this study, and explore whether Story 34's spelling is also distinctive.

Abbreviations

ALB	*Atlas linguistique et ethnographique de la Bourgogne*
FEW	von Wartburg, Walther (1922-2002) *Französisches etymologisches Wörterbuch. Eine Darstellung des galloromanischen Sprachschatzes*
SF	standard French

References

Bertrant, Paule (1979) *Les parlers du Morvan III –Transcription des formes en orthographe française–*. Château-Chinon: Académie du Morvan.

Catach, Nina (1980) *L'orthographe française : traité théorique et pratique avec des travaux d'application et leurs corrigés*. Paris: F. Nathan.

Chagnaud, Clément, Philippe Garat, Paule-Annick Davoine, Elisabetta Carpitelli and Axel Vincent (2017) ShinyDialect: a cartographic tool for spatial interpolation of geolinguistic data, GeoHumanities'17, November 7–10, 2017, Los Angeles Area, CA, USA. doi: https://doi.org/10.1145/3149858.3149864

de Chambure, Eugène (1878) *Glossaire du Morvan; étude sur le langage de cette contrée comparé avec les principaux dialectes ou patois de la France, de la Belgique wallonne, et de la Suisse romande*. Paris: H. CHAMPION, LIBRAIRE, Autun: Dejussieu Père et Fils. https://archive.org/details/glossairedumorva00chamuoft

Dauzat, Albert (1944) *La géographie linguistique*. Paris: Libraire ERNEST.

Guillaume, Alfred (1923) *L'âme du Morvan*. Autun: Maison Vve G. Gervais. https://gallica.bnf.fr/ark:/12148/bpt6k98015409/f1.item.r=L'%C3%A2me%20du%20Morvan

Guillaume, Alfred (1971) *L'âme du Morvan*. Avallon: Air. Graphic.

Ito, Reiko [伊藤玲子] (2020) Nizisseiki tyuuki no burugoonyu tihoo moruvan no onsei tokutyoo to onsei bunpu no bunseki – hutatu no gengotizu wo motiite –『20世紀中期のブルゴーニュ地方モルヴァンの音声特徴と音声分布の分析－2つの言語地図を用いて－』[Analysis of the pronunciation distribution and pronunciation features of Morvan in Burgundy in the Mid-20th Century – using data from two linguistic atlases –]. *Flambeau* 45: 87–105. https://doi.org/10.15026/94460

Ito, Reiko [伊藤玲子] (2021) Burugoonyu tihoo moruvan ni okeru CA-, -CCA, PL-, BL- wo gogen ni motu go nituite『ブルゴーニュ地方モルヴァンにおけるCA-, -CCA, PL-, BL-を語源に持つ語について』[Words having etymon C(A)-, -CC(A), PL-, or BL- in Morvan in Burgundy]. *Studia romanica* 54: 57–66. http://sjsrom.ec-net.jp/studrom/054/ito.pdf

Ito, Reiko [伊藤玲子] (2022) Analyse géolinguistique de la prononciation des pronoms personnels sujets et du verbe « être » au présent de l'indicatif en Bourgogne. *Studia romanica* 55: 45–54. http://sjsrom.ec-net.jp/06.pdf

Régnier, Claude (1979a) *Les Parlers du Morvan I*. Château-Chinon: Académie du Morvan.

Régnier, Claude (1979b) *Les Parlers du Morvan II*. Château-Chinon: Académie du Morvan.

Richard, Jean (1978) *Histoire de la Bourgogne*. Toulouse: Privat.

Straka, Georges and Pierre Gardette (1973) *Les Dialectes romans de France à la lumière des atlas régionaux: Strasbourg, 24–28 mai 1971*. Paris: Éditions du Centre national de la recherche scientifique.

Taverdet, Gérard (1973) Patois et français régional en Bourgogne, *Ethnologie français* 3(3/4): 317–328. https://www.jstor.org/stable/40988262

Taverdet, Gérard (1975) *Atlas linguistique et ethnographique de la Bourgogne I*. Paris: Éditions du Centre national de la recherche scientifique.

Taverdet, Gérard (1977) *Atlas linguistique et ethnographique de la Bourgogne II*. Paris: Éditions du Centre national de la recherche scientifique.

Taverdet, Gérard (1980) *Atlas linguistique et ethnographique de la Bourgogne III*. Paris: Éditions du Centre national de la recherche scientifique.

Taverdet, Gérard (2001) Un écrivain patoisant bourguignon du XXe siècle, Alfred Guillaume, Littératures dialectales de la France : diversité linguistique et convergence des destins. *Bibliothèque de l'École des chartes* 159, t. 1: 209–226. Genève: Librairie Droz Stable. https://www.jstor.org/stable/42957867

Taverdet, Gérard and Danièle Navette-Taverdet (1990) *Dictionnaire du français régional de Bourgogne*. Paris: Éditions.

von Wartburg, Walther (1922-2002) *Französisches etymologisches Wörterbuch. Eine Darstellung des galloromanischen Sprachschatzes* (Vols. 1–25). Bonn: Klopp; Leipzig: Teubner; Basel: Helbing und Lichtenhahn; Bonn/Basel: Zbinden. https://apps.atilf.fr/lecteurFEW/

Publication history

Date received: 2 July 2024
Date accepted: 14 August 2024

Research article

The phenomenon of neg1 omission/retention
− A regional analysis using the *Atlas linguistique de la France* −

SEIMIYA, Takamasa
Doctoral Program, Tokyo University of Foreign Studies

Abstract: In this study, we analyzed the tendencies in the omission and retention of pre-verbal negative adverb (neg1) using the *Atlas linguistique de la France*. The analysis focused on 631 points and 11 maps, conducting statistical analyses from two perspectives: regionality (five linguistic regions) and map type (=morphosyntactic perspective). In the French-speaking regions, neg1 tended to be retained in peripheral areas, while variability in the omission/retention of neg1 was observed around Paris. In the Occitan-speaking regions, except for Gascony, there was a strong tendency for neg1 to be omitted. In the Francoprovençal-speaking regions, neg1 tended to be omitted in Aosta, Italy, and Valais, Switzerland, but it was generally retained in other areas. In the Catalan-speaking regions, neg1 was typically omitted, while in the Ligurian-speaking region, neg1 was retained. However, no statistically significant differences were found between linguistic regions and the omission/retention of neg1. On the other hand, statistically significant differences were observed regarding map type. It was found that neg1 was more likely to be retained when the subject was a determiner phrase (DP) or a proper noun, and more likely to be omitted in constructions like *moi je* [disjunctive form + subject pronoun]. Moreover, in modern spoken French, *ne* tends to be retained when the subject is *nous* or *vous*, whereas this tendency was not observed in negative sentences where the map names included *nous* or *vous*. This discrepancy can be attributed to the fact that the personal pronouns used in the responses obtained from the dialect surveys do not necessarily match those used in the map names. These findings suggest that the morphosyntax of negative sentences has a greater influence on the omission/retention of neg1 than regionality.

Keywords: Atlas linguistique de la France; negation; dialectology; morphosyntax; regional study

1. Introduction

In modern French, negative sentences are expressed by combining the negative particle *ne* with a negative adverb as in (1)b. The negative adverb is also known as *forclusif*, with the most common being *pas*. However, depending on the dialect, different words may be used as negative adverbs. According to Seimiya (2021), who conducted a dialectometrical classification of dialectal points using maps of negative sentences from the *Atlas linguistique de la France* (*ALF*), in addition to the standard

SEIMIYA, Takamasa. 2024. The phenomenon of neg1 omission/retention — A regional analysis using the *Atlas linguistique de la France* —. *Studies in Geolinguistics* 4: 59–94. doi: https://doi.org/10.5281/zenodo.13948562

pas, negative adverbs like *mie* in the east of France, *point* in the north and west of France, *cap* and *brique* in some parts of the south of France, and *nin* in Belgium are used. There are also variations in the meaning of negation depending on the negative adverb, such as *ne...plus* “no longer” and *ne...jamais* “never.”

In spoken language, especially in informal conversations, *ne* is often omitted as in (1)b'. Many sociolinguistic studies, including Ashby (1976, 1981), Coveney (2002), and morphosyntactic studies like Pollock (1989) and Rowlett (1998), have been conducted on the omission of *ne*. From a sociolinguistic perspective, social backgrounds such as age, gender, and place of origin influence the omission of *ne* (Ashby 1976, 1981, 2001; Armstrong and Smith 2002; Coveney 2002). From a morphosyntactic perspective, the type of subject and the content to the right and left of *ne* affect its omission (Ashby 1981; Hansen and Malderez 2004; Meisner and Pomino 2014; Meisner, Tissot and Stark 2015; Moreau 1986).

(1) a. Je chante.
I sing.IND.PRS.1SG
“I sing”

b. Je ne chante pas.
I NEG1 sing.IND.PRS.1SG NEG2
“I do not sing.”

b'. Je chante pas
I sing.IND.PRS.1SG NEG2
“I do not sing.”

In this study, we analyze the omission of negative particle in *ALF*. Section 5.1 briefly outlines the atlas employed in this research, highlighting that the gender and age of the respondents vary depending on the dialectal points. While sociolinguistic analysis is feasible from the perspective of language standardization processes, other forms of analysis may encounter difficulties due to the inherent variability of the data. Nevertheless, because the dialect surveys utilize consistent survey items, it is possible to conduct a geolinguistic analysis from a morphosyntactic perspective. It should be noted that dialect forms do not necessarily use the same word as the standard French *ne*. Therefore, words derived from Latin *non* are referred to as neg1 for the purpose of analysis. However, when referring to standard French or spoken French, *ne* will be used.

2. Features of the omission of *ne* in spoken French from a morphosyntactic perspective

Regarding the relationship between the omission of *ne* and morphosyntax, Meisner and Pomino (2014) provide a detailed analysis based on prior research. They argue that the subject used in a sentence and its phonological form are related to the omission or retention of *ne*. Subjects that tend to facilitate the omission of *ne* include first-person singular, second-person singular, third-person singular and plural pronouns (*je*, *tu*, *il* and *ils*), indefinite pronouns *on*, and cases where a strong form is simultaneously used like *moi je* (Meisner, Robert-Tissot & Stark 2015). On the other hand, *ne* is more likely to be retained with first-person plural and second-person plural pronouns (*nous* and *vous*), determiner phrases (DP), and proper nouns.

Additionally, they note that even with the same pronoun, the likelihood of *ne* omission varies depending on its pronunciation. For instance, they analyze the relationship between the first-person singular subject pronoun *je* and its four phonetic variants in relation to *ne* omission. When *je* is pronounced as [ʒ], [ʃ], or is omitted (∅), *ne* is completely omitted. However, when pronounced as [ʒə], the degree of omission varies depending on the its left and right elements. Their table summarizes the syntactic structures related to *ne* omission and retention. Negative sentences are more likely to omit *ne* if they correspond to the structures under OMISSION, and more likely to retain *ne* if they correspond to the syntactic structures listed under RETENTION.

Table 1 Morphosyntactic Features of *ne* Omission/Retention[1]

	OMISSION	RETENTION
Ashby (1981)	Lexicalization Dependent clause Indicative mood	Non-lexicalization Independent clause Subjunctive mood
Hansen & Malderez (2004) Moreau (1986)	Auxiliary verb Compound tense form	Lexical verb
Armstrong & Smith (2002) Ashby (1981)	Other proclitics	No other proclitics

However, there are some ambiguities concerning the relation between the pronunciation of the subject and the retention of *ne*. For instance, while they state that *ne* is more likely to be retained when *je* is pronounced as [ʒə], it is unclear whether this pronunciation influences the retention of *ne* or if *ne* influences its pronunciation as [ʒə].

[1] Adapted from Meinser & Pomino (2014:25).

Therefore, in this study, we consider the type of subject and the syntactic structure of negative sentences but do not take into account the pronunciation of the subject.

3. Previous studies on negative sentences using *ALF*

There are many previous studies using maps of negative sentence from *ALF*, most of which focus on the types of forclusif. Notable among these are Burnett (2019), Dagnac (2015), and Schwegler (1986).

Burnett (2019) analyzes the types and tendencies of forclusifs used in 22 negative sentences at 150 locations within *ALF*, targeting a total of word 2,989 forms. The survey covers points 1 to 199, including northeastern France, Swiss Romandy, and parts of the Walloon region. In the studied areas, *pas* was predominantly used in the south, *point* and its variant *pont* in the southeast, *mie* in the northwest, and *nen* in the Walloon region (Burnett 2019: 194). Among these forclusifs, *pas* was the most dominant, followed by *mie*, *point*/*pont*, and *nen*.

Dagnac (2015) analyzes the forclusifs used in Picard dialect using Picartext, *ALF*, and *Atlas linguistique et ethnographique picard*. Some *ALF* maps beyond Picardy were also analyzed. While it is unclear whether all *ALF* negative sentence maps were analyzed, *point* was used in Picardy, its variant *pon* in Artois and much of the Pas-de-Calais department, *nin* in parts of the Nord department, and *pas* and its variant *po* in areas where these three regions intersect.

Schwegler (1986), in his paper on the etymology of *cabdorn*, analyzes Map 89 − *dans ce pays, il n'y a pas de sources* "In this country, there is no source" − in limited regions. He plots on a map the dialectal points in southwestern France where *cap* is used as a forclusif. *cap* was frequently used as a forclusif in the departments of Ariège, Pyrénées-Orientales, and Aude.

In contrast, few studies focus on the omission of neg1. Jagueneau (2007) uses 10 *ALF* negative sentence maps as part of his study on the ratio of neg1 omission/retention in the Limousin dialect. He analyzes 1,040 forms from 104 points centered on the Limousin dialect, stating that in Limousin, neg1 is frequently retained, while in surrounding areas, neg1 tends to be omitted.

Moreover, comprehensive studies on *ALF* as a whole are scarce. Seimiya (2021), in his dialectometrical study, uses 9 maps of negative sentences from *ALF* to explore what kind of matrices should be used to reflect the usage and tendencies of both neg1 omission/retention and types of forclusifs used. Interpretation maps for each map are included in the appendix, providing insights into the regionality of neg1 omission/retention and the forclusifs used.

4. Research questions

There are numerous sociolinguistic and morphosyntactic studies on the omission/retention of *ne* in modern spoken French. By comparing the results of various regional spoken language studies, it is clear that there is a certain degree of regionality in the tendencies of *ne* omission/retention. However, due to the heterogeneous quality of the corpora, a complete comparison seems impossible.

Therefore, by using the negative sentence maps in the *ALF*, it is possible to homogeneously analyze the tendencies of *ne* or neg1 omission/retention throughout France. Based on this, this study set the following two research questions:

1. Are there regional tendencies in the neg1 omission/retention?
2. Based on the results of the first question, are there any maps (= negative sentences) where the neg1 is more likely to be omitted/retained?

5. Methods

This chapter describes the analysis targets and methods.

5.1. Targets

5.1.1. Atlas linguistique de la France

This study uses the *ALF* by Jules Gilliéron and Edmond Edmont for analyzing the neg1 omission/retention in negative sentences. *ALF* is dialect data from the early 20th century, collected between 1897 and 1901. The dialect survey was conducted by Edmont alone, ensuring high reliability when comparing the collected data (Kawaguchi et al. 2021). It includes not only France but also the Franco-Romance regions of Belgium, Switzerland, and Italy, as well as the British Crown Dependencies of the Channel Islands, with a total of 638 points[2]. *ALF* covers five major linguistic regions: Langue d'oïl, Langue d'oc, Francoprovençal, Catalan, and Ligurian. The age range of the subjects is broad, from teenagers to those in their 80s, with the most common being in their 50s, followed by those in their 40s and 60s.

[2] Although it is often considered as 639 points because two individuals were surveyed at point 284, this study regards it as 638 points.

5.1.2. Maps

There are 19 maps of negative sentences[3] in *ALF* (Burnett 2019). These include not only *ne...pas* "not" but also *ne...plus* "no longer" and *ne...ni* "neither...nor." Among them, this study analyzes the neg1 omission/retention in *ne...pas*. However, maps were excluded if they met the following two conditions:

Condition 1: Maps surveyed only in part of *ALF*
Condition 2: Negation of infinitive verb

Condition 1 was set to conduct quantitative classification in this study. For example, Map 1650 − *je n'ai pas osé le lui dire* "I didn't dare tell him" − was surveyed in only half of the *ALF* points. Similarly, Map 10 − *n'aie pas peur* "don't be afraid" − was not surveyed in Nord, Somme, Oise, Aisne, Ardennes, or Marne. Including these maps in the analysis could affect the final statistical analysis due to varying numbers of surveyed maps per dialectal point.

Condition 2 is because this study focuses on negative sentences of finite verb. Therefore, Maps 898-1 − *pour ne pas nous* "so that we don't" − and 898-2 − *pour ne pas* "so as not to" − were excluded. Furthermore, Map 896-2 − *ne...pas* "not" − was excluded because infinitives are often used depending on the dialectal point. The 11 maps ultimately chosen to be analyzed in this study are listed below, along with the full sentences where only part of the sentence is mapped[4].

Map 12AB − moi je ne les aide pas

Ils	feront		ce qu'	ils	voudront ;	
they	do.IND.FUT.3PL		REL.ACC.	they	want.IND.FUT.3PL	
moi	je	ne	les	aide	pas.	
me	I	NEG1	them	help.IND.PRS.1SG	NEG2	

"They will do what they want; I do not help them."

Map 89 − il n'y a pas de

Dans	ce	pays,			
in	this	country,			
il	n'	y	a	pas	de sources.
PRONEXPL	NEG1	there	have.IND.PRS.3SG	NEG2	source

"In this country, there is no source."

[3] There are 19 maps, but a single map may contain multiple negative sentences, resulting in a total of 22 types of negative sentences being surveyed in the dialects.
[4] For instance, *Ils feront ce qu'ils voudront; moi je ne les aide pas* is divided into four maps: Map 532 – ils feront, Map 205 – ce qu'ils, Map 1418 – voudront, and Map 12AB – moi je ne les aide pas.

Map 806AB − si nous ne mangeons pas nos prunes

si	nous	ne	mangeons	pas	nos	prunes,
if	we	NEG1	eat.IND.PRS.1PL	NEG2	our	plums
elles	se moisiront		bientôt			
they	go mouldy.IND.FUT.3PL		soon			

“If we do not eat our plums, they will soon rot”

Map 817AB − pourquoi ne vous mariez-vous pas ?

pourquoi	ne	vous mariez-vous	pas ?
why	NEG1	get married.IND.PRS.2PL	NEG2

“Why don’t you get married?”

Map 896-1 − ne…pas

le	roseau	plie	mais	ne	rompt	pas.
the	reed	bend.IND.PRS.3SG	but	NEG1	break.IND.PRS.3SG	NEG2

“The reed bends but does not break.”

Map 897 − qu’ils ne…pas

J’	ai cru	qu’	ils	ne	viendraient	pas.
I	believe.IND.PST.1SG	that	they	NEG1	come.COND.PRS.3PL	NEG2

“I thought that they would not come.”

Map 899 − n’est pas encore

le	blé	est	mûr,				
the	wheat	be.IND.PRS.3SG	ripe				
mais	l’	avoine	n’	est	pas	encore	mûre.
but	the	oat	NEG1	be.IND.PRS.3SG	NEG2	still	ripe

“The wheat is ripe, but the oats are not yet ripe.”

Map 1082 − je ne peux pas

je	ne	peux	pas	perdre,
I	NEG1	can.IND.PRS.1SG	NEG2	lose.INF
ça	c’	est	sûr.	
it	that	be.IND.PRS.3SG	sure	

“I can not lose, that is for sure.”

Map 1083 − on ne peut pas

Par	ce	temps,	on	ne	peut	pas	dormir.
by	this	weather	one	NEG1	can.IND.PRS.3SG	NEG2	sleep.INF

“In this weather, one cannot sleep.”

Map 1352 − mais il ne vaut pas

Celui-ci,	il	est	bon,		
this one	it	be.IND.PRS.3SG	good		
mais	il	ne	vaut	pas	le mien.
but	it	NEG1	be worth.IND.PRS.3SG	NEG2	mine

“This one is good, but it is not as good as mine.”

Map 1409AB − tu ne vois donc pas

Tu	ne	vois	donc	pas	que
you	NEG1	see.IND.PRS.2SG	so	NEG2	that
tu	es	aussi	vieux	que	moi.
you	be.IND.PRS.2SG	also	old	that	me

“Don’t you see that you are as old as I am.”

5.1.3. Points

After selecting the maps for analysis, the word forms of each point at each map were entered into Excel. Of the 638 points in *ALF*, 7 points (261, 281, 282, 291, 396, 874, 989) had missing word forms in at least one map. To facilitate statistical analysis, these 7 points were excluded, resulting in 631 points analyzed in this study. ShinyDialect was used to create the interpretation maps of this study.

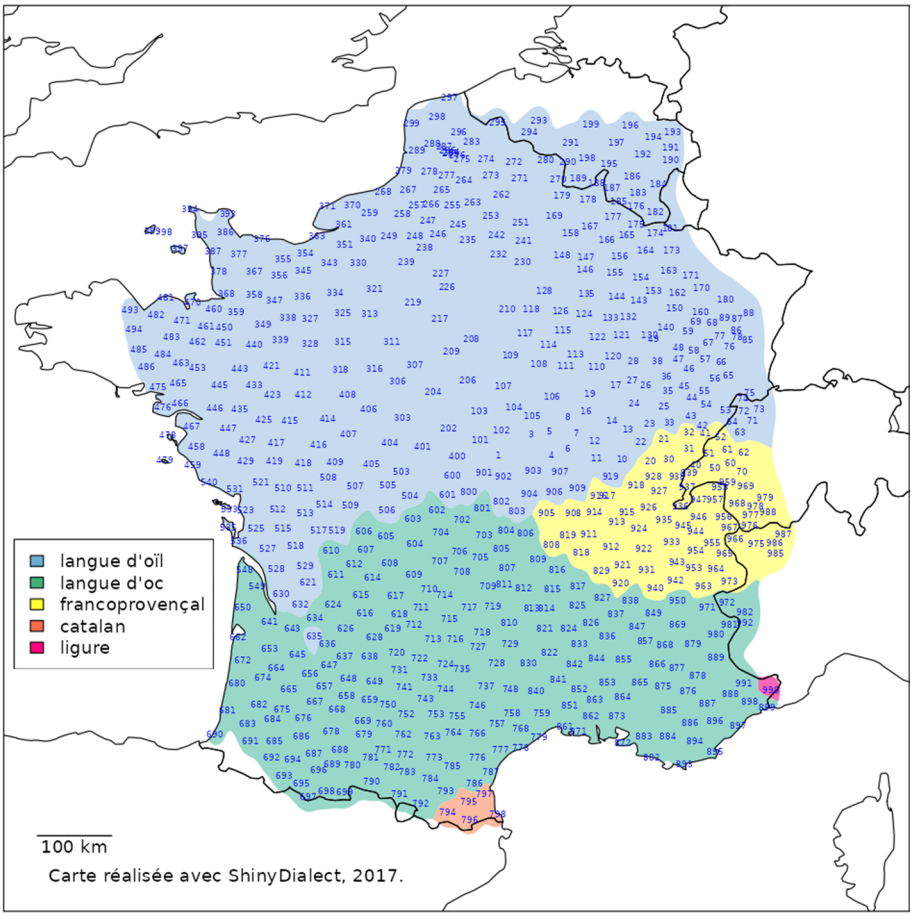


Fig. 1 Distribution of 631 points and language boundary[5]

5.2. Methods

5.2.1. Digitization of map data[6]

The word forms corresponding to standard French *ne* and *pas* on the selected 11 maps were entered into Excel. There were no major issues with the word forms corresponding to the forclusif, *pas*. However, some word forms that were presumed to correspond to *ne* were difficult to judge as negative adverb or other elements at first glance.

For example, in Map 89, the word form '*dy*' in '*dy o pa de*' at point 614 and the word form '*d*' in '*d a pa de*' at point 969, and the word form '*d*' in '*k i d ... ñ*' at point 291

[5] The language boundary lines adopted are those from the SYMILA project at Université Toulouse-II-Jean-Jaurès and Goebl (2018).

[6] The actual word form in alphabet Rousselot-Gilliéron is written in italic within ''.

in Map 897 were challenging. Only when it was difficult to judge based on the target map alone were other maps referred to as needed.

The word forms '*dy*' at point 614 and '*d*' at point 969 in Map 89 were judged not to be negative adverbs based on word forms of *il y a* (there is/are) being confirmed in Maps 103 − *il y a eu* "there was/were" − and 729 − *il y a huit jours* "a week ago" −. On the other hand, the word form '*d*' at point 291 in Map 897 could not be considered anything other than *ne* based on other maps. '*d*' in *ALF* phonetically represents a voiced alveolar plosive [d], similar to the alveolar nasal [n], except for the characteristic that the air does not pass through the nasal cavity. Thus, this '*d*' at this point was judged to be *ne* not pronounced nasally for some reason. Additionally, '*j én*' at point 261 in Map 12 was judged to be a miswritten '*jé n*' based on the first-person singular form '*jé*' in Maps 1082 − *je ne peux pas* "I cannot" − and 1084 − *je…pouvais* "I could" −, and '*én*' was treated as *ne*.

If a negative sentence was constructed using only neg1 without forclusif, it was treated as a point retaining neg1.

5.2.2. Analysis of regionality of neg1 retention/omission

The number of maps on which neg1 was retained out of the 11 maps was counted for each point. Points where neg1 was never retained were labeled as n0, and those where neg1 was retained on all maps were labeled as n11, creating a 12-stage labeling system based on the number of times neg1 was retained. An interpretation map was then created for the neg 1 omission/retention in the 11 maps, analyzing the regionality of neg 1 omission/retention. To analyze whether the tendency of neg1 omission/retention varies by language region, a two-way ANOVA was conducted. The statistical software R4.3.1 was used for the analysis.

5.2.3. Analysis of maps of neg1 retention/omission

After clarifying the regionality of neg1 omission/retention, statistical analysis was conducted on points excluding n0 and n11. By excluding n0, where neg1 was always omitted, and n11, where neg1 was always retained, it was judged that it would be possible to identify maps that were more prone to omitting neg1 or vice versa. To analyze whether the tendency of neg1 omission/retention varies by map (=type of negative sentence), Wilcoxon signed-rank tests were conducted. The same statistical software was used for the analysis.

6. Results

First, in section 6.1, the distribution of neg1 omission/retention in the interpretation maps of the 11 analyzed maps is presented and briefly analyzed. Then, in sections 6.2 and 6.3, the tendencies of neg1 omission/retention are analyzed from two perspectives: regionality and the type of negative sentence.

6.1. Brief analysis of each map

The distribution of neg1 omission/retention in the 11 analyzed maps is presented. Interpretation maps are color-coded with blue for −neg1 points (=omission) and red for +neg1 points (retention).

6.1.1. Map 12

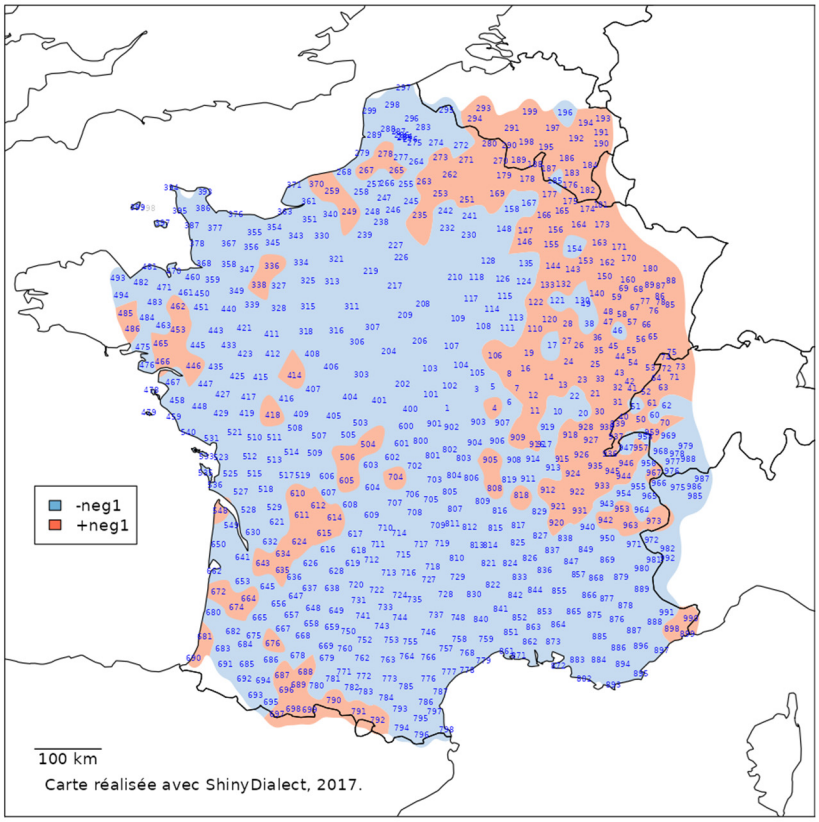


Fig. 2 neg1 omission/retention of Map 12

In Map 12, there are 426 −neg1 points (68%) and 205 +neg1 points (32%), with more points where neg1 is omitted. This map has the most −neg1 points among the analyzed maps. One notable feature of this map is that it is the only map where disjunctive pronoun, the strong form, of a first-person pronoun, is included in the map title. Even in areas where subject pronouns are omitted in other maps, the subject pronoun or the disjunctive form is used in this map.

According to Meisner et al. (2015), in modern spoken French, when the subject combines a disjunctive subject with a subject pronoun like *moi je*, *ne* is more likely to be omitted. Although this data is from about 100 years ago, this map also shows a prominent omission of neg1, indicating a similar tendency.

6.1.2. Map 89

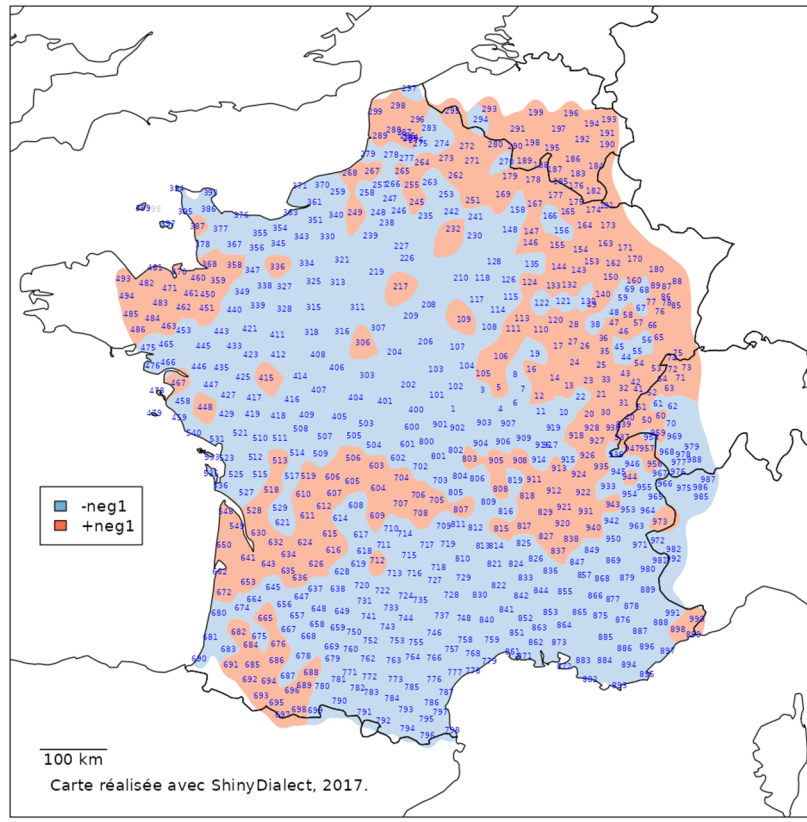


Fig. 3 neg1 omission/retention of Map 89

In this map, there are 377 −neg1 points (60%) and 254 +neg1 points (40%), with more points where neg1 is omitted. This map has the second-highest number of −neg1

points among the analyzed maps. In Ashby (1981), it is stated that in lexicalized expressions such as *il y a* "there is/are," *ne* is more likely to be omitted, showing a similar tendency to modern spoken French.

6.1.3. Map 806

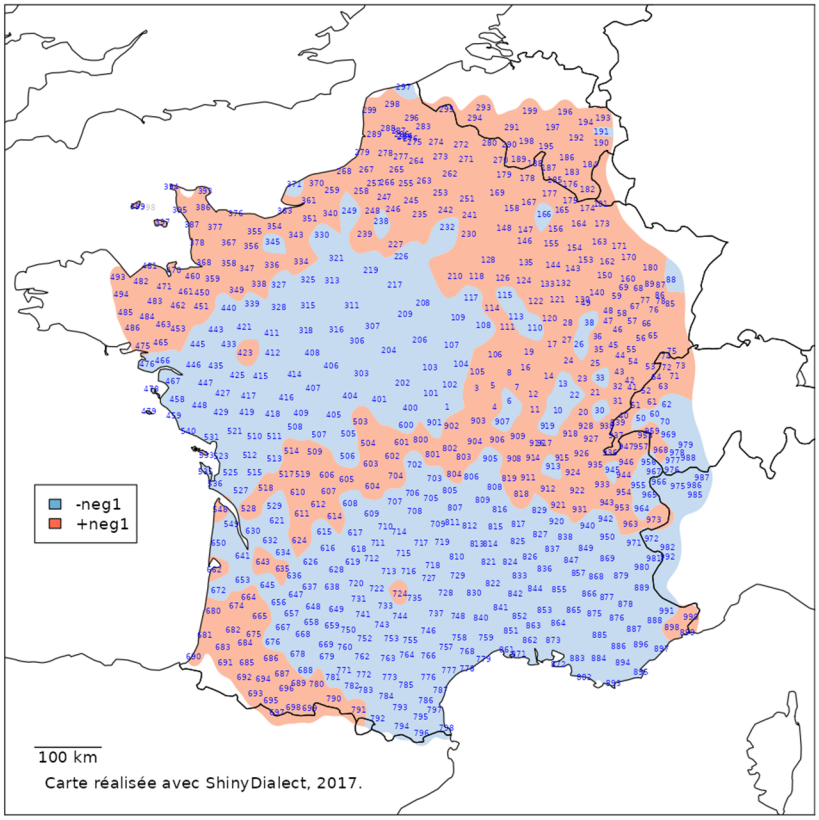


Fig. 4 neg1 omission/retention of Map 806

This is a map where the first-person plural pronoun *nous* is used. In modern spoken French, *nous* is generally used in formal settings where *ne* tends to be retained. However, in this map, there are 323 +neg1 points (51%) and 308 −neg1 points (49%), with a slight predominance of +neg1.

Looking closely at the subject pronouns, *nous* was used in only 137 points, about 20% of the total. The most common was pro-drop with 219 points, followed by *je* in 218 points. An indefinite pronoun *on* widely used instead of *nous* in modern spoken French, was used in 55 points. The contracted form *nous autres* was used in only 2 points.

Table 2 summarizes the subject types and neg1 omission/retention. In pro-drop points, neg1 is more likely to be omitted, while *on* and *nous* as subjects tend to retain neg1. When the subject is the first-person singular *je*, there is no significant difference, but there are more points retaining neg1.

Table N　Subject Type and neg1 omission/retention in Map 806

	je	*on*	*nous*	*nous autres*	∅	Total
−neg1	99	4	44	1	160	308
+neg1	119	51	93	1	59	323
Total	218	55	137	2	219	631

6.1.4. Map 817

Map 817 is the only map of negative interrogative, with the second-person plural pronoun as the subject. There are 341 +neg1 points (54%) and 290 −neg1 points (46%).

Question sentences with interrogatives can have two syntactic structures: [interrogative + V + S] and [interrogative + S + V]. Therefore, the map title *pourquoi ne vous mariez-vous pas*? can also be understood as *pourquoi vous ne vous mariez pas*? In the latter case, the subject pronoun *vous* moves to the left of *ne*, which is not the case in the former.

The table summarizes the syntactic structures and neg1 omission/retention in the 631 analyzed points. Regardless of the word order, neg1 is retained twice as much, and subject inversion does not seem to affect neg1 omission/retention. However, in points where the subject is omitted, neg1 is more likely to be omitted. Over 80% of the points with subject omission are in the Occitan-speaking regions.

Table 3　Syntactic Structure and neg1 omission/retention in Map 817

	int + S + V	int + V + S	int + V	total
−neg	71	70	149	290
+neg	153	141	47	341
total	224	221	196	631

6.1.5. Map 896

This map uses the neg1 in a coordinate clause introduced by the conjunction *mais* “but,” with the same subject, leading to subject omission in the coordinate clause. There are 436 +neg1 points (69%) and 195 −neg1 points (31%), making it the map with the second-highest number of +neg1 points. The distribution shows that in the French-speaking regions, almost all points retain neg1, while in the Occitan-speaking regions, neg1 is omitted in most points.

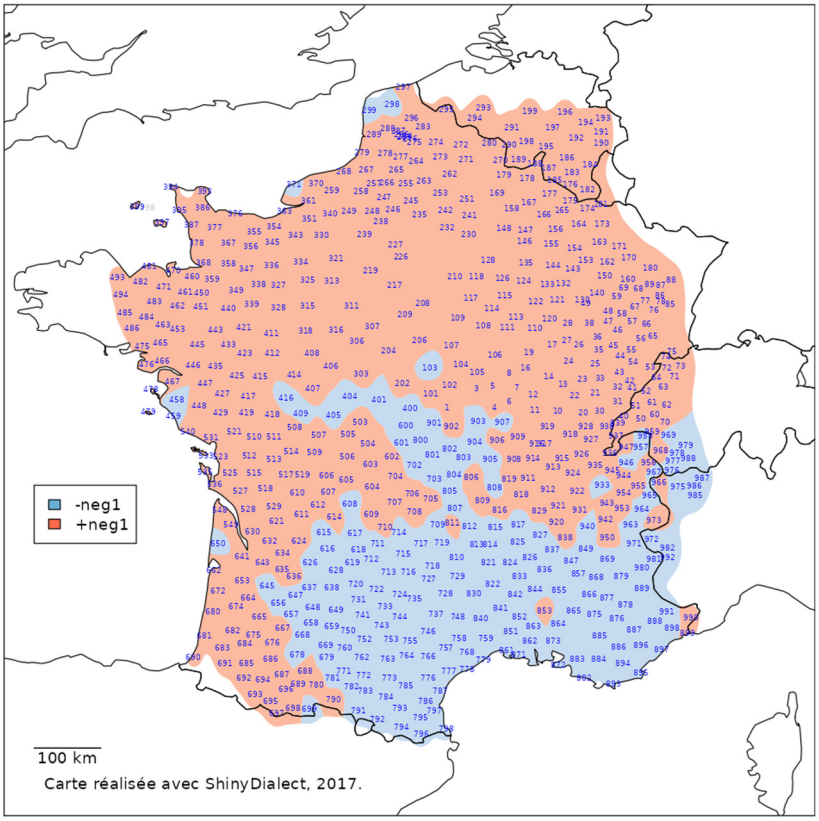


Fig. 6 neg1 omission/retention of Map 896

6.1.6. Map 897

Map 897 is the only map using the third-person plural pronoun. The entire sentence is *j'ai cru qu'ils ne viendraient pas*, with the negative sentence in the subordinate clause and the verb in the present conditional tense, which is significantly different from other maps. There are 416 +neg1 points (66%) and 315 −neg1 points (34%). In the French-speaking regions, neg1 tends to be retained overall, despite some areas around Poitou where neg1 is omitted. In contrast, in the Occitan-speaking regions, neg1 is omitted in almost all points.

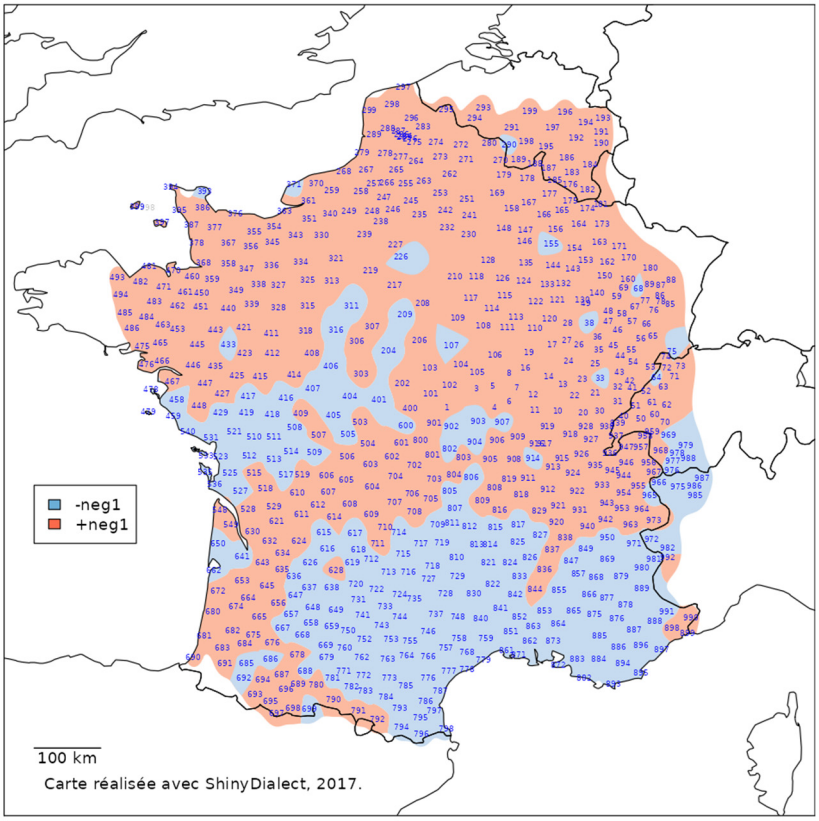


Fig. 7　neg1 omission/retention of Map 897

6.1.7. Map 899

This map includes a negative sentence in a coordinate clause introduced by the conjunction *mais* “but,” but with a different subject in the coordinate clause compared to Map 896. There are 491 +neg1 points (78%) and 140 −neg1 points (22%), making it the map with the highest number of +neg1 locations. Unlike other maps, even in the Occitan-speaking regions, which typically sees neg1 omission, especially in the east, neg1 is retained.

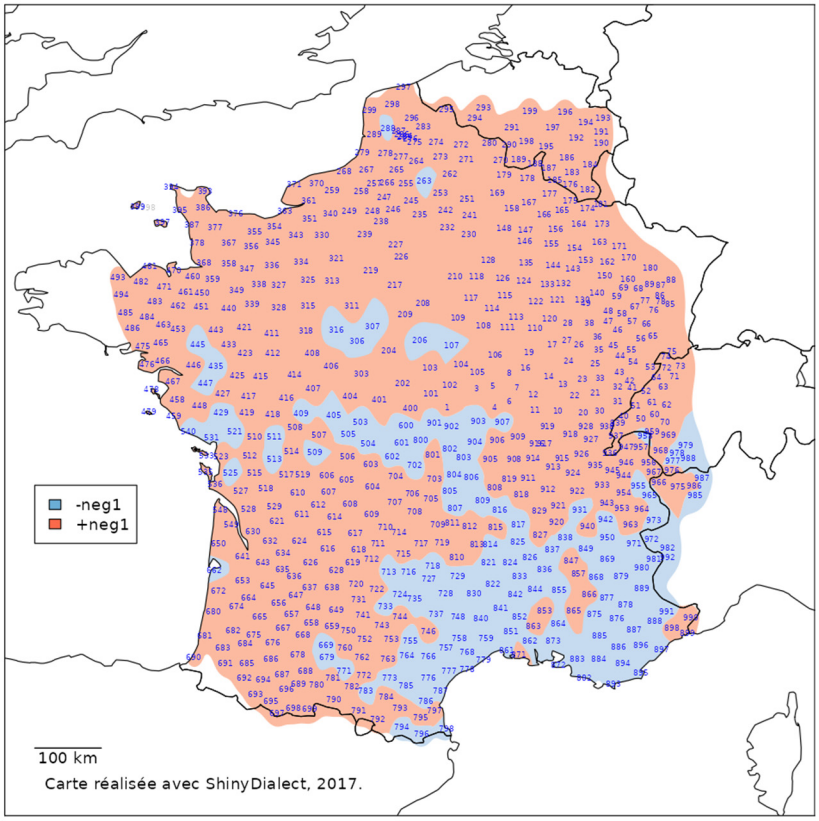


Fig. 8 neg1 omission/retention of Map 899

6.1.8. Map 1082

The negative sentence itself is *je ne peux pas perdre*, but the dialect survey was conducted with the phrase *je ne peux pas perdre, ça c'est sûr*. There are 361 +neg1 points (57%) and 270 −neg1 points (43%).

In modern spoken French, *ne* is generally more likely to be omitted with the first-person singular subject *je*, but this is not necessarily the case in this map. The fact that the verb is *pouvoir* "can" would also be a reason for retention of neg1. In written language, *pouvoir* can form negative sentences without neg2. Indeed, some points formed negative sentences using only neg1. Additionally, the complexity of the right elements, *ça c'est sûr*, may have created an environment where neg1 is more likely to be retained.

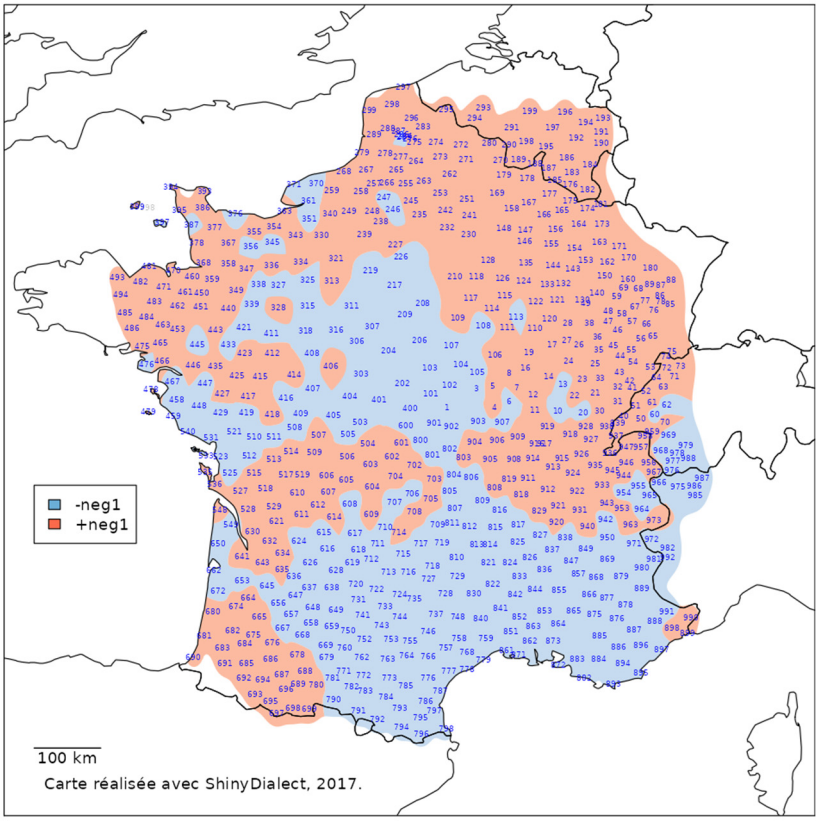


Fig. 9　neg1 omission/retention of Map 1082

6.1.9. Map 1083

This is the only map using the indefinite pronoun *on* as the subject. There are 418 +neg1 points (66%) and 213 −neg1 points (34%). Even in the Occitan and Francoprovençal-speaking regions, which typically see neg1 omission, some locations retain neg1. Particularly in the Occitan-speaking region, the indefinite pronoun *on* is used. According to Ronjat (1937: §774), personal pronouns are not usually used in Occitan, and they are used to avoid ambiguity or for contrast. In most maps analyzed, except for Map 12 and Map 1083, personal pronouns were not used in the Occitan-speaking regions. This suggests that the presence or absence of personal pronouns affects the omission/retention of neg1 in the Occitan-speaking regions.

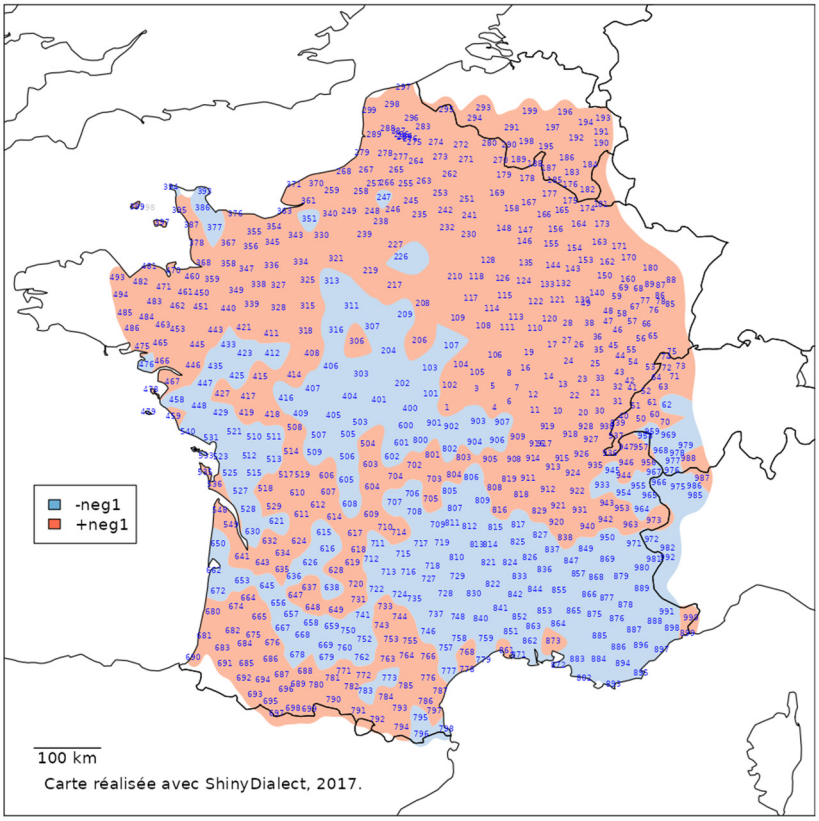


Fig. 10 neg1 omission/retention of Map 1083

6.1.10. Map 1352

This map uses a negative adverb in a coordinate clause introduced by the conjunction *mais* “but.” There are 374 +neg1 points (59%) and 257 −neg1 points (41%). While previous maps of negative sentences in the coordinate clause (Map 896 and Map 899) had many +neg1 points (436 points and 491 points respectively), this map shows different tendencies. Unlike the previous two maps, which had nouns as subjects, this map uses pronouns, indicating that this difference may influence neg1 retention.

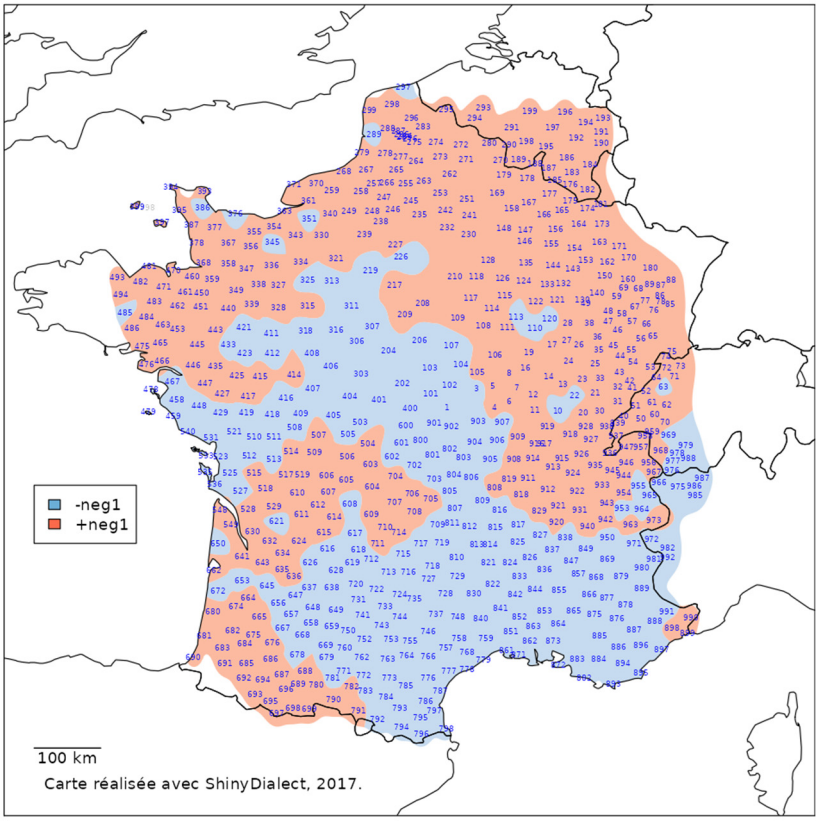


Fig. 11　neg1 omission/retention of Map 1352

6.1.11. Map 1409

This is the only map using the second-person singular pronoun. There are 393 +neg1 points (59%) and 258 −neg1 points (41%), with more +neg1 points. The negative sentence is in the main clause leading the subordinate clause *que tu es aussi vieux que moi*. Ashby (1981) states that *ne* is more likely to be omitted in main clauses, but this is not necessarily the case in this map. Instead, as Meisner et al. (2014) mention, the complexity of the right elements of neg1 seems to influence neg1 retention.

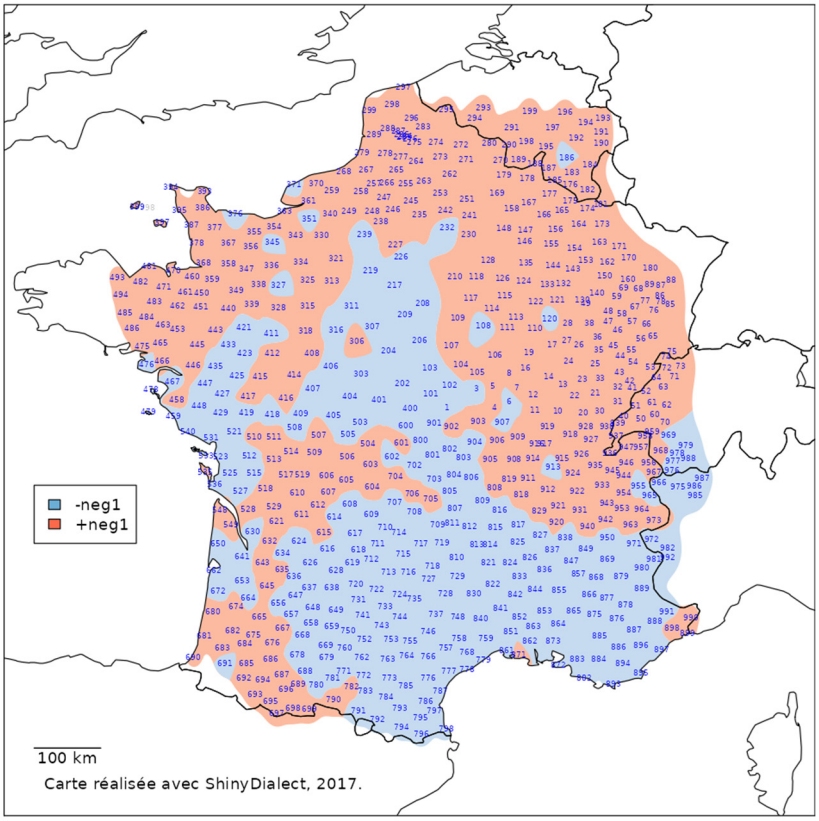


Fig. 12 neg1 omission/retention of Map 1409

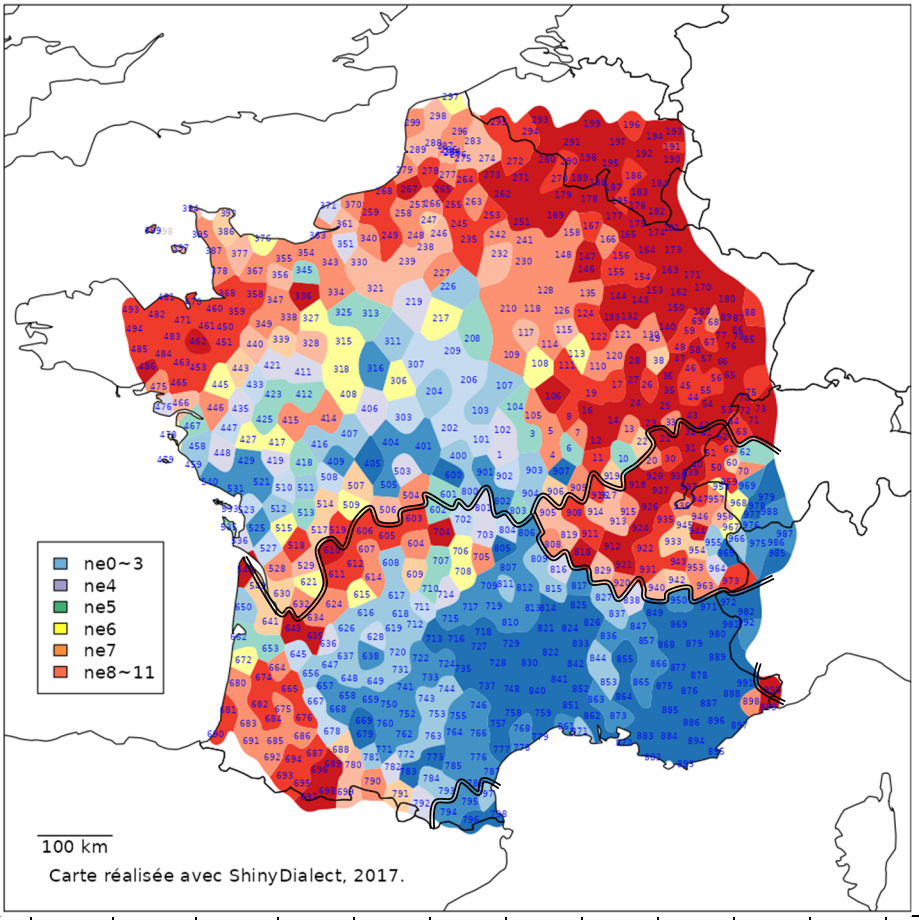
6.2. Analysis of neg1 omission/retention in the entire ALF area

6.2.1. Analysis of neg1 omission/retention by regionality

Figure 13 shows the interpretation map of the number of neg1 retentions in the 11 maps. The double lines in the figure indicate language boundaries. Points where neg1 is omitted/retained in more than 70% of the maps (n0~n3 and n8~n11) are summarized in a single legend.

From the figure, regional tendencies in the omission/retention of neg1 can be observed. In the Occitan-speaking regions (excluding Gascony), the Catalan-speaking regions, the Francoprovençal-speaking regions of Valais in Switzerland and Aosta in Italy, n0~n3 are prevalent. Conversely, in the northern French-speaking regions (including the northwest and northeast), around the Croissant area, excluding the aforementioned regions, in the Francoprovençal-speaking regions, Gascony in the Occitan-speaking regions, and the Ligurian-speaking region, n8~n11 are more common.

Furthermore, many n5 and n6, where the omission/retention of neg1 is about half and half, are distributed between regions with strong tendencies to retain or omit neg1. There are 330 points where neg1 is omitted in more than 70% of the maps (n0~n3), 207 points where neg1 is retained in more than 70% of the maps (n8~n11), and 94 intermediate points (n4~n7). The breakdown of each language region is as shown in Table 4.



	n0	n1	n2	n3	n4	n5	n6	n7	n8	n9	n10	n11
no	73	64	49	21	24	23	24	23	37	87	94	112
%	11.5	10	8	3	4	3.5	4	3.5	6	14	15	17.5

Fig. 13 Summary of neg1 omission/retention in 11 maps

Table 4 Number of Points Classed into n0~n3, n4~n7 and n8~11 in each Language Region

	ALF	French	Occitan	Francoprovençal	Catalan	Ligurian
n0~n3	207	44	145	13	5	0
n4~n7	94	62	43	10	0	0
n8~n11	330	236	22	50	0	1
total	631	342	210	73	5	1

A two-way ANOVA ($p<0.05$) was conducted to investigate the relationship between the overall tendencies of neg1 omission/retention and language regions. Factor (A) was the tendency of neg1 omission (three levels), and Factor (B) was the number of language regions (five regions). The results are shown in Table 5. The p-value for Factor (A) = 0.467 > 0.05, indicating no significant difference in the tendency of neg1 omission/retention. The p-value for Factor (B) = 0.162 > 0.05, also indicating no significant difference in the tendency of language regions. Thus, there is no significant relationship between language regions and the overall tendency of neg1 omission/retention.

Table 5 Result of two-way ANOVA ($p<0.05$)

	variation	df	variance	F-value	p-value	decision
Factor (A)	5572.9	2	2786.5	0.84	0.467	[]
Factor (B)	28928.9	4	7232.2	2.18	0.162	[]
Error	26583.1	8	3322.9			

6.2. Analysis of neg1 omission/retention by map

Table 6 summarizes the number (and percentage) of points where neg1 was retained and omitted in each map. Maps 12 and 89 have over 60% of points where neg1 was omitted. Conversely, maps 896, 897, 899, and 1352 have over 60% of points where neg1 was retained. Maps 1082, 1352, and 1409 have more points retaining neg1, while maps 806 and 817 have about half and half.

To identify maps where neg1 is more likely to be retained or omitted, the analysis focused on points classified as n1~n10, excluding n0 and n11. For example, 64 points are classified as n1. If these 64 points retained neg1 only on specific maps, those maps could be considered more likely to retain neg1.

Table 6 Number and Percentage of neg1 Omisison/Retention in each Map

	12	89	806	817	896	897	899	1082	1083	1352	1409
+neg1	205	254	323	341	436	416	491	361	418	374	373
	(32)	(40)	(51)	(54)	(69)	(66)	(78)	(57)	(66)	(59)	(59)
−neg1	426	377	308	290	195	215	140	270	213	257	258
	(68)	(60)	(49)	(46)	(31)	(34)	(22)	(43)	(34)	(41)	(41)

Table 7 summarizes the relationship between points where neg1 was retained (n1~n10) and the maps where neg1 was retained. The bottom row shows the number of points retaining neg1 excluding the number of n11 points. Row n1, for example, shows the maps where the 64 points classified as n1 retained neg1. For instance, 31 out of 64 points retained neg1 in Map 899, 16 in Map 1083, etc. Column 12 shows the composition of points retaining neg1 in Map 12. Of the 93 points, 30 are classified as n9, and 49 as n10.

Table 7 Number of Points Classed into n1~n10 in each Map: entire *ALF*

	12	89	806	817	896	897	899	1082	1083	1352	1409	total
n1	0	2	2	0	9	4	31	0	16	0	0	64
n2	0	7	3	1	12	9	40	1	19	0	6	49
n3	0	0	3	3	15	12	15	3	5	3	4	21
n4	1	3	3	7	17	16	20	6	15	6	2	24
n5	1	6	6	6	21	19	18	10	12	7	9	23
n6	3	8	4	14	21	20	20	10	14	15	15	24
n7	5	6	9	13	21	17	22	16	14	20	18	23
n8	4	12	21	23	33	35	36	30	34	35	33	37
n9	30	31	73	72	81	82	85	79	85	84	81	87
n10	49	67	87	90	94	90	92	93	92	92	93	94
total	93	142	211	229	324	304	379	249	306	262	261	

Figure 14 is a box plot based on Table 7, showing the distribution of neg1 retention of points classed in n1~n10 for each map. From the table and figure, it is clear that Map 12 has notably fewer points retaining neg1. Similarly, Map 89 also has relatively few points retaining neg1. On the other hand, Maps 896, 897, 899, and 1083 show higher medians, indicating more points retaining neg1. To verify whether the differences in neg1 retention tendencies between maps are statistically significant, a hypothesis test was conducted with the following null hypothesis:

Null Hypothesis: The tendency of neg1 retention is unrelated to the type of map.

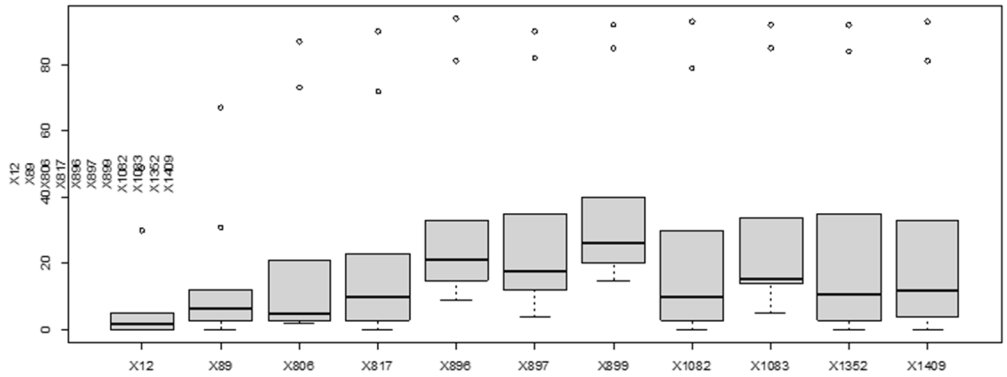


Fig. 14 Box Plot of Number of n1~n10 Points in each Map: entire *ALF*

Wilcoxon signed-rank tests were conducted for all combinations of maps at a significance level of 1%. The results are shown in Table 8. Bold indicates p-values < 0.01, indicating the rejection of the null hypothesis for those combinations. For example, Map 12 shows significant differences with all maps except Map 1352, indicating that Map 12 tends to omit neg1 more than other maps except Map 1352.

Table 8 Results of Wilcoxon signed-rank tests (p-values < 0.01): entire *ALF*[7]

	12	89	806	817	896	897	899	1082	1083	1352	1409
12											
89	**0.009**										
806	**0.006**	0.27									
817	**0.009**	0.05	0.139								
896	**0.006**	**0.006**	**0.006**	**0.006**							
897	**0.006**	**0.006**	**0.006**	**0.009**	0.036						
899	**0.006**	**0.006**	**0.006**	**0.006**	0.137	0.015					
1082	**0.009**	0.047	0.024	0.174	**0.006**	0.012	**0.008**				
1083	**0.006**	**0.006**	**0.006**	**0.009**	0.356	0.959	0.014	0.019			
1352	0.014	0.047	0.05	0.076	0.036	0.058	**0.009**	0.201	0.154		
1409	**0.009**	0.032	0.022	0.057	0.014	0.032	**0.008**	0.16	0.111	0.725	

To visually clarify the results of the Wilcoxon signed-rank tests, the significant differences were converted to 0-1 data[8], and hierarchical clustering (complete linkage, Manhattan distance) was performed. The 11 maps can be broadly divided into two groups based on statistical significance. The box plot and dendrogram are compared to analyze these two groups.

[7] All values were rounded to the fourth decimal place. The same applies to all subsequent analyses.
[8] P-values < 0.01 were marked as 1, and p-values > 0.01 were marked as 0. The same applies to all subsequent analyses.

The first group includes Maps 12, 896, 897, 899, and 1083. This group contains maps with strong tendencies for neg1 omission and retention. Among the 11 maps, Map 12 showed the most significant differences, with statistical significance against all maps except Map 1352. Therefore, Map 12 tends to omit neg1 more than other maps except Map 1352. The next most significant differences were seen in Map 899, showing statistical significance against seven maps. This indicates that Map 899 tends to retain neg1. Maps 896, 897, and 1083 are also statistically significant in retaining neg1 but do not show statistical significance with Map 899, indicating that these three maps are also less likely to omit neg1.

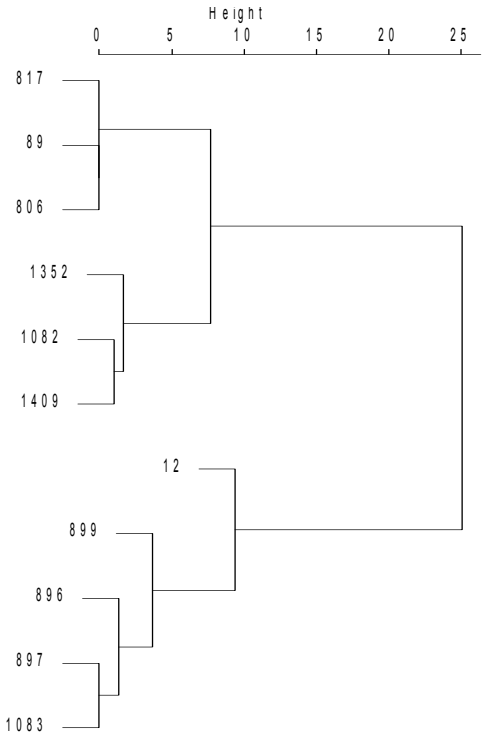


Fig. 15 Result of hierarchical clustering of 11 maps: 631 points

The second group includes Maps 89, 806, 817, 1082, 1352, and 1409. Among these, Maps 89, 806, and 817 show statistical significance with the maps in the first group, indicating that they tend to retain neg1 compared to Map 12 and tend to omit neg1 compared to Maps 896, 897, 899, and 1083. Maps 1082, 1352, and 1409 show statistical significance with three or fewer maps, indicating that these three maps do not have a strong tendency to omit or retain neg1.

6.3. Analysis of neg1 omission/retention by language region

In 6.2, it was clarified that Map 12 tends to omit neg1, while Maps 896, 897, 899, and 1083 tend to retain neg1. To validate these results, similar analyses were conducted for each language region except for Catalan and Ligurian.

6.3.1. Analysis of neg1 omission/retention in the French-speaking regions

Of the 342 points in the French-speaking regions, 236 points (70%) are classified as n8 or higher, indicating a strong tendency to retain neg1. Points with a strong tendency to omit neg1 (n3 or lower) are 44 points (12%), mainly distributed in Poitevin, Berrichon, and Bourbonnais. Points with a wide variation in neg1 retention/omission

(n4~n7) are 62 points (18%). Table 9 summarizes the relationship between points retaining neg1 (n1~n10) and the maps where neg1 was retained in the French-speaking regions. Figure 16 is a box plot based on Table 9.

Table 9 Number of Points Classed into n1~n10 in each Map: 342 Points in French-speaking regions

	12	89	806	817	896	897	899	1082	1083	1352	1409	total
n1	0	0	1	0	7	2	2	0	0	0	0	12
n2	0	0	3	0	7	3	10	1	1	0	5	15
n3	0	0	2	3	14	6	10	1	2	2	2	14
n4	0	1	2	5	14	10	13	4	8	3	0	15
n5	1	3	4	3	15	14	14	9	11	3	8	17
n6	0	4	2	11	17	14	14	7	14	9	10	17
n7	0	1	6	9	13	9	13	11	7	10	12	13
n8	2	7	16	17	26	27	26	20	24	26	25	27
n9	15	17	52	52	60	59	59	56	60	59	60	61
n10	32	43	60	64	64	60	64	64	64	62	63	64

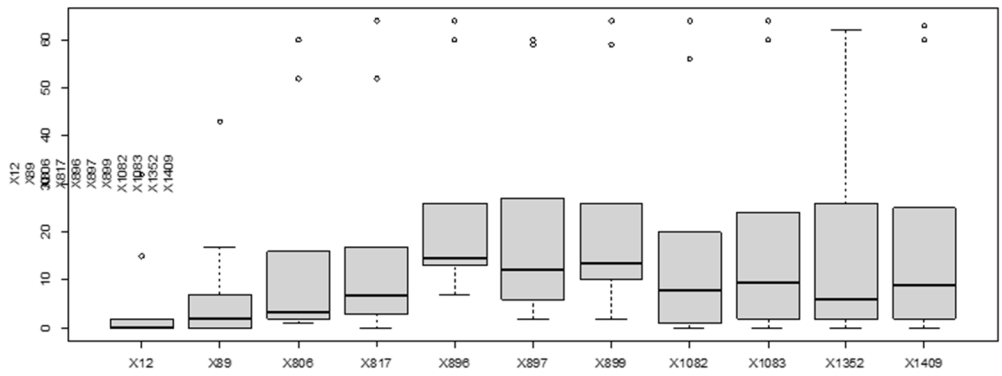


Fig. 16 Box Plot of Number of n1~n10 Points in each Map: 342 Points in French-speaking regions

In the French-speaking regions, Maps 12 and 89 tend to omit neg1 more than other maps. Maps 896 and 899, consistent with the analysis in section 6.2, tend to retain neg1 in the French -speaking regions. Maps 897 and 1083, which showed a tendency to retain neg1 in section 6.2, do not necessarily show this tendency in the French-speaking regions. Table 10 shows the results of statistical hypothesis tests. Bold indicates p-values < 0.01, indicating the rejection of the null hypothesis.

Table 10 Results of Wilcoxon signed-rank tests (p-values < 0.01): 342 Points in French-speaking regions

	12	89	806	817	896	897	899	1082	1083	1352	1409
12											
89	0.022										
806	**0.006**	0.021									
817	0.014	0.022	0.187								
896	**0.006**	**0.006**	**0.006**	**0.009**							
897	**0.002**	**0.002**	0.014	0.043	0.01						
899	**0.006**	**0.006**	**0.006**	**0.009**	0.125	0.056					
1082	**0.009**	**0.009**	0.035	0.399	**0.009**	0.041	**0.009**				
1083	**0.009**	**0.009**	0.038	0.068	0.014	0.135	0.021	0.121			
1352	0.014	0.022	0.122	1	**0.009**	0.044	0.014	0.952	0.206		
1409	0.014	0.013	0.028	0.232	**0.009**	0.22	0.012	0.23	0.799	0.201	

The significant differences were converted to 0-1 data, and hierarchical clustering (complete linkage, Manhattan distance) was performed. The 11 maps can be broadly divided into three groups based on statistical significance. The first group includes Maps 896 and 899. Map 896 shows significant differences with five maps, and Map 899 with six maps. These two maps tend to retain neg1 in the French-speaking regions, consistent with the overall analysis. The second group includes Maps 12 and 89. Map 12 shows significant differences with six maps, and Map 89 with five maps, indicating that Maps 12 and 89 tend to omit neg1. The third group includes the remaining seven maps. Among these, Maps 897 and 1083, which showed a tendency to retain neg1 in the analysis in 6.2, are included.

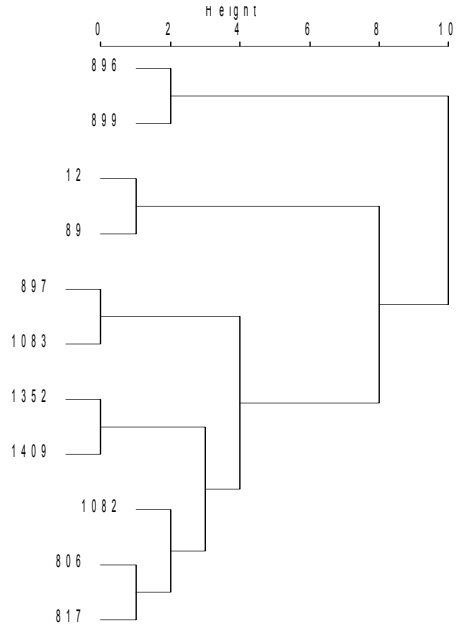


Fig. 17 Result of hierarchical clustering of 11 maps: 342 points in French-speaking region

6.3.2. Analysis of neg1 omission/retention in the Occitan-speaking regions

Of the 210 points in the Occitan-speaking regions, 145 points (69%) are classified as n3 or lower, indicating a strong tendency to omit neg1. Points with a strong tendency to retain neg1 (n8 or higher) are 43 points (21%), and points with wide variation in neg1 omission/retention (n4~n7) are 22 points (10%). Most n4 or higher points are distributed in Limousin, Marchois, and Gascon.

Table 11 summarizes the relationship between points retaining neg1 (n1~n10) and the maps where neg1 was retained in the Occitan-speaking regions. Figure 18 is a box plot based on Table 11.

Table 11 Number of Points Classed into n1~n10 in each Map: 210 Points in Occitan-speaking regions

	12	89	806	817	896	897	899	1082	1083	1352	1409	total
n1	0	2	1	0	2	2	24	0	13	0	0	44
n2	0	7	0	1	3	5	28	0	17	0	1	31
n3	0	0	1	0	1	6	5	2	3	1	2	7
n4	1	2	1	1	3	5	6	2	7	3	1	8
n5	0	3	2	3	5	4	3	1	1	3	0	5
n6	2	4	0	2	3	3	4	1	0	3	2	4
n7	2	4	1	2	4	3	5	3	5	5	1	5
n8	1	0	2	2	1	1	2	2	2	2	1	2
n9	6	10	14	14	15	13	16	14	15	16	11	16
n10	8	12	15	14	16	16	16	16	15	16	16	16

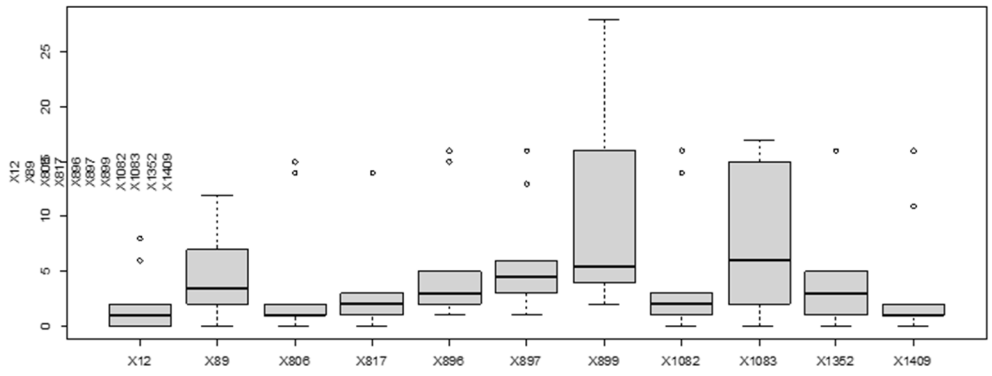


Fig. 18 Box Plot of Number of n1~n10 Points in each Map: 210 Points in Occitan-speaking regions

In the Occitan-speaking regions, the distribution of neg1 omission/retention varies significantly from the French-speaking regions. This is due to the higher prevalence of neg2 in the Occitan-speaking regions. From the box plot, it is clear that Map 12 tends to omit neg1, while Map 899 tends to retain neg1. Map 1083 also shows a tendency to

retain neg1. To verify the statistical significance of the differences in neg1 omission/retention by map, Wilcoxon signed-rank tests were conducted for all combinations of maps at a significance level of 1%. Table 12 shows the results of the statistical hypothesis tests. Bold indicates p-values < 0.01, indicating the rejection of the null hypothesis.

Table 12 Results of Wilcoxon signed-rank tests (p-values < 0.01): 210 Points in Occitan-speaking regions

	12	89	806	817	896	897	899	1082	1083	1352	1409
12											
89	0.015										
806	0.178	0.644									
817	0.058	0.72	0.588								
896	**0.009**	0.228	0.019	0.01							
897	**0.009**	0.151	0.024	0.03	0.524						
899	**0.006**	0.014	**0.009**	0.014	0.042	0.054					
1082	0.037	1	0.24	0.665	0.053	0.042	0.014				
1083	0.014	0.059	0.042	0.075	0.312	0.645	0.029	0.076			
1352	0.014	0.55	0.05	0.049	0.395	0.397	0.059	0.066	0.4		
1409	0.136	0.356	0.774	0.526	0.024	0.013	**0.009**	0.188	0.045	0.076	

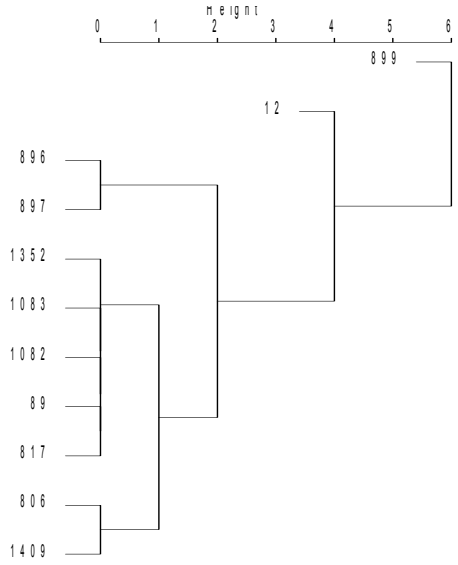


Fig. 19 Result of hierarchical clustering of 11 maps: 210 points in Occitan-speaking regions

The significant differences were converted to 0-1 data, and hierarchical clustering (complete linkage, Manhattan distance) was performed. The 11 maps can be broadly divided into three groups based on statistical significance. The first group includes Map 899. Map 899 shows significant differences with Maps 12, 806, and 1409, which tend to omit neg1. Of the 44 points classified as n1 in the Occitan-speaking regions, 24 retained neg1, and of the 31 points classified as n2, 28 retained neg1, indicating that Map 899 tends to retain neg1. The second group includes Map 12. Map 12 shows significant differences with

Maps 896, 897, and 899, which tend to retain neg1. The box plot distribution shows that Map 12 has very few points retaining neg1, indicating that it tends to omit neg1. The third group includes the remaining nine maps. Among these, Maps 806 and 1409 show significant differences with Map 899, indicating that they tend to omit neg1 compared to Map 899. Maps 896 and 897 show significant differences with Map 12, indicating that they tend to retain neg1 compared to Map 12. However, considering the overall 11 maps, these four maps do not have a strong tendency to omit or retain neg1.

6.3.3. Analysis of neg1 omission/retention in the Francoprovençal-speaking regions

Of the 73 points in the Francoprovençal-speaking regions, 50 points (69%) are classified as n8 or higher, indicating a strong tendency to retain neg1. Points with a strong tendency to omit neg1 (n3 or lower) are 13 points (18%), and points with wide variation in neg1 omission/retention (n4~n7) are 22 points (13%). Points with a strong tendency to omit neg1 are mainly distributed in isolated mountainous areas like Valais in Switzerland and Aosta in Italy.

Table 13 summarizes the relationship between points retaining neg1 (n1~n10) and the maps where neg1 was retained in the Francoprovençal-speaking regions. Figure 20 is a box plot based on Table 13.

Table 13 Number of Points Classed into n1~n10 in each Map: 73 Points in Francoprovençal-speaking regions

	12	89	806	817	896	897	899	1082	1083	1352	1409	total
ne1	0	0	0	0	0	0	4	0	2	0	0	6
ne2	0	0	0	0	2	1	1	0	0	0	0	2
ne3	0	0	0	0	0	0	0	0	0	0	0	0
ne4	0	0	0	1	0	1	1	0	0	0	1	1
ne5	0	0	0	0	1	1	1	0	0	1	1	1
ne6	1	0	2	1	1	3	2	2	0	3	3	3
ne7	3	1	2	2	4	5	4	2	2	5	5	5
ne8	1	5	3	4	6	7	8	8	8	7	7	8
ne9	9	4	7	6	6	10	10	9	10	9	10	9
ne10	9	12	12	12	14	14	12	14	13	14	14	14

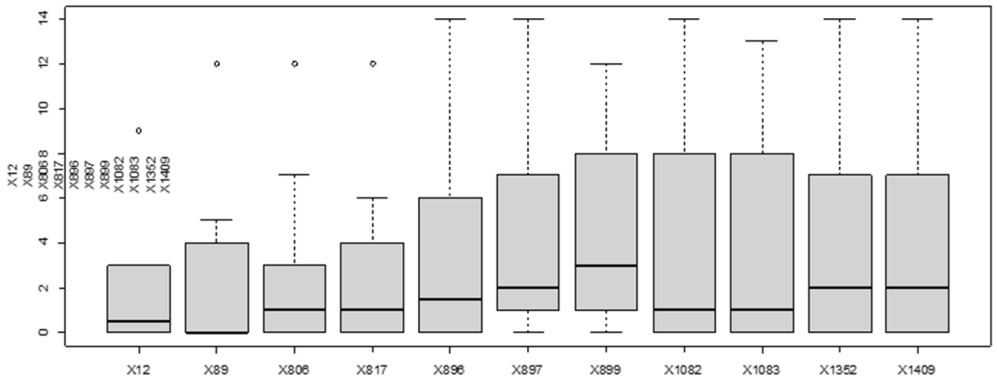


Fig. 20 Box Plot of Number of n1~n10 Points in each Map:
73 Points in Francoprovençal speaking regions

Unlike the French-speaking and Occitan-speaking regions, the Francoprovençal-speaking regions shows less variation in neg1 omission/retention by map. Nonetheless, Map 12, similar to the other regions, tends to omit neg1 more than other maps. To verify the statistical significance of the differences in neg1 omission/retention by map, Wilcoxon signed-rank tests were conducted for all combinations of maps at a significance level of 1%. Table X shows the results of the statistical hypothesis tests. Bold indicates p-values < 0.01, indicating the rejection of the null hypothesis.

Table 14 Results of Wilcoxon signed-rank tests (p-values < 0.01):
73 Points in Frqncoprovençal-speaking regions

	12	89	806	817	896	897	899	1082	1083	1352	1409
12											
89	1										
806	0.586	0.462									
817	0.679	0.203	1								
896	0.206	0.021	0.103	0.066							
897	0.013	0.014	0.013	0.022	0.106						
899	**0.007**	0.014	0.022	0.021	0.232	1					
1082	0.269	0.058	0.174	0.136	0.915	0.071	0.168				
1083	0.203	0.058	0.223	0.202	0.943	0.222	0.056	1			
1352	0.058	0.036	0.035	0.04	0.341	0.149	0.608	0.345	0.546		
1409	0.021	0.022	0.021	0.035	0.242	1	1	0.12	0.301	0.346	

Statistical significance was only observed between Maps 12 and 899. Map 12 tends to omit neg1 compared to Map 899, and Map 899 tends to retain neg1 compared to Map 12. However, compared to the French-speaking and Occitan-speaking regions, there is little difference in the tendencies of neg1 omission/retention by map in the Francoprovençal-speaking regions.

6.4. Summary of the analysis

The analysis revealed that Map 12 tends to omit neg1, while Map 899 tends to retain neg1. In the French-speaking regions, Map 89 also tends to omit neg1, and Map 896 tends to retain neg1, which is a characteristic specific to the French-speaking regions. In the analysis of the entire *ALF* area, Maps 897 and 1083 also showed statistical tendencies to retain neg1, but this is not necessarily the case in the analysis by each language region. These maps are relatively more likely to retain neg1, in the context of *ALF* as whole.

The regional analysis and map analysis indicated that Map 12, the only map using the strong form of the first-person pronoun, *moi*, has the strongest tendency to omit neg1. This tendency is consistent with Meisner et al.'s (2015) findings that strong forms lead to neg1 omission even in *ALF*, data from the early 20th century. Similarly, Map 899 showed a tendency to retain neg1 when the subject position is occupied by a proper noun or nouns other than personal pronoun.

Table 15 Summary of neg1 omission/retention

	ALF	French	Occitan	Francoprovençal
12	−neg1	−neg1	−neg1	−neg1
89		−neg1		
806				
817				
896	+neg1	+neg1		
897	+neg1			
899	+neg1	+neg1	+neg1	+neg1
1082				
1083	+neg1			
1352				
1409				

7. Conclusion

In this study, we analyzed the retention rate of neg1 for each map to elucidate whether there are regions where neg1 tends to omit/retain and whether there are negative sentences where neg1 tends to omit/retain. To this end, we conducted an analysis of 11 maps of negative sentences from the *ALF*.

The findings indicate that in the Occitan-speaking regions, excluding Gascony, the Catalan-speaking regions, the Francoprovençal-speaking regions of the Valais in Switzerland, and the Aosta Valley in Italy, there were numerous instances where neg1 was omitted. Conversely, in the northern French-speaking regions (particularly in the northwest and northeast), around the Croissant area, the Occitan-speaking regions of Gascony, and the Ligurian-speaking region, there were many instances where neg1 was retained. However, no statistically significant differences were observed in the retention/omission of neg1 when considered as language regions, and it became evident that there is variability even within a single language region.

The analysis of each map revealed that the omission of neg1 was statistically prominent in maps where the disjunctive forms were used, whereas the retention of neg1 was prominent in maps where proper nouns were employed. In maps where *nous* and *vous*, which are generally considered to easily retain neg1, were the subjects, no statistically significant differences in the retention of neg1 were observed. However, this may be attributed to discrepancies between the map names and the forms that emerged in the actual dialect surveys. Thus, it can be inferred that there is a significant relationship between the subject and the omission/retention of neg1, as evidenced from the linguistic maps.

Abbreviations

Atlas	
ALF	*Atlas linguistique de la Francce*
Gloss	
ACC	accusative
COND	conditional
FUT	future
IND	indicative
INF	infinitive
NEG	negative
NEG1	pre-verbal negative
NEG2	post-verbal negative
PRONEXPL	impersonal pronoun
PRS	present

PL	plural
REL	relative
SG	singular

Other

−neg1	neg1 omission
+neg1	neg1 retention

Cited references

ALF = Gilliéron, Jules & Edmond Edmont (1902-1910) *Atlas linguistique de la France*. Paris: Honoré Champion.

Armstrong, Nigel & Alan Smith (2002) The influence of linguistic and social factors on the recent decline of French *ne*. *Journal of French language Studies* 12: 23–41.

doi: https://doi.org/10.1017/S0959269502000121

Ashby, William (1976) The loss of the negative morpheme NE in Parisian French. *Lingua* 39: 119–137.

doi: https://doi.org/10.1075/slcs.226.c7

Ashby, William (1981) The loss of the negative particle ne in French: a syntactic change in progress. *Language* 57(3): 674–684. doi: https://doi.org/10.1075/slcs.226.c12

Ashby, William (2001) Un nouveau regard sur la chute du *ne* en français parlé tourangeau. S'agit-il d'un changement en cours? *Journal of French Langauge Studies* 11(1): 1–22.

doi: https://doi.org/10.1017/S0959269501000114

Burnett, Heather (2019) Sentential Negation in North-eastern Gallo-Romance dialects: insights from the Atlas Linguistique de la France. *Journal of French Language Studies* 29: 189–207.

doi: https://doi.org/10.1017/S0959269519000218

Coveney, Aidan (2002) [2nd edition] *Variability in Spoken French. A Sociolinguistic Study of Interrogation and Negation*. Bristol and Portland: Elm Bank Publications.

Dagnac, Anne (2015) *Pas*, *mie*, *point* et autres riens : de la négation verbale en picard. In: Jan Goes & Mariana Pitar (eds) *La negation: études linguistiques, pragmatiqies et didactiques*, 129–152. Arras: Artois Presses Université. URL: https://books.openedition.org/apu/6786

Gilliéron, Jule & Edmond Edmont (1902) Atlas linguistique de la France : Notice servant à l'intelligence des cartes. Paris: Honoré Champion.

Goebl, Hans (2018) La face cachée de la géographie linguistique. Bref aperçu sur les 'cartes muettes' produits pour l'ALF, l'AIS et le FEW. *Revue de linguistique romane* 82: 6–63.

URL: https://goebl.ads.plus.ac.at/people/prof/goebl/docs/Goebl2018_RLR82.pdf

Hansen, Anita & Isabelle Malderez (2002) Le ne de négation en région parisienne : une étude en temps réel. *Langage et Société* 107: 5–30. doi: https://doi.org/10.3917/ls.107.0005

Jagueneau (2007) Négation simple et négation discontinue en occitan limousin. In: Frank Florick (ed) *La négation dans les langues romanes*, 99–116. Amsterdam: John Benjamins. doi: https://doi.org/10.1075/lis.26.06jag

Kawaguchi, Yuji, Reiko Ito, Takamasa Seimiya & Kaho Okawara (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: 146–163. doi: https://doi.org/10.5281/zenodo.5529358

Meisner, Charlotte & Matascha Pomino (2014) Synchronic variation in the expression of French negation: A Distributed Morphology approach. *French Language Studies* 24: 9–28.

doi: https://doi.org/10.1017/S0959269513000355

Meisner, Charlotte, Aurélia Robert-Tissot & Elisabeth Stark (2015) L'absence et la présence du *NE* de négation. In: *Encyclopédie grammaticale du français*. URL : http://encyclogram.fr

Moreau, Marie-Louise (1986) Les séquences préformées : entre les combinaisons libres et les idiomatismes. Le cas de la négation avec ou sans *ne*. *Le français moderne* 54: 137–160.

Pollock, Jean-Yves (1989) Verb Movement, Universal Grammar, and the Structure of IP. *Linguistic Inquiry* 20(3): 365–424. URL: http://www.jstor.org/stable/4178634

Ronjat, Jules (1930–1941) *Grammaire historique des parlers provençaux modernes*. Montpellier: Société des langues romanes.

Rowlett, Paul (1998) *Sentential negation in French*. New York / Oxford: Oxford University Press.

Schwegler, Armin (1986) The *Chanson de Saint Foy*: Etymology of *Cabdorn* (with Cursory Comments on the Localization of Poem). *Romance Philology* 39(3): 285–304.

Seimiya Takamasa (2021) Keiryōhōgengaku ni okeru gokei no keitaitougo-kankei wo han'ei saseta gyōretsu sakusei no kōsatsu - Atlas linguistique de la France no 9 mai no hiteibun no chizu wo mochiite - 『計量方言学における語形の形態統語関係を反映させた行列作成の再考 - Atlas linguistique de la France の9枚の否定文の地図を用いて –』[Reflection on the validity of matrices for analyzing morpho-syntactic relations of words in dialectometry - a study of 9 maps on negation from the Atlas linguistique de la France -]. *Flambeau* 47: 58–76. doi: https://doi.org/10.15026/117134

Publication history

Date received: 10 July 2024
Date accepted: 4 September 2024

Research article

Lexical forms and their distribution of the word for ‘yesterday’ in Yunnan Tibetan

SUZUKI, Hiroyuki
Kyoto University

Abstract: This article describes the word forms for ‘yesterday’ in Khams Tibetan varieties spoken in Yunnan, China, and analyses their distribution. It confirms two types of word forms: Type A, corresponding to Literary Tibetan *kha rtsang* (A1) and its derivations (A2); and Type B, of unclear origin. Type B is isolated to the mBalhag dialect, whereas the remaining varieties belong to Type A. This discussion focuses on the sound shape of Type A owing to this situation. The Nyishe subgroup displays a distinctive feature that involves restructuring the second syllable’s *r*-preinitial into the final consonant of the first syllable. This phenomenon is specific to the subgroup, and this article interprets it as a local innovation resulting from phonological features.*

Keywords: Tibetic; Khams Tibetan; lexical distribution; expansion; syllabic restructure

1. Introduction

This article describes word forms for ‘yesterday’ in Khams Tibetan varieties in the Tibetosphere of Yunnan Province (henceforth, Yunnan Tibetan), and then examines how they are distributed from a geolinguistic viewpoint. According to Tournadre and Suzuki (2023:823), two word forms for ‘yesterday’ are mainly utilised in Tibetic languages: *kha rtsang* and *mdang* (literally meaning ‘last night’) in Literary Tibetan (LT).

In Yunnan Tibetan, *kha rtsang* is widespread with some phonetic modifications or derivational changes. However, etymological information on the form *kha rtsang* is unknown. A syllable-by-syllable analysis has not yet been performed. However, based on derivational morphology, the first syllable *kha* is also used in words, such as *khas*

SUZUKI, Hiroyuki. 2024. Lexical forms and their distribution of the word for ‘yesterday’ in Yunnan Tibetan. *Studies in Geolinguistics* 4: 95–106. doi: https://doi.org/10.5281/zenodo.13948564

* An earlier version of the article was presented at the fifth annual meeting of Geolinguistic Society of Japan (Nanzan University, 2023). I should thank Ray Iwata, Mitsuaki Endo, and Satoko Shirai for their insightful comments. The work was supported by the ILCAA Joint Research Project ‘Studies in Asian and African Geolinguistics’ and JSPS KAKENHI Grant Number JP17H04774.

nyin ‘the day before yesterday’ and *kha nub* ‘last night’, and hence it can be considered to be a core semantic part of ‘the day before the reference point’.

In Literary Tibetan, the regular spelling is *kha rtsang*; however, *kha sang* is also accepted, reflecting the sound reality of several Tibetic languages. In addition, various pronunciations appear, such as *kha rtsa* in Dzongkha and *kha’i sa* in Lhasa. This is also the case in Yunnan Tibetan.

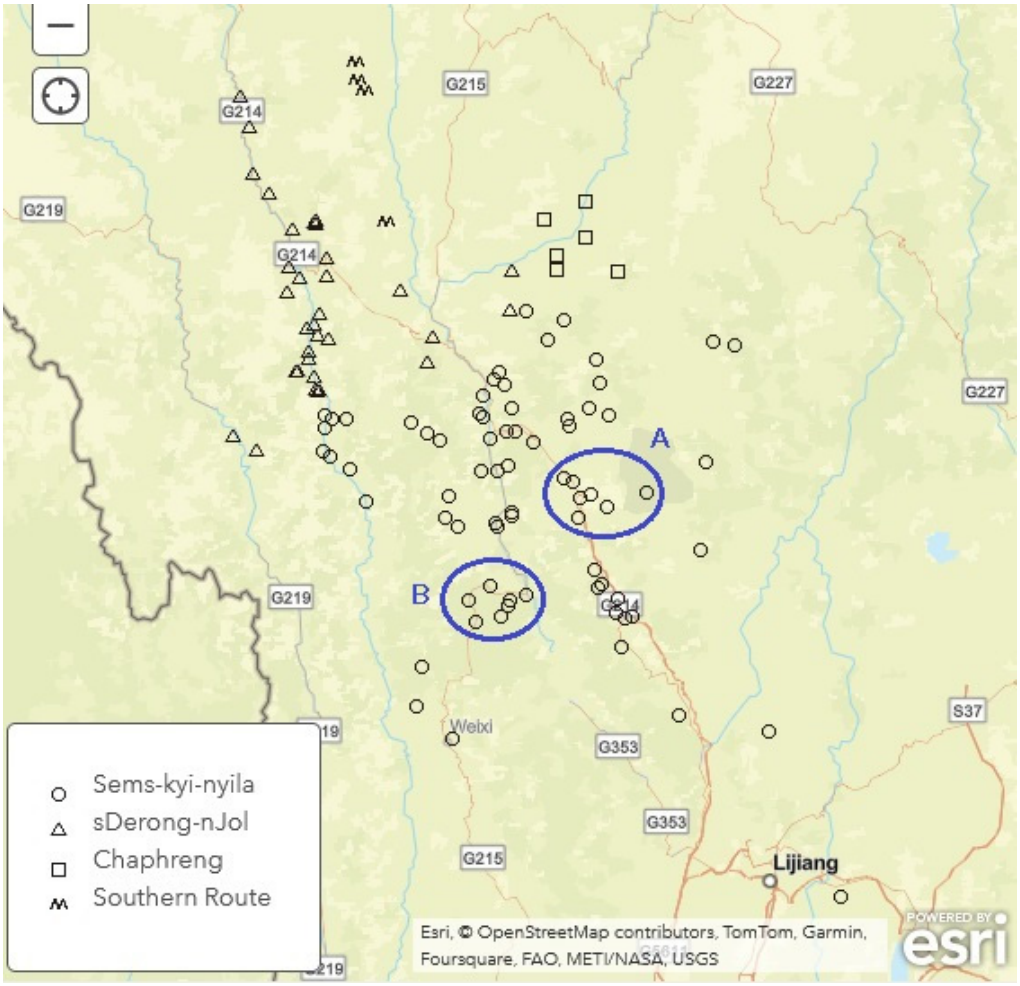


Fig. 1 Classification of Yunnan Tibetan.

Before a discussion of the forms of ‘yesterday’ in Yunnan Tibetan, I display Figure 1, showing the most recent dialectal classification of these varieties (Suzuki 2022b, 2023) with two key areas: A-rGyalthang Town and B-Tacheng Town. Geographical information contains rivers (sky blue) and main traffic roads (brown). This area is in

the south-eastern corner of the Tibetosphere, and other ethnic languages are spoken in contact with Tibetic (see Roche and Suzuki 2018).

2. Word forms

Two forms are distinguished for the word for 'yesterday' in the target varieties: Type A, *kha rtsang* and its derivational forms; and Type B, /ˊda ɬɔ̃/, of an unclear origin. The two forms are classified as follows:

Type A
A1: a disyllabic form corresponding to the LT form (e.g., /ˊkʰa ʰtsɔ̃/ and /ˊkʰəˑ tsɔ̃/).
A2: a trisyllabic form with the first and second syllables' initial as /kʰ/ and /ts/, respectively (e.g., /ˊkʰɛ tsə mA/).
Type B
/ˊda ɬɔ̃/ only.

The above classifications do not distinguish between detailed variations in pronunciation. Type A1 exhibits diversity; however, every sound shape corresponds to LT *kha rtsang*. Type A2 is a word consisting of LT *kha rtsang* and a suffix *ma*. However, the factors related to the addition of the last morpheme are uncertain.

Notably, Tibetic languages in Yunnan are unlikely to use a form corresponding to LT *mdang*. Although Type B contains /d/, it does not correspond to the LT *md* initial since the spelling *md* generally corresponds to a prenasalised initial /ⁿd/ or even a nasal /n/ in Yunnan Tibetan (cf. Suzuki 2016, 2018), whereas its second syllable, /ɬɔ̃/, can correspond to LT *sang*. However, it is difficult to confirm whether Type B is related to LT *sang*, a variant of *rtsang* in LT *kha rtsang*. Moreover, except for Type B, word forms developed in Yunnan or its contacting regions, as well as borrowings, are not attested.

3. Geographical distribution and its analyses

Figure 2 illustrates the geographical distribution of the three types of word forms mentioned above. Type A1 is the most widespread type, whereas Types A2 and B are geographically characterised.

Type A2 is distributed in the south of the map, in Weixi County. It is employed in varieties belonging to the Melung subgroup (in a narrow sense, except for Phongpa; see Suzuki 2024a for the classification) of the Sems-kyi-nyila group. The varieties

spoken in Tacheng Town use different types of word forms for ‘yesterday’ depending on the dialectal groups (Fig. 3). The word forms do not seem to have an influence beyond the subdialectal groups. An interpretation of the case of Phongpa is concerned with the archaicity of Types A1 and A2. When A1 is older than A2, innovation occurs in the central Weixi area of the Melung subgroup. By contrast, if the ABA distribution is applied, then A2 is older than A1; hence, Phongpa’s form A1 is replaced by A2 through a borrowing process from a neighbouring variety.

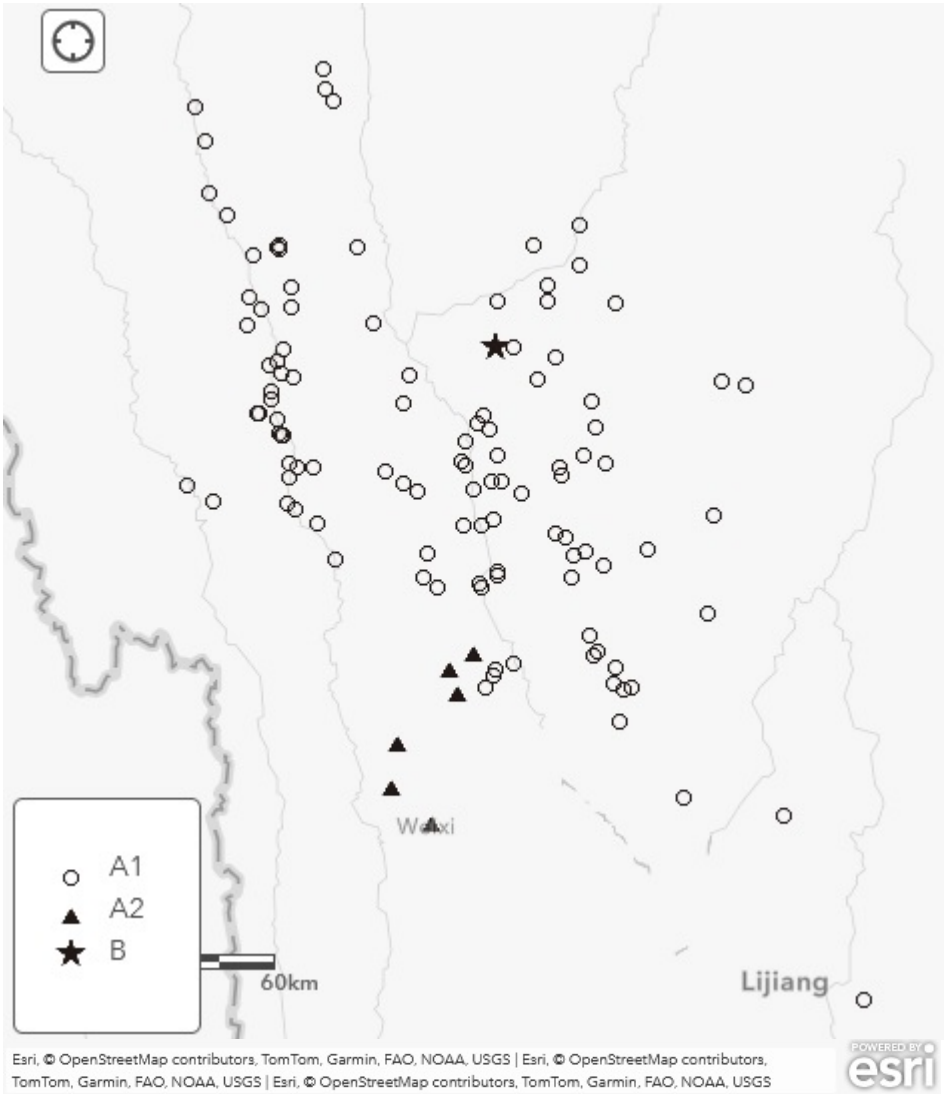


Fig. 2 Word forms for ‘yesterday’.

Type B is isolate at mBalhag. It is difficult to explain its origin owing to the lack of sources of word forms and the migration history of mBalhag Tibetans (see Suzuki 2013).

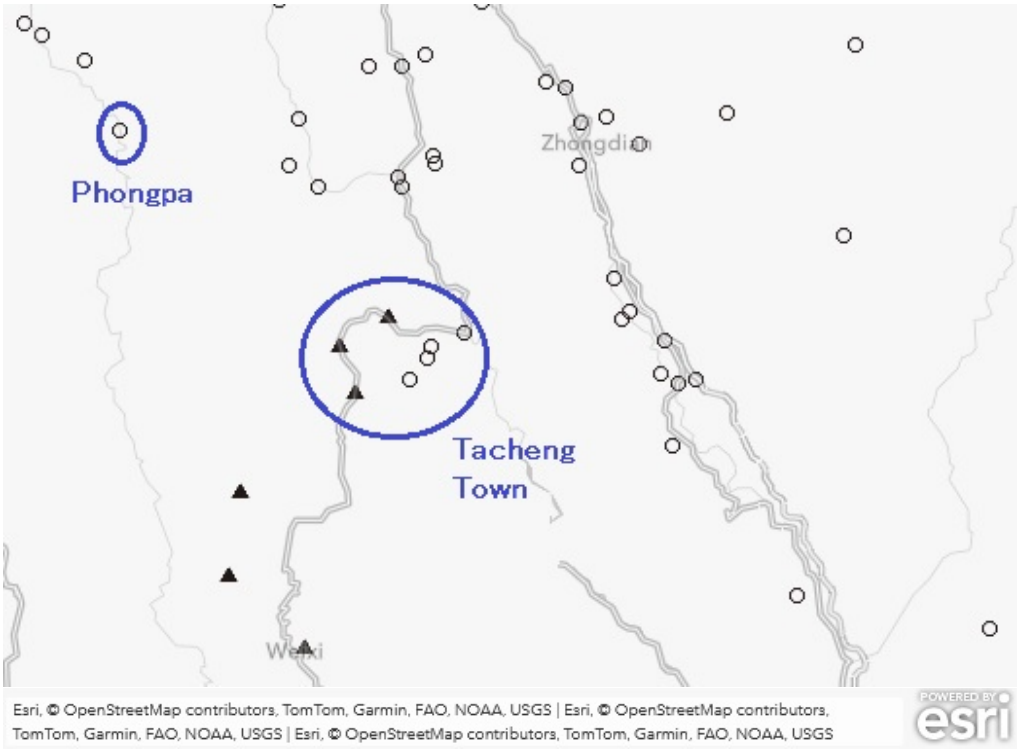


Fig. 3 Word forms for 'yesterday'. Weixi area enlarged.

Type A1 can be further analysed based on its phonological form. The first syllable's rhyme (vowel plus final, if any) is focused on, corresponding to LT *a* in *kha rtsang*, as this part exhibits great variation among the varieties. Hypothesised factors triggering the variation are indicated as a difference in the regularity of the sound correspondence with LT *a#* and a difference in prosodic features (iambic or trochaic; see Suzuki 2022a:299-300). Figure 4 shows the phonological description of the target feature (i.e., the rhyme form of Type A1's first syllable), which is analysed according to the vocalic types with symbols.

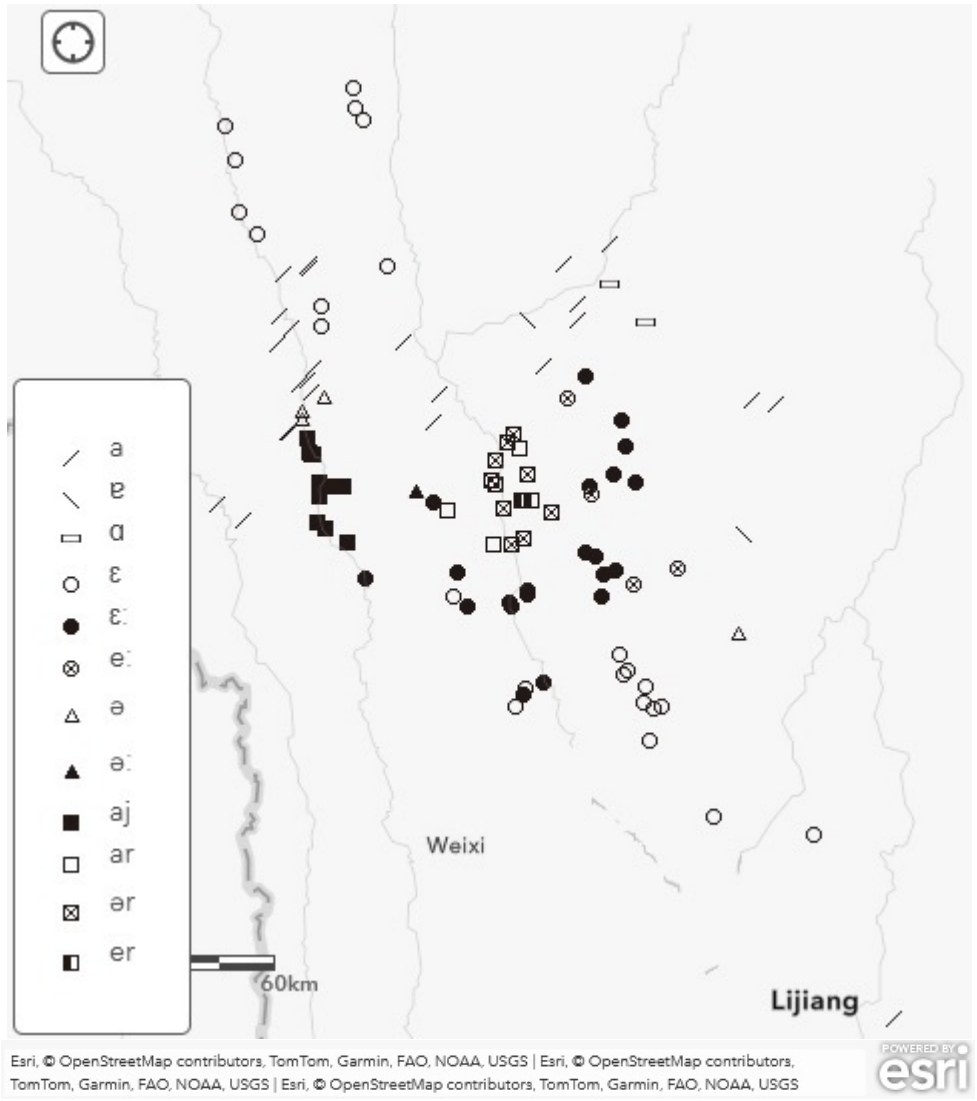


Fig. 4 Type A1's rhyme forms (sound correspondence of *a* in *kha*).

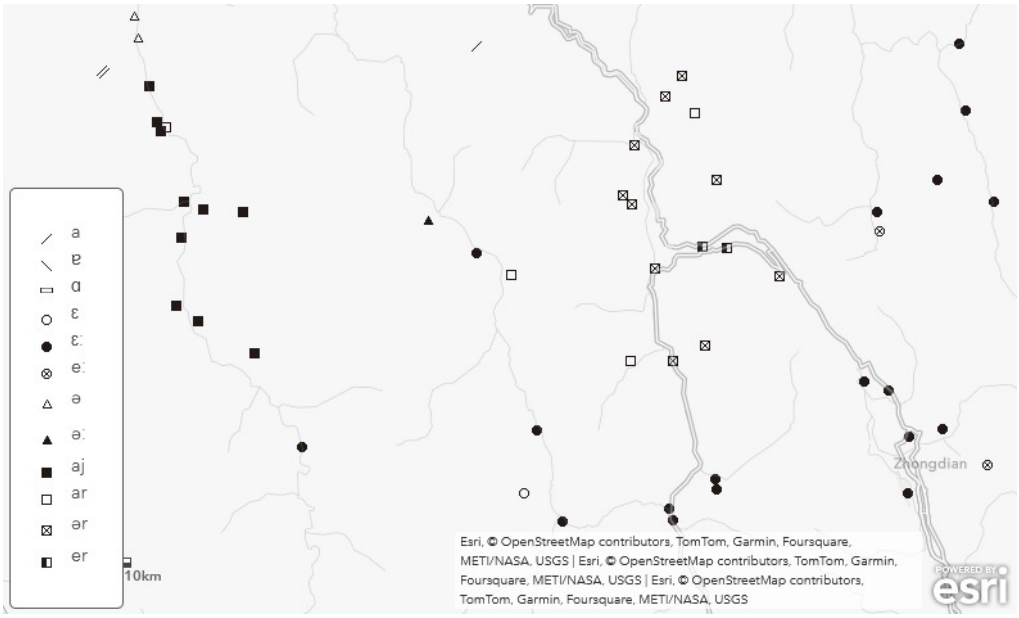


Fig. 5 Type A1's rhyme forms or sound correspondence of *a* in *kha*. Central area enlarged for the distribution of the forms with a final (square symbols).

The first three low-vowel types follow a general sound correspondence with LT *a*# for each variety. Particularly, the sound correspondence with the vowels /ɐ/ and /ɑ/ belongs to the minority within the Tibetic languages (Tournadre and Suzuki 2023:252).

The square symbols represent a type with a final (/r/ and /j/) and are distributed in the central area (along the Jinshajiang River) and to the west (along the Lancangjiang River). This type is striking because a restructuring of two syllables was observed.

Meanwhile, the circle symbols front mid-vowels, in general, do not correspond to LT *a*#, even though the prosodic feature is considered. It is reasonable that rhymes /ar/ and /aj/ (square symbols) evolved into /ɛ/ or /e/ (circle symbols). This process has been observed in various Tibetic varieties of Yunnan (Suzuki 2018:47-48). Long vowels are indicative of an ongoing sound change process whereby a rhyme consisting of a vowel + a final evolves into a short vowel; however, the vowel length is influenced by various prosodic features as well as each variety's phonological system. The historical order of the vowel length cannot always be assessed. It is noteworthy that all of these features appear in the central area (Fig. 5).

Figure 6 shows the simplified types of Type A1's rhyme, reflecting the variations and archetypes shown in Figure 4. The rhyme is classified into the following three types:

A: low vowel type, following the regular sound correspondence with LT *a*#

B: mid-vowel type

C: final's potential existence type

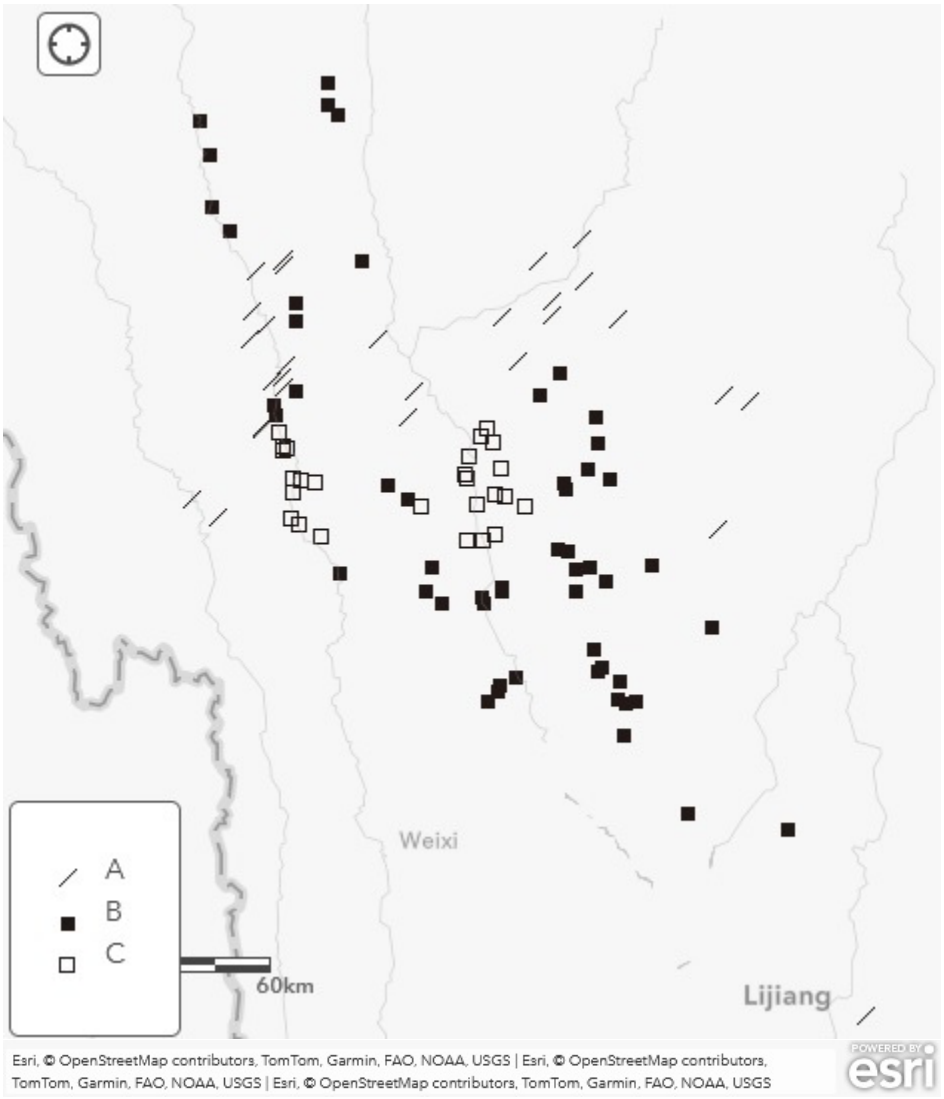


Fig. 6 Simplified classification of Type A1's rhyme.

Type A is unmarked; hence, its distribution is less valuable from the perspective of locality. Nevertheless, it is worth noting that Type A appears in the northern zone, the Chaphreng group, the sPomtserag and Bodgrong subgroups, and part of the nJol subgroup of the sDerong-nJol group.

Type B has two origins. As mentioned earlier, the emergence of mid-vowels originated from a prosodic pattern (iambic) and the evolution of rhymes from Type C.

Type B appears around Type C, for which we consider an ABA-distribution; however, Type C is likely to be an older form than Type B.

Type C includes r-final and j-final. The r-final can change into j-final under certain conditions, following the general tendency of sound changes attested in various Tibetic languages (Tournadre and Suzuki 2023:254). For Type C distributed in the west part, a similar sound change process is attested in some words with a *r*-final, as in *dkar* 'white': ar > a^{ɣ} > aj > ɛ (Suzuki 2011).

After the sound change processes are considered, two hypotheses are proposed:

(1) A > B; and

(2) A > C > B.

The former is a change triggered by the local prosodic feature, that is, /a/ corresponding to LT *kha (rtsang)* changed into another mid vowel, whereas the latter is a change in which the segmental features between syllables are restructured, such as LT *kha rtsang* > *khar tsang. Therefore, these two processes are independent.

Process (1) reflects a natural sound change process, whereas Process (2) is not analysed as a regular sound change owing to the lack of parallel examples. Presently, a similar syllable restructure happens only in the word for 'nineteen' corresponding to LT *bcu dgu* > *bcur gu (Suzuki 2024b). In Yunnan Tibetan, Process (1) belongs to the minority and is found in north-western parts (Fig. 6). This type is likely to be separated from the varieties that use process (2) by Type A (slash), as shown in Figure 7.

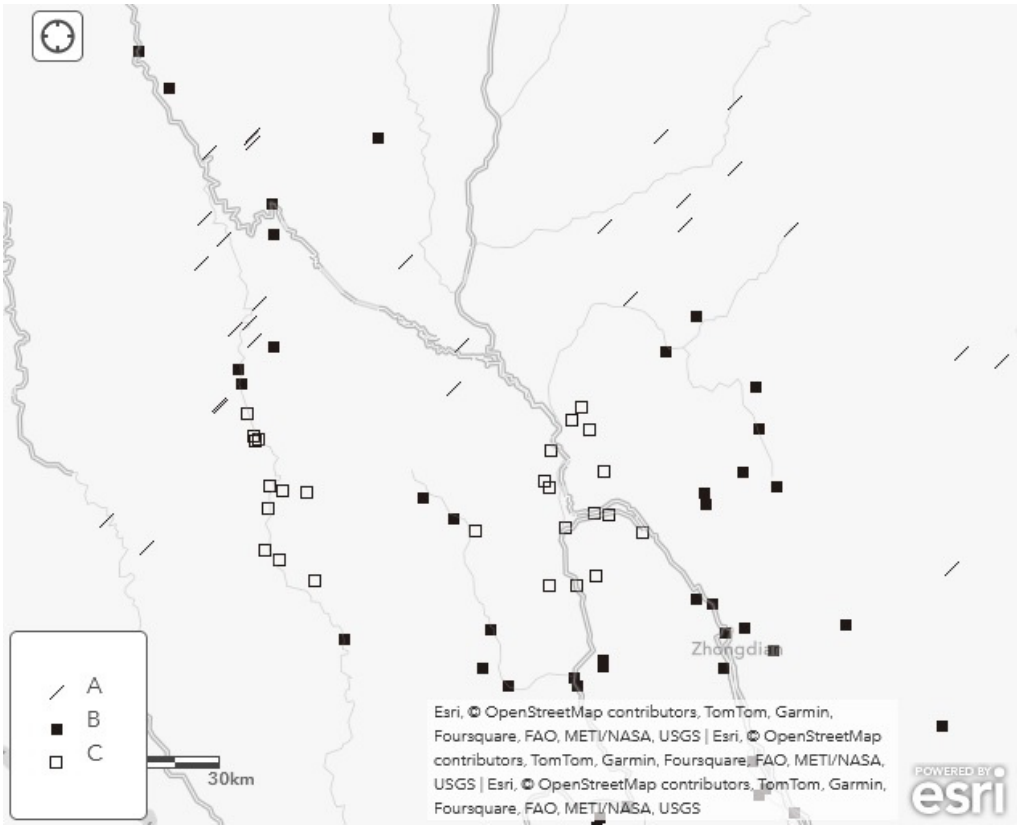


Fig. 7 Simplified classification of Type A1's rhyme (enlarged).

The r-final form of Type C is characterised by varieties in the Nyishe subgroup of rGyalthangic of the Sems-kyi-nyila group (Suzuki 2019, 2024a). These varieties are spoken in the periphery of the rGyalthangic-speaking region; hence, Types B and C exhibit an ABA distribution, although distorted, because Type C is a retention of the older form and Type B, as a new form, expands from the centre (rGyalthang Town, the local economic centre) to its surrounding areas. The ABA-distribution with rGyalthang as the centre is also attested in the word form for 'today' (Suzuki 2023).

The archaicity of Type C maintained in Nyishe is probably owing to its geographical isolation in the steep valley. However, a principal road (National Route 214; G214 in Fig. 1) penetrates the area to connect rGyalthang with nJol. Hence, language contact has become more frequent following the augmentation of local traffic, which has brought changes to the Nyishe varieties. Indeed, the Thangstod variety of Nyishe, spoken in the easternmost area of Nyishe's distribution region in contact with the rGyalthang varieties, exhibits a similarity in the r-final form to the rGyalthang varieties (Suzuki 2019).

4. Conclusion

This article examined the lexical forms for 'yesterday' in Tibetic languages in the Yunnan Tibetosphere. Except for a single variety (mBalhag) using /ˊda ɬɔ̃/, all the varieties employ a word derived from or related to LT *kha rtsang* 'yesterday'. The latter is further classified into two groups: disyllabic and trisyllabic. The trisyllabic form is limited to the Melung subgroup of Weixi County. The disyllabic form is distributed in the widest region and three types are found based on the first syllable's rhyme. These types reflect the change process in which a new form expands from the local centre, rGyalthang, to its periphery, forming an ABA-distribution.

References

Roche, Gerald and Hiroyuki Suzuki (2018) Tibet's minority languages: Diversity and endangerment. *Modern Asian Studies* 52(4): 1227–1278. doi: https://doi.org/10.1017/S0026749X1600072X

Suzuki, Hiroyuki (2009) Naxi bunkaken no Tibettogo Yongsheng ken Daan [Daan] hoogen no hoogen syozoku 納西文化圏のチベット語・永勝県大安[Daan]方言の方言所属 [Dialectal affiliation of the Daan dialect of Tibetic spoken in Yongsheng County in the Naxi cultural area]. *Bulletin of National Museum of Ethnology* 34(1): 167–189. doi: https://doi.org/10.15021/00003919

Suzuki, Hiroyuki (2011) Gagatang Zangyu de yanhua yuanyin yu qi laiyuan 嘎嘎塘藏語的咽化元音與其來源 [Pharyngealised vowels of Gagatang Tibetan and their origin]. *Language and Linguistics* 12(2): 477–500.

Suzuki, Hiroyuki (2011) Zai yinbian guocheng zhong chansheng you xiaoshi de ruanehua yuanyin: Yunnan Deqin Yanmen xiang Guzha Zangyu zhi li 在音变过程中产生又消失的软腭化元音——云南德钦燕门乡谷扎藏语之例 [Velarised vowels produced and disappeared in the course of the sound change: Examples from sGograg Tibetan spoken in Yanmen, Deqin, Yunnan]. *Kyoto University Linguistic Research* 30: 35–49. doi: https://doi.org/10.14989/159068

Suzuki, Hiroyuki (2013) Extraordinary sound development of *s and *z in mBalhag Tibetan (Shangri-La, Yunnan). *Linguistics of the Tibeto-Burman Area* 36(1): 101–110. doi: https://doi.org/10.15144/LTBA-36.1.101

Suzuki, Hiroyuki (2016) Xianggelila Zangyu Yalanghua de biyin xitong 香格里拉藏语亚浪话的鼻音系统 [Nasal system of the gYaglam dialect of Sems-kyi-nyila Tibetan]. *Dongfang Yuyanxue* 16: 114–122.

Suzuki, Hiroyuki (2018) Xianggelila-shi hokubu no Kamutibettogo syohoogen no hoogen tokutyoo to sono keisei 香格里拉市北部のカムチベット語諸方言の方言特徴とその形成 [Dialectal variation and development of Khams Tibetan in Northern Shangri-La Municipality]. *Journal of Asian and African Studies* 95: 5–63. doi: https://doi.org/10.15026/92458

Suzuki, Hiroyuki (2019) Liyong yuyanditu chanming yinbian de kuosan he jiexian: Yi Xianggelila Zangyu de "r-yunwei" yuyin yanbian wei li 利用语言地图阐明音变的扩散和界限：以香格里拉藏语的"r韵尾"语音演变为例 [Exploring a geographical expansion and restriction of sound changes with linguistic maps: A case study on the sound change of the 'r' rhyme in rGyalthang Tibetan]. In Hiroyuki

Suzuki, Keita Kurabe, and Mitsuaki Endo (eds.) *Dongbu Yazhou diliyuyanxue lunwenji*, 1–13. Fuchu: ILCAA. URI: https://publication.aa-ken.jp/sag_mono6_eastern_asian_2019.pdf

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) Dialectal affiliation of Tibetic varieties in gYagrwa within Yunnan Tibetan. *Kyoto University Linguistic Research* 41: 43–68. doi: https://doi.org/10.14989/281541

Suzuki, Hiroyuki (2023) Geolinguistic analysis of word forms for 'today' in Tibetic languages in Yunnan. *Studies in Geolinguistics* 3: 92–98. doi: https://doi.org/10.5281/zenodo.8437061

Suzuki, Hiroyuki (2024a) Shaping rGyalthangic: A historical account of Yunnan Khams. In Masaki Nohara and Takumi Ikeda (ed.) *Grammatical phenomena of Sino-Tibetan languages 6: Typology and historical change*, 87–108. Kyoto: Institute for Research in Humanities, Kyoto University.

Suzuki, Hiroyuki (2024b) Yunnan Kamutibettogo ni okeru suusi ni motoduku tirigengogakuteki bunseki 雲南カムチベット語における数詞に基づく地理言語学的分析 [Geolinguistic analysis based on numerals of Khams Tibetan in Yunnan]. Paper presented at 神戸市外国語大学Research Project B アジア諸言語の接触と変容：通時的・共時的観点からのアプローチ 2023年度研究会.

Suzuki, Hiroyuki (2024c) Evolution of dorsal fricatives in rGyalthangic varieties of Khams Tibetan. *Journal of the Phonetic Society of Japan* 28(2): in press.

Tournadre, Nicolas and Hiroyuki Suzuki (2023) *The Tibetic languages: An introduction to the family of languages derived from Old Tibetan*. Villejuif: LACITO Publications. doi: https://doi.org/10.5281/zenodo.10026628

Publication history

Date received: 7 July 2024
Date accepted: 8 August 2024

Research article

Another contribution to the hypothesis for Austronesian toponyms in Vietnam

TRẦN, Thị Hồng Hạnh
Faculty of Linguistics, USSH, Vietnam National University Hanoi

TRƯƠNG, Nhật Vinh
Faculty of Linguistics, USSH, Vietnam National University Hanoi

Abstract: This paper employs historical phonological evidence and geographical linguistic maps to examine specific toponyms in Vietnam, including Phú Gia, Bà Đanh, Sấu - Giá, and Hà Lầm, Hà Lam, Kha Lâm, Khả Lãm, in order to further investigate the contact between Vietnamese and Austronesian speakers. Two main hypotheses have been posited regarding the origins of Austronesian toponyms, particularly those found in Northern Vietnam. The first suggests that the toponyms resulted from early contact during Austronesian migrations to Southeast Asia. The second posits that contact occurred later, possibly after the 10th century during Đại Việt's conquests of the Champa region, leading to the migration of Cham prisoners into the Northern Delta. Drawing on cultural and historical knowledge, this paper suggests Phú Gia, Bà Đanh, and Sấu - Giá likely emerged from language contact post-10th century. However, further study is needed for toponyms with Lâm/Lam/Lầm/Lãm. The findings also suggest that the presence of Austronesian-speaking residents in Northern Vietnam in antiquity is evident, though their migration timelines and directions remain debated.

Keywords: Austronesian; toponym; linguistic map; migration

1. Introduction

According to archaeological research results, the northern region of Vietnam has been inhabited by humans for a long time. There is substantial consensus among scholars such as Gourou (1936), C.Cœdès (1938) (See Nguyễn Văn Huyên 1944), K.Taylor (1983), Hà Văn Tấn (1997), M.Alves (2021) regarding the view that early Vietnamese residents in the Northern Delta region were not a homogeneous community, but had interactions with diverse communities speaking different languages.

TRẦN, Thị Hồng Hạnh and TRƯƠNG, Nhật Vinh. 2024. Another contribution to the hypothesis for Austronesian toponyms in Vietnam. *Studies in Geolinguistics* 4: 107–118. doi: https://doi.org/10.5281/zenodo.13948576

In his book Les Paysans du delta Tonkinois (The Farmers of the Tonkinese Delta: A Study of Human Geography), P. Gourou (1936) remarked that "researching and investigating toponyms seems to often lead to disappointment; unlike Europe, where toponyms are often ancient and offer valuable insights into successive waves of immigration that have shaped a country, in Tonkin as well as in China, toponyms frequently change and are often arbitrary and artificial creations. (...) Many villages also possess customary names alongside official ones. (...) Some of these customary names bear little resemblance to the official names, whether they are vestiges of ancient toponyms or manifestations of an ancient language—because some of these customary names sometimes assume peculiar forms. Addressing these inquiries necessitates a deep understanding of Vietnamese, Tày, and Chinese languages" (See Gourou 2017, pp. 137-138). This observation by Gourou appears to recognize that the Vietnamese language has been influenced by neighboring languages, particularly Chinese (Sino-Tibetan family) and Tày (Tai-Kadai language family), and that ancient toponyms may serve as evidence of linguistic interactions in antiquity.

In a study of the Phùng Nguyên culture (starting around 2000 BCE), based on the characteristics of stone tools, Hà Văn Tấn (1997) hypothesized "the Phùng Nguyên tribes, the core to form the Vietnamese ethnic community... spoke a dialect of the pro-Austroasiatic language family but were greatly influenced by the two Proto-Tai and Proto-Malayo-Polynesian language families" (Hà Văn Tấn 1997, p.370). But it should be noted that the author is also very cautious when making the above statement because, according to him, "finding the languages of ancient tribes is always difficult" and "documents about Phùng Nguyên culture are still very limited. Many reconstructions are, rather, just hypotheses" (Hà Văn Tấn 1997, p.379).

It can be said that Northern Vietnam is currently a concentrated residence of mainly Austroasiatic speakers, but in fact that is an Austroasiatic community which have undergone a most extraordinary developmental process and contact with other languages, specifically Tai-Kadai, Austronesian (Malayo-Polynesian), Sino-Tibetan and and later European languages. Researching language contact in Vietnam in particular, in Southeast Asia in general with a diachronic approach has never been a simple task and the results and achievements have never satisfied scholars. Thus, studying toponyms can help to shed light on the issues of historical linguistics as well as hypotheses about the population migration in prehistoric and early historical periods. This paper delves into some toponyms in Vietnam to determine the origin of these toponyms and interpret the distributions of those toponyms on maps.The findings may help to enhance our understanding of the language contact between Vietnamese and Austronesian speakers and migration of Austronesian speakers to Northern Vietnam.

2. Data analysis

So far, there have been several published studies discussing the Austronesian toponyms in Vietnam, aiming at not only etymological issues but also the language as well as culture contact. Trần Trí Dõi is one of the authors who pays a lot of attention to this topic. In his studies, (2001a, 2001b, 2005, 2007) (See Trần Trí Dõi 2022) he indicated that toponyms with Cán/Càn/Cờn/Gàn and toponyms Cổ Loa, Cà Lồ, Cửa Lò all originated from Austronesian languages. He also argued that the groups of Cổ Loa, Cà Lồ, Cửa Lò related to the form "kơlo/klo" which means "river mouth, estuary" in Austronesian languages and Cán/Càn/Cờn/Gàn might derive from "ikan" which means "fish" in Austronesian languages. These toponyms, especially river names and estuary names, might have appeared very early and be "the most durable and oldest names compared to other toponyms such as mountain names, names of resident or settlement places, etc.... only where there is water can people live" (Hoàng Thị Châu 2004, p.5).

Trần Thị Hồng Hạnh (2019) studied the Nôm toponyms Sình of the Lại Ân village (Thừa Thiên Huế) which is now used to name the waterfront, local market as well the folk woodcut painting produced here. In this study, this toponym was identified to derive from ching. In Cham language, ching means "to cut, to carve, to engrave". The phonetic form [ciŋ] corresponds to the phonetic form [ʂiŋ^{2}] of Sình, which allows to hypothesize that this village toponym might have originated from a Cham word which refers to the village's famous craft-making wood carving pictures.

In an investigation into the toponym Phú Gia, Trương Nhật Vinh (2020) showed that cultural, geographical information and especially historical linguistics helped to demonstrate that Phú Gia might be a name of Cham origin. Phú Gia 富家 is an ancient village located northwest of West Lake, along the Red River, which is now the area of Phú Gia Ward and Phú Thượng Ward, Tây Hồ district, Hanoi. The village also has a folk name of Kẻ Gạ. Researching historical records and documents, he discovered that the village had three names: Phú Gia, Đa Gia Li and Bà Già. The name Phú Gia appeared as early as the 18th century. Meanwhile, names Đa Gia Li and Bà Già might be used as early as the late 11th century and continued to be used until around the mid-17th century. The resconstructed form of Đa Gia Li might be /*dua ia lich/ or /dua ia ling/, which in Cham language means "a village located between two floodwaters/two rivers that cause damage floods" (and the two rivers that might be referred to the Red River and Thiên Phù River).

Using the same method from an interdisciplinary approach with a focus on historical linguistics, Trương Nhật Vinh (2023) explained the origins and meaning of another toponym Bà Đanh, the name of a famous pagoda in Northern Vietnam. He indicated

that Bà Đanh was the result of the Vietnamization process of the Cham goddess named Po Yang Dari.

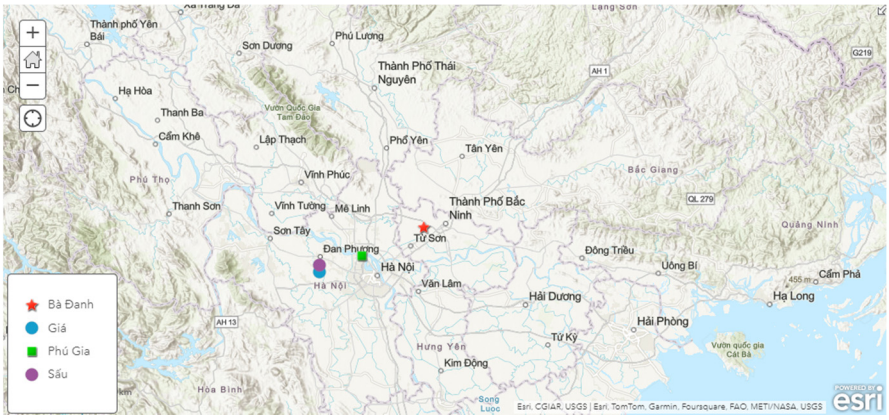
Map 1. The distribution of Bà Đanh toponyms in Northern Vietnam (Trương Nhật Vinh 2023)

In another recent study, Trương Nhật Vinh offered an explanation for the toponyms Sấu-Giá, two clusters of villages located adjacent to each other in Hoài Đức district, Hanoi today (Trương Nhật Vinh 2024). The Sino-Vietnamese names of the villages and communes in the Sấu-Giá region include Yên Sở, Đắc Sở, Dương Liễu, Mậu Hòa, and Quế Dương. These toponyms were documented quite early compared to other Sino-Vietnamese toponyms of villages and communes in the Red River Delta. According to him, the name Giá might be a Vietnamese adaptation of the form *ia/êa, which means "water" in Cham language, indicating a village located near a water source. The geographical location of Yên Sở village along the banks of the Đáy River and the presence of Cổ Sở/Yên Sở wharf today (close to Quán Giá) further support his hypothesis. Additionally, he posited that the toponym Giá, as a Vietnamized name, implies that Giá village was situated in a low-lying area.

In contrast, Vinh suggested that the toponym Sấu village originated from the Vietnamization of the Cham word "chaur", meaning "high pile/dunes of soil". His fieldwork revealed that Sấu village is a large settlement, encompassing both inside and outside the dike. The inner part of the dike features a system of markets, communal houses, pagodas, while the outer part is known as the "new village". The village's terrain is higher compared to surrounding areas. Moreover, the prevalence of water wells built using Cham techniques and the abundance of coconut trees in the village support the notion that Sấu derives from Cham language, reflecting the topographical characteristics of the village situated on higher ground.

In his studies, Trương Nhật Vinh asserted that the commonality among these toponyms - *Phú Gia, Bà Đanh, and Sấu-Giá* - is that they all resulted from language contact between Vietnamese and Austronesian languages (specifically Cham) after the 10th century. There is a consensus among scientists that contact between speakers of Austroasiatic languages (specifically Vietnamese) and Austronesian languages could have occurred in two possible ways. The first possibility is early contact before Vietnamese separated into an independent language. At this stage, contact likely occurred during the migration of Austronesian-speaking populations.

Currently, there are differing opinions about the homeland of Austronesian languages. Some authors argue their origin is in mainland Asia (Kern 1889), specifically Southeastern China (Haudricourt 1954, Bellwood 1984-1985, Blust 1984-1985) (See Hà Văn Tấn 1997). Others suggest their homeland might be the islands of Southeast Asia (Solheim 1984-1985; Meacham 1984-1985). Meanwhile, according to Hà Văn Tấn (1997), their homeland might be in Southeastern China and the coastal areas of Southeast Asia (including Vietnam). He even suggests that "the creators of the Bàu Tró culture may have spoken an Austronesian language" (Hà Văn Tấn 1997, p. 584). The Bàu Tró culture is a late Neolithic culture (3000-2500 BCE) found along the coast from Nghệ An to Thừa Thiên Huế today.



Map 2. The distribution of Sấu – Giá and Phú Gia toponyms

The second possibility is contact after Vietnamese had essentially become an independent language. This occurred during Đại Việt's feudal state military campaigns into the Champa region. As a result of these conflicts, a large number of Cham prisoners of war migrated northward, settled, and assimilated with Vietnamese residents in the Red River Delta. This contact is recorded in Vietnamese history, notably in the 11th century when King Lý Thánh Tông resettled prisoners "from Vĩnh Khang prefecture to Đăng Châu, establishing villages with names reminiscent of ancient Chiêm Thành"

(Ngô Sĩ Liên 2010, p. 165).In our latest studies, we focused on four administrative toponyms: *Hà Lầm* (Quảng Ninh province), *Kha Lâm* (Hải Phòng city), *Khả Lãm* (Nghệ An province) and *Hà Lam* (Quảng Nam province). (Trần Thị Hồng Hạnh & Trương Nhật Vinh 2024). Firstly, in terms of geographical location, there is a coincidence as they are all situated near water sources. Hà Lầm is located near Cửa Lục Bay, a small bay within the Hạ Long Bay system. Four small rivers – Diễn Vọng River, Vũ Oai River, Man River, and Trới River – flow into Cửa Lục Bay and then into Hạ Long Bay. *Kha Lâm* is situated along the banks of Lạch Tray River. *Khả Lãm* (or Hương Lãm) is located along the banks of Cả River. *Hà Lam* is positioned between Trường Giang River and Rù Rì River (also known as Ly Ly River). Therefore, whether they are located in coastal areas (Hà Lầm), plains (Kha Lâm, Hà Lam), or foothills (Khả Lãm / Hương Lãm), these place names are all near rivers. This characteristic is crucial for us to further investigate the linguistic features (phonetic forms and meanings) of these place names. Secondly, these toponyms all consist of two syllables. The second syllables "Lam", "Lầm", "Lãm" are quite similar. There is a slight difference in the first syllable between "Hà" and "Kha/Khả". Trần Trí Dõi (2007) proposed that the second syllable in Sino-Vietnamese toponyms often reflects the phonetic form of the whole proto-Vietic syllable but the first syllable in Sino-Vietnamese toponyms often relates to the initial in proto-Vietic syllable. In our survey of several Austronesian languages, we found that the phonetic forms "hule/halau/halow/hlao" all have meanings related to water or water sources.

Table 1 Forms related to water or water sources in some Austronesian languages

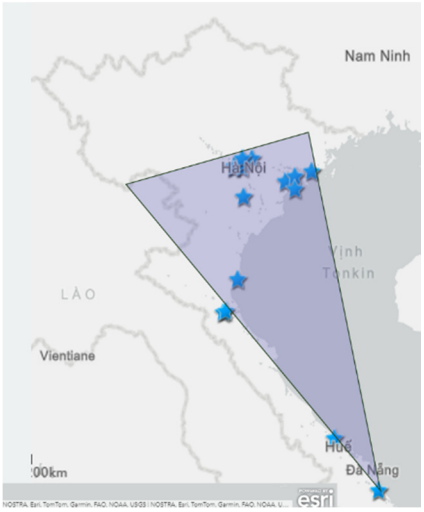
Language	**Word**	**Meaning**
Indonesian	hule/hulu	river source
Raglai	halau	source
Cham	halau/halow	source
Jrai	hlâo	source
Ede	hnoh	stream, source

The phonetic similarities between "hule/halau/halow/hlao" and the toponyms Hà Lầm, Hà Lam, Kha Lâm, and Khả Lãm/Hương Lãm suggest a hypothesis that the first syllable in these toponyms may preserve remnants of the initial part, while the second syllable retains remnants of the latter part of the phonetic forms in Austronesian languages. There is a correspondence in the initial and medial consonant sounds of these phonetic forms. Regarding Kha Lâm and Khả Lãm/Hương Lãm, further discussion is needed to explore the possibility of correspondence between the initial consonant in the first syllable and the initial consonant in the phonetic forms from Austronesian languages. This discussion involves the comparison between /χ/ and /h/. In current Vietnamese phonetics, /χ/ is described as a velar fricative, while /h/ is a

glottal fricative. However, in a historical linguistic study, Nguyễn Tài Cẩn (1995) noted that /h/ in Sino-Vietnamese words originates from two velar fricatives in Chinese, specifically /ɣ/ and /χ/. Nguyễn Tài Cẩn also observed, "In An Nam dịch ngữ, out of 22 cases with initial /h/, it was translated through /χ/ up to 20 times" (Nguyễn Tài Cẩn 1995, p. 100). Therefore, tentative acceptance of the correspondence in the initial consonant can be considered here. The differences in final consonants and tones are understandable and common in the process of Sino-Vietnamese adaptation of toponyms in Vietnam. Furthermore, the meanings of these words in Austronesian languages and the actual geographical locations of these places strongly support our hypothesis. Consequently, it is hypothesized that the toponyms Hà Lầm (Quảng Ninh), Kha Lâm (Hải Phòng), Khả Lãm/Hương Lãm (Nghệ An), and Hà Lam (Quảng Nam) might have originated from Austronesian languages.

Map 3. Distribution of toponyms Hà Lầm, Kha Lâm, Khả Lãm, Hà Lam

Based on the research results mentioned above, the distribution of toponyms hypothesized to have Austronesian origins is represented in maps 3a and 3b below.



Map 3a. Distribution of Austronesian toponyms Map 3b. Distribution of Austronesian toponyms

3. Discussions and suggestions

The distribution shown on map 4a supports Truong Nhat Vinh's claim regarding the origins of the toponyms *Sấu-Giá*, *Phú Gia*, and *Bà Đanh*. These place names likely originated from interactions between Austronesian and Vietnamese communities in Northern Vietnam after the 10th century. In essence, these toponyms can be viewed as manifestations of Austronesian migration from Central Vietnam to the North. However, it is important to note that not all cases necessarily follow to this diffusion pattern. We aim to delve further into the cases of *Kha Lâm* and *Hà Lầm*. In the case of these two toponyms, the distribution depicted in map 3b suggests that they may have originated from a different historical diffusion process. This assumption is supported by additional archaeological evidence.

Map 4. Distribution of toponyms Kha Lâm, Hà Lâm

The ancient village of *Kha Lâm* is located in Nam Sơn ward, Kiến An district, Hải Phòng city. Situated at the base of Đẩu Sơn mountain, which is now part of the expansive Thiên Văn mountain and forest complex, the village is notable for its cultural significance. Within Kha Lâm, a variety of architectural structures including communal houses, temples, and pagodas showcase sophisticated designs. These sites commemorate prominent historical figures who have made significant contributions to the nation, and many of them have been designated as heritage sites. One such example is the Kha Lâm temple dedicated to Princess Chiêu Chinh of the Trần dynasty. *Hà Lầm* is a ward in Hạ Long city, Quảng Ninh province, Vietnam. As mentioned above, *Hà Lầm* area specifically lies on the eastern side of Cửa Lục Bay, a small bay within the Hạ Long Bay system. Both of these toponyms are located within the cultural space of Hạ Long – Hải Phòng, where archaeologists have established that this area is not only a natural wonder but also a cradle of prehistoric civilization. Numerous archaeological research has revealed that stone tools, ceramic containers, stone jewelry, and bones found and collected in Ha Long all date back to the late Neolithic period. Among them, an important study by Heine Geldern from the early 20th century can be mentioned.

Heine Geldern synthesized all contemporary knowledge about surveying archeology, ethnology, and linguistics to build a panorama of Southeast Asia. He placed particular emphasis on the study of cultures and populations, developing a model grounded in diffusionism. Heine Geldern focused on elucidating two pivotal stages in the cultural evolution of Southeast Asia: the Neolithic and Metal Ages. According to his research, the late Neolithic period in Southeast Asia witnessed the emergence of three successive cultural phases among distinct ethnic groups: the cylindrical axe culture (Walzenbeilkultur) among Papuan-speaking populations, the shoulder axe culture (Schulterbeilkultur) among Austronesian-speaking groups, and the quadrilateral axe culture (Vierkantbeilkultur) also among Austronesian speakers. Geldern's scholarly attention was particularly drawn to the migratory patterns of Austronesian peoples originating from China, dispersing across Indochina, Malaysia, and into the maritime regions of Southeast Asia and Oceania. He argued that each of these cultural groups exhibited distinct material and spiritual elements, thereby suggesting that cultural transformations in this region occurred primarily through diffusion and replacement processes (See Phạm Đình Mạnh 2007). While his opinions exerted a profound and enduring influence on contemporary prehistorians and successive generations of his students, they are now considered untenable. However, our special attention is given to his views because they include mention of shoulder axe culture (Schulterbeilkultur) and the quadrilateral axe culture (Vierkantbeilkultur) among the Austronesian speakers. Heine Geldern believed that the owner of the quadrangular ax culture in Southern China was of Austronesian origin. From here they migrated down the Southeast Asian

continent to the Malay peninsula, through Sumatra to Java, to the eastern islands of Indonesia, and then to the Indo-Pacific region.

In Vietnam, these two types of axes, shoulder axe culture (Schulterbeilkultur) and the quadrilateral ax culture (Vierkantbeilkultur), have been discovered in many archaeological sites across the country, including in Hạ Long. At Hạ Long site, archaeologists also found stone axes with shoulders with small shapes and sizes resembling those found in Guangdong, Fujian and Hong Kong. According to Trình Năng Chung (2008), the late Neolithic period in the coastal region of Southeast China is characterized by distinctive cultural artifacts such as printed ceramics, axes, shouldered stone axes, and stepped stone axes. Scholars have identified a close association between this region and the Ha Long culture, which is found in the coastal plains and islands of Bái Tử Long and Hạ Long bays. In Guangxi province, the late Neolithic period is marked by the prevalence of the large stone shovel culture, primarily in the Guinan region. Trình Năng Chung suggested multiple instances of interaction between the Neolithic communities of Guinan and the Hạ Long culture inhabitants. He proposes that these interactions occurred via maritime routes, while with residents of Mai Pha crossing the Kỳ Cùng River and those from Hà Giang using the Bằng River and Gâm River passages (Trình Năng Chung 2008, p. 65). Trình Năng Chung (2008) stated that the stone shovel relics discovered in southern China date back mainly to the late Neolithic period, spanning from approximately 5,000 years ago to the Western Han period (2nd century AD). Vietnamese archaeological records document 37 instances of stone shovel relics across seven northern mountainous provinces and the northeastern coastal region of Vietnam. In the Hạ Long cultural area, the stone shovel are considered products of interaction and exchange. Due to its natural geographical location, this region facilitated extensive multi-directional contacts between the creators of the large stone shovel culture in southern China and the inhabitants of the Hạ Long culture. These interactions primarily occurred via maritime routes and secondarily through overland routes originating from the Lang Son area.

Lương Ninh (2010) asserted that archaeologists generally place the origins of the Austronesian people in a period dating back to 8,000-6,000 years BCE, with some proposing even earlier dates, while the closest evidence suggests a timeframe around 5,000-4,000 years BCE. However, archaeological findings in Vietnam indicate developments dating from approximately 3,000 years BCE, with peak concentrations around 500 years BCE. A. Reid has noted the involvement of the Chamic people in these migrations, but regardless of their origins, whether from the coast of Vietnam or elsewhere, they remained within the Malayo-Polynesian group until their transition to Cham-Malayo-Chamic, which likely occurred no earlier than the beginning of the Common Era, specifically around the 2nd century CE. Many scholars, including

linguist Blust, suggest that the modern Malayo-Chamic groups were present in Vietnam and Aceh from the 4th to 3rd centuries BCE. (Lương Ninh 2010, pg.243).

Due to the limited number of toponyms analyzed in this paper, an exhaustive discussion on the migration of Austronesian peoples to Southeast Asia is not feasible. However, based on the distribution patterns of these two toponyms and supported by archaeological evidence, it is hypothesized that there might be a group of Austronesian inhabitants migrated to Vietnam during the late Neolithic period via maritime routes from southeastern China. However, investigating the anthropology, history, and culture of ethnic groups living in the territory of Vietnam in particular, in Southeast Asia is an interesting topic but cannot be solved by archeology or linguistics alone. It is necessary to apply genetic technology. Therefore, this study, conducted as a case discussion, emphasizes the need for further research to achieve a comprehensive understanding of this topic.

References

Alves, Mark J (2021) The Đông Sơn Speech Community: Evidence for Vietic. *Crossroads* 19(2): 138–174. https://doi.org/10.1163/26662523-bja10002

Gourou, Pierre (1936) *Les paysans du delta Tonkinois: étude de géographie humaine*. Translated into Gourou, Pierre (2016) Người nông dân châu thổ Bắc Kỳ: Nghiên cứu địa lý nhân văn [The farmers of the Tonkinese delta: study of human geography]. Thành phố Hồ Chí Minh: Nhà xuất bản Trẻ.

Hà, Văn Tấn (1997) *Theo dấu các văn hóa cổ* [Following in the footsteps of ancient cultures]. Hà Nội: Nhà xuất bản Khoa học Xã hội.

Hoàng, Thị Châu (2014) *Hợp lưu những dòng suy tư về địa danh, phương ngữ và ngôn ngữ các dân tộc thiểu số* [Confluence of thoughts about places, dialects and languages of ethnic minorities], Hà Nội: Nhà xuất bản Đại học Quốc gia Hà Nội.

Lương, Ninh (2010) Người Nam Á, Nam Đảo và sự hình thành các quốc gia cổ ở Việt Nam, *Kỷ yếu Hội thảo quốc tế Việt Nam học lần thứ ba* [Austroasiatic, Austronesians and the formation of ancient nations in Vietnam, Proceedings of the Third International Conference on Vietnamese Studies], 235–244. Hà Nội: Nhà xuất bản Đại học Quốc gia Hà Nội.

Ngô, Sĩ Liên (2010) *Đại Việt sử ký toàn thư: bản in nội các quan bản mộc bản khắc năm Chính Hòa thứ 18 (1697)* [The Complete Annals of Đại Việt: printed edition, imperial officials' original documents, engraved woodblock edition of the 18th year of Chính Hòa era (1697)."]. Hà Nội: Nhà xuất bản Khoa học xã hội.

Nguyễn, Tài Cẩn (1995) *Giáo trình lịch sử ngữ âm tiếng Việt (sơ thảo)* [Textbook on the History of Vietnamese Phonetics (preliminary)], Hà Nội: Nhà xuất bản Giáo dục.

Nguyễn,Trung Chiến (2015) *Tiếp cận văn hóa biển tiền sử Việt Nam qua các bài nghiên cứu* [Approaching the prehistoric maritime culture of Vietnam through various research studies]. Hà Nội: Nhà xuất bản Khoa học xã hội

Nguyễn, Văn Huyên (1944) *Văn minh Việt Nam* [A civilisation annamite]. Translated by Đỗ Trọng Quang. 2016. Hà Nội: Nhã Nam and Nhà xuất bản Hội Nhà văn.

Phạm, Đức Mạnh (2007) Lịch sử văn hóa vật chất thời tiền sử Đông Nam Á - Một thế kỷ điền dã và liên hiệp nghiên cứu [History of prehistoric material culture in Southeast Asia - A century of fieldwork

and joint research]. *Tạp chí phát triển khoa học và công nghệ* [Journal of Science and technology development] 10(9): 36–50.

Taylor, Keith Weller (1983) *The birth of Vietnam*. University of California Press.

Trần, Quốc Vượng (2004) Dấu vết người Chăm và văn hóa Chăm Pa ở miền Bắc Việt Nam, *Kỷ yếu Hội thảo khoa học quốc tế Việt Nam học lần thứ hai*, 11–28. NXB Đại học Quốc gia Hà Nội.

Trần, Thị Hồng Hạnh (2019) Identifying the origin of toponyms in Vietnam, *Proceedings of the First Annual Meeting of the Geolinguistic Society of Japan*, 91-95. https://doi.org/10.5281/zenodo.4505652

Trần, Thị Hồng Hạnh & Trương, Nhật Vinh (2023) Vấn đề tiếp xúc ngôn ngữ qua một vài cứ liệu địa danh theo hướng gợi mở của Giáo sư Hoàng Tuệ [The issue of language contact through some toponyms data as suggested by Professor Hoang Tue]. *trong Giáo sư Hoàng Tuệ - Ngôn ngữ trong đời sống xã hội* [Professor Hoang Tue - Language in social life], 130–138. Hà Nội: Nhà xuất bản Dân trí.

Trần, Trí Dõi (2001a) Về một vài địa danh, tên riêng gốc Nam Đảo trong vùng Hà Nội xưa. In trong Hà Nội - Những vấn đề ngôn ngữ, văn hóa (Hội ngôn ngữ học Hà Nội) [About some toponyms and proper names of Austronesian origin in the ancient Hanoi area. In Hanoi Language and cultural issues. Edited by Hanoi Linguistics Association]. Hà Nội: Nhà xuất bản Thời Đại.

Trần, Trí Dõi (2001b) Ngôn ngữ và sự phát triển văn hóa xã hội [Language and sociocultural development]. Hà Nội: Nhà xuất bản Văn hóa Thông tin.

Trần, Trí Dõi (2022) Lịch sử ngôn ngữ người Việt – Góp phần tìm hiểu văn hóa Việt Nam [History of Vietnamese language - Contributing to understanding Vietnamese culture]. Hà Nội: Nhà xuất bản Đại học Quốc gia Hà Nội.

Trình, Năng Chung (2008) Mối quan hệ văn hóa giai đoạn hậu kỳ đá mới giữa Bắc Việt Nam và Nam Trung Quốc [Cultural relations during the late Neolithic period between Northern Vietnam and Southern China]. *Tạp chí nghiên cứu Trung Quốc* [*Journal of Chinese Studies*] 8(87): 64–73.

Trương, Nhật Vinh (2020) Tên gọi làng Phú Gia (Hà Nội). *Tạp chí Ngôn ngữ & đời sống* [The name of Phú Gia village. Journal of Linguistics and Life]. 6: 18–25.

Trương, Nhật Vinh (2023) Explanation of the Name "Bà Đanh" Pagoda. In *Researching and Applying Linguistics and Vietnamese Language Studies*, 311–321. Geolinguistic Society of Japan. https://doi.org/10.5281/zenodo.8153120

Trương, Nhật Vinh (2024) Thảo luận về địa danh "Sấu – Giá" (Hoài Đức, Hà Nội). *Tạp chí Ngôn ngữ* [A discussion of the toponyms Sấu - Giá. Journal of Linguistics]. 1: 57–68.

Publication history

Date received: 9 July 2024
Date accepted: 20 August 2024

学術論文

交流度による入声調消失の要因の探求 —中国汾河流域における２種類のデータ分析—

沈力
同志社大学

鄭弯弯
名古屋大学

Exploring the causes of entering tone loss based on interaction density: An analysis on two datasets of the Fen River Basin, China

SHEN, Li
Doshisha University

ZHENG, Wanwan
Nagoya University

Abstract: This study investigated the quantitative changes in the retention of entering tone words in the Lingshi Highlands, which are located in the central Fen River basin of Shanxi Province, China. Our findings suggested that the entering tone is progressively decreasing at varying rates across five dialect regions. However, it remained unclear whether this decrease was due to the influence of the central dialect or the independent development of each local dialect, as well as what factors contributed to the discrepancy in the rate of decrease. To address these issues, this study proposed that the interaction density determines the extent of linguistic change. The interaction density serves as a quantification of the frequency at which people from diverse dialect regions engage in communication with those from central region. The results revealed that the interaction density ranking of each dialect region in the Lingshi Highlands, and the central dialect region aligned with the ranking of number of entering tone words remaining in these regions. This suggested that the disappearance of the entering tone in Middle Chinese was not caused by independent changes within each dialect, but rather by the influence of the central dialect. Furthermore, the hypothesis that the interaction density determines the extent of linguistic change was supported by the data obtained from a follow-up survey in the same Lingshi Highlands in 2022.*

キーワード：汾河流域；方言区；言語変化；入声；交流度
Keywords: Fen River Basin; dialect region; linguistic change; entering tone; interaction density

沈力、鄭弯弯 (2024)「交流度による入声調消失の要因の探求—中国汾河流域における２種類のデータ分析—」『地理言語学研究』4: 119–141. doi: https://doi.org/10.5281/zenodo.13948584
* 本論文は、2018年-2021年・国家社会科学基金一般項目（中国、課題番号18BYY042）と2019年-2022年・日本学術振興会科研費基盤研究B（日本、課題番号JP19H01261）の助成を受けたものであり、それらの研究成果の一部である。

1. はじめに

言語変化の原因がその言語内部にあるのか、他の言語の影響によるものなのかを判断することは、歴史比較言語学において非常に重要な課題である。例えば、A 言語に従来から存在していた現象αが、近年、隣接する B 言語でも見られるようになったとする。このとき、B 言語の現象αは、A 言語から借用されたものなのか、B 言語内部で自発したものなのかを判断する必要がある。A 言語と B 言語がもともと異なる音韻・文法システムを持っていれば、その判断は比較的簡単かもしれない。しかし、方言同士の関係のように、A 言語と B 言語がかつて同様の音韻・文法システムを持っていた場合、その判断は非常に難しくなる。図 1 に示すように、本研究では、まず、中古漢語の入声調（以下「ET」と呼ぶ）は、中国山西省にある汾河流域中部の晋中盆地には依然として存在しているが、汾河流域南部の臨汾盆地には存在せず、舒声調（以下「RT」と呼ぶ）に変わっていることに注目する[1]。

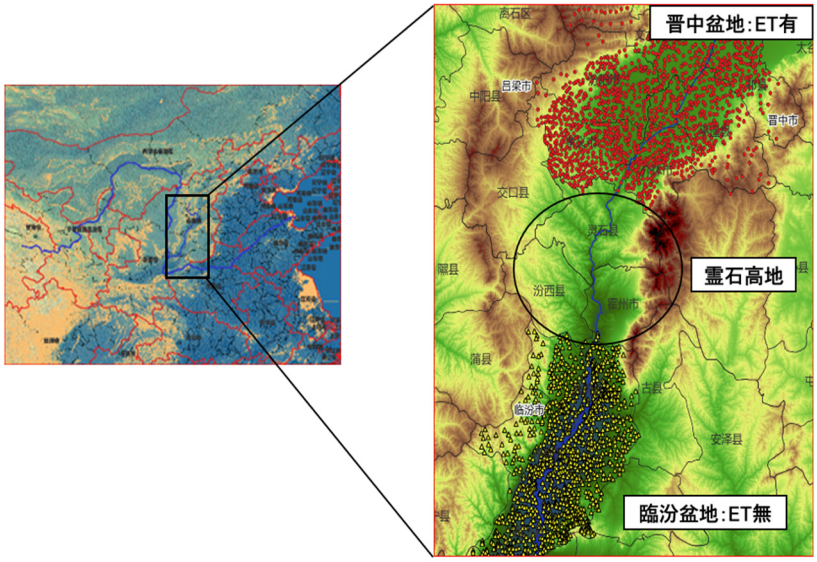


図1：汾河流域における二大盆地のET分布図

図 1 において、左図は、中国における汾河流域諸方言の位置を示しており、右図は、汾河（左図にある青線）でつながった二大盆地における ET の有無を示している。右図上部の赤点で示されている方言群は、ET が保持されている晋方言であり、右図下部の黄色点で示されている方言群は、RT しか持たない

[1] 入声調は英語で Entering Tone、舒声調は英語で Relaxed Tone と呼ばれているため、本稿では前者を ET、後者を RT と略称する。

中原官話である[2]。さらに、図 1 より、霊石高地（霍州市、汾西県、霊石県が含まれる）は、この二大方言地域の中間に位置していることがわかる。ET の有無で対立している晋方言と中原官話は、霊石高地で接触することになるが、このような接触地域において、ET がどの程度残存しているのか、および残存している ET はどのように分布しているのかが問題となる。

筆者は、2011 年に、中古漢語の ET 語[3]（以下「ET 語」と呼ぶ）において ET がどれほど残存しているのかを調査し、調査データに基づいた研究成果を発信している（以下「2011 年データ」と呼ぶ）。また、図 2 に示すように、Shen & Nakano (2015)では、ET の残存数に基づけば、霊石高地の様々な方言は、大きく 5 つの方言に区分できるということを報告している。

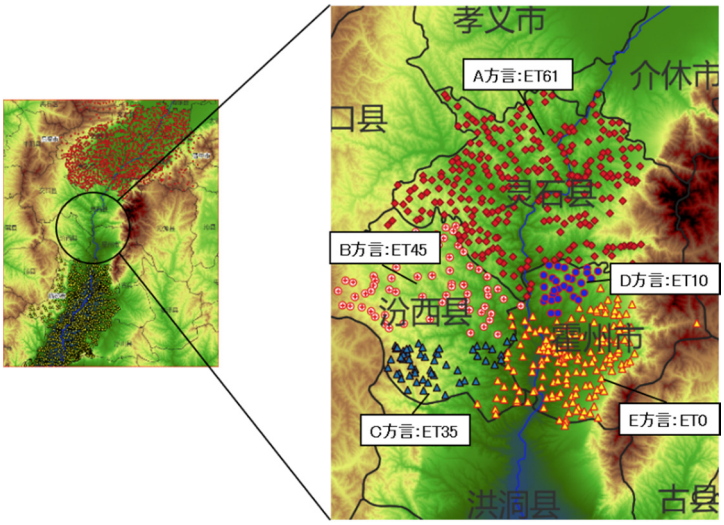


図2： ET残存数による方言区画

図 2 で、A 方言（霊石県）には、小王庄方言、崔家溝方言、軍営坊方言、段純方言、南浦方言、上黄堆方言、城関方言、王禹方言、仁義方言、南関方言の 10 種類の方言が含まれる。これらの方言には、64 の ET 語のうち 61 以上の語で ET が残存している。B 方言（汾西県北部）には、西河方言、勍香方言、永安鎮方言の 3 種類の方言が含まれる。これらの方言には 40 程度の語で ET が残存している。C 方言（汾西県南部）には、邢家要方言があり、35 程度の語で ET が残存している[4]。D 方言（霍州市北部）には、師庄方言があり、ET の残存数は 10 しかない。E 方言（霊石高地の最南端、汾西東南部、霍州市南部）

[2] 晋方言と中原官話の区画に関しては、侯・温（1993）を参照されたい。
[3] 本稿のET語は漢語中古音の入声語を指す。
[4] ET残存数についてShen & Nakano (2015) のデータを使用しているが、計算基準を変更した。Shen & Nakano (2015:86) では「A方言：60、B方言：40、C方言：30、D方言：10、E方言：0」のように十進法のようにまとめたが、今回は実数「A方言：61、B方言：45、C方言：35、D方言：10、E方言：0」のように修正した。

には、汾西下団柏方言、汾西申村方言、霍州城関方言、霍州三教方言の 4 種類の方言が含まれる。これらの方言では、ET が完全に消失している。

前述の A〜E 方言における ET 残存数と分布の関係は、図 3 のようになる。

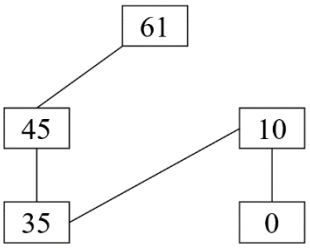


図3： ET残存数の地理的な曲線分布

図 3 より、ET 語の ET は霊石高地で減少しており、その減少傾向は地理的に曲線を描いているということがわかる。ここで 2 つの問題が生じる。第一に、この減少傾向は、霊石高地の様々な方言の内部で自発的に生じたものなのか、中原官話との接触によって生じたものなのかという問題である。第二に、ET の減少は、なぜ直線型ではなく曲線型なのかという問題である。本研究では、交流度の提案に基づいて、これら 2 つの問題を解明する。

2. 霊石高地諸方言におけるET残存の実態

晋中盆地の諸方言と臨汾盆地の諸方言が ET と RT の対立を見せることは、前述の通りである。この二大方言に挟まれた霊石高地の様々な方言における ET の残存状況と分布には、興味深い点がある。本章では、霊石高地の様々な方言における ET 語の残存実態を報告する。

本研究で使用する調査語は、『方言調査字表（3810 字）』（中国社会科学院語言研究所編, 1964）掲載の ET 語（67 語）から選定した 64 語である。中古漢語以降、ET は声母（頭子音）の全濁・次濁・清音の分類に基づいて 3 つに分裂しているため、選定された ET 語はこの分裂の実態を考慮した単語リストとなっている[5]。調査語の詳細を表 1 に示す。

[5] Karlgren (1915−26) によれば、漢語の声調はもともと「平上去入」の四声調体制であったが、後に声母の清濁対立に基づいて陰陽に分裂し、最終的には八声調体制になるという仮説を打ち出している。

表１：64の調査語の分類、語数および漢語中古音[6]

声母類	調査語数	調査語例（漢語中古音）
全濁	18	合(*ɣɒp)、 罰(*bĭwɐt)、毒(*dʻuok)
次濁	9	納(*nɒp)、 月(*ŋĭwɐt)、麦(*mæk)
清音	37	急(*kĭĕp)、 八(*pæt)、 得(*tək)

表 1 より、漢語の頭子音は有声性と共鳴性に基づいて、全濁・次濁・清音の 3 種類に分類される[7]。64 の調査語の中で「合」のような全濁声母を持つ調査語は 18 個、「納」のような次濁声母を持つ調査語は 9 個、「急」のような清音声母を持つ調査語は 37 個ある。

2.1. 2011年データにおけるET語のET残存実態

筆者は、2011年に、霊石高地における606の村落で調査を実施した。調査結果より、霊石高地諸方言のETは、中古声母の全濁・非全濁を基準に分化していること、およびETの残存数によって5つの方言に区分できることが明らかになった[8]。調査結果の詳細は次の通りである。

まず、A方言（霊石県）におけるETの残存状況を表2に示す。

表 2：霊石南浦方言の ET 残存一覧表

1．ET 残存例（63 例）							
全濁 ET	pʻaʔ212u(拔)	sæʔ212u(実)	sæʔ212u (十)	tuʌʔ212u(毒)	xuʌʔ212u(服)	sʌʔ212u(食)	eyʌʔ212u(俗)
	xaʔ212u(合)	xuʌʔ212u(罰)	sæʔ212u(舌)	tuʌʔ212u(読)	tsuaʔ212u(濁)	tiʌʔ212u(笛)	tsʻaʔ212u(雑)
	sʌʔ212u(石)	tsaʔ212u(宅)	piaʔ212u(白)	teyʌʔ212u(局)			
非全濁ET	zuæʔ4u(入)	yʌʔ4u(薬)	miʌʔ4u(麦)	yæʔ4u(月)	miʌʔ4u(滅)	uʌʔ4u(袜)	iəʔ4u(逸)
	nʌʔ4u(納)	tɕiəʔ4u(急)	tɕiəʔ4u(接)	tɕʻiəʔ4u(七)	tɕʻiəʔ4u(切)	tiʌʔ4u(滴)	piəʔ4u(筆)
	piaʔ4u(百)	suaʔ4u(説)	paʔ4u(八)	taʔ4u(搭)	tsaʔ4u(窄)	iəʔ4u(一)	pʻiəʔ4u(匹)
	tʻiəʔ4u(鉄)	ɕiʌʔ4u(錫)	tɕʻyʌʔ4u(曲)	ŋyʌʔ4u(約)	xuʌʔ4u(福)	tʻuʌʔ4u(禿)	tsuʌʔ4u(竹)
	tʌʔ4u(得)	tsʻʌʔ4u(尺)	tsʌʔ4u(織)	eyʌʔ4u(削)	tuʌʔ4u(督)	tɕiʌʔ4u(積)	tsʻuæʔ4u(出)
	sæʔ4u(失)	tɕʻyæʔ4u(缺)	xuaʔ4u(髪)	xaʔ4u(黒)	kaʔ4u(割)	ɕiaʔ4u(歇)	pʻiaʔ4u(拍)
	tsuaʔ4u(桌)	sæʔ212u(湿)	sʌʔ212u(識)				
2．舒声化例（1 例）							
非全濁 ET	liu53(六)						

表 2 より、霊石南浦方言（A 方言）の ET は、中古声母の全濁・非全濁を基準に分化していることが明らかになった。中古で全濁声母を有した ET（以下「全濁 ET」と呼ぶ）の調値は 212u であり、それ以外の ET（以下「非全濁 ET」

[6] 漢語中古音の再建は王（2008）によるものである。

[7] 中古漢語の頭子音はまた全清（無声無気）と次清（無声有気）に分かれているが、声調の分裂には特に関与しないため省略する。

[8] 声調調値は五線譜のように5段階表記法（1[低]～5[高]）に基づいて記述している。また、入声(ET)は中国方言界で数字の下にアンダーラインをつけて短促性を表す習慣があることから、ここでアンダーラインの代わりにuで示すことにする。

と呼ぶ）の調値は 4u である。また、64 の ET 語のうち、舒声化した語は「liu53（六）」の 1 例しかなかった[9]。

次に、B 方言（汾西県北部）における ET の残存実態を表 3 に示す。

表3：汾西永安方言のET残存一覧表

1．ET残存例（45例）							
全濁ET	pʻəʔ3u(拔)	səʔ3u(実)	səʔ3u (十)	tʻuəʔ3u(毒)	fəʔ3u(服)	səʔ3u(食)	ɕyəʔ3u(俗)
	xəʔ3u(合)	fəʔ3u(罰)	səʔ3u(舌)	tʻuəʔ3u(読)	tsuəʔ3u(濁)	tʻiəʔ3u(笛)	tsʻəʔ3u(雑)
	səʔ3u(石)	tsʻəʔ3u(宅)	pʻəʔ3u(白)	tɕʻyəʔ3u(局)			
非全濁ET	vəʔ21u(入)	iəʔ21u(逸)	miəʔ21u(麦)	lyəʔ21u(六)	miəʔ21u(滅)	səʔ21u(識)	fəʔ21u(福)
	tsʻuəʔ21u(出)	səʔ21u(失)	tɕʻiəʔ21u(七)	tɕʻiəʔ21u(切)	tiəʔ21u(滴)	piəʔ21u(筆)	pʻiəʔ21u(匹)
	təʔ21u(得)	ɕiəʔ21u(錫)	tɕʻyəʔ21u(曲)	tsʻəʔ21u(尺)	səʔ21u(湿)	tʻuəʔ21u(禿)	tsuəʔ21u(竹)
	tɕiəʔ21u(積)	xəʔ21u(黒)	tsəʔ12u(織)	suəʔ21u(説)	tɕiəʔ3u(急)	iəʔ3u(一)	
2．舒声化例（19例）							
非全濁ET	na21(納)	yɪ21(月)	iu21(薬)	va21(袜)	tsɪ21(窄)	tsu21(桌)	tʻiɪ21(鉄)
	pɪ21(百)	ɕiu21(削)	pɑ21(八)	ta21(搭)	pʻɪ21(拍)	kəu21(割)	ɕiã21(歇)
	tɕiɪ21(接)	tɕʻyɪ21(缺)	niu21(約)	fa21(髪)	tβ21(督)		

表 3 より、汾西永安方言（B 方言）では、全濁 ET の調値は 3u であり、それ以外の ET の調値は 1u である。また、64 の ET 語のうち、舒声化した語は 19 例ある。B 方言は、A 方言と同様、舒声化したのは非全濁 ET 語のみであるが、それらの ET が陰平調（11 調）に合流していることが特徴である[10]。

次に、C 方言（汾西西南部）における ET 語の残存実態を表 4 に示す。

表4：汾西邢家要方言のET残存一覧表

1．ET残存例（35例）							
全濁ET	səʔ3u(舌)	səʔ3u(実)	səʔ3u (十)	səʔ3u(食)	səʔ3u(石)	fuəʔ3u(服)	tʻuəʔ3u(毒)
	tʻəʔ3u(読)						
非全濁ET	yeʔ1u(月)	iəʔ1u(逸)	iəʔ1u(薬)	miəʔ1u(滅)	tsʻuəʔ1u(出)	səʔ1u(失)	iəʔ1u(一)
	səʔ1u(識)	fuəʔ1u(福)	tɕʻiəʔ1u(七)	tsɛʔ1u(窄)	tiəʔ1u(滴)	piəʔ1u(筆)	tsuəʔ1u(桌)
	təʔ1u(得)	ɕieʔ1u(錫)	tɕʻyəʔ1u(曲)	taʔ1u(搭)	pɛʔ1u(百)	tɕiəʔ1u(急)	tsuəʔ1u(竹)
	tɕiəʔ1u(積)	kəʔ1u(割)	tsəʔ1u(織)	tɕieʔ1u(接)	ioʔ1u(約)	tʻuəʔ3u(禿)	
2．舒声化例（29例）							
全濁ET	pʻa33(拔)	ɕy33(俗)	tʻi33(笛)	tsʻa33(雑)	xə33(合)	tsɛ33(宅)	pʻɛ33(白)
	tɕyo33(局)	fa33(罰)	tsuə33(濁)				
非全濁ET	na11(納)	mɛ11(麦)	lyo11(六)	va11(袜)	və11(入)	sə11(湿)	tʻiə11(鉄)
	tɕʻiɛ11(切)	ɕyo11(削)	pa11(八)	pʻi11(匹)	tsʻə11(尺)	pʻɛ11(拍)	ɕie11(歇)
	xɯ11(黒)	tɕʻyo11(缺)	fɛ11(説)	fa11(髪)	tu11(督)		

[9] 舒声化は、中古漢語ではETであった語がRTになっていることを指す。「liu53（六）」の53調は南浦方言では去声調である。

[10] この地域の諸方言の平声調も全濁声母を持つ平声調と非全濁声母を持つ平声調に分かれている。通常、前者は陽平調、後者は陰平調と呼ばれている。

表 4 より、汾西邢家要方言（C 方言）では、全濁 ET の調値は 3u であり、それ以外の ET の調値は 1u である。また、64 の ET 語のうち、舒声化した語が 29 例ある。C 方言は、A 方言や B 方言と異なり、全濁 ET も非全濁 ET も舒声化しているが、前者は上声調（33 調）に合流し、後者は陰平調（11 調）に合流している。

次に、D 方言（霍州市北部）における ET の残存実態を表 5 に示す。

表5：霍州師庄方言のET残存一覧表

1．ET 残存例（10 例）							
全濁 ET	ʂɤʔ5u (十)	ʂɤʔ5u(食)	ʂɤʔ5u(石)	ʂɤʔ5u(実)			
非全濁 ET	yɤʔ21u(月)	zɣʔ21u(入)	tiɤʔ21u(得)	tsʻuɤʔ21u(出)	ʂɤʔ21u(失)	xuɤʔ21u(髪)	
2．舒声化例（54 例）							
全濁 ET	pʻɑ35(拔)	ɕiu35(俗)	tɕʻi35(笛)	tsʻɑ35(雑)	xɤ35(合)	tsɑi35(宅)	pʻie35(白)
	tɕiu35(局)	fɑ35(罰)	tʂʻuə35(濁)	ʂie35(舌)	tʻu35(毒)	fu35(服)	tu35(読)
非全濁ET	lɑ212(納)	mie212(麦)	liu212(六)	vɑ212(袜)	mie212(滅)	iuɤ212(薬)	kɤ212(割)
	tɕʻie212(切)	ɕiuɤ212(削)	pɑ212(八)	pʻi212(匹)	tʂʻʅ212(尺)	ʂʅ212(湿)	tɕʻie212(鉄)
	xei212(黒)	tɕʻye212(缺)	ʂuɤ212(説)	tʻu212(禿)	tu212(督)	pʻie212(拍)	ɕie212(歇)
	ʂʅ212(識)	fu212(福)	tɕʻi212(七)	tsɤ212(窄)	tie212(滴)	pi212(筆)	tsuɤ212(桌)
	tʂʅ212(織)	ɕi212(錫)	tɕʻiu212(曲)	tɑ212(搭)	pie212(百)	tɕie212(接)	tʂu212(竹)
	tɕi212(積)	i35(一)	tɕi35(急)	ȵiuɤ35(約)	i35(逸)		

表 5 より、霍州師庄方言（D 方言）では、全濁 ET の調値は 5u であり、それ以外の ET の調値は 21u である。また、64 の ET 語のうち、ET は 10 のみ残存している。舒声化した ET のうち、全濁 ET は陽平調（35 調）に合流し、非全濁 ET は陰平調（212 調）に合流している。

最後に、ET が全く残っていない E 方言（霍州南部）における ET の残存実態を表 6 に示す。

表6：霍州城関方言のET残存一覧表

1．ET 残存例（0 例）							
2．舒声化例（64 例）							
全濁ET	pʻɑ35(拔)	ɕiu35(俗)	tɕʻi35(笛)	tsʻɑ35(雑)	xɤ35(合)	tsɑi35(宅)	pʻie35(白)
	tɕiu35(局)	fɑ35(罰)	tsuo35(濁)	ʂie35(舌)	tʻu35(毒)	fu35(服)	tu35(読)
	ʂʅ35 (十)	ʂʅ35(食)	ʂʅ35(石)	ʂʅ35(実)			
非全濁ET	lɑ212(納)	mie212(麦)	liu212(六)	vɑ212(袜)	mie212(滅)	iuɤ212(薬)	ʐu212(入)
	ye212(月)	tie212(得)	tsʻu212(出)	ʂʅ212(失)	fɑ212(髪)	tɕi212(積)	kɤ212(割)
	tɕʻie212(切)	ɕiuɤ212(削)	pɑ212(八)	pʻi212(匹)	tʂʻʅ212(尺)	ʂʅ212(湿)	tɕʻie212(鉄)
	xei212(黒)	tɕʻye212(缺)	ʂuɤ212(説)	tʻu212(禿)	tu212(督)	pʻie212(拍)	ɕie212(歇)
	ʂʅ212(識)	fu212(福)	tɕʻi212(七)	tsɤ212(窄)	tie212(滴)	pi212(筆)	tsuɤ212(桌)
	tʂʅ212(織)	ɕi212(錫)	tɕʻiu212(曲)	tɑ212(搭)	pie212(百)	tɕie212(接)	tʂu212(竹)
	tɕi35(急)	ȵiuɤ35(約)	i35(一)	i35(逸)			

表 6 より、霍州城関方言（E 方言）では、D 方言と同様、全濁 ET は陽平調（35 調）に合流し、非全濁 ET は陰平調（212 調）に合流している。一方、E 方言では、D 方言と異なり、全濁 ET 語の ET も、それ以外の ET 語の ET も残存していない。

以上の調査結果は、表 7 のように一覧できる。

表7：霊石高地諸方言の共通点と相違点

	共通点：声母が全濁か否かによる分裂		相違点：ET 残存数	
	全濁 ET 調値	非全濁 ET 調値	内訳	合計
A 方言	212u（→NA）	4u（→NA）	18+45	63
B 方言	3u（→陽平）	21u（→陰平）	18+27	45
C 方言	3u（→上声）	1u（→陰平）	8+27	35
D 方言	5u（→陽平）	21u（→陰平）	4+6	10
E 方言	NA（→陽平）	NA（→陰平）	0+0	0

表 7 より、霊石高地諸方言には、中古 ET が声母の全濁・非全濁を条件に 2 声調に分裂したという共通点と、ET の残存数が徐々に減少しているという相違点がある。A 方言は少なくとも 61 例[11]、B 方言は 45 例、C 方言は 35 例、D 方言は 10 例、E 方言は 0 例である。

最後に、舒声化した ET がどの舒声調に合流しているかについて、霊石高地のすべての方言では、非全濁 ET が舒声化する場合、必ず陰平（非全濁声母を持つ平声調）に合流している。一方、全濁 ET が舒声化する場合、B、D 方言の ET は陽平（全濁声母を持つ平声調）に合流している。しかし、E 方言の舒声化した調類が B、D 方言と同様であるため、E 方言もかつて同じ基準で ET が分裂していたと考えられる。具体的には、E 方言も全濁 ET が全濁平声調（RT）に、非全濁 ET が非全濁平声調（RT）に合流するという変化規則を持っていたと推測される。また、C 方言の全濁 ET は上声調に合流している。上声調と合流した理由は不明であるが、調値の類似性（3u→33）による可能性がある[12]。

2.2. 霊石高地諸方言と周辺方言との比較

本節では、霊石高地諸方言が、晋中盆地と臨汾盆地の諸方言と共通している点について観察する。まず、臨汾盆地諸方言では、ET は声母の全濁・非全濁を基準として 2 声調に分裂している。その後、分裂した ET は完全に舒声化し、

[11] 南浦方言のET残存数は63であるが、A方言全体から見れば、ETが61例というケースもある。

[12] A方言では舒声化した例が1例しかないため、A方言におけるETの舒声化傾向についての考察は控えておく。

陽平調（24 調）と陰平調（21 調）に合流している。この事実は、次に示す表 8 より一覧できる。なお、表 8 では、臨汾盆地の中心方言である臨汾屯里方言の ET 語の舒声化を示す。

表 8： 臨汾屯里方言の ET 残存一覧表

1．ET 残存例（0 例）							
2．舒声化例（64 例）							
全濁ET	pʻa24(拔)	su24(俗)	tʻi24(笛)	tsʻa24(雑)	xuo24(合)	tʂai24(宅)	pʻe24(白)
	tɕy24(局)	fa24(罰)	tʂuo24(濁)	ʂɤ24(舌)	tu24(毒)	ʂʅ24(実)	xua24(読)
	ʂʅ24(十)	ʂʅ24(石)	ʂʅ21(食)	kuə̃21 (服)			
非全濁ET	na21(納)	me21(麦)	ly21(六)	va21(袜)	mie21(滅)	yo21(薬)	vu21(入)
	ye21(月)	tei21(得)	fu21(出)	ʂʅ21(失)	fa21(髪)	tɕi21(積)	kuo21(割)
	tɕʻie21(切)	ɕye21(削)	pa21(八)	pʻi21(匹)	tʂʻʅ21(尺)	ʂʅ21(湿)	tʻie21(鉄)
	xɯ21(黒)	tɕʻye21(缺)	fuo21(説)	tʻu21(禿)	tu21(督)	pʻe21(拍)	ɕie21(歇)
	ʂʅ21(識)	fu21(福)	tɕʻi21(七)	tʂɤ21(窄)	ti21(滴)	pi21(筆)	pfuo21(桌)
	tʂʅ21(織)	ɕi21(錫)	tɕʻy21(曲)	ta21(搭)	pe21(百)	tɕie21(接)	yo21(約)
	tɕi24(急)	i24(一)	pfu31(竹)	i53(逸)			

表 8 より、臨汾盆地にある屯里方言の ET 語は、霍州南部にある E 方言と同様、完全に舒声化していることがわかる。また、64 の調査語のうち、全濁 ET 語の ET は陽平調（24 調）に、非全濁 ET 語の ET は陰平調（21 調）に合流していることがわかる。霊石高地諸方言の ET 語が、臨汾盆地諸方言と同様の ET 分裂基準を共有することから、霊石高地諸方言は、かつて、臨汾盆地諸方言との類似度が高かったと言えよう。

一方、霊石高地の A 方言区に隣接する晋中盆地諸方言における ET の特徴を見てみよう。晋中盆地諸方言の ET の分裂基準には 2 タイプある。ひとつは、ET 語が中古全濁声母の有無を基準に分裂するタイプである。このタイプには、文水方言（全濁 ET=312u / 非全濁 ET=2u）、祁県方言（全濁 ET=324u / 非全濁 ET=32u）、孝義方言（全濁 ET=312u / 非全濁 ET=3u）が含まれる。もうひとつは、ET 語が有声／無声を基準に分裂するタイプである。この場合、有声には全濁声母のみならず次濁声母（m-、n-、l-などの sonorant）も含まれる。このタイプには、汾陽方言（有声 ET=312u / 無声 ET=2u）、平遥方言（有声 ET=534u / 無声 ET=3u）、介休方言（有声 ET=312u / 無声 ET=12u）が含まれる。霊石高地諸方言の ET 分裂基準は、晋中盆地の汾陽、平遥、介休タイプの方言とは共通していないが、文水、祁県、孝義タイプの方言とは共通していると言える。ここで、晋中盆地における 2 タイプの方言分布図を図 4 に示す。

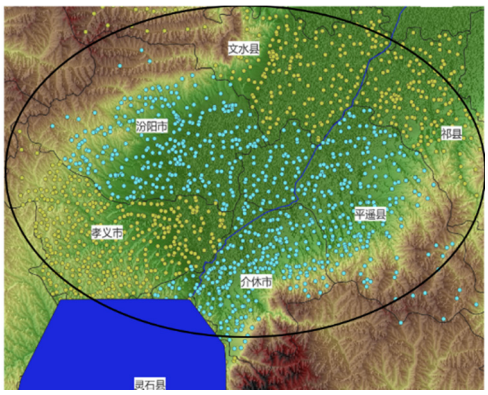


図4：晋中盆地諸方言のET分裂の分布図

図 4 は、晋中盆地諸方言の ET 分裂が 2 つの基準に基づいて分裂していることを示す図である。図 4 より、青点で示された汾陽方言・平遥方言・介休方言の ET は声母の有声・無声を基準に分裂しているのに対して、黄色で示されている分水方言・祁県方言・孝義方言の ET は声母の全濁・非全濁を基準に分裂している。また、晋中盆地の南部に位置する霊石高地の諸方言は、文水方言・祁県方言・孝義方言と同様の基準に基づいて分裂している。孝義城関方言の ET 残存実態を表 9 に示す。

表 9：孝義城関方言の ET 残存一覧表

1．ET 残存例（62 例）							
全濁ET	paʔ312u(拔)	ʂəʔ312u(実)	ʂəʔ312u (十)	tuəʔ312u(毒)	xuəʔ312u(服)	ʂəʔ312u(食)	ʂəʔ312u(石)
	xaʔ312u(合)	xuaʔ312u(罰)	ʂəʔ312u(舌)	tuəʔ312u(読)	tsuəʔ312u(濁)	tiəʔ312u(笛)	tsaʔ312u(雑)
	tsaʔ312u(宅)	piaʔ312u(白)	tɕyəʔ312u(局)	ɕyəʔ3u(俗)			
非全濁ET	uaʔ3u(袜)	iəʔ3u(薬)	miaʔ3u(麦)	yəʔ3u(月)	miəʔ3u(滅)	naʔ3u(納)	tɕiəʔ3u(急)
	tɕiəʔ3u(積)	tsʻuəʔ3u(出)	tɕiəʔ3u(接)	tɕʻiəʔ3u(七)	tɕʻiəʔ3u(切)	tiəʔ3u(滴)	piəʔ3u(筆)
	piaʔ3u(百)	suəʔ4u(説)	paʔ3u(八)	taʔ3u(搭)	tʂaʔ3u(窄)	iəʔ3u(一)	pʻiəʔ3u(匹)
	tʻiəʔ3u(鉄)	ɕiəʔ3u(錫)	tɕʻyəʔ4u(曲)	iaʔ3u(約)	xuəʔ3u(福)	tʻuəʔ3u(禿)	tsuəʔ3u(竹)
	təʔ3u(得)	tʂʻəʔ3u(尺)	tʂəʔ3u(織)	ɕyəʔ3u(削)	tsuaʔ3u(桌)	ʂəʔ3u(湿)	ʂəʔ3u(識)
	ʂəʔ3u(失)	tɕʻyəʔ3u(缺)	xuaʔ3u(髪)	xəʔ3u(黒)	kəʔ3u(割)	ɕiəʔ3u(歇)	pʻiaʔ3u(拍)
	zuəʔ312u(入)	tuəʔ312u(督)					
2．舒声化例（2 例）							
非全濁 ET	liou454(六)	i454(逸)					

表 9 より、晋中盆地に位置し、霊石高地に隣接する孝義城関方言の ET 語は、霊石高地 A 方言と同様、ほとんど舒声化していない。また、ET 語の中で、全濁 ET 語の調値は 312u であり、非全濁 ET 語の調値は 3u である。この事実より、霊石高地諸方言の古 ET の分裂基準は、臨汾盆地諸方言だけでなく、晋中盆地の文水、祁県、孝義タイプの方言とも共通していると言える。

3. 交流度によるET消失の要因の検討

本章では、ET 消失の要因について検討する。言語の変化は、人々の接触の頻度によって影響を受ける。この接触は、空間的および環境的な制約に左右される。そこで、本研究の主要な方法論として、交流度（Interaction Density）を取り上げる。

研究当初、交流度は、「中心方言区と非中心方言区の利便性（徒歩時間）」によって定義された。つまり、非中心方言区から中心方言区への利便性が高いほど、非中心方言は中心方言に似てくるということである。逆に、利便性が低いほど、非中心方言と中心区方言は異なるということである。このような基準に基づき、沈・馮・津村（2009, 2010）では霍州諸方言のデータを分析し、沈・馮・中野（2011a, 2011b）では霊石高地の諸方言のデータを分析している。

その後、交流度は、「人口密度（面積単位あたりの人数）×1 日で到達可能な村数」によって再定義された（cf. 中野・川崎・沈 2013 ; Shen & Nakano 2015 ; 中野 2016）。交流度の本質は、一定の空間内で出会う人数に依存すると判明したためである。しかし、この再定義は、方言地域における社会的接触度をある程度捉えているが、「村数」という概念を導入するための説明的妥当性には欠けている。

本研究では、これまでの定義を踏襲し、改良を加えた次のような計算方法を採用する。

(1) 交流度の計算方法
人口密度×中心地から各方言区への 1 日あたりの到達可能回数
＝1 日あたりの出会い人数

(1)より、人口密度と到達回数は、社会的接触の頻度を示しているが、両者は、それぞれ独自の意味を持っている。前者の人口密度について、多くの人と接触するほど活発なコミュニケーションが広範囲にわたって行われやすくなるため、人口密度は、社会的接触の広がりを測る指標となる。一方、後者の到達回数について、頻繁に接触するほど深い交流が生まれやすくなるため、到達回数は、社会的接触の強度を測る指標となる。このように、社会的接触の頻度、広がり、強度といった三つの側面から人々の交流の度合いを測ることができると考える。

3.1. 人口密度の計算

本節では、対象となる方言区の人口密度の計算プロセスと算出結果の説明を行う。まず、QGIS の凸包ツール(convex tool)を利用し、図 5 に示すような A、B、C、D、E の 5 つの方言地点の区画化とその面積の算出を行う。

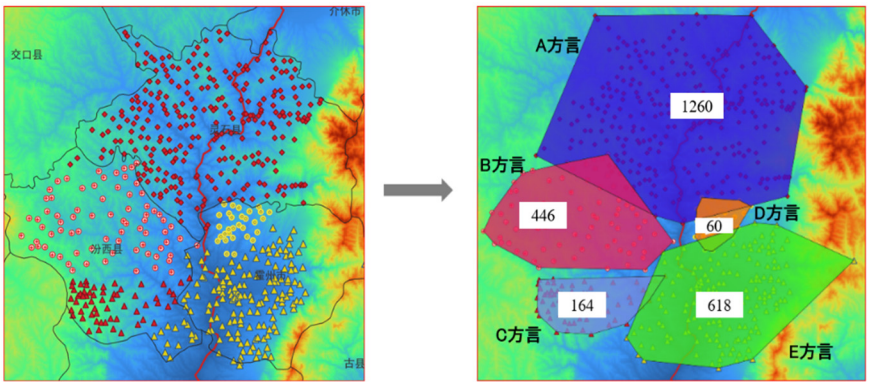


図5：5つの方言区の区画と面積

図 5 は、方言属性に基づいて、霊石高地の 606 村落を 5 つの凸包（各村落を含む最小面積）に区画したことを示している。また、A 方言区の面積は 1,260 km²、B 方言区の面積は 446 km²、C 方言区の面積は 164 km²、D 方言区の面積は 60 km²、E 方言区の面積は 618 km²であることを示している。

次に、図 6 に示すように、各方言区の村落の人口を統計する。

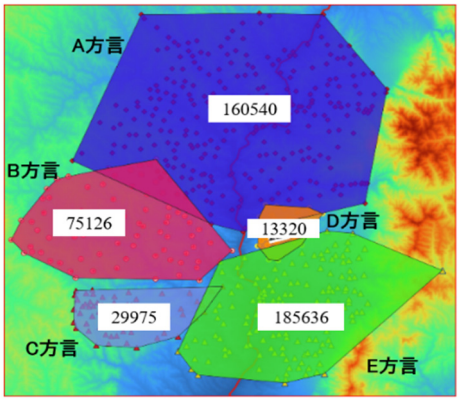


図6：5つの方言区の人口総数

図 6 より、A 方言区の人口総数は 160,540 人、B 方言区の人口総数は 75,126 人、C 方言区の人口総数は 29,975 人、D 方言区の人口総数は 13,320 人、E 方言区の人口総数は 185,636 人である。

そして、次の式を用いて各方言区の人口密度（people/km²）を計算する。

$$\text{人口密度}(\text{people/km}^2) = \frac{\text{人口総数}(\text{people})}{\text{総面積}(\text{km}^2)}$$

算出の結果、図7に示すように、A方言区の人口密度は127人/km²、B方言区の人口密度は168人/km²、C方言区の人口密度は183人/km²、D方言区の人口密度は222人/km²、E方言区の人口密度は300人/km²であることがわかった。

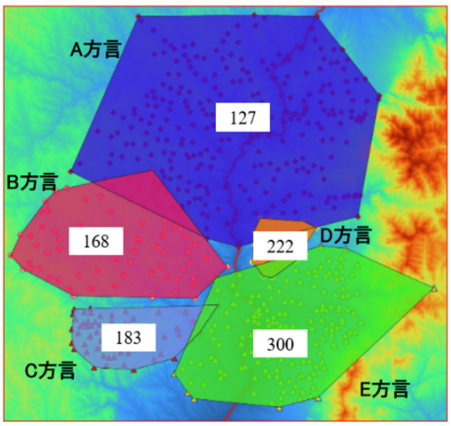


図7：5つの方言区の人口密度

3.2. 中心地から方言区に1日あたり到達可能な回数の計算

本節では、GIS を利用して「中心地から方言区に 1 日あたり到達可能な回数」を計算し、算出結果の説明を行う。まず、人々が村落から移動する際には、必ず道路を利用する必要がある。また、その際の人々の徒歩コストを算出するためには、村落をつなぐ道路網が必要である[13]。そこで本研究では、図 8 のような道路情報と村落情報を結合したデータを作成した。

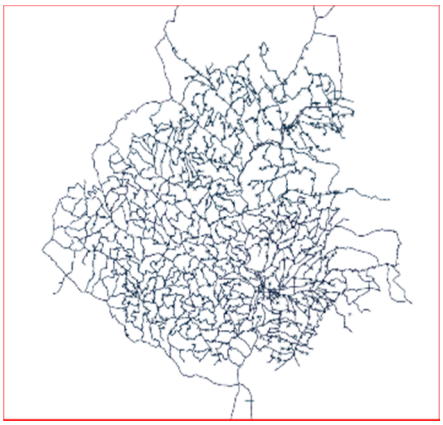


図8：各村落をつなぐ道路網

[13] 道路網を使って徒歩コストを計算すると、実際の歩くルートに基づいた正確なコストが出せる。さらに、道路網を使うことで川を渡るための橋も簡単に考慮できるため、正確な徒歩コストを計算しやすくなる。

図 8 は、GRASS を用いて、霊石高地のすべての道路と 606 の方言地点（村落）を結合したネットワークを示している。

次に、徒歩コストの計算公式を説明する。本研究では、Aitken (1977)および Langmuir (1984)が補正した「歩行時間換算モデル」を採用している。このモデルは、平地を歩く場合、斜面を登る場合、緩斜面や急斜面を下る場合にかかる時間（コスト）はそれぞれ異なるという視点に基づいて作成されている。

(2) 歩行時間換算モデル

$$T = a \times \Delta S + b \times \Delta H_{up} + c \times \Delta H_{moderate\_down} + d \times \Delta H_{steep\_down}$$

(2)より、T は歩行時間、ΔS は移動した水平距離、ΔH_{up}、$\Delta H_{moderate\_down}$、ΔH_{steep_down}は、それぞれ斜面上り、緩斜面下り、急斜面下りにおける垂直変化を表している。本研究では、傾斜角度を次のように定義している：0°以上は斜面上り、-12°から 0°までは緩斜面下り、-12°以下は急斜面下り。なお、これらの距離のデフォルト単位はメートルである。また、徒歩係数(a, b, c, d)は、異なる条件における徒歩速度を考慮したものであり、移動時間を秒単位で計算する際の値である：a=0.72, b=6.0, c=-2.0, d=2.0（秒/m）。

徒歩コストの計算において、霊石高地で最も人口密度が高く、臨汾盆地に最も近いとされるE方言区の霍州市南部の市役所を出発点として選定した。そして、図9に示すように、アメリカNASAの標高データ（地形の高さを示す衛星データ）を使用し[14]、市役所から各方言区の村々への到達に必要な最短時間を算出した。

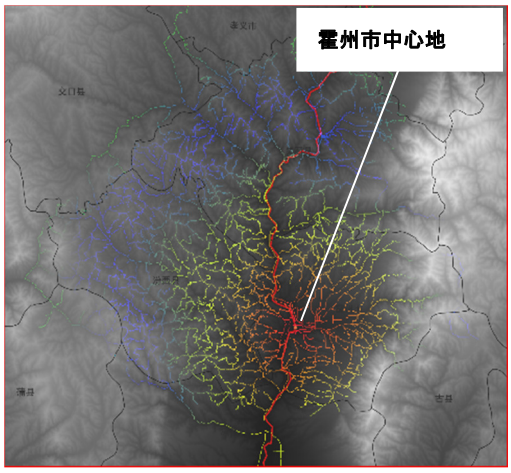


図9：霍州市役所を中心地とする徒歩コスト

[14] 本研究で使用した標高データはNASAのSRTM（Shuttle Radar Topography Mission）で公開しているhgtデータである。公開URL：http://dds.cr.usgs.gov/srtm/

図9では、徒歩コストを4段階に色分けしており、赤色の道路は0-7.1時間、黄色の道路は7.1-14.2時間、青色の道路は14.2-21.4時間、緑色の道路は21.4時間以上かかることを示している。

次に、各方言区への到達平均コストを算出するために、*k*-meansのセントロイドの概念を用いる。セントロイドは、集団内の全データポイントの重心であり、その集団に含まれるすべてのデータポイントの平均位置である。*k*-meansは、データポイントを最も近いセントロイドに割り当てることで、*k*個の集団に分割する方法である。本研究では、各方言区を各集団とするため、グループ化は行わないが、図10に示すように、各集団のセントロイドを用いて集団間の距離を測定する。

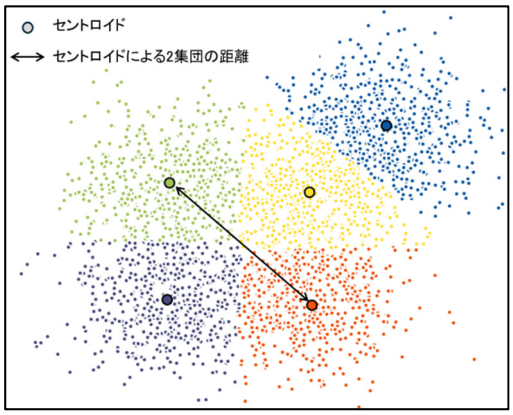


図10：*k*-meansのイメージ図

本研究では、各方言区を一つの集団とみなし、市役所から各方言区のセントロイドまでの到達時間を各方言区への平均コストとする。まず、市役所から各方言区への徒歩コストの合計を求める。

(3) 徒歩コストの合計数(hour)
A区：2820.8h
B区：634.7h
C区：492.7h
D区：130.1h
E区：462.6h

次に、徒歩コスト合計数を各方言区の村落数で割り、各方言区のセントロイドに到達する平均コスト（時間）を求める。

(4) 各方言区への到達平均コスト(hour)
A区：10.14h
B区：8.37h

C区：8.23h
D区：4.34h
E区：2.82h

(4)より、霍州市役所から出発してA区のセントロイドに到達する平均コストは10.14時間であり、これは最も高いコストである。一方、B区のセントロイドに到達する平均コストは8.37時間で、A区よりやや低いコストである。さらに、C区、D区、E区の順に徒歩コストは低くなっていく。そして、24時間を算出された平均コストで割ると、図11に示すような中心地から各方言区までの1日あたりに到達可能な回数が得られた。

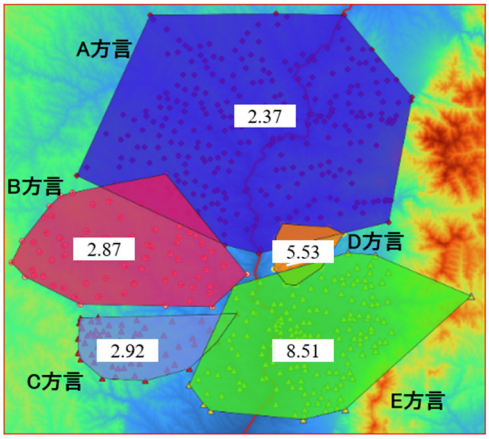


図11：中心地から各方言区に1日あたりの到達可能な回数

図 11 より、市役所を出発点とすると、1 日につき、A 方言区には平均 2 回強（2.37 回 / 24 時間）、B 方言区には平均 3 回弱（2.87 回 / 24 時間）、C 方言区には平均 3 回（2.92 回 / 24 時間）到達できる。一方、D 方言区には平均 5 回強（5.53 回 / 24 時間）、E 方言区には平均 8 回強（8.51 回 / 24 時間）到達できる。

3.3. 交流度に基づく方言伝播の評価

3.1 節では、5 つの方言区の人口密度が算出された。また、3.2 節では、各方言区に 1 日で到達可能な平均回数が算出された。本節では、交流度を算出し、言語変化の要因が交流度によるものかどうかについての評価を行う。

まず、「人口密度×1 日到達可能な回数」に基づいて、「1 日に出会う可能な人数」、すなわち交流度を計算する。

(5) 交流度（各方言区における1日に出会い可能な人数）

A区：5位（301p / 24h）
B区：4位（482p / 24h）
C区：3位（534p / 24h）
D区：2位（1228p / 24h）
E区：1位（2555p / 24h）

(5)より、中心地と各方言区の交流度における順位と、各方言区の ET の減少における順位を比較すれば、霊石高地諸方言における ET 減少の要因を探ることができる。図 12 に示すように、交流度の順位と ET の消失の順位には平行性がある。

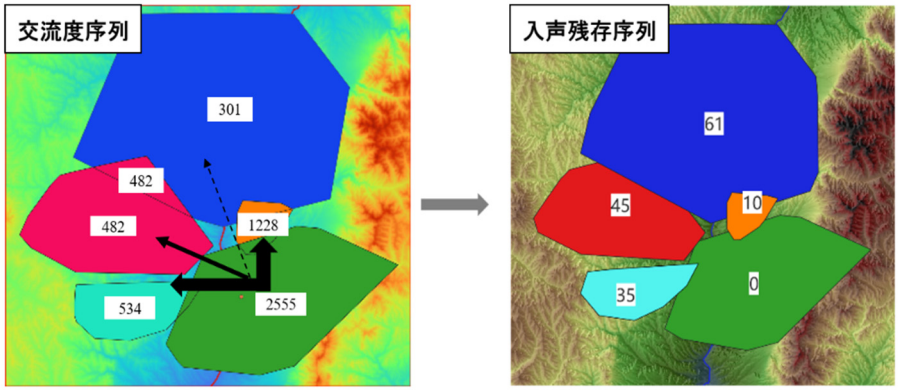


図12：交流度序列とET残存序列の平行性

図 12 より、ET の消失は、各方言内部の要因ではなく、ET を持つ地域と ET を持たない中心方言地域の人々の交流の度合いによって引き起こされた現象であることが示唆される。さらに、ET 消失の曲線は、交流度の高さによって引き起こされたものであると考えられる。

以上の算出結果に基づけば、結論として、霊石高地における曲線型減少は、中原官話である晋南方言との社会的接触度による言語外部の要因によって引き起こされたものであると言える。

4. 2022年データの分析による検証

筆者は、中国国家社会科学基金一般項目(18BYY042, 2018-2021)および日本学術振興会科研費基盤研究 B (19H01261, 2019-2022)を活用して、汾河流域の方言を調査した。この調査には、霊石高地の方言データも含まれている。2011 年のデータでは、30 代から 60 代の地元以外で 5 年以上生活したことがない人々

がインフォーマントであったが、2022 年のデータでは、50 代から 70 代の地元以外で 5 年以上生活したことがない人々がインフォーマントである[15]。つまり、2022 年のデータは、ET 残存に関する 10 年後の情報であり、ET の舒声化は進展することが予測される。また、交流度決定説が妥当であれば、交流度に基づく曲線型減少の傾向に従って、ET が減少または消失すると予測される。本章では、2022 年のデータにおける分析結果を報告する。

4.1. 2022年データにおけるET語のET残存実態

本節では、2022 年においても同様な 64 の ET 語を使用し、霊石高地諸方言に対する追跡調査の結果を報告する。

まず、霊石高地の A 方言では、10 年後も依然として 61～63 程度の ET が保持されており、ET の増減変化は観察されなかった。一方、表 10 に示すように、汾西県永安方言（B 方言）では、舒声化した ET 語が 10 年前より増えていることがわかる。

表10：汾西永安方言のET残存一覧表（2022年）

１．ET残存例（38例）							
全濁ET	səʔ3u(石)	səʔ3u(実)	səʔ3u (十)	tʻuəʔ3u(毒)	fəʔ3u(服)	səʔ3u(食)	ɕyəʔ3u(俗)
	xəʔ3u(合)	pʻəʔ3u(白)	səʔ3u(舌)	tʻuəʔ3u(読)	tsuəʔ3u(濁)	tʻiəʔ3u(笛)	tɕʻyəʔ3u(局)
非全濁ET	vəʔ21u(入)	iəʔ21u(逸)	lyəʔ21u(六)	miəʔ21u(滅)	səʔ21u(失)	tɕiəʔ3u(急)	iəʔ3u(一)
	səʔ21u(識)	fəʔ21u(福)	tɕʻiəʔ21u(七)	tsʻuəʔ21u(出)	tiəʔ21u(滴)	piəʔ21u(筆)	pʻiəʔ21u(匹)
	təʔ21u(得)	ɕiəʔ21u(錫)	tɕʻyəʔ21u(曲)	tsʻəʔ21u(尺)	səʔ21u(湿)	tʻuəʔ21u(禿)	tsuəʔ21u(竹)
	tɕiəʔ21u(積)	xəʔ21u(黒)	tsəʔ21u(織)				
２．舒声化例（26例）							
全濁ET	pʻɑ35(拔)	tsʻɑ35(雑)	fɑ35(罰)	tsɑi35(宅)			
非全濁ET	na21(納)	yɪ21(月)	iu21(薬)	va21(袜)	mɪ21(麦)	tɕʻiɪ21(切)	fi21(説)
	pɪ21(百)	ɕiu21(削)	pɑ21(八)	ta21(搭)	tsɪ21(窄)	tsu21(桌)	tʻiɪ21(鉄)
	tɕiɪ21(接)	tɕʻyɪ21(缺)	niu21(約)	fɑ21(髪)	tβ21(督)	kəu21(割)	ɕiã21(歇)
	pʻɪ21(拍)						

表 10 より、B 方言の ET 語のうち、38 語が ET を保持しており、他の 26 語は舒声化していることがわかる。また、2022 年のデータ（表 10）と 2011 年のデータ（表 3）を比較すると、B 方言の ET を保持している ET 語の残存数は、45 から 38 に減少していることがわかる。

次に、汾西邢家要方言（C 方言）の 2022 年データを表 11 に示す。

[15] 2022年データのインフォーマントの情報について、王・沈（2023）を参照されたい。

表11：汾西邢家要方言のET残存一覧表（2022）

1．ET残存例（16例）							
全濁ET	səʔ3u(石)	fuəʔ3u(服)	tʻuəʔ3u(毒)	tʻəʔ3u(読)	səʔ3u(食)		
非全濁ET	naʔ1u(納)	iəʔ1u(逸)	tɕiəʔ1u(積)	taʔ1u(搭)	tsʻəʔ1u(尺)	tiəʔ1u(滴)	tuəʔ1u(督)
	tsuəʔ1u(竹)	fuəʔ1u(福)	tɕʻyəʔ1u(曲)	tʻəʔ3u(禿)			
2．舒声化例（48例）							
全濁ET	pʻa33(拔)	ɕye33(俗)	tʻie33(笛)	tsʻa33(雑)	xəʔ3u(合)	tsã33(宅)	pʻɪ33(白)
	tɕʻye33(局)	fa33(罰)	tsuə33(濁)	sə33(舌)	sə33(実)	sə33 (十)	
非全濁ET	yɪ11(月)	mɪ11(麦)	lyə11(六)	va11(袜)	və11(入)	iu11(薬)	mie11(滅)
	tɕʻie11(切)	ɕyo11(削)	pa11(八)	pʻiə11(匹)	ɕie11(錫)	sə11(湿)	tʻie11(鉄)
	xɯ11(黒)	tɕʻye11(缺)	fuə11(説)	fa11(髪)	tɕiɪ11(接)	pʻɪ11(拍)	ɕie11(歇)
	tɕie11(急)	kə11(割)	piə11(筆)	tɕʻiə11(七)	sə11(失)	ie11(一)	tsʻuə11(出)
	niu11(約)	tsuə11(桌)	tə11(得)	tsə11(織)	pɪ11(百)	sə11(識)	tsɪ11(窄)

表 11 より、C 方言の ET 語のうち、16 語が ET を保持しており、残りの 48 語は舒声化していることがわかる。また、2022 年データ（表 11）と 2011 年のデータ（表 4）を比較すると、C 方言の ET 語の ET 残存数は、35 から 16 に減少していることがわかる。

次に、霍州北部の霍州師庄方言（D 方言）の ET 語の ET 残存状況を表 12 に再掲する。

表12：霍州師庄方言のET残存一覧表（=表6：霍州城関方言のET残存一覧表）

1．ET 残存例（0 例）							
2．舒声化例（64 例）							
全濁ET	pʻɑ35(拔)	ɕiu35(俗)	tɕʻi35(笛)	tsʻɑ35(雑)	xɤ35(合)	tsɑi35(宅)	pʻie35(白)
	tɕiu35(局)	fɑ35(罰)	tʂʻuə35(濁)	ʂie35(舌)	tʻu35(毒)	fu35(服)	tu35(読)
	ʂʅ35 (十)	ʂʅ35(食)	ʂʅ35(石)	ʂʅ35(実)			
非全濁ET	lɑ212(納)	mie212(麦)	liu212(六)	vɑ212(袜)	mie212(滅)	iuɤ212(薬)	tie212(得)
	ye212(月)	ʐu212(入)	tsʻu212(出)	ʂʅ212(失)	fɑ212(髪)	tɕi212(積)	kɤ212(割)
	tɕʻie212(切)	ɕiuɤ212(削)	pɑ212(八)	pʻi212(匹)	tʂʻʅ212(尺)	ʂʅ212(湿)	tɕʻie212(鉄)
	xei212(黒)	tɕʻye212(缺)	ʂuɤ212(説)	tʻu212(禿)	tu212(督)	pʻie212(拍)	ɕie212(歇)
	ʂʅ212(識)	fu212(福)	tɕʻi212(七)	tsɤ212(窄)	tie212(滴)	pi212(筆)	tsuɤ212(桌)
	tʂʅ212(織)	ɕi212(錫)	tɕʻiu212(曲)	tɑ212(搭)	pie212(百)	tɕie212(接)	tʂu212(竹)
	i35(逸)	tɕi35(急)	ȵiuɤ35(約)	i35(一)			

表 12 より、D 方言では、すべての ET 語が舒声化している。2011 年のデータでは 10 語の ET が残存していたが、2022 年のデータではこれらの ET がそれぞれ 35 調（陽平調）と 212 調（陰平調）に合流している。舒声化後の声調体系（5 声調体系）が霍州城関方言（E 方言）と同様であることから、D 方言

における ET の消失は、中心方言である霍州城関方言の影響を受けて同化された結果であると考えられる。

4.2. 中心方言区との交流度による ET 消失の説明

2011 年のデータと 2022 年のデータの比較より、霊石高地諸方言における ET の消失は、D 方言、C 方言、B 方言、A 方言の順で進行していることがうかがえる。この順序は、本研究の提案である交流度に基づいて、次のように説明することができる。

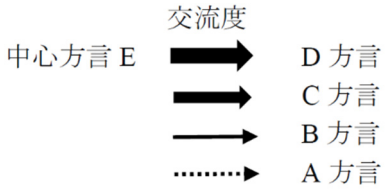


図 13：各方言区から中心方言区との交流度序列

図 13 より、矢印の太さは交流度の高さを示している。つまり、ET の消失は、中心方言 E との交流度が高い D 方言を先頭として進行し、その後を追うように C 方言、B 方言、A 方言の順序で進行しているということである。また、ここ 10 年間における各方言区の ET の減少率を表 13 に示す。

表13：追跡調査によるETの変化

霊石高地方言区	2011データ	2022データ	減少率（%）
E方言区	0	0	
D方言区	10	0	100
C方言区	35	16	54.28
B方言区	45	38	20
A方言区	61	61	0

表 13 より、E 方言との交流度が最も高い D 方言は、E 方言と完全に同化しており、減少率は 100%となっている。E 方言との交流度が 2 番目に高い C 方言は、ET 語の ET が 35 語から 16 語に減少しており、減少率は 54.28%となっている。その次の B 方言は、45 語から 38 語に減少しており、減少率は 20%となっている。そして、E 方言との交流度が最も低い A 方言は、ET 語の ET が減少しておらず、減少率は 0%となっている。

5. 結論と展望

本研究では、中古漢語の ET が存在している晋中盆地（汾河流域中部）と存在していない臨汾盆地（汾河流域南部）の接触地域である霊石高地において、ET の残存状況を記述し、ET の消失の要因として「交流度」を提案した。

まず、両者の盆地には、2 種類の ET が存在していたという共通点がある。この 2 種類の ET は、晋中盆地の方言では、現在も確認できるが、臨汾盆地の方言では 2 種類の平声調に合流している[16]。この事実より、2 種類に分かれていた ET は、その後、それぞれ舒声調に合流したと解釈できる。一方、両者の盆地には、ET の分裂する基準における相違点がある。晋中盆地の方言では、声母に関する 2 つの基準に基づいて ET の分裂が生じる。汾陽、平遥、介休の諸方言では、声母の［全濁・次濁／清音］という対立が基準となり、文水、祁県、孝義の方言では［全濁／次濁・清音］という対立が基準となる。また、臨汾盆地の諸方言では、後者の基準に基づいて ET の分裂が生じる。さらに、臨汾盆地や晋中盆地（文水、祁県、孝義）の諸方言と同様、汾河流域を跨ぐ霊石高地の諸方言も、声母の全濁・非全濁によって ET の分裂が生じる。このような ET の分裂から見れば、歴史的に同じタイプの方言が川沿いに集中していると考えられる。なお、霊石高地における ET の残存数は、「A－B－C－D－E」の順序で段階的に減少している。

本研究では、以上の事実に基づいて、Shen & Nakano (2015)における「交流度」を改良し、中心地域と各方言区の接触の頻度・広がり・強度を測定する指標として新たに提案した。これにより、ET の減少の序列は、霊石高地の各方言区と中心方言区（E 方言）との交流度の序列と一致していることがわかった。結論として、中古漢語における ET の消失は、各方言が独自に変化した結果ではなく、中心方言の影響によるものであると言える。また、2011 年データと 2022 年データを比較した結果、ET は、交流度の序列に基づいて舒声化していることがわかった。具体的には、E 方言との交流度が高い D 方言では、すべての ET が消失し、E 方言と同様、完全に舒声化している。また、ET は、E 方言との交流度の序列に基づいて、C 方言では 54%、B 方言では 20%減少している。一方、E 方言との交流度が低い A 方言では、ET の残存数に変化はない。これらの事実は、本研究における交流度の妥当性を支持していると言える。

最後に「中心方言区との交流度によって言語変化が引き起こされる」という本研究の仮説をより広く検証するためには、追跡調査を行っていく必要がある。例えば、10 年後の ET の残存状況を推測すれば、ET は、A 方言と B 方言では残存しているが、C 方言では、D 方言と同様、完全に消失している可能性が高い。このような追跡調査による発展性については、今後の研究に譲る。

[16] 二大盆地の入声調の特徴の詳細について、王（2003）、喬（2008）を参照されたい。

参考文献/References

Aitken, R. (1977) *Wilderness areas in Scotland*. Unpublished Ph.D. thesis. University of Aberdeen.

侯精一 (1999) 晋語的分区, 侯精一著《現代晋語的研究》, pp.30−43, 商務印書館.

侯精一 (2008) 入声問題−晋語与官話的比較研究之一, 喬全生(主編).《晋方言研究 第三届晋方言国際学術研討会論文集》, pp.1−4, 希望出版社.

侯精一・温端政 (1993)《山西方言調査研究報告》, 山西高校聯合出版社.

Karlgren, Bernhard (高本漢) (1915−26)《中国音韻学研究》, 趙元任、李方桂(訳) (1940) 台湾商務印書館.

Langmuir, E. (1984) *Mountaincraft and leadership*. The Scottish Sports Council/MLTB. Cordee, Leicester.

中野尚美・川崎廣吉・沈力 (2013) GISを用いた言語伝播経路の推定―交流度計算方法の再検討―, 《文化情報学》, Vol. 8, No. 2, pp.14−22, (同志社大学文化情報学会).

中野尚美 (2016) 晋方言・官話方言接触地域における言語の時間的・空間的変化―霊石県諸方言を中心に―, 同志社大学大学院文化情報学研究科, 博士論文.

乔全生 (2008)《晋方言語音史研究》, 中華書局.

沈力・馮良珍・津村宏臣 (2009) 用GIS跟踪霍州話構形法的衰退軌跡, 《東方語言学》No.5, pp. 104−122, 上海教育出版社.

沈力・馮良珍・津村宏臣 (2010) 用GIS分析山西霍州方言元音諧和律的衰退現象―以指小詞変韻為中心―, 《中国語言学報》, pp.137−157. 商務印書館.

沈力・冯良珍・中野尚美 (2011a) 用GIS手段解读混合方言的成因―以霊石高地为例―, 《语言教学与研究》（北京語言大学）, No.5, pp.30−39.

沈力・冯良珍・中野尚美 (2011b) 山西回廊における入声変化の解明―地理情報科学の方法を利用して―, 藤代節 (編) *Dynamics of Eurasian Languages II* (Studies on Languages in Multilingual Areas), Kobe: Kishimoto Press. pp. 75−96.

Shen, Li and Naomi Nakano (2015) A Gradual Path to the Loss of Entering Tone: Case Studies of Jin Dialects in the Lingshi Highlands Shanxi. Dan Xu and Jingqi Fu (eds.) *Space and Quantification in Languages of China*, pp.75−92(251), New York: Springer.

王力 (2008) 《漢語語音史》, 商務印書館.

王臨惠 (2003) 《汾河流域方言的語音特点及其流変》, 中国社会科学出版社.

王堉程・沈力 (2023) 山西中部方言麻韻三等字元音的曲折変化, 《語文研究》, Vol. 166, No. 1, pp.57−65,（山西省社会科学院語言研究所）.

温端政 (1989) 試論山西晋語的入声, 《中国語文》, Vol. 191, No.2. pp.124−127, (中国社会科学院語言研究所).

中国社会科学院語言研究所（編）(1964)《方言調査字表》, 商務印書館.

出版情報

投稿受理日：2024年7月7日
採用決定日：2024年8月20日

Research article (Featured theme: Geolinguistic approaches to linguistic patterns in Asia and Africa)

Introduction: Geolinguistic approaches to linguistic patterns in Asia and Africa

FUKUSHIMA, Chitsuko
University of Niigata Prefecture

Abstract: This introduction provides some background on the geolinguistic approaches to data used by the contributors to this special issue, who were all involved in the project of making the three-volume *Linguistic Atlas of Asia and Africa*. The papers focus on clarifying the distributions of linguistic *patterns*—not linguistic *matter*—in data from Asian and African languages, including sibling term systems, numeral systems, alignment, and stop series.*

Keywords: *Linguistic atlas of Asia and Africa*; geolinguistic approach; linguistic pattern; linguistic matter

The contributors to this special issue were all involved in the project to make the *Linguistic Atlas of Asia and Africa (LAAA), volumes I*, *II*, and *III* (Suzuki et al. 2022, 2023, and Fukushima et al. 2023). The project applied geolinguistic approaches to the linguistic variation in Asia and Africa. Experts on specific languages or language groups collected the data and made linguistic maps of the items of interest for each language or language group, using ArcGIS online. All of the project's participants were guided to understand the expected system types/elements. In addition, the maps the individual researchers created employed the same sets of symbols (assigned beforehand) in most cases, which made it possible to envision the distributions of system types when the maps of different languages were combined.

For the mapping, each language or language group was assigned one of seven colors, as shown in Table 1.

FUKUSHIMA, Chitsuko. 2024. Introduction: Geolinguistic approaches to linguistic patterns in Asia and Africa. *Studies in Geolinguistics* 4: 142–144. doi: https://doi.org/10.5281/zenodo.13948601

* This work was supported by Tokyo University of Foreign Studies, April 2020–March 2023 ILCAA Joint Research Project: Studies in Asian and African Geolinguistics (jrp000256), with Mitsuaki Endo as coordinator; and JSPS KAKENHI grant number JP19K00555. I also thank the members of the 2020–2023 project.

Table 1 The language groups and colors (made by Satoko Shirai)

Language group	Color
Ainu	Red
Andamanese and language isolates (South Asia)	Black
Austroasiatic	Green
Austronesian	Orange
Caucasian	Light blue
Chukotko-Kamchatkan	Black
Dravidian	Red
Hmong-Mien	Orange
Indo-Aryan and Nuristani (South Asia)	Navy
Iranian	Orange
Japonic	Green
Khoe-Kwadi (Kalahari Basin Area)	Brown
Korean	Light blue
Kra-Dai	Navy
Kx'a (Kalahari Basin Area)	Light blue
Mongolic	Green
Niger-Congo	Navy
Nilo-Saharan	Green
Semitic	Brown
Sinitic	Red
Tibeto-Burman	Brown
Tungusic	Orange
Turkic	Light blue
Tuu (Kalahari Basin Area)	Orange
Uralic	Light blue

Linguistic geography was once been called "word geography," and it is often regarded as centering on the history of words. And of course, *LAAA* includes lexical items: animal vocabulary [Rat/Mouse, Chicken, Horse, Dog, Wolf, and Bear] and crop terms [Wheat, Broomcorn Millet (*Panicum miliaceum*), Foxtail Millet (*Setaria italica*), Barnyard Millet (*Echinochloa* Species), Taro, and Yam]. Nevertheless, linguistic geography can also focus on phonological, morphological, or semantic variation. This special issue features four items from the *LAAA*: two lexical systems, those of sibling terms and numerals; one grammatical item, specifically grammatical relations or alignment; and one phonological item, that is, stop series. Note that while the first three systems were assigned symbols for each system type as mentioned above, the stop series were not assigned symbols beforehand.

These items have one common characteristic: they are concerned with what Matras and Sakel (2007) called "linguistic patterns," in contrast to "linguistic matter." These terms are useful to explain distinct types of contact-induced changes. When languages share "linguistic matter," it is due to "the replication of morphological material from the source language" 841), which is usually called "borrowing" and is the result of language contact. You can find many examples of shared linguistic matter in the maps of lexical items, such as those representing the distributions of the animal and crop vocabulary mentioned above. On the other hand, the sharing of "linguistic patterns" refers to "the replication of *usage* patterns from a model language" (841; emphasis added). For example, the sibling term systems occur across geographic space in

linguistic patterns; a type of sibling term system used in a source language is introduced to the surrounding area, and the words of the recipient language are used to occupy the parts of the system. If, for instance, a brother/sister system is introduced from a source language, two local words may be adopted to mean "brother" and "sister." This could also apply to numeral systems. Usage patterns can include grammatical and phonological patterns as well; hence, the maps of alignment and stop series patterns are included as part of this issue's featured theme. They show interesting distributions, which in fact prove that they are also linguistic patterns.

The featured papers all employed the same method to create the maps they include, as follows. The existing maps of each language or language family, made by the individual language/language group experts, were superimposed on each other, again using ArcGIS online. The resulting map included all of the attested types of linguistic patterns of the system in question. In order to show the distribution of each type, the maps of other types were deleted; then a map of each type was produced. The featured articles provide the maps along with the authors' interpretations of the information represented in the maps.

The maps thus produced show the distributions that we might otherwise never have seen and the experts' interpretations clarify the contributions to our understanding of the linguistic geography of Asia and Africa.

References

Fukushima, Chitsuko, Satoko Shirai, Mika Fukazawa, Hiroyuki Suzuki and Mitsuaki Endo (eds.) (2023) *Linguistic atlas of Asia and Africa Vol. 3*. Tokyo: Geolinguistic Society of Japan. https://doi.org/10.5281/zenodo.10223731

Matras, Y. and J. Sakel (2007) Investigating the mechanisms of pattern replication in language convergence. *Studies in Language* 31: 829–65.

Suzuki, Hiroyuki, Mika Fukazawa, Akiko Yokoyama and Mitsuaki Endo (eds.) (2022) *Linguistic atlas of Asia and Africa Vol. 1*. Tokyo: Geolinguistic Society of Japan. https://doi.org/10.5281/zenodo.7118188

Suzuki, Hiroyuki, Kohei Nakazawa and Mitsuaki Endo (eds.) (2023) *Linguistic atlas of Asia and Africa Vol. 2*. Tokyo: Geolinguistic Society of Japan. https://doi.org/10. 5281/zenodo.7754469

Publication history

Date received: 31 August 2024
Date accepted: 17 September 2024

Research article (Featured theme: Geolinguistic approaches to linguistic patterns in Asia and Africa)

Sibling terms in Asia and Africa: A geolinguistic approach to linguistic patterns

FUKUSHIMA, Chitsuko
University of Niigata Prefecture

Abstract: Sibling terms in Asia and Africa are examined by making new maps, each of which shows detailed distributions of each type of sibling term system. The interpretation of the maps strongly supports a previous interpretation, as follows: (1) Change from Type A to Type B and from Type B to Type C can be inferred from the distributions, (2) Type D and Type E expanded relatively recently, and (3) Type F shows the relic, peripheral distributions.*

Keywords: *Linguistic Atlas of Asia and Africa*; sibling system; linguistic pattern

1. Introduction

The purpose of this paper is to apply a geolinguistic approach to sibling term systems in Asia and Africa. The maps published in the *Linguistic Atlas of Asia and Africa III (LAAA III)* are redrawn here to examine the previous interpretation, made in Fukushima (2023d).

Systems of sibling terms are described using three criteria, which are based on distinctions of (1) relative age, (2) sex, and (3) relative sex (Matsumoto 2006 and Murdock 1968), and there is a variety of systems, as shown in Table 1. Types A to E are based on distinctions of relative age and sex. Type A (undifferentiated sibling type) has just one term for ‘sibling’. Type B (relative age type) has two terms, for ‘elder sibling’ and ‘younger sibling’. Type C (skewed age type) has three terms, for ‘elder brother’, ‘elder sister’, and ‘younger sibling’. Type D (age/sex type) has four terms, for ‘elder brother’, ‘elder sister’, ‘younger brother’, and ‘younger sister’. Type E (sex type or brother/sister type) has two terms, for ‘brother’ and ‘sister’. Types F to FE make distinctions of relative sex, which are often combined with other distinctions. These

FUKUSHIMA, Chitsuko. 2024. Sibling terms in Asia and Africa: A geolinguistic approach to linguistic patterns. *Studies in Geolinguistics* 4: 145–155. doi: https://doi.org/10.5281/zenodo.13948605
* This work was supported by Tokyo University of Foreign Studies, April 2020–March 2023 ILCAA Joint Research Project: Studies in Asian and African Geolinguistics (jrp000256), with Mitsuaki Endo as coordinator; and JSPS KAKENHI grant number JP19K00555.

types are typological and regarded as linguistic patterns, each of which is expected to show a different geographical distribution. Therefore, to enable the visual representation of the systems' distributions, a symbol was assigned to each system type before the researchers of each language or language group in the project made the separate linguistic maps.

Table 1 Types of sibling term systems (Fukushima 2023a, 2023c)

	Type	Relative age	Sex	Relative sex	Symbol
A	Undifferentiated sibling type				○
B	Relative age type	+			—
C	Skewed age type	+	+		▽
D	Age/sex type	+ +	++		□
E	Sex type (brother/sister type)		+		◇
F	Relative sex type			+	⊖
FB	Relative sex/age type	+		+	/
FC	Relative sex/skewed age type	+	+	+	◸
FD	Relative sex/age/sex type	+ +	++	+	▱
FE	Relative sex/sex type		+	+	◊

When I first examined the geographical variation of sibling term systems in Asia and Africa as part of the *LAAA III* project in 2023, the maps were not ready for online superposition; thus, I used the rate (percentage) of each system in a language or language family. I could count the numbers of each system by reading the maps or the descriptions of each map. First, a map including pie charts was made using the percentages. See Figure 1.

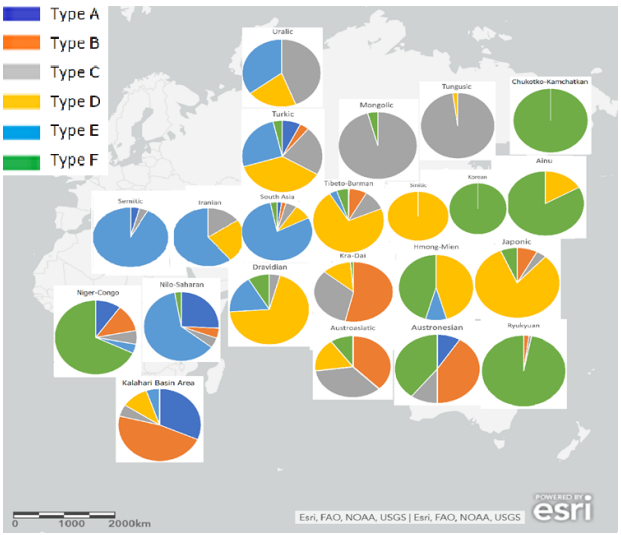


Fig. 1 Types of sibling term systems in Asia and Africa (Fukushima 2023a, 2023b, 2023c, 2023d)

Fig. 2 Type A (Fukushima 2023b, 2023c, 2023d)

Fig. 3 Type B (Fukushima 2023b, 2023c, 2023d)

Fig. 4 Type C (Fukushima 2023b, 2023c, 2023d)

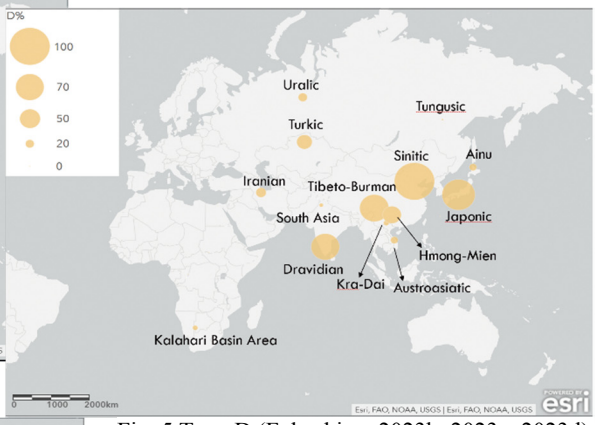


Fig. 5 Type D (Fukushima 2023b, 2023c, 2023d)

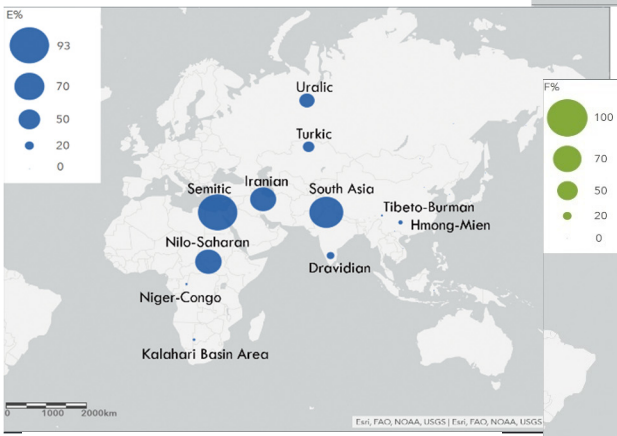


Fig. 6 Type E (Fukushima 2023b, 2023c, 2023d)

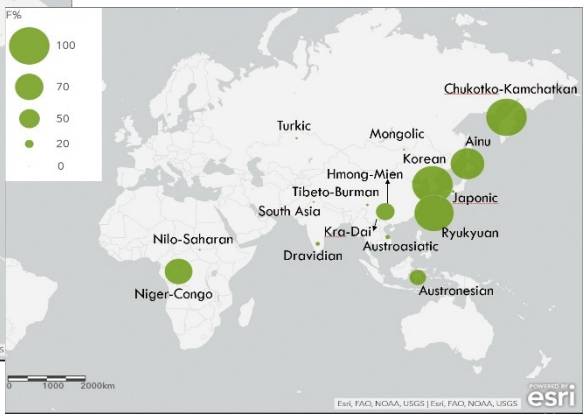


Fig. 7 Type F (Fukushima 2023b, 2023c, 2023d)

Next, maps showing the variation of each system were made. See Figures 2–7 above. Based on these maps, the following interpretation was made (Fukushima 2023d). In this paper, I will verify this interpretation using new maps.

1. Change from Type A to Type B and from Type B to Type C can be inferred from the distributions.
2. Type D and Type E expanded relatively recently.
3. Type F shows the relic, peripheral distributions.

2. Method and results

Separate original maps of sibling term systems were made by experts on specific languages or language groups in Asia and Africa, using ArcGIS online. They were published in the *Linguistic Atlas of Asia and Africa III* (Fukushima et al. 2023). The maps of sibling term systems of each language or language group were superimposed online to produce a map of sibling term systems across Asia and Africa (see Figure 8). The color for each language or language group was decided beforehand; see Table 2.

Based on this map, maps of each type were made by deleting maps that did not show that type, and also deleting symbols of types other than the type at hand. The new maps (Figures 9–14) thus produced are shown and discussed in the following section.

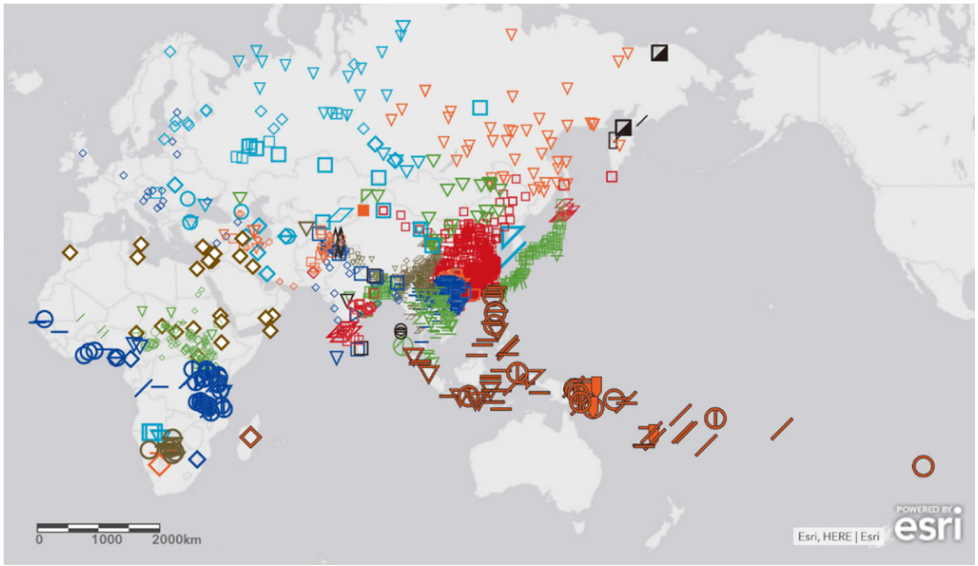


Fig. 8 Sibling term systems in Asia and Africa: all types

Table 2 The language groups and colors (made by Satoko Shirai)

Language group	Color
Ainu	Red
Andamanese and language isolates (South Asia)	Black
Austroasiatic	Green
Austronesian	Orange
Caucasian	Light blue
Chukotko-Kamchatkan	Black
Dravidian	Red
Hmong-Mien	Orange
Indo-Aryan and Nuristani (South Asia)	Navy
Iranian	Orange
Japonic	Green
Khoe-Kwadi (Kalahari Basin Area)	Brown
Korean	Light blue
Kra-Dai	Navy
Kx'a (Kalahari Basin Area)	Light blue
Mongolic	Green
Niger-Congo	Navy
Nilo-Saharan	Green
Semitic	Brown
Sinitic	Red
Tibeto-Burman	Brown
Tungusic	Orange
Turkic	Light blue
Tuu (Kalahari Basin Area)	Orange
Uralic	Light blue

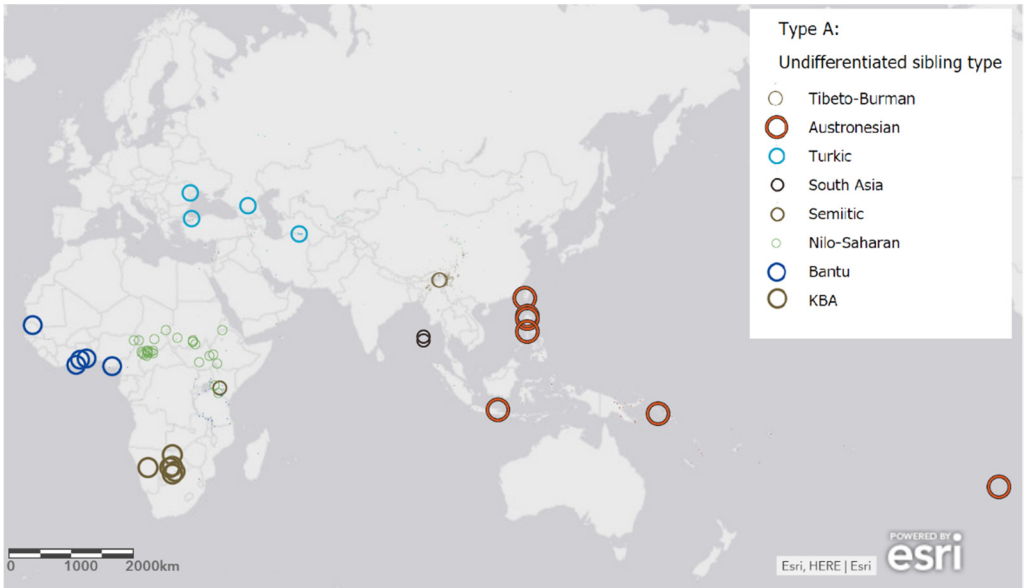


Fig. 9 Type A: Undifferentiated sibling type

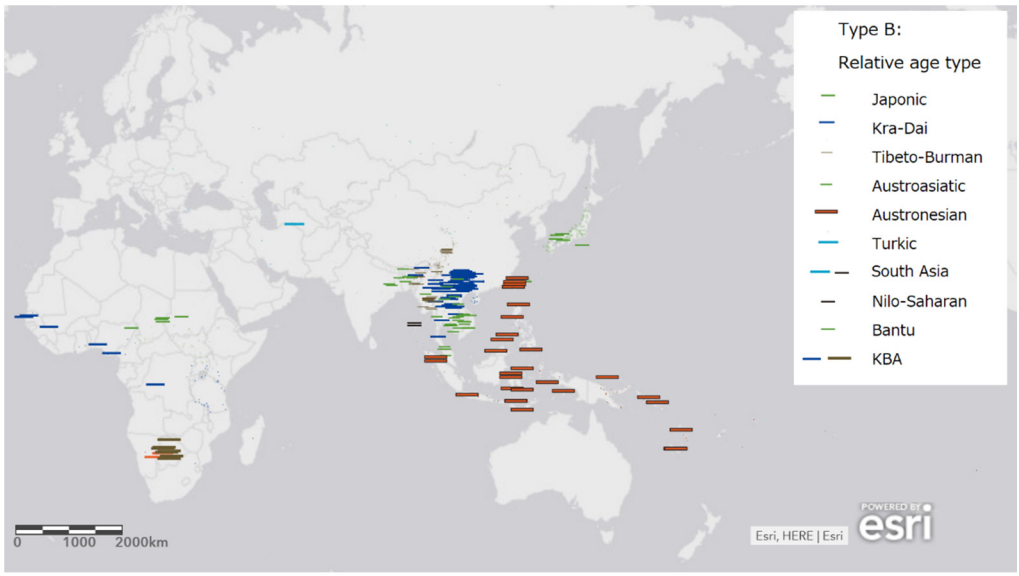


Fig. 10 Type B: Relative age type

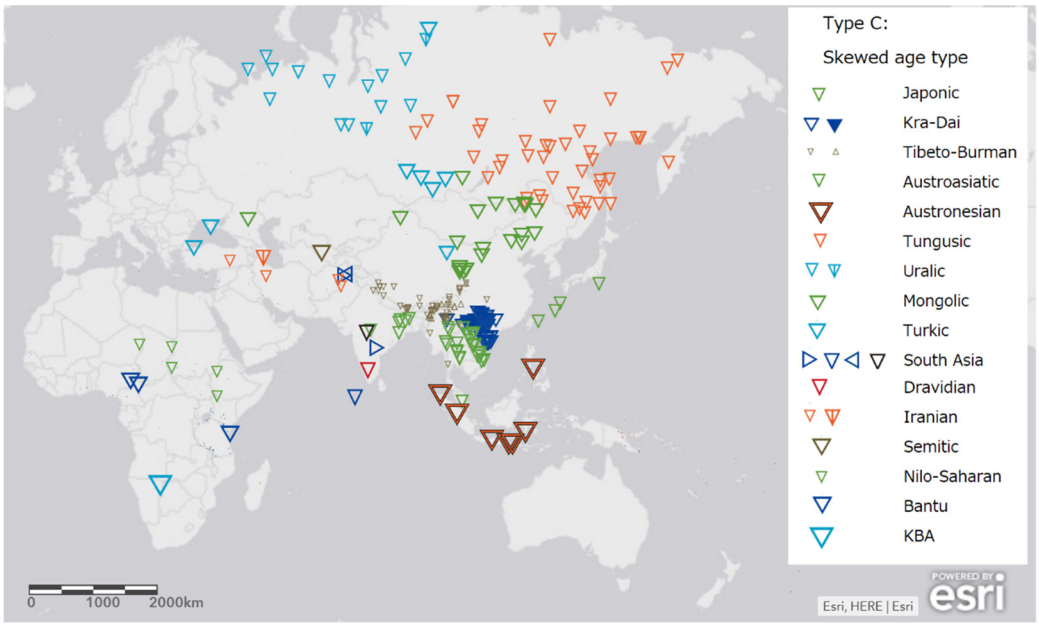


Fig. 11 Type C: Skewed age type

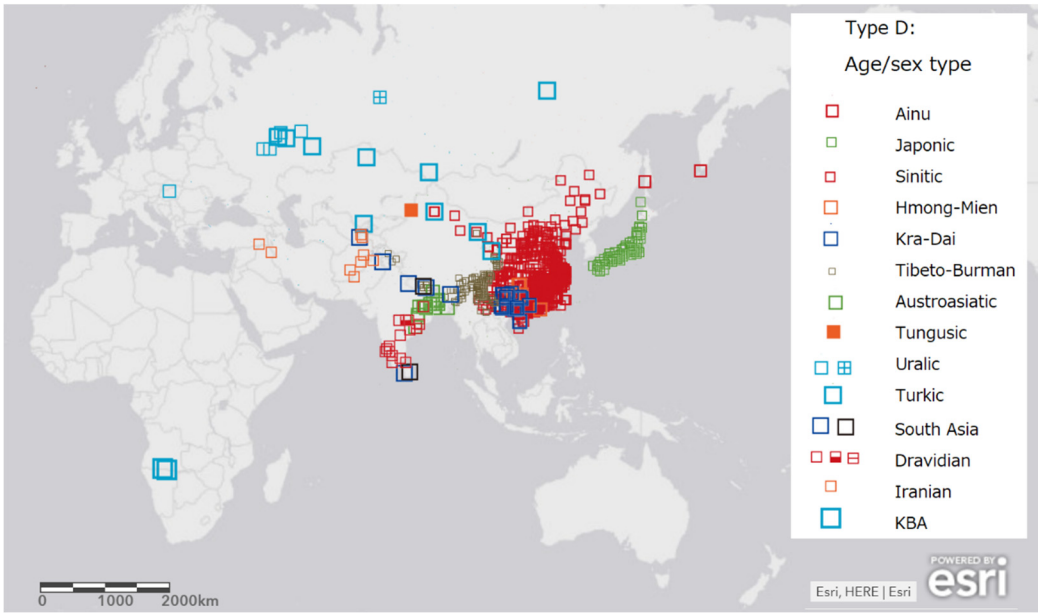


Fig. 12 Type D: Age/sex type

Fig. 13 Type E: Brother/sister type

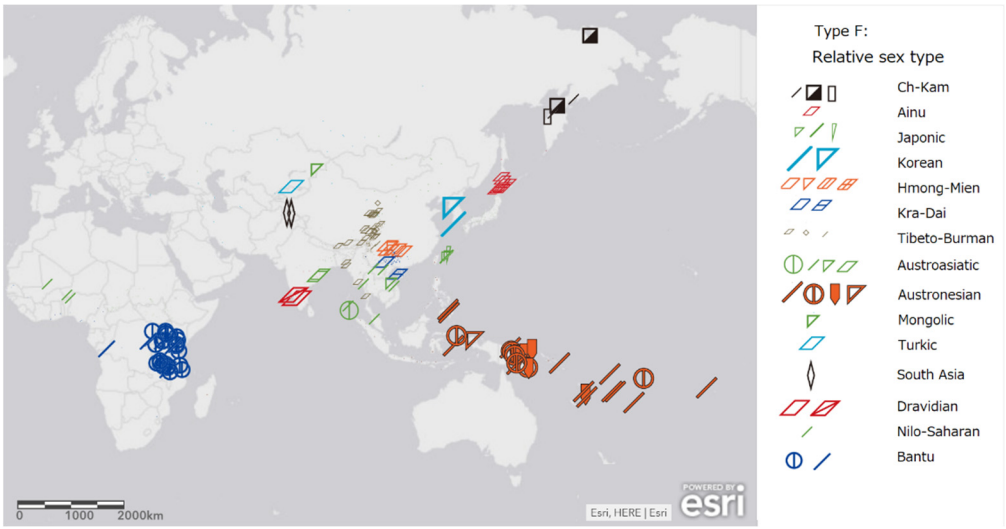


Fig. 14 Type F: Relative sex type

3. Discussion

If we compare these maps with the maps introduced in Section 1 (Figures 2–7), we find that the new maps show detailed, actual distributions of each type: whether the distribution is dense or sporadic, how it expanded, and so forth. Let me examine each map.

Figure 9 shows that Type A is found in the vast area of Southeast Asia and Africa but is rather scattered. Figure 10 shows that Type B is found in areas nearby to Type

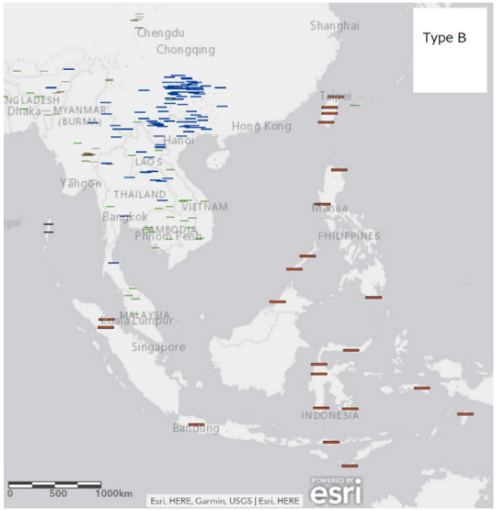


Fig. 15 Type B focusing on Southeast Asia

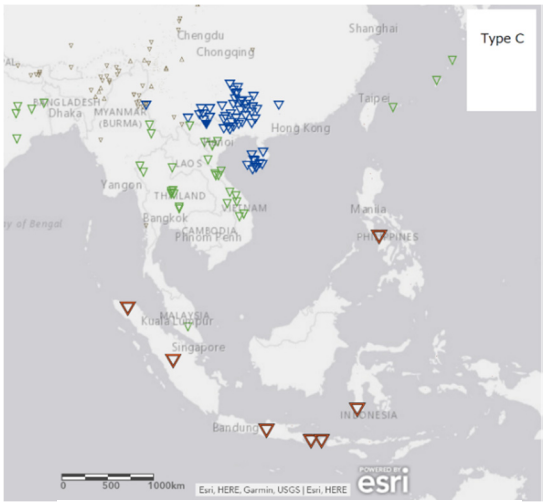


Fig. 16 Type C focusing on Southeast Asia

A, and its distribution is rather dense. Figure 11 shows that Type C is found all over Asia, that its distribution in Indochina is especially dense, and that it is also found in Africa. Also see Figures 15 and 16, the maps of Types B and C, which focus on Southeast Asia. Type C is found in Java and Vietnam, and you can see the routes of Type C in the Tibeto-Burman area, while the distribution of Type B can be seen to be surrounding that of Type C; thus, Type C is newer than Type B. Overall, the change from Type A (undifferentiated sibling type) to Type B (relative age type) and to Type C (skewed age type) is probable as the direction of lexical change, which is supported by the distributions of these three types.

Figure 12 shows that Type D (age/sex type) clearly spread from China, through several routes to South Asia and to Europe. See Figure 17, which focuses on East Asia and shows the routes more clearly. There are a couple of locations in Africa where Type D is used, where it likely developed from Types A, B, and C, maybe due to contact with the brother/sister type.

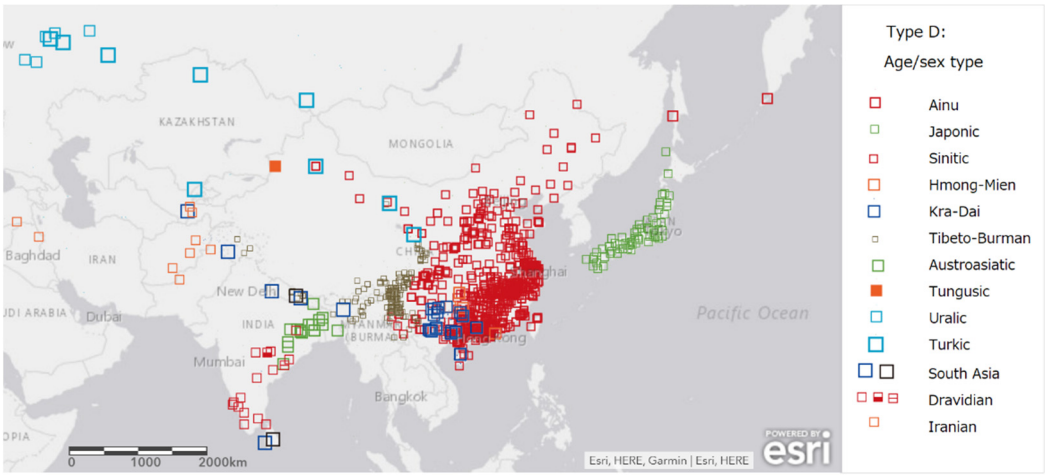


Fig 17 Type D: Age/sex type: focus on East Asia

Figure 13 shows that Type E (brother/sister type) is only found in the western part of the map, which suggests that this type spread from the west.

Figures 12 and 13 suggest that Types D and E spread rather recently, making inroads into the areas of the other types' distributions.

Figure 14 shows that Type F (relative sex type) is found, sporadically, on the peripheries of Asia and Africa, especially in the eastern and southern edges of Eurasia and also in parts of Africa. Many different languages and language families in these areas have this type, suggesting that it may be the oldest type in the area.

4. Conclusion

The newly-made maps strongly support the previous interpretation and illustrate the distributions more clearly.

Notes for copyrights

The members of the 2020–2023 project involved in the investigation of the sibling data are the following linguists: ONO Chikako (Chukotko-Kamchatkan), FUKAZAWA Mika (Ainu), FUKUSHIMA Chitsuko (Japonic), FUKUI Rei (Korean), YAGI Kenji (Sinitic), TAGUCHI Yoshihisa and TANG Baiyan (Hmong-Mien), HIRANO Ayaka, ENDO Mitsuaki, and TOMITA Aika (Kra-Dai), KURABE Keita, EBIHARA Shiho, IWASA Kazue, SHIRAI Satoko, and SUZUKI Hiroyuki (Tibeto-Burman), SHIMIZU Masaaki and MINEGISHI Makoto (Austroasiatic), UTSUMI Atsuko (Austronesian), MATSUMOTO Ryo (Tungusic and Uralic), SAITÔ Yoshio (Mongolic and Turkic), YOSHIOKA Noboru (South Asia), KODAMA Nozomi (Dravidian), IWASAKI Takamasa (Iranian), NAGATO Youichi (Semitic), NAKAO Shuichiro (Nilo-Saharan), SHINAGAWA Daisuke and KOMORI Junko (Niger-Congo), and KIMURA Kimihiko and NAKAGAWA Hirosi (Kalahari Basin Area). I express my sincere thanks to the contributors.

References

Fukushima, Chitsuko (2023a) Overview of the system of ‘sibling’ terms. Paper presented at the 6th meeting of the ILCAA Joint Research Project: Studies in Asian and African Geolinguistics, Tokyo University of Foreign Studies and online, 25–26 March.

Fukushima, Chitsuko (2023b) The geographical variation of the system of ‘sibling’ terms in Asia and Africa. Paper presented at the 10th Congress of the International Society for Dialectology and Geolinguistics (SIDG 10), Bucharest, Romania, 4–8 September.

Fukushima, Chitsuko (2023c) Overview of the system of ‘sibling’ terms. In: Chitsuko Fukushima, Satoko Shirai, Mika Fukazawa, Hiroyuki Suzuki, and Mitsuaki Endo (eds.) *Linguistic Atlas of Asia and Africa III*, 3–8. Tokyo: Geolinguistic Society of Japan. https://doi.org/10.5281/zenodo.10223731

Fukushima, Chitsuko (2023d) Geographical variation of systems of sibling terms in Asia and Africa. *Dialectologia et Geolinguistica* 31: 41–54. De Gruyter. https://doi.org/10.1515/dialect-2023-0003

Fukushima, Chitsuko, Satoko Shirai, Mika Fukazawa, Hiroyuki Suzuki, and Mitsuaki Endo (eds.) (2023) *Linguistic Atlas of Asia and Africa III.* Tokyo: Geolinguistic Society of Japan. https://doi.org/10.5281/zenodo.10223731

Matsumoto, Katsumi [松本克己] (2006) *Sekai Gengo eno Shiza: Rekishi Gengogaku to Gengo Ruikeiron* 『世界言語への視座：歴史言語学と言語類型論』[Outlook on world languages: Historical linguistics and linguistic typology], 391–443. Tokyo: Sanseido.

Murdock, George Peter (1968) Patterns of sibling terminology. *Ethnology* 7(1): 1–24. https://www.jstor.org/stable/3772805

Publication history

Date received: 31 August 2024
Date accepted: 17 September 2024

Research article (Featured theme: Geolinguistic approaches to linguistic patterns in Asia and Africa)

Numeral systems in Asia and Africa: A geolinguistic approach to linguistic patterns

FUKAZAWA, Mika
National Ainu Museum

Abstract: This article presents new maps of the numeral systems in Asia and Africa and provides interpretations, based on articles in the numeral systems of individual language groups/areas and Fukazawa's (2023a) overview in Chapter XVI of *Linguistic atlas of Asia and Africa III*. These new maps offer new interpretations and possibilities regarding the historical changes between quinary, decimal, vigesimal and other systems.*

Keywords: *Linguistic atlas of Asia and Africa*; numeral systems; counting; numbers

1. Introduction

The purpose of this study is to apply a geolinguistic approach to numeral systems in Asia and Africa. Numeral systems are mainly described as quinary, decimal, vigesimal, and other systems, as shown in Table 1. Yasugi (1990, 1995) describes numeral systems based on the principles of the combination of units and bases in a number sequence. When the powers of a base number are multiples of ten, the numeral system is called the decimal system. In the decimal system, the unit numbers range numbers from 1 to 9, and the base numbers are the ranks, such as 10 (B^1), 100 (B^2), and 1000 (B^3); the unit numbers are added or multiplied by the base numbers.

Fukazawa (2023a) presented a simplified map of the distribution of numeral systems in Asia and Africa. In this paper, new maps were drawn where numeral systems that correspond with the Asian and African geographical points can be seen. The numeral systems' linguistic data and original maps in Asia and Africa were produced by the language group and area experts and were already published in *Linguistic Atlas of Asia and Africa III* (Fukushima et al. eds. 2023). To prepare integrating individual maps, Fukazawa (2023a) assigned a symbol to each system and recommended that experts consider dividing numerals into three major categories, namely, 1 to 10, 10 to 20, and

FUKAZAWA, Mika. 2024. Numeral systems in Asia and Africa: A geolinguistic approach to linguistic patterns. *Studies in Geolinguistics* 4: 156–166. doi: https://doi.org/10.5281/zenodo.13948609
* The work was supported by the ILCAA Joint Research Project 'Studies in Asian and African Geolinguistics.'

even 20 or more (cf. Yasugi 1990, 1995), since each category may have a different system within a language. When a dialect or language has multiple systems, then the symbols are represented with a superimposed notation (Table 1). After the following section, these symbols and the original Asian and African maps will be used.

Table 1 Symbols of numeral systems

Symbol	Numeral system (Counting method)
／	Binary type
|	Quinary type
—	Decimal type
○	Vigesimal type
●	Other types, including quaternary and other base number types
N	None type
	Lack of data

Table 2 Language groups and colors

Color	Language group
Red	Ainu; Sinitic; Dravidian
Green	Japonic; Austroasiatic; Mongolic; Nilo-Saharan
Light blue	Korean; Uralic; Turkic; Caucasian; Kx'a
Navy blue	Kra-Dai; Indo-Aryan and Nuristani
Orange	Hmong-Mien; Austronesian; Tungusic; Iranian; Tuu
Broun	Tibeto-Burman; Semitic; Khoe-Kwadi
Black	Chukotko-Kamchatkan; Brushaski, Andamanese and language isolates in South Asia

2. Methods and results

The original numeral system maps of each language group or area have been created using the ArcGIS Online, Esri's web-based mapping software, as shown in Fukushima et al. eds. (2023). Figure 1 shows a new map of numeral systems in Asia and Africa in which the original maps were superimposed on each other, using the ArcGIS Online. Figures 2 and 3 show the enlarged maps of the eastern and western parts of Figure 1, respectively. As can be seen from these maps, the decimal system is widely used in Asia and Africa and has been adopted by many language groups and areas, most likely in the process of modernization. Some language groups have a monotonous distribution with the decimal system, such as Japonic, Korean, Sinitic, Hmong-Mien, Kra-Dai, Tungusic, Uralic, and Semitic. The decimal system must also be found in other areas with no numeral data. Here, it is important to note that few areas use more than three systems simultaneously. This may be due to historical changes in the numeral system. Section 3 will focus on regions that have numeral systems other than a decimal system.

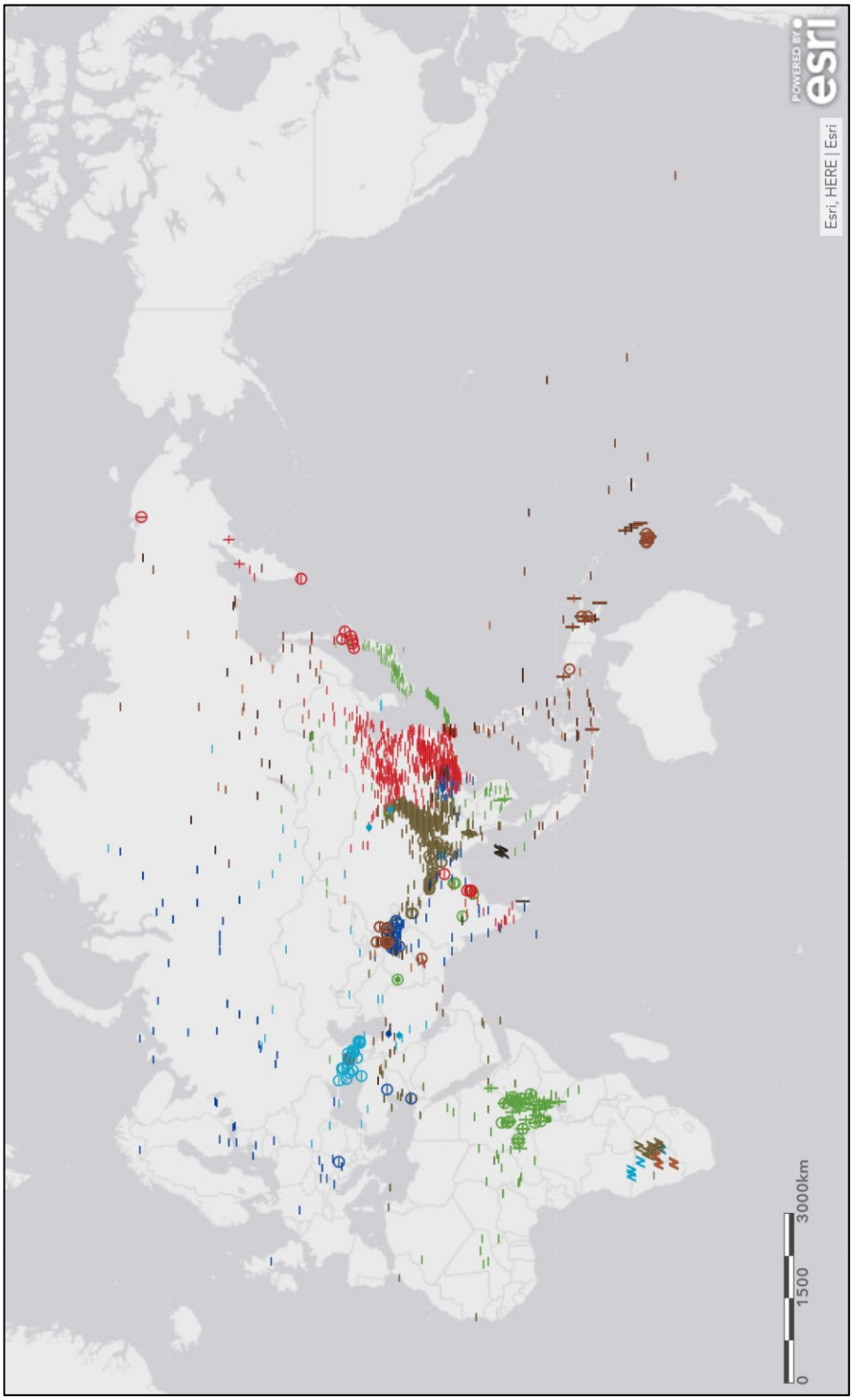


Fig. 1　Numeral systems in Asia and Africa

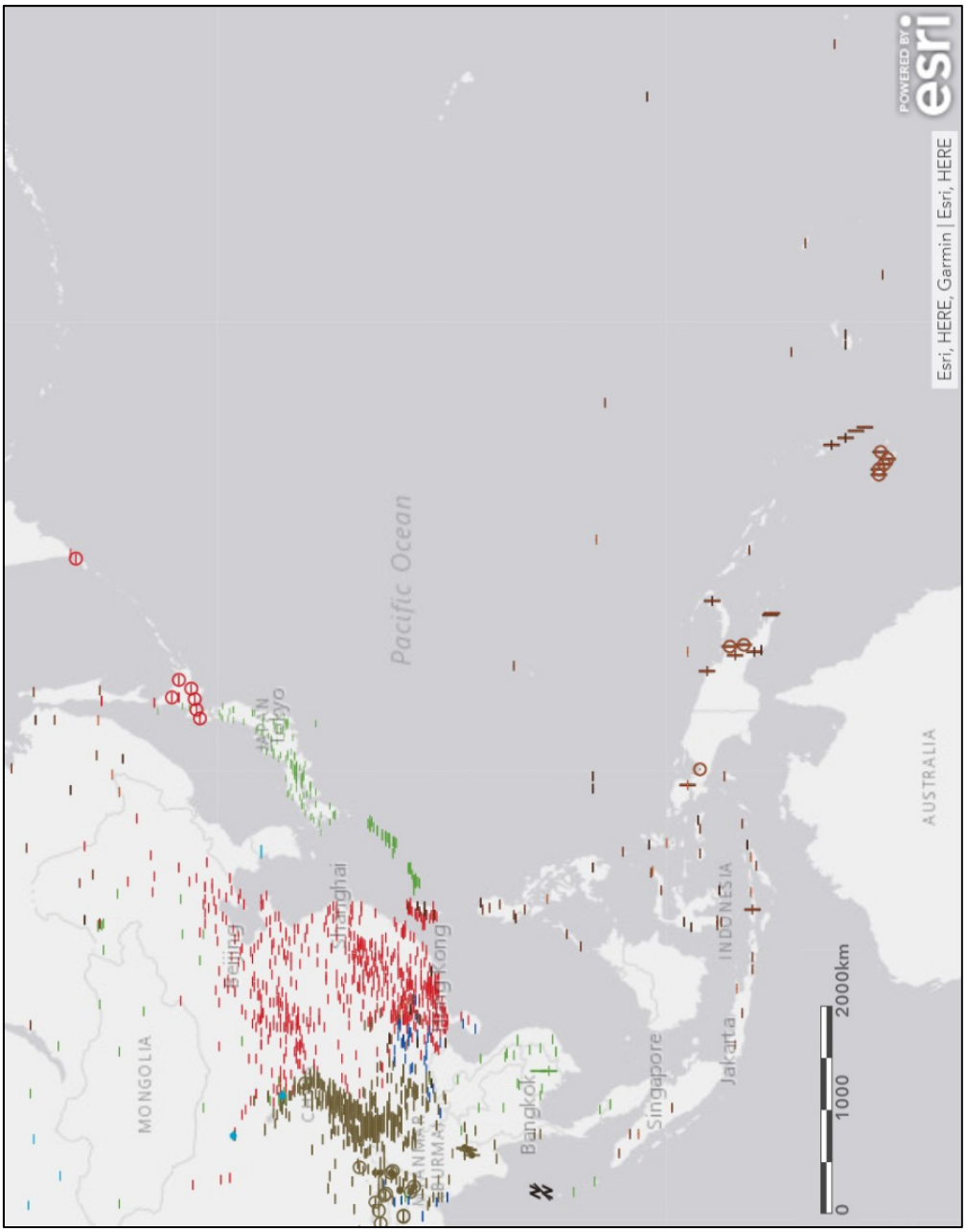


Fig. 2 Numeral systems in Asia and Africa (Eastern)

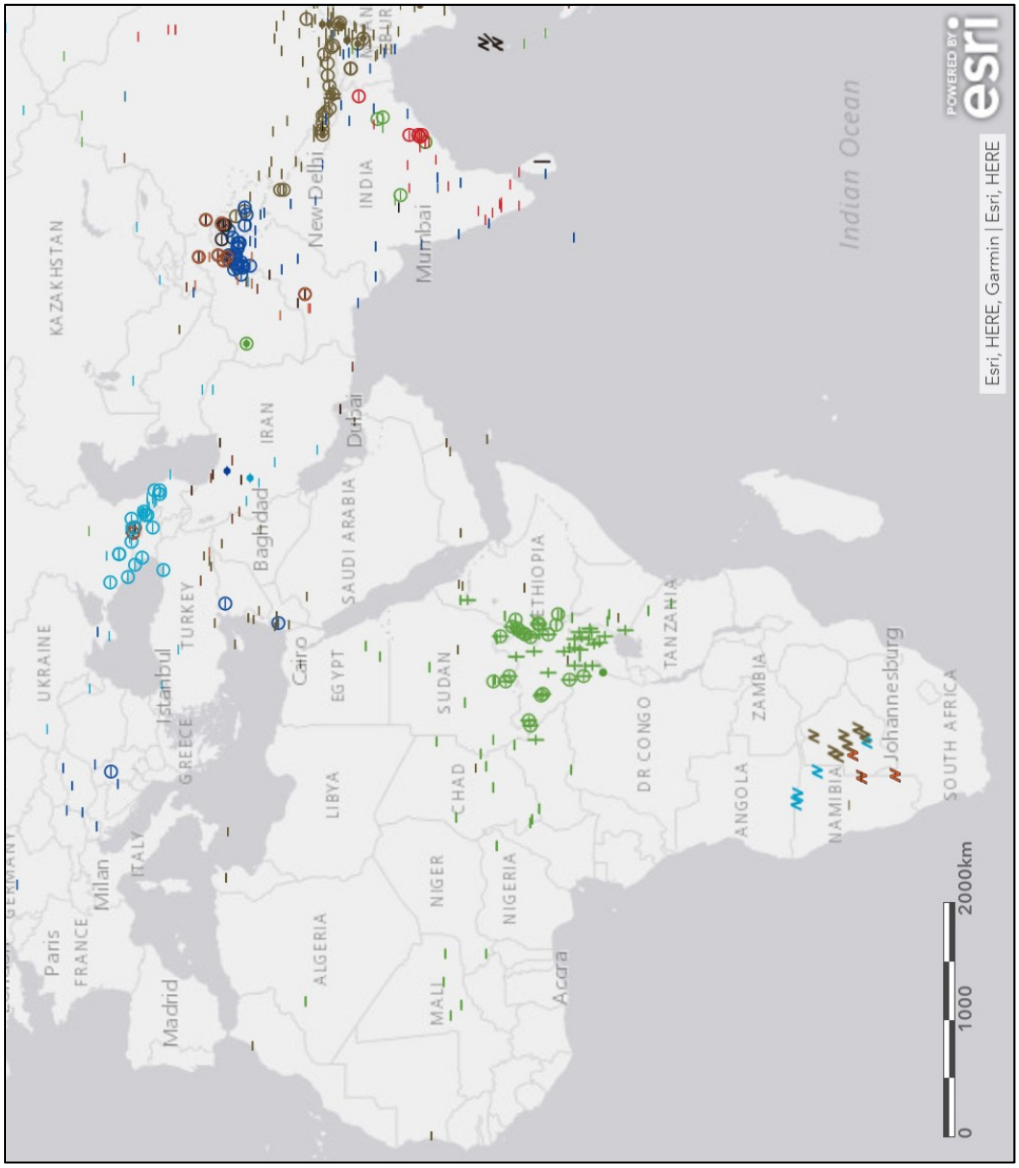


Fig. 3 Numeral systems in Asia and Africa (Western)

3. Discussion

In Figure 4, the decimal symbol, which is widespread throughout Asia and Africa, is removed from the map of Figure 1. Figures 5 and 6 show enlarged maps of the southern and northern parts of Figure 4, respectively. As for the vigesimal system, it is likely that more languages currently use or have used it around the Pacific coast or in the

southern part of Asia and Africa, although the data for expressing numbers larger than 20 are often lacking (see also Fukazawa 2023a). Typical vigesimal languages often use quinary and/or decimal system(s) for counting up to 20. As mentioned in Section 2, fewer areas use more than three systems simultaneously. It is most likely that, in most cases, the more widespread the decimal system has become, the more quinary and vigesimal systems have been eliminated or (partially) replaced by the decimal system.

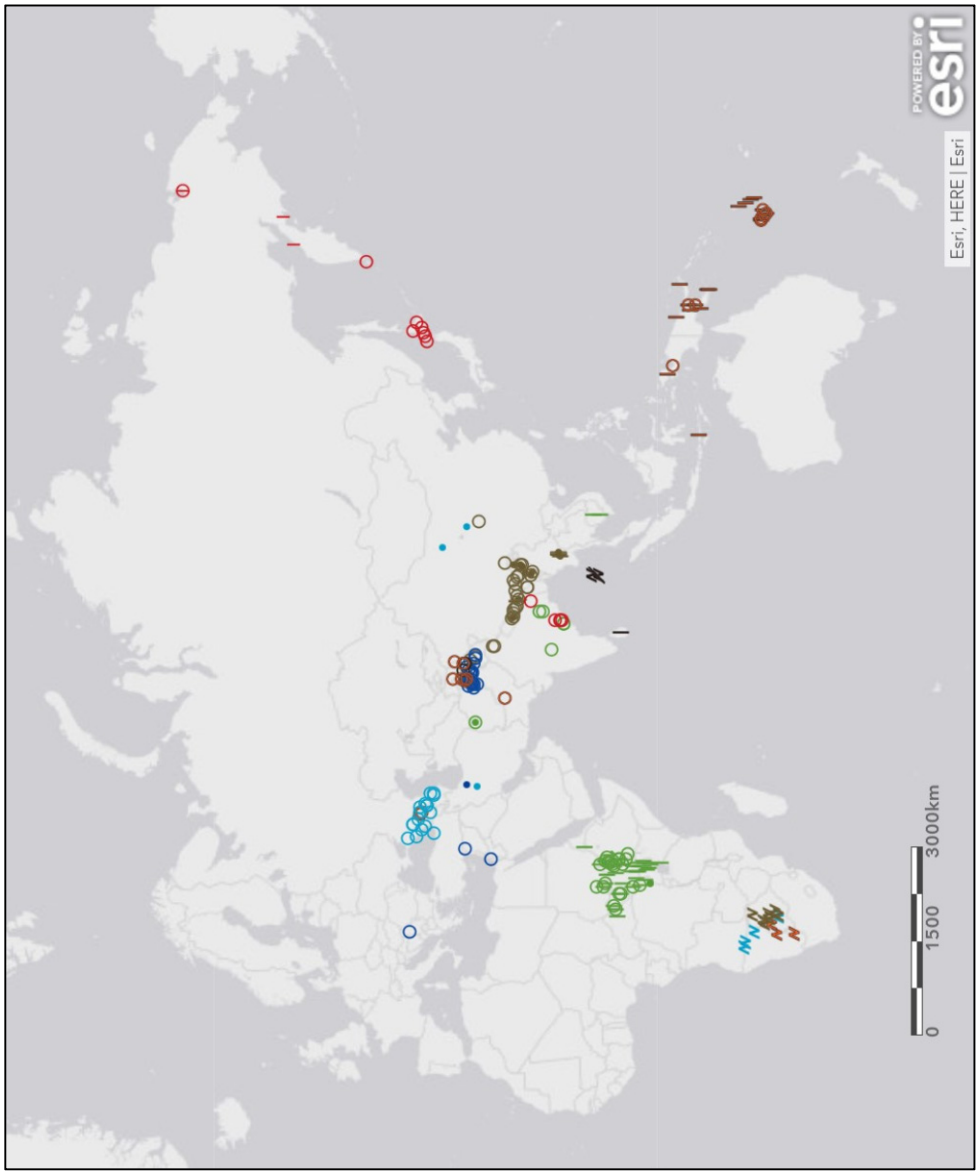


Fig. 4 Numeral systems other than the decimal system in Asia and Africa

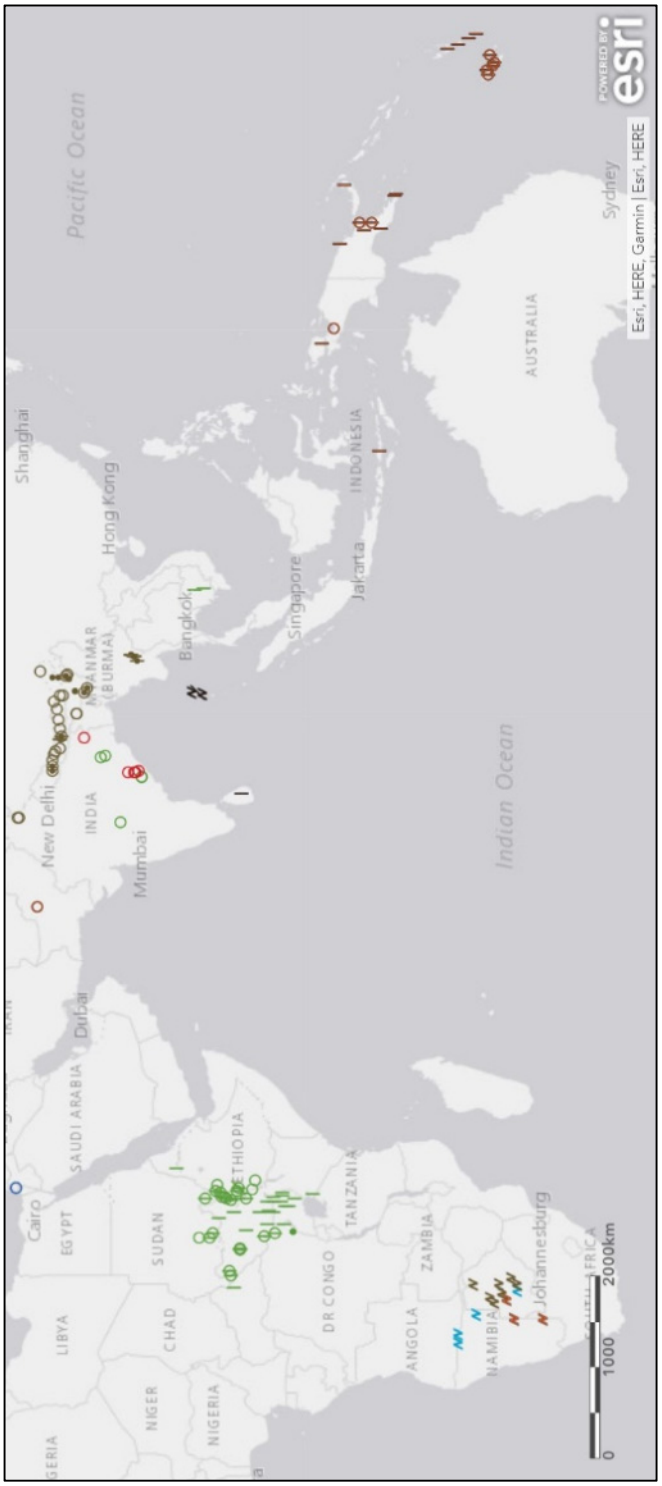


Fig. 5 Numeral systems other than the decimal system in Asia and Africa (Southern)

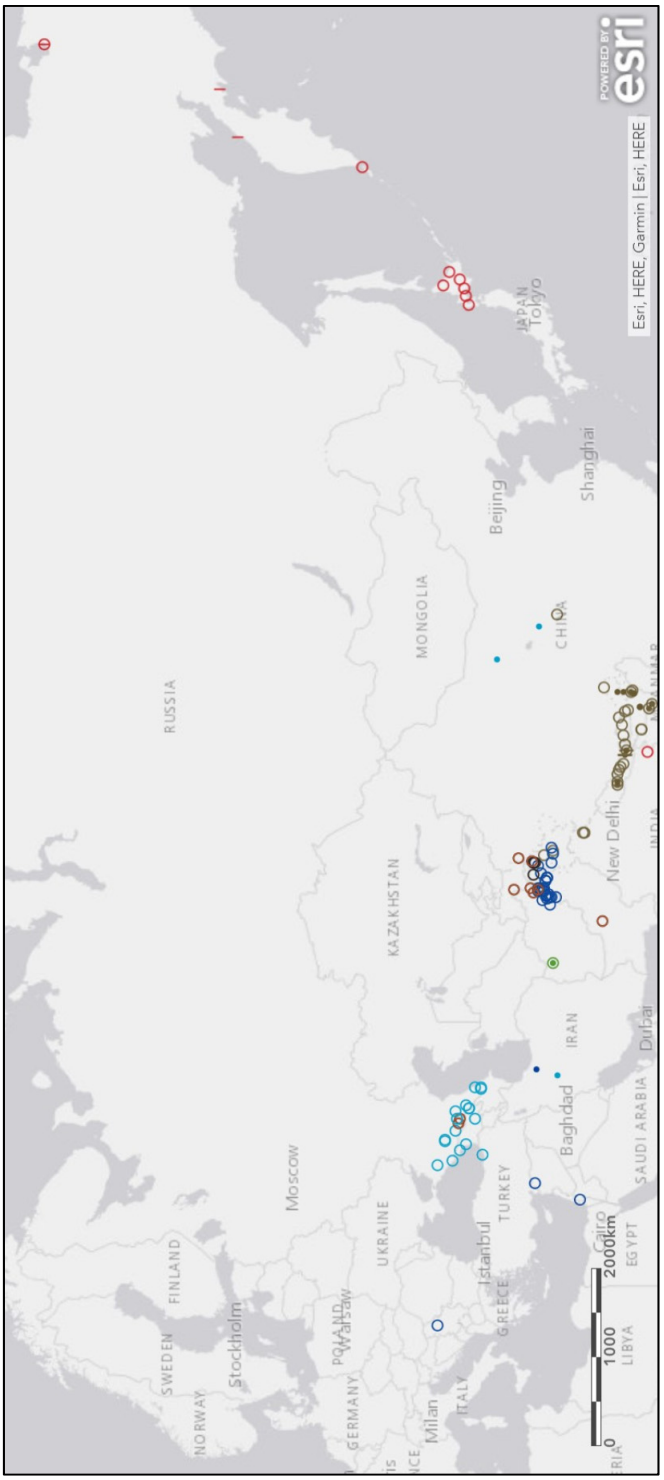


Fig. 6 Numeral systems other than the decimal system in Asia and Africa (Northern)

In language groups and areas where both the vigesimal and quinary systems exist, e.g., Papuan, Chukotko-Kamchatkan, and languages around the White Nile area, the geographical points with the quinary system are more widely distributed than those with the vigesimal system. In these regions, it is possible that the vigesimal system emerged and developed after the quinary system areas. Since these regions are on the periphery of Asia and Africa, and were less affected by the expansion of the decimal system, the vigesimal and quinary systems may actually be older than the decimal system, or the decimal system may be older than the quinary and vigesimal systems. The Proto-Austronesian language is known to use a decimal system in the study of the language reconstruction, and as a result of language contact with Papuan languages, languages on the New Guinea mainland, in New Caledonia, and Vanuatu received the quinary and vigesimal systems (Matsumoto 2006 and Utsumi 2023).

The etymons of the base numbers of 5, 10, 20 are often derived from the common concept of counting five fingers and toes, and the combined uses of the quinary/decimal/vigesimal systems are not only historical contacts, but perhaps also occur in sporadic or remote places coincidentally. The etymon of 5 is 'hand,' that of 10 is 'hands,' and the etymon of 20 is 'man,' namely meaning 'hands and feet.' (Ono 2023; see also Matsumoto 2006, Nakao 2023 and Utsumi 2023). Menninger (2011 [1969]), suggests "In pure counting natural divisions of 5, 10 and 20—'hand, hands, man'—are predominant. The physical representation of this process leads to the earliest numerals used by primitive peoples, which we thus designate as *row-characters*." Fukazawa (2023a, 2023b) mentioned that even in current non-quinary regions, a quinary system is sometimes retained as the etymological fragment of 'hand' for the number '5,' and it can be assumed that the quinary system may have been used, such as *asikne* '5' (*asik* < *aske* 'hand'; *ne* is a copula verb) in the Ainu language.

Binary, quaternary, and other systems, including the no numeral system (None type), appear to be older than the quinary, decimal, and vigesimal systems. In the case where numbers only go up to 3, the language or dialect can be categorized as a None type and would be considered the primitive type before the establishment of the other numeral systems (Matsumoto 2006). The None type examples can be seen in Bea and Jeru in South Asia (Yoshioka 2023), and Tuu, Kx'a and Khoe-Kwadi in the Kalahari Basin Area (Kimura and Nakagawa 2023).

4. Conclusion

In Fukazawa's previous analysis (2023a), the historical changes and transitions of the quinary, decimal, and vigesimal systems were difficult to analyze. However, the new maps of this study provide ways to consider them. Significantly, in the regions where the combined systems of quinary, decimal, and vigesimal systems exist, the etymons of the base numbers 5, 10, and 20 tend to be 'hand,' 'hands,' and 'man.' If the original numeral systems in a language with these base numbers were replaced by the decimal systems of another language, the original systems and forms would be strongly inherited and partially preserved in the language.

Notes for copyrights

The members of the 2020–2023 project involved in the investigation of the numeral systems are the following linguists: ONO Chikako (Chukotko-Kamchatkan), FUKAZAWA Mika (Ainu), Nakazawa Kohei (Japonic), FUKUI Rei (Korean), SUZUKI Fumiaki (Sinitic), TANG Baiyan and TAGUCHI Yoshihisa (Hmong-Mien), TOMITA Aika (Kra-Dai), SHIRAI Satoko, KURABE Keita, EBIHARA Shiho, IWASA Kazue, and SUZUKI Hiroyuki (Tibeto-Burman), MINEGISHI Makoto and SHIMIZU Masaaki (Austroasiatic), UTSUMI Atsuko (Austronesian), MATSUMOTO Ryo (Tungusic and Uralic), SAITÔ Yoshio (Mongolic and Turkic), YOSHIOKA Noboru (South Asia), KODAMA Nozomi (Dravidian), IWASAKI Takamasa (Iranian), SUZUKI Hiroyuki (Caucasian), NAGATO Youichi (Semitic), NAKAO Shuichiro (Nilo-Saharan), and KIMURA Kimihiko and NAKAGAWA Hirosi (Kalahari Basin Area).

References

Fukazawa, Mika (2023a) Numeral systems in Asian and African languages. *Linguistic Atlas of Asia and Africa III*, 143–145. Tokyo: Geolinguistic Society of Japan.

Fukazawa, Mika (2023b) Numeral systems in Ainu. *Linguistic Atlas of Asia and Africa III*, 148–149. Tokyo: Geolinguistic Society of Japan.

Fukushima, Chitsuko, Satoko Shirai, Mika Fukazawa, Hiroyuki Suzuki, and Mitsuaki Endo eds. (2023) *Linguistic Atlas of Asia and Africa III*. Tokyo: Geolinguistic Society of Japan. doi: https://doi.org/10.5281/zenodo.10223731

Kimura, Kimihiko and Hirosi Nakagawa (2023) Numeral systems in the Kalahari Basin Aria. *Linguistic Atlas of Asia and Africa III*, 205–206. Tokyo: Geolinguistic Society of Japan.

Matsumoto, Katsumi [松本克己] (2006) Sekai gengo no suushi-taikei to sono fuhenteki kiban「世界言語の数詞体系とその普遍的基盤」. *Sekai gengo eno shiza: rekishi gengogaku to gengo ruikeiron* 『世界言語への視座：歴史言語学と言語類型論』[A worldwide perspective on languages: Historical linguistics and linguistic typology], 303–312. Tokyo: Sanseido.

Menninger, Karl (1969) *Number words and number symbols*. Cambridge: M.I.T. Press. (Reprinted: Menninger, Karl (2011) *Number words and number symbols: A cultural history of numbers*. New York: Dover Publications)

Nakao, Shuichiro (2023) Numeral systems in Nilo-Saharan. *Linguistic Atlas of Asia and Africa III*, 201–204. Tokyo: Geolinguistic Society of Japan.

Ono, Chikako (2023) Numeral systems in Chukotko-Kamchatkan. *Linguistic Atlas of Asia and Africa III*, 146–147. Tokyo: Geolinguistic Society of Japan.

Utsumi, Atsuko (2023) Numeral Systems in Austronesian. *Linguistic Atlas of Asia and Africa III*, 178–180. Tokyo: Geolinguistic Society of Japan.

Yasugi, Yoshiho [八杉佳穂] (1990) Numeral systems of Middle American Indian languages「中米諸語の数体系」. *Bulletin of the National Museum of Ethnology* 14(3): 519–670. doi: https://doi.org/10.15021/00004297

Yasugi, Yoshiho (1995) *Native Middle American languages: An areal-typological perspective*. Suita: National Museum of Ethnology. doi: https://doi.org/10.15021/00003014

Yoshioka, Noboru (2023) Numeral systems in South Asia. *Linguistic Atlas of Asia and Africa III*, 189–191. Tokyo: Geolinguistic Society of Japan.

Publication history

Date received: 1 September 2024
Date accepted: 29 September 2024

Research article (Featured theme: Geolinguistic approaches to linguistic patterns in Asia and Africa)

Alignment patterns in Asia and Africa: A geolinguistic approach to linguistic patterns

SHIRAI, Satoko
The University of Tokyo

Abstract: This study draws linguistic maps of the alignment patterns in Asian and African languages to analyze the distributions of these patterns and examine their historical development, contrasting intransitive and highly transitive constructions with arguments in equal animacy. The maps that are provided here include an overall map, enlarged-view maps, and maps by type as excerpted from the overall map. The overall map imports the maps provided by the contributors to our previous project. The maps by type clarify the geographical distribution of the alignment patterns. These maps provide examples of the diffusion of linguistic patterns; the alignment patterns may show areal tendencies beyond the differences among genetic groups. Considerably, the ergative-absolutive, tripartite, and transitive patterns show distributions that suggest their histories: the ergative-absolutive type exhibits a continuous distribution in Asia, whereas the tripartite and transitive types are distributed across small but overlapping regions relative to the ergative-absolutive. These distributions suggest the diffusion of the ergative-absolutive type and the later development of tripartite and transitive types. The nominative-accusative patterns without verbal person marking also show an areal tendency in East and Southeast Asia.*

Keywords: *Linguistic atlas of Asia and Africa*; alignment; grammatical relation; linguistic pattern; linguistic area

1. Introduction

This study draws linguistic maps of the alignment patterns found in Asian and African languages to analyze their distributions and examine their historical process. These maps are based on our previous research project, Studies in Asian and African Geolinguistics, ILCAA Joint Research Project No. jrp000256, conducted at the Research Institute for Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies, from April 2020 to March 2023. We published 21 interpretive maps

SHIRAI, Satoko. 2024. Alignment patterns in Asia and Africa: A geolinguistic approach to linguistic patterns. *Studies in Geolinguistics* 4: 167–180. doi: https://doi.org/10.5281/zenodo.13948621

* This work was supported by the following grants: Tokyo University of Foreign Studies, April 2020–March 2023 ILCAA Joint Research Project "Studies in Asian and African Geolinguistics" (jrp000256), with Mitsuaki Endo as coordinator; and JSPS KAKENHI Grant Number JP23K00476. I would also like to thank the members of the 2020–2023 project for providing invaluable data on alignment patterns in each language group and list them at the end of this article.

of alignment patterns that were separately drawn by the contributors, each of whom was responsible for a different language group/area in Fukushima et al. eds. (2023) (*Linguistic Atlas of Asia and Africa Vol. 3*, henceforce LAAA3). Each article in LAAA3 also contains example sentences to precisely show the characteristics of the alignment patterns in each language group. In addition, Shirai (2023) conducted a fundamental analysis of the overall distribution.

Our project reflects the principles of linguistic geography. Linguistic geography, on principle, does not subject the entire system of a given language/dialect (lect) but separately deals with individual linguistic phenomena (Sibata 1969). Therefore, this project deals with the characteristics exhibited by the specifically restricted construction of each lect. Thus, the maps drawn in the project are intended to focus on individual construction across splits in many languages, and maps reflecting other constructions should be separately drawn. We focused on highly transitive constructions featuring arguments in equal animacy, having the following features, as contrasted with intransitive constructions (Shirai 2023: 63):

- The subject and object are equal in the empathy/animacy/person hierarchy: for example, both are in the third person or animals.
- The subject and object are definite, specific, and/or referential.
- The predicate is simple and/or plain in voice and/or mood.
- The predicate is verbal, with high volitionality and/or affectedness.
- The event described by the sentence has already occurred, is finished, or has been completed in the past.
- The information structure and word order are unmarked or most general.

This study differs from the preceding WALS project (Dryer and Haspelmath eds. 2013) in this respect, in addition to the number of points plotted on the maps.

We provide an example of the diffusion of linguistic patterns by illustrating that some alignment patterns are geographically distributed beyond genetic groups. In earlier study on language contact, language structures such as alignment patterns were considered conservative and less likely to diffuse across language groups (e.g., Thomason and Kaufman 1988: 74–75). Later studies, such as those of Ross (2001, 2007), provided counterexamples in which structures or patterns were borrowed without dense lexical borrowing. Matras and Sakel (2007) used the term pattern replication and investigated its mechanism.

This study provides an overall map connecting the previously developed maps in LAAA3 (Fig. 1), along with enlargements that show details of distributions (Figs. 2–4). Consequently, we excerpt feature-specific maps from the overall map (Figs. 5–10). Section 3 discusses what these distributions suggest for language history, and Section 4 draws conclusions.

2. Methods and results

Table 1 illustrates the basic symbols used in the maps. The letters A, S, and P in the header row indicate the subject of a transitive, the sole argument of an intransitive, and the object of a transitive, respectively. In this project, the contributors used these symbols for basic types, being allowed to use modified symbols to express splits. The project provided symbols for types that seemed logically incongruous, as we intended to be exhaustive, but they did not exist. For this reason, these unattested types are shaded, and the symbols are removed from Table 1 below. For details of the classification, please refer to Shirai (2023).

Table 1 Major types and basic symbols (revised version of Table 1 from Shirai 2023)

	Nominal case marking	Verbal person marking	Double marking	No marking
AS|P (Nominative-accusative)	A1	A2	A3 AX3	A4
A|SP (Ergative-absolutive)	B1	B2	B3 BX3	B4
S1|S2 (Active-inactive or split intransitive)	C1	C2	C3 CX3	C4
A|S|P (Tripartite)	D1	D2	D3 DX3	D4
ASP (Neutral)	E1	E2	E3 EX3	E4
More than three splits	☆ F			
Other types	(Unspecified symbol) G			

Symbols are presented in different colors corresponding to each language group; however, the distinguishable number of colors is limited. Therefore, some distant language groups are given the same color. Table 2 shows the language groups and the colors given on the map.

Table 2 The language groups and colors

Language group	Color
Ainu	Red
Andamanese and language isolates (South Asia)	Black
Austroasiatic	Green
Austronesian	Orange
Caucasian	Light blue
Chukotko-Kamchatkan	Black
Dravidian	Red
Hmong-Mien	Orange
Indo-Aryan and Nuristani (South Asia)	Navy
Iranian	Orange
Japonic	Green
Khoe-Kwadi (Kalahari Basin Area)	Brown
Korean	Light blue
Kra-Dai	Navy
Kx'a (Kalahari Basin Area)	Light blue
Mongolic	Green
Niger-Congo	Navy
Nilo-Saharan	Green
Semitic	Brown
Sinitic	Red
Tibeto-Burman	Brown
Tungusic	Orange
Turkic	Light blue
Tuu (Kalahari Basin Area)	Orange
Uralic	Light blue

2.1. Alignment patterns in Asia and Africa: all types

The contributors' maps were imported in the form of layers (Fig. 1). The map shows the alignment patterns of 3,254 points in Asia and Africa, plus parts of Oceania (Austronesian languages) and Europe (Indo-Aryan, Turkic, and Uralic languages). The symbols on the map indicate the principles described above (Tables 1 and 2), except some language groups that use symbols other than those listed above, for which a legend is provided below the map. For details of these classifications, please refer to the articles in LAAA3.

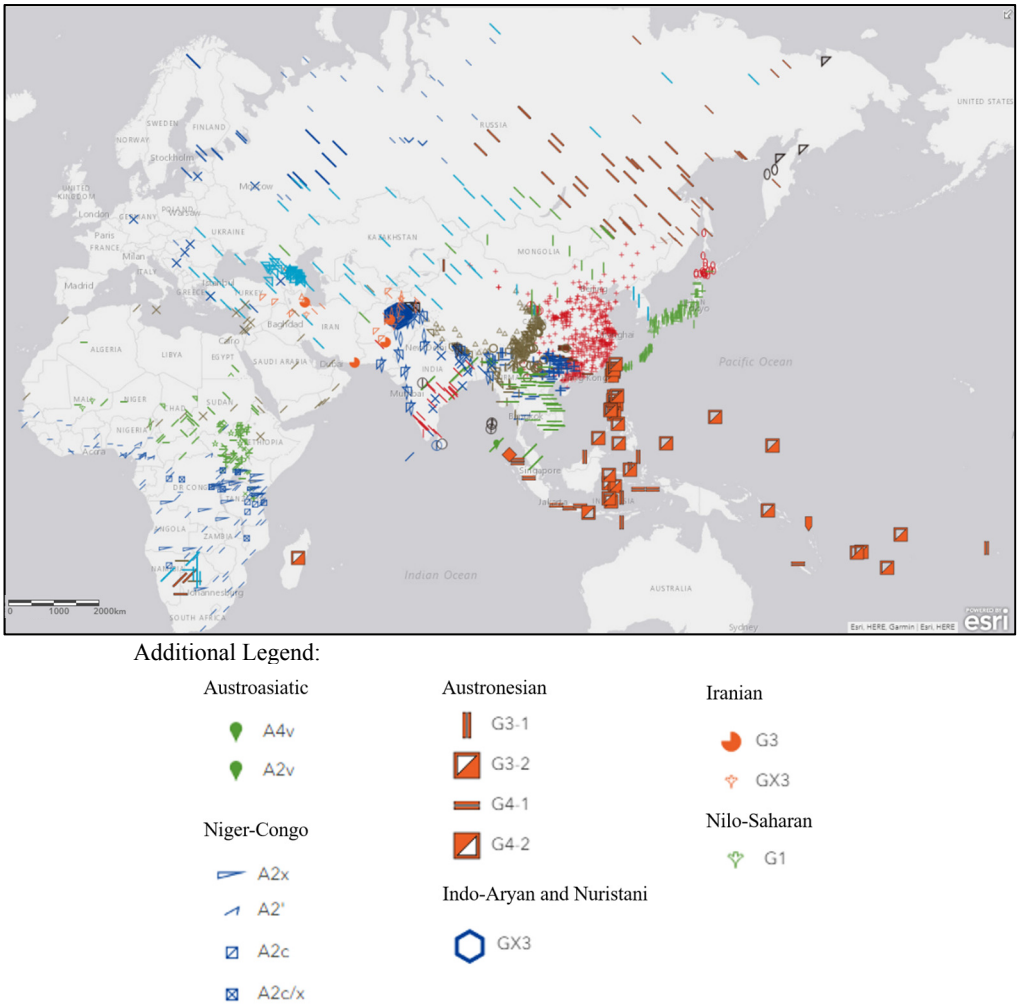


Fig. 1 Alignment patterns in Asia and Africa

The enlarged map shows the distribution of the alignment patterns in areas featuring dense point concentrations in Figs. 2–4, illustrating the distribution in South and Southeast Asia (Fig. 2), Caucasus and Anatolia (Fig. 3), and Central and South Africa (Fig. 4), respectively.

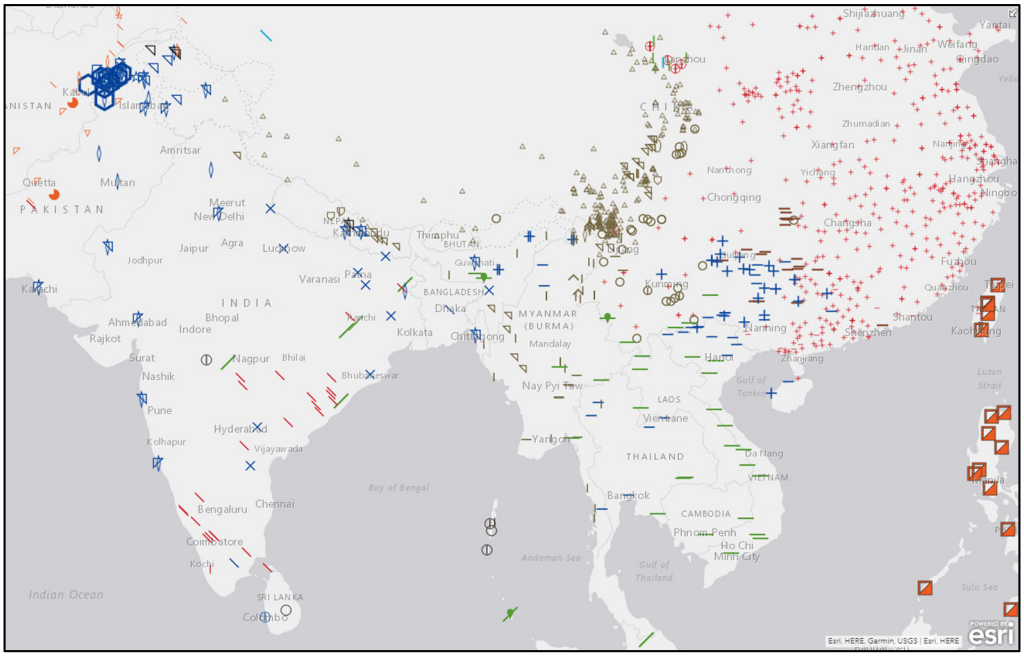


Fig. 2 Enlarged map of alignment patterns: South and Southeast Asia

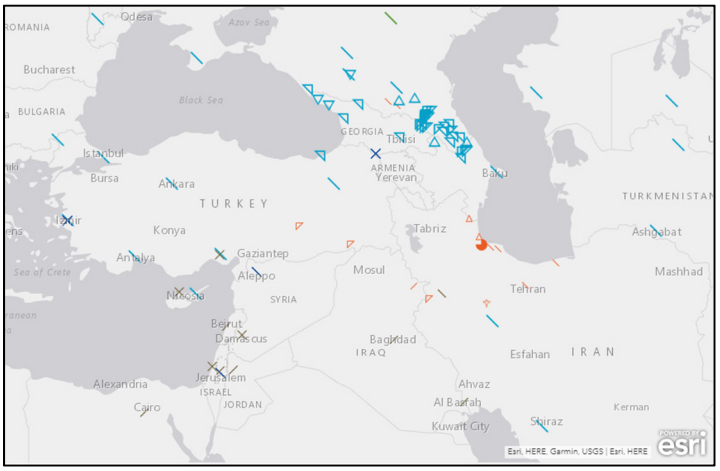


Fig. 3 Enlarged map of alignment patterns: the Caucasus and Anatolia

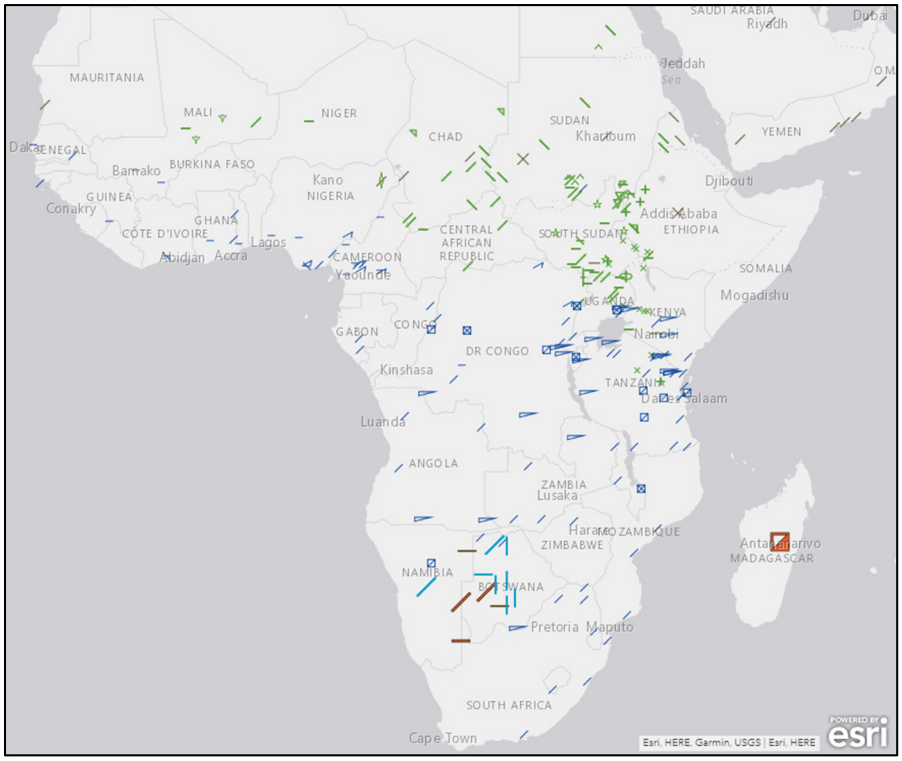


Fig. 4 Enlarged map of alignment patterns: Central and South Africa

2.2. Geographical distribution of each type of alignment

The following maps are provided: nominative-accusative (Fig. 5), ergative-absolutive (Fig. 6), active-inactive (Fig. 7), tripartite (Fig. 8), neutral or hierarchical (Fig. 9), and other types (Fig. 10). Layers are excerpted from each author's maps; however, some symbols have been resized for readability.

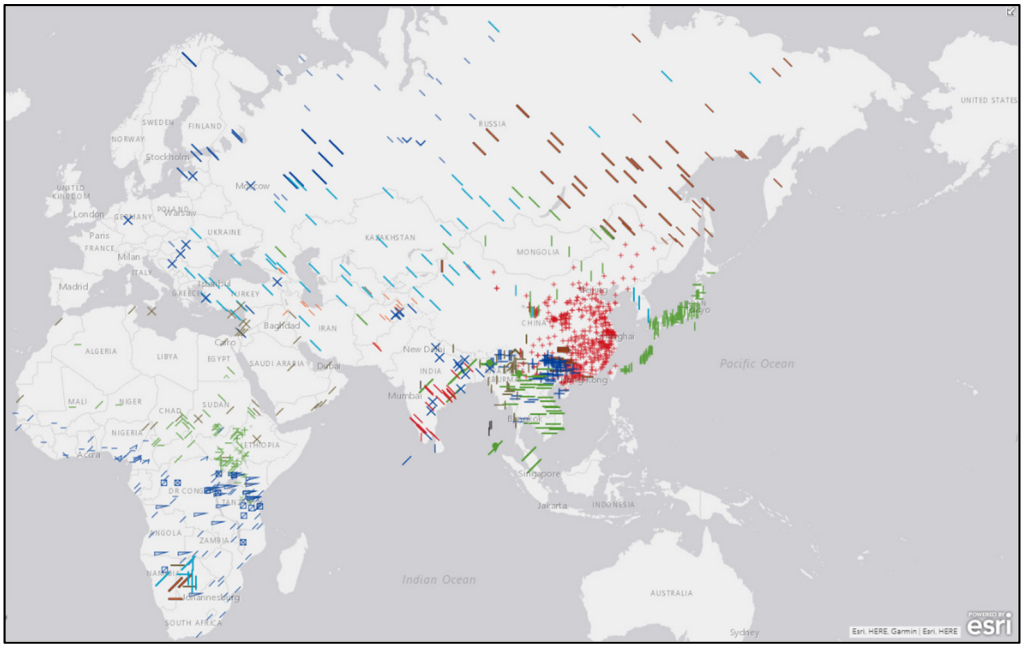


Fig. 5 Alignment patterns in Asia and Africa: Type A (nominative-accusative)

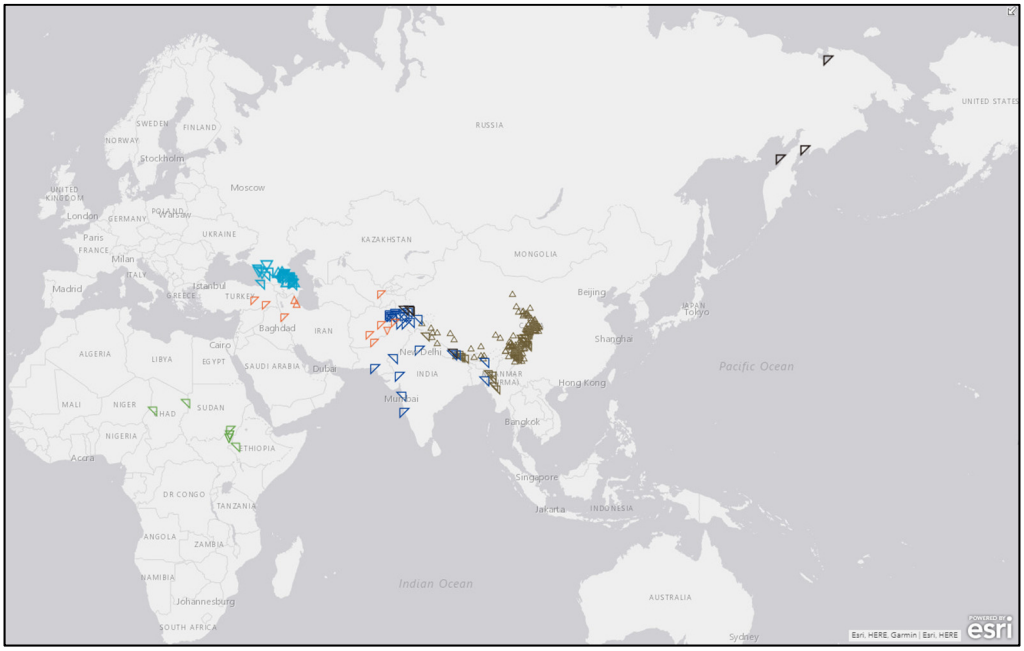


Fig. 6 Alignment patterns in Asia and Africa: Type B (ergative-absolutive)

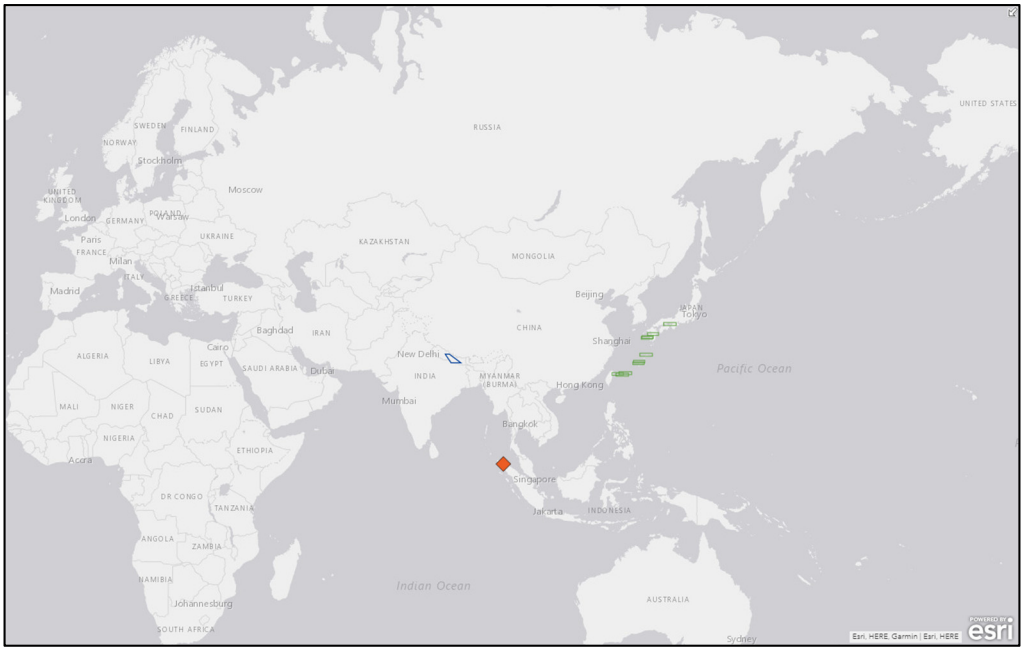


Fig. 7 Alignment patterns in Asia and Africa: Type C (active-inactive or split intransitive)

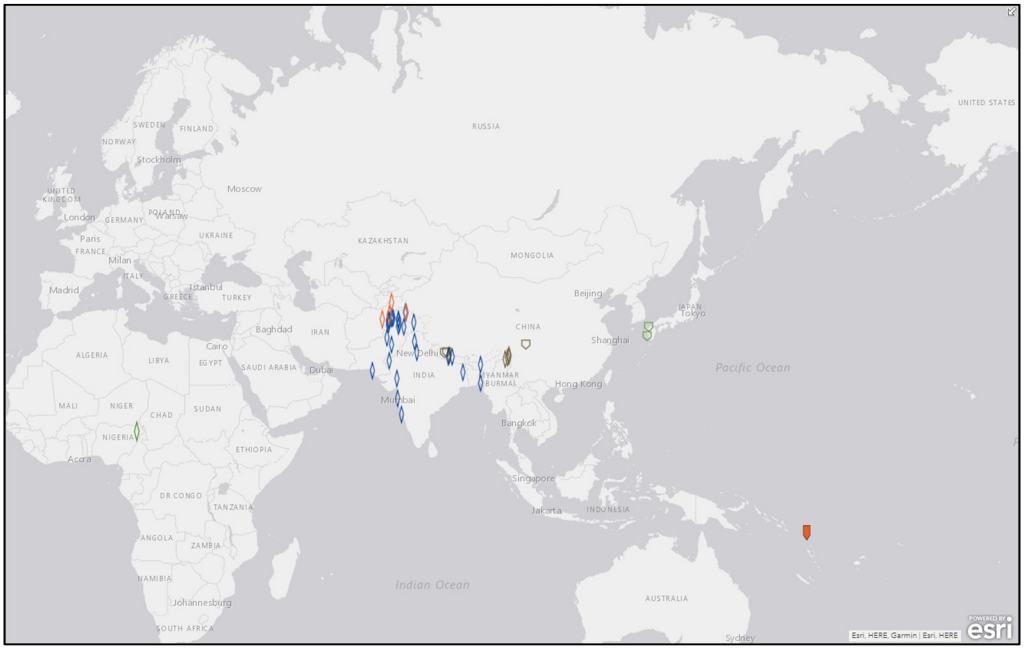


Fig. 8 Alignment patterns in Asia and Africa: Type D (tripartite)

Fig. 9　Alignment patterns in Asia and Africa: Type E (neutral or hierarchical)

Fig. 10　Alignment patterns in Asia and Africa: Types F (complex split) and G (other)

3. Discussion

The maps presented above show certain substantial geographical distributions, suggesting the history of the diffusion of linguistic patterns. Herein, we examine the distributions of each alignment pattern.

Herein, Type A refers to morphological and syntactic nominative-accusative patterns, distinguishing P from A and S. All language groups, except the Chukotko-Kamchatkan, Ainu, Austronesian, and Caucasian languages, involve lects with nominative-accusative patterns. Therefore, its distribution is much broader than that of other types, covering Asia and Africa.

Type A4 (neutral marking but syntactically nominative-accusative, marked by the symbol －) has distribution in peripheral regions of Afro-Eurasia, that is, East and Southeast Asia and West and South Africa, as noted in Shirai (2023). This map provides some additional insights. Type A1 (marked with ｜) is concentrated in the eastern regions, suggesting an areal tendency. Languages that have nominative-accusative verbal person marking (A2, marked with ／, and A3, with ＼) are far more widely distributed than those that do not have it (A1 and A4).

The ergative-absolutive pattern plotted in Fig. 6 presents a continuous distribution (Shirai 2023), except for Chukotko-Kamchatkan (Northeast Asia) and part of the Nilo-Saharan languages (Central Africa). Moreover, ergative-absolutive-type verbal person marking (Types B2 and B3, marked with ▽ and ◸, respectively) is found across the western half of the distribution, except the Chukotko-Kamchatkan languages.

Notably, Suzuki (2021: 58), in an earlier publication of our project, mentioned that Tibetic languages (Tibeto-Burman), Domaaki (Indo-Aryan), and Hunza Burushaski (a language isolate in South Asia) feature a common source for their word forms denoting ‘rice’, although no intense language contact has been reported between these languages. They are all located in a continuous geographical area in which the ergative-absolutive pattern is found in Fig. 6. These facts suggest an early interaction among peoples in the area that diffused the ergative-absolutive type.

As shown in Fig. 7, the split intransitive pattern is absent from Africa and very rare in Asia: one Indo-Aryan language, one Austronesian language, and various dialects of Japanese and Ryukyuan feature this pattern. Due to the few attested lects, discussing their geographical distribution across language groups is difficult.

The tripartite pattern in Asia overlaps with the ergative-absolutive pattern; however, it shows a narrow geographical distribution. As indicated in Fig. 8, languages in continuous regions of western India and the Himalayas’ western, southern, and eastern foothills feature the tripartite pattern. Moreover, in South Asia, most languages with the tripartite pattern also have the ergative-absolutive pattern as a split (Yoshioka 2023;

see also Fig. 2). These distributions suggest that, in this area, the tripartite pattern is newer than the ergative-absolutive pattern. Further, it is also found in Southwestern Japanese dialects, one Nilo-Saharan language, and one Austronesian language.

Fig. 9 presents the distribution of patterns with no case marking or syntactic distinction with respect to whether A or P is treated in the same way as S. This pattern is mainly found in Northeast Asia, Inland China, and South Asia. One Nilo-Saharan language also has this pattern, with a split in the nominative-accusative case marking.

Finally, Fig. 10 presents the distribution of languages with a complex split (Type F) or patterns other than A–E (Type G). We find two substantial patterns in "Other types" (Type G): transitive and symmetrical patterns. The transitive pattern (indicated by the symbols ⬡, ◕, and ✲; the complex split type in Iranian, which is indicated by ☆, also features a split in this pattern) shows a regional distribution around the borders of Afghanistan, Pakistan, and Iran, in addition to Northwest Iran. These languages belong to the Indo-Aryan, Nuristani, and Iranian language groups of the Indo-European family. The distribution overlaps with, but is narrower than, the ergative-absolutive pattern. This distribution indicates that the transitive is also newer than the ergative-absolutive pattern, as Iwasaki (2023: 121) noted. The symmetrical pattern is represented by many Austronesian languages (Utsumi 2023). In addition to these patterns, some Nilo-Saharan languages show a bidirectional pattern (Nakao 2023: 129).

4. Conclusion

This study developed linguistic maps to illustrate alignment patterns in Asia and Africa in relation to the contrast between intransitive and highly transitive constructions and the geolinguistic analysis of the maps. The alignment patterns are motivated by a distinction between grammatical relations and are limited in terms of their options; therefore, the same type may coincidentally occur at different points. However, this study also confirmed the existence of areal tendencies that reflect past diffusion of linguistic patterns. The geographical distributions of the ergative-absolutive, tripartite, and transitive patterns suggest their histories: ergative-absolutive patterns exhibit a continuous distribution around the Himalayas, South Asia, and West Asia, whereas the tripartite and transitive patterns are distributed in small but overlapping regions to the ergative-absolutive. The distribution of the ergative-absolutive partially overlaps with the distribution of a word form found earlier in our project. These facts suggest that the ergative-absolutive type diffused through past interactions, and the tripartite and transitive types arose in the ergative-absolutive patterns. The nominative-accusative patterns that do not show verbal person marking also show an areal tendency.

Notably, this study was not conducted to capture the general characteristics of each language but to focus on the restricted constructions, as mentioned in Section 1, because it is based on the geolinguistic method. Different distributions will be found under different conditions; therefore, investigating constructions under various conditions is necessary to thoroughly understand this issue. Comparing the maps of different features enables us to find more substantial distributions of linguistic features that can be used to tell people's history.

Notes for copyright

The 2020-2023 project members involved in the alignment study are gratefully acknowledged below: ONO Chikako (Chukotko-Kamchatkan), FUKAZAWA Mika (Ainu), KIBE Nobuko, NAKAZAWA Kohei, and YOKOYAMA Akiko (Japonic), FUKUI Rei (Korean), SUZUKI Fumiki (Sinitic), TAGUCHI Yoshihisa (Hmong-Mien), ENDO Mitsuaki, TOMITA Aika, and HIRANO Ayaka (Kra-Dai), SHIRAI Satoko, EBIHARA Shiho, IWASA Kazue, KURABE Keita, and SUZUKI Hiroyuki (Tibeto-Burman), MINEGISHI Makoto and SHIMIZU Masaaki (Austroasiatic), UTSUMI Atsuko (Austronesian), MATSUMOTO Ryo (Tungusic and Uralic), SAITÔ Yoshio (Mongolic and Turkic), YOSHIOKA Noboru (South Asia), KODAMA Nozomi (Dravidian), IWASAKI Takamasa (Iranian), SUZUKI Hiroyuki (Caucasian), NAGATO Youichi (Semitic), NAKAO Shuichiro (Nilo-Saharan), SHINAGAWA Daisuke and KOMORI Junko (Niger-Congo), and KIMURA Kimihiko and NAKAGAWA Hirosi (Kalahari Basin Area).

References

Dryer, Matthew S. and Martin Haspelmath (eds.) (2013) WALS Online (v2020.3) [Data set]. Zenodo. doi: https://doi.org/10.5281/zenodo.7385533, URL: https://wals.info

Fukushima, Chitsuko, Satoko Shirai, Mika Fukazawa, Hiroyuki Suzuki and Mitsuaki Endo (eds.) (2023) *Linguistic Atlas of Asia and Africa III*. Tokyo: Geolinguistic Society of Japan. doi: https://doi.org/10.5281/zenodo.10223731 (LAAA3)

Iwasaki, Takamasa (2023) Alignment in Iranian. In: Chitsuko Fukushima, Satoko Shirai, Mika Fukazawa, Hiroyuki Suzuki and Mitsuaki Endo (eds.) *Linguistic atlas of Asia and Africa III*, 119–122. Tokyo: Geolinguistic Society of Japan.

Matras, Yaron and Jeanette Sakel (2007) Investigating the mechanisms of pattern replication in language convergence. *Studies in Language* 31: 829–865. doi: https://doi.org/10.1075/sl.31.4.05mat

Nakao, Shuichiro (2023) Alignment in Nilo-Saharan. In: Chitsuko Fukushima, Satoko Shirai, Mika Fukazawa, Hiroyuki Suzuki and Mitsuaki Endo (eds.) *Linguistic atlas of Asia and Africa III*, 128–135. Tokyo: Geolinguistic Society of Japan.

Ross, Malcolm (2001) Contact-induced change in Oceanic languages in North-West Melanesia. In: Aikhenvald, Alexandra Y. and R. M. W. Dixon (eds.). *Areal diffusion and genetic inheritance*, 134–166. Oxford: Oxford University Press. doi: https://doi.org/10.1093/oso/9780198299813.003.0006

Ross, Malcolm (2007) Calquing and Metatypy. *Journal of Language Contact* 1: 116–143. doi: https://doi.org/10.1163/000000007792548341

Shirai, Satoko (2023) Alignment in Asia and Africa. In: Chitsuko Fukushima, Satoko Shirai, Mika Fukazawa, Hiroyuki Suzuki and Mitsuaki Endo (eds.) *Linguistic atlas of Asia and Africa III*, 63–66. Tokyo: Geolinguistic Society of Japan.

Sibata, Takesi [柴田武] (1969) *Gengochirigaku no hoohoo*『言語地理学の方法』 [Methods in linguistic geography]. Tokyo: Chikumashobo.

Utsumi, Atsuko (2023) Alignment in Austronesian. In: Chitsuko Fukushima, Satoko Shirai, Mika Fukazawa, Hiroyuki Suzuki and Mitsuaki Endo (eds.) *Linguistic atlas of Asia and Africa III*, 99–103. Tokyo: Geolinguistic Society of Japan.

Yoshioka, Noboru (2023) Alignment in South Asia. In: Chitsuko Fukushima, Satoko Shirai, Mika Fukazawa, Hiroyuki Suzuki and Mitsuaki Endo (eds.) *Linguistic atlas of Asia and Africa III*, 112–116. Tokyo: Geolinguistic Society of Japan.

Publication history

Date received: 30 August 2024
Date accepted: 20 September 2024

Research article (Featured theme: Geolinguistic approaches to linguistic patterns in Asia and Africa)

Stop series in Asia and Africa: A geolinguistic approach to linguistic patterns

SUZUKI, Hiroyuki
Kyoto University

Abstract: In accordance with the overview provided in Chapter XIII "Stop series" of the *Linguistic atlas of Asia and Africa II*, this article presents individual linguistic maps for each feature on the stop series (denti-alveolar sounds), thereby offering further support for the preceding descriptions.*

Keywords: *Linguistic atlas of Asia and Africa*; consonant system; stop series; denti-alveolar; geographical distribution

1. Introduction

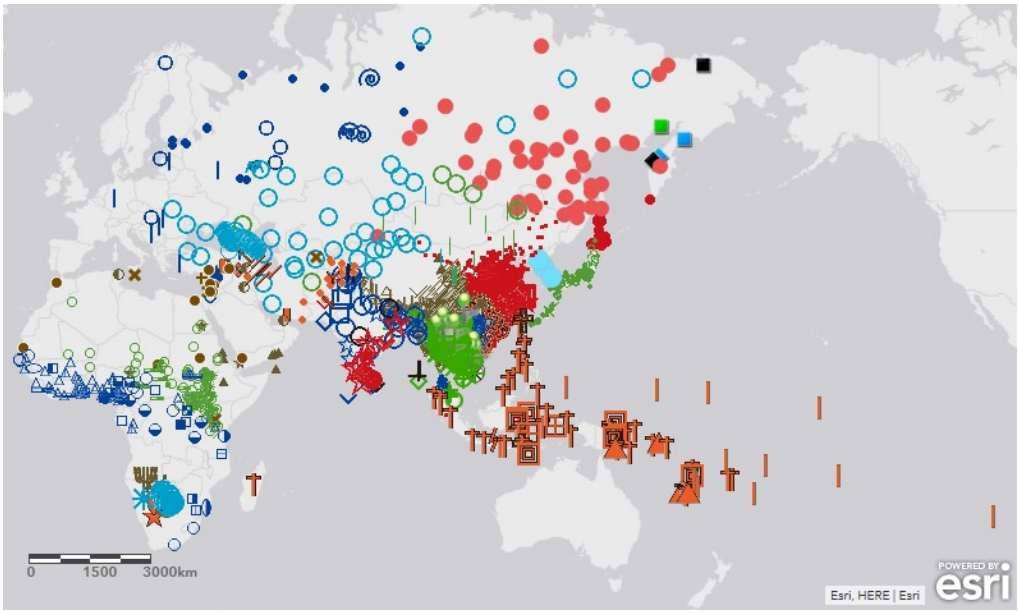
This article provides a more detailed account of the material presented in Suzuki (2022), which is the overview prepared for Chapter XIII of *Linguistic atlas of Asia and Africa II* (LAAA-2; Suzuki et al. 2023). It also incorporates the entire data set on individual maps for each topic addressed in the chapter, thereby, offering a comprehensive visual representation of the information.

In contrast to the other contributions on the present featured theme, including sibling terms (Fukushima 2024), alignment (Shirai 2024), and numeral systems (Fukazawa 2024), the cross-linguistic overview of the stop series topic is challenging due to the need to summarise the data based on a common criterion; see Figure 1 for a reference. The series is represented by dental/denti-alveolar/alveolar (henceforth D/A) plosives and nasals. The project examined a system of D/A stop series that comprised the following components: /th-t-t'-d-dh-ɗ-nd-nt-nth-n-nh/ ([t^h-t-t'-d-$d^ɦ$-ɗ-^{n}d-^{n}t-$^nt^h$-n-$n̥$] for

SUZUKI, Hiroyuki. 2024. Stop series in Asia and Africa: A geolinguistic approach to linguistic patterns. *Studies in Geolinguistics* 4: 181–202. doi: https://doi.org/10.5281/zenodo.13948632

* The work was supported by the ILCAA Joint Research Project 'Studies in Asian and African Geolinguistics' and JSPS KAKENHI Grant Number JP17H04774. I should like to thank Chitsuko Fukushima, Satoko Shirai, and Mitsuaki Endo for designing this article and creating synthesised maps.

a phonetic description). The symbols of each language group are distinguished by different colours, as shown in Table 1.



N.B. See Suzuki et al. (2023) for the detailed legend for each language group.

Fig. 1 The whole dataset of the stop series from LAAA-2.

Table 1 The language groups and colours

Language group	Color	Language group	Color
Ainu	Red	Korean	Light blue
Andamanese and language isolates (South Asia)	Black	Kra-Dai	Navy
		Kx'a (Kalahari Basin Area)	Light blue
Austroasiatic	Green	Mongolic	Green
Austronesian	Orange	Niger-Congo	Navy
Caucasian	Light blue	Nilo-Saharan	Green
Chukotko-Kamchatkan	Black	Semitic	Brown
Dravidian	Red	Sinitic	Red
Hmong-Mien	Orange	Tibeto-Burman	Brown
Indo-Aryan and Nuristani (South Asia)	Navy	Tungusic	Orange
		Turkic	Light blue
Iranian	Orange	Tuu (Kalahari Basin Area)	Orange
Japonic	Green	Uralic	Light blue
Khoe-Kwadi (Kalahari Basin Area)	Brown		

The article deals with the following features: D/A series types, ejective, implosive, pharyngealisation, prenasalisation, voiceless nasal, glottalised stops, and the lack of D/A nasals. The D/A series types include several criteria.

2. D/A series types

Various D/A series types are attested in Asian and African languages. This section examines five types: /t/ only, non-nasal tripartite contrast /th-t-d/, non-nasal quadripartite contrast /th-t-d-dh/, voicing contrast, and aspiration contrasts. Although historical studies indicate that these types are mutually related, here, each description reflects its synchronic status in conjunction with its phonation and suprasegmental features (e.g., Zhu 2010 in Sinitic; Tournadre and Suzuki 2023 in Tibetic).

2.1. /t/ only (bipartite contrast /t-n/)

A two-way distinction represents the minimum D/A stop series system, in which the /t-n/ components are most widely attested in many languages, including Chukotko-Kamchatkan (Chukchi and Koryak), Ainu, Ōgami Ryukyuan of Japonic, Car Nicobarese of Austroasiatic, Austronesian, Khanty and Mansi of Uralic, Chuvash of Turkic, Dravidian, Cypriot Arabic, Pökoot of Nilo-Saharan, and Niger-Congo.

Other two-way distinction patterns are /th-t/ (a limited number of Sinitic varieties, Kra-Dai, Tujia of Tibeto-Burman) and /t-d/ (a limited number of Niger-Congo). As all these features are found in a vast geographical area spanning Asia and Africa, maps are not provided.

2.2. Non-nasal tripartite contrast

A non-nasal tripartite distinction /th-t-d/ is widely attested in several language groups, including Sinitic (particularly Wu varieties), Hmong-Mien, Kra-Dai, Tibeto-Burman, Austroasiatic, Austronesian, Indo-Aryan, Burushaski, Armenian, and Niger-Congo. Of these language groups, some varieties exhibit this distinction with prenasalised features, including /th-t-d-nd/ and /th-t-d-nd-nth/. Figures[1] 2 and 3 illustrate the language groups that demonstrate this non-nasal tripartite distinction. The distribution of these languages is concentrated in East and Southeast Asia.

Other tripartite contrasts comprising voicing, aspiration, prenasalisation, and glottalisation distinctions are also attested, albeit to a lesser extent. These include /t-d-nd/ (Japonic), /t-tˀ-d/ (Japonic), /th-t-tˀ/ (Korean), /th-d-dh/ (Sinitic), and /t-d-dh/ (Dravidian). As all of these are minority patterns, maps are not provided.

[1] For technical reasons, the tiny symbols may persist on the maps. However, their presence does not indicate the actual presence of the corresponding feature.

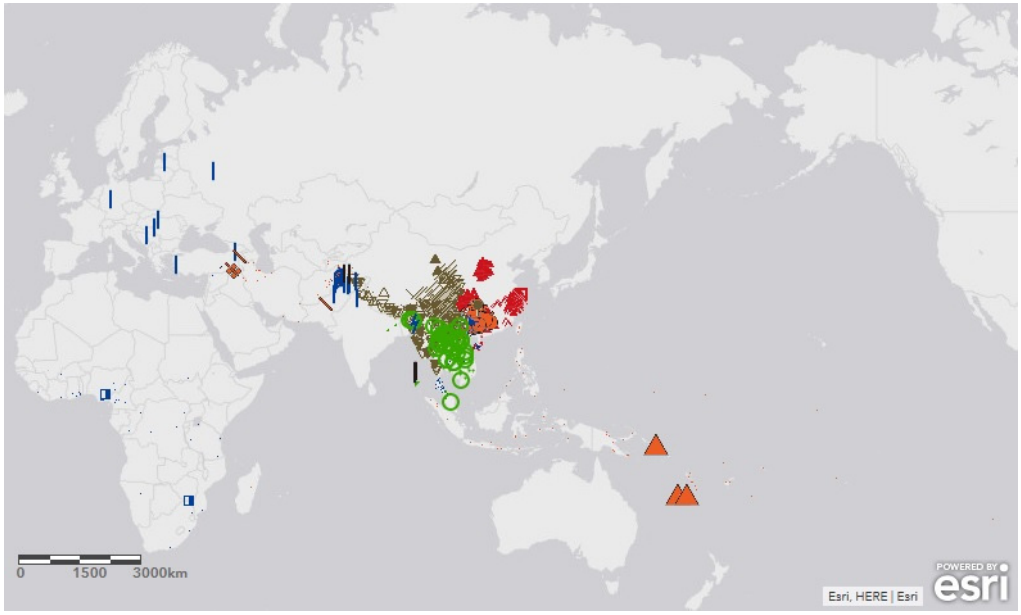


Fig. 2 Distribution of the tripartite contrast /th-t-d/.

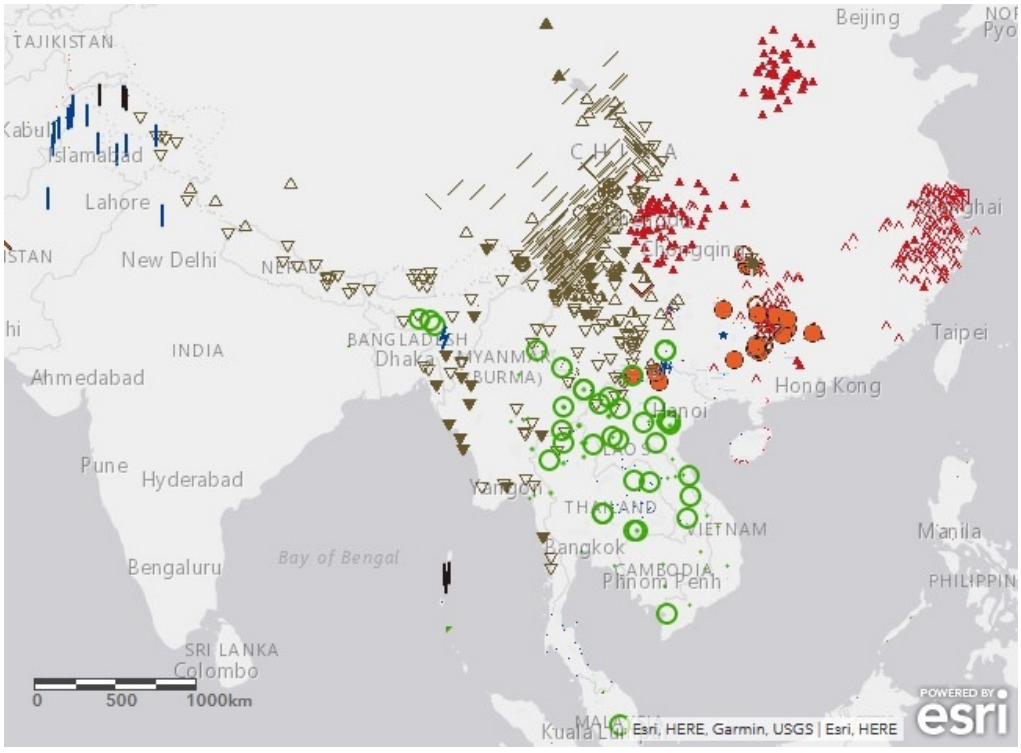


Fig. 3 Distribution of the tripartite contrast /th-t-d/ in Asia.

2.3. Non-nasal quadripartite contrast

A non-nasal quadripartite distinction, /th-t-d-dh/, is attested in several South Asian and Himalayan languages, including Tibeto-Burman, Indo-Aryan, Dravidian, and Iranian. As indicated in the descriptions of the Tibeto-Burman and Dravidian language groups, this series is attributed to Indo-Aryan language contact. The quadripartite distinction is observed in a limited number of major language groups, namely Indo-Aryan and Dravidian. A small number of varieties within the following language groups exhibit the quadripartite distinction: Sinitic, Hmong-Mien, Austroasiatic from eastern Asia and Kx'a (excluding the click series) from southern Africa. Figure 4 illustrates the language groups that exhibit a non-nasal quadripartite distinction.

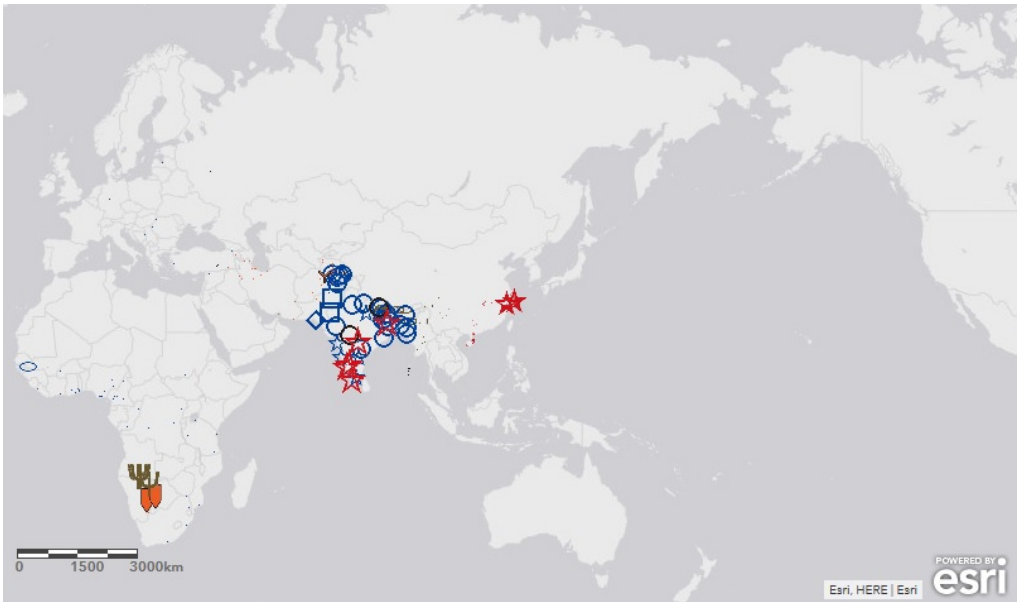


Fig. 4 Distribution of the quadripartite contrast /th-t-d-dh/.

2.4. Voicing contrast

A contrast between voiceless and voiced plosives (i.e., /t-d/) is attested in many language families, including Japonic, Sinitic, Hmong-Mien, Kra-Dai, Tibeto-Burman, Austroasiatic, Austronesian, Tungusic, Uralic, Mongolic, Turkic, Indo-Aryan, Burushaski, Dravidian, Iranian, Armenian, Nilo-Saharan, Niger-Congo, Tu, Kx'a, and Khoe-Kwadi. The data indicates that a majority of these languages exhibit a voicing contrast; hence, Figures 5 and 6 illustrate the language groups **without this contrast**.

It is noteworthy that Liang et al. (2023) address the causal chain relationships between climate, voice quality, and tone in languages in China. The data from Asia and Africa indicate that voicing distinction is a significant feature of languages from low latitudes.

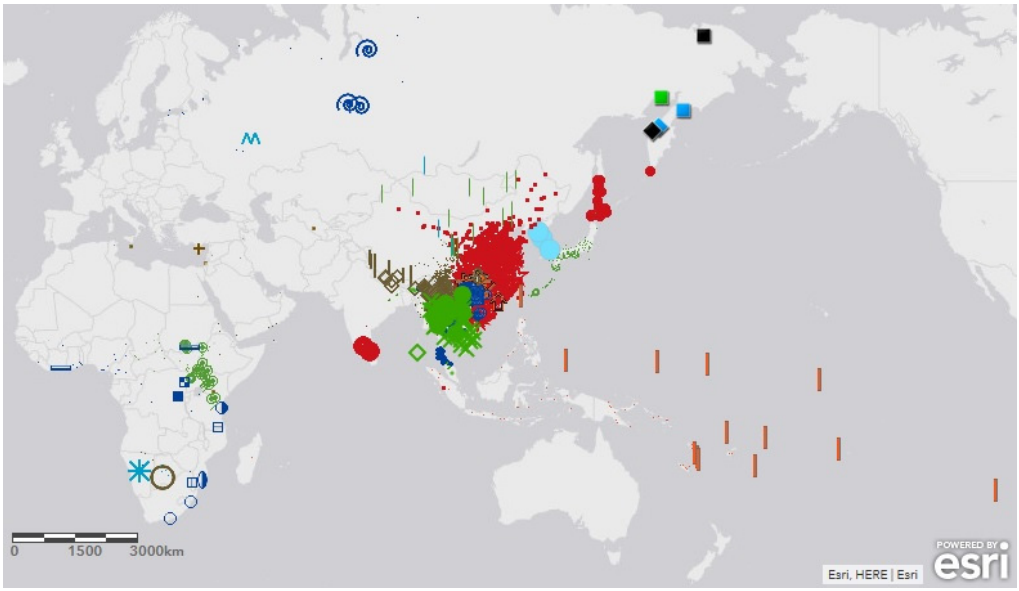


Fig. 5 Distribution of the language groups without a voicing contrast.

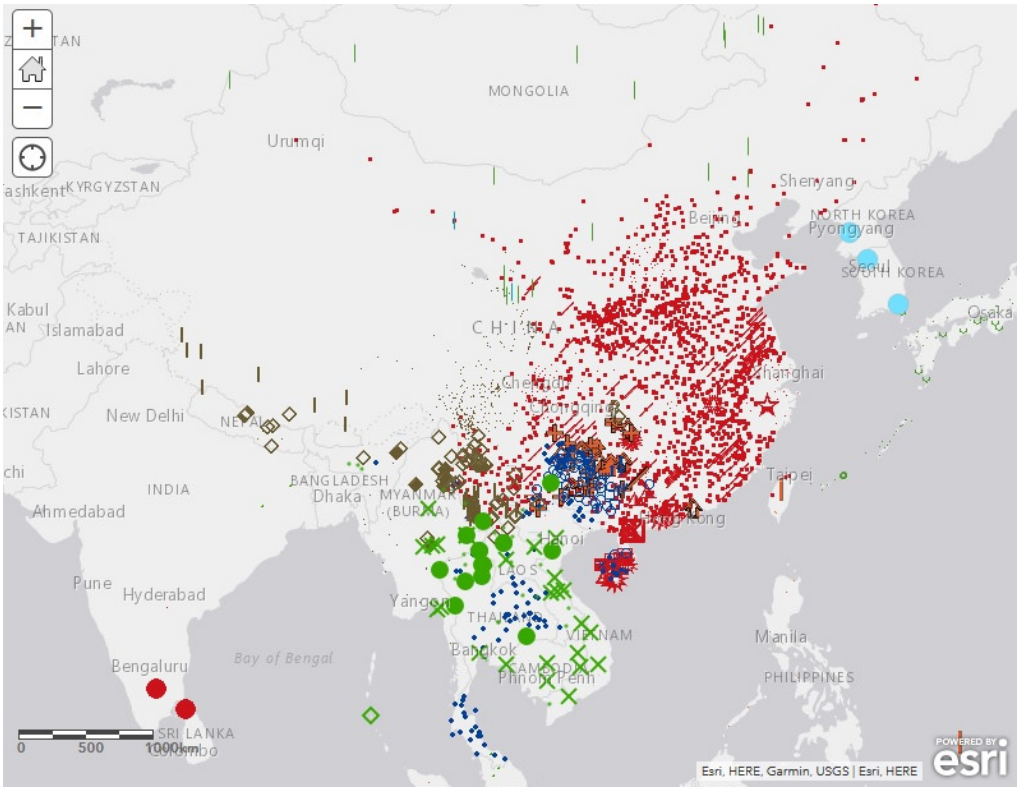


Fig. 6 Distribution of the language groups without a voicing contrast in Asia.

2.5. Aspiration contrast

A contrast between voiceless aspirated and voiceless nonaspirated plosives (i.e., /th-t/) is attested in many language families, including Korean, Sinitic, Hmong-Mien, Kra-Dai, Tibeto-Burman, Austroasiatic, Austronesian, Mongolic, Turkic, Indo-Aryan, Burushaski, Dravidian, Iranian, Armenian, Nilo-Saharan, Niger-Congo, Tu, Kx'a and Khoe-Kwadi. Furthermore, some Saami languages (Uralic) also exhibit aspirated stops (Suzuki 2021). Figure 7 and 8 present the language groups with this contrast, while Figure 9 presents the language groups **without this contrast**.

Furthermore, an aspirated voiced plosive /dh/ (/d$^{\text{ɦ}}$/) is documented in a number of language families, including Sinitic, Tibeto-Burman, Indo-Aryan, Dravidian, Iranian, Niger-Congo, Tuu, and Kx'a. No maps are provided for this feature.

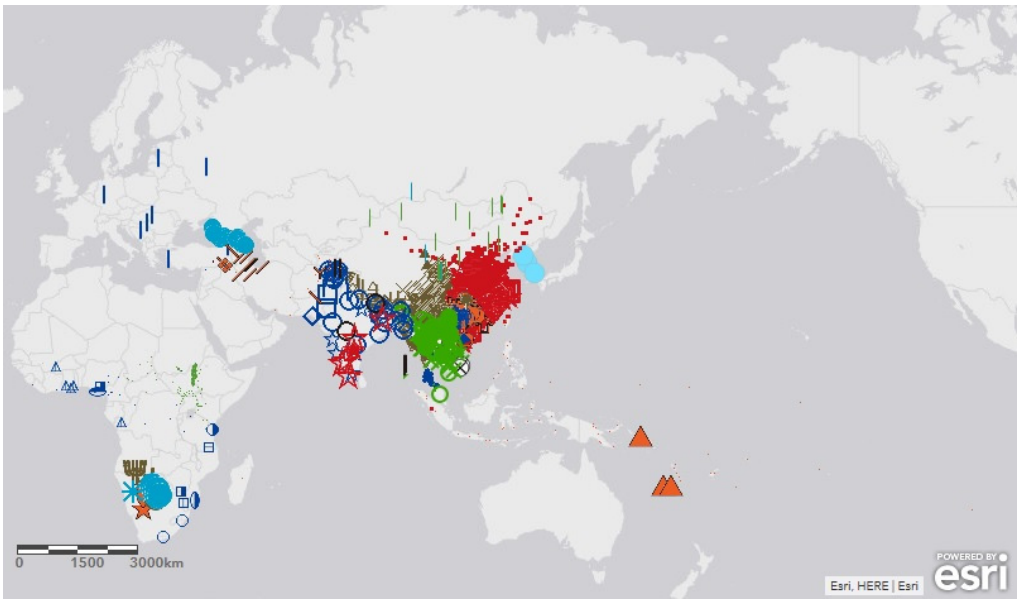


Fig. 7 Distribution of the language groups with an aspiration contrast.

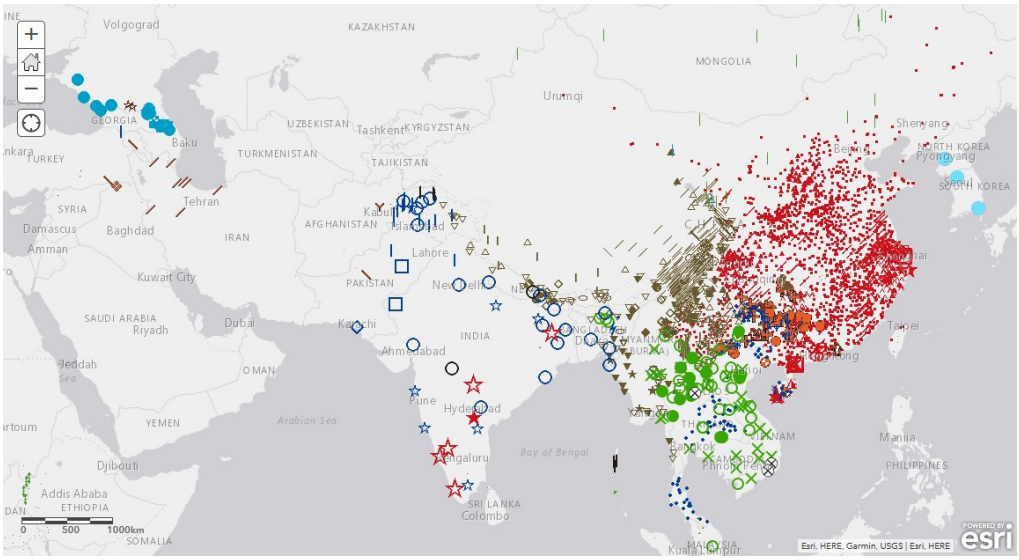


Fig. 8 Distribution of the language groups with an aspiration contrast in Asia.

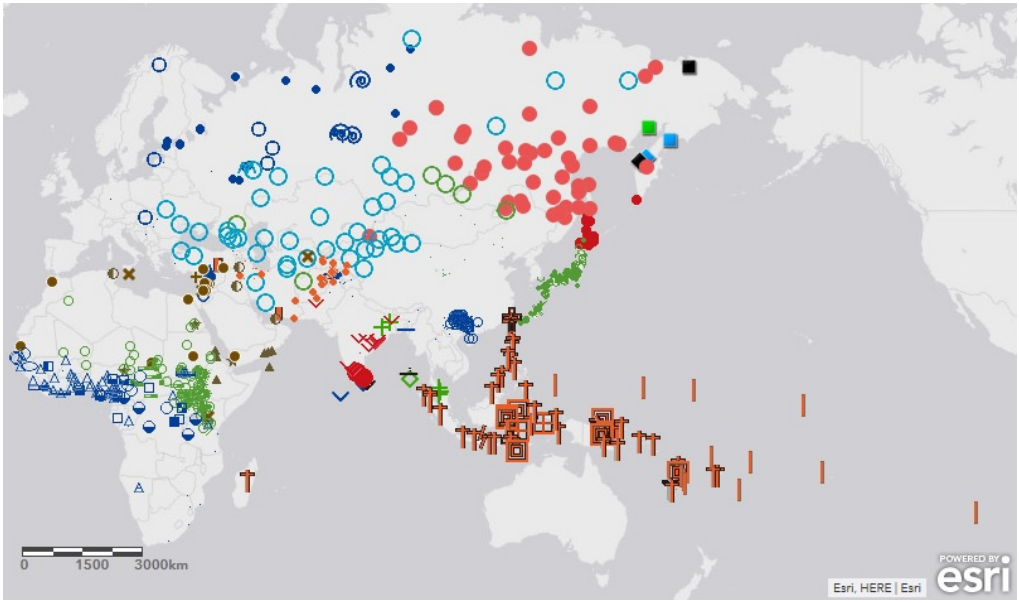


Fig. 9 Distribution of the language groups without an aspiration contrast.

3. Ejective

The ejective sound is documented in a number of geographical locations, including the Caucasus, Ethiopia, easternmost Siberia, the Kalahari Basin Area, and southernmost Africa. It should be noted that the Korean /t'/ is not included in the list

of ejective sounds. For further details, refer to Kim and Duanmu (2004) and Duan and Zhu (2018). In Ethiopia, both Nilo-Saharan and Semitic languages exhibit the presence of an ejective. In the Caucasus region, the ejective plosive is a pervasive feature of a number of Caucasian languages, including Kartvelian, Abkhazo-Adyghean, and Nakho-Daghestanian, as well as the Ossetic (Iranian) languages spoken in that region. Figure 10 illustrates the language groups that possess an ejective phoneme, while Figures 11 and 12 present comprehensive maps of the Caucasus and the Red Sea-Ethiopia region, respectively.

Fig. 10 Distribution of the language groups with an ejective sound.

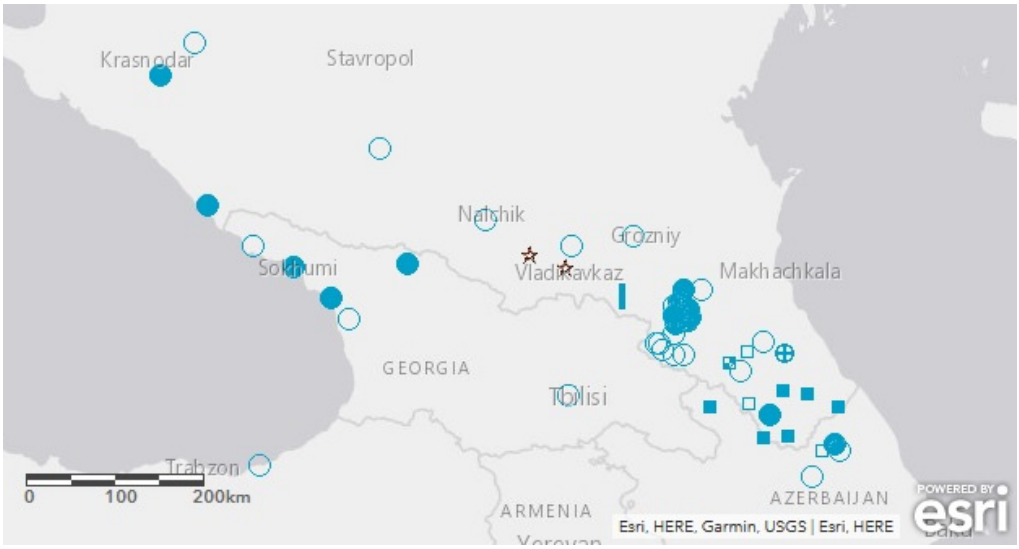


Fig. 11 Distribution of the language groups with an ejective sound in Caucasus.

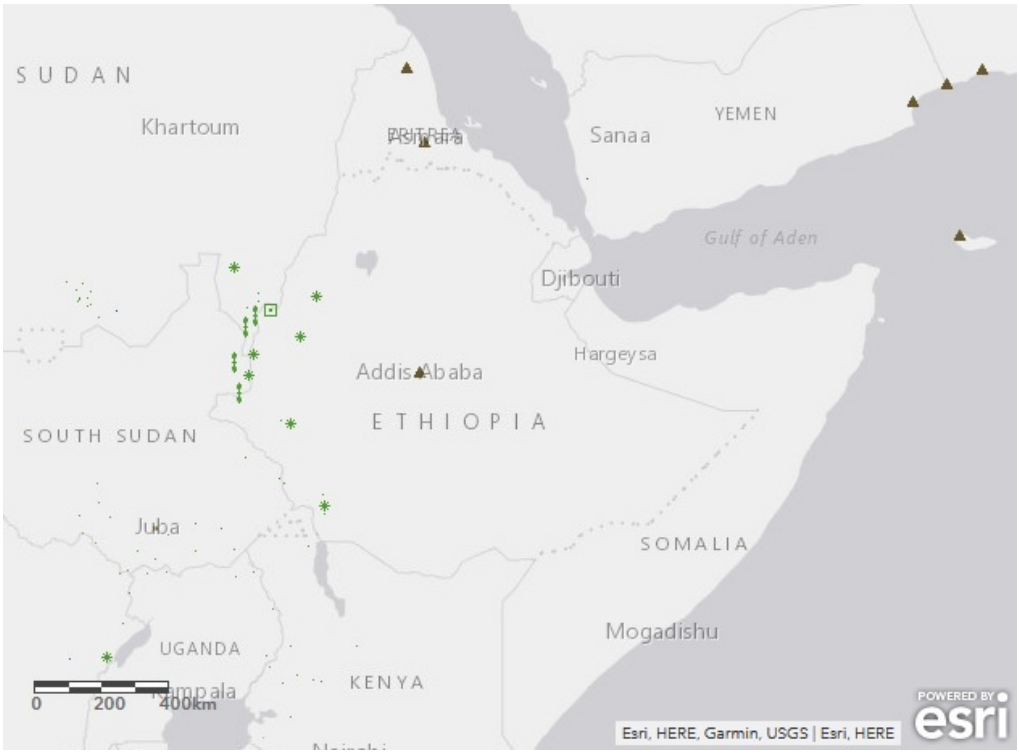


Fig. 12 Distribution of the language groups with an ejective sound in Red Sea-Ethiopia area.

4. Implosive

Implosives are typically voiced; however, a voiceless counterpart has also been identified. The voiced implosive /ɗ/ is attested in a number of language families, including Sinitic, Kra-Dai, Tibeto-Burman, Austroasiatic, Austronesian, Indo-Aryan, Semitic, Nilo-Saharan, and Niger-Congo. The voiceless implosive /ƭ/ is attested exclusively in Niger-Congo as a phonemic status. In some cases, a contact-induced acquisition provides an appropriate explanation. Figure 13 illustrates the language groups that possess an implosive phoneme, while Figures 14 and 15 present detailed maps of East and Southeast Asia and Central Africa, respectively.

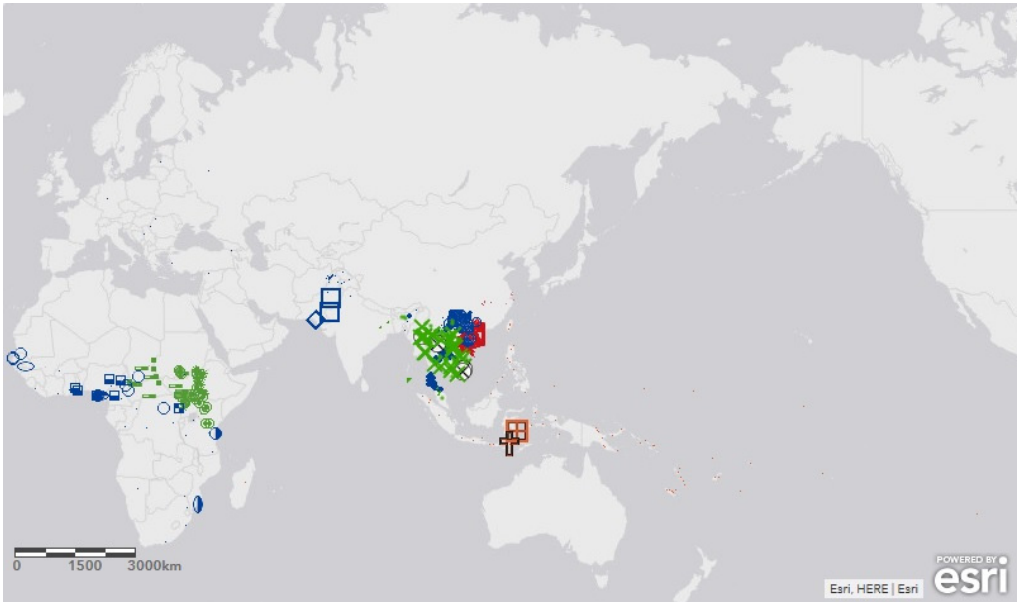


Fig. 13 Distribution of the language groups with an implosive sound.

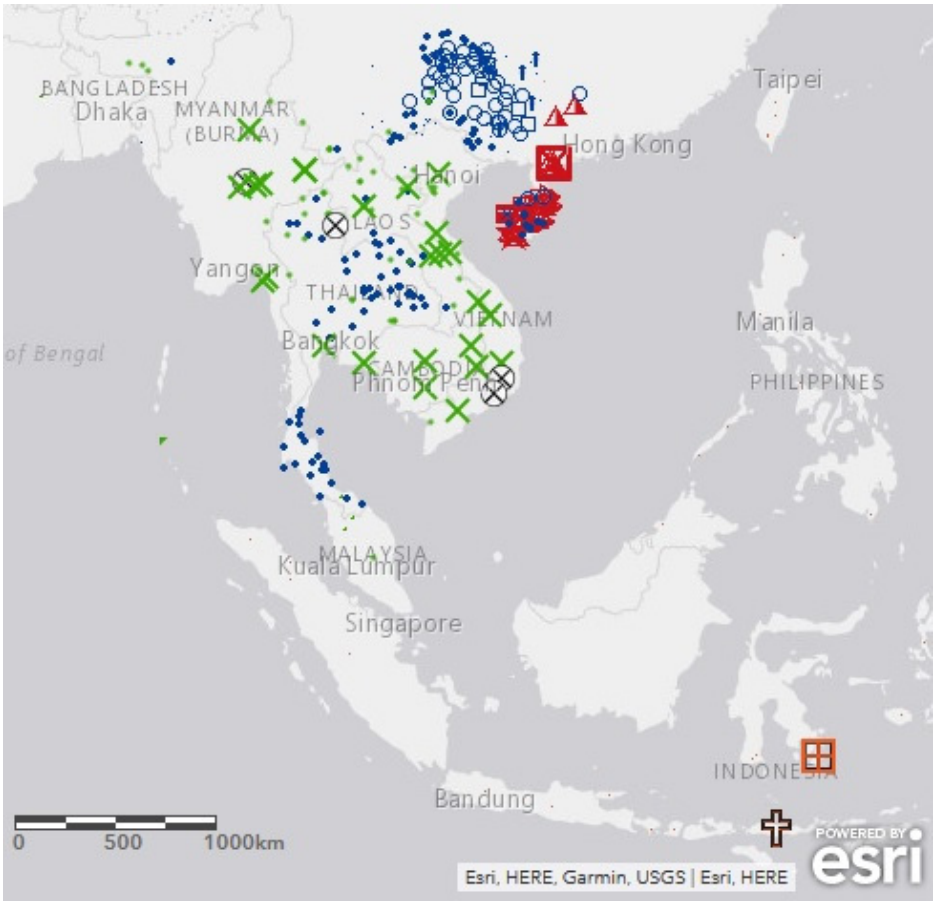


Fig. 14 Distribution of the language groups with an implosive sound in East and Southeast Asia.

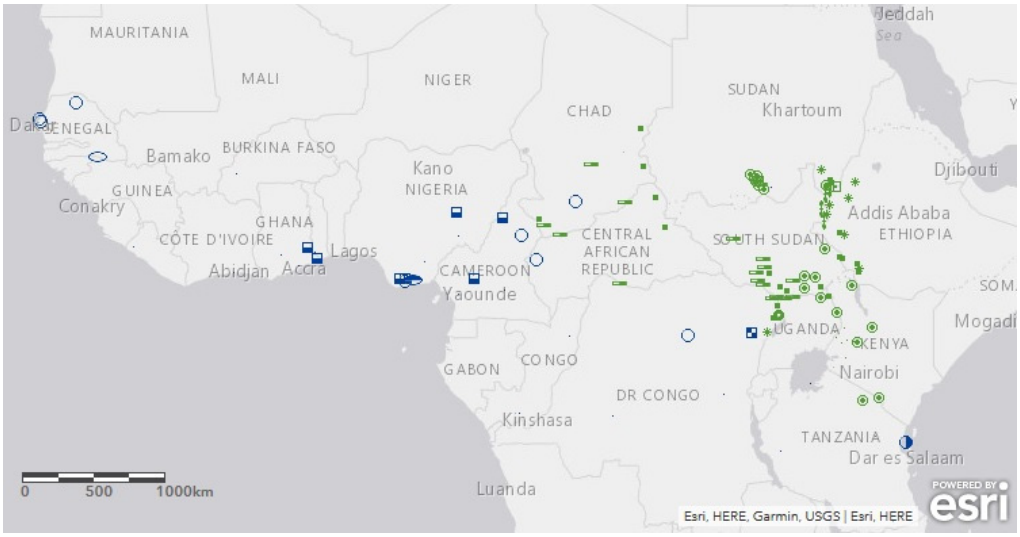


Fig. 15 Distribution of the language groups with an implosive sound in the central Africa.

5. Pharyngealisation

The occurrence of pharyngealised plosives has been documented in a number of languages, including Iranian, Semitic, Indo-Aryan spoken in Middle-East and Nilo-Saharan. The voiced type /dˤ/ is not attested in the majority of Nilo-Saharan languages. A systematic pharyngealised consonant feature, /tˤ-dˤ/, is predominantly attested in Semitic languages. Figure 16 illustrates the language groups that exhibit pharyngealised consonants, with the exception of Nilo-Saharan, for which the pharyngealised sounds were not distinguished in the original map of LAAA-2. Two languages from the Nilo-Saharan language family, Northern Songhay and Sudanese Berta, have been found to possess pharyngealised consonants. Figure 17 presents a detailed map of Middle East.

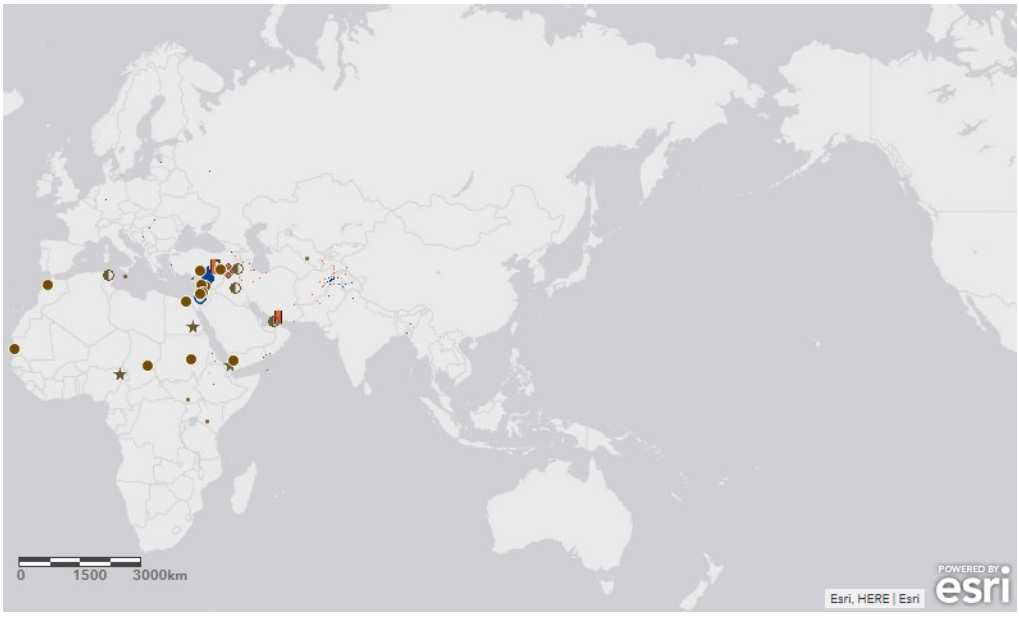


Fig. 16 Distribution of the language groups with pharyngealised consonants.

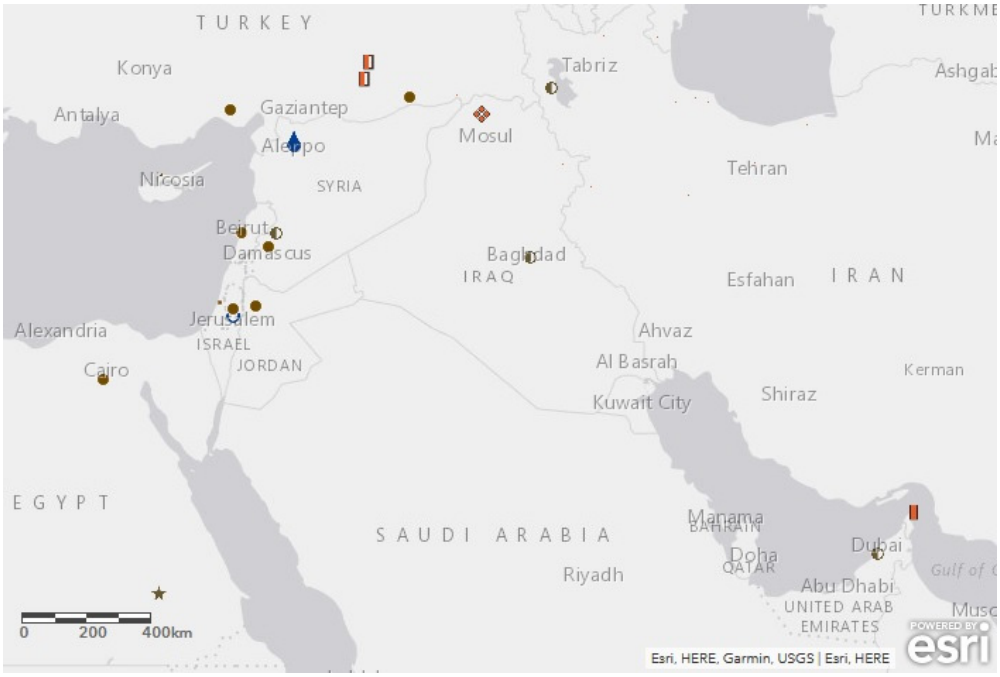


Fig. 17 Distribution of the language groups with pharyngealised consonants in Middle East.

Additionally, the pharyngealised feature is observed in Iranian languages to the west and Nilo-Saharan languages to the south, which are in turn connected to Semitic-speaking regions. The majority of Nilo-Saharan languages with pharyngealised features exhibit the phoneme /tˤ/, as do a number of Semitic and Iranian languages. As indicated in the descriptions of Semitic languages in LAAA-2, pharyngealisation is associated with ejective sounds (see Section 3), which are typically designated as 'emphatic consonants'. These two features are distributed across a continuous geographical area, from the Caucasus to South Africa.

6. Prenasalisation

The occurrence of prenasalised plosives has been documented in several language families, including Japonic, Sinitic, Kra-Dai, Hmong-Mien, Tibeto-Burman, Austroasian, Austronesian, Nilo-Saharan, and Niger-Congo. While voiced prenasalised sounds are pervasive in these languages, it should be noted that Tibeto-Burman, Austronesian, and Niger-Congo also have voiceless (and aspirated) counterparts. Prenasalisation is considered to be a more archaic form in certain language families, such as Japonic, Kra-Dai, and Tibeto-Burman. Conversely, it is seen to be a relatively recent development in other language groups, including Japonic and Sinitic. Languages

that have prenasalised characteristics are predominantly concentrated in East Asia, Austronesian regions, and Central Africa. These features are the result of internal phonological developments rather than language contact acquisition. Figure 18 illustrates the language groups that possess prenasalisation features, and Figure 19 presents a detailed map of East and Southeast Asia.

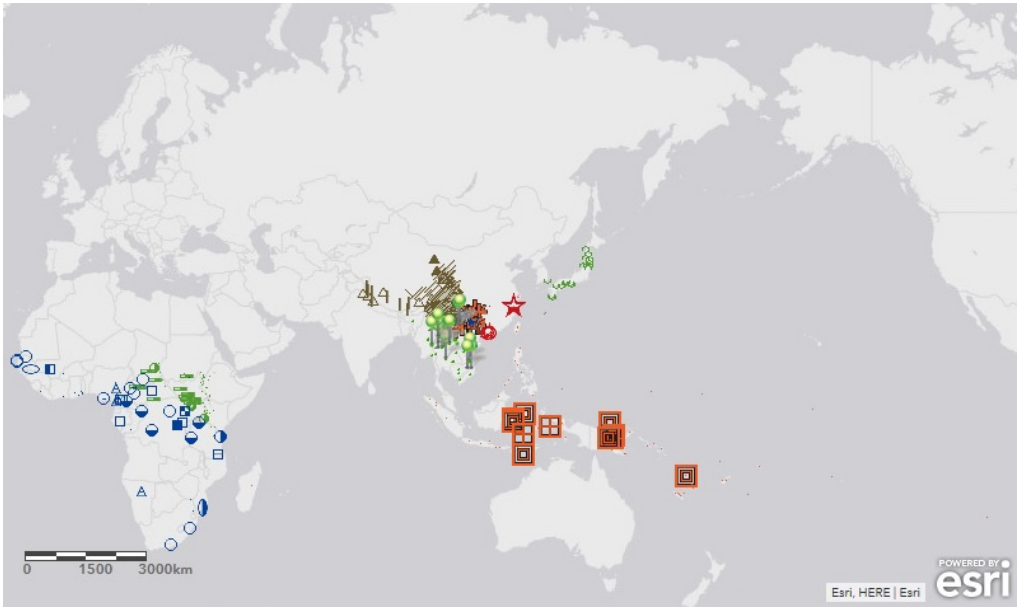


Fig. 18 Distribution of the language groups with prenasalised consonants.

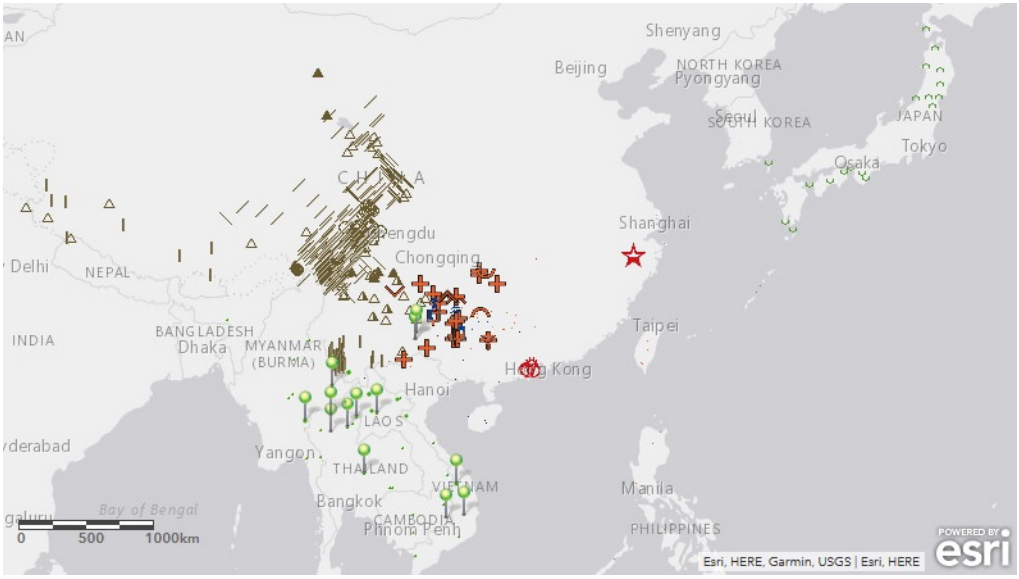


Fig. 19 Distribution of the language groups with prenasalised consonants in East and Southeast Asia.

7. Voiceless nasal

The voiceless nasal /n̥/ is attested in several language families, including Hmong-Mien, Tibeto-Burman, Austroasiatic, Austronesian, and Iranian. Several Saami languages (Uralic), which are not represented on the original LAAA-2 map, also exhibit this feature (Nielsen 1979, Kuruch 1985, Suzuki 2021). The sound is primarily documented in East and Southeast Asia. However, as this is a feature derived from individual sound developments in each language group, it is not considered a regional feature. The Saami languages' case also demonstrates that the given feature is not derived from any regional factors. Figure 20 illustrates the language groups that possess a D/A voiceless nasal, and Figure 21 presents a detailed map of East and Southeast Asia.

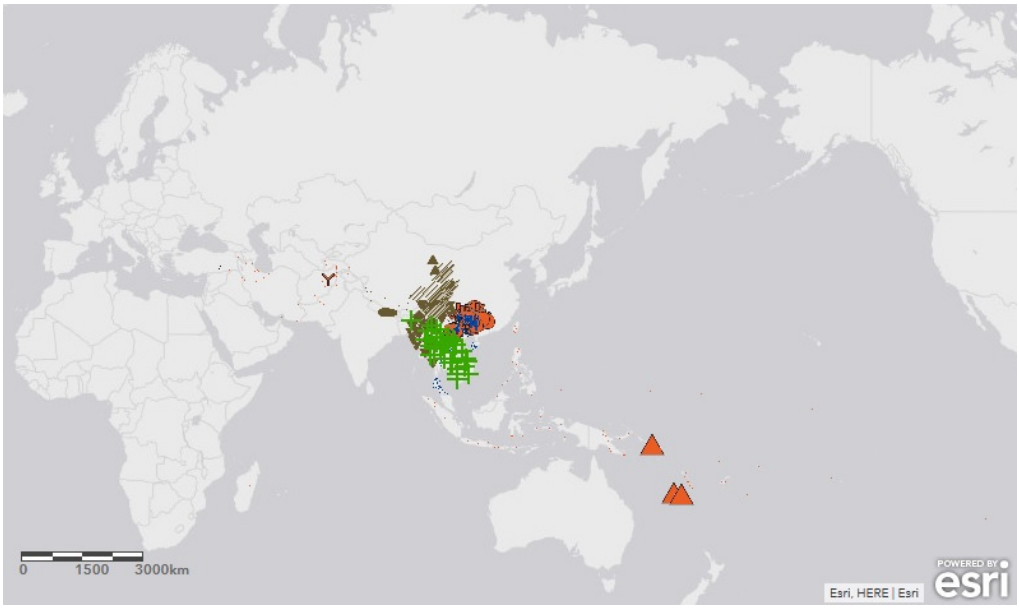


Fig. 20 Distribution of the language groups with voiceless nasals.

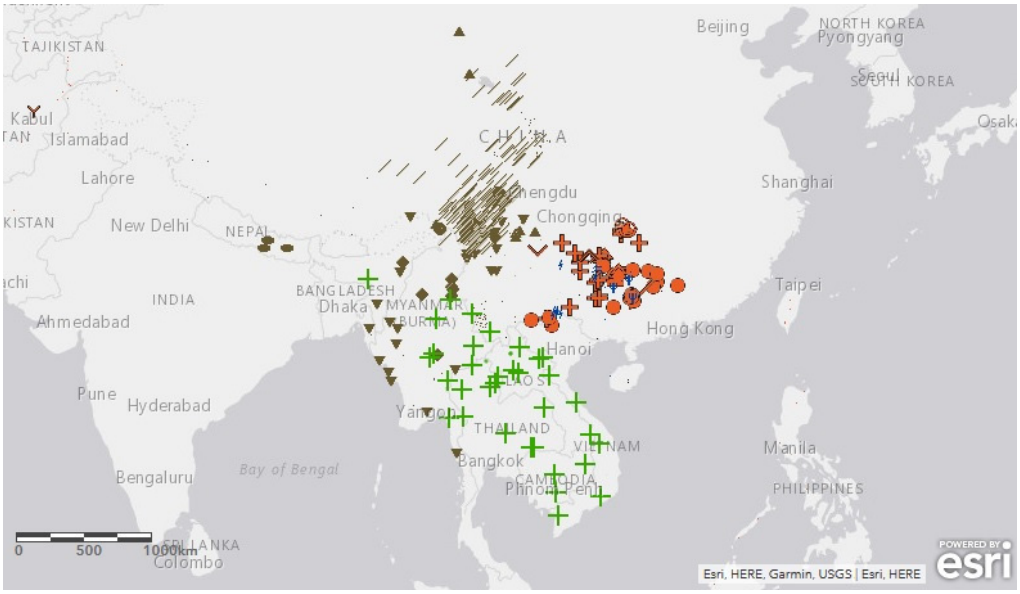


Fig. 21 Distribution of the language groups with voiceless nasals in East and Southeast Asia.

8. Glottalised stops

Glottalisation can be classified into two categories: preglottalised and postglottalised; however, as both are thought to have originated from disparate description conventions, they are depicted together on the map. The glottalisation feature is attested in several language families, including Ryukyuan of Japonic, Sinitic, Kra-Dai, Hmong-Mien, and Indo-Aryan. These sounds are primarily distributed across East and South Asia, with the preglottalised plosive being frequently associated with an implosive (cf. Section 4). Figure 22 illustrates the language groups that possess D/A glottalised stops, and Figure 23 presents a detailed map of East and Southeast Asia.

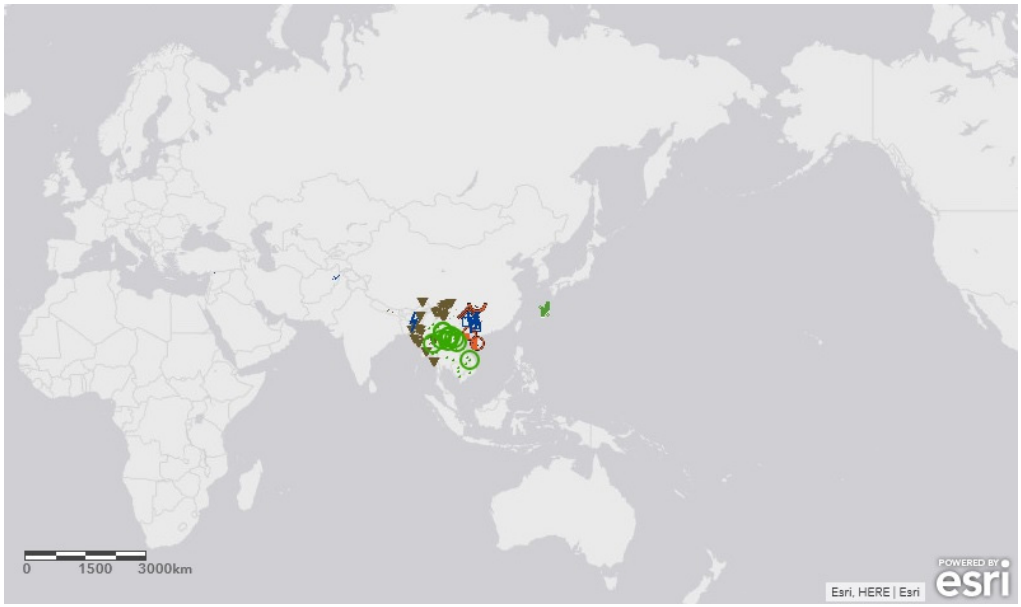


Fig. 22 Distribution of the language groups with glottalised stops.

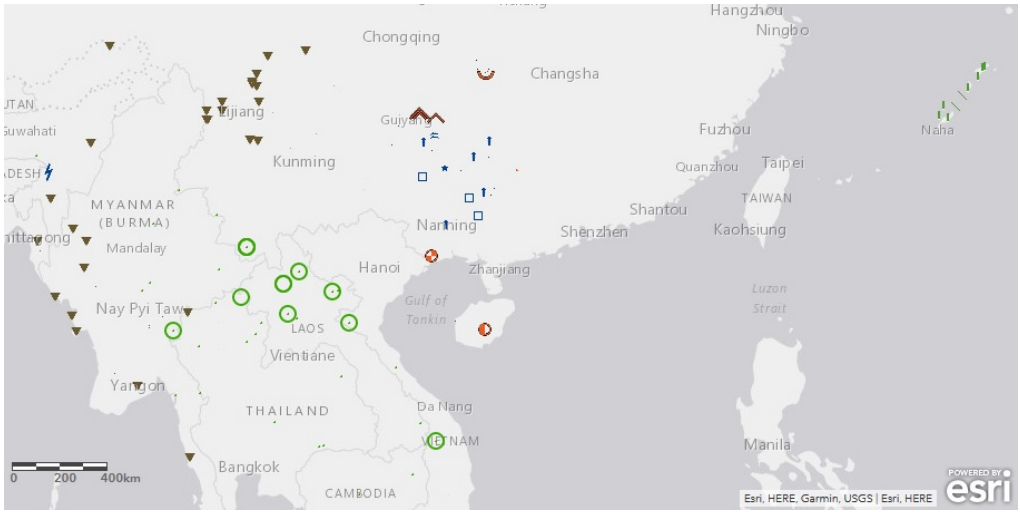


Fig. 23 Distribution of the language groups with glottalised stops in East and Southeast Asia.

9. Lack of D/A nasals

A paucity of D/A nasal sounds is observed in a restricted range of Sinitic, Kra-Dai, Tibeto-Burman, and Niger-Congo languages. In Sinitic and Kra-Dai, the principal factor contributing to the absence of the /n/ sound is the merger of /n/ into /l/. However, it appears that both Sinitic and Kra-Dai independently evolved the merger of /n/ into /l/

as there is no evidence of mutual language contact influences. This feature is also attested in Tibeto-Burman (Tujia), which can be attributed to Sinitic language contact (Southwestern Mandarin). It is important to note that this merger does not result in the absence of the sound [n] but rather its continued presence in the phonetic realisation (see Zhang and Levis 2021). Figure 24 illustrates the language groups that do not possess D/A nasal phonemes and Figure 25 presents a detailed map of East Asia.

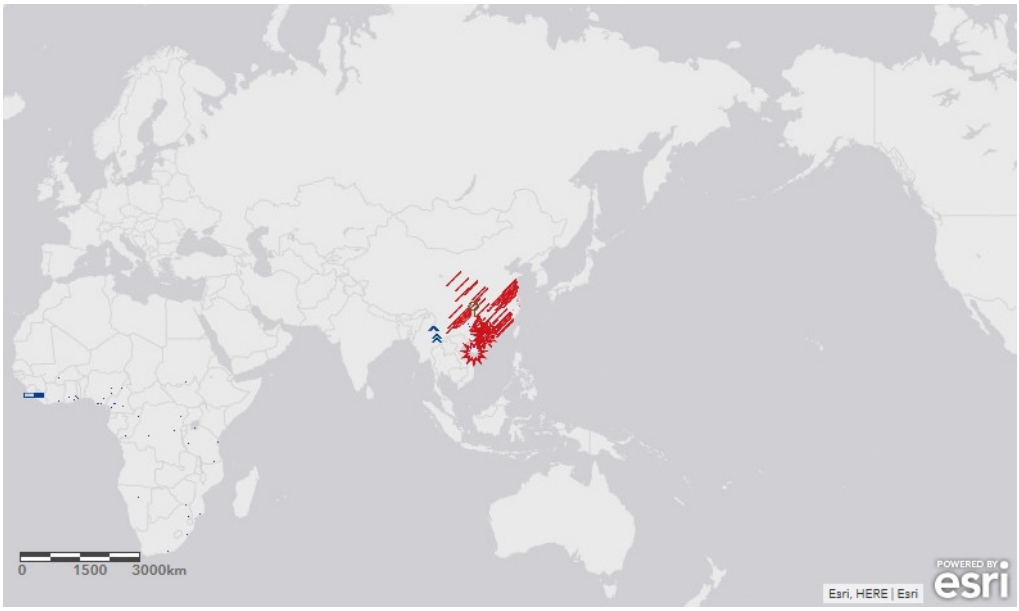


Fig. 24 Distribution of the language groups exhibiting the lack of D/A nasals.

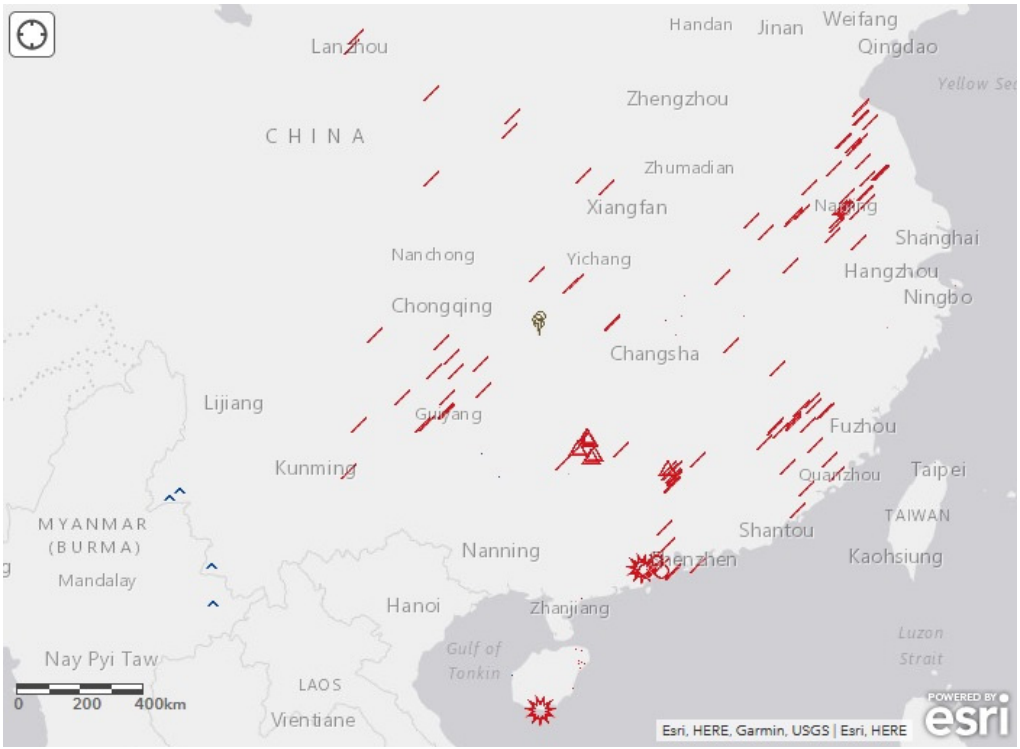


Fig. 25 Distribution of the language groups exhibiting the lack of D/A nasals in East Asia.

10. Conclusion

This article has addressed the 12 dental/denti-alveolar/alveolar stop series features from Chapter XIII Stop series of LAAA-2, with the data that meet the specified conditions from Asia and Africa being presented on individual maps. This article has demonstrated that the synthetic maps produced as part of the LAAA project are a valuable resource for the interpretation of the mutual relationships between language families in specific regions.

Notes for copyrights

The members of the 2020–2023 project involved in the investigation of the stop series are the following linguists: ONO Chikako (Chukotko-Kamchatkan), FUKAZAWA Mika (Ainu), NAKAZAWA Kohei and YOKOYAMA Akiko (Japonic), FUKUI Rei (Korean), YAGI Kenji (Sinitic), TAGUCHI Yoshihisa (Hmong-Mien), ENDO Mitsuaki (Kra-Dai), EBIHARA Shiho, IWASA Kazue, KURABE Keita,

SHIRAI Satoko, and SUZUKI Hiroyuki (Tibeto-Burman), SHIMIZU Masaaki and MINEGISHI Makoto (Austroasiatic), UTSUMI Atsuko (Austronesian), MATSUMOTO Ryo (Tungusic and Uralic), SAITÔ Yoshio (Mongolic and Turkic), YOSHIOKA Noboru (South Asia), KODAMA Nozomi (Dravidian), IWASAKI Takamasa (Iranian and Armenian), SUZUKI Hiroyuki (Caucasian), NAGATO Youichi (Semitic), NAKAO Shuichiro (Nilo-Saharan), SHINAGAWA Daisuke and KOMORI Junko (Niger-Congo), and KIMURA Kimihiko and NAKAGAWA Hirosi (Kalahari Basin Area).

References

Duan, Haifeng [段海凤] & Xiaonong Zhu [朱晓农] (2018) Chaoxianyu de ruan yin fuyin: Cong yuyin shuju dao yinfa fanchou 朝鲜语的软硬辅音：从语音数据到音法范畴 [Fortis and lenis in Korean: From phonetic data to phonological categories]. *Minzu Yuwen* 3: 13–25.

Fukazawa, Mika (2024) Numeral systems in Asia and Africa: A geolinguistic approach to linguistic patterns. *Studies in Geolinguistics* 4: 156–166.

Fukushima, Chitsuko (2024) Sibling terms in Asia and Africa: A geolinguistic approach to linguistic patterns. *Studies in Geolinguistics* 4: 145–155.

Kim, Mi-Ryoung and San Duanmu (2004) 'Tense' and 'lax' stops in Korean. *Journal of East Asian Linguistics* 13(1): 59–104. doi: https://doi.org/10.1023/B:JEAL.0000007344.43938.4e

Kuruch, Rimma D. [Куруч, Р. Д.] (1985) *Саамско-русский словарь*. Москва: Русский язык.

Liang, Yuzhu, Lining Wang, Søren Wichmann, Quansheng Xia, Shuai Wang, Jun Ding, Tianheng Wang, and Qibin Ran (2023) Languages in China link climate, voice quality, and tone in a causal chain. *Humanities and Social Sciences Communications* 10: 453. doi: https://doi.org/10.1057/s41599-023-01969-4

Nielsen, Konrad (1979 [1926-1929]) *Lærebok i lappisk (samisk)* I-III. Oslo: Universitetsforlaget.

Shirai, Satoko (2024) Alignment in Asia and Africa: A geolinguistic approach to linguistic patterns. *Studies in Geolinguistics* 4: 167–180.

Suzuki, Hiroyuki (2021) Stop series in the Saami languages: A geolinguistic approach. *Studies in Asian and African Geolinguistics I—Consonant system—*, 116–123. URI: https://publication.aa-ken.jp/saag1_stop_series_2021.pdf

Suzuki, Hiroyuki (2022) Stop series in Asian and African languages. In Hiroyuki Suzuki, Kohei Nakazawa, and Mitsuaki Endo (eds.) *Linguistic atlas of Asia and Africa II*, 149–154. Tokyo: Geolinguistic Society of Japan. doi: https://doi.org/10.5281/zenodo.7754469

Suzuki, Hiroyuki, Kohei Nakazawa, and Mitsuaki Endo (eds.) (2023) *Linguistic atlas of Asia and Africa II*. Tokyo: Geolinguistic Society of Japan. doi: https://doi.org/10.5281/zenodo.7754469

Tournadre, Nicolas and Hiroyuki Suzuki (2023) *The Tibetic languages: An introduction to the family of languages derived from Old Tibetan*. Villejuif: LACITO Publications. doi: https://doi.org/10.5281/zenodo.10026628

Zhang, Wei and John M. Levis (2021) The Southwestern Mandarin /n/-/l/ merger: Effects on production in Standard Mandarin and English. *Psychology of Language* 6. doi: https://doi.org/10.3389/fcomm.2021.639390

Zhu, Xiaonong [朱晓农] (2010) *Yuyinxue*《语音学》[Phonetics]. Beijing: Shangwu Yinshuguan.

Publication history

Date received: 30 August 2024
Date accepted: 18 September 2024

翻訳

中国少数民族言語における漢語動詞借用の類型

黄成龍
中國社會科學院民族學與人類學研究所
鈴木博之 訳

Mandarin loan verbs in Chinese minority languages

HUANG, Chenglong
Institute of Ethnology and Anthropology, Chinese Academy of Social Sciences
translated by Hiroyuki SUZUKI

Abstract: Lexical borrowing is not uncommon in human languages. The borrowing of verbs is believed to be more difficult than that of nouns. The strategies of how verbs can be borrowed vary in different languages. Since Chinese minority languages have long been in close contact with Mandarin Chinese in history, loan verbs from the latter are attested in minority languages to various extents. This paper analyzes the borrowing strategies of Mandarin loan verbs in minority languages and suggests that there are mainly four types of borrowing, i.e. direct borrowing, verbalizing affixation, coordination with light verbs and structural accommodation. The influence of language phylogeny, morphological paradigm and syntactic structure of verb borrowings as well as their typological features are also discussed.*

キーワード：類型論；少数民族言語；語彙借用；構造の調整
Keywords: typology; minority language; verb borrowing; structural accommodation

黄成龍(2024)「中国少数民族言語における漢語動詞借用の類型」（鈴木博之 訳）『地理言語学研究』4: 203–232. doi: https://doi.org/10.5281/zenodo.13948636

* [原著情報] 黄成龙《中国民族语言借用汉语动词的类型》，载《中国语文》2023年第4期，第387–401页。
[原著注] 本稿は中国国家社科基金重大項目「中国民族言語形態統語論の類型論的研究（中国民族語言形态句法类型学研究）」（18ZDA298）の成果の一部である。本稿の初稿は「言語接触と言語比較国際会議（語言接触与語言比较国际论坛）」（2015年12月6日、上海師範大学）で発表した。会場にて王双成、徐世璇、趙明鳴、邢欣の各氏および参加者からの貴重な意見をいただいた。論文執筆および改訂に際しては、さらに多くの少数民族言語を母語とする研究者や少数民族言語を研究する仲間から関連する言語資料と援助を提供していただいた：喀喇沁モンゴル語（曹道巴特爾）、ウイグル語（阿米娜木）、朝鮮語（千玉花）、エヴェンキ語（烏日格喜樂圖）、キルギス語（宋佳）、東旺チベット語（次林央珍）、アムドチベット語（周毛草、久西杰）、スタウ語（根嘎翁姆）、ムニャ語（達瓦卓瑪）、ゴチャン語（楊曉燕）、ナス彝語（普忠良）、白語（李煊）、リス語（李翔）、ハニ語アカ方言（楊雲）、ナシ語魯甸方言（和菊）、トゥルン語（楊将領）、チュワン語（楊心怡、韋景雲）、水語（韋学純）、石洞侗語（龍仙梅）、湘西ミャオ語（麻秀芝）、ワ語（陳国慶）、プラン語（陶成美）、オーストロネシア諸語（潘家榮）、瓦罕タジク語（侯典峰）。ここに記して謝意を表する。なお、本稿における問題点は筆者が責任を負うものである。
[訳者注] 本稿は原著者である黄成龍氏より翻訳の依頼を受けたものである。英語の題目・要旨・キーワードは原著の英文によった。言語名は国際的に通用する名称に改めた。参考文献の中で漢語のものは、原文通り簡体字で掲げた。また、語例で漢語音の知識を前提とする箇所については、原著者との協議を経て日本語読者のためにピンインを付加した。

1. はじめに

人類が発展してきた歴史とは、人と自然、人と社会、人と人の接触と交流の歴史である。人と社会、人と人の交流の過程において、人類は言語行為または非言語行為を通じて互いに意思疎通と交流してきた。異なる人々が交流する中で、A 語と B 語が接触すると相互に影響し、A 語が B 語の語彙または構造を借用するのは避けられない。このため、ある言語の中の借用語は人類の交流史を研究する際の生きた化石である。借用語は人類の言語の発展と変化によるものであるけれども、人々が借用語の問題に関心を持ったのは百数十年程度の歴史しかない。19 世紀後半にウィットニー（Whitney 1881）が借用語研究を開拓して以来、研究者らの人類の言語または具体的な言語における借用語の研究は絶え間なく続いてきた（Burrow 1946; Haugen 1950; Weinreich 1953; Scotton and Okeju 1973; Gentner 1982; Heath 1984; Poplack and Sankoff 1984; Thomason and Kaufman 1988; Campbell 1993; Thomason 1997, 2001; Matras 1998; Curnow 2001; Aikhavald and Dixon 2001; Ross 2001; Field 2002; Johanson 2002, 2005; Tent and Geraghty 2004; Heine and Kuteva 2005; Sanchez 2005）。タドモル（Tadmor 2009）は世界の 41 種の言語における借用語の研究を通して、語彙の借用は人類の言語における共通性を見せると考え、すなわち語彙の借用は非常に普遍的であるが、しかしながら語彙の借用比率は諸言語において非常に大きな差異を呈している。研究者ら（Haugen 1950; Matras 2008: 48; Tadmor 2009: 61）は、すでに名詞の借用は容易であり、動詞の借用は困難であることに注意を向けている。動詞の借用についての共通性と相違性、およびその原因をさらに詳細に理解するために、多くの研究者が類型論的視点から動詞の借用の類型とその特徴を分析している（Mifsud 1995; Morimoto 2000; Wohlgemuth 2004, 2009; Wichmann and Wohlgemuth 2008; Wohlgemuth 2009）。中国で話される百種以上の少数民族言語において、漢語の動詞はいかに借用されているだろうか、どのような類型的特徴があるだろうか、動詞の借用に共通性と相違性はあるだろうか、といった類の問題は、これまでの言語接触や借用語研究の中で十分に注目されてきたとは言えない。本稿では、『中国の言語』（《中国的語言》）に含まれる約 130 種の言語の系統分類を基礎として、少数民族言語が漢語の動詞を借用する特徴を考察し、漢語の動詞を借用する類型について分類を行い、同時に中国の少数民族言語が漢語の動詞を借用する類型と特徴について内部構造と言語外の解釈を行う。

2. 少数民族言語のが漢語の動詞を借用する方法

ラウファー（Laufer 1916）による「チベット語における借用語」は、中国の少数民族言語の借用語について最も早く研究した論文である。次いで、李方桂（Li 1945）が「ダイ語における古漢語借用語」を発表した。新中国の成立後、中国の少数民族言語の借用語に関する研究は、常にハウゲン（Haugen1950）の枠組みと術語に従ってきた。たとえば、全借用（loanwords）すなわち形式と意味の完全なコピー、半借用（loanblends）と意訳（loanshifts）すなわち意味に限定したコピーで、これには翻訳借用（loan translations）と意味的借用（semantic borrowings）がある。少数民族言語における漢語借用語に関する研究は、ほとんどが借用語の形式と音変化の規則に注目し（李得春 1986；石林 1994；郑贻青 1995；玉珍 1996；朱文旭 1997；闫新红、欧阳伟 2000；沙加尔、徐世璇 2002；曾晓渝 2003；蒋宏军 2010；张建民 2011）、語類の借用（季永海 1985；敏春芳 2012；严木初 2013；石常艳 2015）や文法構造の借用（徐世璇 2014）に注目したものは少ない。現在手元にある言語資料から見ると、中国の少数民族言語が漢語の動詞を借用する方法は、以下の数種類の借用類型に分けられる。

2.1. 直接借用

漢語動詞がいかなる補助成分を付加することなく直接少数民族言語に借用されることを、直接借用またはコピー（direct adoption or copy）と呼ぶ。この種の借用方法は中国国内のクラ・ダイ諸語、ミャオ・ヤオ諸語、オーストロアジア諸語、オーストロネシア諸語、およびチベット・ビルマ諸語ロロ語支、ビルマ語支などを含む少数民族言語の中で最もよく見られる。たとえば、ツァイワ語（徐悉艰、徐桂珍 1984: 170-171）では次のようである。

(1)

jianshe 建设	kjen^{55}ʃe^{31}	建設する
fazhan 发展	fa^{31}tʃan^{51}	発展する
yonghu 拥护	juŋ^{51}xu^{31}	擁護する
biaoyang 表扬	pjau^{51}jaŋ^{31}	表彰する
jiefang 解放	kai^{51}faŋ^{31}	解放する
fandui 反对	fan^{51}tui^{31}	反対する
douzheng 斗争	tu^{31}ts$\underline{\text{i}}$n^{31}	闘争する
qinlüe 侵略	k$^{\text{h}}$jin^{55}ŋjo^{31}	侵略する

いくつかの言語では、漢語動詞を直接借入するのみならず、漢語の「動目句」を借入するときもまた直接借用を採用する。たとえば、雲南省紅河州ロロ語南部方言群ニェス方言（王琳 2021：107）では次のようである。

(2)

jian 煎	tɕa^{31}	煎る
kai dao 开刀	k^{h}ɛ^{33}tau^{33}	手術する
jian baba 煎粑粑	tɕĩ^{55} pa^{55}pa^{55}	蒸しパンの表面を焼く
kang yumi 扛玉米	k^{h}ã^{31}ʑi^{55}mɛ^{31}	とうもろこしを担ぎ運ぶ

以上の例から分かるように、ツァイワ語やロロ語ニェス方言が漢語動詞を借用する場合、直接西南官話から漢語動詞の音形と語義を借入し、その他の付加成分を追加することはない[1]。

さらにクラ・ダイ諸語、ミャオ・ヤオ諸語、オーストロアジア諸語が漢語動詞を借用した語例を以下に掲げる。これらの動目構造はすべて直接借入したものである。

(3)

	chong dian 充电 充電する	shang wang 上网 インターネットに接続する	fa duanxin 发短信 ショートメールを送る
北部チュワン語	ɕuŋ^{24} te:n^{35}	ɕaŋ^{24} va:ŋ^{55}	fa:t^{55}tu:n^{55} θin^{24}
南部チュワン語	ɬuŋ^{33} ti:n^{42}	han^{21} mu:ŋ^{21}	fa:t^{24} ɬən^{33} θik^{33} < fa xinxi 发信息
石洞トン語	suŋ^{44} tjen^{35}	ti^{22} ɕa^{11} < da che 打车	wet^{13} ton^{31}ɕin^{35}
スイ語	tsuŋ^{33}tjen^{11}	saŋ^{11}vaŋ^{55}	fa^{31}tuan^{55}sin^{11}
湘西ミャオ語	ts^{h}oŋ^{44}tjæ^{35}	ʂɑŋ^{35}wɑŋ^{44}	fɒ^{42}tæ^{35}ɕin^{35}
ワ語	tɕ^{h}ɔ̲ŋ ti̲an	saŋ vuaŋ	fa̲$^{?}$ tuaŋ ɕi̲ŋ
プーラン語	tʂ^{h}oŋ^{33}tiɛn^{24}	ʂaŋ^{24}waŋ^{42}	fa^{21}tuan^{42}ɕin^{24}

台湾やフィリピンのオーストロネシア諸語は名詞の借用が多く、動詞の借用は少ない。もし動詞を借入するならば、漢語の音形を直接借用するが、本来の民族語の形態論的特徴に合わせて、オーストロネシア諸語の焦点（focus/voice）標示を付加する[2]。

以上の例から分かるように、借入（目標）言語の形態統語論の類型が漢語の形態統語論の類型と比較的一致を見る場合、クラ・ダイ諸語、ミャオ・ヤオ諸語、オーストロアジア諸語のような言語は、漢語と類似して、みな分析的な言語であり、すべて SOV 型の言語である場合、漢語動詞を借用する方法は単一

[1] 少数民族言語が漢語の語彙を借用するとき、本来の民族語に存在しない音を借入することができる言語もあれば、本来の民族語の音韻規則に基づいて借用語の音形を調整（accomodation）する言語もある。本稿では紙面が限られるため、音韻の調整については扱わず、形態統語論における調整のみ議論する。

[2] オーストロネシア諸語の借入に関する特徴については、北京師範大学珠海校区の潘家榮教授の教えを請うた。

であり、ただ直接借用のいう方法のみを用いる。チベット・ビルマ諸語ロロ語支、ビルマ語支、ジンポー語支のジンポー語は分析的性格もわりと強く、もし単純な漢語動詞を借用するだけであれば、一般的に直接借用を採用する。ただし、もし漢語の「動目句」を借入する場合は、目標言語の語順に基づいて、その構造を賓動構造に調整することになる。

2.2. 動詞化接辞の付加

漢語動詞がいくつかの少数民族言語に借入されるには、本来の少数民族言語における、たとえば接辞や助詞の添加などの動詞化（verbalizer）のような、いくつかの形態統語論的な方法を経て、借用語に動詞の用法を持たせる必要がある。このような方法を間接借用（indirect adoption）または調整（accommodation）という。この間接的な方法は主に大多数のアルタイ系諸言語[訳者注：チュルク諸語、モンゴル諸語、ツングース諸語を含む]や少数のチベット・ビルマ諸語に見られる。

動詞化接辞を付加することはアルタイ系諸言語ではわりと広く分布するが、名詞や形容詞などの語が動詞として用いられるときに、動詞化接辞を付加する必要がある。多くのアルタイ系諸言語が漢語動詞を借用するとき、動詞化の方法を経て、もとの漢語動詞の後ろに本来の少数民族言語の動詞化接辞を付すことになる。

2.2.1. モンゴル語

モンゴル語が借入した漢語借用語は名詞が多いが、ほかに少数の動詞や他の語類もある。漢語動詞を借入するとき、漢語動詞の後ろにモゴル語の付加成分 -l または-n を付す（道布 1983：139-140）[3]。

(4)

duǎ**n**	（< duo 剁）	（詰め物を）ぶつ切りにする
dʒignǎ**n**	（< zheng 蒸）	蒸す
dʒɔɔlǎ**n**	（<zhao 照）	（写真を）撮る
tʃiilǎ**n**	（< chiliang 尺量）	ものさしを用いる
ʃagnǎ**n**	（< jiangli 奖励）	奨励する

2.2.2. モングォル（土族）語

モングォル語の借用語は主に漢語とチベット語からのもので、それらの動詞がモングォル語に借入されたとき、必ず借用語の後ろに助動詞から転化した接

[3] 以下の語例はすべて先行研究からの引用である。接辞の前に“-”を付加したものは動詞の後ろに接辞を付加したことを示す。接辞の後ろに“-”を付加したものは、接辞の後ろにさらに他の接辞（テンス・アスペクト接尾辞、副動詞接尾辞、形容詞化または名詞化接尾辞など）を追加することができる。接辞の前後ともに“-”を付加したものは、その接辞の前後にその他の接辞が現れうることを示す。

辞-gA（個別の借用語では-lA[4]）を付加する。たとえば以下のようである（照那斯图 1981：21-22）[5]。

(5)

quan 劝	tɕyen-**gə-**	勧める
peng 碰	pəŋ-**gə-**	ぶつかる
taolun 讨论	taolyn-**gə-**	討論する
tiaobo 挑拨	tɕiobə-**gə-**	挑発する

2.2.3. マネガチャ（保安）語

マネガチャ語が漢語から動詞を借入するとき、直接動詞とすることはできず、借用語の後ろにマネガチャ語の付加成分-lA（-la, -le, -lo）または-kA（-kə）や-gA（-gə）を加える必要がある。-lA, -kA, -gA はいずれも借用した漢語語彙が動詞であることを示すが、互いに交替はできない。これらの中で-lA を用いるものが最も多く、-kA, -gA を用いるのは少ない。たとえば以下のようである（布和、刘照雄 1982：83-84）。

(6)

bang 绑	vaŋ**la-**	ひもで結ぶ
jiu 救	dʑiuu**la-**	救う
piping 批评	pipin**la-**	叱る
pa 爬	paa**la-**	登る
xian 献	ɕin**le-**	差し上げる
guan 管	go**lo-**	管理する
ban 搬	ban**kə-**	運んで移動する
gua 刮	guaa**kə-**	剥く
bang 帮	baŋ**kə-**	助ける
zhan 占	dzaŋ**kə-**	占める
ta 踏	taa**gə-**	踏む
ca 擦	saa**gə-**	こする
tan 弹	taŋ**gə-**	弾く
tui 推	tuii**gə-**	押す

[4] アルタイ系諸言語では通常母音調和の現象があるため、本稿では大文字のAで母音を代表し、具体的な母音表記は実際の発音を示す。

[5] モンゴォル語がチベット語動詞を借用するときもまた-gAを付加する。たとえばkabda-**gə-**「会話する」、ndziər-**gə-**「くっつく」（照那斯图1981：22）。

2.2.4. シラ・ユグル語

シラ・ユグル語が漢語動詞を借入するとき、漢語動詞の後ろに一律同言語の名詞から動詞を派生させる付加成分-lA（-la, -le, -lo）を加えて動詞語幹を成立させなければならない。この種の動詞語幹と固有の動詞語幹は、形態論的にも統語論的にも機能は全く同じである。たとえば以下のようである（照那斯图 1981：82-83）。

(7)

yonghu 拥护	juŋχu**la-**	擁護する
cheng 称	tʃəŋ**la-**	量る
piping 批评	pipiŋ**la-**	叱る
jiefang 解放	dʒeɸaŋ**la-**	解放する
tugai 土改	tugai**la-**	土地改革する
bangzhu 帮助	baŋgu**la-**	助ける
tuanjie 团结	tuandʒe**le-**	団結する
die 叠	dee**le-**	重ねる
lian 炼	(gaŋ)len**le-**	（鋼を）製錬する
jian 捡	(sai)dʒem**le-**	（おかずを）つまむ

（同時に構造を調整しているもの）

zhaung 装	dʒoŋ**lo-**	詰める
guan 管	gon**lo-**	管理する
bo 簸	bo**lo-**	揺れる
zuan 钻	dzoŋ**lo-**	穴を開ける

2.2.5. カザフ語

カザフ語の漢語借用語も名詞が多く、さらに少量の量詞があるが、動詞の数は極めて少ない。漢語動詞を借用するとき、漢語動詞の後ろに動詞を構成する付加成分-dɑ、-le を付す。たとえば(8)のようである。もし名詞がカザフ語に借入されると、借入ののち動詞を構成する付加成分-lɑn-、den-、-lɑ-を付す。たとえば(9)のようである（耿世民、李增祥 1985：178）。

(8)

pipan 批判	pipɑn**dɑ-**	批判する
pizhun 批准	pij**le-**	批准する

(9)

gongshe 公社	goŋʃə	公社	goŋʃə-**lɑn-**	公社化する
geming 革命	gəmiŋ	革命	gəmiŋ-**den-**	革命化する
laoyi 劳役	lɑw	劳役	lɑw-**lɑ-**	劳役用の馬に乗る

2.2.6. サリグ・ユグル語

漢語動詞がサリグ・ユグル語に借入されたとき、一律同言語に固有の付加成分-lA, -nA を付し、各語に動詞の効能を持たせる必要がある。たとえば以下のようである（陈宗振、雷选春 1985：55-56）。

(10)

bian 编	dɑl**ɑ-**	（かご、座席を）編む
dui 对	due**lɑ-**	対質する
guai 拐	gue**lɑ-**	くじく
kaihuai 开会	kexue**lɑ-**	会議する
jianshe 建设	dʒenʂə**lɑ-**	建設する
daoxie 道谢	ʃe**le-**	感謝する
qian 鸽	tʃe**le-**	鳥がくちばしでつつく
pie 撇	pi**le-**	（あくを）すくい取る
ji 掖	ji**le-**	持ち歩く
xuanju 选举	ʃyendʒy**le-**	選挙する
jing 敬	dʒin**ne-**	（酒で）敬意を表す
biaoyang 表扬	piojɑŋ**nɑ-**	表彰する
piping 批评	pipiŋ**nɑ-**	叱る
jiefang 解放	dʒefɑŋ**nɑ-**	解放する

2.2.7. トゥヴァ語

漢語動詞がトゥヴァ語に借入されたとき、もとの動詞の後ろにトゥヴァ語の動詞化接辞 lA-が必要となる。たとえば以下のようである（吴宏伟 1999: 163）。

(11)

du 堵	du**la-**	詰まる
jie 借	dʒe**le-**	借りる

2.2.8. 満洲語

満洲語が漢語動詞を借用すると、一般的に動詞化接辞-nA, -lA を付す。たとえば以下のようである（王庆丰 2005：32-33）。

(12)

jiang 讲	giɑŋ**nɑ**	話す
cheng 称	gin**lə**	（はかりで）量る
zhaoxiang 照相	dʐo**lə**	写真を撮る
kaihui 开会	kɑixui**lə**	会議する

2.2.9. シベ語

漢語動詞がシベ語に借入されたとき、一般に同言語の動詞化接辞-lA を付す。たとえば以下のようである（李树兰、仲谦 1986：37）。

(13)

fenxi 分析	fənɕi**lə**m	分析する
bao 包	bau**lu**m	（生産、仕事を）一手に引き受ける
baozha 爆炸	baudʐa**lə**m	爆発する

2.2.10. オロチョン語

オロチョン語が漢語動詞を借用すると、一般的に動詞化接辞-lAA（-laa, -ləə）が必要である。たとえば以下のようである（胡增益 1986：204-205）。

(14)

shengchan 生产	ʃəŋtʃan**laa-**	生産する
kaihuai 开会	kajxue**laa-**	会議する
xuanju 选举	ʃuan**la-**	選挙する
xuanchuan 宣传	ʃuantʃuan**laa-**	宣伝する
biaoyang 表扬	beewjaŋ**laa-**	表彰する
jiefang 解放	dʒeeɸaŋ**laa-**	解放する
yonghu 拥护	juŋku**ləə-**	擁護する
piping 批评	pipiŋ**ləə-**	叱る
duo 剁	doo**ləə-**	（肉を）ぶつ切りにする

2.2.11. エヴェンキ語

オロチョン語は漢語動詞を借用することが少ない。もし漢語動詞を借用するならば、一般的に漢語動詞の後ろに動詞化接辞-lArAn を付す必要がある。たとえば以下のようである（語例は烏日格喜樂圖博士による）。

(15)

xuanju 选举	shyan**laran**	選挙する
fasong 发送	faa**laran**	発送する
yonghu 拥护	junhu**leren**	擁護する

2.2.12. ヘジェ語

ヘジェ語が漢語動詞を借用するとき、一般的に漢語動詞の後ろに動詞化接辞-lA を付す必要がある。たとえば以下のようである（金莉娜、程亚恒 2019：74）。

(16)

gai 改	kai**lə**	改める
suan 算	suan**lə**	計算する
lingdao 领导	liŋtau**lə**	導く
dongyuan 动员	tuŋyan**lə**	動員する
biaoyang 表扬	piaujaŋ**la**	表彰する
piping 批评	$p^hip^hiŋ$**la**	叱る
douzheng 斗争	təutʂəŋ**la**	闘争する
luosuo 啰嗦	losu**ləi**	長々としゃべる

以上に掲げた例から分かるように、アルタイ諸語のモンゴル諸語のモンゴル語、モングォル語、マネガチャ語、シラ・ユグル語、チュルク諸語のカザフ語、サリグ・ユグル語、トゥヴァ語、満洲・ツングース諸語の満洲語、シベ語、エヴェンキ語、オロチョン語、ヘジェ語など 12 種のアルタイ諸語は、漢語動詞を借用するとき本来の漢語動詞の後ろに各民族語の動詞化接辞を付す。これらは表 1 のようにまとめられる。

表 1：添加される動詞化接辞

語族	言語	動詞化接辞
モンゴル諸語	モンゴル語	-l/-n
	モングォル語	-gə/-la
	マネガチャ語	-la/-le/-lo;-ka/-gə
	シラ・ユグル語	-la/-le/-lo
チュルク諸語	カザフ語	-dɑ/-le
	サリグ・ユグル語	-lɑ/-le/-ne/-nɑ
	トゥヴァ語	-la/-le
満洲・ツングース諸語	満洲語	-nɑ/-lə
	シベ語	-lə/-lu
	エヴェンキ語	-laran/-leren
	オロチョン語	-laa/-ləə
	ヘジェ語	-la/-lə/-ləi

2.3. 軽動詞の付加

漢語動詞が少数民族言語に借入されるとき、いくつかのチベット・ビルマ諸語、朝鮮語、タジク語は漢語動詞を名詞とし、本来の漢語動詞の後ろに各民族言語の「する」「打つ、放つ」「置く」などの意味を表す動詞を付加する。この現象は、研究者によっては「する（do）」による迂言法（Jäger 2004）と呼ばれるが、ここではこれを軽動詞（light verb）添加法と呼ぶことにする[6]。

[6] 軽動詞（light verb）という概念は、最初にデンマークの言語学者イェスペルセン（Jerspersen 1954: 24）が1954年に提出したものである。この概念が提出されることで、たとえば“take a look”という構造におけるtakeの統語的な属性が定められることになった。このタイプの構造におけるtakeが担う語義はlookに比べて軽い

2.3.1. 朝鮮語

漢語動詞が朝鮮語に借入されたのち、必ず固有の動詞-hata「する」と連用することで、はじめて動詞になれる。たとえば以下のようである。

(17)

fangwen 访问	paŋmun **hata**	訪問する
laodong 劳动	rotoŋ **hata**	労働する
jingguo 经过	kyəŋkwa **hata**	経過する
bianhua 变化	pyənhwa **hata**	経過する
shiyong 使用	sayoŋ **hata**	使用する

2.3.2. ダクパ語

漢語動詞またはチベット語動詞がダクパ語に借入されたのち、必ず固有の動詞 ja^{35}「する」と連用することで、はじめて動詞になれる。たとえば以下のようである（陆绍尊 2002：334）。

(18)

piping 批评	$p^hi^{55}p^hiŋ^{55}\mathbf{ja^{35}}$	叱る
チベット語借用語[訳者注：rtsi]	$tsi^{53}\mathbf{ja^{35}}$	計算する
チベット語借用語[訳者注：tshod]	$ts^hø^{53}\mathbf{ja^{35}}$	心配する

2.3.3. 四土ギャロン語

漢語動詞が四土ギャロン語チョクツェ方言に借入されたのち、固有の動詞を構成する成分である ka-pa「する」、kɐ-lɐt「打つ、放つ」、ka-ta「置いておく」などを伴って初めて独立した動詞を形成することができる。この３つの動詞構成成分は生産性がある（林向荣 1993: 138）。

(19)

a.	xiang 想	ʃiaŋ **ka-pa**	思う
	zhunbei 准备	tʂuən pi **ka-pa**	準備する
	zao 造	ts^hou **ka-pa**	製造する
	suanzhang 算账	swɐntʂɐŋ **ka-pa**	帳簿を計算する
b.	shua 刷	ʃwaʃwa **kɐ-lɐt**	刷る
	fangjia 放假	faŋtɕa **kɐ-lɐt**	休暇になる
c.	zheng 蒸	tʃən **ka-ta**	蒸す

ため、この種の構造におけるtakeなどの語が軽動詞と画定された。1990年代にチョムスキーが軽動詞の概念を変形生成文法研究に導入し、現在のように海外の言語類型論界ですでに広くこの術語が用いられている。このことから、本稿でもまたこの術語を使用する。

2.3.4. クロスキャプ語

漢語動詞またはチベット語動詞がクロスキャプ語に借入されたのち、必ず固有の動詞 vi^{33}「する、作る」と連用することで、はじめて動詞になれる。たとえば以下のようである（黄布凡 2007：140-141）。

(20)

a.	changxi 唱戏	tʂ^{h}ɑŋ^{55}ɕi^{55}**vi^{33}**	歌舞劇を演じる	
	kaihui 开会	k^{h}e^{33}χi^{53}**vi^{33}**	会議する	
b.	チベット語借用語[訳者注：'chams]	ntɕ^{h}am^{53}**vi^{33}**		チャムを踊る
	チベット語借用語[訳者注：re mos]	rɛl^{33}mu^{53}**vi^{33}**		順番にする

2.3.5. ゴチャン語

ゴチャン語が漢語から借入した動詞が、もし単音節であれば直接借用し、2音節であれば軽動詞 pe^{353}「する」を付加する。たとえば以下のようである。

(21)

baomi 保密	pau^{53} mi^{55} **pe^{353}**	秘密にする
fangjia 放假	xuɔŋ^{31} tɕa^{53} **pe^{353}**	休暇になる
fanfa 犯法	xuɛ̃^{31} xa^{55} **pe^{353}**	法を犯す
shangwang 上网	sɔŋ^{31} wɔŋ^{53} **pe^{353}**	インターネットに接続する

2.3.6. トゥルン語

トゥルン語が漢語から借入した動詞はすべて2音節である。これらの2音節動詞はトゥルン語の中で使用されるとき、みなトゥルン語の語構成の成分 wa^{53} を付加する必要がある[7]。たとえば以下のようである（孙宏开 1982：46-47）。

(22)

canjia 参加	tsɑntɕɑ^{55}**wɑ^{53}**	参加する
gongzuo 工作	kuŋ^{55}tsɔ^{53}**wɑ^{53}**	仕事する
baozheng 保证	pɔ^{55}tsɯŋ^{53}**wɑ^{53}**	保証する
jieshao 介绍	kɑi^{55}sɔ^{55}**wɑ^{53}**	紹介する
xiaomie 消灭	ɕɔ^{55}mje^{31}**wɑ^{53}**	消滅する
yikao 依靠	i^{55}kɔ^{55}**wɑ^{53}**	頼る

2.3.7. タジク語

タジク語はインド・ヨーロッパ語族イラン語派に属する。漢語の動詞的な概念を借用すると、瓦罕タジク語では意訳するかウイグル語を借用するか、また

[7] 楊将領博士に確認を求めたところ、もし否定を表す場合、否定標識はwa^{53}の前に現れる。ことことから分かるように、wa^{53}は真正な動詞である。

はウイグル語＋瓦罕タジク語の組み合わせを用いるかの方法で表す。現在のところ、以下の比較的典型的な漢語動詞が収集できただけである。漢語動詞の後ろにこの民族言語における軽動詞 xak「する」を付加する。たとえば次のようである[8]。

(23)

geli 隔离	gəli **xak**	隔離する
saomiao 扫描	sawmjɔw **xak**	QR コードを読み取る
liudiao 流调	liwdjaw **xak**	流行疾病の疫学調査をする
hesuan 核酸	χəswan **xak**	PCR検査をする（他動詞）、 PCR検査を受ける（自動詞）

以上に掲げた例から分かるように、漢語の元の動詞の後ろに軽動詞を付加する言語は、主にチベット・ビルマ諸語のダクパ語、羌語支の四土ギャロン語、クロスキャプ語およびゴチャン語、ジンポー語支のトゥルン語、さらにアルタイ諸語の朝鮮語及びインド・ヨーロッパ語族イラン語派のタジク語がある。軽動詞を添加する言語は表２のようにまとめられる。

表２：軽動詞の添加される動詞化接辞

語派	語支	言語	軽動詞
チベット・ビルマ語派	チベット語支	ダクパ語	ja^{35}「する」
	羌語支	四土ギャロン語	ka-pa「する」、kɐ-lɐt「打つ、放つ」、ka-ta「置いたままにする」
		クロスキャプ語	vi^{33}「する、作る」
		ゴチャン語	pe^{353}「する」（２音節語）
	ジンポー語支	トゥルン語	wɑ^{53}「する」
満洲・ツングース語派		朝鮮語	hatɑ「する」
イラン語派		タジク語	xak「する」

2.4. 二種類の間接的な方法（接辞と軽動詞）

モンゴル諸語のサンタ語、康家語、チュルク諸語のウイグル語、キルギス語、サラル語、チベット・ビルマ諸語のルマ語は、２種類の間接的な方法を用いる。もし借入した漢語動詞が単音節であれば、動詞化接辞または助詞を付加するという方法を用いる一方、もし借入した漢語動詞が非単音節であれば、一般に各民族言語の軽動詞「する」、「放す」または「あたる、称する」を付加する。

[8] 瓦罕タジク語の語例は、南開大学外国語学院の侯典峰博士の提供による。

2.4.1. サンタ語

サンタ語が漢語から借入した動詞が単音節のときは、もとの借用語の後ろにサンタ語の動詞を派生させる付加成分-la, -liə, - dʑi, -ji 付加する。たとえば(24)のようである。借用した漢語動詞が 2 音節のときは、漢語動詞の後ろにサンタ語の軽動詞 giə「する」を付加する。たとえば(25)のようである。

(24)

gao 搞	gao**la-**	する
bang 绑	ban**la-**	結ぶ
bei 背	bə**iliə-**	背負う
jie 接	dʑiə**liə-**	迎える
zhan 粘	dʐan**dʑi-**	ひっつく
bian 编	biən**dʑi-**	編む
la 拉	la**ji-**	引っ張る
dui 堆	dui**ji-**	積む

(25)

jiefang 解放	dʑiəfan**giə-**	解放する
fazhan 发展	fadʐan**giə**	発展する
bangzhu 帮助	bandʐu**giə-**	助ける
kaihui 开会	kaixui**giə-**	会議する

2.4.2. 康家語

漢語動詞が康家語に借入されたとき、漢語動詞の後ろに康家語の動詞付加成分-gA「する、やる」（斯琴朝克图 1999: 289）及び-lA（斯琴朝克图 1999: 66, 303, 304）を付加する。

(26)

ran 染	ʐaŋ**gi-**	染める
dang 当	daŋ**gi-**	なる
kaihui 开会	kaiχui**gi-**	会議する
kao 靠	kau**gi-**	寄りかかる
quan 劝	tʃyen**gi-**	勧める
qifu 欺负	tʃifʉ**gi-**	いじめる
rang 让	ʐaŋ**la-**	させる
suo 锁	sʉdzɯ**la-**	鍵をかける
zhuang 装	dʒɔ**la-**	詰め込む
zhao 照	dʐau**la-**	写真を撮る
zai 栽	dzai**la-**	植えつける、移植する

2.4.3. ウイグル語

ウイグル語が借用した漢語動詞が単音節のとき、動詞化接辞-lA を付加する。借用したのが２音節であれば、一般的に軽動詞 qil-「する」を付加する。ウイグル語に借入された漢語単音節動詞は、さらにまず漢語助詞「了」を付加し、次に軽動詞を加えることができる。ウイグル語の口語資料から見ると、現在のところ軽動詞を付加する現象が多数を占める。

(27)

suo 锁　so**la-**　鍵をかける
cui 催　süy**lä-**　催促する
ショートメールを送る
duanšin　fa　le　qil-
duanxin 短信　fa 发　le 了　する
充電する
tok　čoŋ　le　qil
電気　chong 充　le 了　する
カードをスワイプする
šua　le　qil-
shua 刷 le 了　する
shua　le　qil-
shua 刷 le 了　する

ウイグル語は軽動詞 qil-「する」を付加する以外に、さらに軽動詞 bol-「あたる、称する」を付加することができる。たとえば以下のようである。

(28)

U　tüyšü/ tüyšyu / tuixiu　**boldi**.
彼/彼女 tuixiu 退休 [退職する]　あたる/称する
「彼/彼女は退職した。」

2.4.4. サラル語

サラル語における漢語単音節動詞は同民族言語の語形成の付加成分-lA を付加して動詞を構成する。たとえば(29)のようである。２音節動詞の場合は同民族言語の軽動詞 et-「する」と連用する。たとえば(30)のようである（林莲云 1985：24）。

(29)

bao 包　bo**lɑ-**　包む
han 焊　xɑn**lɑ**　はんだづけする

die 叠	de**le**	折りたたむ
jie 接	ze**le-**	接続する

(30)

baogao 报告	+ et-「する」→	bogo	**et-**	報告する
fandui 反对	+ et-「する」→	fɑnduj	**et-**	反対する
canjia 参加	+ et-「する」→	tsɑndʒɑ	**et-**	参加する
shenchan 生产	+ et-「する」→	ʂəŋtʂɑn	**et-**	生産する

2.4.5. キルギス語

キルギス語は漢語動詞を借用するのが少ない。もし漢語動詞を借用するときは、単音節動詞には動詞化接辞-lA を付加し、非単音節動詞には軽動詞 qəl-「する」、qoj-「放つ」を添加する。

(31)

jiayou 加油	maj dʒa**la**		がんばる
saoma 扫码	armima-nə saolə **qəl**		コードを読み取る（命令法）
dabao fan 打包饭	tamaq-də	dabao	**qəl-**
	ごはん-ACC	dabao 打包	軽動詞「する」-
	食べ残したごはんをパックに入れる		
fakuan 罚款	pakan[9]	**qoj-**	
	fakuan 罚款	軽動詞「放つ」-	
	罰金を科す		

2.4.6. ルマ語

漢語単音節動詞がルマ語に借入されたのち、固有の語形成の助詞-t^hɑ を付加することで、はじめて動詞を構成する（孙宏开 1981: 69-71; 刘光坤 1998: 270; 黄布凡、周发成 2006: 207-271 参照）。たとえば榮紅ルマ語について以下のようである。

(32)

kao 靠	khɑu-**t^hɑ**	寄りかかる
han 焊	χɑn-**t^hɑ**	はんだづけする
jie 接	tɕe-**t^hɑ**	接続する
ji 挤	tɕi-**t^hɑ**	押し込む
song 送	soŋ-**t^hɑ**	送る
bao 包	pɑu-**t^hɑ**	包む

[9] キルギス語は/f/の音素を持たないため、借用語中に/f/の音素があれば、通常は/p/で置換される。

漢語２音節動詞がルマ語に借入されたのち、必ず同民族言語の軽動詞 pə～pu「する」と連用する（孙宏开 1981: 69-71; 刘光坤 1998: 271; LaPolla with Huang 2003: 36; 黄成龙 2006: 94 参照）。たとえば榮紅ルマ語について以下のようである。

(33)

baogao 报告	＋pə「する」→	pɑukɑu	**pə**	報告する
fandui 反对	＋pə「する」→	fɑndui	**pu**	反対する
zhunbei 准备	＋pə「する」→	tʂunpəi	**pə**	準備する

漢語動詞がいくつかの少数民族言語に借入されたのち、単音節、２音節の違いに基づいて異なる借用方法が採用される。これは主にアルタイ語族モンゴル語派のサンタ語と康家語、チュルク語派のウイグル語、キルギス語、サラル語、シナ・チベット語族チベット・ビルマ語派のルマ語の５種の言語に認められる。表３のようにまとめられる。

表３：２種類の間接的な方法を用いる諸言語

語族	語派	言語	単音節	２音節
アルタイ語族	モンゴル語派	康家語	-la/-le　-gi/-gə「する」	
		サンタ語	-la/-liə/-dzi/-ji	giə-「する」
	チュルク語派	ウイグル語	-la	qil-「する」、bol-「あたる、称する」
		キルギス語	-la	qəl-「する」、qoj-「放つ」
		サラル語	-la/-le	et-「する」
汉藏語系	チベット・ビルマ語派	ルマ語	-t^hɑ	pə/pu「する」

2.5. 構造の調整

白語の語順は漢語の影響を受けて SOV から SVO 構造に変化した。漢語の動目構造が白語に借用されたのちもなお動目構造を用いる。たとえば以下のようである。

(34)

chong dian 充电	ts^hoŋ^{44} ti^{55}	充電する
jiao che 叫车	ɣɯ^{35} ts^hɛ^{44}	車を呼ぶ
shangwang 上网	sa^{55}va^{31}	インターネットにつなぐ
zha yangyu 炸洋芋	tse^{35}ia^{42}ɣ^{55}	じゃがいもを揚げる
fa duanxin 发短信	fa^{35}tua^{31}ɕiɯ^{55}	ショートメールを送る
xia xiangqi 下象棋	ɕia^{55}ɕia^{55}tɕ^{h}i^{42}	象棋を指す

しかしながら、漢語の動目構造がいくつかの SOV 型の少数民族語に借入されたとき、借入した目標言語に固有の語順の規則に従い改めて組み合わされ、その成分について「賓動構造」に調整される。たとえば南部ビジ語では次のようである（徐世璇 2007）。

(35)

chou jin 抽筋	tɕĩ^{33}	+	tsʰa^{35}	筋肉がつる
	jin 筋		chou 抽	
	筋肉		引っ張る	
kai hui 开会	huei^{33}	+	kʰai^{13}	会議する
	hui 会		kai 开	
	会議		開く	
nian jing 念经	tɕĩ^{33}	+	niã^{13}	お経を読む
	jing 经		nian 念	
	お経		声に出して読む	
da dunr 打盹儿	ŋã^{55}pi^{13}	+	tsʰuã^{33}	うとうとする
	yanpi 眼皮		chuan 穿	
	まぶた		貫く	

漢語の動目構造がルマ語に借入されたあともまた、ルマ語に固有の語順の規則に従い改めて組み合わされ、その成分について調整され、さらに動詞化助詞 tʰɑ を加える。たとえば以下のようである。

(36)

pao cha 泡茶	tʂʰɑ	+	pʰɑu	+ tʰɑ	→	tʂʰɑ pʰɑu-tʰɑ
お茶を入れる	cha 茶		pao 泡	助詞		
	お茶		入れる			
suan zhang 算账	tʂɑn	+	suɑn	+ tʰɑ	→	tʂɑn suɑn-tʰɑ
勘定する	zhang 账		suan 算	助詞		
	帳簿		計算する			
kai che 开车	tʂʰetsə	+	kʰai	+ tʰɑ	→	tʂʰetsə kʰai-tʰɑ
車を運転する	chezi 车子		kai 开	助詞		
	車		運転する			

チベット・ビルマ諸語は白語を除いてみな SOV 言語であるため、多くの言語が漢語の動目式の句を借用したときもまた、各民族語の語順に基づいて動目構造を賓動構造に調整する。たとえば以下のようである。

(37)

東旺チベット語：	zha yangyu 炸洋芋 じゃがいもを揚げる	$jo^{24}ji^{55}$ $tʂa^{35}$
	xia xiangqi 下象棋 象棋を指す	$ɕã^{35}tɕ^{h}i^{41}$ $ɕa^{35}$
	fa duanxin 发短信 ショートメールを送る	$tuã^{41}ɕĩ^{24}$ fa^{35}
ナス彝語：	chong dian 充电 充電する	$tian^{55}tʂ^{h}o^{33}tse^{55}$
	da che 打车 車を呼ぶ	$ts^{h}ə^{33}ndhu^{21}$
	zha yangyu 炸洋芋 じゃがいもを揚げる	$ʑa^{11}ʑu^{55}tʂa^{21}$
アカ・ハニ語：	chong dian 充电 充電する	$tɛ^{55}$ $tsuŋ^{33}$
	da chuzuche 打出租车 車を呼ぶ	$ts^{h}u^{33}tsu^{33}ts^{h}ɤ^{33}$ ta^{55}
	fa duanxin 发短信 ショートメールを送る	$tua^{31}ɕi^{55}$ fa^{33}
福貢リス語：	chong dian 充电 充電する	tie^{35} $ts^{h}o^{55}$
	da motuo 打摩托 バイクを呼ぶ	$mo^{33}do^{33}$ $k^{h}u^{33}$
	fa xinxi 发信息 ショートメールを送る（させる）	$ɕi^{35}ɕi^{31}$ $fɑ^{31}$
魯甸ナシ語：	chong dian 充电 充電する	$tiɛ^{55}$ $tʂ^{h}o^{33}$
	da chezi 打车子 車を呼ぶ	$tʂ^{h}ə^{33}$ tsi^{33} $tɑ^{13}$
	fa duanxin 发短信 ショートメールを送る	$tuɛ^{33}ɕi^{55}$ $fɑ^{13}$
ゴチャン語：	chong dian 充电 充電する	$tiɛ̃^{353}$ $wu^{31}ts^{h}uŋ^{55}$
	da chezi 打车子 車を呼ぶ	$ts^{h}e^{55}$ $tsɿ^{31}$ ta^{55}
	kai fapiao 开发票 領収書を切る	$xua^{55}piau^{55}$ $khai^{55}$

ムニャ語[10]：	chong dian 充电 充電する	dian mətə
	da qiche 打汽车 車を呼ぶ	$tɕ^hit ʂ^he\ k^hələ$
	kai fapiao 开发票 領収書を切る	faphiew tudzi̱
雲南サドゥ語[11]：	da zhang 打仗 戦いに行く	$tsa^{55}tɛ^{313}$
	kai qiche 开汽车 車を運転する	$tɕ^hi^{55}ts^hə^{33}\ k^hɛ^{33}$
	fan zui 犯罪 罪を犯す	$tsui^{55}fa^{55}$

漢語の動目式の動詞句が部分的な少数民族言語に借入されたあと、形式上は直接動目句を借入したのちさらに動詞をコピーしてはじめて完全な支配的な意味を構成することができる。たとえば白宏ハニ語の例は以下のようである（徐世璇 2009）。

(38)

chang ge 唱歌 歌を歌う	$ts^ha^{55}ko^{33}$ +	chang 唱 歌う	ts^ha^{55} →	$ts^ha^{55}ko^{33}$ **ts^ha^{55}**
tiao wu 跳舞 踊りを踊る	$t^hiɔ^{55}v^{33}$ +	tiao 跳 踊る	$t^hiɔ^{55}$ →	$t^hiɔ^{55}v^{33}$ **$t^hiɔ^{55}$**
kai hui 开会 会議する	$k^hɛ^{55}xui^{55}$ +	kai 开 開く	$k^hɛ^{55}$ →	$k^hɛ^{55}xui^{55}$ **$k^hɛ^{55}$**
dian han 电焊 はんだづけする	$tiɛ^{55}xaŋ^{55}$ +	han 焊 はんだづけする	$xaŋ^{55}$ →	$tiɛ^{55}xaŋ^{55}$ **$xaŋ^{55}$**

漢語の動目構造が喀喇沁モンゴル語に借入されたとき、もし目的語が単音節のときは語順に変化はなく、なお動目構造を保持する。もし目的語が２音節名詞のときは語順に変化があり、同民族語に固有の賓動構造に調整される。たとえば以下のようである。

(39)

chong dian 充电	$tʃ^hʉŋtɛnna$	充電する
da che 打车	$tətʃ^hə:nə$	車を呼ぶ

[10] ムニャ語は漢語名詞の借用が比較的多く、ここに挙げた例は名詞のみを借用し、動詞は民族語である。また、語順は賓動構造である。

[11] 许鲜明 ほか(2019:51)を参照。

shangwang 上网	ʂaŋwaŋna	インターネットにつなぐ
zha tudou 炸土豆	tʰʉtoʉ tʂa:nə	じゃがいもを揚げる
fa duanxin 发短信	tuanʃin fa:nə	ショートメールを送る

トゥルン語が漢語の動目句を借入したのち、漢語の語順を保持し、さらに同民族語の軽動詞 wɑ[53]「する」を付加する。たとえば以下のようである（孙宏开 1982：47）。

(40) shang ke 上课 + wɑ[53]「する」→ saŋ[55]kɔ[55]**wɑ[53]** 授業する
dai tou 带头 + wɑ[53]「する」→ tɑi[55]tɯ[31]**wɑ[53]** 率先する
pai dui 排队 + wɑ[53]「する」→ pɑi[55]tui[55]**wɑ[53]** 列に並ぶ
fang jia 放假 + wɑ[53]「する」→ faŋ[55]tɕɑ[53]**wɑ[53]** 休暇になる
chang ge 唱歌 + wɑ[53]「する」→ tsɑŋ[31]kɔ[55]**wɑ[53]** 歌を歌う
jing li 敬礼 + wɑ[53]「する」→ tɕin[55]li[53]**wɑ[53]** 敬礼する

以上のように、借用方法から見ると、接辞を付加する方法、また軽動詞を付加する方法、構造の調整というのは、ただ SOV 型で形態が豊富な膠着型と屈折型の言語にのみあてはまる。たとえば、アルタイ諸語、チベット・ビルマ諸語羌語支、チベット語支及びジンポー語支のいくつかの言語は、漢語の形態統語論との差異が比較的大きく、語順も漢語と異なるため、すべて直接借用を採用せず、間接借用（目標言語の接辞の付加または軽動詞の付加）と構造の調整を採用する。すなわち、漢語動詞の「動目構造」を目標言語の「目動構造」に調整し、そののち目標言語の動詞形態標識を付加することになる。

3. 動詞借用の類型論的特徴

言語類型論の研究では、ある種の語類は容易に借用されやすく、ある種の語類はあまり容易に借用されないことがすでに実証されている。このため、語彙の借用は同等ではない。ムイスケン（Muysken 1981）は語類の借用難易度の階層（integrated hierarchy）を提出した。この階層は、ウィンフォード（Winford 2003: 51）やマトラス（Matras 2008: 70）によって以下のように言い直されている：

名詞 > 形容詞 > **動詞** > 前置詞 > 並列接続詞 > 量詞 >
限定詞 > 自由代名詞 > 附属代名詞 > 従属接続詞

以上の階層は左側に行けば行くほど借用が容易であり、右へ行けば行くほど借用が困難であることを示している。これは、名詞が最も容易に借用され、次

いで形容詞、動詞となり、最も右側の従属接続詞が最も借用されにくいということである。

モラヴチク（Moravcsik）は最も早く動詞の借用の類型について通言語的に考察した研究者である。同氏はすでに 40 年前に動詞の借用の類型について通言語的に比較している（Moravcsik 1975）。ミフスドは記述言語学と比較言語学の方法を用いてマルタ語の動詞借用について研究を行った（Mifsud 1995）。2005 年、イギリスのマンチェスター大学のマトラスとサケルが「マンチェスター文法借用」プロジェクトを開始し、2008 年に『通言語的視点から見た文法借用』（Matras and Sakel 2008）を編集出版した。2003 年、ドイツのライプツィヒ大学進化人類学研究所のハスペルマートとタドモルが主宰となって「借用語類型論」プロジェクトを実施し、2009 年に２人で『世界の言語の借用語比較ハンドブック』（Haspelmath and Tadmor 2009）を編集出版した。2004 年以来、ヴィヒマンとヴォールゲムートは「借用語類型論」プロジェクトを利用して約 250 種の言語における動詞借用の状況を収集し、2008 年に共著として「類型論的視点から見た動詞借用」を発表した（Wichmann and Wohlgemuth 2008）。この中で著者らは動詞借用の方法を、軽動詞戦略（the light verb strategy）、間接挿入（indirect insertion）、直接挿入（direct insertion）、体系の転移（paradigm transfer）の４種に帰納した。ヴォールゲムートはさらに「借用語類型論データベース」を用いて博士論文を完成させ、2009 年に De Gruyter 出版社から単著『動詞借用の類型論』（Wohlgemuth 2009）を出版した。同氏は通言語的観点から直接借用、間接借用、軽動詞の方法、語形変化、語義借用、その他、識別不能、借用なしなどに分け、その分布は表４に示すようになる（Wohlgemuth 2009: 147）：

表４：動詞の借用手段の分布表

借用方法	言語	語派	語族	順位
直接借用	207	91	49	1
間接借用	86	42	22	3
軽動詞の方法	104	60	35	2
語形変化	3	2	2	6
その他	3	3	2	6
語義借用	18	15	12	4
識別不能	6	6	6	5
借用動詞なし	3	3	3	-

表４から見えるように、動詞借用の各種方法の分布は一様ではなく、直接借用、間接借用、軽動詞の方法の分布が相対的に広く、その中で直接借用の頻度が最も広く、49 の語族、91 の語派、207 種の言語に分布する。次に軽動詞の方法が 35 の語族、60 の語派、104 種の言語に分布し、３位には間接借用が 22 の語族、42 の語派、86 種の言語に分布する。４位になる語義借用は 12 の語

族、15 の語派、18 種の言語に分布する。その他の４種の借用方法である語形変化、識別不能、動詞借用なし、その他の借用方法の分布は非常に低い。

ヴォールゲムート（Wohlgemuth 2009: 157）はさらに８種の動詞借用方法の出現頻度について統計を取っている。統計から判明したのは、表５に示すようなそれぞれの借用方法の頻度である：

表５：動詞の借用手段の頻度

借用方法	言語	パーセンテージ	例	例のパーセンテージ
直接借用	207	58.8%	309	52.5%
間接借用	86	24.4%	121	20.6%
軽動詞の方法	104	29.5%	140	23.8%
語形変化	3	0.8%	3	0.5%
その他	30	8.5%	15	2.6%
総計	352	-	588	100%

表５から分かるように、動詞借用の各種方法の出現頻度もまた一様ではなく、直接借用、間接借用、軽動詞の方法の出現頻度が相対的に高く、その中で直接借用の頻度が最も高い。次に軽動詞の方法が高く、その次に間接借用となり、その他の借用方法の出現頻度は非常に低い。

中国少数民族言語が漢語動詞を借用する方法の分布から見ると、ヴォールゲムートの通言語的統計データといくつかの点で異なりがある。中国少数民族言語が漢語動詞を借用する場合、直接借用の方法を採用する分布が最高で、接辞の付加（間接的な方法）が２位となり、軽動詞を付加する方法を採用するのが３位となる。

中国少数民族言語が漢語動詞を借用するとき、主に直接借用、接辞の添加、軽動詞の付加、構造の調整という４種の方法を採用する。それぞれの借用方法の特徴は表６のようになる。

表６：動詞の借用手段の頻度

借用手段	形式の借入	意味の借入	目標言語との緊密度	動詞の借入状況
直接借用	あり	あり	緊密でない	単純
接辞の付加	あり	あり	緊密である	複雑
軽動詞	あり	あり	あまり緊密でない	複雑
構造の調整	あり	あり	緊密である	複雑

以上４種の借用類型は、漢語動詞の形式と意味がみな借入されているが、少数民族言語との緊密度（degree of integration）、完全な動詞（a full verb）であるかどうか、動詞の借入状況の面で差異がある。直接漢語動詞を借用する場合、単に形式と意味がともに借入されるだけでなく、少数民族言語の中でも完全な動詞として用いられるため、少数民族言語との結合が緊密でなく、借用方法も

比較的単純である。接辞の付加は漢語動詞の形式と意味を借用するとはいえ、少数民族言語において単独で動詞になれず、少数民族言語の動詞接辞をを付加することで複合動詞（complex verb）を構成する必要があり、少数民族言語との結合が緊密であるため、借用方法は複雑である。漢語動詞の後ろに軽動詞を付加する方法は、動詞が不完全な動詞（a non-full verb）であるために、それ自身では述語になれず、少数民族言語の軽動詞を付加する必要があり、複雑な述部（a complex predicate）の一部となって、漢語から借用した動詞との結合は緊密でないが、借用方法はなお複雑である。構造の調整も形式と意味をすべて借入しているものの、少数民族言語の語順規則へと調整する必要があり、少数民族言語との結合は緊密で、借用方法もまた複雑である。構造の調整という方法は、現在のところいくつかの SOV 型言語の中にのみ見つかる。

本研究から分かるように、ある言語ではただ１種の動詞借用の方法しかなく、別の言語では２種の借用の方法があり、また別の言語では３種の借用の方法がある。詳細は表７を参照。

表７：中国少数民族言語が動詞を借用する常用の手段

言語				動詞借用の手段			
語族	語派	語支	言語	直接借用	構造の調整	軽動詞の付加	接辞の付加
シナチベット語族	侗台語派	全語支	全言語	+	-	-	-
	ミャオ・ヤオ語派	全語支	全言語	+	-	-	-
オーストロアジア語族	モン・クメール語派	全語支	全言語	+	-	-	-
オーストロネシア語族			全言語	+	-	-	-
シナチベット語族	チベット・ビルマ語派	チベット語支	チベット語	+	+	-	-
		チベット語支	ダクパ語	-	+	+	
		彝語支	多数の言語	+	+	-	-
		ビルマ語支	多数の言語	+	+	-	-
		ジンポー語支	トゥルン語	+	+	-	-
		羌語支	スタウ語	+	+	-	-
		羌語支	ゴチャン語	+	+	+	-

		羌語支	四土ギャロン語	-	+	+	-
		羌語支	クロスキャプ語	-	+	+	-
		羌語支	ルマ語	-	+	+	+
アルタイ語族	チュルク語派	西匈語支	ウイグル語	-	+	+	+
		东匈語支	キルギス語	-	+	+	+
		西匈語支	カザフ語	-	-	-	+
		东匈語支	サリグ・ユグル語	-	-	-	+
			トゥヴァ語	-	-	-	+
	モンゴル語派		モンゴル語	-	+	-	+
			シラ・ユグル語	-	+	+	-
			モングォル語、マネガチャ語	-	-	?? [12]	+
	満洲・ツングース語派	２語支	５種の言語	-	-	-	+
			朝鮮語	-	-	+	-
印欧語族	イラン語派		瓦罕タジク語	-	-	+	-

表７から分かるように、シナ・チベット語族侗台語派、ミャオ・ヤオ語派、オーストロネシア語族モン・クメール語派、台湾のオーストロネシア語族は、漢語動詞を借用するとき直接借用を採用している。アルタイ語族チュルク語派のカザフ語、サリグ・ユグル語、トゥヴァ語、満洲・ツングース語派は、漢語動詞を借用するとき各少数民族語の動詞化接辞を付加することのみ採用している。朝鮮語と瓦罕タジク語は軽動詞の方法のみ採用している。多くのチベット・ビルマ諸語は直接借用と構造の調整という２種の方法を採用している。チベット・ビルマ語派羌語支の四土ギャロン語、クロスキャプ語、モンゴル語派のシラ・ユグル語は軽動詞と構造の調整という２種の方法を採用している。モンゴル語は動詞化接辞の付加と構造の調整という２種の方法を採用している。ゴチャン語は直接借用、軽動詞、構造の調整という３種の方法を採用している。アルタイ語族チュルク語派のウイグル語、キルギス語、シナ・チベット語族チベット・ビルマ語派のルマ語はすべて接辞/助詞の付加、軽動詞、構造の調整という３種の方法を採用している。

[12] ２つの疑問符は、先行研究にはモングォル語とマネガチャ語が軽動詞の方法を採用しているかどうか言及がないため、現段階ではこの２種の言語がこの方法を採用しているかどうかは不確定であることを示す。

4. 結び

本稿では、主に少数民族言語が漢語動詞を借用する方法及びその類型論的特徴について議論し分析した。少数民族言語が漢語動詞を借用するとき、主に直接借用、動詞化接辞の付加、軽動詞の付加、および構造の調整という４種の方法を採用している。直接借用は中国国内の民族言語の中で最もよく見られ、主に西南、中南地区の侗台諸語、ミャオ・ヤオ諸語、オーストロアジア語族モン・クメール諸語、台湾のオーストロネシア諸語、および大多数のチベット・ビルマ諸語に認められる。動詞化接辞の付加は、主に北方のアルタイ諸語に認められる。軽動詞を付加する方法は、主にシナ・チベット語族のダクパ語、四土ギャロン語、クロスキャプ語、トゥルン語など少数のチベット・ビルマ諸語、アルタイ系満洲・ツングース諸語の朝鮮語およびインド・ヨーロッパ語族イラン語派をタジク語に認められる。動詞化接辞の付加と軽動詞の方法を同時に採用するものは、主に北方のアルタイ系モンゴル諸語の康家語とサンタ語、チュルク諸語のウイグル語、キルギス語、サラル語、およびシナ・チベット語族のルマ語に認められる。中国少数民族言語が漢語動詞を借入するとき、異なる借用方法を採用するとはいえ、借入した漢語動詞が目標言語の中で対応する語が存在しないため、多くは空白を埋めることになる。

少数民族言語が漢語動詞を借用するときに採用する異なる方法の分布と頻度に基づくと、少数民族言語による漢語動詞の借用には、次のような特徴がある：（1）内部一致性：アルタイ系諸言語が漢語動詞を借用するときはみな直接借用をとらず、動詞化接辞を付加するか、軽動詞を付加するか、またはこれら両者の方法を兼用するかのいずれかで、言語系統内部での一致性を見せる；（2）地域性：東南、華南、および西南地区の大多数の言語における構造は類型的に全く異なっているが、ほとんどが漢語動詞を直接借用し、その地域性が突出している；（3）音節構造の差異：６種の言語が２種以上の借用方法を採用しているが、どの方法を採用するかは漢語動詞の音節によって決まっている。単音節の漢語動詞を借入するときは動詞化接辞または助詞を付加する方法を採用し、もし借入するのが非単音節の漢語動詞であるならば、一般的に軽動詞の方法を採用する；（4）言語構造の類型の差異：少数民族言語による動詞借用の類型は各言語の内部構造と一定の関係がある。たとえば北方のアルタイ系諸言語、西南の形態論的に複雑な少数のチベット・ビルマ諸語、朝鮮語さえもすべて膠着語であり、みな漢語動詞を直接借用することはない。漢語と構造の類型が類似している分析型の言語における借用方法は比較的単一で、みな漢語動詞を直接借用する。漢語と構造の類型が比較的大きく異なっていて、形態論が比較的複雑な膠着語が漢語動詞を借用する場合は、単に間接的な方法を採用するだけでなく、いくつかの言語は２、３種の漢語動詞借用の方法を採用している。タジク語が漢語動詞を借用することは少なく、基本的にウイグル語動詞

を借用するが、それでも直接ウイグル語動詞を借用することはなく、軽動詞の方法を採用する。このため、漢語動詞の借用方法は言語構造と密接に関係しているといえる。過去の研究では SOV 型の言語が漢語の動目式動詞句を借用する特徴について十分に注目してこなかったが、今後は漢語の動目式動詞句が少数民族言語に借入される際の特徴、分布、頻度について、さらに詳細な調査研究が望まれる。

参考文献/References

布和、刘照雄 (1982) 《保安语简志》，民族出版社。
陈宗振、雷选春 (1985) 《西部裕固语简志》，民族出版社。
道布 (1983) 《蒙古语简志》，民族出版社。
高尔锵 (1985) 《塔吉克语简志》，民族出版社。
耿世民、李增祥 (1985) 《哈萨克语简志》，民族出版社。
胡增益 (1986) 《鄂伦春语简志》，民族出版社。
黄布凡、周发成 (2006) 《羌语研究》，四川人民出版社。
黄布凡 (2007) 《拉坞戎语研究》，民族出版社。
黄成龙 (2006) 《蒲溪羌语研究》，民族出版社。
蒋宏军 (2010) 《哈萨克语中的汉语借词》，《新疆大学学报》第2期。
季永海 (1985) 《论满语中的汉语借词》，《满语研究》第1期。
金莉娜、程亚恒 (2019) 《黑龙江同江赫哲语》，商务印书馆。
李得春 (1986) 《关于朝鲜语里的汉语借词》，《延边大学学报》第2期。
李树兰、仲谦 (1986) 《锡伯语简志》，民族出版社。
林莲云 (1985) 《撒拉语简志》，民族出版社。
林向荣 (1993) 《嘉戎语研究》，四川民族出版社。
刘光坤 (1998) 《麻窝羌语研究》，四川民族出版社。
刘照雄 (1981) 《东乡语简志》，民族出版社。
陆绍尊 (2002) 《门巴语方言研究》，民族出版社。
敏春芳 (2012) 《东乡语汉语借词研究》，《汉语史研究集刊》第15辑。
沙加尔、徐世璇 (2002) 《哈尼语中汉语借词的历史层次》，《中国语文》第1期。
斯琴朝克图 (1999) 《康家语研究》上海远东出版社。
石常艳 (2015) 《撒都语的借词类型与特点》，《玉溪师范学院学报》第6期。
石林 (1994) 《侗语中汉语新借词的读音》，《民族语文》第5期。
孙宏开 (1981) 《羌语简志》，民族出版社。
孙宏开 (1982) 《独龙语简志》，民族出版社。
孙宏开、胡增益、黄行（主编） (2007) 《中国的语言》，商务印书馆。

吴宏伟 (1999) 《图瓦语研究》，上海远东出版社。

许鲜明、白碧波、尹明 (2019) 《云南玉溪撒都语》，商务印书馆。

徐世璇 (2007) 《从南部土家语的特殊构词看语言接触的深层影响》，《东方语言学》第1期。

徐世璇 (2009) 《语言接触与词汇系统的重整》，中央民族大学学术报告， 2009年12月。

徐世璇 (2014) 《南部土家语中汉语借词的特点》，《百色学院学报》第1期。

徐悉艰、徐桂珍 (1984) 《景颇族语言简志（载瓦语）》，民族出版社。

闫新红、欧阳伟 (2000) 《试论维吾尔语中的汉语借词》，《喀什师范学院学报》第3期。

严木初 (2013) 《论嘉戎语中的汉语借词》，《阿坝师范学院学报》第1期。

玉珍 (1996) 《藏语安多方言同仁话中的汉语借词》，《中国藏学》第1期。

曾晓渝 (2003) 《论水语里的近、现代汉语借词》，《语言研究》第2期。

郑贻青 (1995) 《回辉话中的汉语借词及汉字读音》，《民族语文》第5期。

张建民 (2011) 《夏河藏语中的汉借词与汉语西北方音》，《中国藏学》第2期。

照那斯图 (1981) 《土族语简志》，民族出版社。

照那斯图 (1981) 《东部裕固语简志》，民族出版社。

朱文旭 (1997) 《凉山彝语中的汉语借词》，《民族语文》第4期。

Aikhenvald, Alexandra Y. and R.M.W. Dixon (eds.) (2001) *Areal diffusion and genetic inheritance: Problems in comparative linguistics*. Oxford: Oxford University Press. doi: https://doi.org/10.1093/oso/9780198299813.001.0001

Curnow, Timothy J. (2001) What language features can be 'borrowed'? In Alexandra Aikhenvald and R. M. W. Dixon (eds.) *Areal diffusion and genetic inheritance: Problems in comparative linguistics*, 412–436. Oxford: Oxford University Press. doi: https://doi.org/10.1093/oso/9780198299813.003.0015

Campbell, Lyle (1993) On proposed universals of grammatical borrowing. In Robert J. Jeffers and Henk Aertsen (eds.) *Historical Linguistics 1989: Papers from the 9th International Conference on Historical Linguistics*, 91–109. Amsterdam / Philadelphia: John Benjamins. doi: https://doi.org/10.1075/cilt.106.08cam

Field, Fredric (2002) *Linguistic borrowing in bilingual contexts*. Amsterdam / Philadelphia: John Benjamins. doi: https://doi.org/10.1075/slcs.62

Gentner, Dedre (1982) Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. In S.A. Kuczaj (ed.) *Language development, Vol. 2: Language, thought and culture*, 301–334. Hillsdale, NJ: Erlbaum.

Haugen, Einar (1950) The analysis of linguistic borrowing. *Language* 26: 210–231. doi: https://doi.org/10.2307/410058

Haspelmath, Martin (2008) Loanword typology: Steps toward a systematic cross-linguistic study of lexical borrowability. In Thomas Stolz, Dik Bakker and Rosa S. Palomo (eds.) *Aspects of language contact: New theoretical, methodological and empirical findings with special focus on Romancisation processes*, 43–62. Berlin / New York: Mouton de Gruyter. doi: https://doi.org/10.1515/97831102060 43.43

Haspelmath, Martin and Uri Tadmor (eds.) (2009) *Loanwords in the world's languages: A comparative handbook*. Berlin: Mouton de Gruyter. doi: https://doi.org/10.1515/9783110218442

Heath, Jeffrey (1984) Language contact and language change. *Annual Review of Anthropology* 13: 367–384. doi: https://doi.org/10.1146/annurev.an.13.100184.002055

Heine, Bernd and Tania Kuteva (2005) *Language contact and grammatical change*. Cambridge: Cambridge University Press. doi: https://doi.org/10.1017/CBO9780511614132

Johanson, Lars (2002) *Structural factors in Turkic language contacts*. London: Curzon.

Johanson, Lars (2005) On copying grammatical meaning. S*prachtypologie und Universalienforschung (STUF)* 58: 75–83.

LaPolla, Randy J. with Chenglong Huang (2003) *A grammar of Qiang with annotated texts and glossary*. Berlin: Mouton de Gruyter. doi: https://doi.org/10.1515/9783110197273

Laufer, Berthold (1916) Loan-words in Tibetan. *T'oung Pao* 17(4/5): 403–552. doi: https://doi.org/10.1163/156853216X00157

Li, Fang-kuei (1945) Some Old Chinese loan words in the Tai languages. *Harvard Journal of Asiatic Studies* 8(3/4): 333–342. doi: https://doi.org/10.2307/2717820

Matras, Yaron (1998) Utterance modifiers and universals of grammatical borrowing. *Linguistics* 36(2): 281–331. doi: https://doi.org/10.1515/ling.1998.36.2.281

Matras, Yaron (2008) The borrowability of structural categories. In Yaron Matras and Jeanette Sakel (eds.) *Grammatical borrowing in cross-linguistic perspective*, 31–73. Berlin / New York: Mouton de Gruyter. doi: https://doi.org/10.1515/9783110199192.31

Matras, Yaron and Jeanette Sakel (2007) Investigating the mechanisms of pattern-replication in language convergence. *Studies in Language* 31(4): 829–865. doi: https://doi.org/10.1075/sl.31.4.05mat

Matras, Yaron and Jeanette Sakel (eds.) (2008) *Grammatical borrowing in cross-linguistic perspective*. Berlin / New York: Mouton de Gruyter. doi: https://doi.org/10.1515/9783110199192

Mifsud, Manwel (1995) *Loan verbs in Maltese: A descriptive and comparative study*. Leiden: Brill. doi: https://doi.org/10.1163/9789004348387

Moravcsik, Edith (1975) Borrowed verbs. *Wiener linguistische Gazette* 8:3–30.

Morimoto, Yukiko (2000) Loan words and their implications for the categorial status of verbal nouns. *Proceedings of the 25th Annual Meeting of the Berkeley Linguistics Society*, 371–382. doi: https://doi.org/10.3765/bls.v25i1.1192

Muysken, Pieter (1981) Halfway between Quechua and Spanish: The case for relexification. In Arnold Highfield and Albert Valdman (eds.) *Historicity and variation in Creole studies*, 52–78. Ann Arbor: Karoma.

Poplack, Sandra and David Sankoff (1984) Borrowing: the synchrony of integration. *Linguistics* 22: 99–136. doi: https://doi.org/10.1515/ling.1984.22.1.99

Ross, Malcolm (2001) Contact-induced change in Oceanic languages in north-west Melanesia. In Alexandra Y. Aikhenvald and R.M.W. Dixon (eds.) *Areal diffusion and genetic inheritance*, 134–166. Oxford: Oxford University Press. doi: https://doi.org/10.1093/oso/9780198299813.003.0006

Sanchez, Tara Savannah (2005) Constraints on structual borrowing in a multilingual contact situation. PhD Dissertation, University of Pennsylvania. URI: https://hdl.handle.net/20.500.14332/37644

Scotton, Carol M. and John Okeju (1973) Neighbors and lexical borrowings. *Language* 49(4): 871–889. doi: https://doi.org/10.2307/412066

Tadmor, Uri (2009) Loanwords in the world's languages: Findings and results. In Martin Haspelmath and Uri Tadmor (eds.) *Loanwords in the world's languages: A comparative handbook*, 55–75. Berlin: Mouton de Gruyter. doi: https://doi.org/10.1515/9783110218442.55

Tent, Jan and Paul A.Geraghty (eds.) (2004) *Borrowing: A pacific perspective*, 215–232. Canberra: Pacific Linguistics (Pacific Linguistics; 548). doi: https://doi.org/10.15144/PL-548

Thomason, Sarah G. (ed.) (1997) *Contact languages: A wider perspective*. Amsterdam: John Benjamins. doi: https://doi.org/10.1075/cll.17

Thomason, Sarah G. (2001) *Language contact. An introduction*. Edinburgh: Edinburgh University Press.

Thomason, Sarah G. and Terrence Kaufman (1988) *Language contact, creolization, and genetic linguistics*. California: University of California Press. doi: https://doi.org/10.1525/9780520912793

Weinreich, Uriel (1979[1953]) *Languages in contact: Findings and problems* (9th printing). New York: Linguistic Circle of New York.

Whitney, William Dwight (1881) On mixture in language. *Transactions of the American Philosophical Association* 12: 5–26. doi: https://doi.org/10.2307/2935666

Wichmann, Søren and Jan Wohlgemuth (2008) Loan verbs in a typological perspective. In Thomas Stolz, Dik Bakker and Rosa S. Palomo (eds.) *Aspects of language contact: New theoretical, methodological and empirical findings with special focus on Romanisation processes*, 89–122. Berlin: Mouton de Gruyter. doi: https://doi.org/10.1515/9783110206043.89

Winford, Donald (2003) *An introduction to contact linguistics*. Oxford: Blackwell.

Wohlgemuth, Jan (2009) *A typology of verbal borrowings*. Berlin: Mouton de Gruyter. doi: https://doi.org/10.1515/9783110219340

出版情報

投稿受理日：2024年7月10日
採用決定日：2024年9月11日

彙報
2023年度（2023年4月1日～2024年3月31日）

Memoranda and personalia 2023 (April 1, 2023 – March 31, 2024)

1. 全体

4 月 3 日にオンラインで新旧世話人引継ぎ会議を開催した。世話人交代に伴う新体制を確認するとともに、前代表世話人の遠藤光暁氏から会務の概要、分担及び注意点について説明があった。

6 月 10 日の大会開催日にハイブリッド方式（対面/オンライン）で世話人会議を開催し、年間を通じた会の運営・活動方針を討議した。下記はその決定事項。

1. 2024 年度の大会は大阪大学箕面キャンパスにて 6 月 1 日に開催すること。開催日を 10 日程度前倒しするのは、学会誌への投稿を促進するため。
2. 大会募集要項、学会誌投稿要領、モノグラフ投稿要領等の規定をわかりやすく提示するための改訂を加えるとともに、ホームページに英語版を加えること。

2. 大会関係（集会担当）

1. 4 件の発表申込があり、すべて受理した。このほか、遠藤光暁氏(青山学院大学)と Philippe del Giudice 氏(コートダジュール大学)を招待講演者として依頼した。
2. 大会開催校の鈴木史己氏（集会担当世話人）から大会プログラムと要旨集の原案が提示され、世話人会として承認した（4 月 26 日）。
3. 6 月 10 日、南山大学において、第 5 回大会を開催した。4 年ぶりに対面方式を復活させ、オンライン方式とのハイブリッドで実施した。参加者数は 32 名(対面 16 名、オンライン 16 名)であった。
4. 大会終了後、会計報告が鈴木史己氏より総務に提出された。
5. 大会終了後、要旨集が zenodo で公開された（6 月 20 日）。
6. 大会終了後、集会担当世話人の鈴木史己、清水政明両氏を中心に 2024 年度大会の運営について検討された。会場使用料、機材使用料等を支払う必要があり、学会の負担が重くなることから、できるだけ一日開催とすること。また、大会参加費は、2023 年度一般 1,000 円、学生無料であったが、コロナ以前の第 1 回大会は、一般 2,000 円、学生 1,000 円であったことに鑑み、一般 1,500 円、学生無料とする方針が決まった。

3. 学会誌関係（編集担当）

1. 『地理言語学研究３』には依頼原稿も含め、10 件の投稿があった。
2. 編集担当より、編集の報告があり（10 月 2 日）、世話人会議で承認した。
3. この号では、第 55 回国際漢蔵語学会（ICSTLL-55）ワークショップでの発表論文 4 篇を特集「シナ＝チベット語族への地理言語学的アプローチ」としてまとめることが提案され、世話人会議で承認した。但しこの措置は今後恒常的に「特集」を組むことを意味せず、必要な場合はその都度検討することとする。
4. PDF 版は zenodo で、また紙版は MyISBN で公開された。

4. モノグラフ関係（編集担当）

下記のモノグラフの公刊が承認され（括弧内は世話人会の承認日）、いずれも PDF 版は zenodo で、また紙版は MyISBN で公開された。

- Monograph No.5　Mark J. ALVES et al. eds. *Researching and Applying Linguistics and Vietnamese Language Studies*　(7 月 15 日)
- Monograph No.6　Trinh Cam Lan et al. eds. *Proceedings of the Fifth International Conference on Asian Geolinguistics*　(9 月 20 日)
- Monograph No.7　Chitsuko Fukushima et al. eds. *Linguistic Atlas of Asia and Africa III*　(同名の既刊Ⅰ、Ⅱの続編)　(11 月 29 日)

5. 学会諸文書の改訂と英訳（総務・WEB担当）

1. 6 月 1 日世話人会議（ハイブリッド）での決定を承け、7 月 1 日 の世話人会議（メール審議）において、「大会募集要項」「学会誌投稿要領」「モノグラフ投稿要領」の改定案が提案された。メール審議の上、一部修正の上承認した（7 月 9 日）。また、引き続き英語訳に取り掛かることとした。
2. 学会各種文書の英訳がネイティブチェックを経て完成したことから、先に承認した募集要項、投稿要領の改定案に加えて、会則等も含めて、日本語本文及び英訳を提示し、メール審議を行った。一部修正の上承認した（9 月 14 日）。
3. WEB 担当より、英語版を含むホームページの更新準備が整ったとの報告があり、総務担当世話人及び世話人会議でのチェックを経て、一部修正の上、更新した（12 月 27 日）。

6. 会計報告（総務・会計監査担当）

下記、会計監査担当世話人の監査を経た（2024 年 4 月 10 日付）後、世話人会で承認された。

収入

2022年度繰越金	31,362円
2023年度大会収入	22,794円
合計	54,156円

支出

学会HPサーバー年間使用料（さくらインターネット）	5,788円
学会HP英語化のための翻訳、校閲謝金（2年分）	20,000円
学会誌紙版公開手数料（MyISBN/デザインエッグ）	5,478円
振込手数料（330円×3回）	990円
次年度繰越金	21,900円
合　計	54,156円

正誤表

Errata et corrigenda

第 3 号/Volume 3

page, section, line	errata	corrigenda
p. 16, sec. 3.1, l. 3	0.9978	0.9133
p. 33, sec. 3.2, l. 9	0.9348	0.9491