Linguistic Atlas of Asia and Africa

II

Edited by
Hiroyuki Suzuki, Kohei Nakazawa, and Mitsuaki Endo

Geolinguistic Society of Japan

ISSN 2436-6471

Linguistic Atlas of Asia and Africa

II

EDITED BY
Hiroyuki Suzuki,
Kohei Nakazawa,
and Mitsuaki Endo

Geolinguistic Society of Japan
2023

Studies in Geolinguistics, Monograph series, No. 4
ISSN 2436-6471

Linguistic Atlas of Asia and Africa II, edited by Hiroyuki Suzuki, Kohei Nakazawa, and Mitsuaki Endo, 2023

Cover photo: Crop field in Zungchu (Songpan) © 2006 Tshewang nGyurmé

First published 2023
doi: https://doi.org/10.5281/zenodo.7754469

Published by: Geolinguistic Society of Japan
Website: https://geolinguistics.sakura.ne.jp/
Office address: Room G1305, Aoyama Gakuin Univeristy, 4-4-25 Shibuya, Shibuya-ku, Tokyo

TABLE OF CONTENTS

Contributors

EBIHARA Shiho	Tokyo University of Foreign Studies
ENDO Mitsuaki	Aoyama Gakuin University
FUKAZAWA Mika	National Ainu Museum
FUKUI Rei	The University of Tokyo
HIRANO Ayaka	Osaka University
IWASA Kazue	Nagoya University of Foreign Studies
IWASAKI Takamasa	Kyoto University
KIMURA Kimihiko	Tokyo University of Foreign Studies
KODAMA Nozomi	Kumamoto University
KOMORI Junko	Osaka University
KURABE Keita	Tokyo University of Foreign Studies
MATSUMOTO Ryo	Kobe City University of Foreign Studies
MINEGISHI Makoto	Tokyo University of Foreign Studies
NAGATO Youichi	J. F. Oberlin University Graduate Division
NAKAGAWA Hirosi	Tokyo University of Foreign Studies
NAKAO Shuichiro	Osaka Univeristy
NAKAZAWA Kohei	Shinshu University
ONO Chikako	Hokkai-Gakuen University
SAITÔ Yoshio	Takushoku University
SHIMIZU Masaaki	Osaka University
SHINAGAWA Daisuke	Tokyo University of Foreign Studies
SHIRAI Satoko	The University of Tokyo
SUZUKI Fumiki	Nanzan University
SUZUKI Hiroyuki	Kyoto University
TAGUCHI Yoshihisa	Chiba Univeristy
TOMITA Aika	Osaka Shoin Women's Univeristy
UTSUMI Atsuko	Meisei University
YAGI Kenji	Kokushikan University
YOKOYAMA Akiko	Tokyo University of Foreign Studies
YOSHIOKA Noboru	National Museum of Ethnology

❦

Acknowledgements

This volume is a part of the results of the ILCAA Joint Project 'Studies in Asian and African Geolinguistics' 2020–2022. Chapter 13 is a slightly corrected version of *Studies in Asian and African Geolinguistics I*, published by ILCAA in 2021. The articles are partially supported by JSPS Grants-in-Aid for Scientific Research (KAKENHI) JP18H05510, JP21H00354, JP20J40023, JP19K20801, JP18K12392, JP17H04774, JP18H00670, JP19H01254, JP20H00011, JP16H01925, JP20H04480, JP19K00567, JP18H03577, JP19K13187, JP20J40127, JP21K12993, JP22H04929, and JP22K18249, respectively.

❦

Abbreviations

#	heuristically reconstructed (Nilo-Saharan) root
[±ATR]	the 'advanced tongue root' feature
Akk	Akkadian
CDIAL	*A comparative dictionary of Indo-Aryan languages* (Turner 1985)
DEDR	*A Dravidian etymological dictionary* (Burrow & Emeneau 1984)
ES	Eastern Sudanic
IA	Indo-Aryan
LB	Lolo-Burman
n	the northern branch of Eastern Sudanic
PIE	Proto-Indo-European
PIr	Proto-Iranian
PKC	Proto-Kuki-Chin
PLB	Proto-Lolo-Burman
PTB	Proto-Tibeto-Burman
s	the southern branch of Eastern Sudanic
STEDT	*The Sino-Tibetan etymological dictionary and thesaurus*
TB	Tibeto-Burman
WrT	Written Tibetan

Chapter VII

Wheat

‘Wheat’ in Asian and African languages

Wheat was one of the first crops that humans cultivated, about 10,000 years ago, and along with corn (maize) and rice, it is one of the most important grains for human consumption today (Wheat Flour Institute 1976).

The cultivation of wheat is believed to have originated in central Eurasia from the Caucasus to Mesopotamia. Wheat cultivation developed around the Mediterranean Sea and then spread westward to Europe and eastward to Asia (*Nipponica*).

Table 1 shows the principal forms in our data.

Table 1: Main word forms for ‘wheat’.

Languages	Word forms
Koryak	*pʃenitsa* (< Rus.), *eliɣeqlevalqəl*
Ainu	*múnki ~ munki ~ muŋi*
Japonic	Japanese KOMUGI type Ryukyuan INAMUGI type
Korean	*mir* < MK *mirh*
Sinitic	*mai* 麦, *maizi* 麦子
Kra-Dai	mak 麦 type, mi:n 麵 type, su 黍 type, etc.
Tibeto-Burman	PTB *m-grwa Coronal initial type /m/-initial type
Austroasiatic	*gahũ*, *gohom*, *gɔhɔm*. *gandɔm* Khmer *srəv saaləy* Vietnamese *lúa mì*
Austronesian	
Tungusic	MAIS, OWEES, PULIMPU
Uralic	NIS-, VEHNA, TO-, Š-/Č-, BUZA
Mongolic	bugdai type
Turkic	bugdai type
Indo-Aryan	gōdhūma type kaṇika type
Dravidian	Tamil *kōtumai* Brahui *xolum* Kurukh *gohom*, Malto *gohme*
Iranian	PIr. *gantuma- PIr. *yava-
Caucasian	*xorbal-i*, *kobal-i*, etc. *q’ir*, *muq’e*, *buq’*, *rizq’i*, etc. *šibaratlw’*, *cibatl’ab*, etc.
Semitic	ħinṭa type qamħ type bar type ʃenraj type
Nilo-Saharan	Nile Nubian *illee* Western Nilotic Nuer *lap* Kwegu *nyamcele*, Me’en *sarguwo*, Suri/Tirma *goso*
Bantu	

Some languages lack words for wheat (e.g. Austronesian and Bantu). Other languages obviously borrowed words from other languages such as Koryak (from Russian), Ainu (from Japonic), Kra-Dai (from Sinitic), and Austroasiatic and Dravidian (from Indo-Aryan). In the areas where these languages are spoken, wheat has begun to be cultivated fairly recently, and it is thought that not only the plant but also the plant name were introduced from neighboring languages. The Mongolic and Turkic forms are also considered to be loanwords from one to the other.

Japonic *mugi* (< *mogi*?), Korean *mir* (< MK *mirh*), Sinitic *mai* (< MC *meak* < OC *m-rˁək), and Caucasian *muq’e* all start with /m/, and except for Caucasian, they are geographically close, and Japonic and Korean probably borrowed from Sinitic. Since the origin of wheat is near the Middle East, it is possible that Sinitic also borrowed from another language (cf. Caucasian *muq’e*).

Sinitic *mai* and Japonic *mugi* are general terms for ‘wheat’ and ‘barley’. Therefore, it is necessary to consider words for ‘barley’ as well. In addition, forms such as Manchu

mərə ‘buckwheat’ would also need to be compared.

Caucasian and Semitic have forms with no apparent relationship to forms of other languages; most Nilo-Saharan languages use loanwords for wheat, but a few unique words are found that have no apparent relationship to other languages. This may be because wheat was cultivated early in these areas.

It is interesting that Siamese ข้าวสาลี *khaaw khaaw saalii* and Lao ເຂົ້າສາລີ *khao sā lee* in Kra-Dai and Ryukyuan INAMUGI have a similar structure such that these words contain the morpheme for ‘rice’.

(NAKAZAWA Kohei)

‘Wheat’ in Chukotko-Kamchatkan

The Chukotko-Kamchatkan people (Chukchi, Alutor, Kerek, Koryak and Itelmen) did not engaged in agriculture before the arrival of the Russian. Their traditional means of livelihood were gathering, hunting, fishing and nomadism. Therefore, most of the agricultural vocabularies in Chukotko-Kamchatkan were derived from Russian.

‘Wheat’ is documented in Koryak as *pʃenitsa,* a word borrowed from Russian. There is also a word for ‘wheat’ in Koryak *eliɣeqlevalqəl* translated as ‘what becomes white bread’(Zhukova 1967).

(ONO Chikako)

☐A. *pʃenitsa* type

Figure 7.1.1: ‘Wheat’ in Chukotko-Kamchatkan.

‘Wheat’ in Ainu

Remains of wheat were unearthed from the Satsumon Culture sites in Hokkaido, which is estimated around at least the 9th–10th centuries. Barley remains were unearthed from Satsumon and Okhotsk culture sites in Hokkaido, which is estimated around the 8th century. However, since Ainu were basically hunter-gatherers around the 13th century, they have few original word records for wheat and barley. In “Moshiogusa 藻汐草” (1792), the first Ainu-Japanese dictionary, by Uehara Kumajirō, the term メングロ [menguro] was collected as “麦”, which means ‘wheat or barley.’ It corresponds to *menkuru* or *menkur* in the Shizunai dialect; however, the meaning is recorded as ‘broomcorn millet (Panicum miliaceum)’ (Watanabe et al. 1984; Watanabe et al. 1995). Although Batchelor (1938) collected the word *mungi* as the term for ‘wheat,’ it might be ‘barley.’

Hattori (1964) recorded *múnki* and *muŋi* as ‘wheat and barley,’ which were borrowed from Japanese 麦 *mugi*. Kayano (2002) said, “Zensuke Kobayashi, a man of mainland Japan, brought barley grains to the region of Saru river basin in the late 1800s, and was nicknamed *munki caca*, which means ‘uncle barley,’ by the Ainu people.”

(FUKAZAWA Mika)

A. *múnki* type

○ *múnki* ~ *munki* ~ *muŋi*

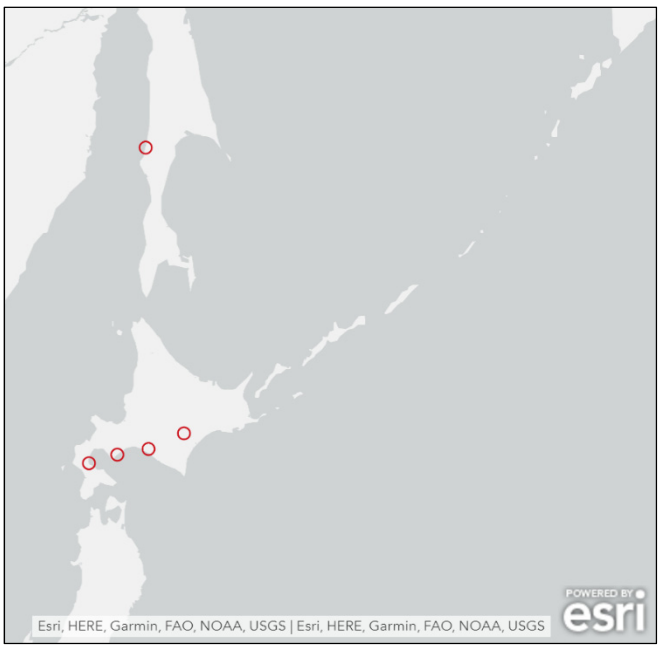


Figure 7.2.1: ‘Wheat’ or ‘barley’ in Ainu.

‘Wheat’ in Japonic

The major types of wheat are KOMUGI (*komugi*, *komɯŋi*, *komogi*, *komoʒi*, *komuɴ*, *komuʔ*, *kwaamugi*, …), which literally means ‘small wheat’, found in mainland Japan, and INAMUGI (*ʔinamugi*, *inamudʒi*, *inamui*, *njaamugi*, *njaamu^{d}ʒi*, *naamugi*, *ɴnamugï*, *ɴnamudzï*, *mnamugï*, *mnamudzï*, *ɴnamuɴ*, *ʔijamui*, …), found in the Ryukyus. MUGI (*mugi*, *mudʒi*, *mugï*, *mudzï*, *muɴ*, *mui*, …), a general term for wheat, barley, oat, and various grains, is also found in Southern Ryukyuan and other areas. MUGI often refers to wheat in Japanese and Ryukyuan. INAMUGI in Ryukyuan, which literally means ‘rice wheat’, might have been derived from the similarity between wheat and rice. While barley is commonly referred to as *oomugi* ‘big *mugi*’, oat is *karasumugi* ‘crow’s *mugi*’. Names for other forms of wheat are KOBAKU (*kobaku*) (BAKU is a Sino-Japanese reading of <麦> MUGI), KONAMUGI (*konamugi*) ‘powder wheat’ (derived from the processing of wheat into flour), and UDONMUGI (*udoɴmugi*) ‘noodle wheat’ or DANGOMUGI (*daɴgomugi*, *dagomugi*) ‘dumpling wheat’, which derive from what the wheat is used for.

The oldest attested words for wheat in Japanese are *komugi* ‘small *mugi*’ and *mamugi* ‘true *mugi*’. The latter word was probably used because it was the most common grain. However, since the KOMUGI type can be found only in mainland Japan, except for some points in the Northern Ryukyus, the Proto-Japonic form would be **mugi*, and it is thought to have been called *ko-mugi* or *ma-mugi* in mainland Japan and **ina-mugi* in the Ryukyus to distinguish it from barley or oat.

Wheat was brought to Japan from China during the Yayoi period. **mugi* is thus considered a loanword from Mainland China, and reflects a form such as */mluək/, which is the older pronunciation than the Sino-Japanese reading /mjaku/ or /baku/. The Korean word /mir/ < Middle Korean /mirh^{H}/ may have had the same origin.

In *Honzō Wamyō* (a Japanese book of medicinal herbs), *karasumogi* <加良須毛岐> is attested, suggesting that the archaic form for *mugi* might be *mogi*. If so, we must compare /mogi/ with other languages.

(NAKAZAWA Kohei and YOKOYAMA Akiko)

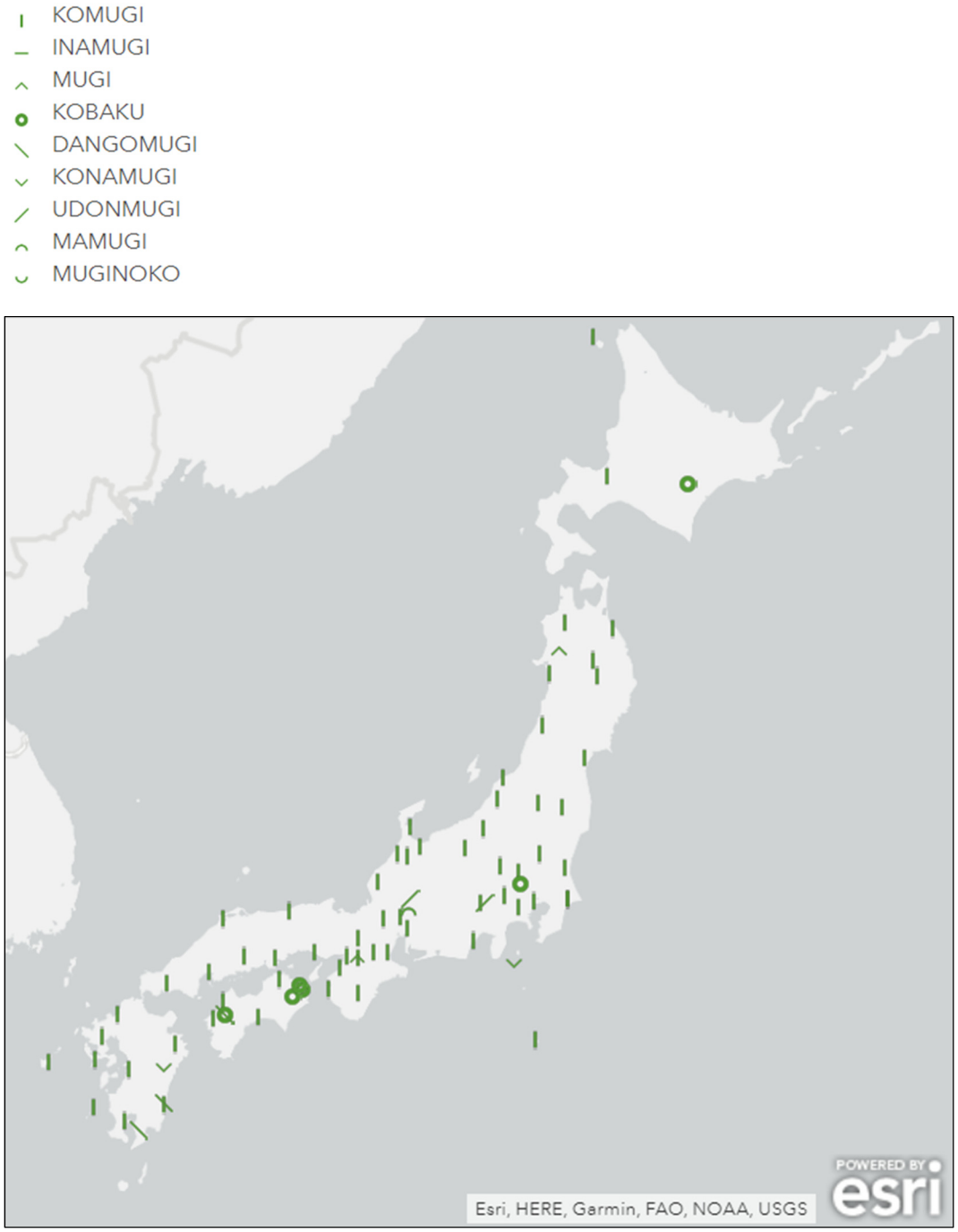


Figure 7.3.1: ‘Wheat’ in mainland Japan.

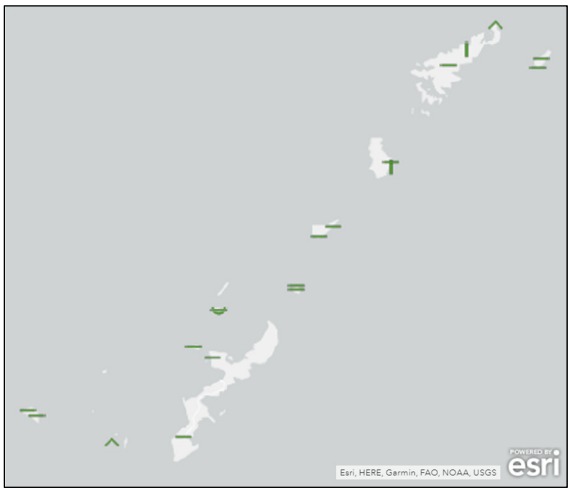


Figure 7.3.2: ‘Wheat’ in Northern Ryukyu Islands.

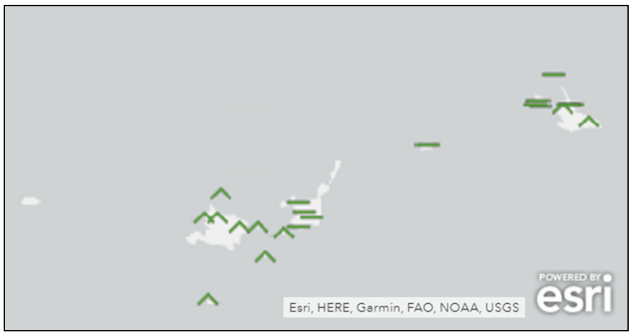


Figure 7.3.3: ‘Wheat’ in Southern Ryukyu Islands.

‘Wheat’ in Korean

The following table shows Modern and Middle Korean terms for ‘wheat’. Middle Korean tone is shown in parentheses.

	Modern Korean	Middle Korean (15-16c.)
wheat (小麦)	mir	mirh (H)

As can be seen in this table, Modern and Middle Korean forms are almost the same except for the fact that Middle Korean form has a word-final *h* which is lost in modern forms.

If we go back the history of this languages, we have older forms recorded in the *Jīlín lèishì* (鶏林類事) written in the 12th century and in *Cháoxiǎnguǎn yìyǔ* (朝鮮館訳語) compiled at the beginning of the 15th century.

	鶏林類事	朝鮮館訳語
wheat (小麦)	麥曰密	麥 冊閔

In *Jīlín lèishì* (鶏林類事), ‘wheat’ is represented by the character ‘密’ which apparently corresponds to MK *mirh*.

In *Cháoxiǎnguǎn yìyǔ* (朝鮮館訳語), ‘wheat’ is recorded as ‘冊閔’, which seems to correspond to a complex word *chʌr-mirh* (lit. sticky wheat).

Generally speaking, earlier forms recorded in these two documents are in accordance with the Middle Korean forms.

Dialectal differences are not so great for this item.

(FUKUI Rei)

mir

Figure 7.4.1: ‘Wheat’ in Korean.

‘Wheat’ in Sinitic

Almost all the word forms for ‘wheat’ share the stem *mai* 麦. We classified them into two categories by word formation type: A stem (+ suffix) type and B modifier type.

A-1 *mai* 麦 (monosyllabic) and A-2 *maizi* 麦子 (stem + suffix) are distributed in the whole area of Sinitic languages. While these forms are also used as the generic terms for the wheat family (e.g., barley, oats, rye, etc.), it usually means ‘wheat.’ The A and B-1 *xiaomai* 小麦 types are likely to collocate especially in the northern areas. The *mai* 麦 (麥) character appeared early in the oracle bones of the Shang dynasty. According to Baxter and Sagart (2014), the Old Chinese (OC) form is *m-rˤək (OC) > meak (Middle Chinese).

B-1 *xiaomai* 小麦 (‘small wheat’) contrasts with “*damai* 大麦” (‘big wheat’), meaning barley. B-1 tends to be distributed in northern and middle China, where people must distinguish between wheat and barley (see Table 1).

Table 1

place		word form
North	Ha’erbin	小麦[ɕiau^{213} mai^{52}]/麦子[mai^{52} tsə^{0}]
	Beijing	小麦[ɕiau^{214} mai^{51}]/麦子[mai^{51} tsɿ^{0}]
	Xi’an	小麦[ɕiɔ^{53} mei^{21}]/麦[mei^{21}]/麦子[mei^{21} tsɿ^{0}]
Middle	Nanjing	小麦[ɕiɔ^{22} mɛʔ^{5}]/麦子[mɛʔ^{5} tsɿ^{0}]
	Suzhou	小麦[siæ^{51} maʔ^{523}]
South	Xiamen	麦[beʔ^{5}]
	Meixian	麦□[mak^{5} ɛ^{31}]
	Guangzhou	麦[mɐk^{2}]

B-2 *ximai* 细麦 is distributed in the Hunan and Jiangxi provinces.

Jiangyong (江永): 细麦[si^{21} mu^{33}]

Wuyuan (婺源): 细麦[si^{35} bo^{51}]

This form is regarded as a dialectal variation of B-1 *xiaomai* 小麦, which reflects the word form for ‘small’ in these areas.

B-3 *hemai* 禾麦 is distributed in the Guangdong and Guangxi provinces.

Xinyi (信宜): 禾麦[wɔ^{11} mak^{22}]

Beihai (北海): 禾麦[wɔ^{21} mɐk^{22}]

The head component *he* 禾 means ‘rice plant’ in these areas. Its word formation may not be the modifier type but the parallel type.

B-4 *mianmai* 面麦 is distributed in the Hunan and Guangxi provinces.

Lianyuan (涟源): 面麦[mi^{11} mɔ^{55}]

Loudi (娄底): 面麦子[mĩ^{11} mɔ^{35} tsɿ^{0}]

Nanning (南宁): 面麦[min^{22} miɐk^{23}]

Mian 面 (麵) (‘wheat flour’) represents the wheat usage.

(SUZUKI Fumiki)

A. stem (+ suffix) type
- A-1 *mai* 麦 (monosyllabic)
- A-2 *maizi* 麦子

B. modifier type
- B-1 *xiaomai* 小麦
- B-2 *ximai* 细麦
- B-3 *hemai* 禾麦
- B-4 *mianmai* 面麦
- B-5 others: 黄麦, 谷麦, 春麦

— C. others: 夏粮, 梭子

Figure 7.5.1: ‘Wheat’ in Sinitic.

‘Wheat’ in Kra-Dai

Wheat is not cultivated in Kra-Dai-speaking areas. It may be that flour only arrived recently, perhaps even in the 20th century. For this reason, most Kra-Dai dialects use Sinitic loanwords for wheat.

In Siamese and Lao, type F, including a sanskrit loanword *saalii*, for ‘wheat’ is used. The first element *khaaw* means ‘rice’. Rice is the main food in this area, hence wheat is treated as a type of rice, and the second element determines its subclassification.

The same type of formation is seen in the southern and western areas, denoted by a red circle in Figure 7.7.1.

Within China’s borders, type A with square symbols is distributed in peripheral areas. This type exhibits with the Sinitic loanword *mak*, meaning ‘wheat’. Type A1, which has a -k ending is distributed in the Guangxi Zhuang Autonomous Region, as its donor Sinitic dialects are Yue, Pinghua, Hakka, etc., which preserve the -k ending.

In Yunnan and Guizhou provinces, type A2, without the -k ending, is distributed, as its donor dialect is Southwestern Mandarin, which dropped it.

Type A5 includes diminutive suffixes such as ʔi^{2} 儿 or tsɯ^{2} 子, etc. Type A6 is a reduplicated form, which can also be treated as a diminutive form.

Type A3 includes ka:ŋ^{3}, which may be a Sinitic loanword 糠 meaning ‘bran’. The meaning of type A4, vaŋ^{33}, is *Setaria italica*. Another word, ɬi:u^{3}, which is included in type A7, also means *Setaria italica*. The meaning of tai$^{35\text{-}11}$ in type A8 is unknown.

Type B, which involves the use of triangle symbols, consists of the Sinitic loanword mi:n 麵, meaning ‘flour’. Type B2 includes ha:ŋ^{24}, which may be 糠, meaning ‘bran’. The distributed area is smaller than that of type A; hence, type B was brought later.

Type C shows a mixture of types A and B; C1 is a portmanteau of types A and B, whereas C2 is a portmanteau of types B and A. As type C’s area of distribution falls between types A and B, this type is a newly emerged example proceeding from contact between types A and B. However, type C2 is found among Sinitic dialects (see the section on ‘wheat’ in Sinitic by F. Suzuki in this volume), it is also possible that type C2 was brought directly from the Sinitic dialect located in the eastern area of Zhuang.

Type D includes the Sinitic loanword su 黍 *Panicum miliaceum*. Moreover, mpaŋ^{31} in type D2 means *Setaria italica*. Type E includes θyt^{9} 术, which should be a Sinitic loanword referring to some type of *Atractylodes ovata*.

The formation of types G–J is difficult to infer, as they are scattered over a small area and their meanings and origins are unknown.

(ENDO Mitsuaki, TOMITA Aika, and HIRANO Ayaka)

A: mak 麦 wheat type

A1: me:k[8], mɐk[8], hau[4] me:k[8], khau[3] mɐk[10], ɣəu[4]mɐk[10]

A2: xɑu[3]mə[3], xau[3]mə[3]

A3: me:k[10] ka:ŋ[3], me:k[10] ka:ŋ[6], me:k[8] ka:ŋ[3]

A4: mə[53]vaŋ[33]

A5: me:k[8] ʔi[2], mə[4]tsɯ[2], me[3]tsɯ[2], mə[31]tsi[44], me[32]tsə[33]

A6: mɐk[8]mɐk[8']

A7: ɬi:u[3] me:k[10]

A8: tai[35-11] mɐk[23]

B: mi:n 麵 flour type

B1: miən[6], mi:n[6],hau[4] mian[6], khau[3] mia[6], hau[4] mi:n[6], gau[4] mi:n[6], mi:n[6] θi:n[5], miən[6] θa:n[1]

B2: ha:ŋ[24]mian[53]

C: mixed type

C1: me:k[8] mi:n[6], me:k[10] mi:n[6], mik[10] mi:n[6], mak[8]mi:n[6], me:k[8] mian[6], me:k[8] mjən[6]

C2: mi:n[6] mai[4], mi:n[6] me:k[8]

D: su 黍 panicum miliaceum

D1: su[55], khau[3] ɕu[1], ɣau[4] ɕiu[1]

D2: su[55]mpaŋ[31]

E: 术 θyt[9] wɔ[4]

F: ข้าวสาลี, ເຂົ້າສາລີ

G: əu[1]mo[6], həu[3]mo[6], hau[4] mo[6] nai[6]

H: doŋ[45], dzɑŋ[31]

I: khɐu[3]tʂha[3]

J: zɿ[31]

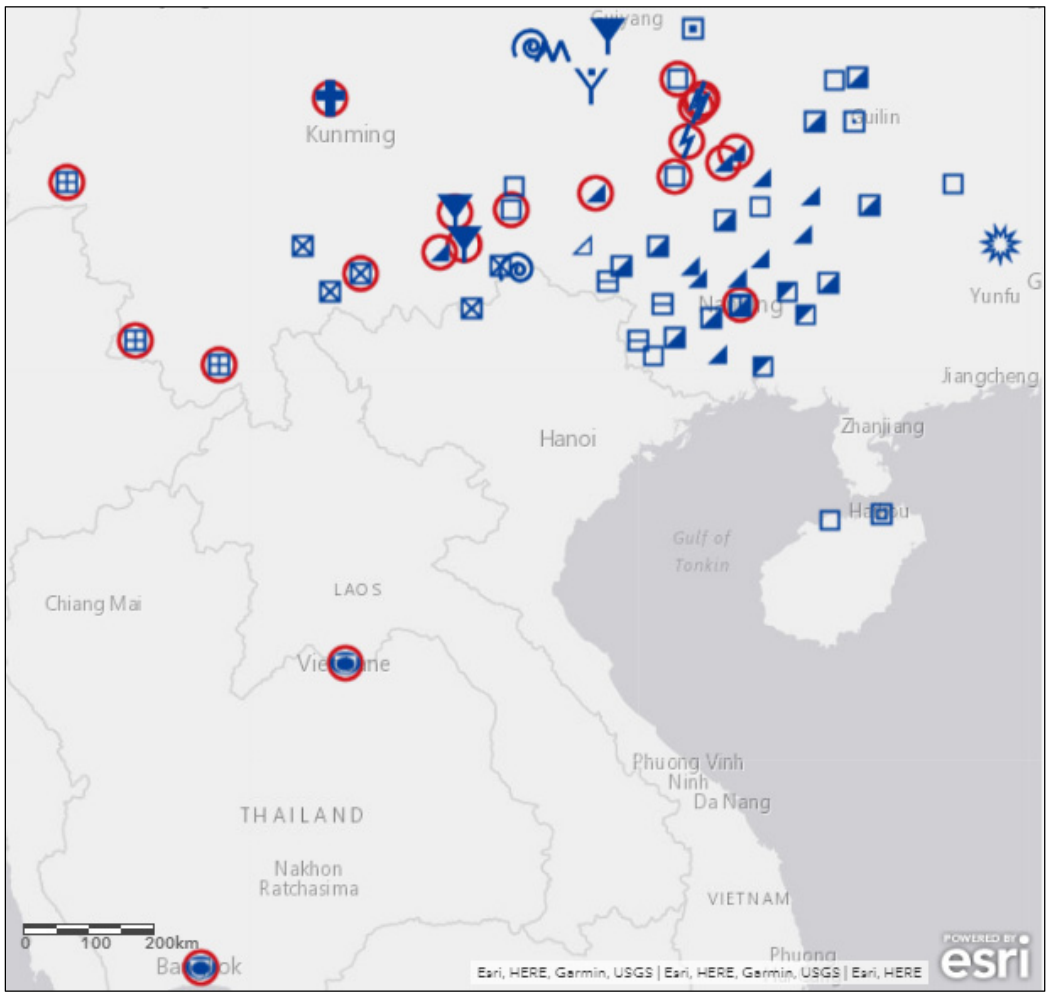


Figure 7.7.1: ‘Wheat’ in Kra-Dai.

‘Wheat’ in Tibeto-Burman

We classify the root and word form for ‘wheat’ in TB languages and varieties into five main types. Types A, B and C are derived from or at least related to proto-language word forms. Type D is a TB root, the roots of which do not originally denote ‘wheat’. Type E includes loans from Sinitic and Indic languages. Type X groups together various independent stems that do not fall under the types above.

Type A includes various forms derived from PTB *m-grwa. Type A1 potentially reflects the most archaic feature, found in only a few Tibetic varieties in the eastern Tibetosphere; Type A2 includes TGTM *grwa + WrT *gro*, found in TGTM languages and most Tibetic languages, as well as Eastern Bodish languages; Type A3 includes a velar voiced initial /g, rg, gl/, found in several rGyalrongic and Tani languages; Type A4 features an initial such as /j/ and /l/, found in many rGyalrongic and Qiangic languages; Type A5 potentially includes a uvular initial /q/, found in several rGyalrongic and Qiangic languages; and Type A6 has a syllable with a /g/-initial + a nasal final, found in Bodo-Garo and surrounding languages. This type is mainly distributed across the whole Tibetosphere, to the north of the Himalayas. Some languages in Nepal and Bangladesh also have Type A, especially A6.

Type B includes various forms with coronal initials. Type B1 includes a /z, dz/-initial; Type B2 includes an /s, ʃ, ʂ, ɕ/-initial, as well as featuring a form with a /x/-initial, which is provisionally classified in this group due to its typological relationship with /ɕ/. This type is mainly found in the Lolo-Burmese languages of Yunnan and surrounding areas. Type B1 appears to be derived from Burmic *zu^3, whereas, according to the analysis of Bradley (2011), Type B2 is derived from Burmic *$ša^3$. Hence, there is another potential approach to allow Type B1 to be entirely distinguished from Type B2.

Type C includes various forms with a /m/-initial. Type C1 only has one syllable, with a /m/-initial only; Type C2 is a compound including a syllable with a /m/-initial. This type, provisionally analysed as being related to Sinitic *mai*, is mainly found in Bai and Tujia, which show a strong influence from the Sinitic languages. Type C is distinct from Type E (E1), which is a loanword type.

Type D includes word forms that underwent a semantic change. Type D1 is forms derived from WrT *nas* ‘highland barley’; Type D2 is forms derived from WrT *so ba* ‘barley’; Type D3 is forms derived from WrT *sngo* ‘green, blue’; and Type D4 are forms derived from WrT *dkar po* ‘white’. All of the subtypes are found in Tibetic languages. Type D1 appears at the northern and southern edges of the eastern Tibetosphere; Type D2, is only found in Yunnan; Type D3 appears at two points (rNgawa and Ragwo) from Amdo and Khams, respectively. Type D4 is found in Rongbrag, rGyalthang and Dzongkha.

Type E groups loanwords together. Type E1 uses a Sinitic loan, such as *maizi* and *xiaomai*; Type E2 uses an Indic word, derived from Sanskrit *bhūjambū*.

Obviously, Type E1 appears more commonly in present-day China, particularly in the south-western corner of Yunnan. Type E2 appears in Myanmar. The Rawang form is a loan from Burmese, and that word derives from Indic for ‘wheat’.

Type X contains a range of various independent stems, found in languages spoken in Myanmar, India and Nepal, such as Pa-O, Thulung, Newari, Kulung and Jinghpaw.

Bradley (2011:140) suggests three TB etyma for ‘wheat’: *ša^{3}, WrT *gro* and *gom/goŋ; however, our analysis finds only two main types (A and B), which correspond to the last two in Bradley’s count, and *ša^{3}, respectively. The distribution of these two forms is clear; the south-eastern area of the TB languages (Type B) and the remaining area (Type A).

In addition to the overall distribution, Type C is mainly attested within the territory of Type B. We relate Type C to the Sinitc word *mai*, but Bradley (2011:136) states that the Sinitic form *mai* ‘appears to be a loan from Indo-European, possibly via Turkic’. Hence, if our analysis is correct, the varieties deriving from Type C received the Sinitic lexical form after the word was borrowed into Sinitic borrowing.

In the Tibetosphere, other crops (barley, rye and highland barley) than wheat are widely cultivated. The distribution of Type A2 is spread widely, and its form corresponds to Literary Tibetan, but in the areas where the wheat grows instead of the highland barley (Yunnan), the word form for ‘wheat’ is substituted for with other crops’ terms, as shown in Type D. It appears in Tibetic languages, particularly in varieties spoken at lower altitude. Type D1 is, in most cases found in riverside areas of Yunnan. Speakers rarely encounter the ‘highland barley’ species known as *qingke* in Chinese. It is highly plausible and persuasive that the form corresponding to WrT *nas* ‘highland barley’ experienced a semantic change into ‘wheat’. Type D3 is related to the colour of wheat as a plant, and Type D4 is related to the colour of the flour. In general, highland barley and rye is more brown than the wheat is, although Tibetans often classify highland barley into types based on its colour to specify the taste and origin of *rtsam pa* (Tournadre & Suzuki 2022).

Type E1 is mainly found in varieties of Lahu, and Type E2 is found in Burmese. This is probably related to the late introduction of wheat cultivation, due to the inhospitable climate to wheat.

To conclude, Types A and B each have a distribution territory, and their chronological order of dissemination is unlikely to be determined. Compared with these types, the others are all newcomers. Observing the distribution of Types D and E, we find a tendency for non-proto forms to be used in regions where wheat is not familiar. Additionally, Type D suggests that ancient Tibetic languages have already had many lexically diverse crop terms regarding wheat and barley species, some of which have experienced semantic changes.

(SUZUKI Hiroyuki, EBIHARA Shiho, IWASA Kazue, KURABE Keita, SHIRAI Satoko)

A. PTB *m-grwa
- ■ A1: *^{n}ɖo, ^{n}ɖu*, etc.
- ○ A2: *ʈo, kro, co, tɕo, cço, tʂ^{h}o^{12}*, etc.
- ⊕ A3: *rgak, rgo^{ʕ}, lgap, ka lie*, etc.
- ● A4: *lɯ, ^{l}zi, ji, jɔ*, etc.
- ◐ A5: *kaj, qø^{24}, ʁuə^{241}*, etc.
- ◒ A6: *gom, goŋ*, etc.

B. Coronal initial type
- ▯ B1: *zu^{33}, zɿ^{33} zɛ^{31}, ^{n}dza, tso:dza*, etc.
- ▮ B2: *sa^{ʕ}, ʃua^{44}, ʂa^{44}, xa^{33}*, etc.

C. /m/-initial type
- ☆ C1: *mɯ*, etc.
- C2: *se mɯ*, etc.

D. Semantic change type
- ▽ D1: *nẽ:, nje:, nɛ:*, etc.
- △ D2: *s^{h}o wa*
- ⁄ D3: *^{ɣ}ŋ̊o, ŋ̊o*
- ＼ D4: *kə rə, ka ra, kâ, kar*

E. Loanword type
- E1: Sinitic; *maizi, xiaomai, mɯ tsi*
- E2: Indic; *jǒu, jouɴ^{M}, zong*
- ▬ X. miscellaneous

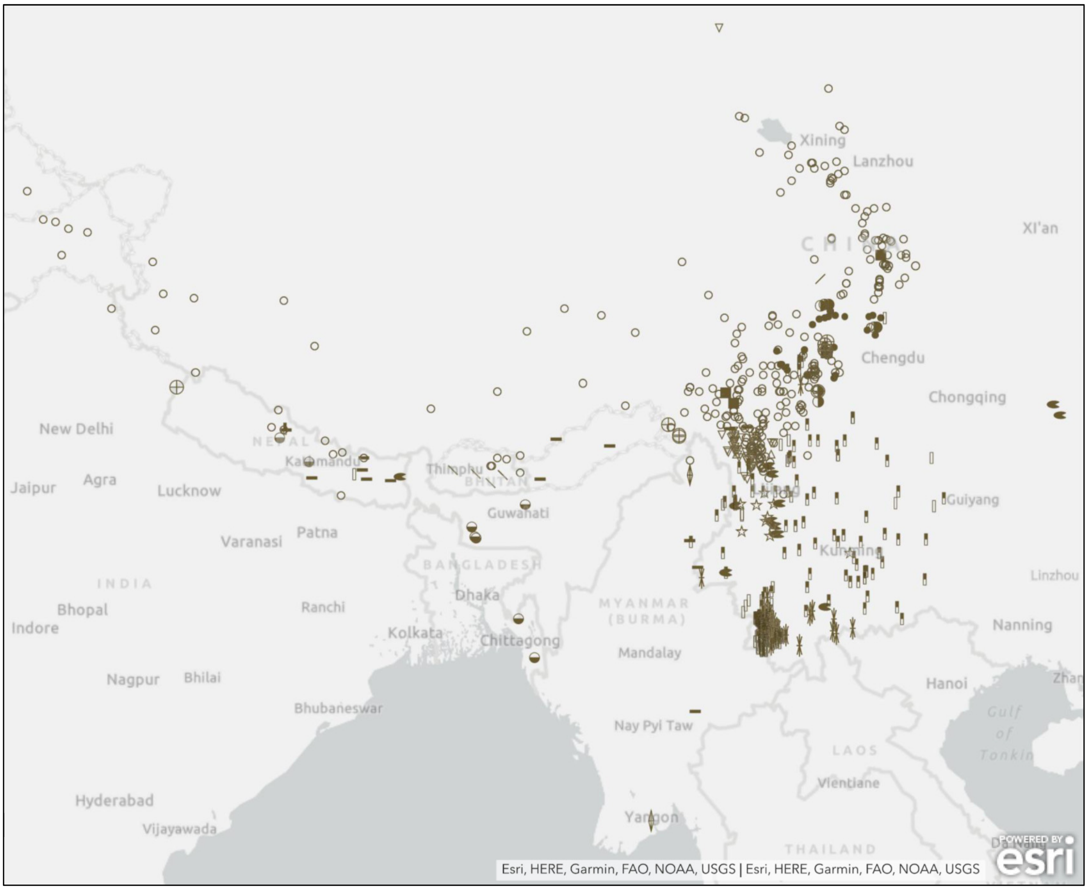


Figure 7.8.1: ‘Wheat’ in Tibeto-Burman.

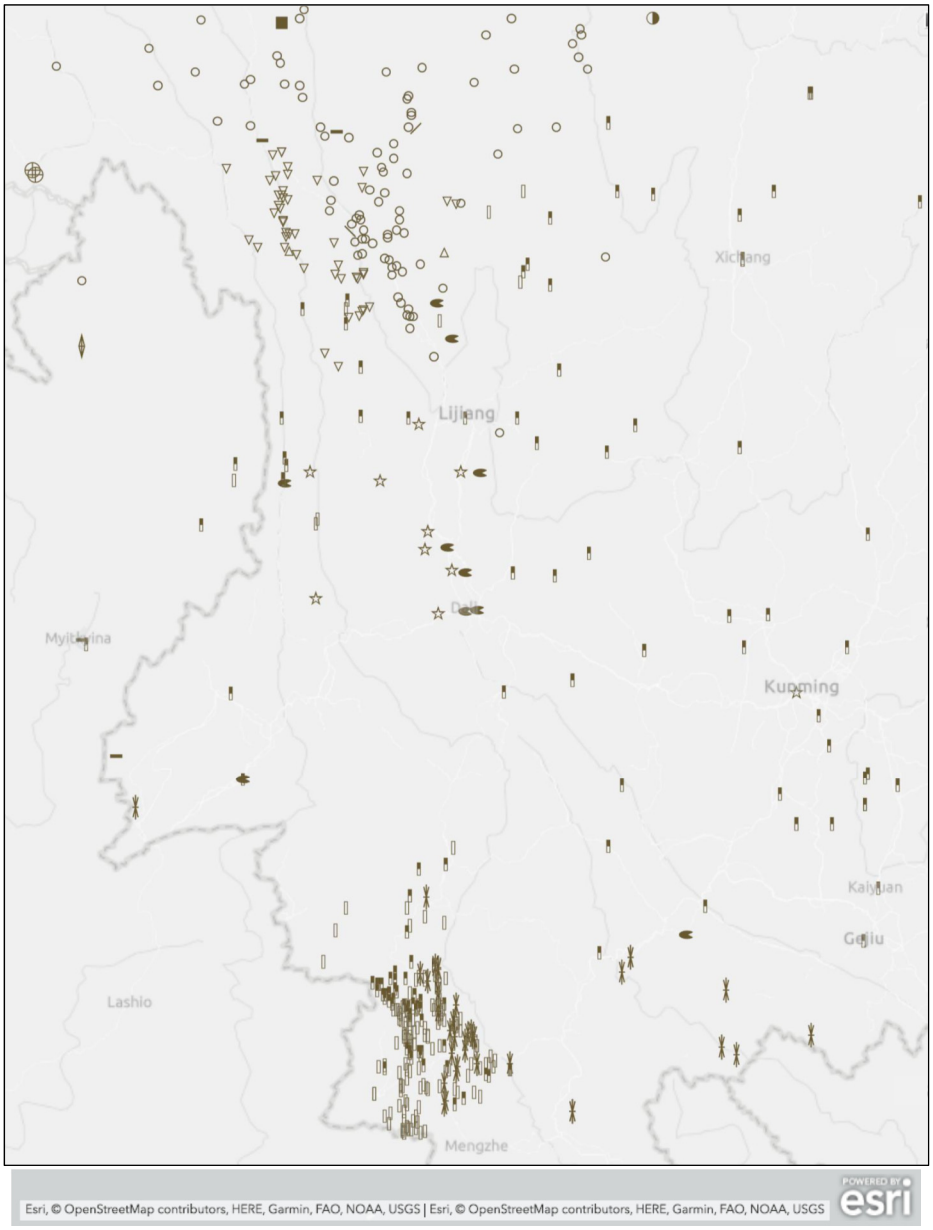


Figure 7.8.2: ‘Wheat’ in Tibeto-Burman (enlarged version).

‘Wheat’ in Austroasiatic

Wheat was not cultivated until recently, as rice is the staple food in most of the Austroasiatic language areas.

Etymons for ‘wheat’ are classified into the following four types, A through D.

Of these, type C in Khmer and D in Vietnamese in mainland Southeast Asia are probably of recent origin after the introduction of wheat to the area.

Type A: (*gahũ, gohom, gɔhɔm, gandɔm)*

These forms are derived from a Persian loan via Hindi. They are distributed in the Indian subcontinent, and the Malay Peninsula.

Type B: (*kew, kɛ-huːn)*

kew in Khasi is found in Northeast India, and *kɛ-huːn* is found in Car Nicobarese, in the Andaman Sea. They are considered to be devoiced variants of type A.

Type C: (*saaləy)*

saaləy in Khmer is a loan from Sanskrit *śāliḥ* “rice”.

Type D: (*lúa mì, tro mì)*

lúa mì in Vietnamese is a compound made of *lúa* “rice plant” and *mì* “noodle.” *tro mì* in Pacoh is probably a compound made of an unknown plant name and *mì* “noodle,” a loan from adjacent Vietnamese.

Etymons for ‘wheat’ are not found in the rest of the Mon-Khmer languages.

(MINEGISHI Makoto, and SHIMIZU Masaaki)

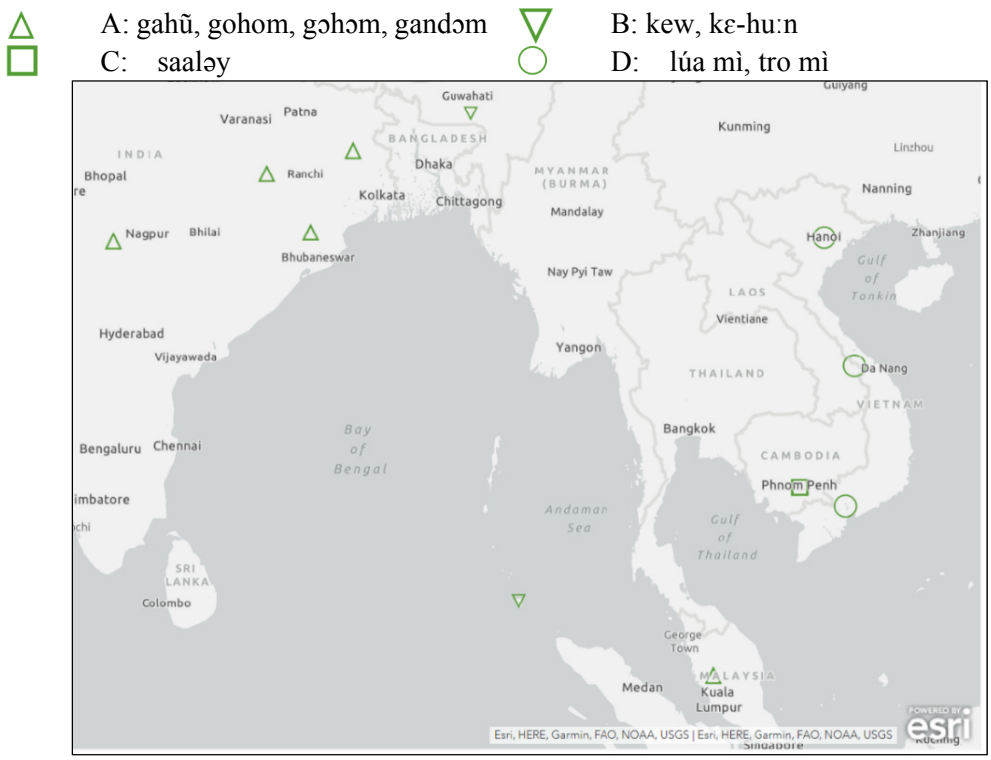


Figure 7.9.1: ‘Wheat’ in Austroasiatic.

‘Wheat’ in Austronesian

Wheat has not been cultivated in the Austronesian world for very long; its recency explains the absence of original terms for this crop. The Austronesian languages encompassing words for “wheat” have generally adopted a loan form from varied source languages. The data utilized for this study revealed that the most frequently found loan forms included the Persian *gandum*, the Spanish/Portuguese *trigo*, and the English “wheat.”

A TRIGO: Loan forms from *trigo*, originally a Spanish or Portuguese term, spread in the Philippines and some parts of Sulawesi, such as Tagalog /*trīgo*/, Aklanon /*trīgoh*/, Kagayanen and Gorontalo /*trigu*/, Bangingi Sama /*tidigu*/, and Dobel /*tarigu*/.
B GANDUM: This loan term from Persian is found in some parts of Indonesia. The form /*gandom*/ is used in Acehnese, Minangkabau, Indonesian, Sundanese, Javanese, and Sawai. The word becomes /*ga^{n}do*/ in Da’a and /*gandoŋ*/ in Bugis.
C WHEAT: Numerous languages in the Pacific islands employ the English loan form. The borrowed terms include /*uit*/ (Manam), /*witi*/ (Kilivia, Eastern Fijian), /*wit*/ in Lewo, /*uiti*/ in Rotuman, and /*uite*/ in Tongan.
D: A French loan word is found in A'jië as /*bele*/.
E: Other forms borrowed from unknown origins include /*mogi*/ (Yami), / *aturɔy*/ (Palawan), /*vari-m-bazaha*/ (Malagasy Merina), *batara kosoaŋ* (Konjo), and /*sana*/ (Samoan). The Tahitian /*titona*/, Rotuman /*saitō*/, and Samoan /*saito*/ could have originated in the Greek word /*sītos*/. All these terms are categorized as “other forms” in this paper.

Type A forms, supposedly adopted from Portuguese or Spanish, are distributed in the Philippines and some parts of Sulawesi. Indonesian and Sundanese also respectively use the form /*tərigu*/ and /*tarigu*/, in addition to the Persian loan form /*gandum*/. Buru employs the form /*trigu*/ in addition to /*tapong*/, the word that also denotes “powder.”

Type B comprises loan forms from Persian, which are found in some areas of Indonesia, such as Maluku, Sulawesi, Sumatra, and Java.

Type C and type D respectively signify loan forms from English and French and are found in the Pacific islands.

Other forms do not evince specific areas of distribution.

(UTSUMI Atsuko)

A: *trīgo, trīgoh, trigu, tidigu, tarigu*

B: *gandom, ga^{n}do, gandoŋ*

C: *uit, witi, wit, uiti, uite*

D: *bele*

E: mogi, aturɔy, vari-m-bazaha, batara kosoaŋ, sana, titona, saitō, saito

Type A and Type B:

Type A and Tepung ‘powder’

Type C and Others

Figure 7.10.1: ‘Wheat’ in Austronesian languages in Taiwan and Philippines.

Figure 7.10.2: ‘Wheat’ in Austronesian languages in Indonesia.

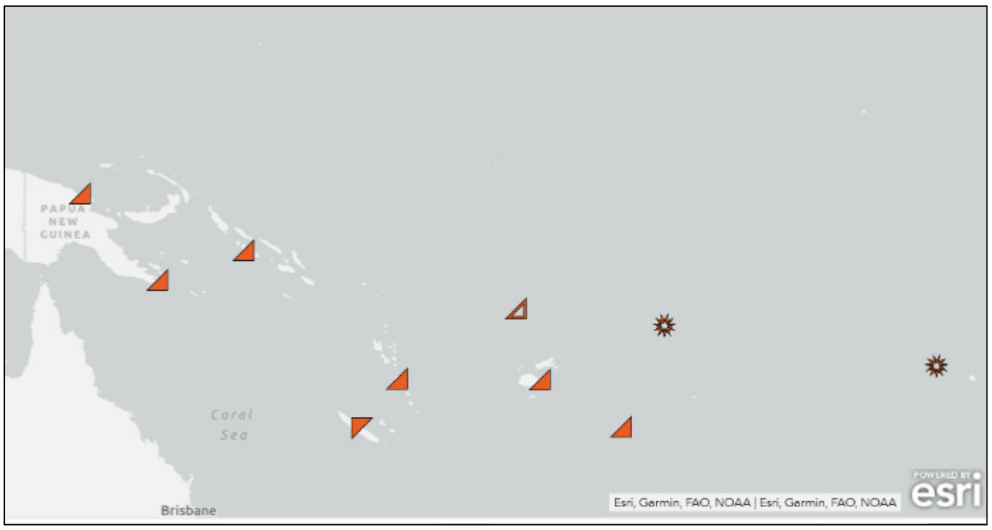


Figure 7.10.3: ‘Wheat’ in Austronesian languages in Papua and Pacific Islands.

‘Wheat’ in Tungusic

The word for ‘麦子*màizi*’, which includes wheat 小麦, barley, oats, rye, is found, but they are all Chinese Tungusic.

Orochon	*majsa*
Ewenke (Solon dialect)	*pulimpu*
Ewenke (Tungus dialect)	*majs*
Ewenke (Yakut diakect)	*owees*
Hezhe	*maisə*
Sibe	*mais*

The form *majs, majsa, maisə* must have been borrowed from Chinese, *owees* may be from Russian *овёс* ‘oat’. The another form *pulimpu* might be also borrowed from other languages, but it is unknown.

(MATSUMOTO Ryo)

Type A MAIS
Type B OWEES
Type C PULIMPU

Figure 7.11.1: ‘Wheat’ in Tungusic.

‘Wheat’ in Uralic

Uralic languages have several forms for the word ‘wheat’, which seem phonologically unrelated between each other. In the northeastern area where Samoyed and Ugric peoples live, there is no agricultural culture. Therefor they have no word for ‘wheat’ expect the loanword from Russian for the selling products.

The forms are classified into 5 types as followings:

Type A *NIS-*
Type B *VEHNA*
Type C *TO-*
Type D *Š-/Č-*
Type E *BUZA*

To each type belong languages as below:

Type A *NIS-*: Saami, Estonian *nisu,* Karelian, Veps *nižu,* Livonian *nizzõz*

Type B *VEHNA*: Finnish, Ingrian *vehnä*

Type C *TO-*: Moksha *tozer,* Erzya *tovzjuro*

Type D *Š-/Č-*: Mari *šydaŋ,* Udmurt *čabej,* Komi *šobdi*

Type E *BUZA:* Hungarian *búza*

It is not clear for me why such different forms are observed, but they are geographically distributed in appropriate areas, not randomly.

(MATSUMOTO Ryo)

Type A *NIS-*
Type B *VEHNA*
Type C *TO-*
Type D *Š-/Č-*
Type E *BUZA*

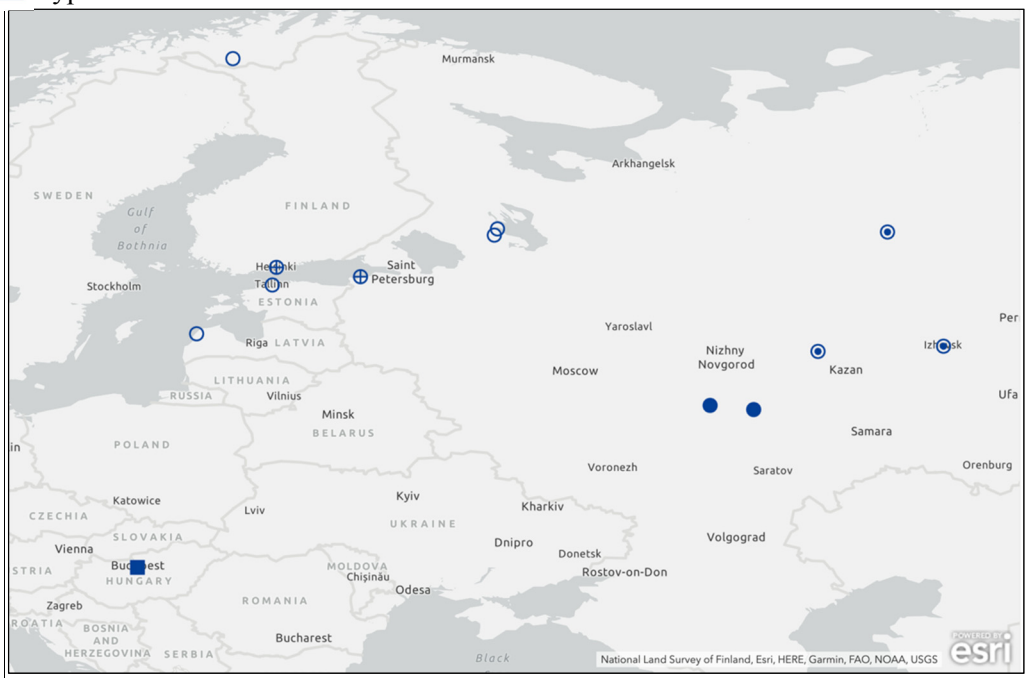


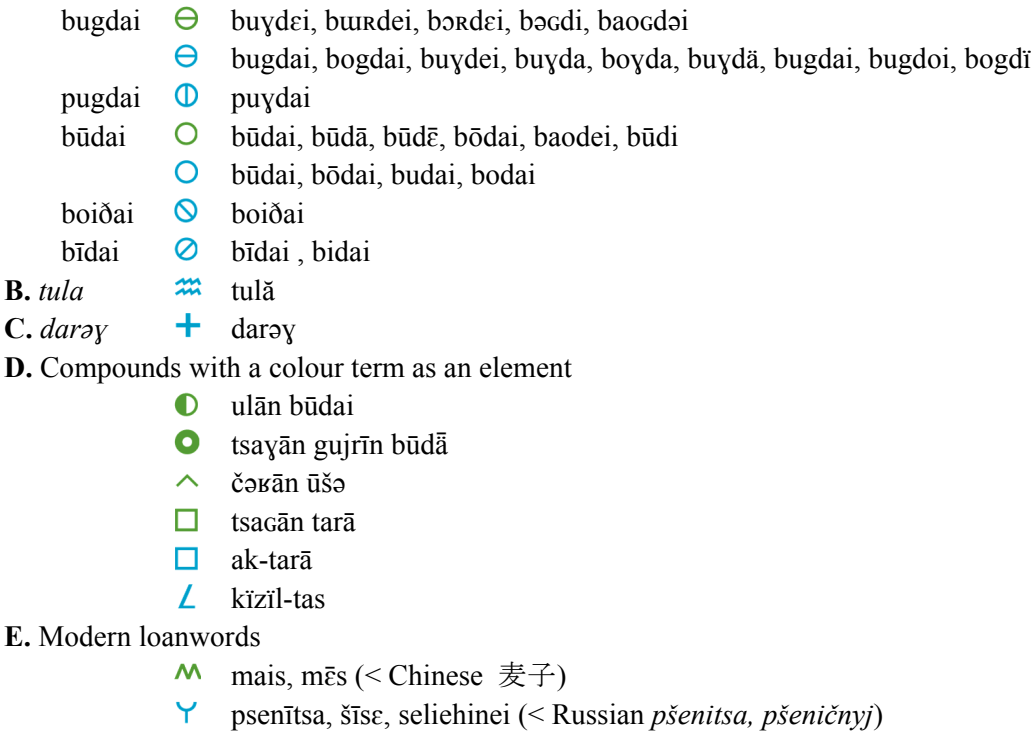
Figure 7.12.1: ‘Wheat’ in Uralic.

‘Wheat’ in Mongolic and Turkic

Mongolic and Turkic people in steppe and taiga do not cultivate the land, but even they know wheat. Most Mongolic and Turkic languages use *bugdai*-type words for ‘wheat.’ Other forms, except modern loanwords from Chinese and Russian, include the following: *čəʁān ūšə* in Shira Yughur (*čəʁān* ‘white,’ *ūšə* ‘food’), *darəɣ* in Sarïg Yughur, *kïzïl-tas* and *ak-tarā* in Tuvan (*kïzïl* ‘red,’ *ak* ‘white’), *tulă* in Chuvash (cf. Old Turkic *talkān* ‘crushed parched grain’), etc. The forms *darəɣ* and *tarā* are both from Old Turkic *tarïg* ‘grain.’

The forms other than *bugdai*-type words are found only in peripheral regions.

(SAITÔ Yoshio)



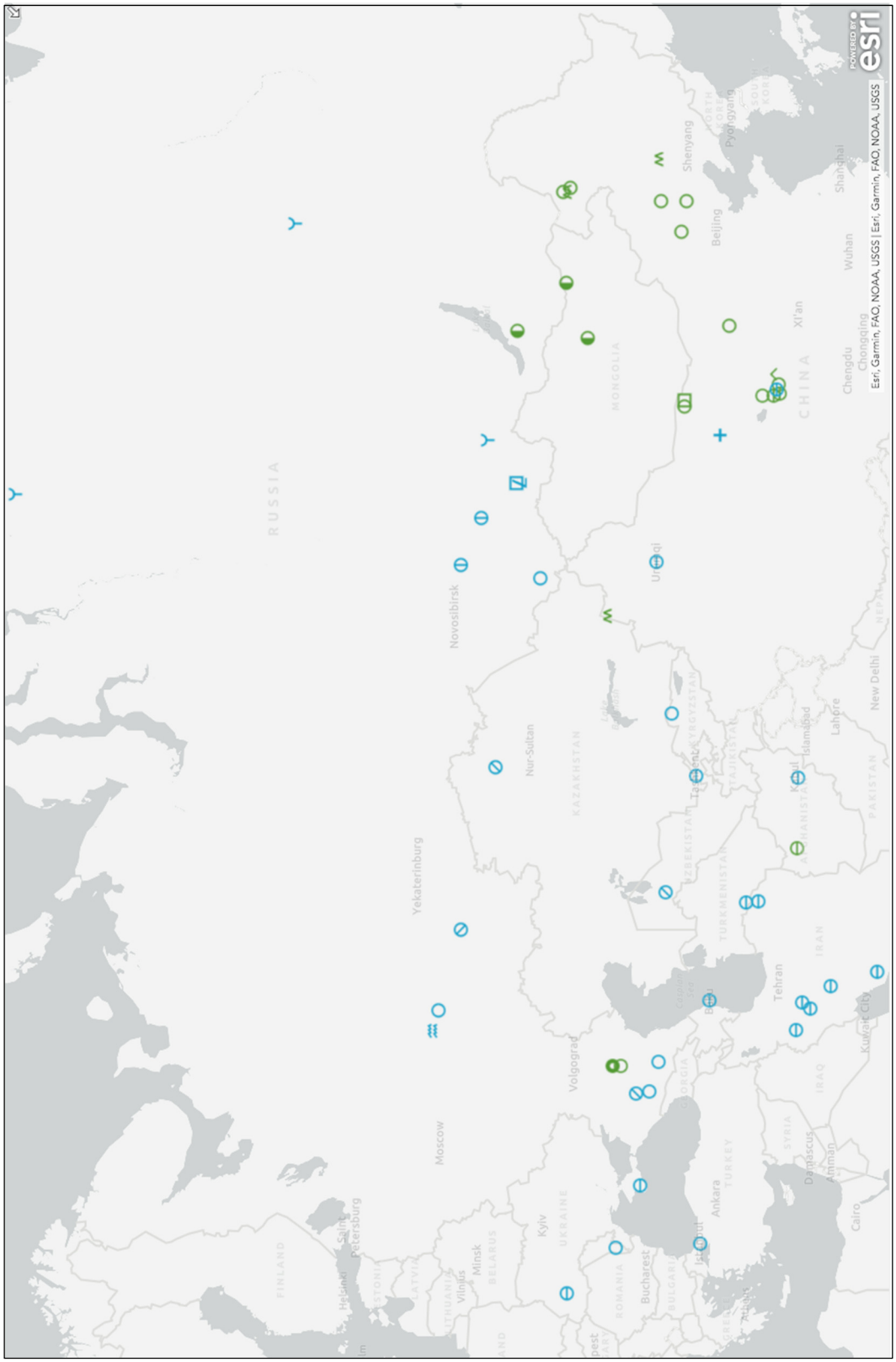


Figure 7.13.1: ‘Wheat’ in Mongolic and Turkic.

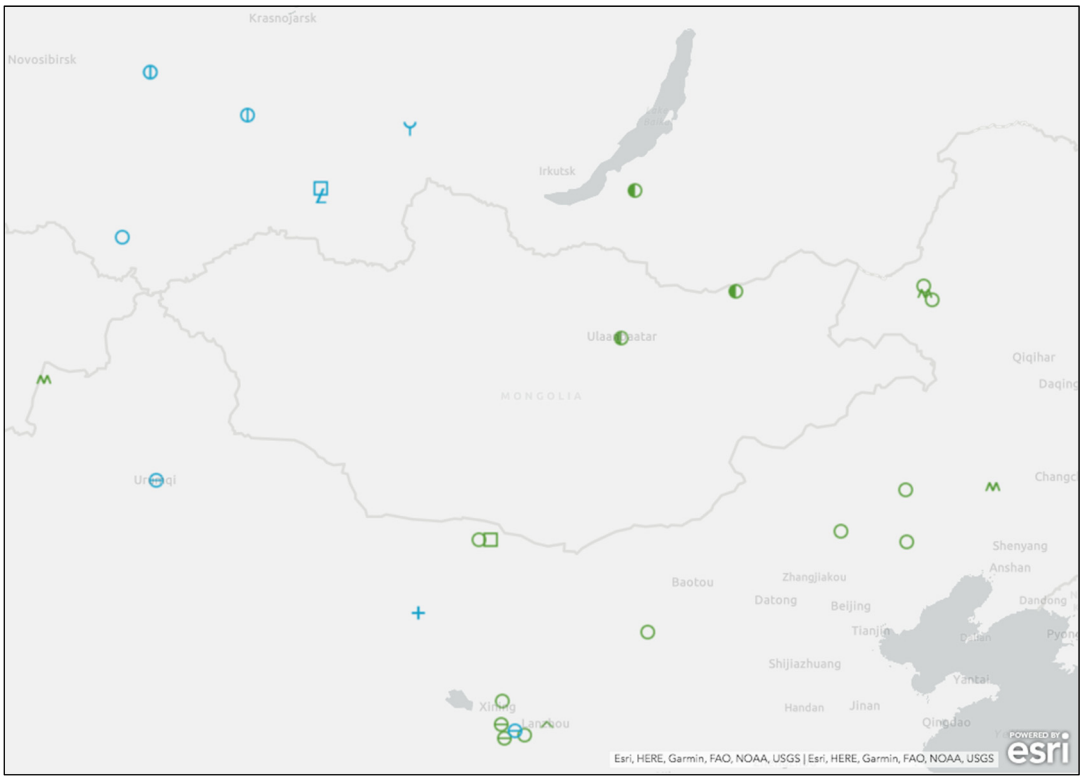


Figure 7.13.2: ‘Wheat’ in Mongolic and Turkic. (The Mongolian Plateau and its vicinities magnified)

‘Wheat’ in South Asia

I describe the languages of Indo-Aryan (IA), Nuristani, and language isolates in South Asia.

There are two major types of word forms – A) *gōdhūma*, B) *kaṇika* – and a minor one, C) *gur*.

The distribution of ‘wheat’ words is mainly the type A overwhelmingly from the Indian subcontinent to Europe, with a little the type B in the northwestern part. The type C is only for Burushaski.

The most major type is *gōdhūma*. This type is derived from Sanskrit *gōdhū́ma* गोधूम ‘wheat’, derived from Proto-Indo-Iranian **gandʰúHmas* ~ **gantúHmas*. Forms of this type are used by Indo-Aryan and Nuristani languages in South Asia. Most forms are derived via Sanskrit, while Urdu *gandum* گندم is a loanword from Persian. Whereas Hindi *gēhū̃* inherits the Sanskrit term. Their choice of the words, is the product of Islamisation, and Sanskritasation. In some Romani languages the form has changed into *giv*, and furthermore *djiv*. And then Selice Romani has confused the form with the different word *jiv* ‘snow’, which is derived from Sanskrit *híma* हिम, ultimately from PIE **ǵhimós* ‘cold, frost’. So they now calls ‘wheat’ as *šužo jiv*, which literally means ‘clean snow’, to distinguish between each other. Dhivehi *godan* and Vaagri Boli *ghawnādānā* are accompanied by the morphemes *-dan* and *-dānā* derived from Sanskrit *dhānya* धान्य ‘grain, corn, rice’. As for *dhānya*, Zargari Romani in Iran *dāhān* ‘wheat’ is a simple descendant of it. (See also ‘Foxtail millet in South Asia’.)

The *kaṇika* type appear only in IA languages. They are found in the northwestern part of the subcontinent, and exactly the languages belong to the northwest group of the IA branch. The origin of the type is the Sanskrit word *kaṇika* कणिक ‘grain’. Many languages in this group retain the retroflex nasal consonant /ɳ/ (ṇ) unlike to most modern IA languages such as Hindi and Bengali; now the sound can be seen in these offspring forms.

The *gur* type appears only in Burushaski languages. Berger (1998: 161) suggests that this word might be derived from Tibetan *gro* སྒྲོ ‘wheat’, but it is not well attested.

For ‘wheat’ Sindhi employs the form *burra* ﺑﺮ, its etymon is possibly **birr* ~ **bar* ‘wheat’ for the South Arabian languages of the Semitic family (see further Nagato’s chapter in this volume). In Panjabi they employ the word *darō* ਦਰੋ, it may be from Sanskrit *dhū́rvā* धूर्वा ‘*Cynodon dactyon* (bermuda grass, ギョウギシバ)’. There are two other forms – Kusunda *khərwi* and Sinhala *tiriṇgu* තිරිඟු – on Figure 7.14.1, but the etyma are unclear for me.

(YOSHIOKA Noboru)

A. *gōdhūma* type (74) 0

gōdi, gandum, gõ̆m, gōom, gõm, gõ̤m, gɔm, gúum, gũ̆m, guōmə, gam, gamo, gomú, gũ̆, gōhũ̆, gēhũ̆, gahū, gahũ̆, gé́sũ̆, gišu, geúñ, gewũ, geu, gíũ, giũh, gĩ̃ũ̃, giv, gaũ, gãv, gavaṇi, guy, ghóom, ghẽ̃hu, ghə̃w, ghaũ, giñyu, gyũ̆, ɠẽ̃hūn, djiv, ùtilde

[+*dhānya* (2)]

godan, ghawnādānā

[+X]

šužo jiv

B. *kaṇika* type (12)

kaṇika, kaṇak, kaṇk, kəṇk, kiniḳh, qaṇaq

C. *gur* type (3) ^

gur

F. others

[*dhānya*] dāhān; burra; darō; khərwi; tiriŋgu

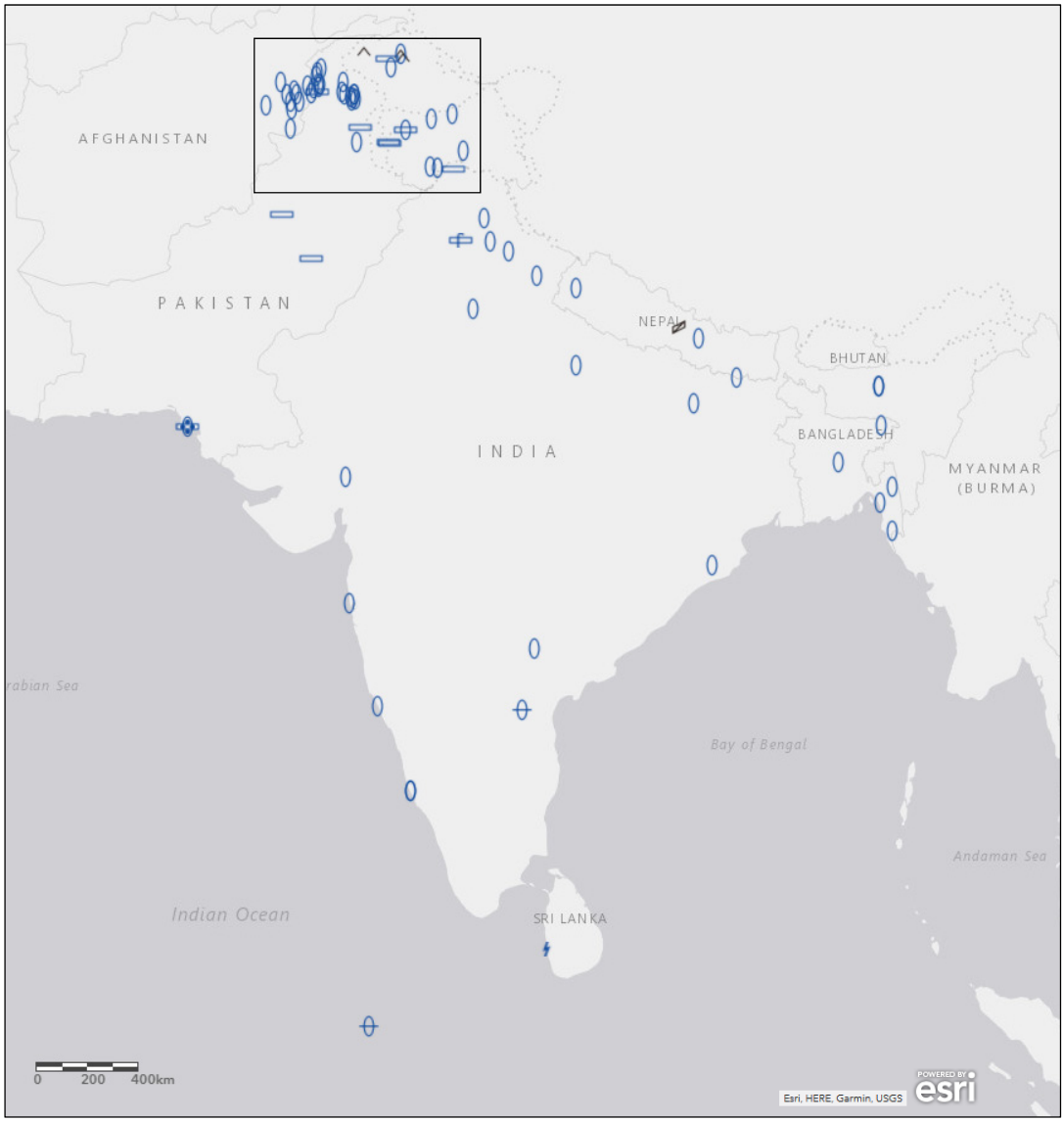


Figure 7.14.1: ‘Wheat’ in SA: Indo-Aryan, Nuristani (both in navy blue), and language isolates (those in black).

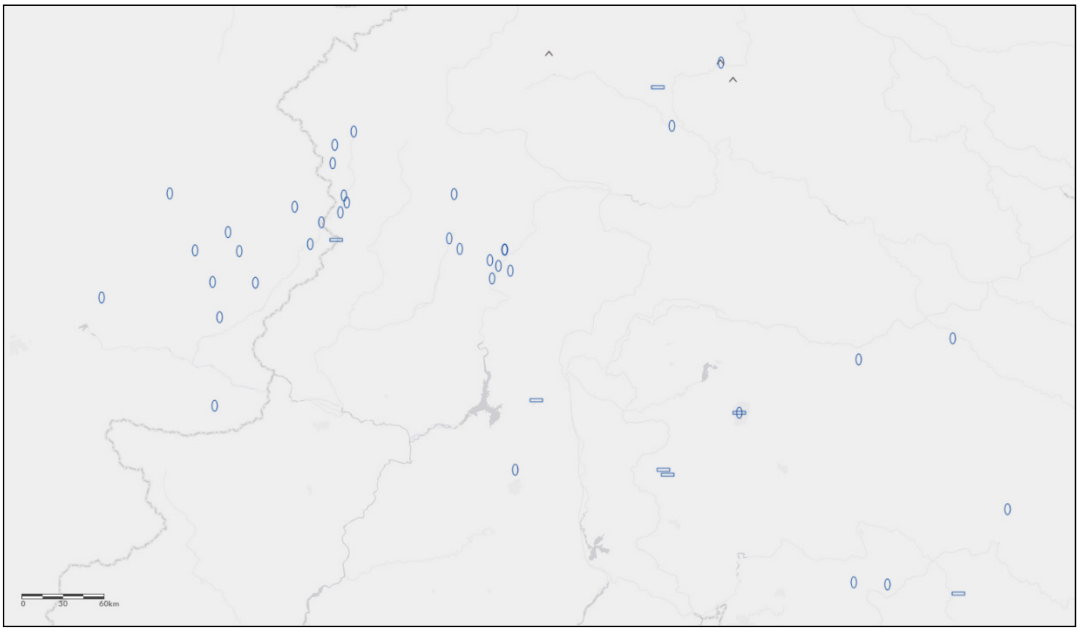


Figure 7.14.2: Types for ‘wheat’ in the northern part (the area encloed by the rectangle in Figure 7.14.1).

Figure 7.14.3: Types for ‘wheat’ in Indo-Aryan languages outside South Asia.

‘Wheat’ in Dravidian

No Dravidian etymon is found for ‘wheat’. All languages employ a loanword of an Indo-Aryan source. Absence of this etymon for the crop which must have been known to the Harappans but is not attested in the contemporary Southern Neolithic sites in the Deccan indicates geographic diffusion of Dravidian speakers on the subcontinent antedated the introduction of wheat.

Words with the intervocalic /*dh*/ retained in Peninsular India are borrowed from *gōdhūma* in Sanskrit, the common literary language in pre-modern India. Tamil *kōtumai* is an orthographic modification for transcription of Sanskrit voiced consonants into Tamil script.

The Brahui form *xolum* is from an unknown Indo-Aryan source. The initial *x*- indicates it is an early borrowing preceding the characteristic sound change PDr, *k > x.

Kurukh *gohom* and Malto *gohme* are borrowing from New Indo-Aryan languages in their present location.

(KODAMA Nozomi)

gōdhi
kōtumai, kōtampu, gōdhuma
xōlum
gohom, gohme

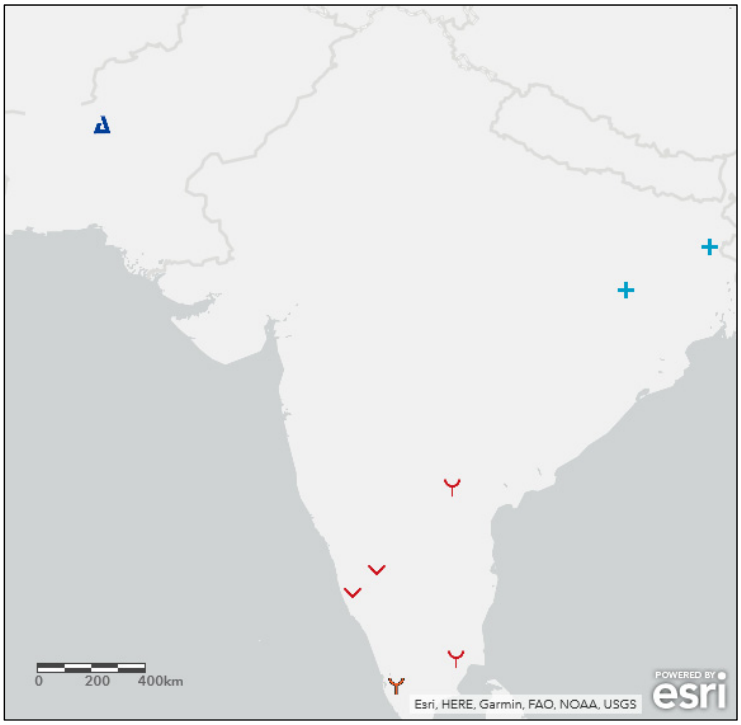


Figure 7.15.1: ‘Wheat’ in Dravidian.

‘Wheat’ in Iranian

Wheat is one of the most important crops in the region where Iranian languages are spoken. There are three major types of the words for ‘wheat’ in Iranian languages.

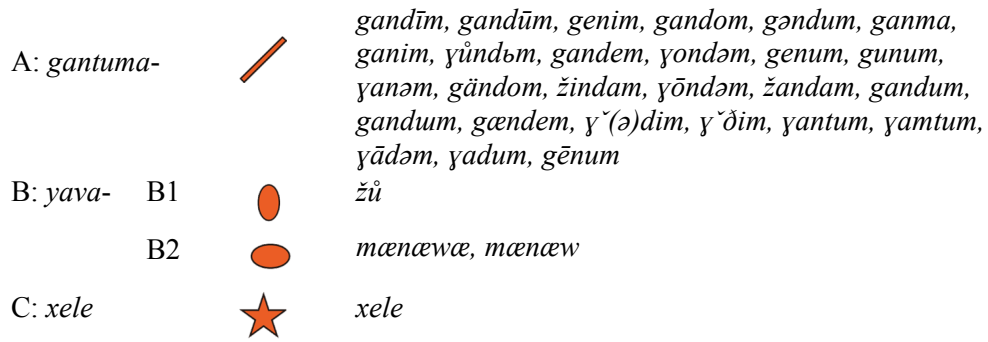
Almost all Iranian languages fall within Type A, which is inherited from the Proto-Iranian forms **gantu(H)ma-* or **gandu(H)ma-* (cf. Avestan *gantuma-*). They and their Indo-Aryan cognates such as Sanskrit *gōdhūma-* show no regular sound correspondences, thus the common Indo-Iranian proto-form cannot be reconstructed by these examples (see also Rastorgueva and Ėdel’man 2007: 150).

Type B, derived from Old Iranian **yava-* ‘barley’, is either a simplex (B1) or a compound (B2). B1 is found in Yazgulami spoken in a remote area of Tajikistan. This language also has a term *ɣʷont* ‘roasted and dried wheat (flour)’, apparently inherited from Type A **gantuma-* with a slight semantic change.

B2 is distributed in Ossetic (Iron and Digoron) spoken in the Caucasus. According to Abaev (1973: 92), the Ossetic words for ‘wheat’ *mænæw(æ)* consist of *mæ(n)* ‘my’ and *jæw* ‘millet’. The latter historically changed its referent from ‘barley’, ‘grain’, then to ‘millet’.

Type C *xele* is observed in Zazaki spoken in Eastern Turkey. Its etymology possibly dates back ultimately to Arabic غلة (*ɣallat*) ‘grain’. Central Kurdish has a cognate word *xele* ‘corn, maize’ for this.

(IWASAKI Takamasa)



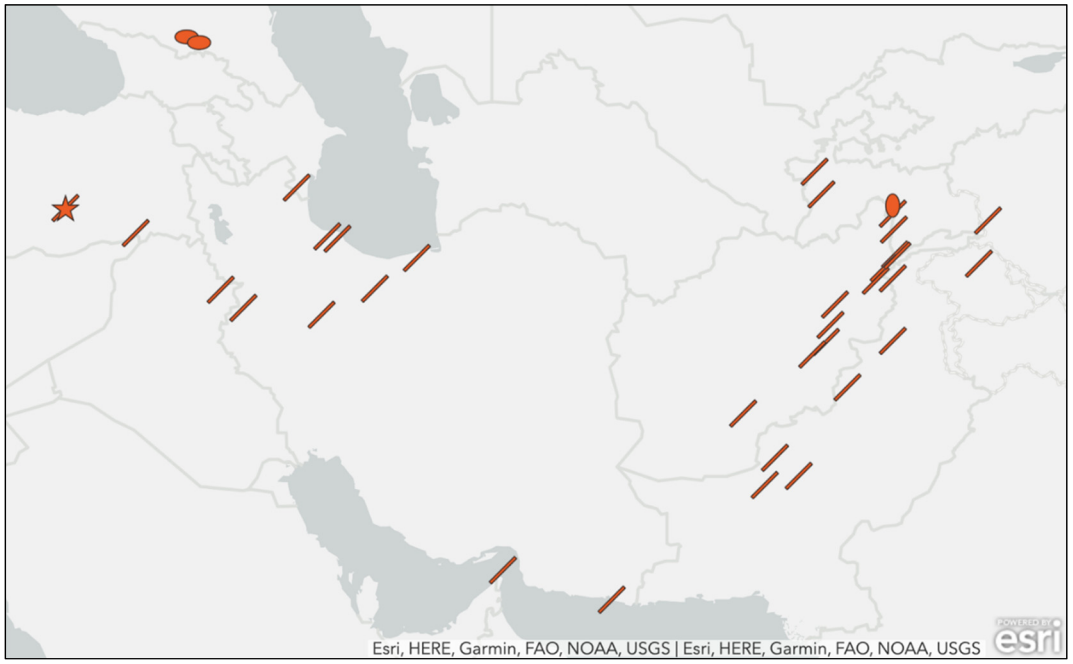


Figure 7.16.1: ‘Wheat’ in Iranian.

‘Wheat’ in Caucasian languages

Caucasian languages exhibit multiple roots that designate ‘wheat’. I classify them by origin into nine groups and one miscellaneous type. Some forms are related to the word for ‘bread’ in the given language.

Type A appears in Adyge and all Kartvelian languages except for Svan. Type B is mainly found in Abkhazo-Adyghean languages. Type C exists in certain Nakh languages (Chechen and Ingush). Type D is distributed in the north of the Northeast Caucasian-speaking area. Type E is found at the southern periphery of the Northeast Caucasian-speaking area. Types F and G appear in part of the north of the Northeast Caucasian-speaking area, in Bagvalal, Tindi and Chamalal and in Tsez, Khvarshi and Hinukh, respectively. Type H is found in Avar, Botlikh and Khinalug. Type I is a loan from Azerbaijani. Type X is a miscellaneous category, containing the remaining roots.

At present, it is challenging to interpret this distribution of word forms.

(SUZUKI Hiroyuki)

A: *xorbal-i*, *kobal-i*, etc.
B: *a-čʲa*, *čiʕʷra*, *kʷəcɨ*.
C: *jalta*.
D: *q’ir*, *muq’e*, *buq’*, *rizq’i*, etc.
E: *fu*.
F: *šibaratlʷ’*, *cibatl’ab*, etc.
G: *at’*, *arɯn*.
H: *t’oršel*, *uršal*, *pšä*.
I: *täχilʲ*.
X: others; *malpxw*, *ʁawə*, *k’av*, *bataχu*, *ruzi*, *guni*, etc.

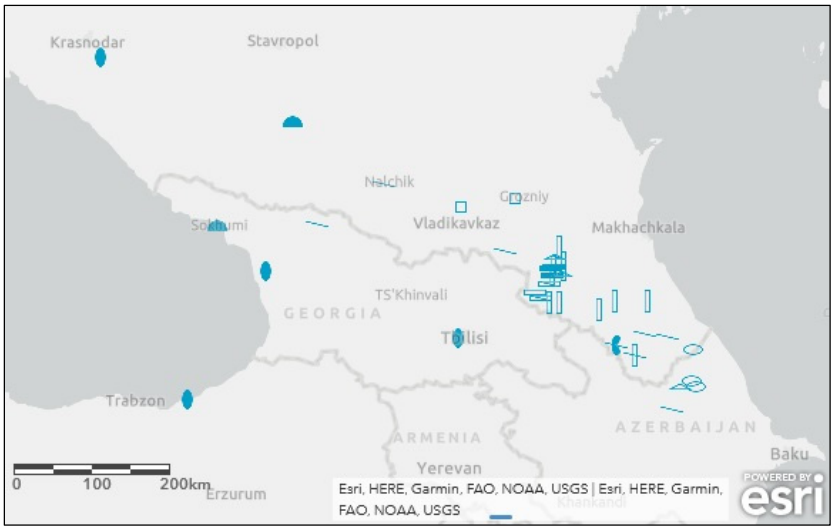


Figure 7.17.1: ‘Wheat’ in Caucasian languages.

‘Wheat’ in Semitic

A-1. *ħinṭa* type is found in Mesopotamia. The Arabic forms of this type are *ħinṭa* (Çukurova in Turky), *ħunṭa* (Iraqi, Bəħzāni, Basra, Tell Kēf), *ħönṭa* (Babylon), *ħənṭa* (Mossul). And *ħunṭa* in Uzbekistan Arabic derived from Mesopotamian Arabic. These may be borrowed from Aramaic.

A-2. In Aramaic and Hebrew **nṭ* changed to *ṭṭ*, thus Aramaic *ḥeṭṭe* (Hertevin), *xəṭṭe* (Khabur), *ħeṭṭθa* (Ma’lula) and Hebrew *χiṭṭa* חטה < Biblical Hebrew *ḥiṭṭa* חטה. cf. Ugaritic *ħ-ṭ-t* 𐎈𐎉𐎚, Syriac *ḥiṭṭā* or *ḥeṭṭō* ܚܛܬܐ ‘a grain of wheat’.

B. *qamħ* type is found in Arabic of Arabian peninsula and Africa. It is probably borrowed from Aramaic. The oldest forms are Biblical Hebrew *qémaḥ* קֶמַח ‘fine meal (of barley or wheat)’, Aramaic *qamḥa* קַמְחָא, Syriac *qamḥō* ܩܡܚܐ ‘flour, meal’, Ge’ez *qamḥa* ቀምሐ ‘to eat grain or other fodder, graze’ or *qamḥ* ቀምሕ ‘produce, yield, fruit, leguminous plant’.

B-1. The voiced *g* for ق /q/ is the feature of the nomadic or rural dialect. In Africa *gamħ* (Hassaniya Arabic in Mali, Tripoli of Libya, Djerba of Tunisia), *gəmħ* (Khemisset of Morocco, Tunisia, Algeria) and in the Sudan belt *gamiħ* (Sudan), *gameh* (Chad), *game’* (Juba).

In Arabian peninsula *gamħ* (Mekka in Saudi), *gumħ* (Bahrain). In San’a of Yemen *gamħ* is found but *birr* is more used.

In Palestine and the surrounding area *gamħ* (Arāmša in Galilee, Negev, Sinai), *gəmħ* (Irbid of Jordan), *gamiħ* (Amman).

B-2. The voiceless *ʔ*, *q* or *k* is the urban dialect feature: *ʔamħ* (Cairo, Lebanon, Tlemen of Algeria), *ʔamh* (Malta *ħ* > *h*), *ʔaməħ* (Damascus), *ʔamiħ* (Jerusalem), *ʔaməħ* (Tripoli of Lebanon), *ʔamaħ* (Sūr).

B-3. The urban voiceless *q* or *k* forms are *qamħ* (Masqaṭ of Oman, Abu Dhabi, Djidjelli of Algeria), *qəmħ* (Algier), *kamħ* (Ramallah in Palestine). *kampx* of Cypriot Ar. has an epenthesis consonant *-p-*.

C. *bar* type is distributed in the southern part of the Arabian Peninsula: *bàr* (Hobyot), *bohr* (Jibbali) of South Arabian languages and *burr* (Gulf Arabic) *birr* (Yemeni Arabic). Probably Arabic *burr* or *birr* is borrowed from the South Arabian languages. Hebrew *bar* ‘grain’ is probably not related to it.

D. *ʃenrai* or *sɨndaj* type are found in Ethiopic. *ʃenrai* (Tigre of Eritrea) and *sɨrnaj* (Tigrinya of Eritrea, ስርናይ) are related directly to *ʃerna:j* (Ge’ez ሥርናይ) with metathesis of *n* and *r*, and *ʃ* > *s* respectivly.

And *sɨndaj* (Tigrinya of Rayya, Nothern Ethiopia) and *sɨnde* (Amharic, ስንዴ) are also probably derived from the Ge’ez form.

There seems to be no cognate of the Akkadian *kibtu* 𒆠𒅁𒌅 in the modern Semitic. cf. Sumerian *GIG* 𒄀 ‘wheat’.

(NAGATO Youichi)

■ A-1: *ħinṭa*, *ħunṭa*, *ħənṭa*, *ħōnṭa*
□ A-2: *ḥeṭṭe*, *ḥeṭe*, *ħeṭṭθa*, *xəṭṭe*, *χita*
▲ B-1: *gamħ*, *gamiħ*, *gəmħ*, *gameh*, *game’*
△ B-2: *ʔamħ*, *ʔamh*, *ʔamiħ*
◭ B-3: *qəmħ*, *kampx*
○ C: *bàr*, *birr*, *burr, bohr*
★ D: *sɨndaj*, *sɨnde*, *sirnaj*, *ʃenraj*

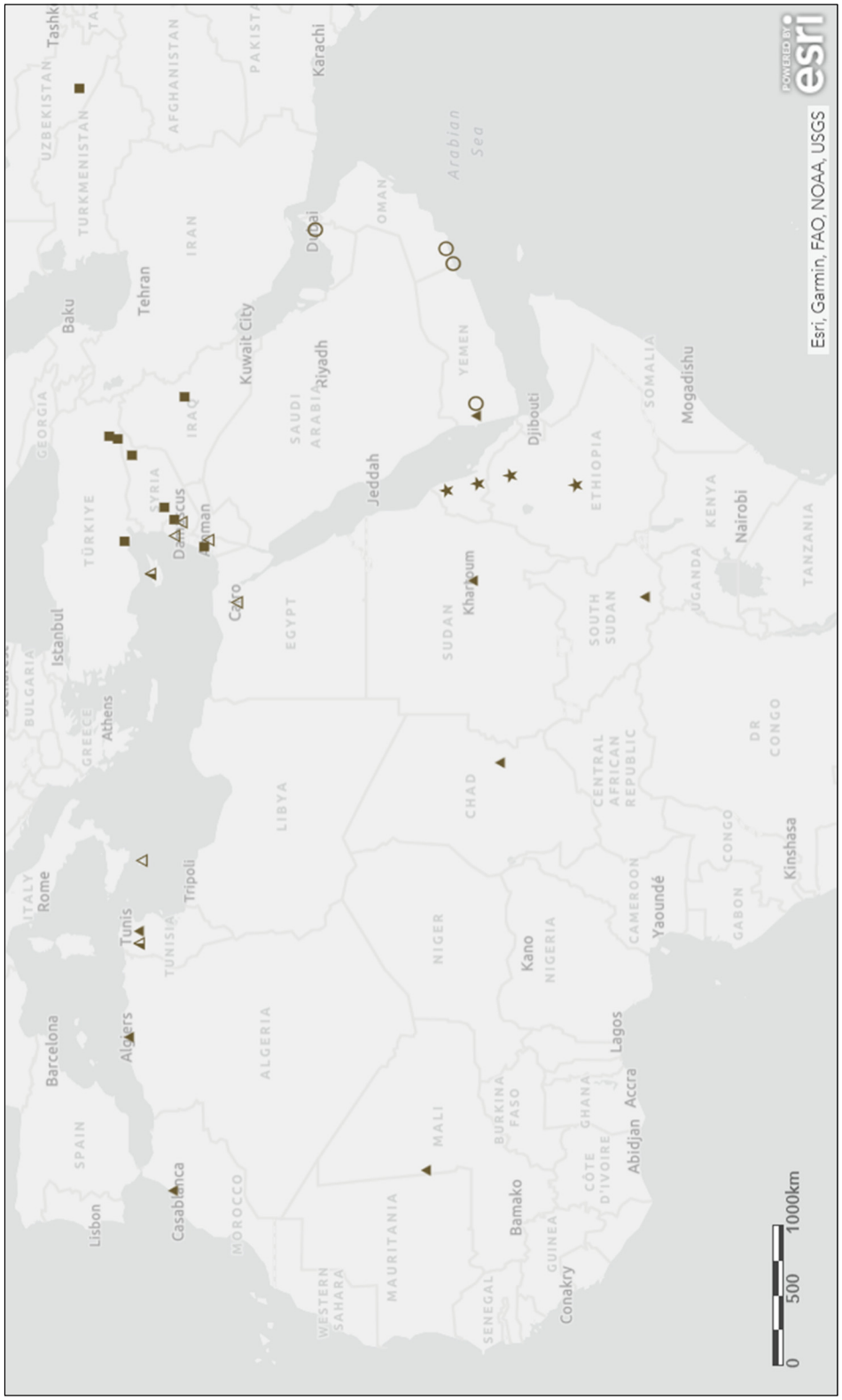


Figure 7.18.1: ‘Wheat’ in Semitic

‘Wheat’ in Nilo-Saharan

Most Nilo-Saharan languages do not attest native terms for ‘wheat’. Some languages attest a loanword from a local *lingua franca*, namely Arabic (Type A, via Sudanese-Chadian Arabic *(al-)gamiḥ* or Written Arabic *(al-)qamḥ*), Swahili (Type B, *ngano* ultimately from Indo-Aryan, e.g., *gēhũ̃*, Gujarati *ghaũ*) or Amharic (Type C, *sənde*). An old source (Muratori 1948) gives a periphrastic phrase ‘foreign (aliens’) millet’ (Type D) for three languages spoken in South Sudan (Bari, Lotuko and Acoli), but these terms seem to be obsolete.

A few languages attest native terms for ‘wheat’. Western Nilotic Nuer attests *lap* (Type E), which could be related to a root for ‘rice’ or ‘sorghum’ in some other Western Nilotic and Surmic languages. Nile Nubian languages attest the common root *illee* (Type F). Three Ethiopian Surmic languages attest different terms (Type G, Kwegu *nyamcele*, Me’en *sarguwo*, Suri/Tirma *goso*).

(NAKAO Shuichiro)

- A: Arabic borrowing
- B: Swahili borrowing
- C: Amharic borrowing
- D: ‘foreign millet’
- E: Nuer *laap*
- F: Nubian *illee*
- G: Surmic (different terms)

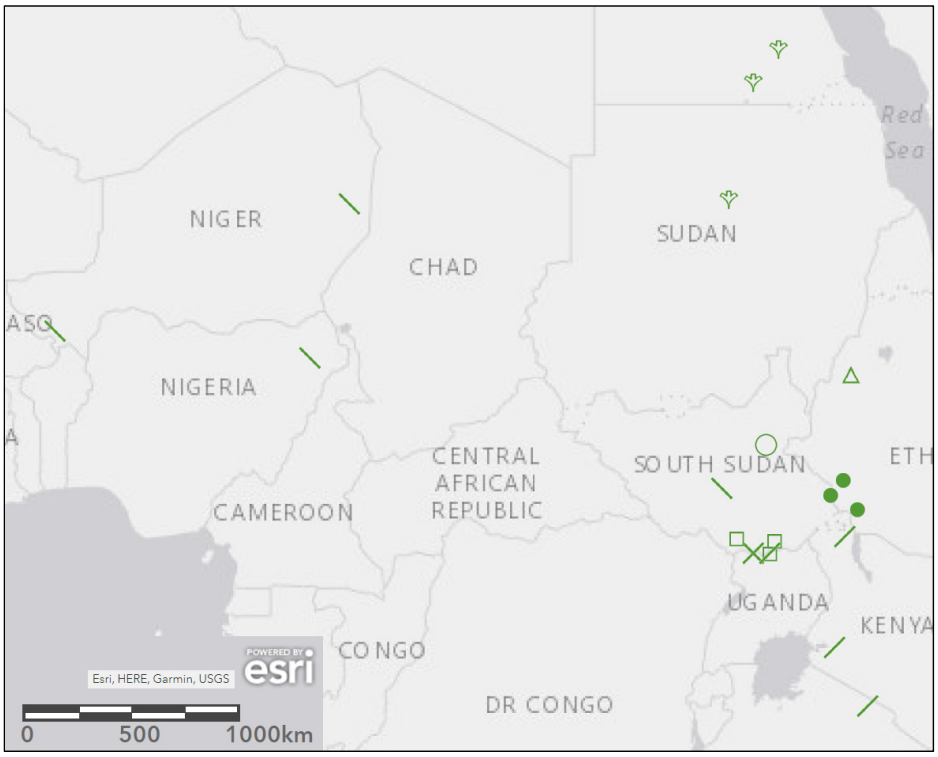


Figure 7.19.1: ‘Wheat’ in Nilo-Saharan.

Chapter VIII

Broomcorn millet

‘Broomcorn millet’ in Asian and African languages

Broomcorn millet, a.k.a. proso millet or common millet (*Panicum miliaceum*), was first domesticated in around 10,000 BP in Northern China (Lu et al. 2009) and is widely cultivated in Eurasia and spread worldwide.

Table 1 shows the principal forms in our data: Ainu, Korean, Japonic, Sinitic, Kra-Dai, Tibeto-Burman, Uralic, Mongolic, Turkic, Indo-Aryan, Nuristani, Dravidian, and Iranian. Data from languages in Africa are unavailable.

Table 1: Main word forms for ‘broomcorn millet’.

Languages	Word forms
Ainu	*sipúskep* *menkur(u)* *oya amam*
Japonic	*kibi/kimi* 黍
Korean	*kicaŋ* *cicaŋ*
Sinitic	*shu* 黍 type *ji* 稷 type *mei* 糜 type
Kra-Dai	*Fang* type
Tibeto-Burman	PLB **tsap* PLB **C-lu-k* *jon*
Uralic	*hirssi* *jamks* *tar* *proso* (< Russian *npoco*)
Mongolic	*amū* type *pisɣə̄* *budā* type *arzan* (< Persian *arzan*)
Turkic	*tarïk* type *dügü* *vir*
	sokba type *arzan* (< Persian *arzan*) proso (< Russian *npoco*)
Indo-Aryan & Nuristani	*cīna(ka)* Proto-IA **Harnuṣ* *ṭ/kaṅgunī* *karaz* *lāva* *priyaṅgu* *śyāmāka*
Dravidian	*varaku* *baragu* *varigelu*, *varaga*
Iranian	PIr. **(h)ardzana-* *pindʒ* *garis* PIr **yavas*

It is noticeable that the Chinese character for ‘broomcorn millet’ is from that of Sinitic *shu*, which is mainly used in the northern Sinitic area.

Several languages use the derivation forms from the cover term ‘crop’ or ‘grain’, which is the case for Ainu *amam* ‘grains’. The *Fang* type in Kra-Dai is linked to the ‘foxtail millet’ and ‘barnyard millet’ crops. The PLB forms **tsap* and **C-lu-k* in Tibeto-Burman are also used for ‘foxtail millet’ and ‘barnyard millet’, respectively.

Loanwords are confirmed in Uralic, Mongolic, and Turkic from Persian and Russian.

(SUZUKI Hiroyuki)

‘Broomcorn millet’ in Ainu

In “Moshiogusa 藻汐草” (1792), the first Ainu-Japanese dictionary, by Uehara Kumajirō, “シプシケ” [ʃipuʃike] was recorded as ‘foxtail millet (*Setaria italica*)’. However, I estimate for this term that the phonological representation is /sipuskep/, which means ‘broomcorn millet (*Panicum miliaceum*)’ in the Saru dialect. The word *sipúskep* consists of the verb *sipúske* ‘swell’ and the classifier *-p* ‘thing’ (Tamura 1996).

For ‘broomcorn millet (*Panicum miliaceum*),’ the word *menkuru* or *menkur* is used in the Shizunai dialect (Watanabe et al. 1984; Watanabe et al. 1994), and the word *oyá amám* (lit. ‘other grains’) is used in the Mukawa dialect (Watanabe et al. 1991) (see also ‘Foxtail millet’ in Ainu).

(FUKAZAWA Mika)

A. *sipúskep*
B. *menkur(u)*
C. *oya amam*

Figure 8.2.1: ‘Broomcorn millet’ in Ainu.

'*Panicum miliaceum*' in Japonic

KIBI (*kibi*, *kiᵐbï*, *kibɯ*, *tʃibe*, *kiʔ*, *kiibi*, …) and KIMI (*kimi*, *kïmï*, *kïm*, *tsïm*, *kiɴ*, …) are word forms for *Panicum miliaceum* or proso millet that are widely distributed in mainland Japan and the Ryukyus. Also, several dialectal forms such as INAKIBI (*inakibi*, *inakimi*, *inekibi*, *inarikibi*) 'rice millet', AWAKIBI (*awakibi*, *awakimi*) 'millet of millet', KOKIBI (*kokibi*, *kokimi*, *kogimi*, *kokkimi*, *kogiɴ*) 'small millet', TOOKIBI (*tookibi, cjookibi, tookiɴ, tokibi, tokiʔ*, …) 'Chinese millet', and HOOKIKIBI (*hookikibi*, *hookikimi*, *hookibi*, *hokkibi*, *hokibi*) 'millet for broom' exist. -KIBI represents both -*kibi* and -*kimi*, except for KIBI.

The oldest attested forms for proso millet in Japanese are *kimi* and *kibi*, which are distributed both in mainland Japan and the Ryukyus, and are considered to be Proto-Japonic form(s). More precisely, *kimi* seems to be older than *kibi*, based on the center-versus-periphery theory (*Hōgen-Shūken-Ron*), since attestations of *kimi* are distributed outside thouse of *kibi*.

It is said that proso millet was introduced into Japan after foxtail millet and barnyard grass. Therefore, it is presumed that KIBI/KIMI is a relatively new word. In addition, KIBI/KIMI refers to 'high millet, great millet, corn, and sugar cane', and the map may include KIBI/KIMI, which refers to not only proso millet, but also high millet and other corns. Cognate words with KIBI/KIMI in other languages are unknown.

(NAKAZAWA Kohei and YOKOYAMA Akiko)

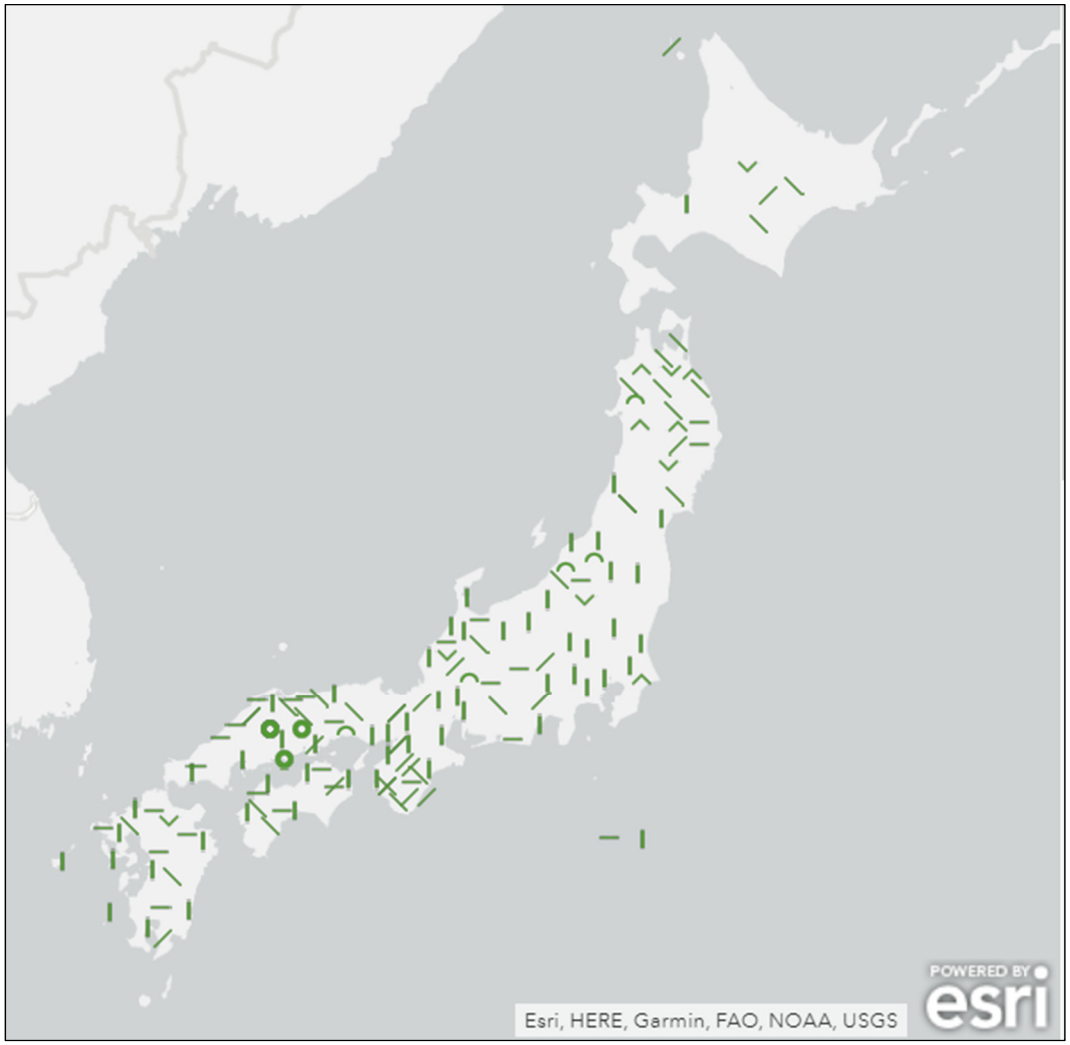


Figure 8.3.1: '*Panicum miliaceum*' in mainland Japan.

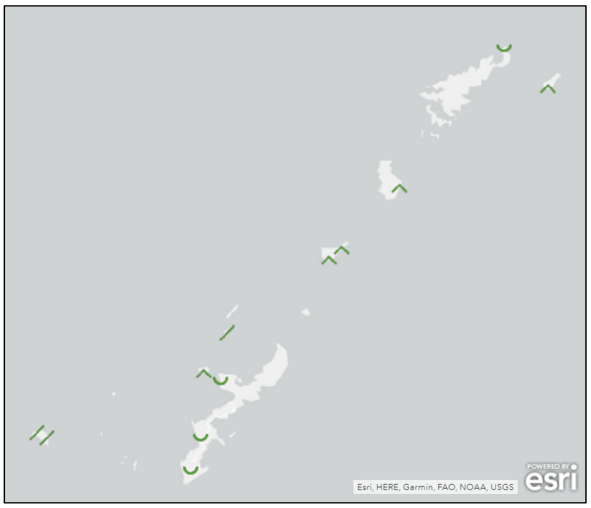


Figure 8.3.2: '*Panicum miliaceum*' in Northern Ryukyu Islands.

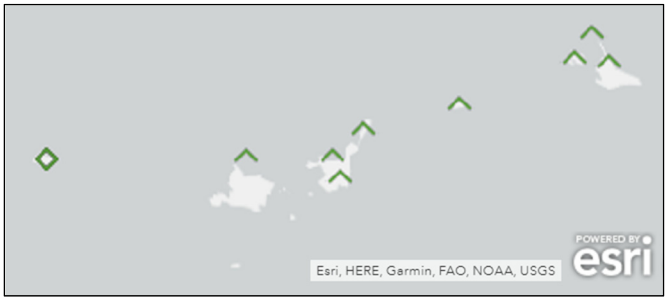


Figure 8.3.3: '*Panicum miliaceum*' in Southern Ryukyu Islands.

‘*Panicum miliaceum*’ in Korean

The following table shows Modern and Middle Korean terms for ‘panicum miliaceum’. Middle Korean tones are shown in parentheses.

	Modern Korean	Middle Korean (15-16c.)
Panicum miliaceum (キビ)	kicaŋ	kicaŋ (LL)

As can be seen in this table, Modern and Middle Korean forms are almost the same except for the fact that Middle Korean form has a distinctive LL pitch.

This word was not recorded in *Jīlín lèishì* (鷄林類事, 12th century), but recorded in *Cháoxiǎnguǎn yìyǔ* (朝鮮館訳語, early 15th century) as follows:

	鷄林類事	朝鮮館訳語
Panicum miliaceum	–	黍米 吉雜色二

In this document, ‘*Panicum miliaceum*’ is recorded as ‘吉雜色二’ which seems to correspond to a complex word *kicaŋ-psʌr*.

As for the modern forms, we have two main types: one retaining original *kicaŋ*, and the other changing its initial *k* to a palatalized form *cicaŋ or cicɔŋ*.

A1: *kicaŋ*
A2: *cicaŋ, cicɔŋ*

Type A1 appears in the central dialects and A2 in southern dialects. The next map shows a simplified version of data recorded in Ogura (1944). Unfortunately, data for northern dialects are not recorded for this item.

(FUKUI Rei)

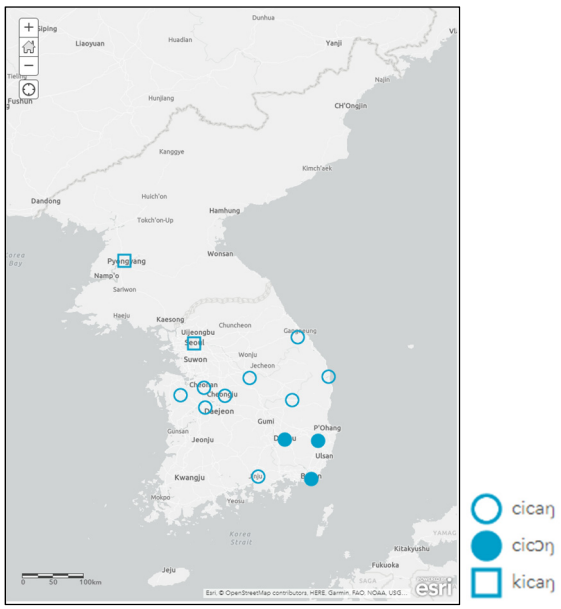


Figure 8.4.1: ‘*Panicum miliaceum*’ in Korean.

‘Broomcorn millet’ in Sinitic

Broomcorn millet, primarily cultivated in northern China, can be classified into two types by glutinosity. Word forms for ‘broomcorn millet’ are often not listed in the reports of southern dialects. Since Sinitic languages distinguish between the plant name and the husked grains of broomcorn millet, we made a map for each term.

Figure 8.5.1. shows the broomcorn millet plant name, classified by stem into three types: A *shu* 黍, B *ji* 穄, and C *mei* 糜.

A-1 *shu* 黍 is the generic form for ‘broomcorn millet’ in Old Chinese. It is often used exclusively to mean ‘glutinous broomcorn millet.’ B-1 *ji* 穄 and C-1 *mei* 糜 are also found in Old Chinese, but they mean ‘nonglutinous broomcorn millet’. *Mei* 糜 is a dialectal variation of *ji* 穄 (Huang 2016). Table 1 shows the reconstruction of Old Chinese forms by Baxter and Sagart (2014).

Table 1

	OC	MC	referent
shu 黍	s-t^{h}aʔ	syoX	generic; glutinous
ji 穄	[ts][a][t]-s	tsjejH	nonglutinous
mei 糜	C.m(r)aj	mje	nonglutinous

The A *shu* 黍 type is distributed in eastern Huabei, while the C *mei* 糜 type is distributed in western Huabei, showing an east–west contraposition.

The A *shu* 黍 type denotes ‘glutinous broomcorn millet,’ but the A-3 *jishu* 稷黍 and A-4 *meishu* 糜黍 subtypes modified by *ji* 稷 or *mei* 糜 denote nonglutinous broomcorn millet. In contrast, the C *mei* 糜 type denotes nonglutinous broomcorn millet, but the C-2 *ruanmeizi* 软糜子 (‘soft broomcorn millet’) subtype denotes glutinous broomcorn millet.

Figure 8.5.2. shows the word forms for ‘husked grains of broomcorn millet.’ All the word forms share the stem *mi* 米 meaning ‘husked seed.’ We classified them by their referents into three types: A *huang* 黄 type, B glutinous type, and C nonglutinous type.

The A *huang* 黄 type is distributed in Huabei. B glutinous type distributed in eastern Huabei corresponds to the A *shu* 黍 type in the Figure 8.5.1. The referents of A *huang* 黄 type vary depending on dialect.

Plant name / Husked grains

Lanzhou (兰州): 糜子[mi^{51} tsɿ^{0}]
/黄米[xuɔŋ^{51} mi^{0}]: nonglutinous

Jingle (静乐): 1) 糜子[mi^{33} tsəʔ^{4}]
/黄米[xuɤ^{33} mi^{314}]: nonglutinous
2) 黍子[fu^{314} tsəʔ^{4}]
/软米[vɛ̃^{35} mi^{314}]: glutinous

Shuozhou (朔州): 1) 糜子[mi^{212} zəʔ^{0}]
/糜子米[mi^{212} tsəʔ^{44} mi^{53}]: nonglutinous
2) 黍子[ʂu^{53} zəʔ^{0}]
/黄米[xuɒ^{212} mi^{0}]: glutinous

Juxian (莒县): 黍子[θu^{13} tsɿ^{0}]
/黄米[xuaŋ^{53} mi^{13}]:glutinous
黏米[nian^{53} mi^{55}]: glutinous

(SUZUKI Fumiki)

A. *shu* 黍 type
- A-1 *shu* 黍, *shuzi* 黍子
- A-2 *ruanshu* 软黍
- A-3 *jishu* 稷黍(nonglutinous)
- A-4 *meishu* 糜黍(nonglutinous)

B. *ji* 穄 type
- B-1 *ji* 穄, *jizi* 穄子(nonglutinous)
- B-2 *huangji* 黄穄

C. *mei* 糜 type
- C-1 *mei* 糜, *meizi* 糜子(nonglutinous)
- C-2 *ruanmeizi* 软糜子

D. others: 黄秥, 黄狗粟, 穄子

Figure 8.5.1: ‘Broomcorn millet’ (plant) in Sinitic.

A. *huang* 黄 type
- A-1 *huangmi* 黄米
- A-2 *dahuangmi* 大黄米
- A-3 *xiaohuangmi* 小黄米

B. glutinous type
- B-1 *shumi* 黍米, *shuzimi* 黍子米
- B-2 *nianmi* 黏米
- B-3 *ruanmi* 软米

C. nonglutinous type
- *meizimi* 糜子米

Figure 8.5.2: ‘Broomcorn millet’ (husked grain) in Sinitic.

'*Panicum miliaceum*' in Kra-Dai

'*Panicum miliaceum*' has only one word form (type A). Type A is further divided into three types according to initial consonants.

A3, with the initial consonant *ph-*, is distributed in close proximity to A4-subtype (i.e. *ph-* type) of Setaria italica. For this reason, the A-type of *Panicum miliaceum* may be the same as the A-type of *Setaria italica*, which goes back to the Proto-Tai *f- of Li (1977). However, *Panicum miliaceum* and *Setaria italica* do not overlap in data points, so it is unknown what word forms of *Setaria italica* are represented at these places.

(TOMITA Aika)

A. *Fang* type
- A-1 *f-* type
  faaŋ^{3}
- A-2 *v-* type
  viaŋ^{3}
- A-3 w- type
  phaaŋ^{3}, p^{h}a:ŋ^{C2}

Figure 8.7.1: 'Panicum miliaceum' in Kra-Dai.

‘Broomcorn millet’ in Tibeto-Burman

Among the 26 collected items of the word form for ‘broomcorn millet’ (*Panicum miliaceum*) in TB languages and varieties, we find two major types and five minor ones.

Type A includes forms derived from PLB *tsap ‘millet’, attested in LB languages with /ts^{h}, tɕ^{h}, s/-initials.

Type B is derived from PLB *C-lu-k ‘millet’, attested in Lisu and Hani.

The following are minor types. Type C only appears in Lolopo, and it may be related to the /l/-initial of Type B. Type D only appears in Ronghong Rma, related to *ʂpa* (Mawo Rma) ‘foxtail millet’. Type E appears in Naic languages and has an unclear origin. Type F is in Kurtö, Bumthang and Khengkha, possibly related to *yó* (Zaiwa) ‘barnyard millet’. Type G is only in Rawang, related to *luŋšin* (Atsi) ‘foxtail millet’. Type X contains various independent stems, found in languages spoken in Jino, Jinghpaw and Kathmandu Newari.

Looking at Types A, B and C, it can be concluded that, based on their limited distribution, Type A is older than Types B and C.

(SUZUKI Hiroyuki, EBIHARA Shiho, IWASA Kazue, KURABE Keita, SHIRAI Satoko)

□ A. PLB *tsap: *tshi*^{55}, *tɕhy*^{21}, *hshe̲*^{55}.
● B. PLB *C-lu-k: *sa*^{33} *la̲*^{31}, *lɔ*^{31} *lø*^{55}.
/ C. *no*^{55}.
⬭ D. *ʂpa*
✳ E. *dzuꜜluꜜ*, *dzuꜜnjɤꜜ*.
⬮ F. *jon, yon*.
◁ G. *lʏ̀mbóngshí*.
▬ X. *khu*^{42} *jo*^{44} *to*^{44} *mi*^{44}, *ɛəgyi*, *satiwā*.

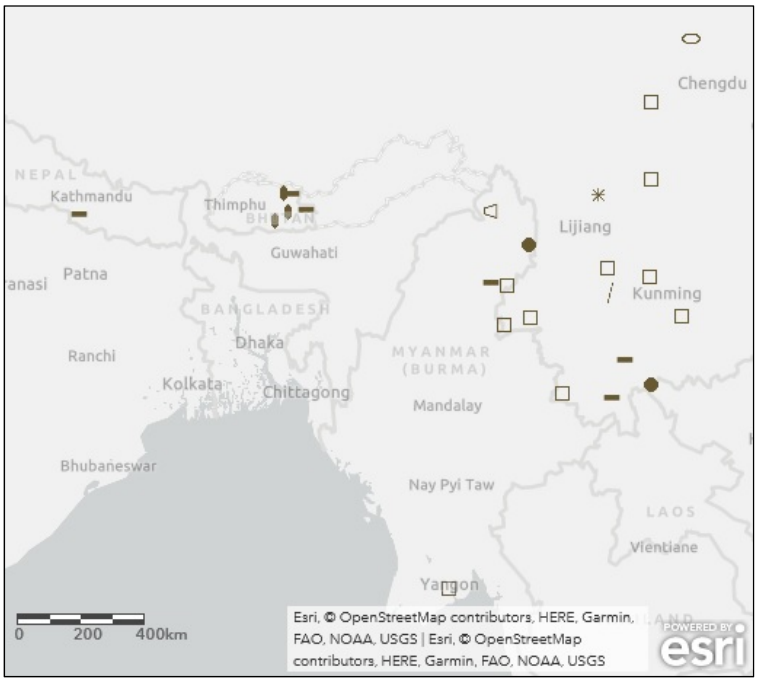


Figure 8.8.1: ‘Broomcorn millet’ in Tibeto-Burman.

‘(Proso) millet’ in Uralic

The word for ‘millet’ is less observed than ‘wheat’. No word in the northeastern area - Samoyedic and Ugric - as in the case of ‘wheat, and in some Finno-Baltic and Saami.

Type A *hirssi* in Finnish, Estonian
Type B *jamks* in Mordvin
Type C *tar* in Mari
Type D *köles* in Hungarian
Type F borrowing from Ru. *proso* in Komi

(MATSUMOTO Ryo)

- Type A *hirssi*
- Type B *jamks*
- Type C *tar*
- Type D *köles*
- Type F borrowing from Ru. *proso*

Figure 8.12.1: ‘(Proso) millet’ in Uralic.

‘Broomcorn millet’ in Mongolic and Turkic

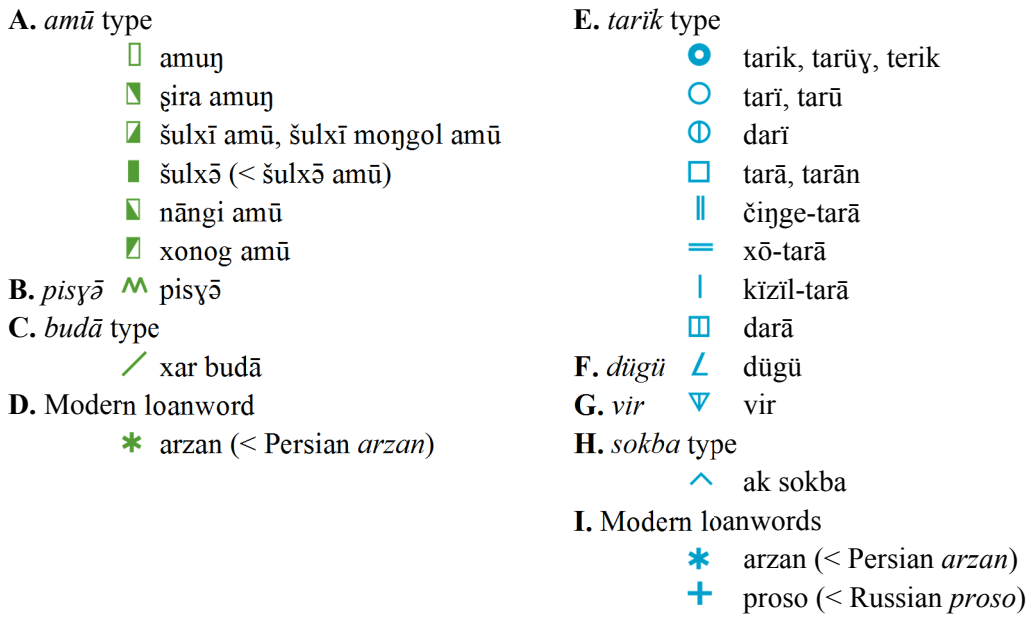
Both Mongolic and Turkic people seem to be familiar with broomcorn millet to a certain extent, but they use completely different words (except the Persian loanword *arzan*).

The Mongolic *amū* and Turkic *tarïk* types are the dominant forms (cf. Written Mongol *amun* ‘grain,’ Old Turkic *tarïg* ‘grain’). Most other forms are found in peripheral regions: *pisɣə* in Dagur, *vir* (< *üren* ‘seed’) in Chuvash, *dügü* in Salar, *ak sokba* (*ak* ‘white,’ *sokba* ‘cereal’) in Sarïg Yughur. Some languages have more than one term for the crop: *šulxī moŋgol amū* (*šulxī* ‘slobbery’), *xonog amū* and *xar budā* (*xar* ‘black,’ *budā* ‘grain’) in Khalkha Mongol, *šulxī amū* and *nāngi amū* (*nāngi* ‘sticky’) in Chakhar Mongol, *tarā* (‘grain’), *čiŋge-tarā* (*čiŋge* ‘fine’), *xō-tarā* (‘millet with a drooping spike’) and *kïzïl-tarā* (*kïzïl* ‘red’) in Tuvan, etc.

In general, a word for ‘grain’ is used to denote the millet they eat. Crops called by the same word may differ depending on the region.

The Salar form *dügü* and the Uighur form *terik* can denote both broomcorn millet and foxtail millet. This may mean that they do not always clearly distinguish between them.

(SAITÔ Yoshio)



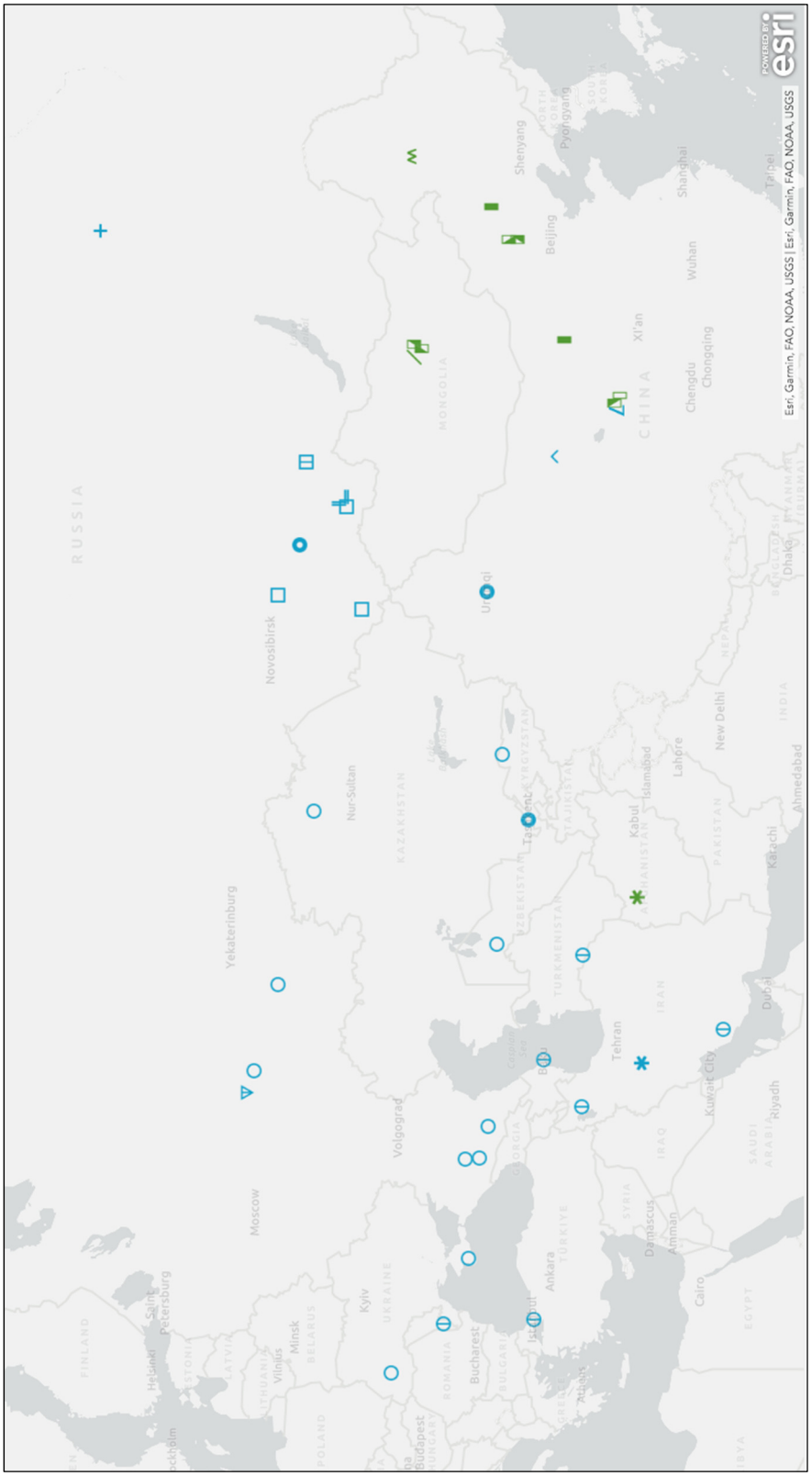


Figure 8.13.1: ‘Broomcorn millet’ in Mongolic and Turkic.

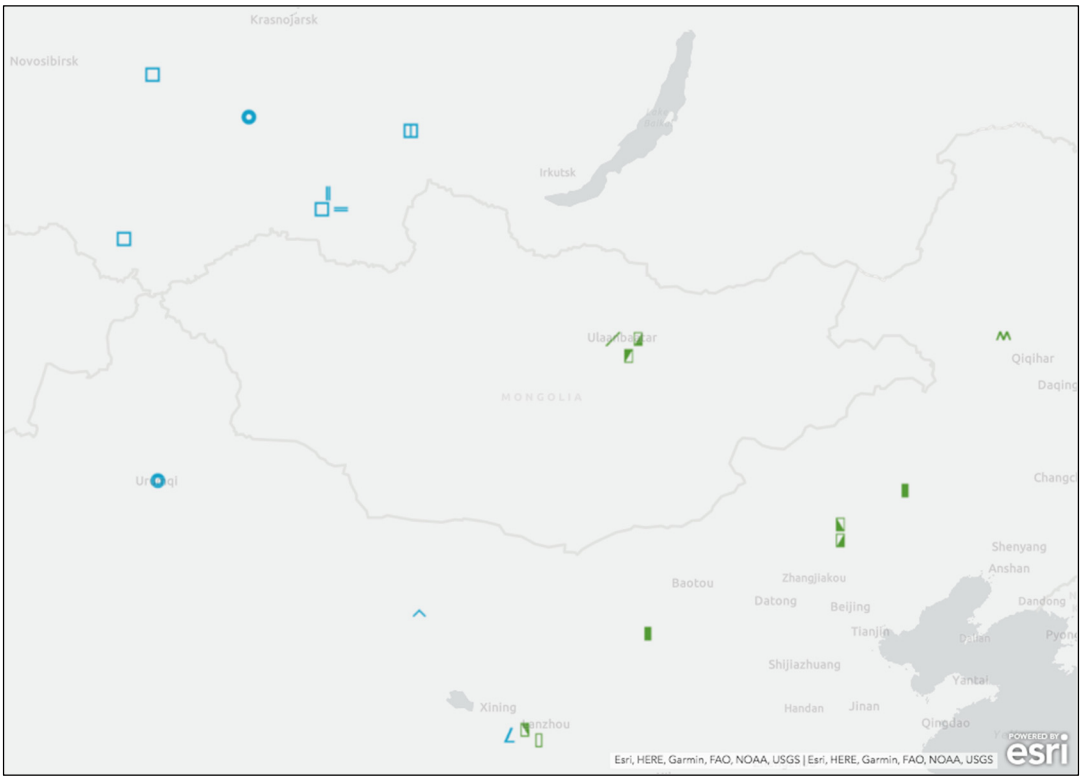


Figure 8.13.2: ‘Broomcorn millet’ in Mongolic and Turkic (The Mongolian Plateau and its vicinities magnified).

‘Proso millet (*Panicum miliaceum*)’ in South Asia

I treat the languages of Indo-Aryan (IA) and Nuristani of the Indo-Iranian branch of IE, and language isolates in South Asia.

There are six relatively major categories of word forms – A) *cīna(ka)*, B) *aṇu(ni)*, C) *ṭ/kaṅgunī*, D) *karaz*, and E) *lāva* – with minor categories – F) *priyaṅgu* and G) *śyāmāka*. Here I targeted ‘proso millet’ or simply ‘millet’, not ‘a kind of millet’, when a dictionary does not well specify the referent with scientific names in detail.

The distribution of ‘proso millet’ words is hard to draw the shape. The type A is plotted mainly in the plains in Pakistan, northern India, and Bangladesh. In Karakoram-Hindukush, on the one hand, IA and Nuristani languages employ Types B.

The most major type is *cīna(ka)*. This type is derived from Sanskrit *čīna* चीन ‘Panicum miliaceum’ with or without the suffix *-ka*. I’m not sure the term is related to the ideophone *čīna* चीन ‘China’.

The *aṇu(ni)* type appears in IA and Nuristani, which are concentrated in the area of the Karakoram and Hindukush mountain ranges. The original Sanskrit form *áṇu* अणु refers to ‘the grain plant Panicum milliaceum, atom, particle’, which is inherited from Proto-IA **Harnuṣ*. And its derived word *aṇuni* अणुनि refers to ‘millet’.

The third type *ṭ/kaṅgunī* is used in IA languages mainly in the Himalayan range. Sanskrit *kaŋgunī* कङ्गुनी means ‘Setaria italica, Celastrus paniculatus’, which deeply correlates with **kaŋkunī ~ ṭaŋgunī* ‘a panic grain, id.’. This type is well found in the terms of ‘foxtail millet’ (further see the chapter).

Next, the *karaz* type can be seen only in IA languages. The etymon of the forms is not clear and the type name is decided by me for now. Turner (1966) suggests *kā́śa* काश ‘grass’ or *grāsa* ग्रास ‘mouthful’, however I feel the modern forms do not allow us to fix the etymon easily.

The type E of *lāva* is found in Nuristani languages in Hindukush and Vaagri Boli in southern India. The word *lāva* लाव in Sanskrit originally meant ‘reaping, cutting’. The *priyaṅgu* type for IA languages is descendant of Sanskrit *priyáṅgu* प्रियंगु ‘*Panicum italicum*’, and modern *śyāmāka* terms are from Sanskrit *śyāmā́ka* श्यामाक ‘*Panicum frumentaceum*’. The former type is detected in ‘foxtail millet’, and the latter in ‘barnyard millet’. In the same way, Sanskrit *maḍaka* मडक ‘the small grain *Eleusine corocana* (finger millet, シコクビエ)’ has descendant forms in Himalaya.

varī type for Dhivehi *urā* ތުއްރާ and *yavanāla* type for Hindi *junhār* जुनहार are related to other kinds of millet. As for the former one, see also ‘Barnyard millet in South Asia’. The latter originally refers ‘sorghum’. Kalderaš Romani *mohóro* may be from European terms (< Latin *milium* < PIE **melh₂-* ‘to grind, crush’), while ‘Standard’ Romani *khurmi* is a descendant of Sanskrit *karambhá* करमभ ‘groat, gruel’. In Ashkuňu, they employ *pisã*, which is originated in Sanskrit *pēṣaṇa* पेषण ‘crushing / grinding of grain’. Nihali *oro* maybe a loanword from Marathi *varō* वलो ‘kind of grass, grain’ or Korku *oro* ‘sorghum’ (cf. Mundari *iṛi* ‘id.’).

(YOSHIOKA Noboru)

A. *cīna(ka)* type (12)
čĭ̃nẵ, čēnā, čeno, čin, čīṇẵ, čiṇā, čīṇō

B. *aṇu(ni)* type (9)
ā́ṇu, aṛ́n, ärín, éṛin, awṛ̃ĭ̃, áriin, ạín, hạín, oḷíin

C. *ṭ/kaṅgunī* type (6)
kaŋgunī, kã̃guni, kaŋgṛī, kauṇī, kauṇe, kōṇī

D. *karaz* type (6)
kāraz, kāraž, karāc, kaaríi, garaṣ

E. *lāva* type (5)
lāw, ṛov, ʁāwõ̃, hāwõ̃

F. *priyaṅgu* type (4)
piyaᵑgu, piŋga, péreŋ, çiŋ

G. *śyāmāka* type (4)
sa̍̃mẵ, sa̍̃vā, sāĭ̃

H. others
[*bayˌ* (3)] bayˌ, ba; [*maḍaka* (2)] maṛuwā, moruwa; [*raṅ* (2)] rāko, ran; [*varī*] urā; [*yavanāla*] junhār; barag, gõdlī, khurmi, mohóro, oro, pisã

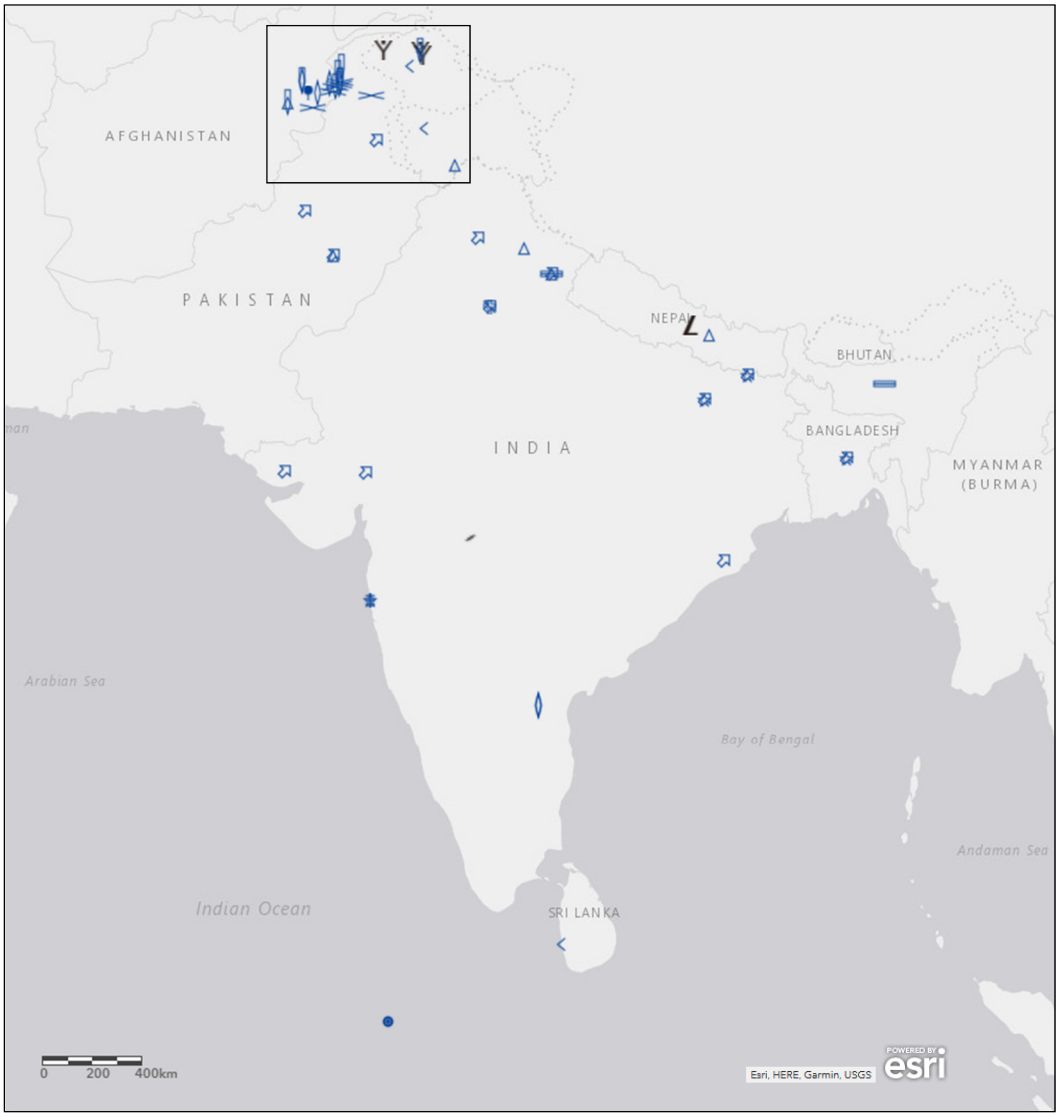


Figure 8.14.1: ‘Proso millet’ in SA: Indo-Aryan, Nuristani (both in navy blue), and language isolates (those in black).

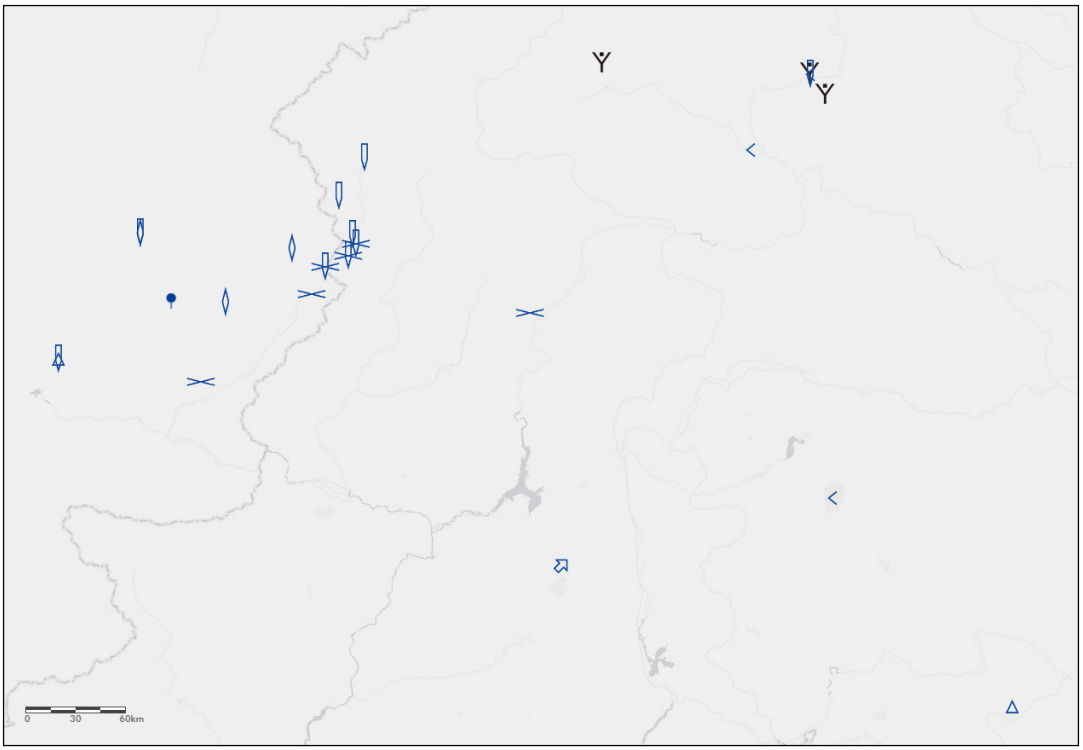


Figure 8.14.2: Types for 'proso millet' in the northern part (the area encloed by the rectangle in Figure 8.14.1).

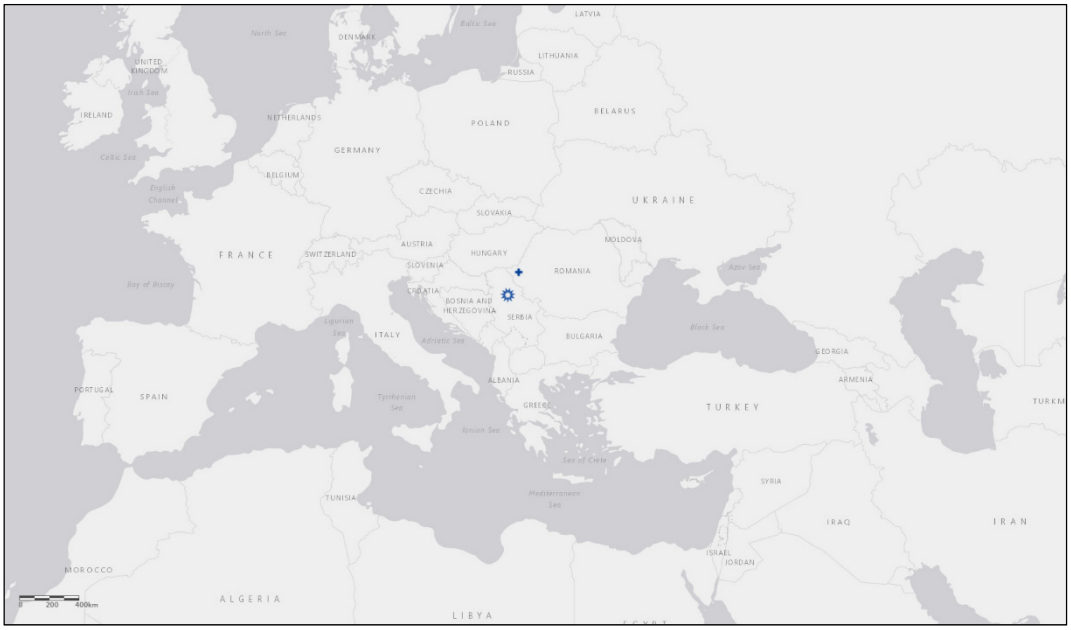


Figure 8.14.3: Types for 'proso millet' in Indo-Aryan languages outside South Asia.

‘Broomcorn millet’ in Dravidian

Information on this crop is scanty in DEDR. The etymon entry #5260 lists cognates in the four major literary languages, of which only Telugu *varagalu* (pl.) and varigelu (pl.) are specifically glossed as *Panicum miliaceum*.

Tamil *varaku* and Kannada *baragu* appear to be a generic term for some millets which include ‘broomcorn millet’, ‘barnyard millet’ and, in Tamil, ‘kodo millet’. The use of Malayalam *varaku* for ‘broomcorn millet’ is doubtful.

(KODAMA Nozomi)

varaku
baragu
varigelu, varaga

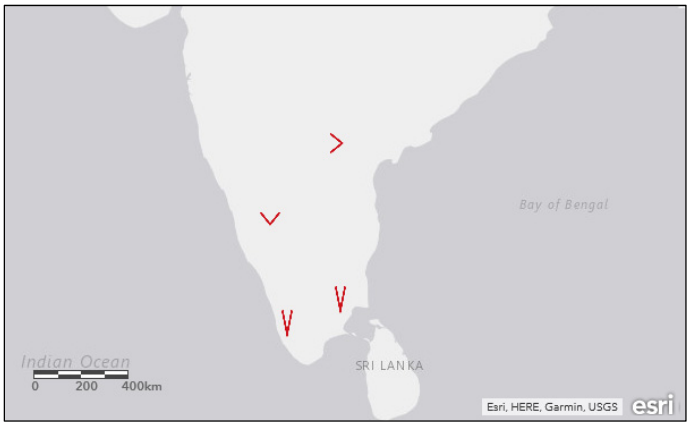


Figure 8.15.1: ‘Broomcorn millet’ in Dravidian.

‘Broomcorn millet’ in Iranian

Although millet is no longer as important as the staple crops wheat and rice in the Iranian language area, various terms for ‘broomcorn millet (*Panicum miliaceum*)’ are found throughout the region.

I have classified the Iranian terms for ‘broomcorn millet’ into 7 types (A through G). However, in some cases, it is difficult to distinguish whether the term meant ‘broomcorn millet’ or ‘foxtail millet (*Setaria Italica*)’.

Type A is the most widespread within the Iranian world. It dates back to Proto-Iranian *(h)ardzana- ‘millet, a kind of plant’. Type B is distributed in the northern Pamir area. This type is thought to be a loanword from Indic (cf. Kashmiri *piŋga* and Sanskrit *priyaṅgu*).

Type C is a confusing word because many other Iranian languages use this term to refer to ‘Foxtail millet’. Kurdish dialects have word forms of this type in addition to those of Type A (*herzin*).

Types D through G are scattered in the peripheral area of the Iranian spoken region. Every type is only seen in one language/ dialect. Type D is derived from Old Iranian **yava-* ‘barley’ (further from Proto-Indo-Iranian **yavas*). This type is observed in the Iron dialect of Ossetic. Type E *xarban* form is observed only in Yazgulami. Type F is derived from Old Iranian **xwarna-*, or **hwāra-*, which originally meant ‘food (in general)’ (cf. Persian خوراک *xorâk* ‘food’). This type is confined to the Digoron dialect of Ossetic. Zazaki has a Type G word form which is probably related to Persian گال *gâl* ‘a kind of millet’.

(IWASAKI Takamasa)

A: **(h)ardzana-* *arzun, herzin, ærzæn, arzan, arzən, úžḍьn, yūrzən, ažan, ârzən, žḍən, wužḍän, arzun, yɯrzn, yirz(n), irz(n), yúrzun*

B: *pindʒ* *pĭndʒ, pendʒ*

C: *garis* *garis*

D: **yava-* *jæw*

E: *xarban*

F: *xwar*

G: *gilgil*

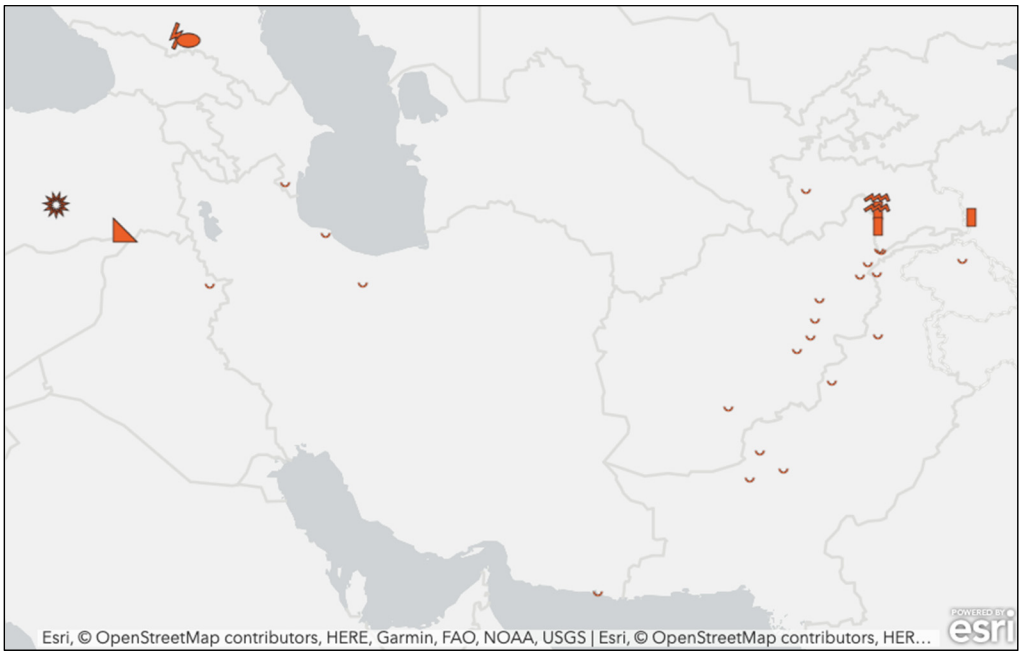


Figure 8.16.1: ‘Broomcorn millet’ in Iranian.

Chapter IX

Foxtail millet

‘Foxtail millet’ in Asian and African languages

Foxtail millet (*Setaria italica*) was domesticated in approximately 16,000 YBP, was a recognised crop in around 9000–10,000 YBP, and became popular in Northern China in about 5000–6000 YBP (Diao and Jia 2017). Now it is cultivated all over the world.

Table 1 shows the principal forms in our data: Ainu, Japonic, Korean, Sinitic, Hmong-Mien, Kra-Dai, Tibeto-Burman, Mongolic, Turkic, Indo-Aryan, Nuristani, Burushaski, Dravidian, and Iranian. Data from languages in Africa are unavailable.

Table 1: Main word forms for ‘foxtail millet’.

Languages	Word forms
Ainu	*múnciro* type *siru amam* *awá* (< Japonic *awa* 粟)
Japonic	*awa* 粟 *kumago* *tookibi*
Korean	*co* type *sɔsuk* type *susu* type
Sinitic	*su* 粟 type *gu* 谷 type *mi* 米 type
Hmong-Mien	*nau* *tsai* *noŋ* *tshaŋ*
Kra-Dai	*Fang* type *ɣung* type *la (u)* type *Sai* type *mung tso* type
Tibeto-Burman	PTB **khrəy* PLB **tsap* *muŋ*^{55}
	no^{35} use of other crops terms xiaomi, guzi (< Sinitic)
Mongolic	*narim* type *xonog* type
Turkic	*konok* type *tüge* type *sokba* type
Indo-Aryan & Nuristani	*ṭ/kaṅgunī* *kāśa* *cīna(ka)*
Burushaski	*čha*
Dravidian	*tinai* *kural* *ārgu* *paṇḍi*
Iranian	PIr. **gawarsa-* *qūnoq* (< Uzbek qo‘noq)

It is noticeable that the Chinese character for ‘foxtail millet’ in Japonic is derived from that of Sinitic *su*, which is mainly used in the southern Sinitic area.

Mongolic *xonog* is a cognate with Turkic *konok*. The Sinitic form *mi* is also used for ‘rice’ and the Proto-Tibeto-Burman form **khrəy* is also used for ‘barnyard millet’. Word forms that originally denote ‘rice’ are also used in some South Asian languages.

Japonic *awa* may be considered a cognate with Proto-Turkic **arpa* ‘barley’.

A lexical borrowing of the word for ‘foxtail millet’ is observed in some varieties of Ainu (from Japonic), some Hmong-Mien and Tibeto-Burman languages (from Sinitic), and Iranian (feom Uzbek, Turkic).

(SUZUKI Hiroyuki)

‘Foxtail millet’ in Ainu

The terms for ‘foxtail millet (*Setaria italica*),’ *múnciro* and *mínciro*, are original forms, and *awa* is a loanword from Japanese (Chiri 1976[1953]; Hattori 1964).

“Moshiogusa 藻汐草” (1792) includes four words, “ムジロ,” “シプシケ,” “キテナアマム,” and “ムリタンネ,” in the entry for *awa* 粟 ‘foxtail millet (*Setaria italica*)’. The word “ムジロ” [mudziro] can be assumed to be the word *múnciro* ‘foxtail millet’ and the word “ムリタンネ” [muritanne] is probably mistaken for “ムリクンネ” [murikunne] ‘foxtail millet’. The other two words may be confused with the words for other types of millet: e.g., “シプシケ” [ʃipuʃike] means ‘Broomcorn millet (*Panicum miliaceum*)’. The word “キテナアマム” [kitena amamu] is not found in the modern dictionaries (see also ‘Broomcorn millet’ in Ainu).

The word *amám* means ‘grains,’ including rice plant, foxtail millet, barnyard millet, and broomcorn millet. Fukazawa (2021) mentioned that the words *X amám* sometimes occur to distinguish different grains, as in Table 1. Foxtail millet is called *sirún amam ~ siru amam* (lit. ‘grain in this land’), *tóyta amám* (lit. ‘cultivated grain’), and *áynu amam* (lit. ‘Ainu’s grain’) in Hokkaido, and *mancuu amam* (lit. ‘grain in Manchuria 満州’) in Sakhalin. In contrast, a rice plant is called *síamam ~ siyámam*, which consists of the prefix *sí-* ‘real; true; the very’ and *amám* ‘grain,’ and *tonó amam* in ‘Japanese people’s grain.’

(FUKAZAWA Mika)

Table 1: The word for rice plant and three varieties of millet (modified from Fukazawa 2021).

<table>
<tr><th>Kinds of grain</th><th>amám</th><th colspan="2">X amám</th><th>Original form</th></tr>
<tr><td>Rice plant</td><td rowspan="4">amám</td><td colspan="2">síamam ~ siyámam
tonó amam</td><td>—</td></tr>
<tr><td rowspan="2">Foxtail millet
(Setaria italica)</td><td rowspan="3">tóyta amám
aynu amam</td><td>sirú(n) amam</td><td rowspan="2">múnciro ~ minciro
awá</td></tr>
<tr><td>mancuu amam</td></tr>
<tr><td>Barnyard millet
(Panicum Crusgalli L. var. frumentaceum Trin)</td><td>ayús amam ~
ayus ámam</td><td>piyápa</td></tr>
<tr><td>Broomcorn millet
(Panicum miliaceum)</td><td>oyá amam</td><td>sipuskep
menkur(u)</td></tr>
</table>

A. *múnciro* type

A-1. *múnciro ~ munciro*

A-2. *minciro*

B. *siru amam ~ sirún amam*

C. *mancuu amam*

D. *awá*

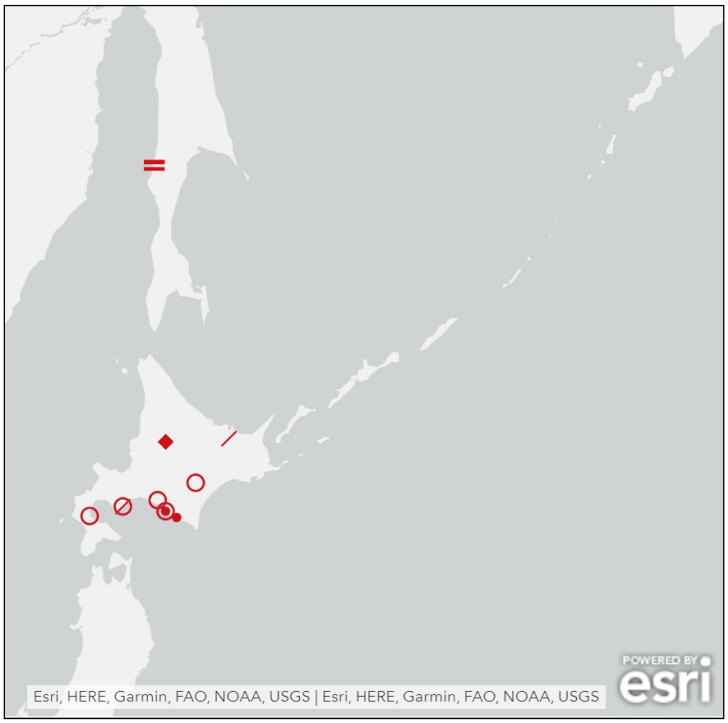


Figure 9.2.1: ‘Foxtail millet’ in Ainu.

'*Setaria italica*' in Japonic

Setaria italica or foxtail millet, is known as AWA (*awa*, *aa*, *ʔawaa*, *ʔoo*, *ʔaɴ*, *woa*, …) in both mainland Japan and the Ryukyus. There are also a few other forms such as AWAGOME (*awagome*, *awaɴkome*, *awa no kome*, *oogome*, …) 'rice of millet' in mainland Japan, and KUMAGO, AWAKIBI 'millet of millet', and YAMAAWA 'mountain millet' distributed in Western Japan.

The oldest attested form for foxtail millet in Japanese is *apa*, which is also considered to be the proto-Japonic form because the AWA type is widely distributed in both mainland Japan and the Ryukyus. In Japanese and Ryukyuan, the intervocalic (*)/p/ changed to [w]. At first glance, the lack of variation indicates that AWA was a relatively new loanword. However, AWA (/apa/ > /awa/) has experienced sound changes as seen in Naze *ʔo*, Yoron *oo*, Shuri *ʔaa*, and Hachijo *woa* (cf. Naze *to(o)ra*, Yoron *toora*, Shuri *taara*, Hachijo *toara* < *tawara 'bales'). It is thus obvious that AWA is a native word in various dialects, and not borrowed from other dialects. Foxtail millet is one of the "five grains" (soybeans, wheat, broomcorn, foxtail millet, and rice) and has been cultivated over a long period of time in Japan. The lack of variation may be due to the fact that foxtail millet was (at least once) a part of the very basic vocabulary, like *kusa* 'grass' and *ame* 'rain'. Therefore, this word for foxtail millet has not been replaced with other words.

AWA may be cognate with Proto-Turkic *arpa 'barley' (cf. Mongolian *арвай* 'barley', Manchu *arfa* 'barley, oats').

(NAKAZAWA Kohei and YOKOYAMA Akiko)

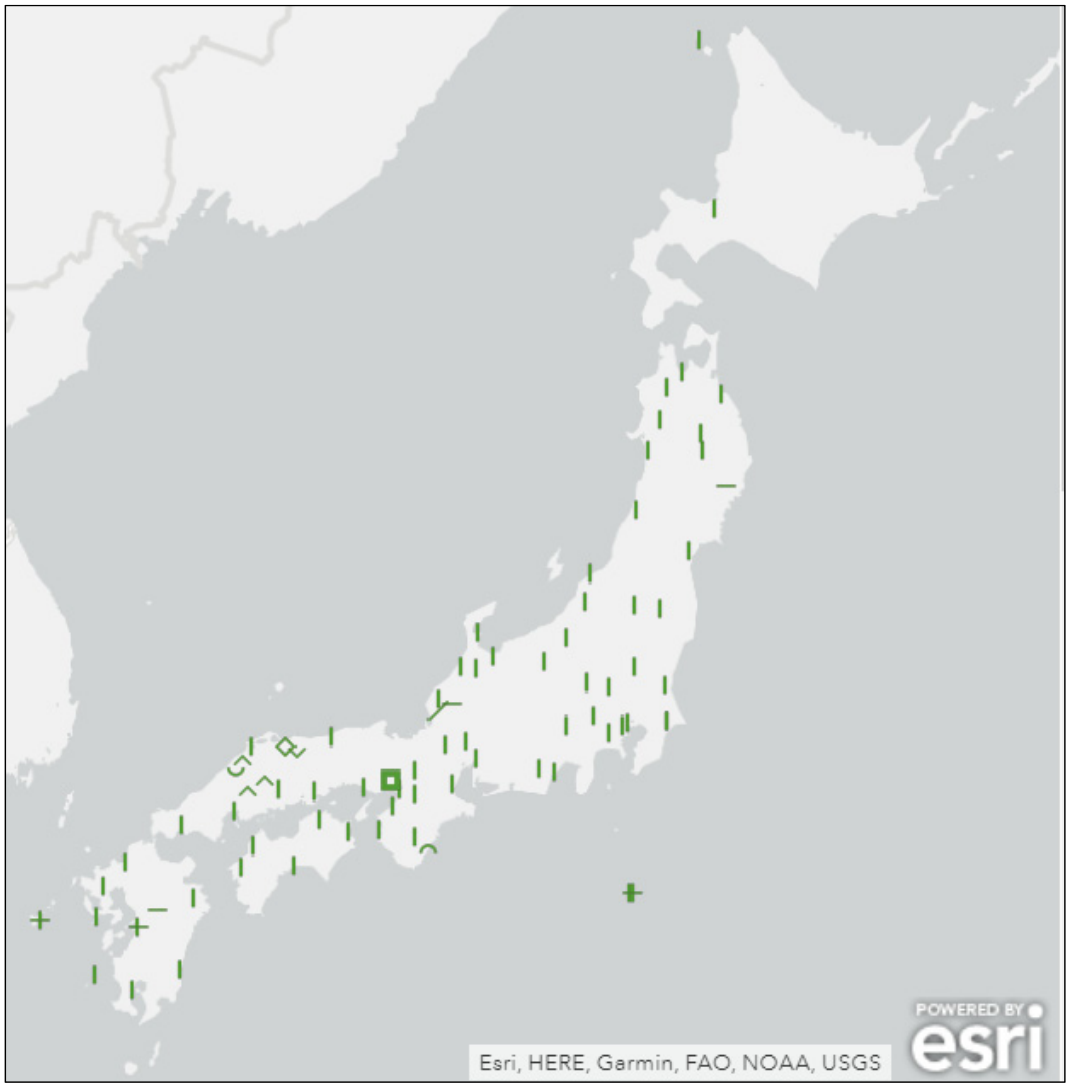


Figure 9.3.1: '*Setaria italica*' in mainland Japan.

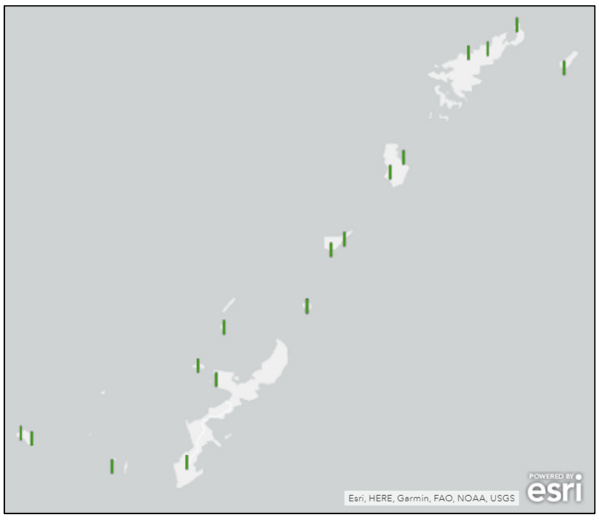


Figure 9.3.2: '*Setaria italica*' in Northern Ryukyu Islands.

Figure 9.3.3: '*Setaria italica*' in Southern Ryukyu Islands.

'*Setaria italica*' in Korean

The following table shows Modern and Middle Korean terms for '*Setaria italica*'. Middle Korean tones are shown in parentheses.

	Modern Korean	Middle Korean (15-16c.)
Setaria italica (アワ)	co	coh (L)

As can be seen in this table, Modern and Middle Korean forms are almost the same except for the fact that Middle Korean form has a final *-h* and distinctive L pitch.

If we go back the history of this languages, we have older forms recorded in the *Jīlín lèishì* (鷄林類事) written in the 12th century and in *Cháoxiǎnguǎn yìyǔ* (朝鮮館訳語) compiled at the beginning of the 15th century.

	鷄林類事	朝鮮館訳語
Setaria italica	粟曰田菩薩	粟米 左色二

In *Jīlín lèishì* (鷄林類事), 'setaria italica' is renderd as compound word '田菩薩', of which the first character seems to corresponds to MK *coh* and the rest is the word for rice (MK *psʌr*).

In *Cháoxiǎnguǎn yìyǔ* (朝鮮館訳語), 'setaria italica' is recorded as '左色二', which seems to correspond to *co-psʌr*.

Generally speaking, earlier forms recorded in these two documents are in accordance with the Middle Korean forms. And the origin of this word is generelly considered to be Sinitic or Sino-Korean morpheme '粟'.

As for modern forms, we use the data recorded in Ogura (1944) and they can be classified as follows:

A: *co* type. *co, coi, copsal, cipi*
B: *sɔsuk* type. B1 *sɔsuk*, B2 *sisik*, B3 *susuk*, B4 *sɨsuk*
C: *susu* type. C1 *susu*, C2 *cansu*
D: other forms. *cɛ*
(N.B. In Ogura (1944), the locations for the basic form *co* is not specified, because it is "used in many places", so that it does not appear in the map.)

A type forms are made by attaching a suffix to the basic morpheme *co*. B type forms are basic form *sɔsuk* and its phonetic varieties. C types are similar to the B type but *susu* is originally a term for a different kind of crop so that treated separately. As for the form in D, its origin and etymological relations to other types are unknown.

The map below was made based on the data recorded in Ogura (1944) . This map shows a clear case of north vs. south distribution. Northern dialects uses A type (*co* and it's varieties) and Southern dialects use B and C type forms. The same pattern was found in the case of the item 'rice'.

(FUKUI Rei)

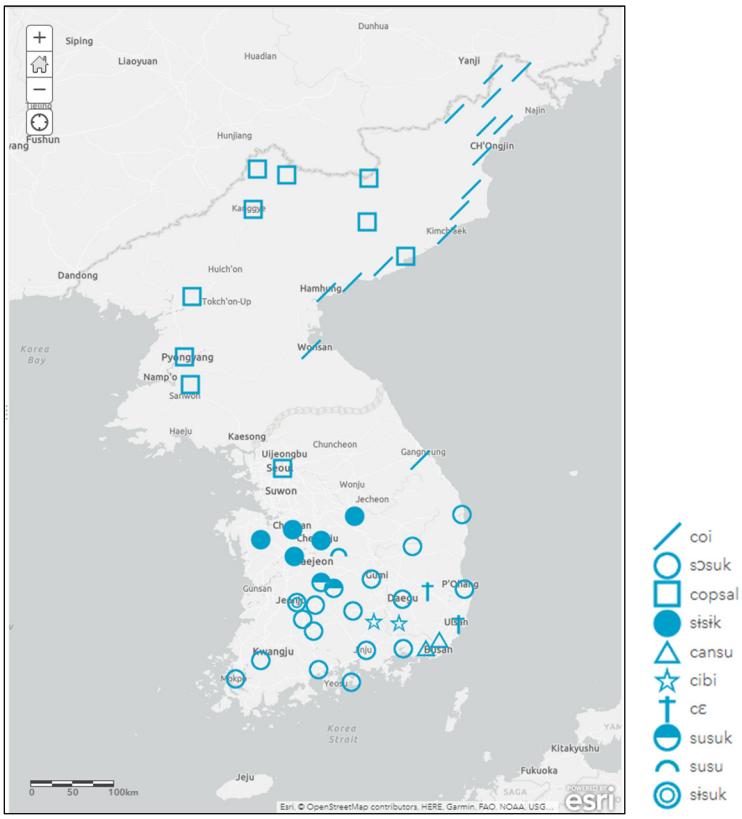


Figure 9.4.1: ‘*Setaria italica*’ in Korean.

‘Foxtail millet’ in Sinitic

Since Sinitic languages distinguish between the plant name and the husked grains of foxtail millet, we made a map for each term.

Figure 9.5.1. shows the plant name for foxtail millet. We classified the word forms into three types by their stems: A *su* 粟, B *gu* 谷, and C *mi* 米.

The A *su* 粟 type is distributed in the south and the B *gu* 谷 type in the north, showing a south–north contraposition. A-3 *sugu* 粟谷 is the contamination form of the A *su* 粟 and B *gu* 谷 types and is distributed in the two types’ contact area. The C *mi* 米 type shows a sporadic distribution in the southwestern part of China.

Figure 9.5.2. shows the word forms for the husked grains of foxtail millet. The classification method is the same as Figure 9.5.1.

The A *su* 粟 type is distributed in the south and the C *mi* 米 type in the north, showing a south–north contraposition.

Table 1

	Plant name	Husked grains
Beijing	谷子[ku^{214} tsɿ^{0}]	小米儿[ɕiau^{35} mir^{214}]
Wuyuan (婺源)	粟[sa^{51}]	粟米[sa^{51} bi^{31}]
Guangzhou	粟[ʃʊk^{5}], 狗尾粟[kɐu^{35} mei^{23} ʃʊk^{5}]	粟[ʃʊk^{5}], 狗尾粟[kɐu^{35} mei^{23} ʃʊk^{5}]

The earliest word form for ‘foxtail millet’ in Old Chinese is *he* 禾, which appeared in oracle bones. *Ji* 稷 is also an exclusive form denoting ‘foxtail millet.’ However, *he* 禾 and *ji* 稷 do not appear in modern dialects except as a compound form, such as A-4 *suhe* 粟禾. A *su* 粟 originally identified the unhusked grains of foxtail millet but eventually became a generic name for the foxtail millet plant. B *gu* 谷 (穀) was a generic form for grains in Old Chinese and diminished in meaning to denote ‘foxtail millet’ after Middle Chinese. C *mi* 米 in Old Chinese means ‘husked grains of foxtail millet’ (Huang 2016, Wang 2011). Table 2 shows the reconstruction of Old Chinese forms by Baxter and Sagart (2014).

Table 2

	OC	MC	referent
he 禾	[ɢ]^{ʕ}oj	hwa	foxtail millet (generic)
ji 稷	tsik	[ts]ək	foxtail millet (generic)
su 粟	[s]ok	sjowk	unhusked grains of foxtail millet
gu 穀	[k]^{ʕ}ok	kuwk	grains (generic)
mi 米	(C.)m^{ʕ}[e]jʔ	mejX	husked foxtail millet

(SUZUKI Fumiki)

A. *su* 粟 type
- A-1 *su* 粟, *suzi* 粟子
- A-2 *sumi* 粟米
- A-3 *sugu* 粟谷
- A-4 *suhe* 粟禾
- A-5 *suji* 粟穄
- A-6 *huangsu* 黄粟
- A-7 *gouweisu* 狗尾粟
- A-8 others: 穭粟, 细粟, 猫尾粟

B. *gu* 谷 type
- B-1 *gu* 谷, *guzi* 谷子
- B-2 *gusui* 谷穗

C. *mi* 米 type
- C-1 *xiaomi* 小米
- C-2 *ximi* 细米

D. others
- D-1 *tai* (稊, 秕), *quanweitai* 犬尾 tai
- D-2 *gouweir* 狗尾儿
- D-3 others: 穄子, 黍仔, 稻子

Figure 9.5.1: ‘Foxtail millet’ (plant) in Sinitic.

A. *su* 粟 type
- A-1 *su* 粟, *suzi* 粟仔
- A-2 *sumi* 粟米, *suzimi* 粟仔米
- A-3 *sugu* 粟谷
- A-4 *huangsu* 黄粟
- A-5 *huangsumi* 黄粟米
- A-6 *gouweisu* 狗尾粟
- A-7 others: 猫尾粟, 粟穄, 粟仁

B. *gu* 谷 type
- B-1 *guzi* 谷子
- B-2 *gumi* 谷米

C. *mi* 米 type
- C-1 *mi* 米
- C-2 *xiaomi* 小米, *xiaomizi* 小米子
- C-3 *ximi* 细米, *huangximi* 黄细米

D. others
- D-1 *shu* 黍, *shuzi* 黍仔
- D-2 *jizi* 穄子, *huangji* 黄穄
- D-3 *tai* (稊, 秿)
- D-4 others: 犬尾秫, 禾子

Figure 9.5.2: ‘Foxtail millet’ (husked grain) in Sinitic.

'*Setaria italica*' in Hmong-Mien

We examine the geographical distribution of the forms of *SETARIA ITALICA*. First, based on comparative evidence, we categorize the forms into several types.

There are ten types in *SETARIA ITALICA*: A: *nau*; B: *tsai*; C: *noŋ*; D: *tshaŋ*; E: *nau tshaŋ*; F: *maŋ*; G: *qo*; H: *pɔ lɔ ka tɔ*; I: *seu*; J: *ɛɔ mi*.

Type A exhibits the widest distribution geographically and phylogenetically in Hmongic, thus suggesting that this type is the most archaic in the Hmongic branch. Since Type C is only observed in North Hmongic, and Type D is only observed in West Hmongic, they must be an innovation in each group. Interestingly, one of the West Hmongic, Dananshan, exhibits both Type A and Type D. This suggests that Type A is the older form, coexisting with an innovative form, Type D. Type B has the widest distribution in Mienic, and must be the oldest form in Mienic. This form likely is a loanword from Chinese 粱 (Ratliff 2010).

(TAGUCHI Yoshihisa)

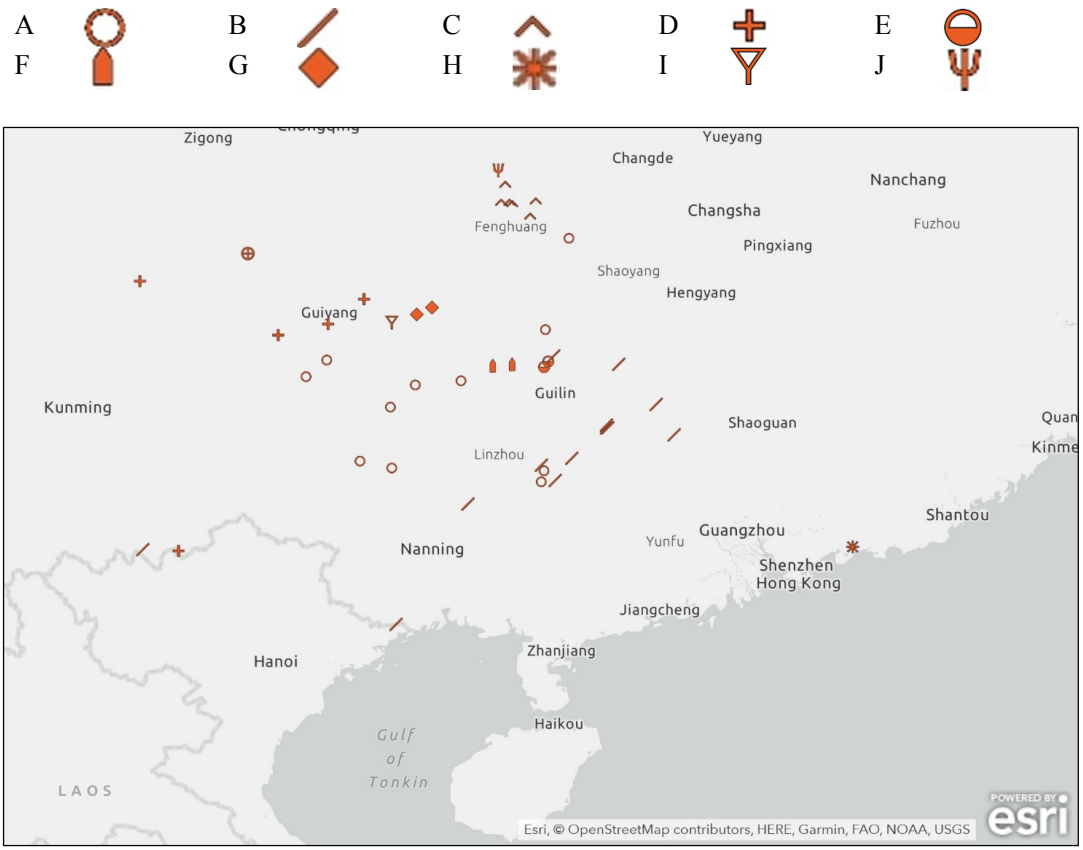


Figure 9.6.1: '*Setaria italica*' in Hmong-Mien.

‘*Setaria italica*’ in Kra-Dai

‘*Setaria italica*’ has word forms of types A through J. Of these, Type A has a very large number of variants. Type A can be divided into four subtypes according to initial consonants.

Type A is a word form group distributed over a wide area, including Hainan Island. It is classified into four subtypes from A1 to A4 according to initial consonants. Figure 1 is distribution of the Type A subtypes. In this figure, the subtypes are represented by different symbols. Tomita (1997) contains *(khâao) fâaŋ* as a word referring to *Setaria italica*, and Type A is the word form corresponding to this word. The Type A initial consonants have a phonological correspondence with the Proto-Tai **f-* of Li (1977). Proto-Tai **f-* is represented as *ph-* in Central Tai. Type A also has many word forms in which ‘rice’ appears as the first element.

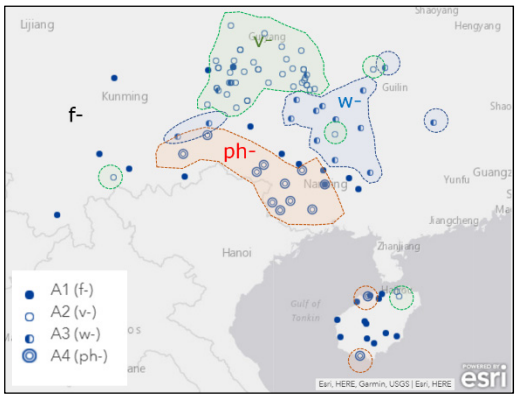


Figure 1: Type A of ‘*Setaria italica*’.

Type B is a word form with initial consonants *γ-* or *h-*, and is found in some Buyi and Zhuang languages, all of which are found in Northern Tai. The phonetic correspondence is similar to that of Li's (1977) Proto-Tai **γ-*, but it is not certain because the Type B word form does not appear in Central Tai or Southwestern Tai. Also, according to Li (1977), in Buyi the **γ-* in Proto-Tai is represented by *v-* in front of the rounded vowel. There are many points of Buyi language classified as A2 subtype, which is an initial consonant *v-*. So, word forms that are originally the same as type B may be included in A2.

As for Types C through F, these word forms are distributed in the Kra branch. The word structure of the multi-syllabic word form is not clear, but the first element of type E would correspond to “rice” in Gelao, which belongs to the Kra branch.

Types G and H are distributed on Hainan Island. Type I is found in Tai Na and Tai Lue, Yunnan. The word structure is ‘rice + tail + dog’, similar to the English name 'foxtail millet' and the Chinese word 狗尾粟 of *Setaria italica*. The J type is thought to be a reading of the Chinese 狗尾.

(TOMITA Aika)

A. *Fang* type

- A-1 *fang* type
  hau^{4} fiəŋ^{3}, hau^{4} fəːŋ^{3}, hau^{4} fɯːŋ^{3}, hau^{4} fɯəŋ^{3}, hau^{4} fiaŋ^{3}, hau^{4} fɯaŋ^{3}, hau^{4} fuːŋ^{3}, ʔo^{3}fuŋ^{3}, fuŋ^{3}, xau^{3}faːŋ^{3}, khɐu^{3}faŋ^{3}, khau^{3}faːŋ^{3}, khau^{3}faŋ^{3}, feːŋ^{3}, fiaŋ^{3}, fhiaŋ^{3}, feːŋ^{3}ʔuːt^{7}, feːŋ^{3}tshut^{7}-pha^{4}, feːŋ^{3} mut^{7}, fɛŋ^{3}, faŋ^{3}
- A-2 *vang* type
  hau^{4} vuːŋ^{4}, hau^{4} vəːŋ^{3}, hɑu^{4}vɯːŋ^{3}, vɑŋ^{3}, ɦau^{4}vɯːŋ^{3}, ɣau^{4}vuŋ^{3}, vuːŋ^{3}, viŋ^{3}, vuŋ^{3}, hau^{4}vuŋ^{3}, vɯːŋ^{3}, khau^{1}vaŋ^{1}, əu^{3}viŋ^{6}, həu^{3}viːŋ^{3}, ɣəu^{4}vjɛŋ^{4}, vaŋ^{4}
- A-3 w*ang* type
  hau^{4} wɯəŋ^{3}, ɣau^{4} wuŋ^{3}, gau^{4} wuːŋ^{3}, hau^{4} wəŋ^{3}, hau^{4} wuəŋ^{3}, hau^{4} wuːŋ^{3}, hau^{4} wiːŋ^{1}, wuːŋ^{3}, wɑŋ^{3}, ɣau^{4}wɑŋ^{3}
- A-4 *phang* type
  hau^{4} phluːŋ^{3}, hau^{4} phlɯːŋ^{3}, phyːŋ^{3} luːŋ^{2}, hau^{4} phiːŋ^{3}, khau^{3} phaːŋ^{3}, phaːŋ^{3}, pʰaŋ^{3}

B. *yung* type

- B-1 *yung* type
  ɣau^{4} ɣuːŋ^{3}
- B-2 *hung* type
  hoŋ^{3}, huŋ^{3}, huːŋ^{3}
- C. *la (u)* type
  ljou^{35}, la^{312}, le^{31}i^{55}, haːŋ^{24}la^{33}, lau^{55}tshan^{13}pau^{33}
- D. *Sai* type
  łai^{5}, hu^{4} lui^{6} łai^{5}, sai^{5}, tshai^{31}
- E. *mung tso* type
  məˀ^{53}tso^{31}, mɯŋ^{42}tɕhi^{35}
- F. *qa mia* type
  qa^{33}miɑ^{33}, qym^{3}pɛ^{22}
- G. *bang* type
  ɓaŋ^{4}
- H. *pua* type
  pɯa^{1}
- I. *xau hang ma* (rice+tail+dog) type
  xau^{3} haːŋ^{1} ma^{1}, xɑu^{3}haŋ^{1}ma^{1}
- J. *kou me* (狗尾) type
  kou$^{31\text{-}11}$ me^{23}

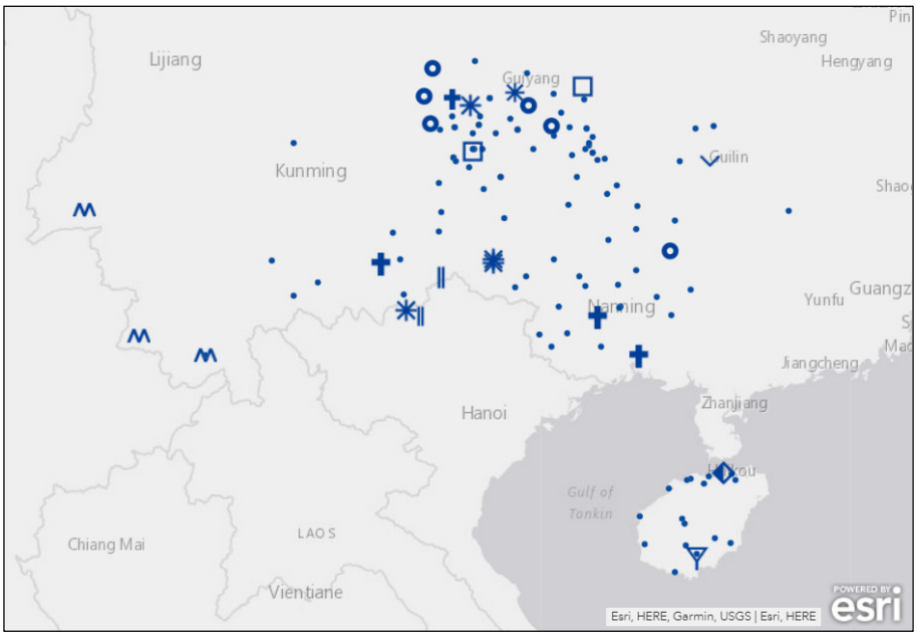


Figure 9.7.1: ‘*Setaria italica*’ in Kra-Dai.

‘Foxtail millet’ in Tibeto-Burman

Among 71 collected items of the word form for ‘foxtail millet’ (*Setaria italica*) in TB languages and varieties, we find five main types. Types A, B and E have proto-level etyma. Types F, G and L are related to either semantic change or borrowing.

Type A includes forms derived from PTB *khrəy ‘MILLET / RICE’, mainly attested in Tibetic languages.

Type B includes forms derived from PLB *tsap ‘MILLET’, mainly attested in Lolo-Burmese languages.

Type C includes various forms with a /m/-initial, attested in the eastern area, scattered.

Type D includes various forms with a /n/-initial, attested in some dialects of Bai.

Type E includes forms derived from Proto-Tani *jak ‘MILLET (foxtail)’, attested in the languages of Arunachal Pradesh and its surroundings.

Type F includes languages that use a form common to either ‘broomcorn millet’ or ‘barnyard millet’. Due to the limitation of the data, we do not provide further analysis of this lexical commonality.

Type G includes word forms that have undergone a large semantic change. The change depends on the language. We find the following cases: a form derived from *drus ma* ‘polished grain’ (Baima); *rgáɣ* (Wobzi Khroskyabs), used for ‘wheat’ in other languages; and *’bru’u* (Zhollam) denoting ‘small crop’ literally. The last one is considered as a calque of a Sinitic word *xiaomi*.

Type H is a category using a Sinitic loanword. Two lexical forms are attested: *xiaomi* and *guzi*.

Type X contains various independent stems. The word forms, which are of unclear origin, appear not to be related to each other.

Note that it is possible to distinguish ‘plant’ from ‘grain’, as happens in Mawo Rma, in two forms, Types F and H, where the latter is supposed to denote ‘millet grain’.

Types A and B are often used for ‘broomcorn millet’ (*Panicum miliaceum*) or ‘barnyard millet’ (*Echinochloa* species) in several languages. This suggests that these languages originally possessed a single term for these plants. However, the data from the secondary sources complicate the analysis, as no rigorous classification has been done for the word form for ‘foxtail millet’; the sources often describe the target term as ‘*xiaomi*’ in Chinese and ‘millet’ in English. These questions then influence our view of Type F as well.

Referring to the geographical distribution in south-western China, Type B, found peripherally, can be identified as an earlier form than Types C and D. However, all proposals for lexical history remain provisional due to the insufficient size of the dataset.

(SUZUKI Hiroyuki, EBIHARA Shiho, IWASA Kazue, KURABE Keita, SHIRAI Satoko)

■ A. PTB *khrəy: *ʈʰe, tsʰe, kʰrə, kʰə*.
□ B. PLB *tsap: tshe, tsaq, tshei, tɕa*^{35}.
\ C. *se*^{31}*pɛ*^{42}*me*^{33}, *muŋ*^{55}, *món, jĩ*^{33}*mĩ*^{53}.
/ D. *no*^{35}, *nou*^{35}, *ŋo*^{31}*mĩ*^{31}, *nɛɹ*.
● E. *aŋ jie, ja*.
⬭ F. terms for other crops.
⬯ G. semantic changes.
◁ H. Chinese loan.
▬ X. *ran, wəta, ḍusi, pyáuɴ*.

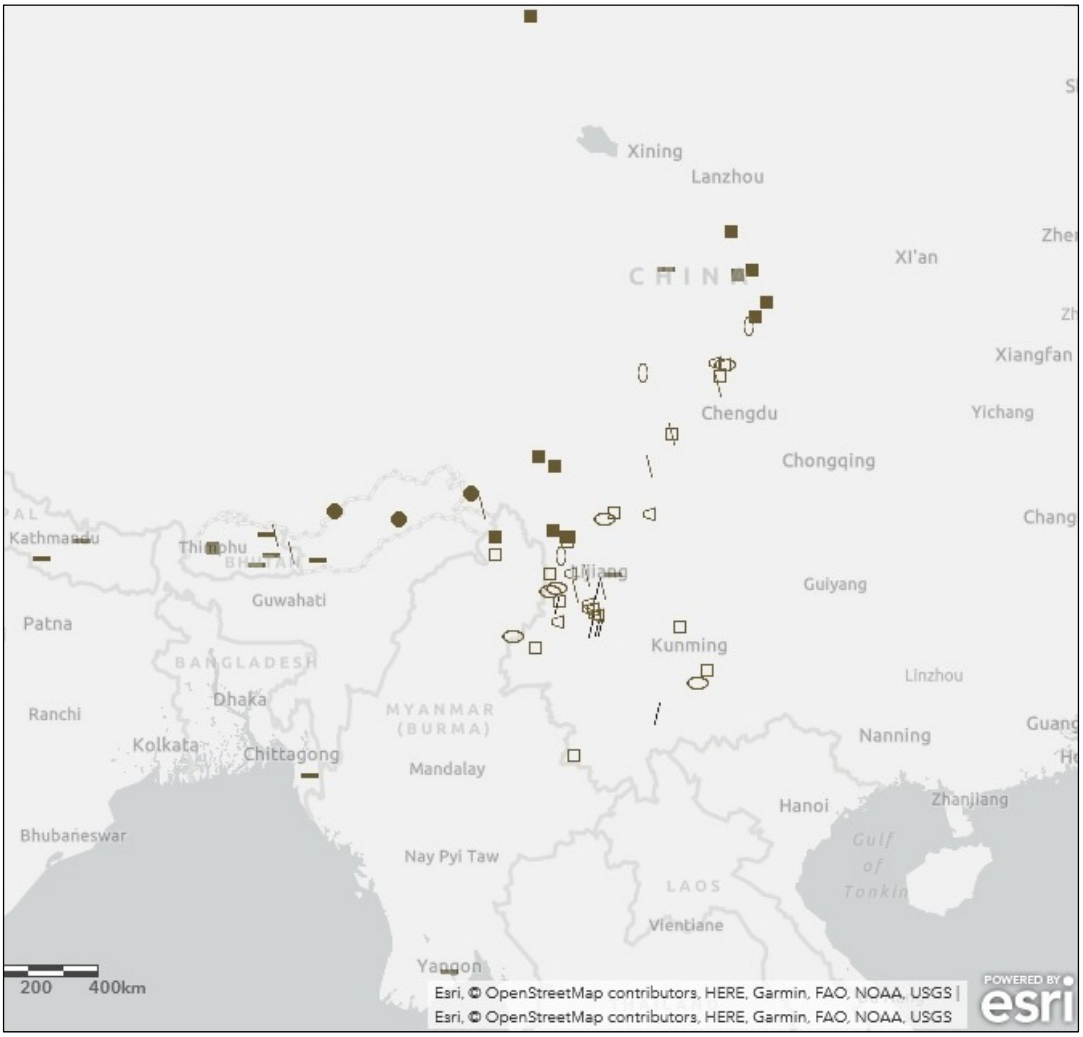


Figure 9.8.1: ‘Foxtail millet’ in Tibeto-Burman.

‘Foxtail millet’ in Mongolic and Turkic

For foxtail millet, the Mongolic languages use *narim* in easternmost regions and both *narim* and *xonog* in other areas. Some Turkic languages in Central Asia use the form *konok*, which is a cognate with Mongolic *xonog* (< *konog*).

Some languages have more than one term for the crop: *narim*, *narimīn čagān* (*čagān* ‘white’), *narim budā* (*budā* ‘grain’) in Chakhar Mongol, *dügü* and *konax* in Salar, *sök*, *tüge* and *terik* in Modern Uighur, *šigī sokba* (*šigī* ‘thin’), *ʂitɢuzuruk šige sokba* (*ʂitɢuzuruk* ‘dog tail’), *sarəɣ sokba* (*sarəɣ* ‘yellow’) in Sarïg Yughur, etc.

The Salar form *dügü* and the Uighur form *terik* can denote both foxtail millet and broomcorn millet. This may mean that the crops are not always clearly distinguished.

(SAITÔ Yoshio)

A. *narim* type


næram, nɛrəm, narim, nærjam, narēm, nɛr

B. *xonog* type


xonog, ɢono

C. Compound words


narimīn čagān


xonog narim


nɛr badā, narim budā, narim budā


xonog budā

šar budā, šir budā

D. *konok* type


konok, konak, konax

E. *tüge* type


tüge


dügü

F. *sokba* type

šigī sokba


ʂitɢuzuruk šige sokba


sarəɣ sokba

sök

G. *terik*


terik

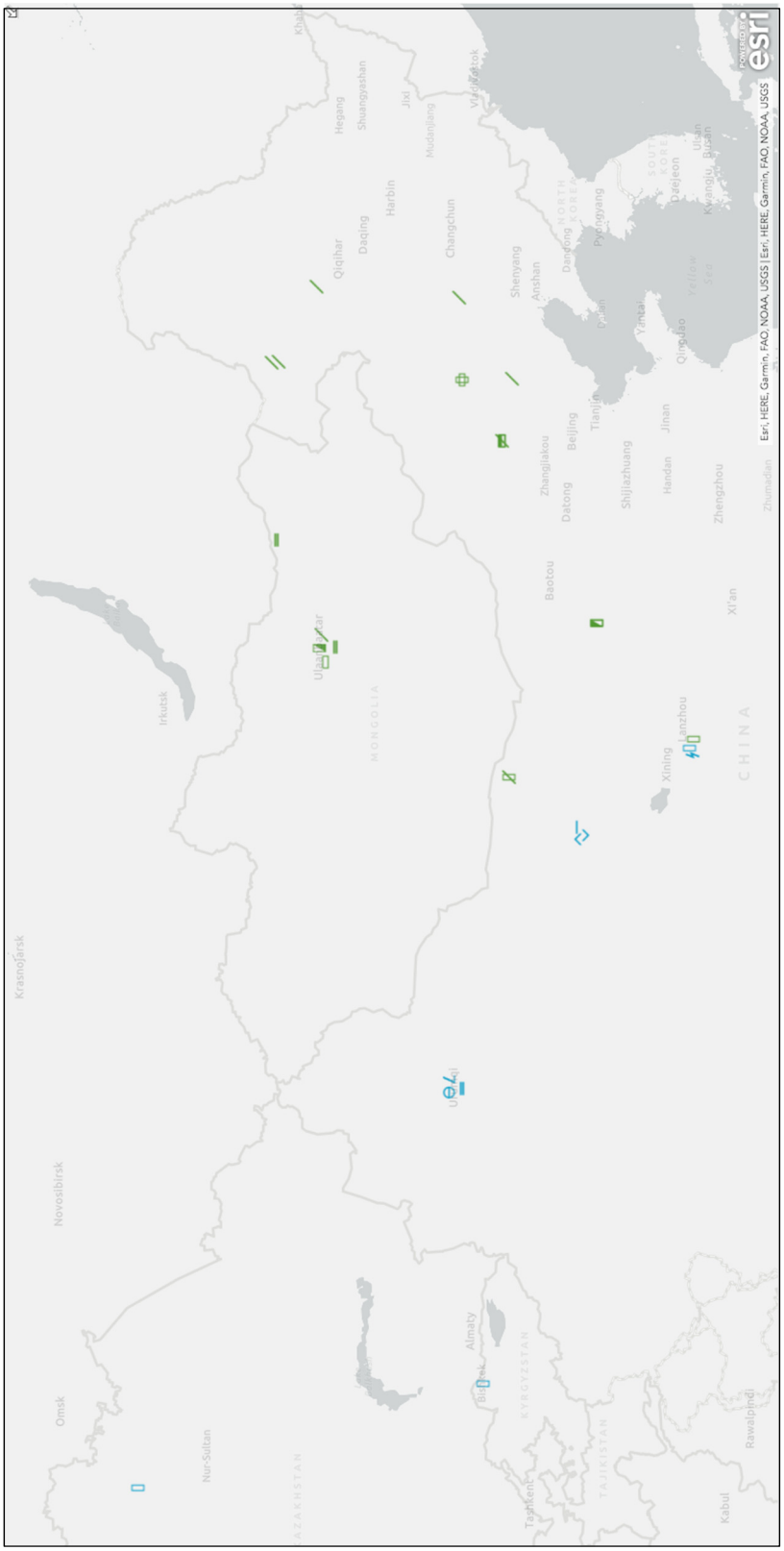


Figure 9.13.1: ‘Foxtail millet’ in Mongolic and Turkic.

‘Foxtail millet (*Setaria italica*)’ in South Asia

I describe the languages of the Indo-Aryan (IA) and Nuristani branches, and Burushaski (isolate) in South Asia.

There are a major category of word forms for ‘foxtail millet’ – A) *ṭ/kaṅgunī* – and three minor categories B) *kāśa*, C) *cīna(ka)*, and D) *čha*.

Indo-Aryan languages south of latitude 34° N on the mainland employ Type A. Type B can be seen in Nuristani languages. Type C is detected in geographically peripheral inland IA languages. Type D is only for Burushaski lects.

The most major type is *ṭ/kaṅgunī*, which is seen in ‘Proso millet in South Asia’. This type is derived from Sanskrit *kaŋgunī* कङ्गुनी ‘*Setaria italica*, *Celastrus paniculatus*’. Most modern forms have the initial consonant /k/ as well as the Sanskrit term above, while Oriya *ṭaŋgɔṇɔ* ଟାଙ୍ଗଣ, Gaya Magahi *ṭãgun*, and Southwestern Magahi *ṭãguni* retain /ṭ/ in initial similarly to the reconstructed Sanskrit form **ṭaŋgunī*. The Assamese term *kɔni dhan* কণী ধান is a composition of *kaṅgunī* and *dhānya* meaning literally ‘foxtail millet grain’.

The *kāśa* type appears only in some Nuristani languages, and therefore the distribution is narrow, only in the Nuristan province of Afghanistan. The Sanskrit form *kā́śa* काश means ‘a grass used for mats, *Saccharum spontaneum* (wild sugarcane, ワセオバナ)’. Turner (1966) points out the similarity between *kāśa* forms and the *karaz* forms, I mention in the chapter ‘Proso millet in South Asia’ and Khowar *grac* ‘foxtail millet’, and suggests all of them have the identical etymon, that is, *grāsa* ‘mouthful’, which has some descendants with crop referents such as Woṭapuri *gāsə* ‘cooked rice’. However the initial /gr/ of some words regularly remain the cluster including a liquid sound in Nuristani langauge; For example, derived from the Sanskrit *grā́ma* ग्राम ‘troop, village’, Western Kati *grām* (Konow 1913: 123) and Ashkuňu *glam* have /gr/ and /gl/ clusters as well as old Khowar *gram* has /gr/. And so Western Kati *kacɔ* ‘foxtail millet’ and Ashkuňu *kāc* ‘id.’ cannot be of the same etymon of Khowar *grac* ‘id.’ in principle.

The third type *cīna(ka)* is of descendant forms derived from Sanskrit *čīna* चीन ‘*Panicum miliaceum*’. This is also found in some terms for ‘proso millet’. For more information, see also the chapter.

Type D of *čha* is for Burushaski. It shows the regular corresponding of consonants, the aspirated fricative /čh/ in Eastern Burushaski versus the unaspirated /č/ in Western Burushaski.

Besides them, languages in South Asia have some more words for ‘foxtail millet’. Marathi *rāḷē* राळे and Urdu *rālā* رالا are derived from Sanskrit *rāhala* राहल ‘a kind of pulse’. *piŋga* in Kashmiri is derived from Sanskrit *priyáŋgu* प्रियंगु ‘*Panicum italicum*’, which originally means the plant, that is ‘foxtail millet’, but modern IA languages rather use cognate terms for ‘proso millet’. Dhivehi *kudibat* ކުޑިބަތް literally means ‘small rice’. And the last, Kalasha *šilī* is inherited from Sanskrit *sītiya* सीतिय ‘ploughed, corn’.

(YOSHIOKA Noboru)

A. *ṭ/kaṅgunī* type (17) △
kaŋgnī, kaŋgṇī, ṭaŋgɔṇɔ, kãgnī, kāgunī, ṭã̃guni, ṭã̃gun, kaŋgu, kaŋg, kã̃g, kaun, kākan, kiriŋghu
[+*dhānya*]
kɔni dhan

B. *kāśa* type (4) ▷
kacɔ, kacó, kẵc

C. *cīna(ka)* type (3) ⛳
čina, čīn, čīṇō

D. *čha* type (3) ❙
čha, ča

E. others
[*rāhala* (2)] rālā, rāḷē; [*dhānya* + X] tana-hāl; [*karaz*] grac; [*priyaṅgu*] piŋga; kudibat; šilí

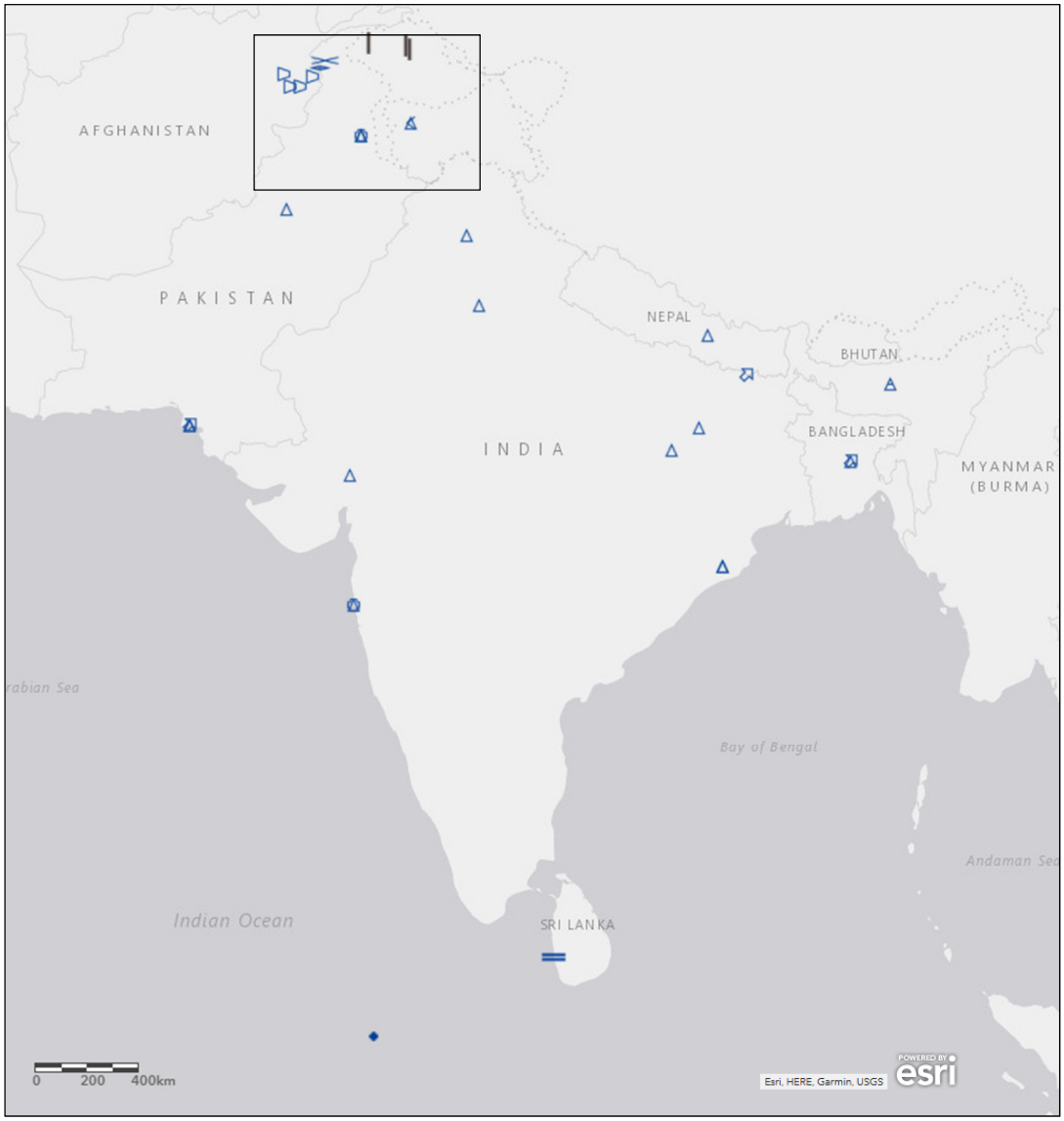


Figure 9.14.1: ‘Foxtail millet’ in SA: Indo-Aryan, Nuristani (both in navy blue), and Burushaski (in black).

‘Foxtail millet’ in Dravidian

Foxtail millet is one of the earliest crops attested in the Southern Neolithic sites in the dry zone of peninsular India. Five etyma for this crop listed in DEDR may reflect the long history of cultivation.

Reflexes of the etymon #2163 are distributed across peninsular India, although its reflexes are obsolete in SDr literary languages. The other four etyma are confined to a particular clade, #3712 and #3265 in SDr, #195 in CDr, #379 in SCDr.

The etymon #3265 has reflexes meaning spike and ear of any grain in Kannada and Kota, and its reflex in Tamil refer to several species of millet. Reflexes of #379 in Kannada and Telugu refer to a different species of millet, *Paspalum scrobiculatum.* Irula *paṇḍi*, recorded by Kamil Zvelebil, is of unknown origin.

(KODAMA Nozomi)

tinai/nuvaṇai/kural,navaṇi/kural
kural, korralu, koṟale, korle
korly, koyḷ
koṟeŋ
kueri/ārka
ārgu, ārku
ayk, aykil
paṇḍi
tina

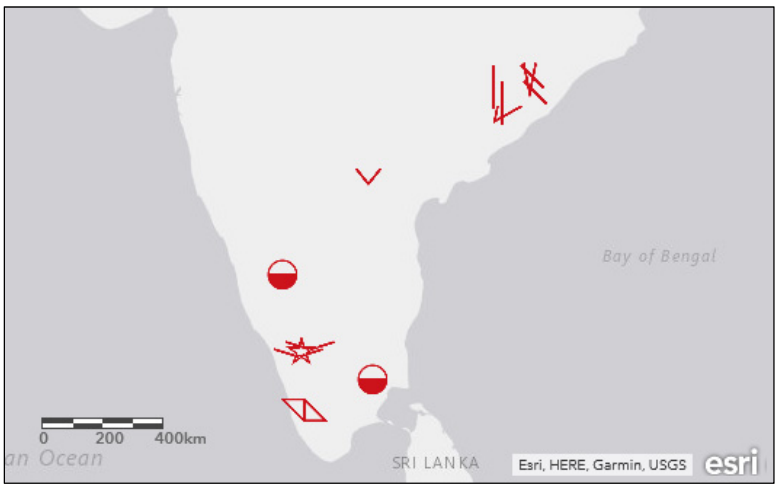


Figure 9.15.1: ‘Foxtail millet’ in Dravidian.

‘Foxtail millet’ in Iranian

Foxtail millet is not as popular as other grains such as rice and wheat. Geographically, the languages which have the terms for ‘foxtail millet’ concentrate in the northeastern region.

The terms ‘foxtail millet’ are divided into three types. Type A, whose forms seem to be descendants of Proto-Iranian **gawarsa-*, is overwhelmingly dominant within the Iranian world. Type B is observed only in Yazgulami. This type of word is ambiguous. Morgenstierne (1974: 98) describes this word as ‘Foxtail millet’, whereas Ėdel’man (1971: 306) counts it as ‘Broomcorn millet’. See the section on ‘Broomcorn millet’ in the present volume.

Type C, a loanword from Uzbek *qo‘noq* ‘millet’, is found in Tajik, which has been influenced by the Turkic language.

(IWASAKI Takamasa)

A: **gawarsa-* — *gāwars, gâš, ɣoxˇt, dʒäwāwsk, gâvars, dʒuwōxˇč, govars*
B: *xarban* — *xarban*
C: *qūnoq* — *qūnoq*

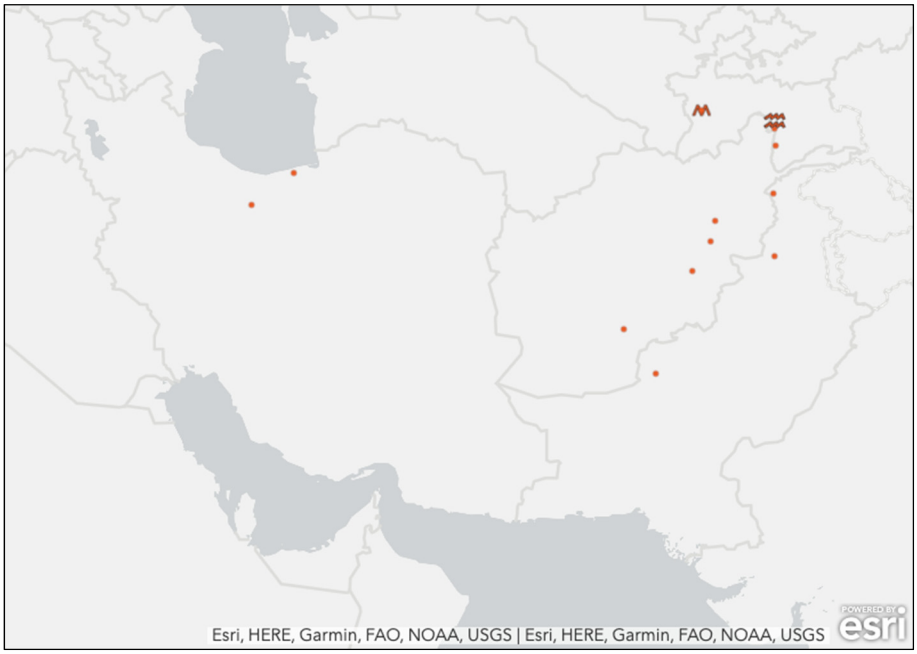


Figure 9.16.1: ‘Foxtail millet’ in Iranian.

Chapter X

Barnyard millet

‘Barnyard millet’ in Asian and African languages

The English term ‘barnyard millet’ principally denotes two *Echinochloa* species: *Echinochloa esculenta* ‘Japanese barnyard millet’ and *Echinochloa frumentacea* ‘Indian barnyard millet’. The former originated from and is widely cultivated in East Asia, while the latter is grown in South Asia but is uncertain in its domestication period.

Table 1 shows the principal forms in our data: Ainu, Japonic, Korean, Sinitic, Kra-Dai, Tibeto-Burman, Indo-Aryan, and Dravidian. Data from languages in Africa are unavailable. Depending on the source, we cannot determine whether the entry ‘barnyard millet’ in the sources was recorded as ‘Japanese barnyard millet’ or ‘Indian barnyard millet’millet’; nevertheless, we ignored the difference between them when we produced and analysed the linguistic maps.

Table 1: Main word forms for ‘barnyard millet’.

Languages	Word forms
Ainu	*piyápa* *ayús amam*
Japonic	*hie* 稗 *karasumugi* 烏麦 *karasumai* 烏米
Korean	*p^{h}i* 稗
Sinitic	*bai* 稗 *you* 莠
Kra-Dai	*fang* type *ɣung* type *vəi* type
Tibeto-Burman	PTB **khrəy* PTB **C-lu-k* *yó* *pe.ɹ^{42}* (< Chinese *bai* 稗) *vi^{33}*
Indo-Aryan	*śyāmāka* type *kōdrava* type *cīna(ka)* type *varī* type
Dravidian	*kutiraivāli* *ūdara*

Noticeably, the Chinese character for ‘barnyard millet’ in Japonic is from that of Sinitic *bai*, which is distributed throughout the Sinitic-speaking area. The Japonic lexical form *hie* may be a cognate with the pronunciation of Sinitic *bai*. Korean also uses this Sinitic form *p^{h}i*.

(SUZUKI Hiroyuki)

‘Barnyard millet’ in Ainu

In the Ainu oral literature, *amám* ‘grains’ were stolen by the progenitor of the Ainu from the place of either *kamúy mosir* ‘the land of gods,’ or *sísam mosir* ‘the mainland of Japan’ (Yoneda 1995). In fact, barnyard millet (*Panicum Crusgalli L. var. frumentaceum Trin*) is called *piyápa* and *ayúsamam ~ ayusámam* (lit. ‘thorny grain’), and one of the subspecies (*Panicum Crusgalli L. var. submutica mey.*) is called *sitá amam ~ setá amam* (lit. ‘dog grain’, or ‘unworthy grain’) (Chiri 1976[1953]; Hattori 1964; and Sawai 1989). In “Moshiogusa 藻汐草” (1792), the item of *hie* 稗 ‘barnyard millet’ contains the three following terms: “アイラシアマム (probably mistaken for アイウシアマム)” [aiuʃi amamu], “ツナシアマム” [tsunaʃi amamu], “ビヤバ” [bijaba], for which were estimated the word forms *ayus amam*, *tunas amam* (lit. ‘fast grain’), and *piyapa* (see also ‘Foxtail millet’ in Ainu).

(FUKAZAWA Mika)

A. *piyápa* type
piyápa ~ piyapa
B. *ayús amam* type
ayús amam ~ ayus ámam
C. *tóyta amam*
D. *mancuu amam*

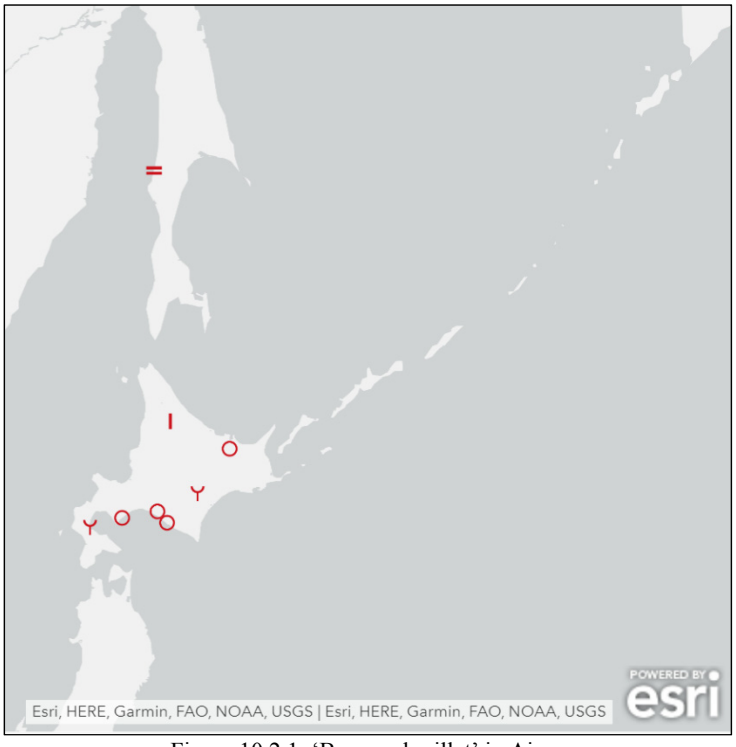


Figure 10.2.1: ‘Barnyard millet’ in Ainu.

'*Echinochloa frumentacea*' in Japonic

The major type of *Echinochloa frumentacea* or barnyard grass in mainland Japan is HIE (*hie*, *hije*, *hee*, *fie*, *hjee*, *hii*, *fee*, *he*, *hje*, *pii*, *fii*, …), while the forms in the Ryukyus vary, including KARASUMUGI (*garaʃi nu muɴ*, *garasamuɴ*) 'crow's wheat' and KARASUMAI (*garasimai*) 'crow's rice'.

The oldest attested form of barnyard grass in Japanese is *piye*, which may also be the proto-Japonic form. In addition, it seems that the Ryukyuan people did not use barnyard grass actively because there were few responses to questions on the maps. Some HIE forms in the Ryukyus must be borrowed from Japanese *hie*, since they do not reflect regular sound changes. Ryukyuan KARASUMUGI and KARASUMAI literally mean 'wheat/rice that crows eat' (cf. *karasumugi* 'crow's wheat' for oats in Japanese). These forms are considered newer than HIE.

HIE may be cognate with Chinese *bài* 'millet'.

(NAKAZAWA Kohei and YOKOYAMA Akiko)

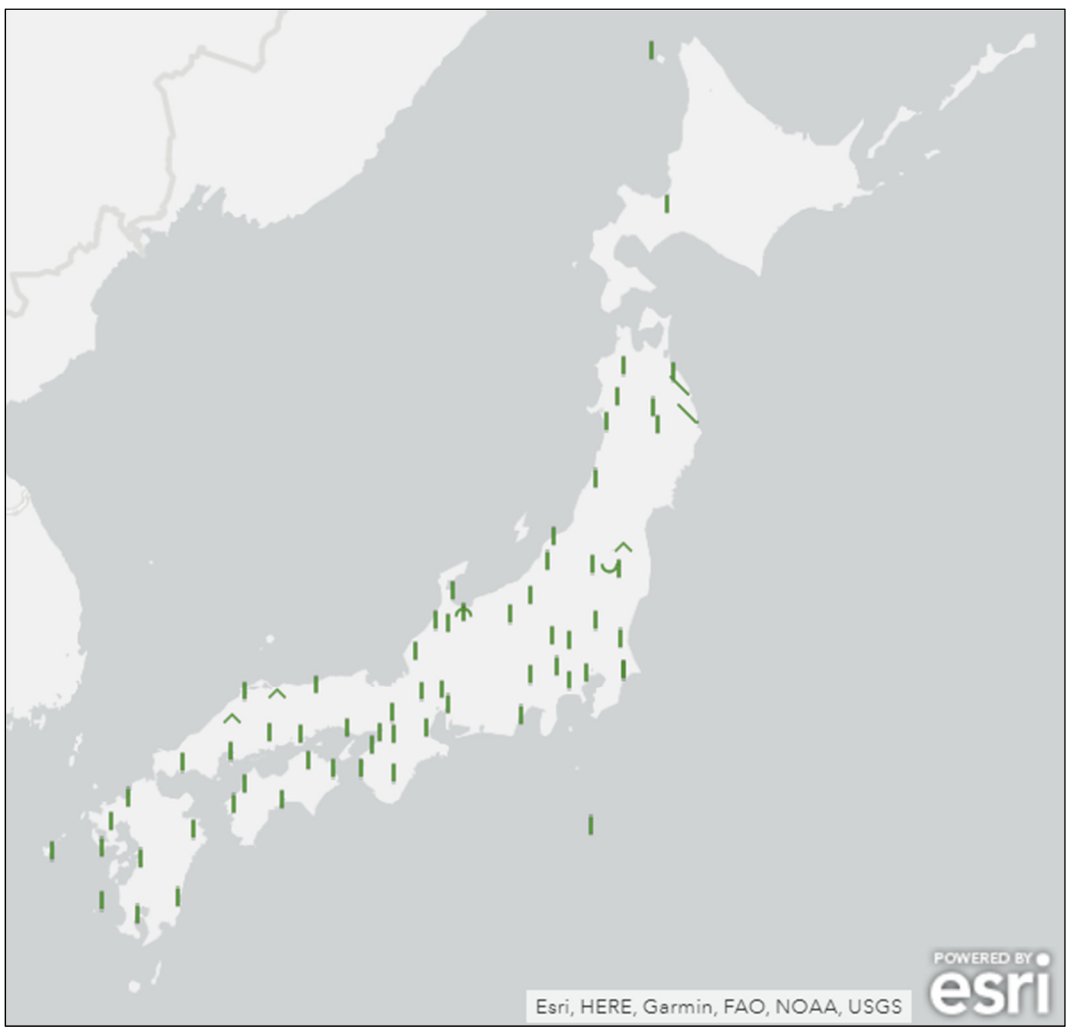


Figure 10.3.1: '*Echinochloa frumentacea*' in mainland Japan.

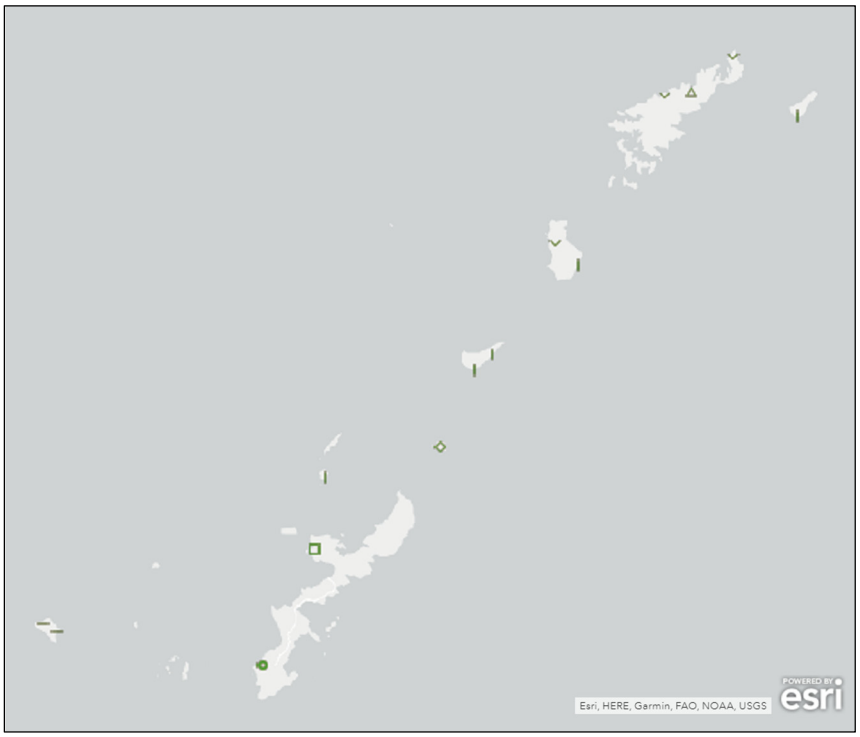


Figure 10.3.2: '*Echinochloa frumentacea*' in Northern Ryukyu Islands.

Figure 10.3.3: '*Echinochloa frumentacea*' in Southern Ryukyu Islands.

‘*Echinochloa frumentacea*’ in Korean

The following table shows Modern and Middle Korean terms for ‘echinochloa frumentacea’. Middle Korean tone is shown in parentheses.

	Modern Korean	Middle Korean (15-16c.)
Echinochloa frumentacea (ヒエ)	p^hi	p^hi (H)

As can be seen in this table, Modern and Middle Korean forms are almost the same except for the fact that Middle Korean form has a distinctive pitch (high tone).

The origin of the word for ‘echinochloa frumentacea’ is generally believed to be the Sino-Korean morpheme ‘稗’.

Dialectal differences are not so great. Only a few words made of *p^hi* plus a suffix, such as *p^hi-madɨi*, *p^hi-nadʒi* are found in northern dialects but these marginal forms are omitted in the map.

(FUKUI Rei)

○ p^hi

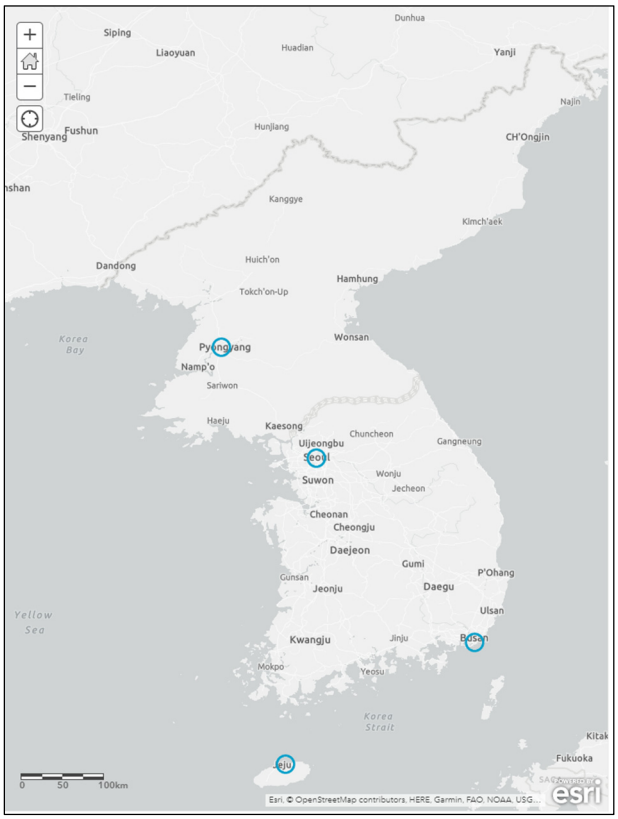


Figure 10.4.1: ‘*Echinochloa frumentacea*’ in Korean.

‘Barnyard millet’ in Sinitic

We classified the word forms for ‘barnyard millet’ into three types by stem: A *bai* 稗, B *you* 莠 and C others.

The A-1 monosyllabic *bai* 稗 is mainly distributed in the Fujian, Guangdong, and Taiwan areas. The word form for ‘barnyard millet’ in Old Chinese is the monosyllabic *bai* 稗, reconstructed as *C.[b]ˤre-s (OC) > beaH (Middle Chinese) by Baxter and Sagart (2014).

The most frequent form among them all is A-2 *baizi* 稗子, distributed in the whole area of Sinitic languages. *Zi* 子 can be interpreted as a diminutive suffix, but sometimes has the concrete meaning of ‘seed.’ Some dialects use A-2 *baizi* 稗子 to mean ‘seeds of barnyard millet.’ (see Table 1)

Table 1

	Plant name	Seeds
Nantong (南通)	稗[pʰa^{213}]	稗子[pʰa^{213} tsɿ^{0}]
Hangzhou (杭州)	稗草[bo^{13} tsʰɔ^{53}]	稗子[bo^{13} tsɿ^{53}

There are special terms for when barnyard millet is viewed as a weed that grows in rice fields rather than as a crop, such as A-3 *baicao* 稗草 (‘barnyard grass’) and A-4 *maobai* 毛稗 (‘fuzzy barnyard millet’).

The other stems also originated from weeds that grow in fields. For example, the B *you* 莠 type comes from the word forms for ‘green bristle-grass’ and is distributed throughout the Shandong province. C-1 *caozi* 草子, meaning ‘grass,’ is distributed in the Jiangxi and Guangdong provinces. C-2 *canzi* 穇子, distributed in the Shandong province, is shifted from word forms for ‘Finger millet.’

(SUZUKI Fumiki)

A. *bai* 稗 type
- A-1 *bai* 稗 (monosyllabic)
- A-2 *baizi* 稗子
- A-3 *baicao* 稗草
- A-4 *maobai* 毛稗
- A-5 *shuibaizi* 水稗子
- A-6 *daobaizi* 稻稗子

B. *you* 莠 type
- B-1 *youzi* 莠子
- B-2 *yougu* 莠谷, *guyouzi* 谷莠子
- B-3 *daoyouzi* 稻莠子

C. others
- C-1 *caozi* 草子
- C-2 *canzi* 稺子, *maocanzi* 毛稺子
- C-3 others: 野禾, 假秧, 茅稻

Figure 10.5.1: ‘Barnyard millet’ in Sinitic.

'*Echinochloa esculenta*' in Kra-Dai

The word forms of '*Echinochloa esculenta*' can be divided into Types A through F.

The word form of Type A is very similar to Type A of *Setaria italica*, but notice that the number of variants is less than in *Setaria italica*. Type A appears to have a similar word form to Type A of *Setaria italica*, but differs in several points. First, *Echinochloa esculenta* lacks the word form of the initial consonant *ph-* and instead has a fairly widespread distribution of the word form *w-*. The points in Central Tai that had the initial consonant *ph-* in Setaria italica are The word form of the initial consonant *w-*. Second, there is a very limited distribution of word forms beginning with *f-*, which is dominant in *Setaria italica*. And *Echinochloa esculenta* is relatively homogeneous with few variants of the Type A word form. In addition, the tone number is 1, which differs from Type A of *Setaria italica*, which has tone 3.

Type B also has a similar word form to Type B of *Setaria italica*, but its distribution is more southward among the Northern Tai points, and it has a different tone.

Types C through F are distinctive forms, apparently different from the word forms of Setaria italica, and they are found in the Kra branch and on the Hainan Island.

(TOMITA Aika)

A. *Fang* type

● A-1 *fang* type
faŋ[53], mə[53]faŋ[24]

○ A-2 *vang* type
vaŋ[1], vɑŋ[1], vɐŋ[1], vaŋ[24], bɔt[7]vɔŋ[1], vuaŋ[55]i[44]

◐ A-3 *wang* type
waŋ[1], ɕe[6] waŋ[1], ko[1] waŋ[1]

B. *ɣung* type

▲ B-1 *ɣ-* type
ɣau[3] ɣuŋ[1]

◭ B-2 *x-* type
xuŋ[1]

△ B-3 *h-* type
hɔːŋ[1]

▼ B-4 *ŋ-* type
ko[1] ŋaŋ[1]

▽ B-5 *j-* type
ko[1] joŋ[1]

□ C. *vəi* type
vəi[24], vei[35]dəɯ[33], lu[31]pəɯ[13]vəɯ[31]

\+ D. *laŋ* type
laŋ[1], θiaŋ[1]

∠ E. *mak bat* type
mɐk[8]ʼbat[7]’

∿ F. *qau li* type
qaːu[322]li[33]

Figure 10.7.1: ‘*Echinochloa esculenta*’ in Kra-Dai.

‘Barnyard millet’ in Tibeto-Burman

Among the 53 instances of the word form for ‘barnyard millet’ or specifically ‘*Echinochloa* species’ collected for the TB languages and varieties, we find four main types. See STEDT for the proto forms.

Type A includes forms related to PTB *khrəy ‘MILLET / RICE’, appearing in Kman and Bai (Shiping).

Type B is derived from PLB *C-lu-k ‘millet’, found only in Burmese (Yangon).

Type C only contains *yó* (Zaiwa), of an unknown origin.

Type D includes various forms with labial sounds such as /p/ (D1) and /v/ (D2), which may be a cognate (or potentially a loan) of Sinitic *bai* ‘*Echinochloa* species’. D1 mainly appears in various dialects of Bai, and D2 appears in Loloish languages.

Our data show that the distribution of Type D in China is restricted; hence, it is plausible that a relationship exists between Type D and Sinitic *bai*. Because forms derived from Types A and B are also used for ‘foxtail millet’ in other TB languages, Types A and B for ‘*Echinochloa* species’ may be a relatively recent development for the sake of semantic precision.

(SUZUKI Hiroyuki, EBIHARA Shiho, IWASA Kazue, KURABE Keita, SHIRAI Satoko)

■ A. PTB *khrəy: *a kuɛr, koŋ*^{31}*ɹa*^{55}.
● B. PTB *C-lu-k: *lú*.
⬮ C. *yó*.

D. labial initial type
/ D1. *peɹ*^{42}, *se*^{31}*le*^{35}*pɛ*^{42}, *tshɔ̃*^{33}*tʃŋ*^{33}*pø*^{53}.
\ D2. *vi*^{33}, *vi*^{22}, *vi*^{55}, *ve*^{13}.

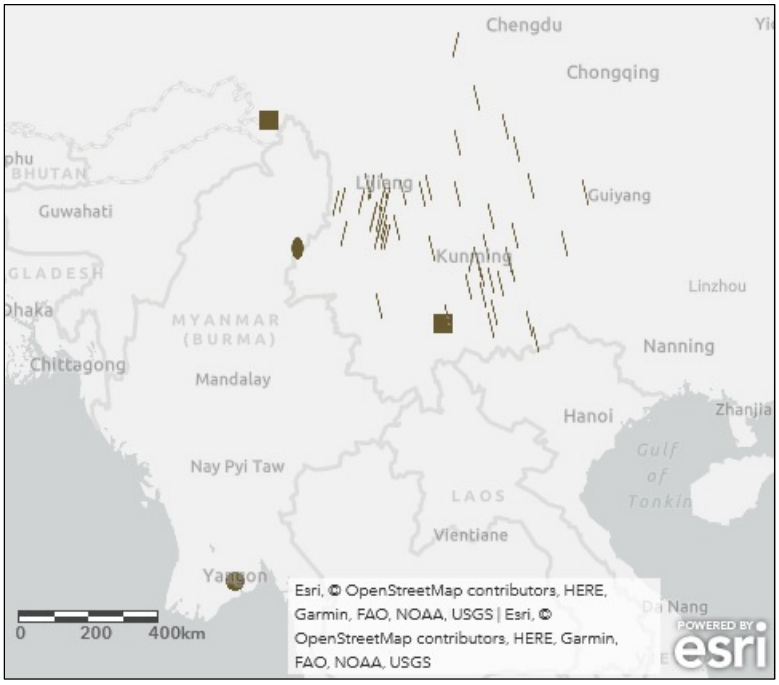


Figure 1.1.1: ‘Barnyard millet’ in Tibeto-Burman.

‘Barnyard millet (*Echinochloa frumentacea*)’ in South Asia

I describe only the languages of Indo-Aryan languages in South Asia here.

A small number of languages in South Asia have terms for ‘barnyard millet’. There are a major category of word forms – A) *śyāmāka* – and three minor categories B) *kōdrava*, C) *cīna(ka)*, and D) *varī*.

As for the distribution of ‘barnyard millet’ words, there are only the Indo-Aryan languages located south of latitude 34° N on the mainland. Of course, south of lat. 19° N they mainly speak the Dravidian languages. There is not detected any language having the name of ‘barnyard millet’ on the islands of South Asia.

The most major type is *śyāmāka*. This type is derived from Sanskrit *śyāmāka* श्यामाक ‘*Panicum frumentaceum*’ as I touched once in the part of ‘proso millet’. Forms of this type are used in the western part of the subcontinent. The Oriya form *suẽ čāuḷa* ଶୁଆ ଚାଉଳ is a compound with *čāuḷa* ଚାଉଳ ‘rice’, which is inherited of Sanskrit *taṇḍulá* तंडुल ‘grain, esp. rice, after threshing and winnowing’. Some other IA languages employ *śyāmāka* terms to refer to ‘*Panicum frumentaceum roxburghii* = *Sorghum bicolor* (great millet, Indian millet, sorghum, モロコシ)’ or ‘*Echinochloa colona* (jungle rice, コヒメビエ)’.

The *kōdrava* type appears in languages belonging to the eastern group of IA. The Sanskrit form is *kōdrava* कोद्रव and refers to ‘*Paspalum scrobiculatum* (kodo millet, スズメノコビエ)’, so some modern IA terms as its descendant refer the plant. And the English name ‘kodo millet’ also comes from Nepali inherited term *kōdō* कोदो ‘id.’.

The *cīna(ka)* type C is detected for this millet, too. They call so in and around the Bihar state of India.

The fourth type *varī* can be seen on the northern coast of the Arabian Sea in India. Its etymon is Sanskrit *varī* वरी ‘*Asparagus racemosus* (シャタバリ), a kind of grain’.

(YOSHIOKA Noboru)

A. *śyāmāka* type (6) ☆
sə̃wak, sə̃wā, sə̆vā, sə̃ə̃, šōl
[+*taṇḍula*]
suə̃ čāuḷā
B. *kōdrava* type (3) ⛉
kodua, kodo
C. *cīna(ka)* type (2) ⛳
čīnh
D. *varī* type (2) ●
varī, warāi

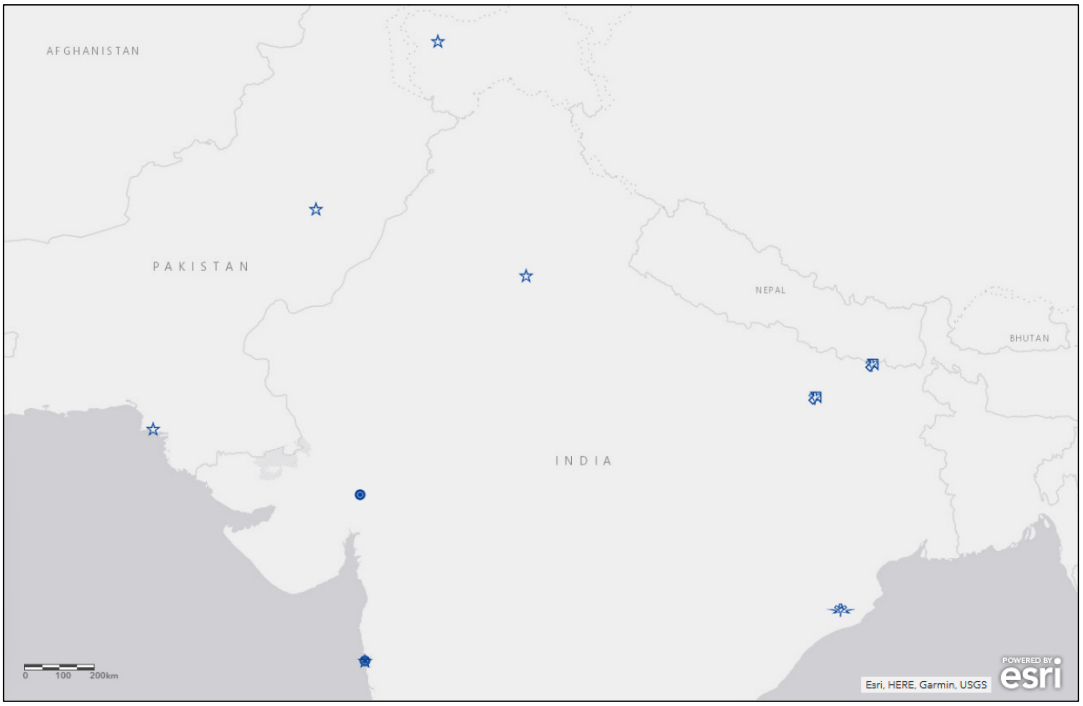


Figure 10.14.1. ‘Barnyard millet’ in SA: Indo-Aryan.

‘Barnyard millet’ in Dravidian

Information on Indian barnyard millet, *Echinochloa frumentacea*, is limited. Tamil and Malayalam forms mean ‘horse tail’. Telugu *ūdara* or *ūda* may refer to its wild ancestor, *Echinochloa colona.*

(KODAMA Nozomi)

- *kutiraivāli, kutiravāli*
- *ūdara*

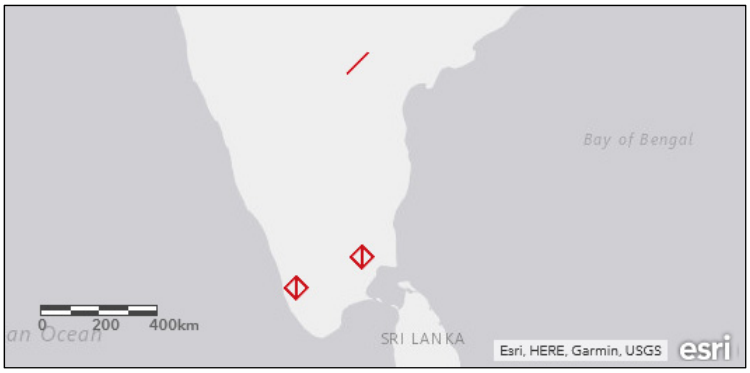


Figure 10.15.1: ‘Barnyard millet’ in Dravidian.

Chapter XI

Taro

‘Taro’ in Asian and African languages

Taro (*Colocasia esculenta*) is likely to be native to Southern India and Southeast Asia and is now cultivated in tropical areas worldwide.

Table 1 shows the principal forms in our data: Japonic, Korean, Sinitic, Hmong-Mien, Kra-Dai, Tibeto-Burman, Austroasiatic, Austronesian, Indo-Aryan, Dravidian, and Nilo-Saharan. The data are limited to languages spoken in the regions where taro is cultivated.

Table 1: Main word forms for ‘taro’.

Languages	Word forms
Japonic	*satoimo* 里芋 *imoko* *zuikiimo* *hataimo* 畑芋 *koimo* 子芋 *ieimo* 家芋
Korean	*t^{h}oran*
Sinitic	*yu* 芋 *shu* 薯 *shao* 苕
Hmong-Mien	*wo* *hou*
Kra-Dai	sa(k) type ke-/ge- type p(h)uu(k) type j- type zung type v-type r-type
Tibeto-Burman	PTB *m-n(w)ay PLB *blum^{2} PKC *ɓaa PTB *kywəy *yutou*, *yuzi* (< Chinese *yu* 芋)
Austroasiatic	traaw sᵊroʔ^{1} krao bɔːn^{1} môn ʔoːj
Austronesian	CALI KALADI GABI TALES
Indo-Aryan	*ālu(ka)* type < PIE *h_2eHlu *kacu* < Skt. *kaču* Forms < Skt. *tarūṭa* Forms < Skt. *karkaṭa*
Dravidian	*cēmpu, cēvu, kēsave*-type *kiyub*-type *hōpa*-type *kanda:*
Nilo-Saharan	*nyalamba* *mayuni* (< Ugandan Bantu)

Languages in areas where taro is considered native, such as Kra-Dai, Austroasiatic, Indo-Aryan, and Dravidian, have various roots for the word denoting ‘taro’. Contrarily, languages such as Japonic, Sinitic, and Tibeto-Burman have a single root for ‘tuber’ (*imo*, *yu*, and PTB *kywəy, respectively) from which various specific nouns are derived.

Although the forms of Hmong-Mien represent the most archaic status, they are borrowing from Sinitic *yu*.

The Indo-Aryan or Sanskrit *kaču* type and the Dravidian *cēmpu* type are mutually related, with the latter being acquired as an early borrowing.

It is observed that the word for a specific tuber can be borrowed as a word for another plant a lexical loan. For example, Indo-Aryan *ālu* ‘esculent root of *Amorphophallus campanulatus*’ is borrowed into Lhasa Tibetan as a loan denoting ‘potato’.

(SUZUKI Hiroyuki)

‘Taro’ in Japonic

The major type of taro in mainland Japan is SATOIMO (*satoimo*, *sadoemo*, *satuumu*, …) ‘village tuber’. Moreover, many other types are also used, such as IMO(NO)KO (*imonoko*, *imoɴko*, *imoko*, *imonokokko*, *emonogo*, …) ‘tuber’s child’, ZUIKIIMO (*zuikiimo*, *zujukiimo*, *zɯkiimo*, *zïïkiimo*, *zïkeimo*, *dzïgïemo*, *sugiɴ-imo*, *zuiki*, *zïgi*, …) ‘stem tuber’, HATA(KE)IMO (*hatakeimo*, *hataimo*, *hadaimo*, *ataimo*, *pateeumu*, …) ‘field tuber’, TAIMO (*taimo*, *taammu*, *taamm*, *taaɴ*, *tooɴ*, *taanimu*, *taaniumu*, …) ‘(rice) field tuber’, KOIMO (*koimo*, *kookoimo*, *kozooimo*) ‘small/child tuber’, IE(NO)IMO (*ienoimo*, *eimo*, *enaimo*, *enoimo*) ‘house tuber’, EGUIMO (*eguimo*, *egaimo*, *jogoimo*, *juguimo*, *uguimo*) ‘harsh-taste tuber’, TADAIMO (*tadaimo*) ‘common tuber’, TOO(NO)IMO (*toonoimo*, *tooimo*, *tonoimo*, *toimo*) ‘Chinese tuber,’ HAIMO (*haimo*, *faemo*, *haeemo*) ‘leaf tuber,’ MAIMO (*maimo*) ‘true tuber’, and BAKAIMO (*bakaimo*) ‘stupid tuber’. In the Ryukyus, forms such as SATOIMO and MUZI (*mudzï*, *muudzï*, *muzu*, *muɴtsu*, *muda*, …) are used; the latter type appears only in the Ryukyus.

The oldest attested forms for taro in Japanese are *ipe tu imo* and *ipe no imo*, both of which literally mean ‘tuber of house’. These forms would have been derived from the fact that taro is cultivated near houses while yam grows wild in the mountains. It seems that *imo* and *umo* originally referred to taro, and even in modern dialects, IMO (*imo*, *umu*, …) refers more to taro than to yam. In modern dialects, IMO is often used to refer to potatoes or sweet potatoes. The existence of more forms referring to taro than to yam suggest that taro was originally called IMO, and when new IMO members such as potatoes and sweet potatoes were brought to Japan, various forms for taro were created to distinguish taro from new members of the IMO class.

It is important to consider whether *imo* or *umo* is older when searching for cognates with *imo* or *umo* in other languages. Since the oldest attested form for taro in Japanese is *umo*, and the form of proto-Ryukyuan is also **umo*, the proto Japonic form is generally considered **umo*, but *imo* > *umo* seems more natural from the viewpoint of sound change. It is thus necessary to determine the proto Japonic form for taro when we compare Japonic with other languages.

(NAKAZAWA Kohei and YOKOYAMA Akiko)

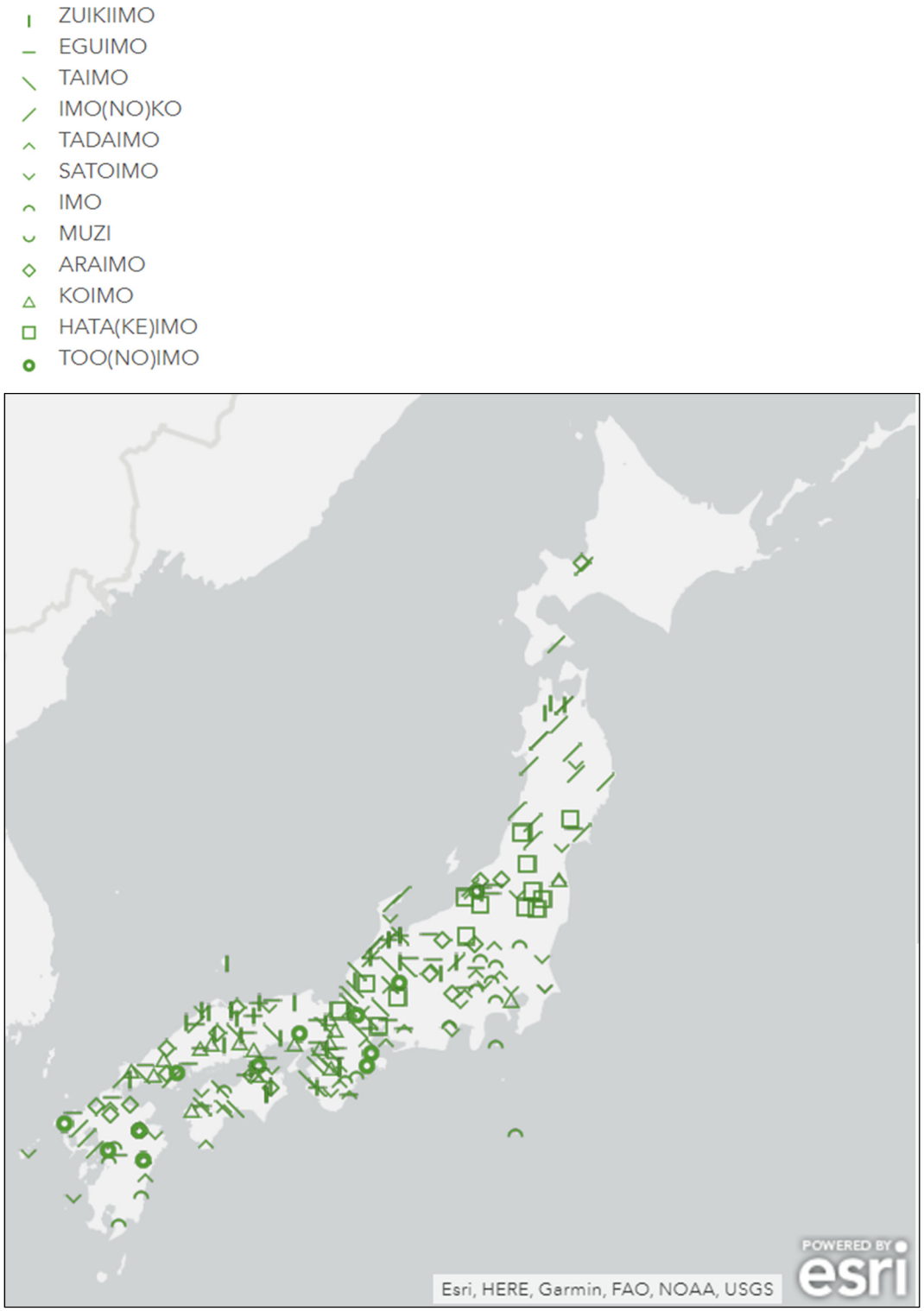


Figure 11.3.1: ‘Taro’ in mainland Japan.

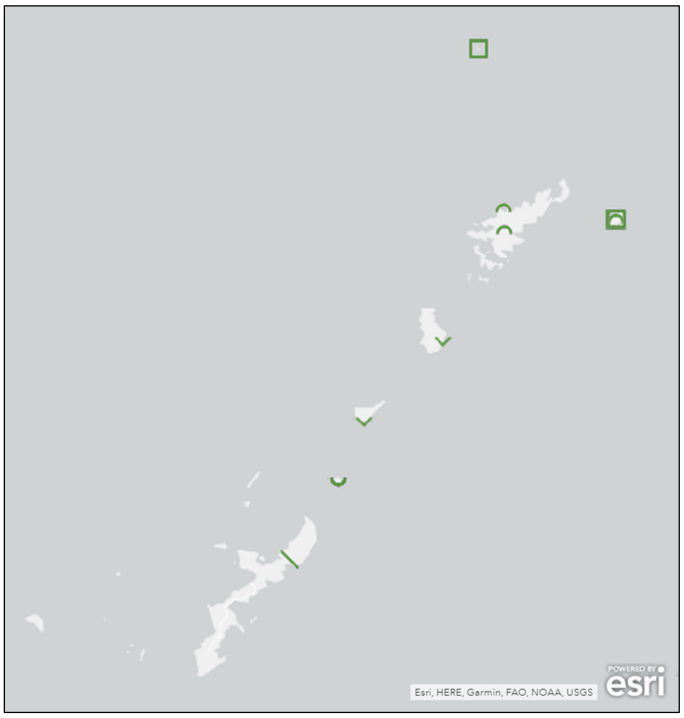


Figure 11.3.2: ‘Taro’ in Northern Ryukyu Islands.

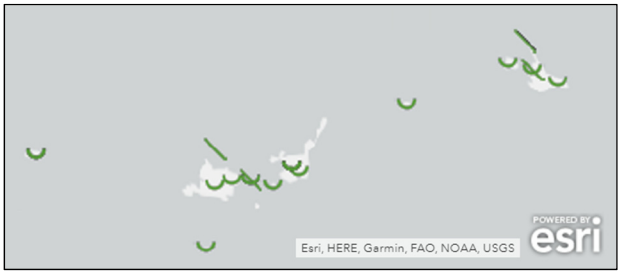


Figure 11.3.3: ‘Taro’ in Southern Ryukyu Islands.

‘Taro’ in Korean

The following table shows Modern and Middle Korean terms for ‘taro’. Middle Korean tone is shown in parentheses.

	Modern Korean	Middle Korean (15-16c.)
taro (里芋)	*t*^{h}*oran*	*t*^{h}*oran* (H-)

As can be seen in this table, Modern and Middle Korean forms are almost the same except for the fact that Middle Korean form has a distinctive pitch (high tone) in the initial syllable.

The origin of the word for ‘taro’ *t*^{h}*oran* is generally believed to be a construction based on Sino-Korean morphemes ‘土卵’ (*t*^{h}*oran*).

This word was not recorded either in *Jīlín lèishì* (雞林類事, 12th century) or in *Cháoxiǎnguǎn yìyǔ* (朝鮮館訳語, 15th century).

Dialectal differences are not so great. Generally speaking, ‘Taro’ seems less known and less consumed in Korea than in Japan.

(FUKUI Rei)

○ t^{h}oran

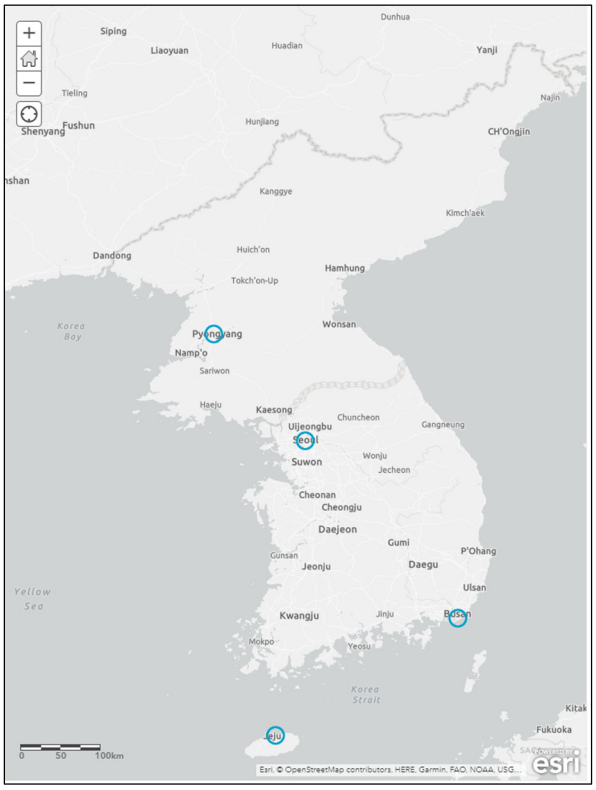


Figure 11.4.1: ‘Taro’ in Korean.

‘Taro’ in Sinitic

Taro was introduced to China early in the tenth century BC and cultivated primarily in southern China. Thus, most northern dialect reports do not include word forms for ‘taro.’

Almost all the word forms for ‘taro’ share the stem *yu* 芋, while the word forms for ‘yam’ have the stem *shu* 薯 (see ‘Yam’ in Sinitic). We can classify them into two categories by word formation type: A stem (+ suffix) type and B modifier type.

A-1 *yu* 芋 (monosyllabic) is mainly distributed in the Fujian, Guangdong, and Taiwan areas. The monosyllabic *yu* appeared in Old Chinese literature. Baxter and Sagart (2014) reconstructed it as *[G]^{w}(r)a-s (OC) > hjuH (Middle Chinese).

The most frequent form among them all is A-2 *yutou* 芋头, distributed in the whole area of Sinitic languages. A-3 *yunai* 芋艿 shows a concentrated distribution in the lower reaches of the Yangtse River. A-4 *yuzi* 芋子 is sparsely distributed in southern China. A-2 *yutou* 芋头 and A-3 *yunai* 芋艿 are derived from ‘mother tuber,’ while A-4 *yuzi* 芋子 is derived from ‘daughter tuber.’ (The southern dialects often distinguish between the ‘mother’ and ‘daughter’ tuber.)

Ningbo (宁波): 芋艿[ɦɲn^{22} na^{44}]: generic
芋艿子[ɦɲn^{22} na^{44} tsɿ^{55}]: daughter
Lichuan (黎川): 芋[y^{13}]: generic
芋头[y^{13} hɛu^{35}]: mother
芋仔[y^{13} tsɿ^{53}]: daughter
Nanchang (南昌): 芋头[i^{11} tʰɛu^{0}]: generic
芋头婆仂[i^{11} tʰɛu^{35} pʰo^{35} li^{0}]: mother
芋头崽嗨[i^{11} tʰɛu^{35} tsʰai^{213} tɛ^{0}]: daughter

In Lichuan and Nanchang, ‘mother tuber’ is differentiated by the additional suffixes *tou* 头 (頭 ‘head’) and *po* 婆 (‘mother’). ‘Daughter tuber’ is differentiated by the additional suffixes *zi* (子, 仔) and *zai* 崽.

The B modifier type is sparsely distributed in the southeast, although the B-2 *mao* 毛 type (*maoyutou* 毛芋头, *maoyunai* 毛芋艿) is concentrated in the Shandong and Jiangsu provinces. The B type forms consists of the stem *yu* 芋 and modifiers derived from various characteristics of taro. For example, B-1 *shan* 山 (‘mountain’) indicates its growing place; B-2 *mao* 毛 (‘fur’) describes its fuzzy roots; and B-3 *xiang* 香 (‘sweet-smelling’) describes its scent.

(SUZUKI Fumiki)

A. stem (+ suffix) type
- A-1 *yu* 芋 (monosyllabic)
- A-2 *yutou* 芋头
- A-3 *yunai* 芋艿, *yunaitou* 芋艿头
- A-4 *yuzi* 芋子
- A-5 *yuluan* 芋卵

B. modifier type
- B-1 *shan* 山 type: 山芋, 山芋艿
- B-2 *mao* 毛 type: 毛芋头, 毛芋艿
- B-3 *xiang* 香 type: 香芋头
- B-4 others: 白芋, 有皮芋

C. others: *shu* 薯, *shao* 苕

Figure 11.5.1: ‘Taro’ in Sinitic.

‘Taro’ in Hmong-Mien

We examine the geographical distribution of the forms of TARO. First, based on comparative evidence, we categorize the forms into several types.

There are nine types in TARO: A: *wo*; B: *hou*; C: *qau*; D: *kwai ɕwei*; E: *i deu*; F: *vu tu*; G: *y teu*; H: *zi du*; I: *ʔzu*.

Since Type A has the widest distribution in Hmongic, and Type B has the widest distribution in Mienic, they must represent the most archaic state for the entry in each branch. Both types are borrowing from Chinese 芋, likely from different sources. Type E to Type I are more recent Chinese loanwords from different sources.

(TAGUCHI Yoshihisa)

Figure 11.6.1: ‘Taro’ in Hmong-Mien.

‘Taro’ in Kra-Dai

The word ‘taro’ in Kra-Dai can be classified into A-Q categories. The most popular type is Type C. It is derived from Proto-Tai *phlïek^{D}/*phrïek^{D} reconstructed by Li (1977) or *pruək^{D} reconstructed by Pittayaporn (2009). The Type C forms are distributed widely in the Kra-Dai area except to the Hainan Island. This category can be classified into six subtypes. The two large subtypes are p-type (C-1) and ph-type (C-2). Type C-1 is distributed in the Northern Tai area, and Type C-2 is distributed in the Central and Southwestern Tai area, respectively. These two subtypes lack the consonant cluster of Proto-Tai. Types C-3 to C-5 have consonant clusters such as pj- (C-3), pl-/pr- (C-4), or phl- (C-5). While the Proto-Tai *phl-/*phr- (or *pr-) has been changed into pj- in the Northern Tai area, modern varieties in the Central and Southwestern Tai area have lacked such consonant cluster. Regarding C-6, based on comparing etyma, it is obvious that *phl-/*phr- (or *pr-) in Proto-Tai has changed into f- in Black Tai.

Type A forms are found in the northern region of the Hainan Island, that is, the Be language area. A modern variety of Gelao (belongs to the Kra branch) in Guizhou also has a Type A form.

Type B also shows the coincidence among the Hainan Island and the Kra branch languages. Both Types B-1 and B-2 are distributed in the Hainan Island and around the Kra area.

Besides Types A and B, we can find some correspondences between the Hainan Island and the Kra branch area. Concretely, Type F in the Kra area and Type G in the Hainan Island have the correspondence of voiced/voiceless initial.

Regarding word formation, ‘taro’ in Kra-Dai follows the other word in some modern varieties, such as in Figure 11.7.1. The most popular type is <LA(K) + ‘taro’> type. It is distributed in the Northern Tai area. On the other hand, <MA(K) + ‘taro’> type is distributed mainly in the northern Hainan Island. It seems that the former elements in these types are derived from ‘root’ and ‘fruit’, respectively.

Figure 11.7.1: Word formation of ‘taro’ in Kra-Dai.

(HIRANO Ayaka)

■ A sa(k) type
mak^{8} sak^{7}, sak^{7}, mak^{8} sak^{8}, ma^{2} sak^{7}, taŋ^{31} sa^{35}

B ke-/ge- type

◆ B-1 voiced type
ge:ʔ^{7}, ge:ʔ^{8}, ge:k^{7}, ʔɣa:k^{7}

◇ B-2 voiceless type
he:ʔ^{8}, xeʔ^{8}, kɯ^{5}, xɯ^{2}, qu^{31}, qa^{13}

C p(h)uu(k) type

● C-1 p- type
pɯ:k^{7}, pi:k^{9}, pɯə^{5}, pə^{3}, pɯ:k^{9}, pik^{7}, pu:k^{7}, pɯʔ^{7}, pɯ^{5}, pɯak^{9}, pə:k^{7}, pə:k^{8}, pə:k^{9}, pə^{5}, pəʔ^{7}, piak^{9}, pie^{5}, pie^{5}, pɯ^{3}, pɯak^{2}, pɯə^{1}, pɯək^{9}, pɯk^{9}, py:k^{9}, kuk^{8} pɯ:k^{7}, lak^{8} pi:k^{9}, lak^{8} pu:k^{9}, lɯk^{8} pu:k^{9}, lə^{2} pɯk^{9}, lɯk^{8} pɯə^{5}, lɯʔ^{8} pi:ʔ^{7}, lɯʔ^{8} pɯ:ʔ^{7}, man^{2} pɯk^{8}

◓ C-2 ph- type
phək^{9}, phɤ:k^{2}, phɯak^{2}, phɤk^{2}, phɯ:k^{9}, phɯək^{3}, phy:k^{9}, phyɯɯk^{2}, phaɯ^{5}, phek^{9}, phə:k^{7}, phə^{3}, phəɯ^{5}, phɤ:k^{3}, phia^{5}, phik^{2}, phik^{9}, phɯ:k^{7}, phɯk^{9}, phy:k^{7}, ma:k^{9} phə^{4}, ho^{1} phək^{9}

◉ C-3 pj- type
ləɯ^{2} pjəɯ^{7}, pjo:k^{9}, lɯk^{8} pju:k^{9}, lak^{8} pjɯ:k^{9}

◎ C-4 pl-/pr- type
plə:k^{9}, pliək^{9}, plɯak^{1}, plɯak^{2}, plɯak^{5}, plɯək^{9}, prɯ:k^{9}

⊖ C-5 phl- type
phlu:k^{9}, phlɯ:k^{9}

⊕ C-6 fɯaʔ^{2}

D j- type
ja:k^{7}, ja:k^{9}, ji:k^{31}, ʔjek^{7}

E zung type
ləʔ^{8} zəŋ^{3}, lu^{55} zuŋ^{55}, lɯk^{8} zɯm^{4}, te^{6} zɯŋ^{4}

▲ F v- type
və˞^{31}, vo^{44}

△ G fau^{4}

★ H ð- type
ma^{0} ðiak^{55}, ðuə^{24}

I r- type
la^{4} rjɛk^{9}, la^{4} rɛk^{9}, ra:ʔ^{7}

J ʐ- type
la^{13} ʐa^{13}, ʐo^{45}

K lɔ^{31} ɯɯ^{33} ʂɔ^{33} ju^{33}

+ L ɕɯam^{2}

✱ M hɔi^{4} ou^{2}

N lak^{8} ŋa:ŋ^{6}

⋈ O lək^{8} jəm^{4}

P luŋ^{31} ŋu^{53}

Q tɕi^{34} taŋ^{31} ʐyi^{22}

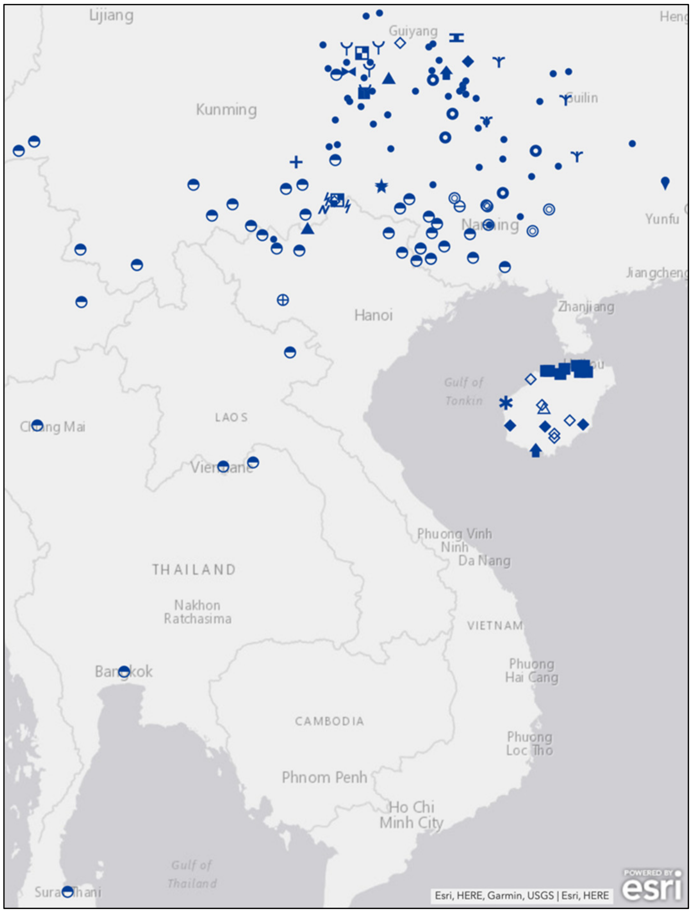


Figure 11.7.2: ‘Taro’ in Kra-Dai.

‘Taro’ in Tibeto-Burman

Among the 124 items of the word forms for ‘taro’ that have been collected in TB languages and varieties, we find eight types that show one loan type and miscellaneous. All of the proto forms are based on STEDT.

Type A includes forms derived from PTB *m-n(w)ay ‘YAM / TARO’.

Type B includes various forms related to PLB *blum^{2} ‘TARO’, provisionally classified into five subcategories. Type B1 groups forms including a /bi/-syllable, attested in Loloish, Prinmi (Taoba), Tujia and Kurtö; Type B2 includes /pi (t^{h}v)/-like forms, attested in Bai; Type B3 is forms including /p, p^{h}/+a following syllable, attested in Loloish and Burmish; Type B4 categorises forms including a /d/-initial, attested in Loloish; and B5 is forms including /bu p^{h}a/, attested in Loloish.

Types C and D include forms derived from PKC *ɓaa ‘YAM’ and PTB *kywəy ‘YAM’, respectively. These types also appear in the word forms for ‘yam’, as shown in the meaning of the proto forms.

Types E–H group sets of words of an unknown origin. Type E appears in Anong, Bai (Luobenzhuo), Pwo Karen and Dumi. Type F appears in Lololish languages. Type G appears in Galo, Atong and Hayu. Type H appears in Idu, Songlin and Geba.

Type I categorises the Sinitic loanwords such as *yutou* and *yuzi*. These appear generally in the north-eastern part of the target region, where TB languages are in contact with Sinitic. In addition, ‘taro’ is generally not cultivated in the distribution area of Type I.

Type X contains several independent stems. These are mainly found in Arunachal Pradesh and Nagaland, India.

Most types exhibit specific geographical distribution. Although Type B includes five subgroups, no clear chronological order was found among them. However, Type B1 is likely to be prior to Type B2 due to its initial form and word form, which consists of the stem only. Due to its limited distribution, it may also be that Type F is a newly acquired form in the Loloish languages.

To conclude, it is too early to assign a complete historical network among the word forms for ‘taro’ in TB languages based on the present dataset.

(SUZUKI Hiroyuki, EBIHARA Shiho, IWASA Kazue, KURABE Keita, SHIRAI Satoko)

■ A. PTB *m-n(w)ay: *muɛ*, *mɛ*, etc.

B. PLB *blum[2]

○ B1: *bi*, *ni[1] bi[1]*, *lo[35]bɛ[35]*, *bjo*.

⊕ B2: *pi[21] tʰv[55]*, *pi[42] tʰuɛɹ[55]*, etc.

● B3: *pɛ[53] si[11] phu[33]*, *bø[31] si[31]*, etc.

◑ B4: *ŋo[33]*, *dɔ[33]*, *ŋɖŋ[33]*, *dɛ[33] mo[21]*, etc.

◒ B5: *a[33] bu[33] pha[33]*, *a[33] bu[33] phɛ[33]*, etc.

▯ C. PKC *ɓaa: *baay*, *bǎal*, etc.

☆ D. PTB *kywəy: *gwe*, etc,

E. *khu dʒu*, *khoksi*, *xúthī*, etc.

F. *tʂho[33]*, *a[55] tsho[33]*, etc.

G. *rã:pi*, *ring*, etc.

H. *sūdī*.

I. Chinese loan type

▽ X. *ɔ[44] gi[44] ɔ[44]*, *dzünuo*, *i-ŋe*, etc.

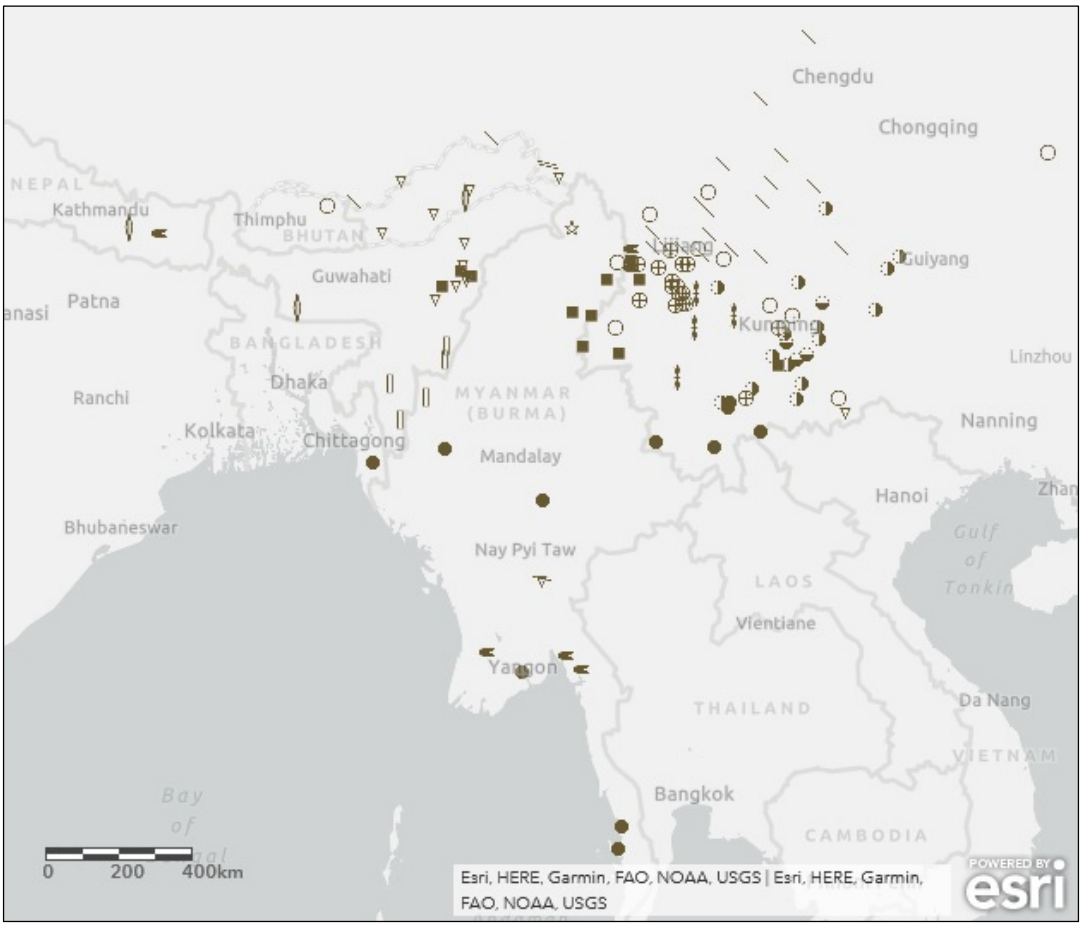


Figure 11.8.1: ‘Taro’ in Tibeto-Burman.

‘Taro’ in Austroasiatic

The data for Austroasiatic is mainly based on the online databases, *Mon-Khmer Etymological Dictionary* and *Munda Etymological Dictionary*. As the latter provides no etymon for “taro”, Map 11.9.1 shows only the data for Mon-Khmer.

Type A: (traaw, traw, (ʔa)rạ̈ːw, ʔarạ̈w, rạ̈aw~ʔarạ̈aw, raw)

Type A is in southeastern Indochina, from southern Vietnam, and Cambodia to northeastern Thailand.

Type B: (sᵊroʔ^{1}, ˉsəroʔ, srɔː, sɔː^{4}, thộ̈ː, θoː^{1} khoai sọ)

Type B is in the mountainous part of northern Indochina: from northern Vietnam and Laos to southern China. Besides, another etymon is isolated in central Myanmar.

Type C: (cʰru)

Type C is found only in the state of Meghalaya in Assam, northeast India.

Focusing on the similarities in the initial consonants and rhymes, Types A, B, and C may constitute one superior type which is the most dominant, and probably the oldest type in the area.

Type D: (krao, kᵊro^{1}, kʰreː)

Among etymons of Type D, Mon ***krao***, and Danaw ***kᵊro*** are in the Myanmar Basin, relatively close to each other. However, Chong of Kasong ***kʰreː*** is in the distant seacoast of Thailand. Thus, the similarity in the first two and the last one may be coincidental.

Type E: (bɔːn^{1}, boːn, voːn^{1}, bɔn)

Type E of the Vietic branch is in Vietnam, and that of the Aslian branch is in the distant Malay Peninsula; thus, their similarity thus looks only coincidental.

Type F: (môn, moːn^{1})

Type F is represented with Vietnamese ***môn,*** which is broadly distributed in the eastern coast of Indochina.

Type G: (buːm traw, buːm hala, buːm, buom, buom giek, buom giɛk)

Type G is concentrated in southern Laos. The common syllabic pattern, initial ***/b-/*** and final ***/-m/*** may denote “generic tuber”, followed by some modifying words. Among the modifying words, ***traw*** in ***buːm traw*** of the Alak language is the same as an etymon of Type A.

Type H: (poŋ, pố̈ng, pụ̈ŋ, póung)

Type H is in southern Laos and central Vietnam.

Type I: (ʔoːj, ʔaoj, ʔaol)

Type I is in the coastal areas of Thailand and Cambodia.

Type J: (ʔɛic, ʔəic, ʔeic, ʔaʔik, ʔek)

Type J is in the mountainous border of Kunming, Myanmar, and northern Thailand.

Type K: (taːk bum, taːk gjɔːŋ, taːk hlak)

Type K is found only in Nyaheun in southern Laos, and **taːk** may denote generic tuber followed by a modifying noun.

In addition, the other sporadic forms are found: **tadóːk, gaŋ, kladiʔ, waŋ, səlow, ɣóp, kwaaj, rbɔl,** and **pliː**

(MINEGISHI Makoto, and SHIMIZU Masaaki)

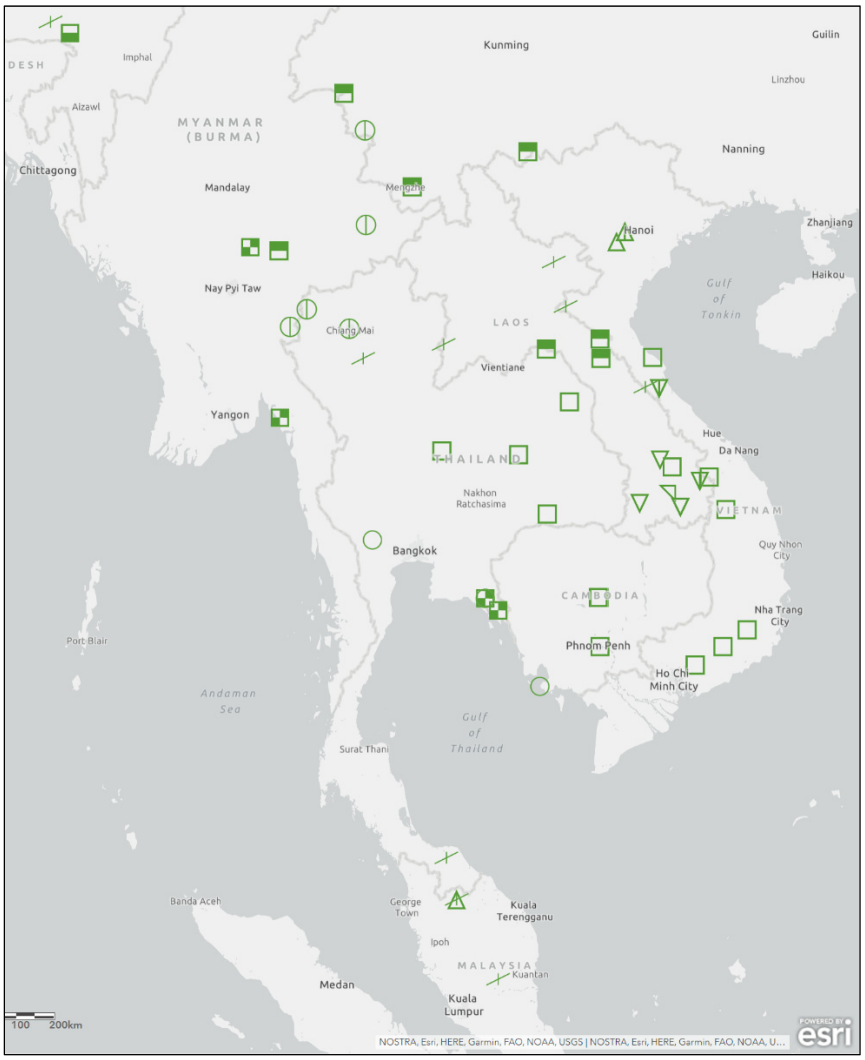


Figure 11.9.1: ‘Taro’ in Mon-Khmer.

- A: traaw, traw, (ʔa)rạ:w, ʔarạw, rạaw~ʔarạaw, raw
- B: sᵊroʔ[1], ˉsəroʔ, srɔ:, sɔ:[4], thộ:, θo:[1] khoai sọ
- C: cʰru
- D: krao, kᵊro[1], kʰre:
- E: bɔ:n[1], bo:n, vo:n[1], bɔn
- F: môn, mo:n[1]
- G: bu:m traw, bu:m hala, bu:m, buom, buọm giek, buọm giɛk
- H: poŋ, pống, pụŋ, póung
- I: ʔo:j, ʔaoj, ʔaol
- J: ʔɛic, ʔəic, ʔeic, ʔaʔik, ʔek
- K: ta:k bum, ta:k gjɔ:ŋ, ta:k hlak
- Others

‘Taro’ in Austronesian

The taro (*Colocasia esculenta*) has been cultivated since ancient times in Southeast Asia and across the Pacific, likely predating the arrival of Austronesian immigration, and it is among the most important plants for everyday consumption, across the Austronesian world, eaten in Taiwan, inland Southeast Asia, the Philippines, Indonesia, Papua New Guinea, and across vast area of Oceania. Although Taro has been replaced by rice in much of Southeast Asia, it remains a staple in the islands of Melanesia and Polynesia, where rice was not introduced until very recently. Tarot is a polymorphic crop, and it is difficult to tell domesticated tarot from wild. The data used in this paper contain a small number of forms, which suggests that confusion with closely related crops is not widespread.

The reconstructed Proto-Austronesian term for taro is **cali*, which became **tales* in Proto-Malayo–Polynesian. Another reconstructed term, from Proto Western Malayo–Polynesian, is **kaladi*.

A CALI: Forms considered to be innovations of Proto-Austronesian **cali* include /sali/ (Seediq), /soli/ (Yam, Itbayat), /tali/ (Amis), /ihari/ (Thao), and /tai/ (Bunun).
B KALADI: Innovations in Proto Western Malayo–Polynesian **kaladi* are found predominantly in Philippine languages, such as, /kaladi/ (Agutayanen, Sabah Bisaya, Iban, Wolio), /keladi/ (Malay, Balinese), /koladi/ (Mongondow), and /kolai/ (Tonsawang).
C GABI: The form /gábi/ is found in Tagalog, Aklanon, and Cebuano, among others.
D TALES: Forms that are innovated from **tales* are most frequently found in the Austronesian languages.

Forms that are very similar to the proto-form are found in /talas/ (Indonesian, Minangkabau,Tetun), /tales/ (Javanese), /täläs/ (Palawan), /tálus/ (Hanunó'o), /talis/ (Aorlan Tagbanwa), and /taleus/ (Rejang, Sundanese).

The final consonant /s/ is missing in the remaining the languages. Gorontalo, Marovo, Motu, Tongan, Samoan, Tuvaluan, and Niue exhibit the form /talo/, Nias /talõ/; Maori, Anuta, and Tahitian have /taro/, Roti has /tale/, and Paamese has /taro/. The first consonant changes into /k/ in Hawaian /kalo/, and into /d/ in Nukuoro and Eastern Fijian /dalo/. The first consonant is missing, giving /alo/ in Lau and Kwaio and /aro/ in 'Āre'āre, Arosi, and Bauro. The second vowel is missing in Erromanga /tal/. In Rennellese, the second consonant changes into /g/, as /tago/.

Other forms include /tar-kura/ in Rotuman and /na-tale/ in Nakanamanga.
E: Other forms include /wōt/ (Marchallese) and /ohd/ (Ponapean).

Type A forms are almost exclusively found in Taiwan, with the exception of Itbayat, found in the Philippines. Types B and C are predominantly found in the Philippines, but type C is also found Malay-speaking areas (Malay Peninsula and in Bali and Sulawesi). Type D forms are

distributed across the vastest areas, including Sumatra, Java, the Nusa Tenggara islands, Papua New Guinea, and Pacific islands. Other forms are also found in Pacific islands.

(UTSUMI Atsuko)

- A: *sali, soli, tali, ihari, tai*
- B: *kaladi*, *keladi*, *koladi*, *kolai*
- C: *gábi*
- D: *talas, tales, täläs, tálus, talis taleus*, *talo, talõ, taro, tale, taro, kalo, dalo, alo, aro, tal, tago, tar-kura, na-tale*
- E: *wōt, ohd*

Figure 11.10.1: ‘Taro’ in Austronesian languages in Taiwan and Philippines.

Figure 11.10.2: ‘Taro’ in Austronesian languages in Indonesia.

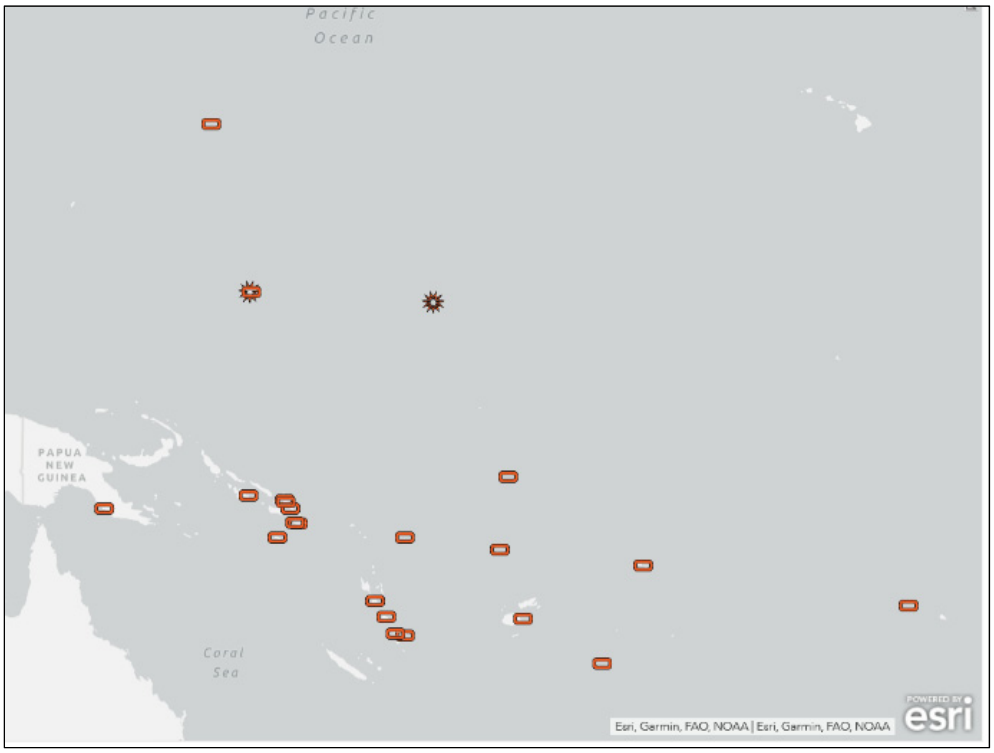


Figure 11.10.3: ‘Taro’ in Austronesian languages in Papua and Pacific Islands.

‘Taro’ in South Asia

I describe only the languages of Indo-Aryan (IA) for ‘taro’, unlike to other terms. Because there are no terms in the other languages there.

There are two categories of word forms – A) *ālu(ka)* and B) *kacu*.

The distribution of ‘taro’ words is not so wide. On the one hand, the type A is widely employed except in the east of the Indian subcontinent, on the other hand, Type B is detected there.

The most major type is of *ālu(ka)*. This type is derived from Sanskrit *ālu* आलु ‘esculent root of *Amorphophallus campanulatus* (ゾウコンニャク)’, and then ‘potato’, or *āluka* आलुक (*-ka* is used for diminutive or adjectiviser), which are inherited from Proto-Indo-Iranian **HaHlu*, ultimately Proto-Indo-European **h₂eHlu* ‘edible root’. So this type is cognate with Latin *alium* ‘onion, garlic’. Some languages use forms derived from the compound of *ālu(ka)* with Sanskrit words *piṇḍa* पिंड ‘lump, clod’ (see ‘Yam’ in South Asia) or *kaču* कचु (see Type B below). For the Dhivehi term *oḷu ala* ދ ދ is a composition with the adjective *oḷu* ‘tamed’.

The *kacu* type appears in the languages of the eastern part of India and Bangladesh as well as in Hindustani (i.e., Hindi-Urdu). The original Sanskrit form is *kaču* कचु ‘the esculent root of *Colocasia esculenta*’, which is recorded with a synonym ‘*Arum colocasia*’ in Turner (1966: 129). The eastern IA languages employ the *kacu* bare forms for ‘taro’, whereas Hindustani says *kačālū* कचालू / کچالو as a compound with *ālū* आलू / آلو of the *ālu* type, which primarily means ‘potato’ in the language. The region of bare *kacu* forms seems one of the birthplaces of cultivated taro from where it has spread abroad.

Besides the types, there remains some other words for ‘taro’. *tārō* تارو in Urdu is derived from Sanskrit *tarūṭa* तरूट ‘lotus root’, and some modern words of the etymon can be seen in ‘Yam in South Asia’. Of course the similarity between the Urdu term and the English name taro is a casual coincidence. Nepali *karkalī* कर्कली may be related to either Sanskrit *karkaṭa* कर्कट ‘name of various plants, curved root of a plant’ or *karkāru* कर्कारु ‘the gourd *Beninkasa cerifera*’ (? < √*kar* ‘dig’). And Oriya *saru* ସାରୁ seems derived from Mundari *saru* ‘*Colocasia antiquorum*, *Alocasia indica*’ (Osada 1995: 36) or Sanskrit *saru* सरु ‘fine’.

(YOSHIOKA Noboru)

A. *ālu(ka)* type (12) □
aḷū, aḷvī, ala, aruī, aravī, arubī, aḍuā
[+x]
oḷu ala
[*kacu*+ (2)] (see type B below)
[*piṇḍa*+]
piṇḍālū

B. *kacu* (6) ／
kɔču, kɔsu, xosu
[+*ālu* (2)]
kačālū

C. other
[*tarūṭa*] tārō; [*karkar*] karkalī; saru

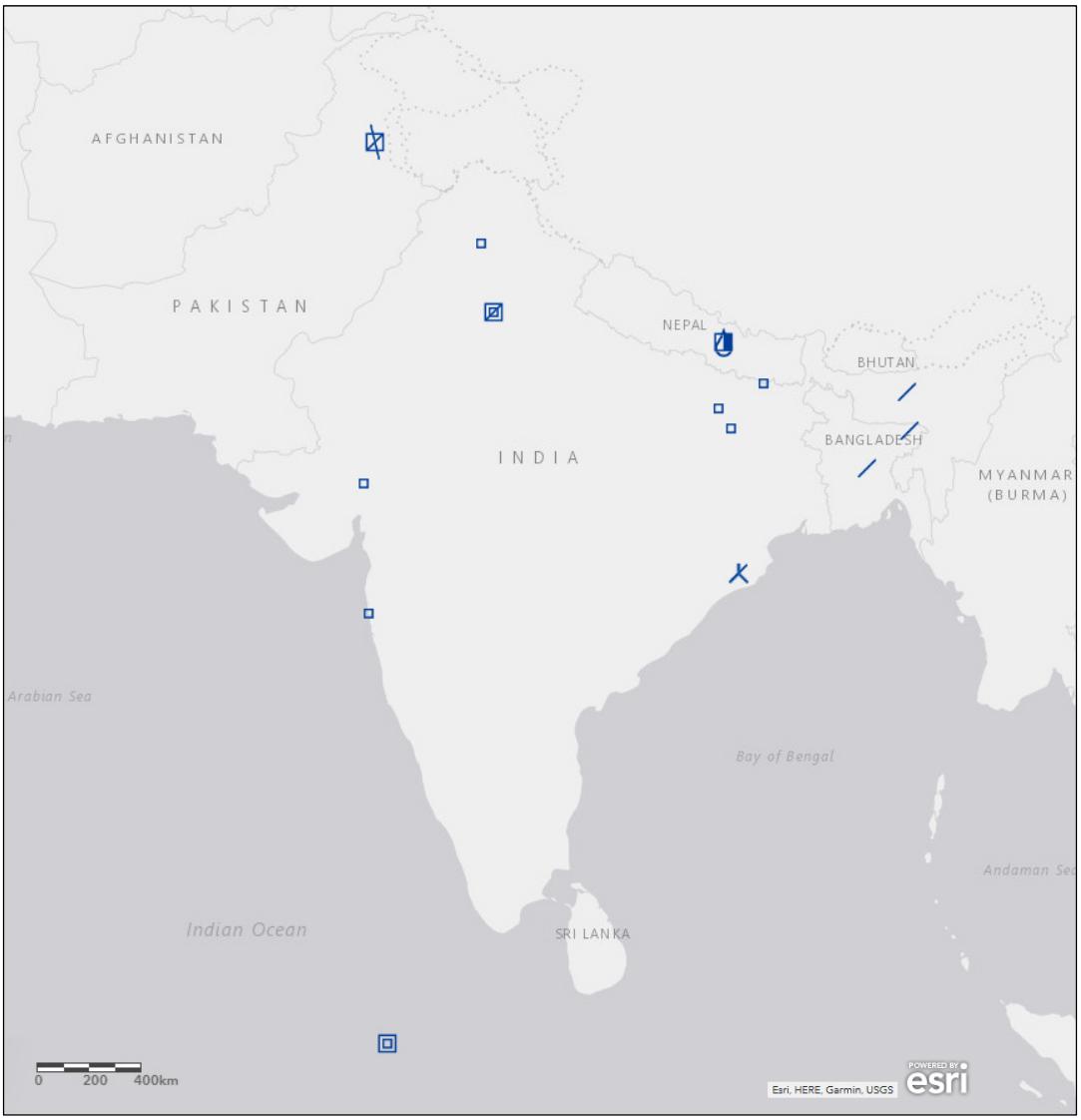


Figure 11.14.1: ‘Taro’ in Indo-Aryan.

‘Taro’ in Dravidian

Burrow and Emeneau’s (1984) DEDR classifies Dravidian words meaning *Colocasia antiquorum,* a kind of Taro native to India, into two entries for the etyma #2004 and #2881.

The former is related to Sanskrit forms for the crop such as *kemuka-, kecuka-, kevūka, kacu-, kacvī-,* probably early borrowings from Dravidian. The latter includes the words exclusively found in a closely related clade of the South Central Dravidian, with the combination of consonants relatable to the former except that the back vowel does not justify the preceding palatalization of the initial consonant. If these two etyma can be ultimately ascribed to a single etymon, the greater variation may reflect a long history of this etymon for the crop.

DEDR #2004 has a Kurukh word *kisgō* glossed as ‘yam’, which is not shown on the map for ‘Taro’. Instead another word *kanda:* glossed as ‘edible root, kind of Taro, potato’ in Kobayashi and Tirkey (2017) is shown on the map.

(KODAMA Nozomi)

cēmpu, cēmpai
cēvu, tēvu
kēsave, kĕ̄su, kesa, kesavu
cēma
sōmbu
kiyub, kībi
hōpa, hom, hūpu
kanda:

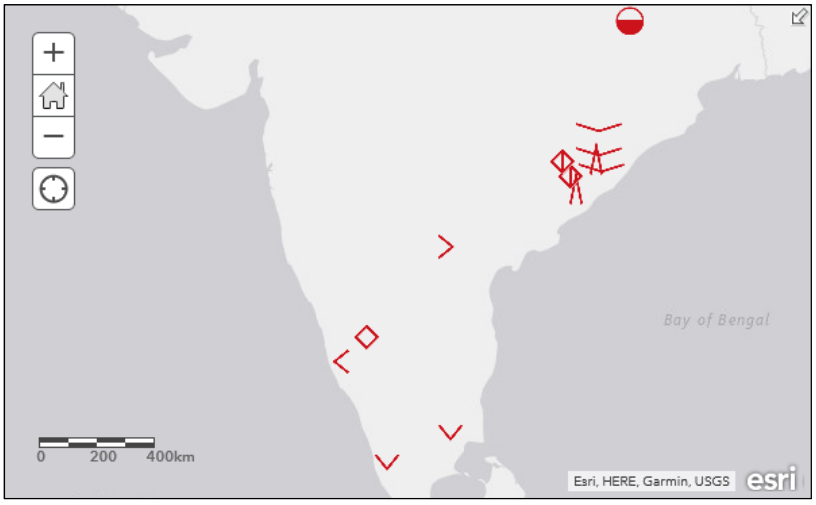


Figure 11.15.1: ‘Taro’ in Dravidian.

‘Taro’ in Nilo-Saharan

Among about 100 Nilo-Saharan languages surveyed, only two languages attest a term for ‘taro’: Kumam, a Southern Lwo (Western Nilotic) language in Uganda and Furu/Bagiro, a Central Sudanic (Sara-Bongo-Bagirmi branch) language in the Democratic Republic of the Congo.

The Kumam term *mayuni* is a borrowing from Ugandan Bantu (e.g., Luganda, plural, *(a)ma-yuni*). The etymology of the Furu term *nyalamba* is unknown.

(NAKAO Shuichiro)

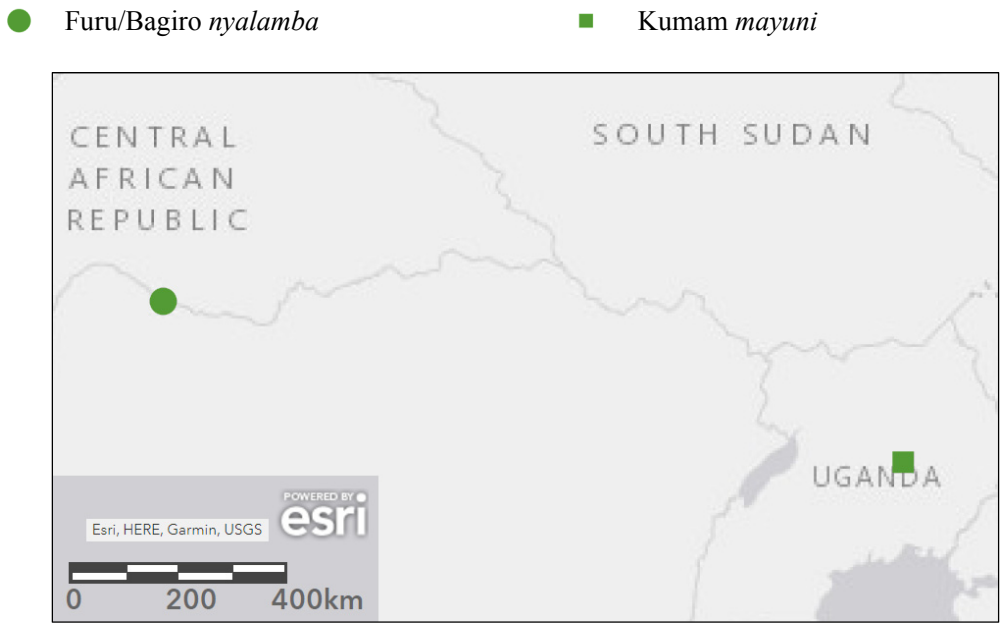


Figure 11.19.1: ‘Taro’ in Nilo-Saharan.

Chapter XII

Yam

‘Yam’ in Asian and African languages

Yam is the common name for some *Dioscorea* plant species and is widely cultivated and consumed around the globe.

Table 1 shows the principal forms in our data: Japonic, Korean, Sinitic, Tibeto-Burman, Austroasiatic, Austronesian, Indo-Aryan, Andamanese, Dravidian, and Nilo-Saharan. Although yam is a cover term for various species, our sources generally do not distinguish one from another. Therefore, our dataset deals with ‘yam’ regardless of genetic variation. The data available in our analysis is limited.

Table 1: Main word forms for ‘yam’.

Languages	Word forms
Japonic	*yamaimo* 山芋 *zinenzyo* 自然薯 *tororoimo* *nagaimo* 長芋
Korean	*ma*
Sinitic	*shu* 薯 *shanyao* 山药
Tibeto-Burman	PTB *m-n(w)ay PTB *kywəy PTB *g/s-rwa
Austroasiatic	*kwai* *kʲjoŋ* *traaw* *se*/*səɹ* *pɔŋ* *waŋ*
Austronesian	UBI SAKUT TALUKO KAI/KOKO DAM
Indo-Aryan	*ālu(ka)*-type *mīnō*-type *tarūṭa*-type *čāda*-type
Andamanese	*konmo*-type
Dravidian	*kiẓaṅku*-type *geṇasu*-type *kisgō*-type
Nilo-Saharan	*#badho* *#manga* *doya* (< Hausa)

In several languages, ‘yam’ and ‘taro’ are mutually related in the lexical root, as in Japonic (*imo*), Sinitic (*shu*), and Tibeto-Burman (PTB *m-n(w)ay, *kywəy).

(SUZUKI Hiroyuki)

‘Yam’ in Japonic

As for the word forms of yam, in addition to YAMAIMO (*jamaimo*, *jamaemo*, *jamɛɛmo*, *jamauɴ*, *jamamm*,…) and YAMANOIMO (*jamanoimo*, *jamaneimo*) ‘mountain tuber’, several terms are used in Japanese, including ZINENZYO (*dʒineɴdʒo*, *dʒineɴdʒoo*, …) ‘naturally growing tuber’, TORORIMO (*tororoimo*, *tororoemo*, *toroimo*, *tororo*, …) ‘tuber used for *tororoziru* soup’, derived from the way it is eaten, NAGAIMO (*nagaimo*, *naŋaemo*, …) ‘long tuber’ from China, and CUKUNEIMO, which refers to a specific breed. In some Southern Ryukyuan languages, it is simply called IMO (*ʔun*, *ʔoon*) ‘tuber’.

Other forms of yam in Japonic are BOOIMO (*booimo*, *booʔuɴ*) ‘stick tuber’, NOIMO ‘wild tuber’, and KAYAIMO (*kajauɴ*, *kajooɴ*, *kajaɴ*). Similar to YAMA(NO)IMO, NOIMO was named for its natural growth in the wild.

The oldest attested forms for yam in Japanese are *yama tu imo* and *yama no imo*, which literally mean ‘mountain tuber’, because it grows naturally in the mountains. ZINENZYO is a Sino-Japanese word and a newer form than the YAMA(NO)IMO. Since *imo* refers to both yam and taro, *yama tu imo* or *yama no imo* came to be used to distinguish between yam and taro.

(NAKAZAWA Kohei and
YOKOYAMA Akiko)

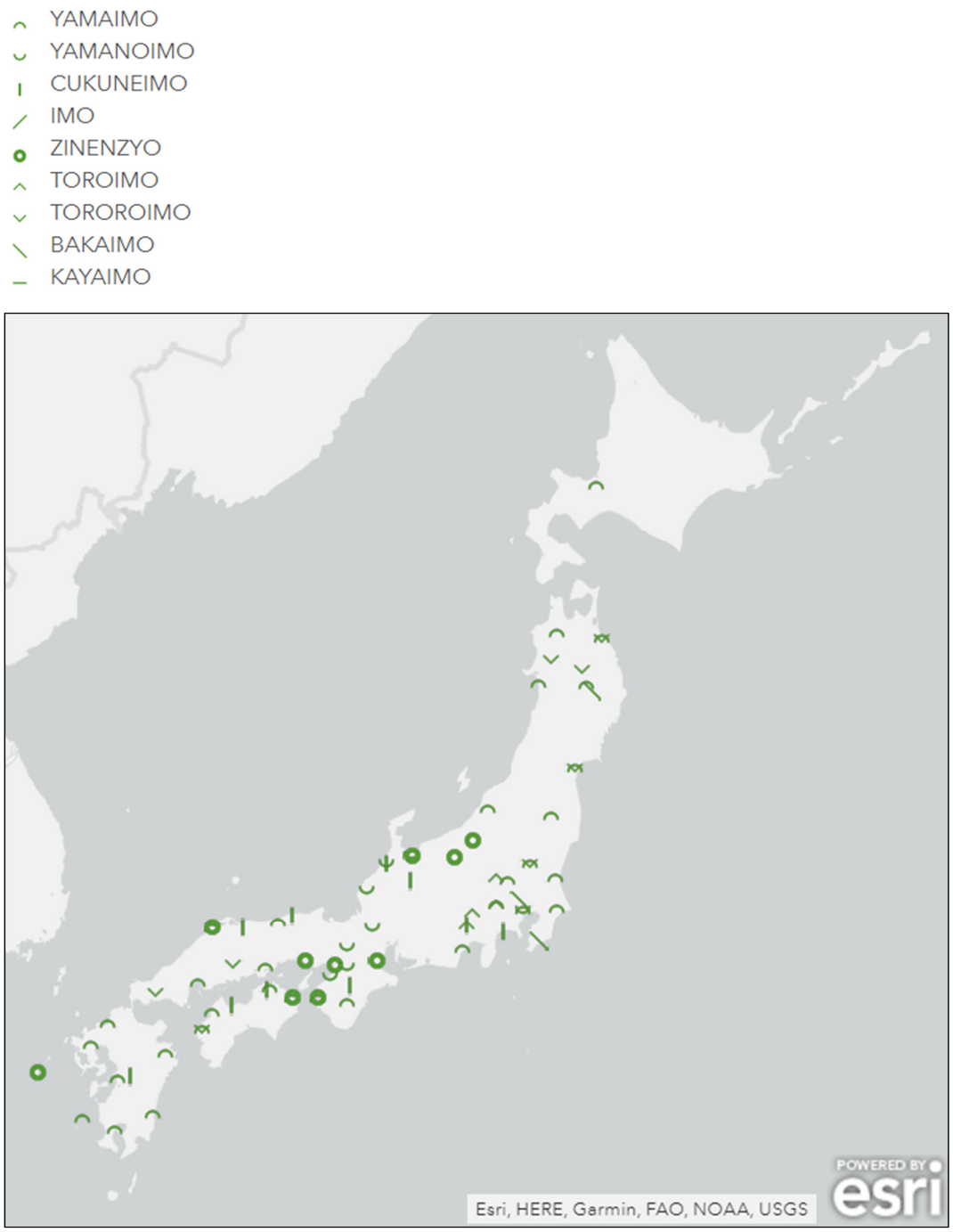


Figure 12.3.1: ‘Yam’ in mainland Japan.

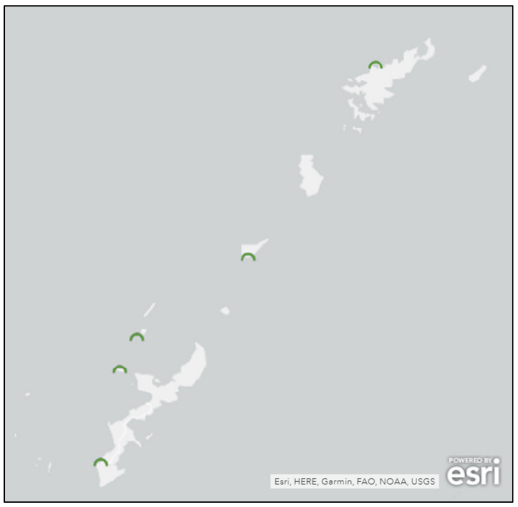


Figure 12.3.2: ‘Yam’ in Northern Ryukyu Islands.

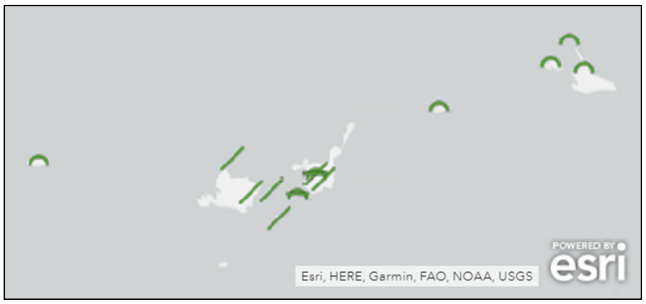


Figure 12.3.3: ‘Yam’ in Southern Ryukyu Islands.

‘Yam’ in Korean

The following table shows Modern and Middle Korean terms for ‘yam’. Middle Korean tone is shown in parentheses.

	Modern Korean	Middle Korean (15-16c.)
yam (山芋)	*ma*	*mah* (H)

As can be seen in this table, Modern and Middle Korean forms are almost the same except for the fact that Middle Korean form has a final *-h* and distinctive pitch (high tone).

The origin of the word for ‘yam’ *ma* is unknown. Also, this word was not recorded either in *Jīlín lèishì* (鷄林類事, 12th century) or in *Cháoxiǎnguǎn yìyǔ* (朝鮮館訳語, 15th century).

Dialectal differences are not so great. Generally speaking, ‘Yam’ seems less consumed in Korea than in Japan.

(FUKUI Rei)

ma

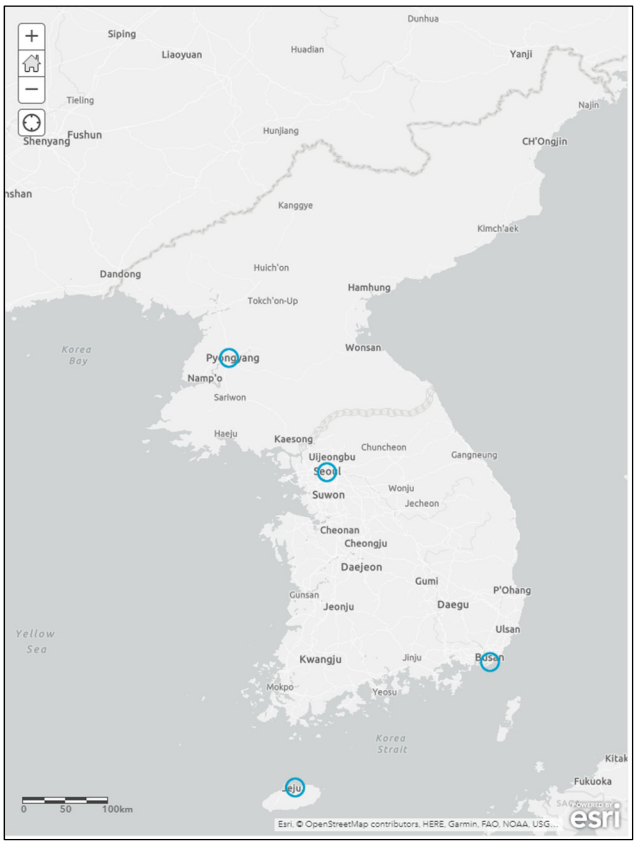


Figure 12.4.1: ‘Yam’ in Korean.

‘Yam’ in Sinitic

Word forms for ‘yam’ have two main stem types: A *shu* 薯 and B *shanyao* 山药. *Yao* 药 always combines with *shan* 山; therefore, we regard the whole *shanyao* 山药 as one stem.

Type A *shu* 薯 is distributed in southern China, while Type B *shanyao* 山药 is distributed in the north, showing a south–north contraposition. A-6 *shanyaoshu* 山药薯 is the contamination form of Type A *shu* 薯 and Type B *shanyao* 山药 types.

Considering word formation, we can classify Type A *shu* 薯 into the stem types A-1 and 2 and the modifier types A-3, 4, and 5. The modifiers are derived from characteristics such as *shan* 山 (‘mountain’), *bai* 白 (‘white’), and *jiaoban* 脚板 (‘sole of the foot’).

The Old Chinese form for ‘yam’ is A-1 *shuyu* 薯蓣. “*Shuyu* 薯蓣” changed into B-1 *shanyao* 山药 by two naming taboos:

1) “*shuyu* 薯蓣” was replaced by “*shuyao* 薯药” because of the Tang dynasty Daizong (代宗)’s naming taboo against the given name Yu 豫.
   shuyu 薯蓣 > *shuyao* 薯药
2) “*shuyao* 薯药” changed into “*shanyao* 山药” because of the Song dynasty Yingzong (英宗)’s naming taboo against the given name Shu 曙.
   shuyao 薯药 > *shanyao* 山药

Hence, Type B *shanyao* 山药 type can also be regarded as a variation of Type A *shu* 薯 type. The word forms for ‘yam’ share the stem *shu* 薯, whereas almost all the word forms for ‘taro’ share the stem *yu* 芋 (see ‘Taro’ in Sinitic).

The referents of “*shanyao* 山药” vary depending on the dialect. When it lacks an additional component, it means ‘potato’ in north Shanxi, ‘sweet potato’ in south Hebei, and ‘yam’ in south Shanxi, showing a geographic complementary distribution. In north Shanxi, the word forms for ‘yam’ add the modifier “*chang* 长” (‘long’). However, the word forms for ‘sweet potato’ adopt another form “*Hongshu* 红薯” (red potato). In south Hebei, the word forms for ‘yam’ add the modifier “*bai* 白” (‘white’), and the word forms for ‘potato’ add the suffix “*dan* 蛋” (‘egg’) to avoid the homonymic collision. Table 1 shows the conceptual model for “*shanyao* 山药.”

Table 1

	Shanxi		south
	north	south	Hebei
potato	***shanyao*** **山药**	*shanyaodan* 山药蛋	
sweet potato	*Hongshu* 红薯		***shanyao*** **山药**
yam	*changshanyao* 长山药	***shanyao*** **山药**	*baishanyao* 白山药

(SUZUKI Fumiki)

A. *shu* 薯 type
- A-1 *shuyu* 薯蓣
- A-2 *shu* 薯, *shuzi* 薯仔
- A-3 *shanshu* 山薯
- A-4 *baishu* 白薯
- A-5 *jiaobanshu* 脚板薯
- A-6 *shanyaoshu* 山药薯
- A-7 others: 大薯, 毛薯, 铁薯

B. *shanyao* 山药 type
- B-1 *shanyao* 山药
- B-2 *shanyaodan* 山药蛋
- B-3 *changshanyao* 长山药
- B-4 *caishanyao* 菜山药
- B-5 *maoshanyao* 毛山药
- B-6 others: 白山药, 麻山药

C. others
- C-1 *shao* 苕 type: 苕, 脚板苕
- C-2 *huaishan* 淮山
- C-3 others: 麻芋子, 山芳, 洋山蔓

Figure 12.5.1: 'Yam' in Sinitic.

‘Yam’ in Tibeto-Burman

Among the 63 word forms for ‘yam’ collected in the TB languages and varieties, we find six types. Types A–D are derived from different proto-level forms (see STEDT).

Type A includes forms derived from PTB *m-n(w)ay ‘YAM / TARO’. Type B includes forms derived from PTB *kywəy ‘YAM’. Type C includes forms derived from PTB *g/s-rwa ‘TARO / YAM / TUBER’. Type D includes forms derived from PKC *ɓaa ‘YAM’.

Type E includes *ɬa^{31}ba^{35}* and *dzuꞁʁoꞁ-ɬiꞁbiꟻ*, found in Naic (Shuhing and Na).

Type F includes the Burmese term *myauʔʔû*, literally ‘monkey’s egg; bulb’, and its loans. It is found in Myanmar.

Additionally, Type X groups together various independent stems.

The distributions of Types A and B are divided around the Myanmar-India border, where also Types C and D appear. It is challenging to hypothesise a history of this lexical distribution based on the current data.

(SUZUKI Hiroyuki, EBIHARA Shiho, IWASA Kazue, KURABE Keita, SHIRAI Satoko)

- □ A. PTB *m-n(w)ay: *mui^{31}*, *nwɛ^{33}*, etc.
- ○ B. PTB *kywəy: *'ron ki*, *xi*, *ki*, etc.
- ⬮ C. PTB *g/s-rwa: *ta*, *tharem*.
- ▯ D. PKC *ɓaa: *báa*, *bâa*.
- ◆ E. *ɬa^{31}ba^{35}*, *dzuꞁʁoꞁ-ɬiꞁbiꟻ*
- ♈ F. *myauʔʔû*, *myúʔ*
- ▽ X. *ŋem ʈhaŋ*, *sintaq ya*, *ɛ̃^{33}da^{55}da^{55}*, *puɯ^{33} ɛ^{33} ŋɛ^{33} mɐ^{33}*, etc.

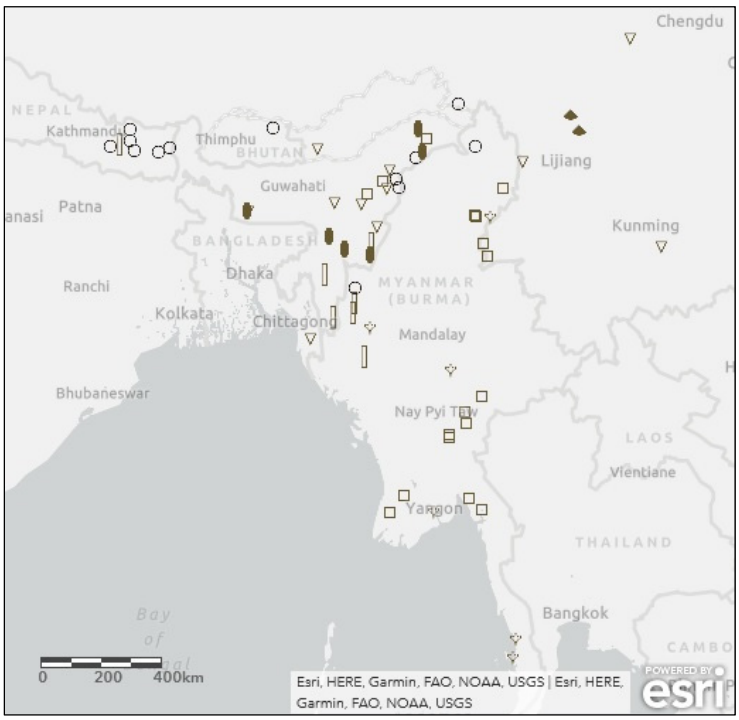


Figure 12.8.1: ‘Yam’ in Tibeto-Burman.

‘Yam’ in Austroasiatic

Among the etymons for yam in Austroasiatic languages, only those of Mon-Khmer languages are available. Most of them lack scientific names; thus, it is not clear whether they refer to the same species. Some of them may be a generic term denoting ‘tuber,’ such as in the case of Vietnamese. The following classification for ‘yam’ is therefore a tentative one.

Etymons for ‘yam’ are classified into the following eight types, A through H, based on the initial consonants.

Few similarities are found across the language branches. What follows is the types, their distinctive phonemic features, the etymons, and the geographic distributions.

Type A, with an initial /k(h)-w/ *(kwai, kwaːj, khwáaj, khoai mõn, kuói)*

kwai ‘yam’ is found in Mon in Southern Myanmar; *kwaːj* ‘yam tuber,’ in Northern Vietnam; *khwáaj* ‘potato, yam, cassava,’ in Central Thailand; *khoai mỡ* ‘yam,’ in Vietnam, and *kuói* ‘yam,’ in the Malay Peninsula.

Considering their distribution in the central area of mainland Southeast Asia, the similarity in their forms may be due to their genetic relation or to borrowing between languages.

Type B, with an initial /k/ (*kʲjoŋ, kloa, kətàt, ko-peŋ*)

kʲjoŋ is found in Central Myanmar, and *kloa* ‘wild yam’ and *kətàt* ‘a kind of yam’ are found in Southern Myanmar. As the change from /l/ to /j/ is a plausible one in the area, they are considered to be cognates. On the other hand, their resemblance to *ko-peŋ* in the distant Nicobar Islands seems only coincidental.

Type C, with an initial /t/ *(traaw, takøb, tə-ki-niːˀ-ə̃)*

traaw “*Colocasia antiquorum*” is found in Khmer in Cambodia, and *takøb* “wild yam” and *tə-ki-niːˀ-ə̃* “yam (the best variety)” are both found in the Nicobar Islands.

Type D, with an initial /s/ *(se, səɹ, si-nɛːˀ-mɛh)*

se/*səɹ* “sweet potato, yam, a sweet starchy tuber” is found in the Malay Peninsula, and *si-nɛːˀ-mɛh* “a variety of yam” is found in the Nicobar Islands.

Type E, with an initial /p/ (*pɔŋ, (ʔa)pɔŋ, pʰan ʃnreʔ*)

pɔŋ “yam (Dioscorea esculeuta)” is found in Southern Laos; *(ʔa)pɔŋ* “potato, yam, and some other edible roots,” in Northeast Thailand; and *pʰan ʃnreʔ*, in Northeast India.

Type F, with an initial /w/ *(waŋ, wean)*

waŋ “yam” is found in Northeast India, and *wean* “jungle yam” is found in the Nicobar Islands.

Type G, with an initial /m/ (*man ʔatûŋ*)

man ʔatûŋ “a kind of yam,” found in Northeast Thailand, is a loan from *man* in Thai.

Type H: The residual etymons, such as *bʌkɔɪ* and *gaʔ* in the Malay Peninsula.

(MINEGISHI Makoto, and
SHIMIZU Masaaki)

A: kwai, kwaːj, khwáaj, khoai mɔ̃n, kuói
B: kʰjoŋ, kətàt, kloa, ko-peŋ
C: traaw, takøb, tə-ki-niː-ə̃
D: se, səɪ, si-nɛː-mɛh
E: pɔŋ, (ʔa)pɔŋ, pʰan ʃnreʔ
F: waŋ, wean
G: man ʔatûŋ
H: Other miscellaneous forms

Figure 12.9.1: ‘Yam’ in Mon-Khmer.

‘Yam’ in Austronesian

The yam is a variety of plants that are native to a range of tropical and subtropical areas. It is an important staple food in Austronesian-speaking areas. Various species of yams are cultivated and eaten, most frequently the ube (*Dioscorea alata*) and the lesser yam (*Dioscorea esculenat*). The other varieties are not usually considered to be food for everyday consumption, as they contain toxins that require adequate processing to remove, which may take too much time.

When collecting vocabulary for fauna and flora, difficulty often arises in identifying the item that a lexical item refers to. A list of words that seem to refer to ube (*Dioscorea alata*), which is the principal variety of yam in the Austronesian world, always retain the possibility of confusion with similar plants, which cannot be eliminated from the data of this paper. It should be noted that the analysis given below were conducted in light of the probable deficiency of the data. Many different word forms are found for yam in Austronesian languages, which indicates its long history of cultivation and its importance as a staple food in the area, as well as the possible confusion with similar plants in the larger group of yams.

A UBI form: Word forms that are akin to /ube/ or /ubi/ are most frequently used forms within the Austronesian languages. The middle consonant alternates among /β/, /v/, /w/, and /f/, and the long consonant /ū/ may substitute for the vowel /u/ ; in some languages the semiopen vowel /o/ or /e/ may perform this role. The most typical form is /ubi/, as found in Indonesian, Balinese, Bantik, and Bangini sama. In Javanese, the word is pronounced as /uwi/, in Sundanese /hui/, and in Gorontalo /wiwiʔ /. In the Oceanic languages, the forms /uvi/ (Da’a, Manggarai and Ngada), /ufi/ (Roti, Samoan, Mele-Fila, Tahitian), and /ʔuhi, ʔufi/ (Rotuman, Tongan, and Rapanui) are found in the Pacific. Other forms include /up/ (Tolai), /owi/ (Wolio), /nufi/ (Maringe), and /yui/ (Lewo).

B SAKUT form: Word forms that have word-initial /s/ and word-final /t/ sound belong to type B. The form /sakut/ is found in Murut, /suhat/ in Batak Toba, and /šakuɣ / in Atayal.

C TALUKO form: Word forms that have three or more syllables and an alveolar or uvular consonant in the last syllable. In Tsou the word form is /taʔrućo/, in Rukai /baḷiloko/, and in Bugis /lame kalokoʔ/.

D KAI/KOKO: Word forms that have the word-initial /k/ are grouped together as KAI/KOKO forms. Lau and Kwaio us the form /kai/ to denote yam. The similar forms /ku/ and / kūk/ are found in Xârâcùù and Nemi, respectively. Nengone has the form /koko/ for the yam.

E DAM: The two Dam languages have similar forms, both containing the consonants /d/ and /m/. Raga has the form /damu/ and Port Sandwich /na-^{n}d^{r}am/.

F Other forms: Other attested forms do not have any similarity with one another, including the following range: /pulungan/ (Palawan), /tila/ (Madurese), /ambon/ (Sasak), /tuka/ (Sika), /maŋa-t/ (Buru), /k^{w}al tulir/ (Dobel), /sɛsul/ Sawai), /kana/ (Nyindrou), /uaŋai/ (Manam), /dabel/

(Takia), /gamu/ (Dami), /bīdi/ (Mbula), /ame/ (Yabem), /eni/ (Kaulong), /go/ (Buang), /ganaŋ/ (Adzera), /taitu/ (Kilivila), /woida/ (Tawala), /maho/ (Motu), /lama/ (Mekeo), /marihi/ (Roviana), /auh/ (Paamese), /nuk/ (Kwamera), /nup/ (North Tanna), /û/ (Cémuhî), /mʌu/ (A'jië), /kēp/ (Ponapean), /sepa/ (Woleaian), and /βičua/ (Western Fijian).

A loan form from English, /iām^{w}/ is found in Marshallese.

Type A forms are spread across the widest area, spreading over vast ranges of Southeast Asia, including Taiwan, the Philippines, and Indonesia. Languages in the Solomon islands and the Oceanic area also demonstrate one of Type A forms.

Type B forms are found in Taiwan, Northern Borneo, and Northern Sumatra. Type C forms are found in Taiwan and Southern Sulawesi. Type D and type E forms are only found in Oceanic languages. There is no specific distribution area for other forms.

(UTSUMI Atsuko)

A: *ubi, uwi, hui, wiwiʔ, uvi, ufi, ʔuhi, ʔufi, up, owi, nufi, yui*

B: *sakut, suhat, šakuɣ*

C: *taʔruću, baḷiloko, lame kaloko?*

D: *kai, ku, kūk, koko*

E: *damu, na-^{n}ḏ^{r}am*

F: *pulungan, tila, ambon, tuka, maŋa-t, k^{w}al tulir, sɛsul, kana, uaŋai, dabel, gamu, bīdi, ame, eni, go, ganaŋ, taitu, woida, maho, lama, marihi, auh, nuk, nup, û, mʌu, kēp, sepa, βičua, iām^{w}*

Figure 12.10.1: 'Yam' in Austronesian languages in Taiwan and Philippines.

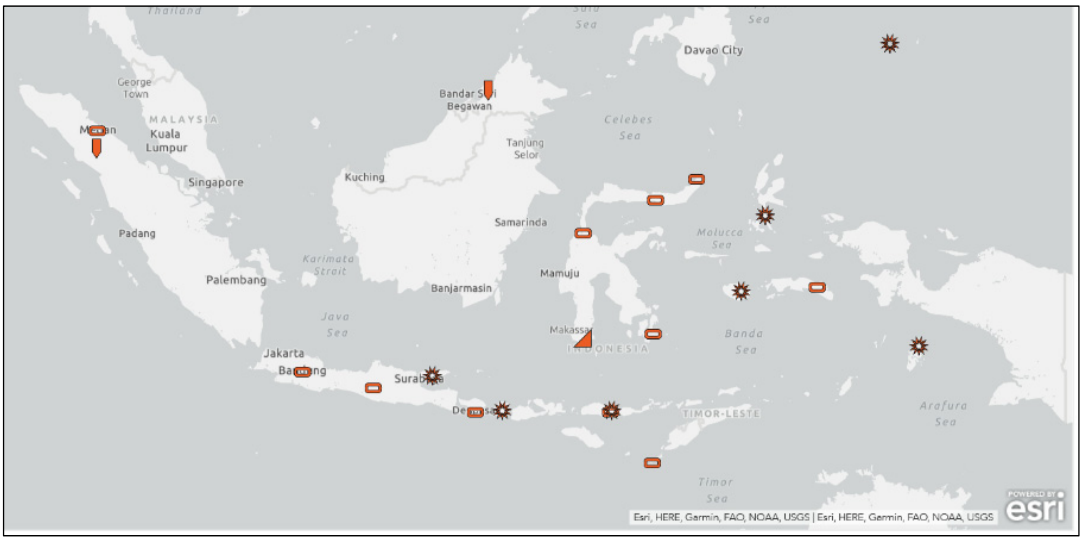


Figure 12.10.2: ‘Yam’ in Austronesian languages in Indonesia.

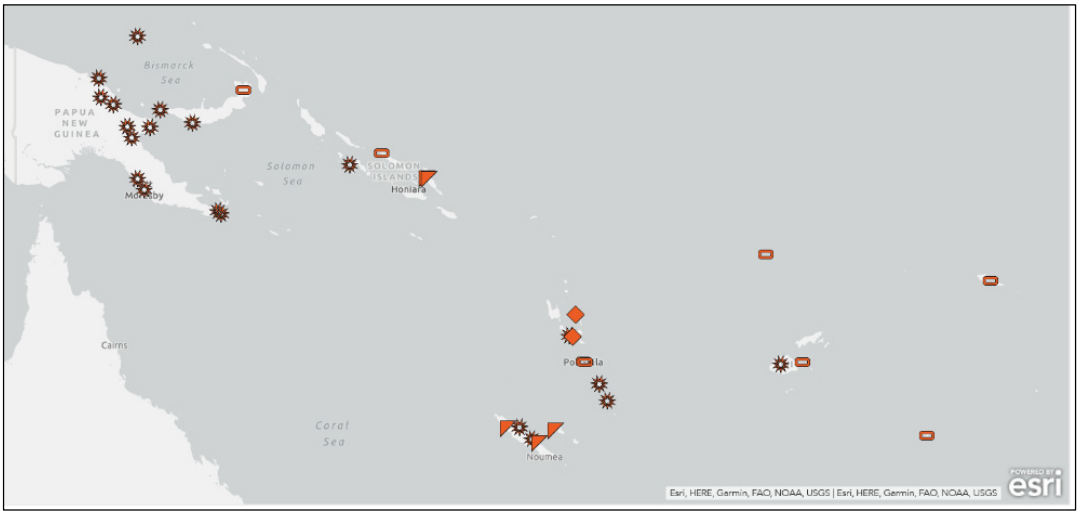


Figure 12.10.3: ‘Yam’ in Austronesian languages in Papua and Pacific Islands.

‘Yam’ in South Asia

I describe the languages of Indo-Aryan (IA), Andamanese, and language isolates in South Asia.

There are two major categories of word forms – A) *ālu(ka)* and B) *konmo* – and three minor categories C) *mīnō*, D) *tarūṭa*, and E) *čāda*. When a language has several words for something like yams, I targeted only Dioscorea terms.

The distribution of ‘yam’ words is relatively simple. On the one hand, Indo-Aryan languages and languages having contact with them employ the types A. On the other hand, on the distant islands, Andamanese languages employ B, C, and E types.

The most major type is *ālu(ka)*, similarly to the cases in ‘Taro in South Asia’. Forms of this type are mainly used by Indo-Aryan languages, but also Nihali and Vedda, which are language isolates, employs as loanwords from IA. Some languages use an *ālu* word to mean both ‘taro’ and ‘yam’ (and ‘potato’, too). Kalasha *kačaḷú* and Vedda *kaṭuvalapojja* are forms of *kacu+ālu* composition which originally referred ‘taro’ (see ‘Taro in South Asia’). Besides them, they compose *ālu* with *piṇḍa*, *rakta*, or something else in many languages. The part *piṇḍa* is the Sanskrit word *piṇḍa* पिंड ‘lump, clod’, and so the total meaning is something like ‘edible lump root’. There is a *piṇḍa+ālu* form for ‘taro’, too: cf. Nepali *piṇḍālū* पिंडालू. While *rakta*, that is from Sanskrit *rakta* रक्त means ‘coloured, dyed, red, blood’, and then the sense of *rakta+ālu* is ‘coloured/red edible root’. For example, Hindi *ratālū* रतालू refers ‘*Dioscorea alata* (purple yam, ダイジョ)’. The literal meaning of Dhivehi *kaṭṭala* ކައްޓަލަ ‘*Dioscorea pentaphylla*, *Dioscorea purpurea*’ is ‘thorned *ālu*’, and the one of Marathi *gōrāḍū* गोराडू is ‘fine / white *ālu*’.

The *konmo* type appears in all the Andamanese languages. Meanwhile in most Andamanese languages, they have either the *mīnō*-type or the *čāda*-type words, with a so complex distribution that we cannot draw an isogloss.

Next, the *tarūṭa* type is used in Nepali and Panjabi. Forms of this type are inherited from Sanskrit *tarūṭa* ‘lotus root’ (see also ‘Taro in South Asia’).

Looking at the remaining sporadic ‘yam’ vocabulary in South Asia, *kurukuru* ކުރުކުރު ‘dioscorea’ in Dhivehi is certainly derived from *kuru* ކުރު ‘short’. In Panjabi they use *jimī̃ kand* ਜਿਮੀਂ ਕੰਦ, which is literally translated as ‘belonging to ground’.

Kusunda has three words for ‘yam’: the first is *yebu* which is used for all kinds of ‘yam’; the second is *datči* (literally ‘Dutch’, maybe) only for ‘*Dioscorea daemonia*’; and the last one is *qom* specifically used for ‘*Dioscorea sativa*’.

(YOSHIOKA Noboru)

A. *ālu(ka)* type (26) □
ằlu, ala, aḷu, alaya
[+x (4)]
kaṭṭala, khamaruā, gōrāḍū, kōnphaḷ
[*kacu*+ (2)]
kačaḷú, kaṭuvalapojja
[*piṇḍa*+ (7)]
piṇḍālū, piṇḍaḷu, pēṇḍāḷū, pinālu, pĩṛālu, hiritalapojja
[*rakta*+ (7)]
ratālū, ratāḷu, ratār, latru

B. *konmo* type (6) ○
kɔnmɔ̃̀, kɔnmu, kune, gōnōda, kharŋe
C. *mīnō* type (3) ▽
mɨ̃́nɔ̃̀
D. *tarūṭa* type (2) \
taraṛī, tarul
E. *čāda* type (2) ◂
čātīda, čārō
E. others
[*karkar*] kurukuru; baáde dhaggwa; datči; jimɨ̃ kand; kaŋgar; krich; munyátiko; qom; yebu

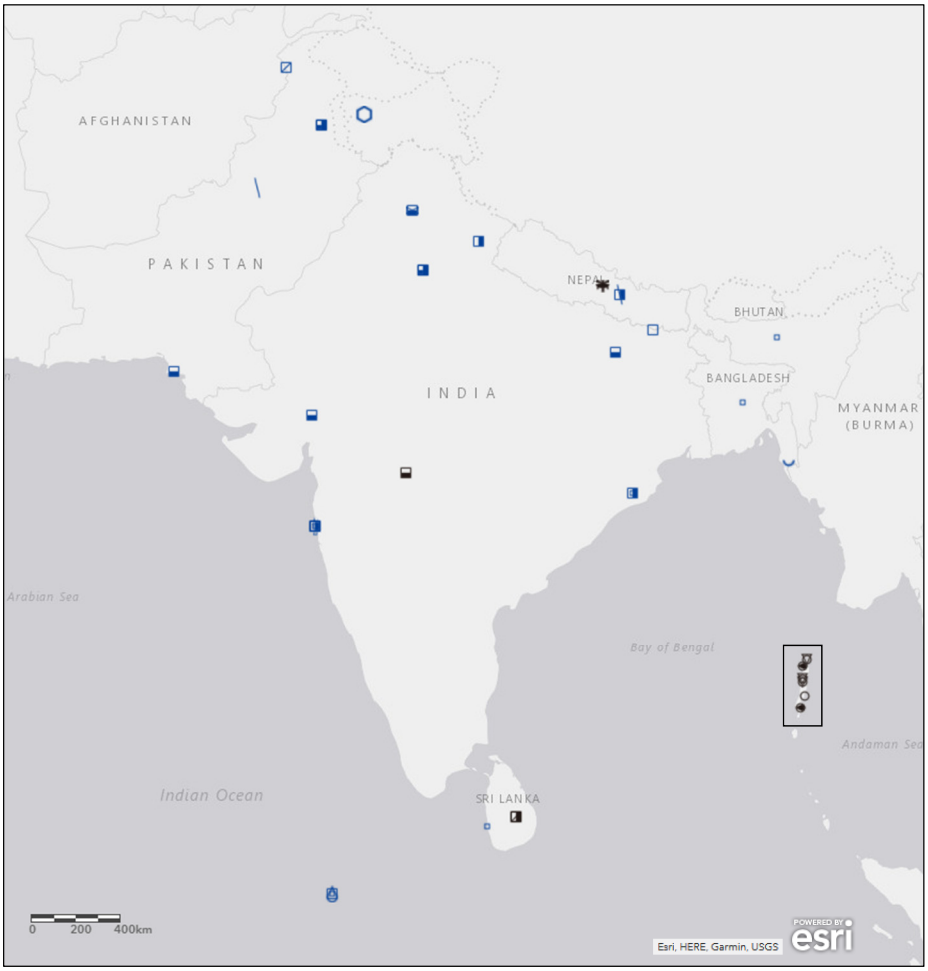


Figure 12.14.1: ‘Yam’ in SA: Indo-Aryan (both in navy blue), Andamanese, and language isolates (those in black).

Figure 12.14.2: Two types of ‘Yam’ on the Andaman archipelago (the area encloed by the rectangle in Figure 12.14.1).

Figure 12.14.3. ‘Yam’ in Kalderaš Romani.

‘Tuber’ in Dravidian

Although some literary Dravidian languages do have a generic term for the various species of the genus *Dioscorea*, such as *peṇḍalam* in Telugu, which is a borrowing from Indo-Aryan *piṇḍālu,* the various edible roots of ‘yam’ are usually covered by a more generic term for a ‘tuber’ which may include more recent introduction such as ‘potato’ or ‘sweet potato’ or ‘cassava’.

The map shows reflexes in the two DEDR etymon entries #1578 and #1683. The former are widely distributed across peninsular India while the latter are found exclusively in the closely related clade of South Central Dravidian. The reflexes *geṇasu* and *genusu* in Kannada and Telugu, which have other generic terms for ‘tuber’ such as *gaḍḍe* or *dumpa*, appear to have specifically referred to *Dioscorea esculenta* ‘lesser Yam’, although almost replaced by ‘sweet potato’ in current usage. Other reflexes retain the generic usage and often form a compound for more specific edible roots.

Kurukh *kisgō* is glossed as ‘yam’ in DEDR entry #2004 for ‘taro’. This may reflect a common confusion between *Colocasia* and *Dioscorea* among foreign lexicographers.

(KODAMA Nozomi)

kiẓaṅku, kiẓaṅṅu, kiṛinj māṭi
kecaŋku
kereṅgů
geṇasu, genusu, geṇagi, gereŋgi
kirre
kuṇi, kuna, kunna, kūṇi, kune
kisgō

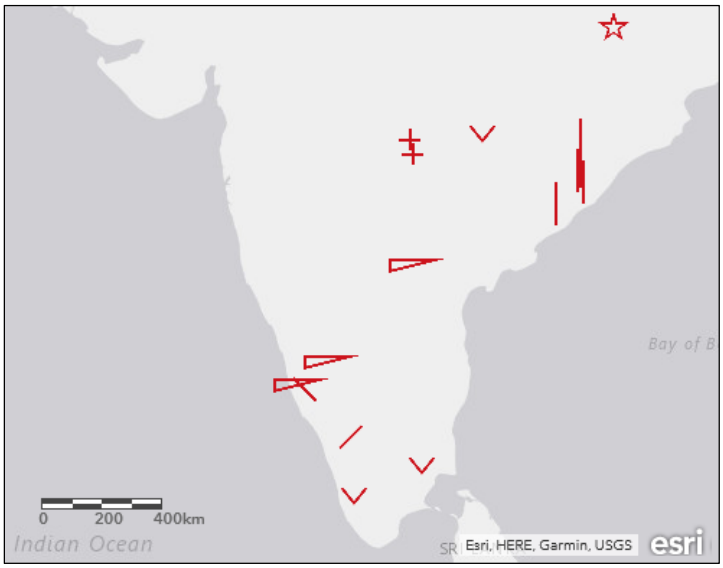


Figure 12.15.1: ‘Tuber’ in Dravidian.

‘Yam’ in Nilo-Saharan

Two roots are heuristically reconstructed (marked by #) for ‘yam’ in Nilo-Saharan.

Type A *#badho* is commonly attested in Surmic (Murle *badhoc*, Suri/Tirma *bayoy*) and Western Nilotic (Dinka *badha*, Alur *obadho*). Two Central Sudanic languages spoken in South Sudan, Ma’di (*bado*) and Beli (*moto*) attest the same root, probably due to contact with Western Nilotic.

Type B is mostly attested in Central Sudanic languages (Jur-Modo *manga*, Gula *manga*, Sar *mang*) with the notable exception of Western Nilotic Reel (*amaŋ*).

Saharan Kanuri attests *doya*, a borrowing from Chadic Hausa (Type C) in addition to another native term *bərma*. Such other native terms (Type D) occur in only one language, and their etymologies are unknown.

(NAKAO Shuichiro)

A: #badho
B: #manga
C: Hausa borrowing
D: the other native terms

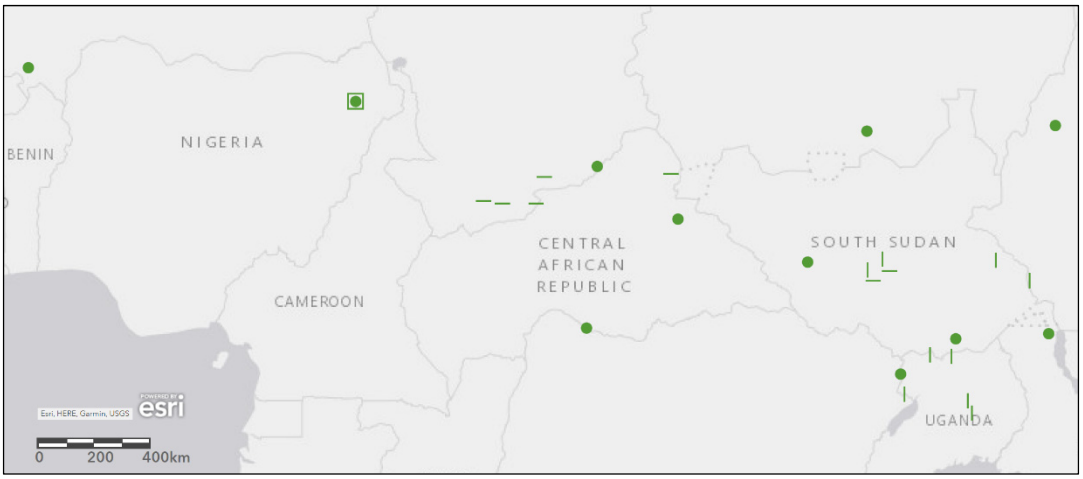


Figure 12.19.1: ‘Yam’ in Nilo-Saharan.

Chapter XIII

Stop series

Stop series in Asian and African languages

This project overviewed the stop series sound system in more than 2500 Asian and African languages (including regiolects and vernaculars if available), with the main focus being on dental/denti-alveolar/alveolar (henceforth D/A) plosives, and nasal sounds.

1. Data components

The D/A stop series was primarily selected for this project. This series sought to determine the most complex patterns in the articulatory positions in most languages, dialects, and varieties. While some languages may distinguish dental and alveolar sound, this distinction was secondary. Dental sounds typically appear in Indo-Iranian and Dravidian languages, denti-alveolar sounds appear in Sinitic and Tibeto-Burman languages (see Zhu's 2010 system), and alveolar sounds appear in Japanese. Note that these articulatory positions are more complex when there are fricatives.

The project examined a system of D/A stop series, based on the following components /th-t-t'-d-dh-ɗ-nd-nt-nth-n-nh/ (e.g., [t^h-t-t'-d-$d^ɦ$-ɗ-^{n}d-^{n}t-$^nt^h$-n-n̥] for a phonetic description). Plosives and nasals were included as /n/ can be regarded as a *nasalised* stop in phonetics. Note that there are also nasalised fricatives; for example, see 'rhinoglottophilia' by Matisoff (1975) and 'nasalised the aspiration' in Suzuki (2015). However, although affricates are members of the stop, they were excluded. Other sounds, such as /d'/ [t̪] (Tibeto-Burman), /ʔn/ [^{ʔ}n] (Hmong-Mien; Austroasiatic), /ʔd/ [^{ʔ}d] (Austroasiatic), /'t/ (Korean), and /$t^ʕ$, $d^ʕ$/ (Semitic) were also properly counted as data for the project. Ejectives and clicks were also included even though their geographical distribution is limited. Ejectives are pervasive in Caucasian languages, of which types such as /t^h-t'-d-n/ (Kartuli) /t^h-t-t'-d-n/ (Lezgi) are attested (Klimov 1994). Clicks (sounds including a dental /ǀ/) are found in the Kalahari Basin Area and can be combined with voicing, aspiration, and ejective features.

In other words, using this model, the historical changes and the plosive typologies were examined based primarily on the phonation or laryngeal features, and the geographical distributions encoded. Non-pulmonic sounds were also included in the description to elucidate the potential correlations and interactions between non-pulmonic and pulmonic sounds synchronically and diachronically and to determine their geographical distribution in more detail than in previous works such as WALS (Dryer and Haspelmath eds. 2013). This approach also differed from the theoretical, typological analysis in Duanmu (2016).

Prenasalisation was also included as a potential feature to trace some crucial sound changes in phonological systems; however, the preaspiration and postnasalisation 'series' by Maddieson (1984) were excluded. Preaspiration has a crucial function in the consonant system in several languages, such as Tibetic (Tibeto-Burman) and Saami (Uralic). The preaspirated consonants in these languages

are respectively derived from consonant clusters and long consonants (Suzuki 2011b; Korhonen 1981; Iosad 2022).

There were also challenges. For example, there were discrepancies in the traditional and individual phonological analysis preferences and the phonetic notation customs (despite the existence of the International Phonetic Alphabet and its extended edition, extIPA; Ball et al. 2018). For example, it has been disputed whether /d'/ [t̪] (Tibetic) is an independent consonant phoneme or a consonant /t/ with a breathy suprasegmental (tonal) feature. The classification terminologies also differed, such as the use of 'fortis/lenis' rather than 'voiceless/voiced'. In several Uralic languages, the plosives voicing contrast is understood as fortis/lenis, with the 't' and 'd' described as /t/ [t^{h}, t] and /d/ [t, d̥] in Northern Saami. In this case, it was debatable whether /t-d/ (as in Nielsen 1979) or /t^{h}-t/ should be used (as in Nickel 1994) in the project. It was, therefore, necessary to explain the invisible phonetic features in the phonological analysis.

Table 1 shows the principal forms in our data, namely, Ainu, Japonic, Sinitic, Kra-Dai, Tibeto-Burman, Mongolic, Turkic, Indo-Aryan, Nuristani, and Dravidian. Data from langauges in Africa are unavailavle.

2. Types of the D/A stop series

A two-way distinction is the minimum D/A stop series system, in which the /t-n/ components are most widely attested in languages such as in Ainu, Japonic, Austroasiatic, Austronesian, Uralic, Turkic, Arabic, Nilo-Saharan, and Niger-Congo. The striking features are summarised below following the language families and groups presented in the project.

Voicing contrast

A contrast between voiceless and voiced plosives is attested in Japonic, Sinitic, Hmong-Mien, Kra-Dai, Tibeto-Burman, Austroasiatic, Austronesian, Tungusic, Uralic, Mongolic, Turkic, Indo-Aryan, Burushaski, Dravidian, Iranian, Armenian, Nilo-Saharan, Niger-Congo, Tuu, Kx'a, and Khoe-Kwadi.

Aspiration contrast

A contrast between voiceless aspirated and voiceless nonaspirated plosives is attested in Korean, Sinitic, Hmong-Mien, Kra-Dai, Tibeto-Burman, Austroasiatic, Austronesian, Mongolic, Turkic, Indo-Aryan, Burushaski, Dravidian, Iranian, Armenian, Nilo-Saharan, Niger-Congo, Tuu, Kx'a, and Khoe-Kwadi.

In Iranian languages, an aspirated feature has been derived from a voiceless sound in the voicing contrast. In these cases, the contrasts between the voicing and aspiration are mutually related. As suggested in the Sinitic and Tibeto-Burman language descriptions, a part of the words that have aspirated features is derived from the voicing contrast.

The aspirated voiced plosive /dh/ (/d^{ɦ}/) is attested in Sinitic, Tibeto-Burman, Indo-Aryan, Dravidian, Iranian, Niger-Congo, Tuu, and Kx'a. Languages with this phoneme tending to have a voiceless aspirated counterpart; however, it is not a prerequisite, as seen in Sinitic and Dravidian.

Contrasts consisting of plosive voicing and aspiration combinations

A tripartite contrast /th-t-d/ is widely attested in Sinitic, Hmong-Mien, Kra-Dai, Tibeto-Burman, Austroasiatic, Austronesian, Indo-Aryan, Burushaski, Armenian, and Niger-Congo.

Other tripartite contrasts comprising voicing and aspiration distinctions are marginally attested, such as the /th-d-dh/ (Sinitic) and /t-d-dh/ (Dravidian).

Indo-Aryan is a typical language that has a quadripartite contrast of plosives, such as /th-t-d-dh/. This type is also attested in Kx'a.

Ejectives

Ejective sounds are restricted to a voiceless feature /t'/ in the languages mentioned in the project, except for Tuu, which has a click voiced ejective /gǃ'/. /t'/ is attested in Chukotko-Kamchatkan, Iranian, Semitic, Nilo-Saharan, Niger-Congo, Tuu, Kx'a, and Khoe-Kwadi. In Semitic languages, the ejective /t'/ can be related to the emphatic *t*, which appears as a pharyngealised feature /tˤ/ in many Arabic languages.

Implosives

Implosives are usually voiced; however, a voiceless counterpart was also found. The voiced implosive /ɗ/ is attested in Sinitic, Kra-Dai, Tibeto-Burman, Austroasiatic, Austronesian, Indo-Aryan, Semitic, Nilo-Saharan, and Niger-Congo. The voiceless implosive /ƭ/ is attested only in Niger-Congo as a phonemic status.

As suggested by Li (1977), implosive sounds are related to glottalised sounds such as /ʔd/ in Kra-Dai, and is also possibly true in Sinitic and Austroasiatic. Shuichiro Nakao (p.c.) suggested that it is possible that the /ɗ/ in Semitic, Nilo-Saharan, and Niger-Congo languages spoken near Lake Chad is phonetically realised as [ʔd].

Prenasalisation /ⁿd-ⁿt-ⁿtʰ/

Prenasalised plosives are attested in Japonic, Sinitic, Kra-Dai, Tibeto-Burman, Austronesian, Nilo-Saharan, and Niger-Congo. While the voiced prenasalised sound is pervasive in these languages, Tibeto-Burman, Austronesian, and Niger-Congo also have voiceless (and aspirated) counterparts.

Prenasalisation is both posited as a more archaic form (Japonic, Kra-Dai, Tibeto-Burman, etc.) and a newly emerged form (Japonic, Sinitic, etc.). De-prenasalisation (*ⁿd* > *d*) is attested in Japonic and Tibeto-Burman, and progressive assimilation (*ⁿd* > *n*) is also attested in Tibeto-Burman.

Pharyngealisation /tˤ-dˤ/

Pharyngealised plosives are attested in Iranian, Semitic, and Nilo-Saharan. The voiced type /dˤ/ is not attested in the majority of Nilo-Saharan.

Voiceless nasal /n̥/

A voiceless nasal /n̥/ is attested in Hmong-Mien, Tibeto-Burman, Austro-asiatic, Austronesian, Uralic (Saami), and Iranian. In Tibeto-Burman languages, voiceless nasals have a clear origin derived from a consonant cluster of /s/ and a nasal (Matisoff 2015).

Glottalised stops

Glottalisation has two types: preglottalised and postglottalised. A preglottalised plosive /ʔd/ is attested in

Sinitic, Kra-Dai, Hmong-Mien, and Indo-Aryan. A preglottalised nasal /ʔn/ is attested in Japonic (Ryukyuan), Hmong-Mien, Kra-Dai, and Tibeto-Burman, and a postglottalised stop /dˀ/ is found in Indo-Aryan (Bishnupriya).

Other than the mentioned features, several marginal features were also found. The Korean /'t/ is so striking that its phonetic status is still being debated (cf. Kim and Duanmu 2004, Duan and Zhu 2018). An aspirated nasal /nɦ/ is attested in Indo-Aryan, and while a lack of nasal sounds was noted in some Sinitic, Kra-Dai, and Tibeto-Burman languages that originated from a merger from /n/ into /l/, it does not mean that all nasals are lacking in each sound system. Various click sounds are found in the language families of the Kalahari Basin Area, namely Tuu, Kx'a, and Khoe-Kwadi.

3. Geographical relationships over language families

The description here focuses on the features characterised by the cross-linguistic geographical distribution described earlier.

Aspiration+voicing quadripartite plosive series

The /th-t-d-dh/ series is attested in languages spoken in South Asia and around the Himalayas, such as Tibeto-Burman, Indo-Aryan, Dravidian, and Iranian. As suggested in the Tibeto-Burman and Dravidian language group descriptions, this series is attributed to Indo-Aryan language contact. It is also noteworthy that the quadripartite plosive series appeared in the Brāhmī script system (third century BCE; see Machida 2001).

The same series is also attested marginally in Sinitic languages, but is not related to Indo-Aryan.

Ejective /t'/

The ejective sound is found around the Caucasus, in Ethiopia, in easternmost Siberia, the Kalahari Basin Area, and southernmost Africa. In Ethiopia, both Nilo–Saharan and Semitic languages have an ejective. As suggested in the Semitic language description, ancient Semitic languages that were distributed in Mesopotamia and Syria, such as Akkadian and Ugaritic, had ejectives. In the Caucasus region, the ejective plosive is pervasive in Caucasian languages (Kartvelian, Abkhazo-Adyghean, and Nakho-Daghestanian; see Alekseev 1999) as well as in the Ossetic (Iranian) languages in that region.

Implosive /ɗ/

An implosive /ɗ/ is attested in various language groups. Some cases have not been analysed as genetic features but as contact-induced acquisition. For example, the /ɗ/ attested in Tibeto-Burman languages is a feature that was acquired through Austroasiatic language contact.

Pharyngealisation

A systematical pharyngealised consonant feature /tˤ-dˤ/ is mostly attested in Semitic languages. This feature also expands to Iranian to the west and Nilo-Saharan to the south that connect to Semitic-speaking regions. Nilo-Saharan languages with pharyngealised features mainly possess /tˤ/

as do some Semitic and Iranian languages. As suggested in the Semitic language descriptions, pharyngealisation is related to ejective sounds, which are generally called 'emphatic consonants'.

As a reference, pharyngealised sounds have also been attested on vowels in Tibeto-Burman (see Evans 2006; and Suzuki 2011a). The sounds are often related to velarisation, uvularisation (Gong 2019) and retroflex (Suzuki 2013). In Tibetic languages, the pharyngealised sounds are derived from a consonant /r/ preceding a vowel.

Historically, pharyngealisation is reconstructed in Old Chinese (Baxter and Sagart 2014); however, no pharyngealised sounds remain in modern Sinitic languages.

Prenasalisation

Languages with prenasalised features are mainly found in East Asia, Austronesian areas, and middle Africa. These features appeared due to internal phonological development rather than language contact acquisition.

Voiceless nasal /n̥/

This sound is principally attested in East and Southeast Asia. However, as this is a feature derived from individual sound development processes in each language group, it is not considered a regional feature.

Glottalised stops

Preglottalised stops (both plosives and nasals) are mainly found from East Asia to South Asia, with the preglottalised plosive often being related to an implosive. Preglottalised nasals are found in the Ryukyu islands, and as suggested in Japonic descriptions, are attributed to internal sound change processes.

Lack of D/A nasals

A lack of D/A nasal sounds is attested in some Sinitic, Kra-Dai, Tibeto-Burman, and Niger-Congo languages, the first three of which are spoken in East and Southeast Asia. However, it appears that both Sinitic and Kra-Dai independently developed a 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 was because of Sinitic language contact with Sinitic (Southwestern Mandarin).

The present analysis revealed detailed regional connections between the striking phonological features within and beyond language groups. By drawing up linguistic maps, it is possible to assess how language contact occurred and functioned in given areas.

(SUZUKI Hiroyuki)

Stop series in Chukotko-Kamchatkan

1. Classification

1.1. Chukotko-Kamchatkan languages

Chukotko-Kamchatkan languages have six vowels: /i, e, a, o, u, ə/ (Kurebito et al. 2001).

Chukchi has 14 consonants: /p/, /t/, /k/, /q/, /ʔ/, /s/, /ɣ/, /ɬ/[ɬʲ], /w/, /j/, /r/, /m/, /n/, and /ŋ/ (*ibid.*). The alveolar plosive /t/ is voiceless, and there is no distinction between aspirated and unaspirated stops. In Chukchi, /t/ appears in word-initial, medial, and final position: *tawtaw* 'bark', *ɬiɬit* 'mittens' (Skorik 1961, Inenlikej 2006).

Koryak has 18 consonants: /p/, /t/, /tʲ/[tʲ], /k/, /q/, /ʔ/, /c/[ʧ], /v/, /j/[ʒ~j̵], /ɣ/, /ʕ/, /m/, /n/, /nʲ/, /ŋ/, /l/, /lʲ/, and /w/ (Kurebito et al. 2001). The alveolar plosive, nasal, and lateral in Koryak have a non-palatalized/palatalized opposition: /t/-/tʲ/, /n/-/nʲ/ and /l/-/lʲ/ (Kurebito 2009).

Alutor has 18 consonants: /p/, /t/, /k/, /q/, /ʔ/, /tsʲ/ (sʲ), /r/, /v/, /ɣ/, /ʕ/, /m/, /n/, /nʲ/, /ŋ/, /l/, /lʲ/, /w/, and /j/. The Alutor alveolar nasal /n/ and lateral /l/ show a non-palatalized/palatalized opposition: /n/-/nʲ/ and /l/-/lʲ/ (Kurebito et al. 2001).

Itelmen has 26 consonants: /p/ [p], /p'/ [p'], /t/ [t], /t'/ [t'], /k/ [k], /k'/ [k'], /q/ [q], /q'/ [q'], /c/ [ʧ], /c'/ [ʧ'] , /ʍ/ [ɸ~ʍ~x], /w/ [β~w~ɣ], /s/ [s~ʃ], /z/ [z~ʒ], /x/ [x], /χ/ [χ], /m/ [m], /n/ [n], /nʲ/ [nʲ], /ŋ/ [ŋ], /l/ [l], /lʲ/ [lʲ], /ɬ/ [ɬʲ], /r/ [r], /j/ [j], and /ʔ/ [ʔ] (Ono 2020). In Itelmen, /t/ appears in word-initial, medial, and final position: *tuzaʔn* 'you (pl.)', *itχ* 'they, them', *səmt* 'earth, ground'.

Itelmen plosives and affricates have the non-ejective/ejective opposition /p/-/p'/, /t/-/t'/, /k/-/k'/, /q/-/q'/, and /c/-/c'/. /t'/ also occurs in word-initial, medial, and final position: *t'ot'ot'* 'sandpiper'.

Personal pronoun 'you (pl.)' in Chukotko-Kamchatkan

Ch.	Kor.	Alu.	Itl.
turi	tuju	turu(wwi)	tuzaʔn

1.2. Nivkh

Nivkh has 6 vowels /i, ɨ, e, a, o, u/ and 32 (Amur dialect) or 33 (Sakhalin dialect) consonants: /p/, /pʰ/, /b/, /t/, /tʰ/, /d/, /ʧ/, /ʧʰ/, /ʤ/, /k/, /kʰ/, /g/, /q/, /qʰ/, /ɢ/, /f/, /ř/ [r̥], /r/, /s/, /z/, /x/, /ɣ/, /χ/, /ʁ/, /m/, /n/, /ɲ/, /ŋ/, /l/, /j/, /h/, and /v/[v~w] (Sakhalin dialect) (Siraishi & Tangiku 2015, Grudzeva 1997).

Nivkh plosives show an aspirated/unaspirated opposition. Nivkh also shows certain consonant alternations at morpheme boundaries: a) morpheme-initial plosives fricativize following a vowel, a glide, or a plosive; and b) morpheme-initial fricatives are realized as plosives following a fricative or a nasal (Shiraishi 2010).

2. Geographical distribution

See Figure 13.1.1.

(ONO Chikako)

- t-n
- t-t'-n
- t^h-t-d-n

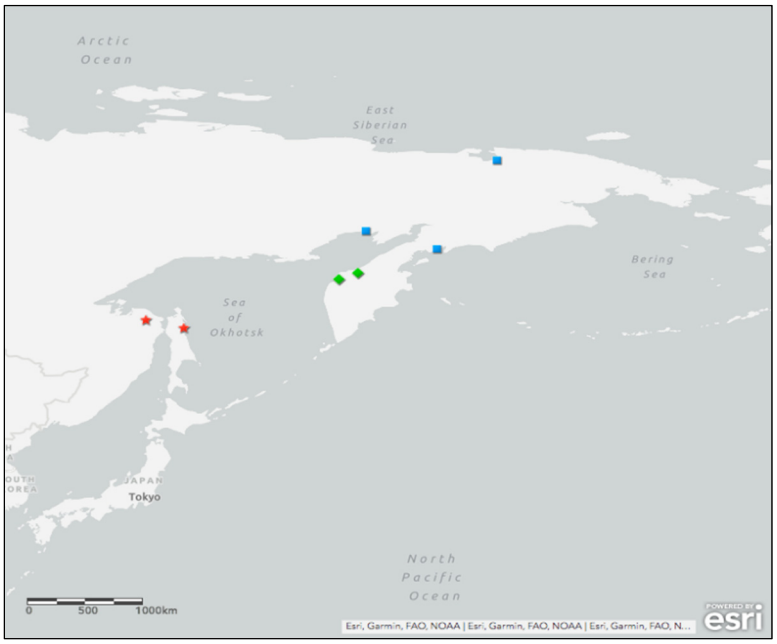


Figure 13.1.1: Stop series in Paleoasian.

Stop series in Ainu

1. Classification of stop series

The Ainu language has five vowels, /i, e, a, o, u/, and eleven (or twelve including a glottal stop, /ʻ/ [ʔ]) consonants, /p/ [p, b], /t/ [t, d], /k/ [k, g], /c/ [ʧ, ʦ, ʤ, ʣ], /s/ [ʃ, s], /m/, /n/, /r/ [ɾ], /w/, /y/ [j], and /h/ [h, x].

The (denti-)alveolar plosive /t/ has no voiced/ voiceless opposition and no aspirated/unaspirated distinction (Tamura 2000: 21). In Hokkaido Ainu dialects, when /t/ occurs syllable-finally, it ends with the closure of the articulatory organs as shown by [t̚]. In most Sakhalin dialects, the coda /-t/ have historically changed to /-h/ [-x].

In the 1792 Japanese-Ainu dictionary "Moshiogusa," which was compiled by a native Japanese translator, the word for 'seashore' was recorded as "オタシヤム" [otaʃam]; however, it was also recorded as "オダシヤム" [odaʃam] in the handwritten copy. Since Japanese has voiced/voiceless opposition, in this case represented by /t/ and /d/, there are often different katakana characters for writing the Ainu /t/ plosive in Japanese materials.

For the dialects of the Kuril Islands, the only existing materials were written by some explorers around the 19th century. We can see both the letters t- and d- for /t/:

WORD	Material A	Material B	
'hand'	*tek*	*dēk*	/tek/
'wing'	*teikup*	*dīkkūp*	/tekup/
'two'	*túup*	*dūpk*	/tup/

(Murayama 1971: 44)

(Murayama compiled the written materials of northern Kuril Ainu. Material A is quoted from Krascheninnikov, S. P., *Vocabularium latino-curilice- chuhachtscha-Kamtschtzice-ukinice*, and material B is Klaproth, J., *Asia Polyglotta* in 1823.)

Ainu also has a nasal stop /n/, which may optionally velarize [ŋ] before [k]. An informant of Ochiho dialect seemed to pronounce the nasal /n/ as [ɴ] in Hattori and Chiri (1960), although that may have been influenced by his second native language of Japanese: *cinkew* [ʧiɴkeɯ] for 'root', *ahto ran* [ɑxto ɾɑɴ] for 'it rains,' *kunne* [kuɴne] for 'black,' and so on.

2. Geographical distribution

See Figure 13.2.1.

(FUKAZAWA Mika)

- A. t-n type

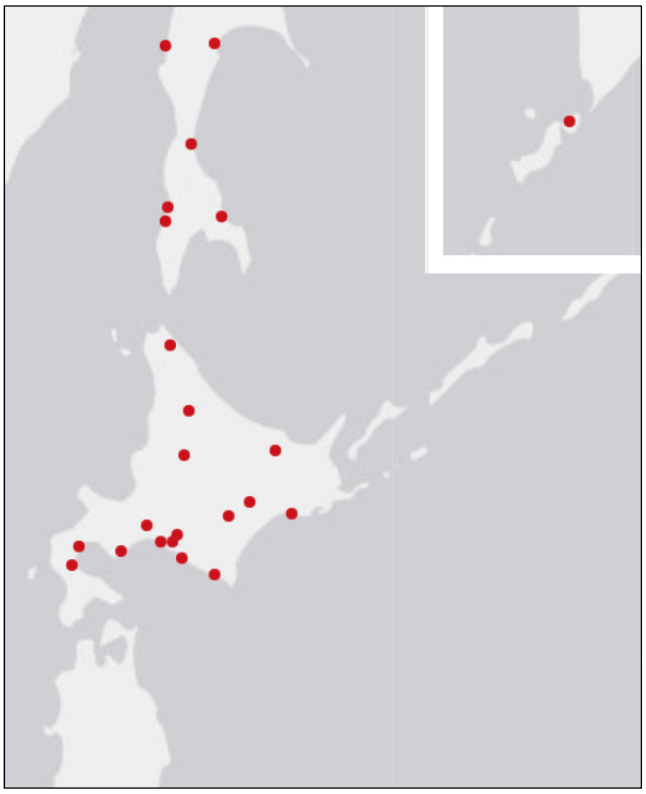


Figure 13.2.1: Stop series in Ainu.

Stop series in Japonic

1. Classification

In the maps, the synchronic types of stop series in Japonic (Japanese and Ryukyuan) are classified into seven categories:

A: t-ⁿd-n, type with prenasalized voiced obstruents
B: t-d-ⁿd-n, type with distinctive prenasalization in the (voiced) obstruents
C: t-d-n, type without prenasalization in the voiced obstruents
D: t-tˀ-d-n-nˀ, type with distinctive glottalization in both the (voiceless) obstruents and nasals
E: t-tˀ-d-n, type with distinctive glottalization in the (voiceless) obstruents
F: t-d-n-nˀ, type with distinctive glottalization in the nasals
G: t-n, type with no voiced obstruents

2. Geographical distribution and interpretation

On the mainland side, Type A is spread across the Kii Peninsula, Shikoku and the area around Kyushu; Type B in the Tohoku region; and Type C in most of the remaining regions. In the Ryukyu Islands, types with distinctive glottalization—D, E, and F are distributed in the Northern Ryukyus from Kikai island to northern Okinawa. Type C is distributed in the Southern Ryukyus (except for Type E in Yonaguni and Type G in Ōgami island) and the southern Okinawa.

Regarding the types with prenasalized obstruents (A and B) and those without (C, D, E, F, and G), we naturally assume that the former underwent phonetic changes to the latter through denasalization. Regarding Type A and B, B—which has more phonemes—appears to be older than A at first glance. However, when focusing on the phonological environment, /d/ of Tohoku dialects (Type B) evolved from intervocalic */-t-/, such as that within *mado* 'target' < *mato; it is distinguished from /ⁿd/, such as that within *maⁿdo* 'window.' In other words, because it is in Type B that the allophones [t] ~ [d] have phonemized, Type A, where this split has not occurred, is older than B. Most of the mainland dialects have shifted directly from A to C through denasalization, but in C of Tochigi, Ibaraki, and part of Iwate, */-t-/ > [-d-] occurred as in B; thus, it is presumed that those areas changed in the order of A > B > C.

In the Ryukyu Islands as well, the proto-system is thought to be Type A: in the Kohama dialect of the Southern Ryukyus, the cluster [nd] corresponds to standard Japanese [d], such as *junda* 'branch' (Jpn. *eda*) and *sundi* 'sleeve' (Jpn. *sode*). In the Northern Ryukyus, the change from Type A to C was followed by a change to D, with glottalized consonants, and then E (loss of glottalized nasals) or F (loss of glottalized obstruents). Glottalized consonants developed from compensation for the loss of preceding syllables, such as *taː* (< *ta) 'rice field' vs. *tˀai* (< *putari) 'two people,' *naː* (< *mipa) 'garden' vs. *nˀa* (< *ima) 'already' in Ie dialect.

In the southern Okinawa and in Southern Ryukyus, Type C has spread, with the exceptions of Ōgami and Yonaguni.

However, C of the southern Okinawa and that of Southern Ryukyus have different processes. In the southern Okinawa, after passing from C to D, E, and F, the glottalized sounds were lost again in the shift to C. Evidence for this theory is found in Shuri dialect, which is one of the southern Okinawa dialects: /m/ and /N/ (moraic nasal) are distinguished from glottalized /mˀ/ and /Nˀ/, and geminate obstruents, such as *ttɕu* 'person' (< *pito), also demonstrate the one-time glottalization. Conversely, there are no traces of the glottalization in the Southern Ryukyus, except for Yonaguni; hence, there was a direct move from A to C in the same way as mainland dialects. Regarding Yonaguni's glottalization, those originating from sound reduction such as *tˀa* 'tongue' (< *sita) (cf. *ta* 'rice field') can be observed as in the Northern Ryukyus; however, the *C > Cˀ / __$V_{[+high]}$, which also happened in the Northern Ryukyus, does not occur; the glottalization occurred through parallel change. The Ōgami dialect is of a rare type (G) that changed from Type C to voiced obstruents becoming voiceless and merging with the voiceless obstruents.

The summary of historical changes is as follows:

A = A (mainland surrounding area)
A > B (Tohoku)
A > B > C (part of East Japan)
A > C (most of the mainland and Southern Ryukyus)
A > C > D (part of the Northern Ryukyus)
A > C > D > E or F (")
A > C > D > E or F > C (")
A > C > E (Yonaguni)
A > C > G (Ōgami)

(NAKAZAWA Kohei and YOKOYAMA Akiko)

A: t-ⁿd-n
B: t-d-ⁿd-n
C: t-d-n
D: t-tˀ-d-n-nˀ
E: t-tˀ-d-n
F: t-d-n-nˀ
G: t-n

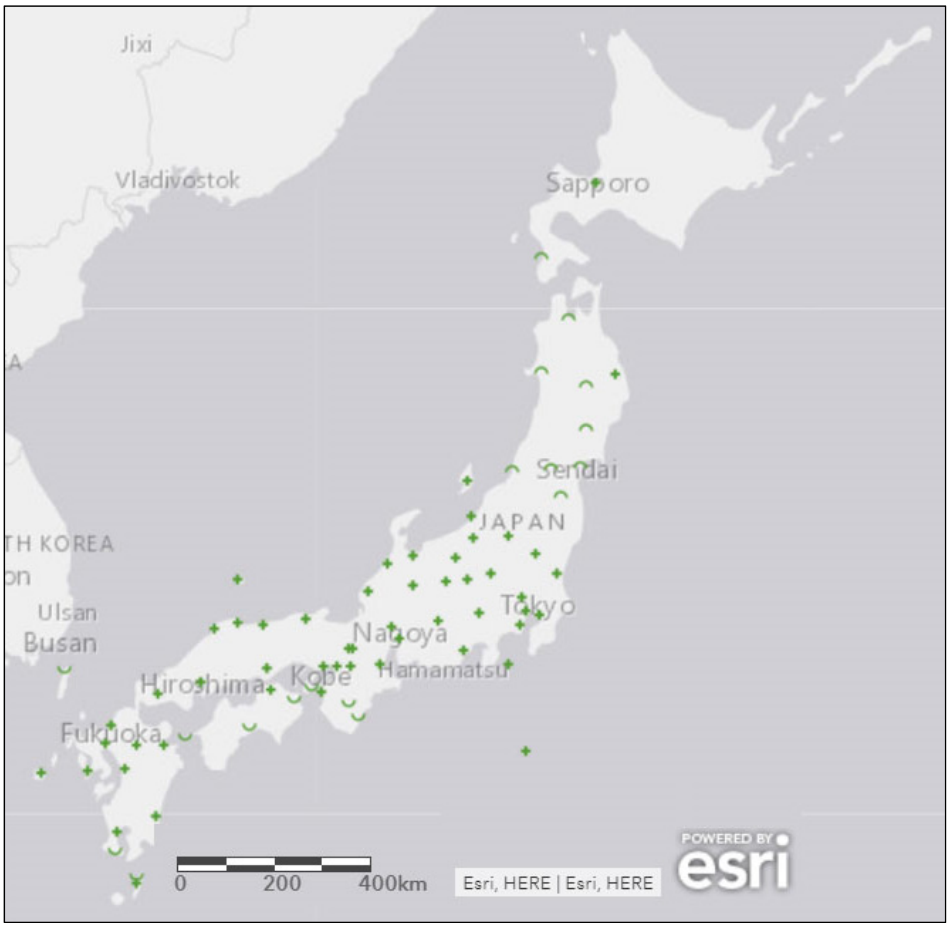


Figure 13.3.1: Stop series in mainland Japan.

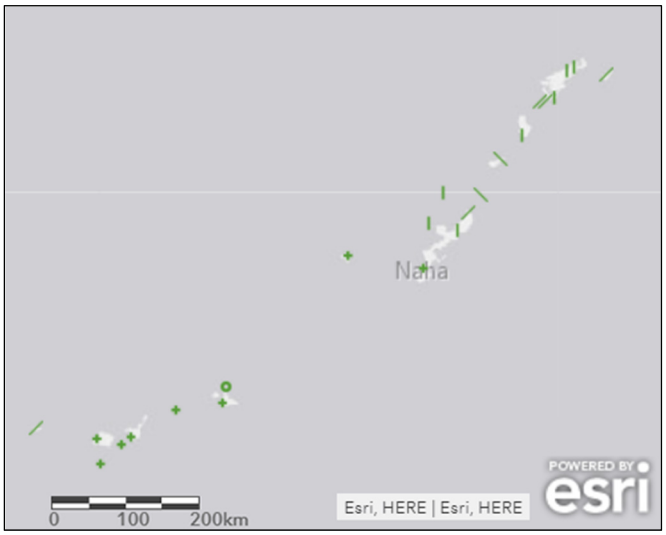


Figure 13.3.2: Stop series in Ryukyu Islands.

Stop series in Korean

1. Classification

It is well known that the Korean language has three oral stops and a nasal stop.

(1) t^h – t – t' –n

(2) Examples:

t^hal 'mask', tal 'moon', t'al 'daughter', nal 'day'

The three oral stops have been called variously according to authors. The following table summarizes terms used to denote these three stops found in a few recent publications.

	t^h	t	t'
Kagaya (1974)	aspirated	lax	forced
Ladefoged and Maddieson (1996)	aspirated	unaspirated	stiff voice
IPA Handbook (1999)	aspirated	lenis	fortis
Lee and Ramsey (2011)	aspirate	plain	reinforced

Traditionally, the three oral stops have been described phonetically as follows:

Aspirate: Characterised by a strong aspiration. Kagaya (1974) observes positive abduction of the vocal folds and heightened subglottal pressure.

Plain: Slightly aspirated initially and (sometimes) voiced intervocalically. With no positive laryngeal gestures.

Reinforced: Voiceless unaspirated. Kagaya (1974) observes a complete adduction of vocal folds before the explosion, stiffening of vocal folds and increasing subglottal pressure and/or lowering of the glottis, and so on.

2. Descriptions on recent Seoul speakers

It has been reported that young Seoul speakers pronounce initial aspirates and plain stops with almost the same amount of aspiration and the distinction between these two types are maintained by a high pitch associated with aspirates (for example, Silva (2006)). This can be called another case of tonogenesis. However, we have to keep in mind that aspirate and plain stops are maintained as such in intervocalic positions so that the merger is not complete yet.

3. Nasal stop

It has been observed that an initial nasal is sometimes pronounced something like [n^d], with the loss of nasality at the release of the oral closure. Such pronunciations can be heard as a voiced stop for speakers of a language having the initial voicing contrast.

There are no dialectal differences except for minor phonetic details. For examples, some dialects, such as the Kyŏngsang dialects, are said to show a smaller amount of aspiration for plain stops.

Historically, reinforced stops are developed from consonant clusters. In Middle Korean we have initial consonant

clusters like sp-, st-, sk-, pt-, ps-, pc-, pst-, psk-, all developed into a reinforced stop.

However, we have in fact pronunciations which seem quite similar to modern reinforced stops in Middle Korean. Such cases appear not within a lexically simple morpheme, appearing only medially in a specific combination of morphemes, and in such cases they used a symbol for the glottal stop or a geminate.

(FUKUI Rei)

$t^h - t - t' - n$

Figure 13.4.1: Stop series in Korean.

Stop series in Sinitic

1. Classification

We collected the published data of 2343 Chinese dialects. Classification is shown below.

A1. /t/ type
Sanya; Jiangmen
B1. /t^h-t/ type
Hefei; Nanjing; Guiyang
B2. /t-^{n}d/ type
Xinhui
B3. /t-n/ type
Duchang; Qionghai; Dongkou
C1. /t^h-t-d/ type
Quanzhou; Xingan; Guanyang
C2. /t^h-t-^{n}d/ type
Doumen; Taishan; Kaiping
C3. /t^h-t-n/
Beijing; Nantong; Taiyuan
C4. /t^h-d-n/
Nanhui
C5. / t^h-d̰-n / type
Wuchuan
C6. /t-d-n/ type
Wenchang; Xingzi; Yueyang;
C.7 /t-d^h-n/ type
Tongcheng; Pingjiang; Chibi
C8. /t- d̰-n/ type
Changhua; Dongfang; Qiongzhong
D1. /t^h-t-d-n/ type
Shanghai; Wenzhou; Yongzhou
D2. /t^h-t-d^h-n/ type
Xuancheng; Shaoxing; Shanghai
D3. /t^h-t-d̰-n/ type
Tengxian; Hezhou; Ledong
D4. /t^h-t-d-n/ type
Taigu; Yanzhou; Ziying
D5. /t^h-d-d^h-n/ type
Songjiang; Yongkang
D6. /t-d-dh-n/ type
Yueyang
D7. /t-d-d̰-n / type
Wenchang; Wanning
E1. /t^h-t-d-d^h-n/ type
Zhengfang; Zhongjiang
E2. /t^h-t-d-d̰-n/ type
Chongpo
E3. /t^h-t-d-n-n^h/ type
Chongming; Jiading; Yangshuo
F1. /t^h-t-d-d^h-n-n^h/ type
Dinghai
G1. /t^h-t-d-d^h-nd-n-n^h / type
Yiwu

2. Geographical distribution and interpretation

The most common type of Sinitic language is C3, which also corresponds to standard Chinese. This type is widely distributed throughout China (Figure 1). (Chao (1968), however, states that in standard Chinese, t actually corresponds to d̰, and is realized as d in the pronunciation of a word.)

The second most common type is D1, which contains d in addition to t, t^h and n. This type is characteristic of Wu and Old Xiang dialect. Centering on the Yangtze River basin, D1 is widely distributed in the southern area (Figure 2). (Chao (1928) also notes that in Wu dialect, d is generally unvoiced, with voiced airflow like [$t^ɦ$] at the beginning of a word, while it becomes voiced when placed between vowels.)

Types E1, E2, and F1 have four variations of t. They are sporadically distributed in Yangtze River basin and in the Hainan island. In Chongming, /t-t^h-n(ʔn)/ and /d-n^h(ɦn)/ both form

complementary distribution by tone. In Zhongjiang, however, the appearance of d^h is only limited to tone III (52), while minimal pairs of t, t^h, d, and d^h exist (Table 1).

Table 1: Four variations of t

E3: Chongming: Zhang (2009) 刀tɔ^{55}, 滔$t^hɔ^{55}$, 逃dɔ^{24}, 拿ʔnɔ^{55}, 挠ɦnɔ^{24}
E1: Zhongjiang: Cui (1996) 逮tai^{21}, 踏t^ha^{21}, 貸dai^{21}, 台d^hai^{21}

Types A1 and B3 have only one kind of t, and the presence of these types indicates that large-scale changes of the initial consonant system are underway in these areas. They are distributed in Jiangxi, Hunan, Guangdong and Hainan. In Sanya and Jiangmen, *t^h is merged with x, and *n is merged with l; therefore only t remains. In Duchang and Dongkou, t^h has merged with l and x respectively (Table 2).

Table 2: Types possessing only one kind of t

B3: Duchang: 刀tau　套lau 桃lau　(Lu 2007)
B3: Dongkou: 刀tau　套xau 桃xau (Long 2008)

Implosive ɗ is distributed in the Hainan island, Guangdong, and Guangxi province.

(YAGI Kenji)

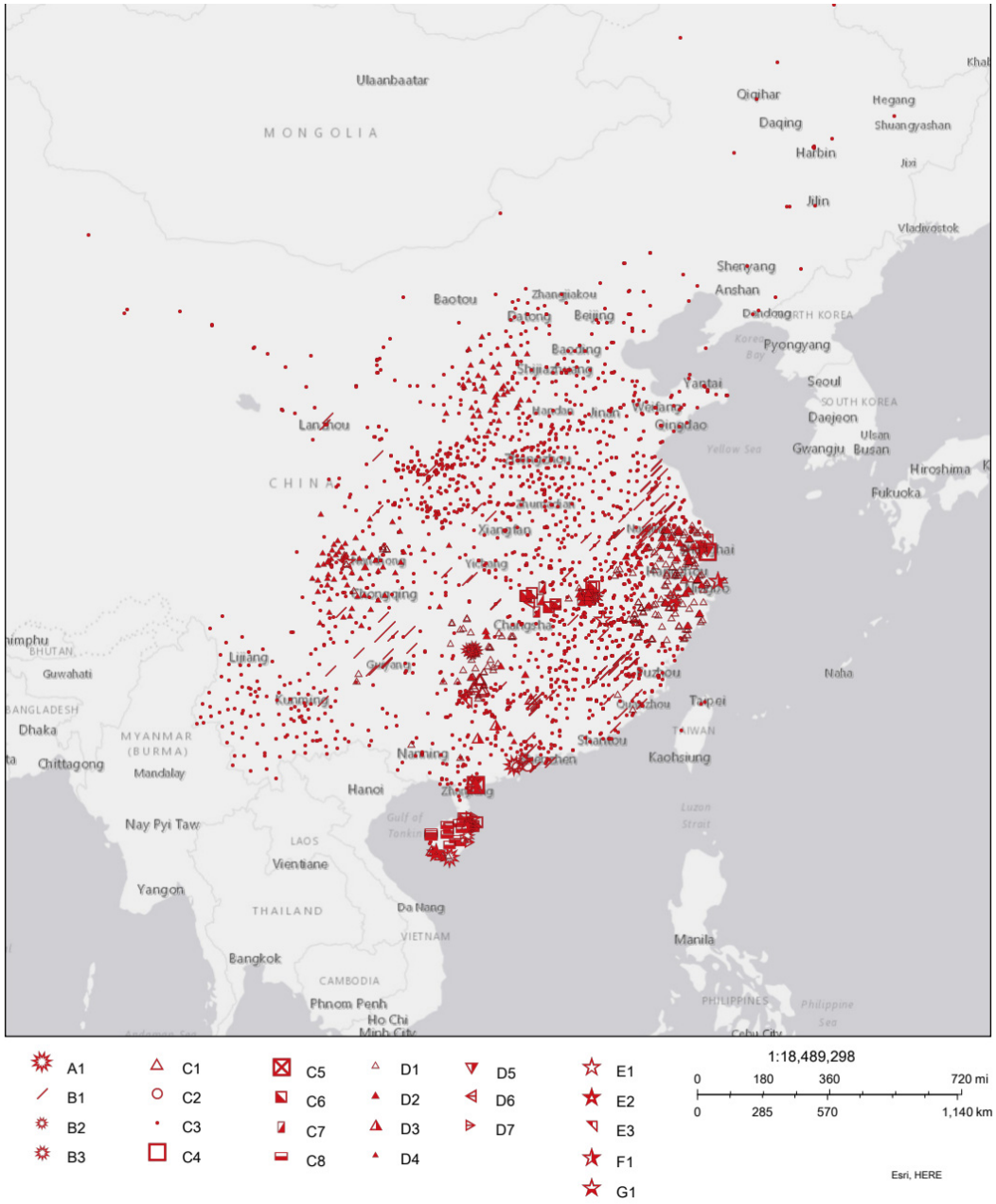


Figure 13.5.1: Stop series in Sinitic.

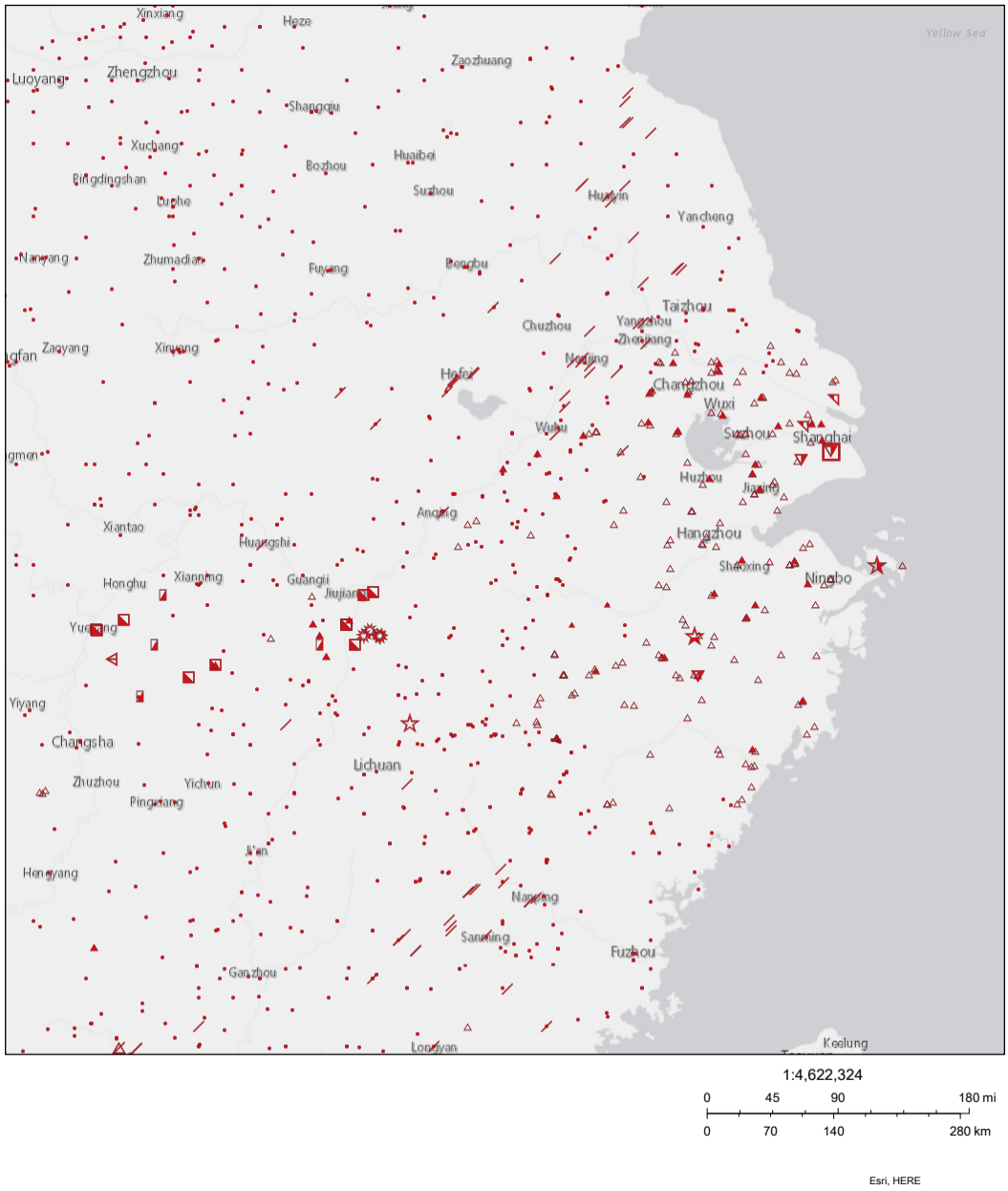


Figure 13.5.2: Stop series in Sinitic (east central area).

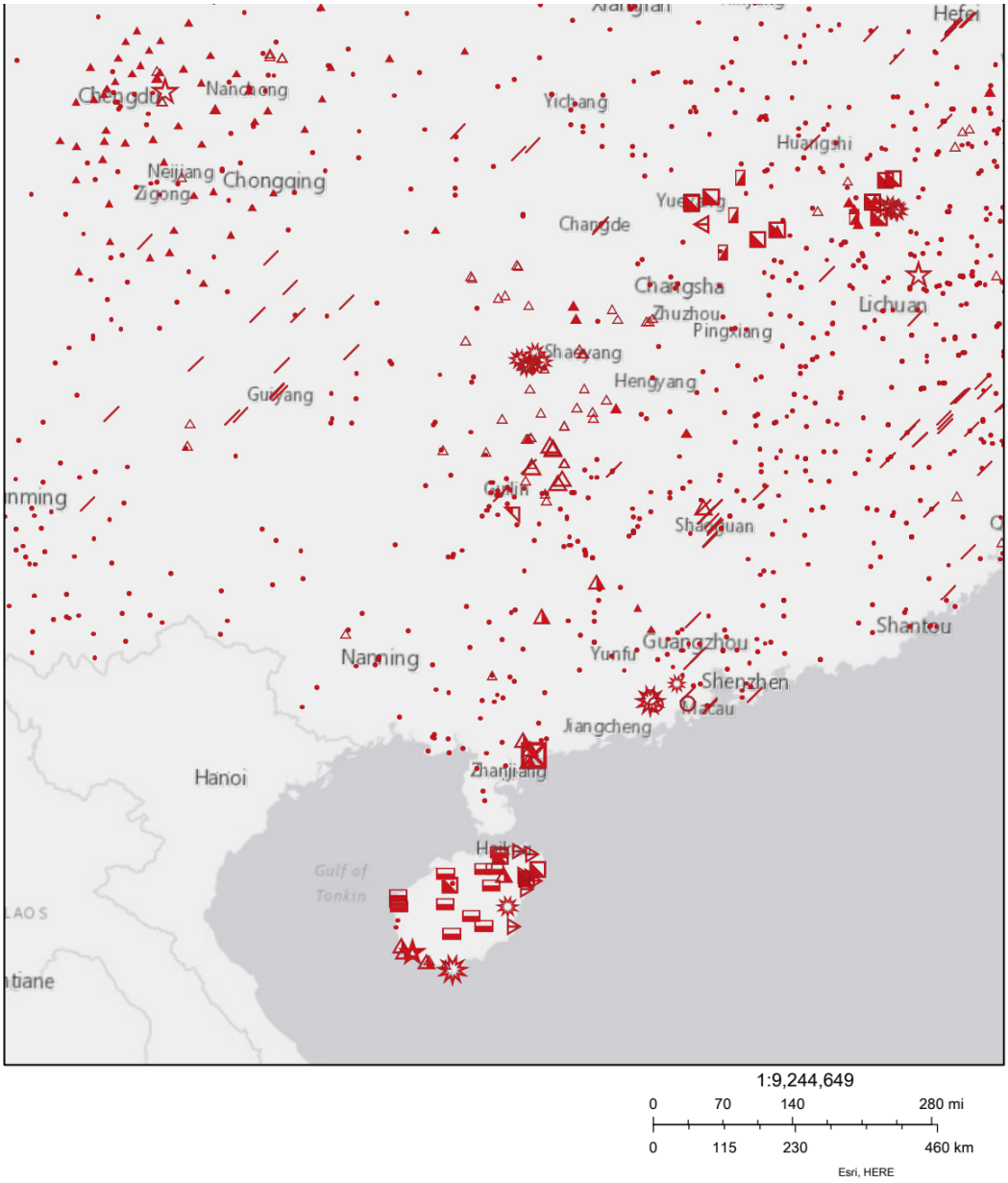


Figure 13.5.3: Stop series in Sinitic (southern area).

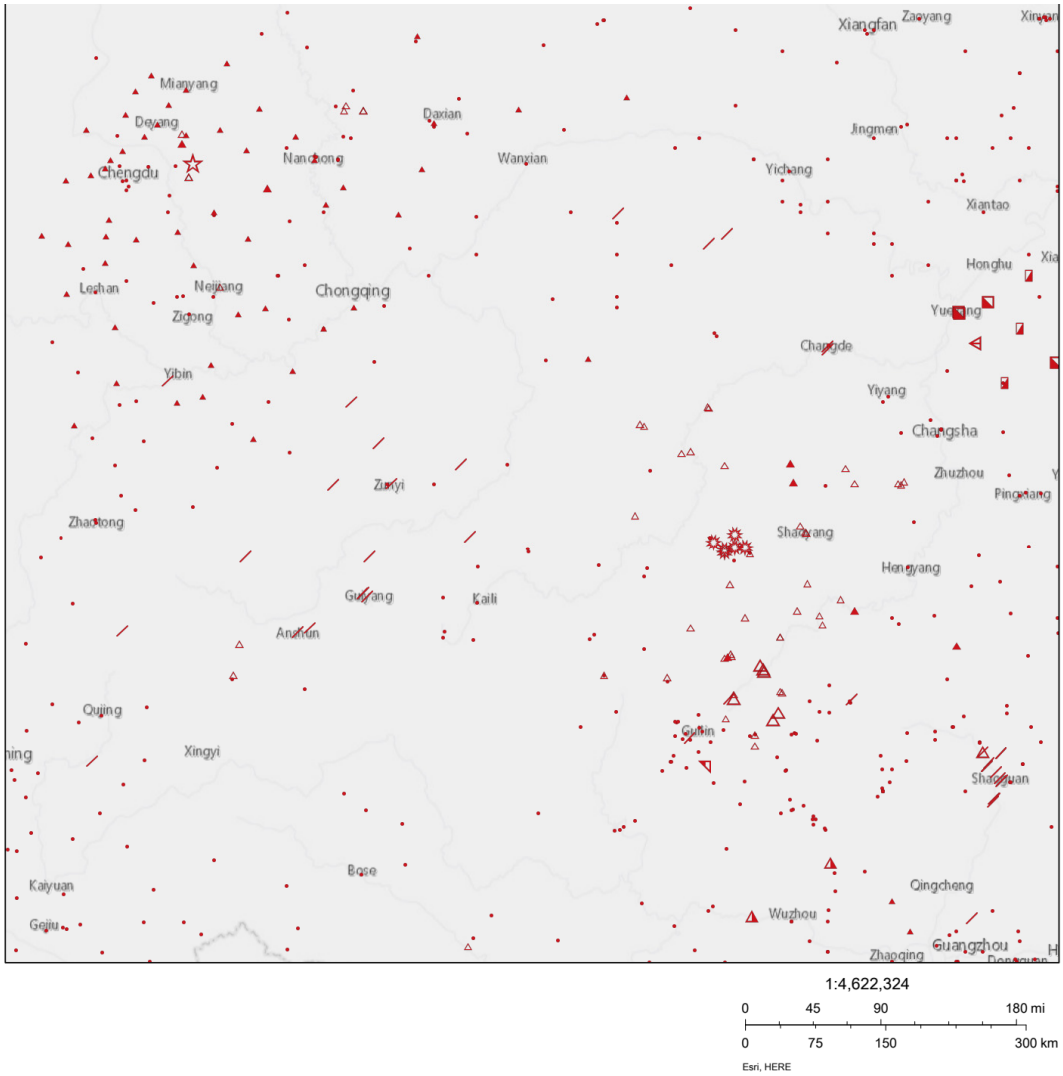


Figure 13.5.4: Stop series in Sinitic (south central area).

Stop series in Hmong-Mien

1. Classification

Based on the comparative evidence, types of consonant series are classified into 18 categories. It is believed that Proto-Hmong-Mien has the following consonant series at the initial position of a syllable: th-t-d-nd-nt-nth-n-nh-ʔn. The first three consonants (th-t-d) are plain plosives, the next three (nd-nt-nth) are prenasalized plosives, and the last three (n-nh-ʔn) are nasals. Further, lects are classified according to their position in the diagram illustrating the historical order of the phonological changes that the proto-consonant series has undergone (Figure 13.6.1). The first change—which is represented by Type B in Figure 13.6.1—is a merger of voiced and voiceless consonants, i.e., *t and *d, *nt and *nd, and *n and ʔn. Five of the 18 types have not undergone the merger. These are classified as subcategories of Type A (A1 to A5). All the other types—the descendants of Type B—have undergone this change. Types C–G signify consonant series that have experienced aspiration and/or prenasalization loss. Type H represents a stage wherein the Type B consonant series undergoes loss of prenasalization and voicing of the plosive (e.g., nt > d).

If it is known that the consonant series of two lects originate from different sources, then these are classified as two different types even if the patterns of their consonant series are the same. The 18 types are indicated below ('0' indicates a gap compared with the consonant series of Proto-Hmong-Mien).

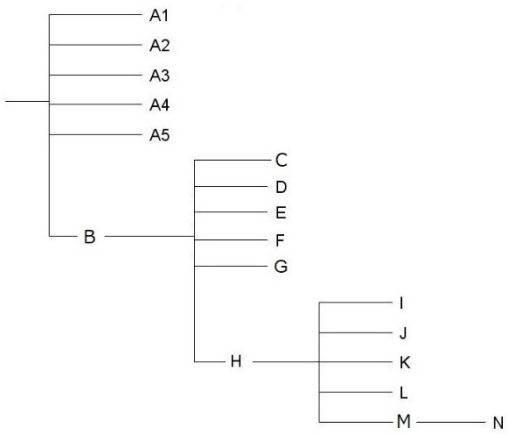
A
- A1: th-t-0-nd-nt-nth-n-nh-ʔn
- A2: th-t-d-nd-nt-nth-n-nh-0
- A3: th-t-tɦ-ntɦ-nt-0-n-nh-nɦ
- A4: th-t-d-0-0-0-n-nh-ʔn
- A5: th-t-d-0-0-0-n-0-0

B: th-t-0-0-nt-nth-n-nh-0
C: th-t-0-0-0-0-n-nh-0
D: th-t-0-0-0-0-n-0-0
E: th-t-0-0-nt-0-n-nh-0
F: th-t-0-0-nt-nth-n-0-0
G: 0-t-0-0-nt-0-n-0-0
H: th-t-d-dh-0-0-0-n-nh-0
I: 0-t-d-0-0-0-n-0-0
J: 0-t-d-ʔt-0-0-0-n-0-0
K: th-t-d-ʔd-0-0-0-n-0-0
L: th-d-ʔt-0-0-0-n-0-0
M: th-t-d-0-0-0-n-nh-0
N: th-t-d-0-0-0-n-0-0

2. Geographical distribution and interpretation

Type A—which exhibits an archaic state—is spread across the northwestern part of the distribution area. The lects that belong to this type constitute relic areas. Type H and its descendants (Types I–N, Figure 13.6.1) are distributed across the eastern and southern parts of the area. Most lects that belong to these types are Mienic languages and Northern Hmongic languages (aka, Xiangxi dialects). They represent more innovative states.

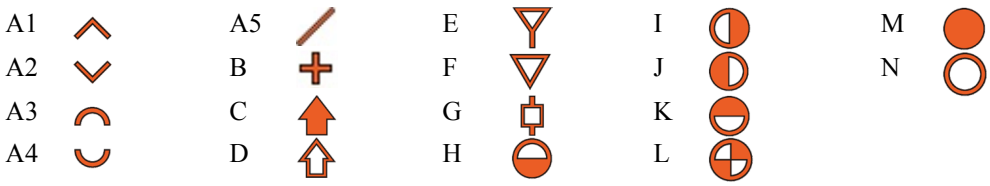
Figure 13.6.3 depicts the distribution of prenasalization in the consonant series. It indicates that prenasalization is observed in the northern and western parts of the area, including the relic area mentioned above. In



the eastern and southern parts, prenasalization tends to drop with or without making the following obstruent element voiced.

(TAGUCHI Yoshihisa)

Figure 13.6.1: xxx



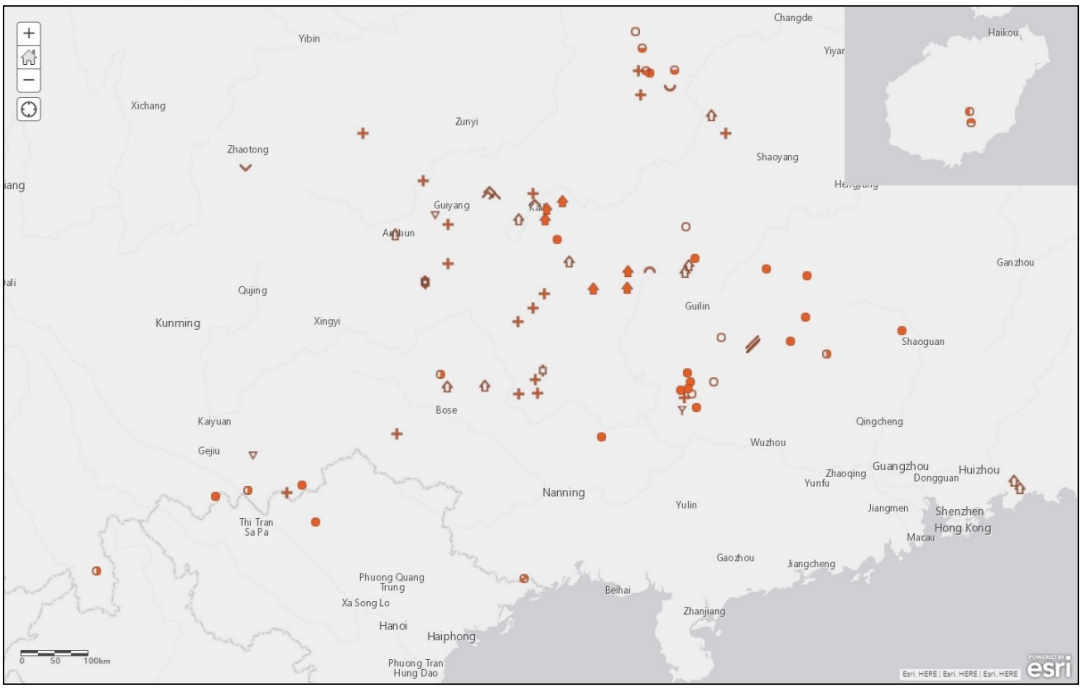


Figure 13.6.2: Stop series in Hmong-Mien.

prenasalization + no prenasalization ○

Figure 13.6.3: Prenasalization in Hmong-Mien.

Stop series in Kra-Dai

1. Classification

There are 13 types of initial dental stop series in Kra-Dai.

A: th-t-ɗ-nd-n-nh-ʔn
B: th-t-d-ɗ-n-nh
C: th-t-d-nd-n-ʔn
D: th-t-d-n-nh
E: th-t-ɗ-n-ʔn
F: th-t-n-nh
G: th-t-n-nt
H: th-t-ɗ-n
I: th-t-n
J: th-ɗ-n
K: th-t
L: t-ɗ-n-ʔn
M: t-ɗ-n

2. Geographical distribution and interpretation

Type H, which is indicated by a small dot, is the most widespread variety, as it is found across the whole Kra-Dai area and includes Bangkok Siamese and Lunchow Zhuang. The corresponding rule between proto-Tai and Type H is as follows (Li, 1977):

proto-Tai	*t-	*th-	*d-	*ʔd-	*n-	*hn-
Siamese	t-	th-	th-	d-	n-	n-
Lungchow	t-	th-	t-	d-	n-	n-

Li (1977:107) described the phonetic nature of *ʔd-, as follows: “This consonant is preserved as a preglottalized consonant ʔd- in Wu-ming, but is represented by d- in most dialects – at least so transcribed. It is generally pronounced with some laryngeal stricture and depression, and may even be implosive in the pronunciation of certain speakers.” In this study, [ɗ] is used to indicate this sound, but it can also be transcribed as d-, as mentioned above. It is noteworthy that the pure voiced consonant *d- has disappeared in almost all Kra-Dai dialects, with lower series of each tone emerging in its place. This has also occurred in a majority of Sinitic dialects and it appears that voiced initial consonants easily disappear in tonal languages.

The descendants of *ʔd- occur in the upper tones, and this series lacks a velar counterpart. These properties indicate that the voicing is due to later innovations caused by a change into an implosive.

Voiceless nasals also existed in proto-Tai, which in the modern Siamese orthography are indicated by clusters preceded by h- that occur in the upper tones and are merged with the ordinary nasals.

Type M, which is indicated by the round symbol, is the second most frequent type. Type M, in which the aspirates merge to non-aspirates, is found in Northern Zhuang, Southern Buyi, and Be. While the sound change from Type H: th-t-ɗ-n to Type M: t-ɗ-n is widespread in these areas, some scholars have postulated a reversed direction of change, claiming that the aspirates emerged later.

There are four other types occurring in the Tai branch. Type E: th-t-ɗ-n-ʔn and Type L: t-ɗ-n-ʔn, which are more conservative and have glottalized nasals, are distributed in the east of the Kra-Dai area next to the non-Tai branches, as shown in the map below. These glottalized nasals are possibly retention of an archaic

distinction. Type L also has a de-aspiration and in Type I: th-t-n, the *ʔd- has changed to n- and in Type K: th-t, the n- has changed to l-.

In the southern group, Hlai and Be from Hainan island have experienced the same changes as in the Tai branch, that is, Type H > Type M. There is also a Type J: th-ɗ-n in the Cun language, which is adjacent to Hlai. In this type, the t- is lacking and is replaced by tθ-, which may have been because of a change from t- > tθ-.

The northern Kra-Dai, Kra, Lakkia, and Kam-Sui groups have more and less complicated systems. For example, the Sui language has the most complicated Type A: th-t-ɗ-nd-n-nh-ʔn system, and there is also a pre-nasal voiced stop in Type C: th-t-d-nd-n-ʔn. Voiceless nasals are preserved in Type A and Type B: th-t-d-ɗ-n-nh, D: th-t-d-n-nh, and F: th-t-n-nh, and in Type G: th-t-n-nt, there is a voiceless stop after the nasal. The geographical distribution of these conservative types is scattered, and it is hard to tell why they occurred from a comparative linguistic point of view.

Cognate words are relatively difficult to find between the Tai and non-Tai branches; therefore, the sound correspondences are also less stable than in the Tai branch.

(ENDO Mitsuaki)

A: th-t-ɗ-nd-n-nh-ʔn
B: th-t-d-ɗ-n-nh
C: th-t-d-nd-n-ʔn
D: th-t-d-n-nh
E: th-t-ɗ-n-ʔn
F: th-t-n-nh
G: th-t-n-nt
H: th-t-ɗ-n
I: th-t-n
J: th-ɗ-n
K: th-t
L: t-ɗ-n-ʔn
M: t-ɗ-n

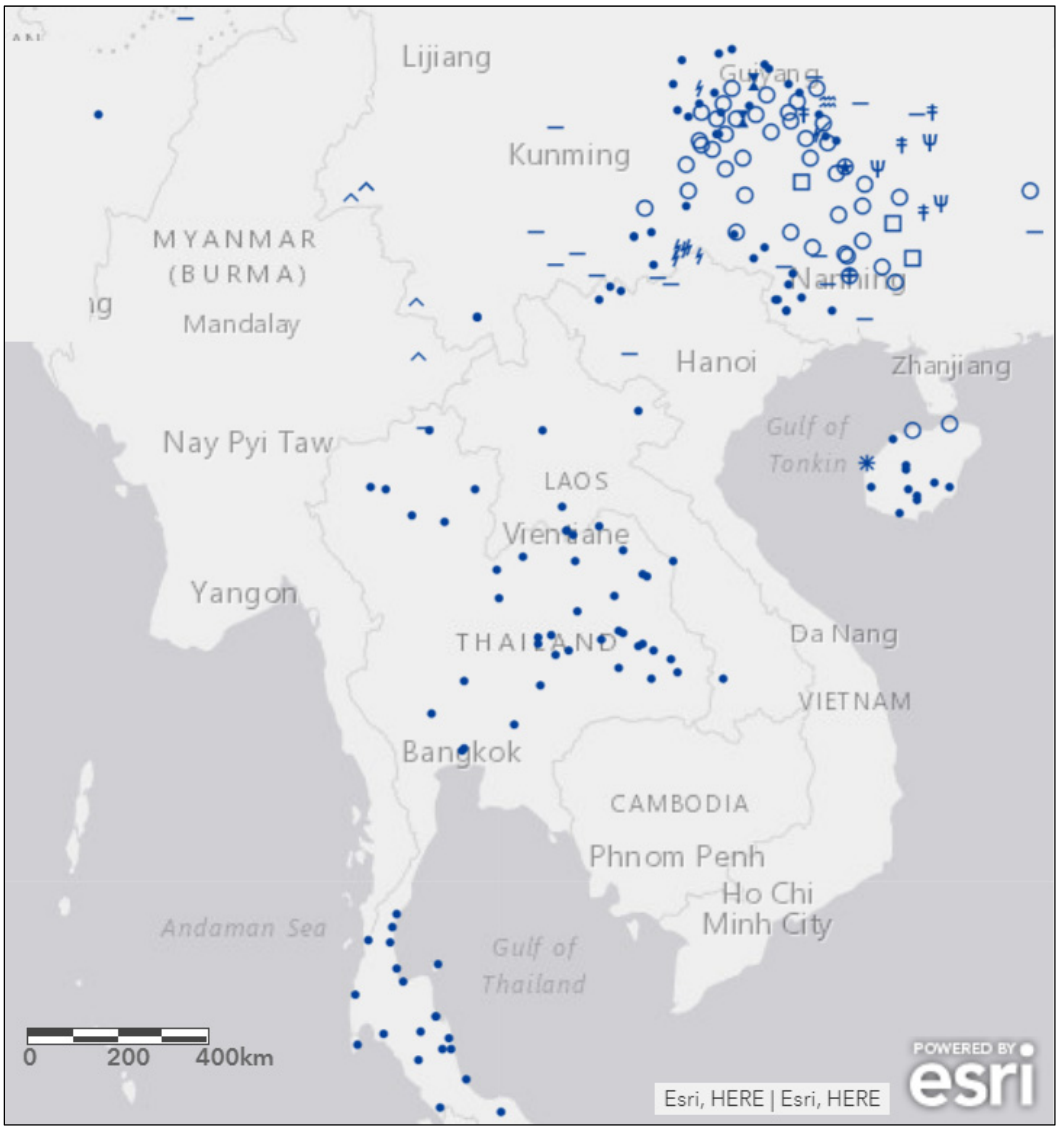


Figure 13.7.1: Stop series in Kra-Dai.

Stop series in Tibeto-Burman

1. Classification

A. /t-d/ type (/t-d-n/ only)
Trung; Puroik; Bangru; Galo (Tani)

B. /t^h-t-d/ type
B1a /t^h-t-d-^{n}d-^{n}t-$^n t^h$-n-n̥/ type
Tibetic (Zulong, mPhagri); nDrapa (Ngwirdei); Zakhring
B1b /t^h-t-d-^{n}d-^{n}t-$^n t^h$-n/ type
Tibetic (Babzo); rGyalrongic (Situ, bTsanlha, Khroskyabs, sTau, Nyagrong Minyag)
B2a /t^h-t-d-^{n}d-$^n t^h$-n-n̥/ type
Majority of dialects of Tibetic languages in Khams; nDrapa (Mätro); Choyu; Lhagang Choyu; Lamo; Larong sMar; Drag-yab sMar; Lizu
B2b /t^h-t-d-^{n}d-$^n t^h$-n/ type
Tibetic (sKyangtshang, Bragkhoglung, Phyugtsi, Daan); Namuyi (Dzolo); Ersu
B3a /t^h-t-d-^{n}d-n-n̥/ type
Tibetic (mDungnag); Betsi Choyu; Shuhing; Yi Northern (Senza); Songlin
B3b /t^h-t-d-^{n}d-n/ type
Tibetic (Chabcha, Mangra, Brag-g.yab, rTsamda, Limi, Tabo); Pema; Basum; Darmdo Minyag; Yi Eastern (Nersu, Nipu); Naxi; Malimasa; Dao; Selibu (Shuimofang)
B4 /t^h-t-d-$^n t^h$-n/ type
Alo; Yi Eastern
B5a /t^h-t-d-n-n̥(ʔn)/ type
Lidim; Laluba; Lalu; Northern Prinmi; Central Prinmi; Burmese (Yangon, Yaw); Daai Chin
B5b /t^h-t-d-n/ type
Tibetic (gTsangbawa, kLurtse, Ladaks, Balti, Khumbu, Chocha-ngacha); Rmaic (Mawo, Ronghong, Longxi, Taoping); nGochang (Qianxi); Yongning Na; Yi Western (Lalo, Lipo); Yi Southern (Narsu, Nesu); Yi Central; Axi; Lisu; Burmese (Palaw, Myeik); Kaman; Idu; Hayu; Dolakha Newar; Chantyal; Kinnauri
B6 /t^h-t-d/ type
Tujia (Tanxi, Xiaqieji, Xiadu)

C. /t^h-t-d-t‘(d’/$d^{\tilde{n}}$)/ type
Ca /t^h-t-d-$d^{\tilde{n}}$-n-n̥/ type
Kathmandu Newar; Camling
Cb /t^h-t-d-t‘-n/ type
Tibetic (Denjongke, Dzongkha, Brokpa); Wambule Rai

D. /t^h-t/ type
D1 /t^h-t-^{n}d-n/ type
Lahu; Kucong; Tibetic (Lhasa, Largyab, Shigatse, Ruthog)
D2a /t^h-t-n-n̥/ type
Ganan; Ao; Xiandao; Taungyo Burmese
D2b /t^h-t-n/ type (including /t^h-d-n/)
Azha; Bai; Zozo; Hani (Biyue, Shuigui); Jino; Zaiwa; Jinghpaw; Kadu; Selibu (Longwangbian); Tujia (Xianren); Phom; Manang
D3 /t^h-t/ type
Tujia (Pojiao, Tasha, Laxidong)

E. others (with an implosive sound)
Karenic (Bwe, Geba, Manu, Kyonpyaw Pwo); Asho Chin; Cak

2. Geographical distribution and interpretation

The reconstruction of proto-Tibeto-Burman (PTB; Matisoff 2003) includes a bipartite system of ‘voiceless’ *t and ‘voiced’ *d in plosives. This is common with Sinitic (Old Chinese; Baxter & Sagart 2014). Hence, this bipartite system (Type A) takes first position in the present classification as the

most archaic form, though we do not confirm that the system reflects the reality of archaic forms. Referring to the sound development attested principally in Tibetic languages, we list the types as follows: a tripartite system, voiceless aspirated, voiceless non-aspirated, and voiced (Type B); a quadripartite system (Type C); and another bipartite system, aspirated and non-aspirated (Type D). Second, the number of prenasalisations is classified (e.g., B1 & B2); finally, the nasals are considered (e.g., B1a & B1b). The following types are in chronological order.

/n̥/ (or voiceless nasals) is mainly derived from an *s prefix, of which the evidence, in most cases, remains in Written Tibetan forms as well as rGyalrongic languages. The latter (B1b) still maintains a consonant cluster /sn/ instead of /n̥/.

Type A is marginally found. In our data, several languages distributed between Bhutan and Northwestern Yunnan are classified into this type.

Type B exhibits the widest distribution, which nearly covers the whole TB area. It is first subclassified based on the prenasalised pattern: Tripartite (Type B1), bipartite (Type B2), voiced only (Type B3), voiceless only (Type B4), and no prenasalisation (Type B5). Note that we find a restriction of the appearance of prenasalised forms. There are reports of several languages, such as Ladaks and Balti, in which prenasalisation only appears in word-medial position. However, we do not reflect this case in the classification or the maps.

Some discrepancies due to different conventions of transcription are unified into a representative one for simplicity. For example, 'th-t-d-nt-nth-n' in Ringmo Tibetan is unified into Type B2b. The transcription 'ʔn' in Laluba is considered as bearing a close status to /n̥/, although we need confirmation.

Type B1 is mainly distributed in the Ethnic Corridor of West Sichuan. Note that the Tibetic languages with Type B1 are derived from those with Type B2 due to individual innovations. We also find it in Zakhring, which has had strong language contacts with Khams Tibetan (B2a) and Kaman (B5b), spoken in Dzayul (Tibet).

Type B2 is found in the eastern Tibetosphere. There is a discrepancy between previous works and our description regarding the existence of the prenasalised voiceless aspirated /ⁿtʰ/ in Khams Tibetan. We follow a description that recognises this sound as a part of the system.

Type B3 is marginally found: Tibetic languages spoken in the northeasternmost and southeasternmost areas, the Yi Northern and Eastern groups, as well as Sinitic-based 'mixed languages' such as Dao and Selibu.

Type B4 is found in Yi Eastern in the small easternmost area. It seems that this type is rare.

Type B5 is mainly distributed in the south of the Tibeto-Burman linguistic sphere: from Guizhou to Yunnan, Arunachal Pradesh, Bhutan, and Ladakh.

Types B6 and D3 lack the nasal sound; however, this is due to a merger of [n] with /l/. This phenomenon is widely attested in Southwestern Mandarin (Cao 2008).

Type C is found on the southern side of the Himalayas, namely, Nepal, Sikkim, and Bhutan. It can have two patterns of origin:

adaptation of the Indic sound system and a transitional status between Types B and D. These two are not classified in the maps. For the latter origin, we find several ways of representing the *fourth* feature, including /t‘/, /d’/, and /d$^{\text{fi}}$/, which all represent a breathy sound. Some studies have described it as ‘murmur voice’; at the present stage, we consider both ‘breathy’ and ‘murmur’ in a single unit. There is an analysis of these types of phonation as part of suprasegmental realisations and thus not in the consonant system. In this case, the breathy feature would also appear resonant.

Considering its historical position, we find that this phoneme is derived from a voiced simplex *d in Dzongkha and Denjongke and is to be merged with /t$^{\text{h}}$/ as attested in Lhasa Tibetan, whereas in Brokpa, it seems that /d$^{\text{fi}}$/ is derived from complex initials of which the main initial is *d; a similar phenomenon is attested in dPalskyid Tibetan (B2b).

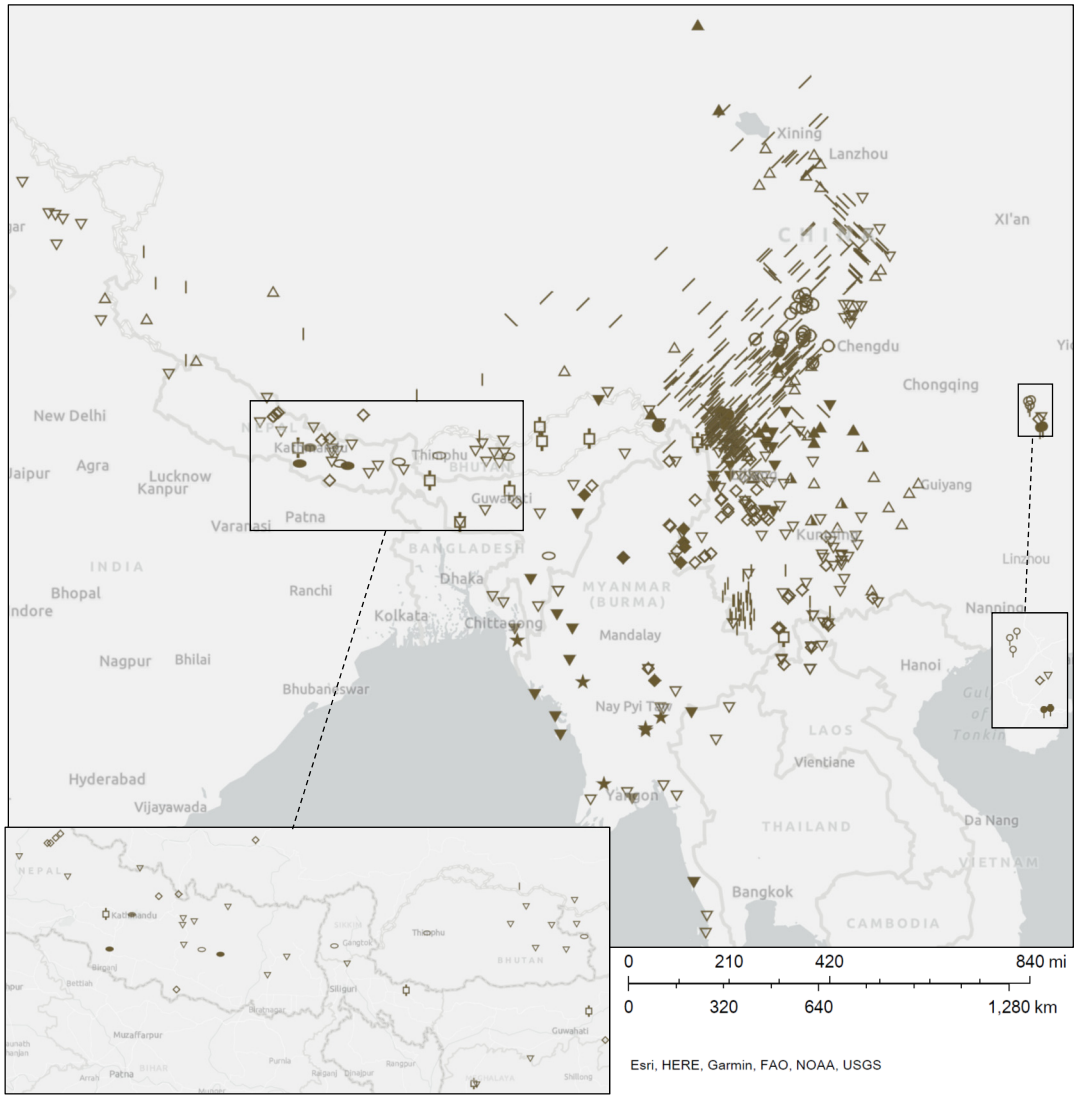
Type D is found in the Tibet Plateau, scattered, as well as in the border area of China (Yunnan) and Laos. Type D1 seems similar to a subtype of Type B. However, a prenasalised sound /$^{\text{n}}$d/ is not regarded as a substitute for a simple /d/, regardless of its phonological status. Lahu’s phonological description is /t$^{\text{h}}$-t-d-n/, but its phonetic realisation is [t$^{\text{h}}$-t-$^{\text{n}}$d-n]. We follow the latter for the present analysis. Moreover, observing the tendency of sound change in Lhasa Tibetan, we can see Type D1 going to merge into Type D2.

Type E is a group possessing an implosive /ɗ/. The languages of Type E are spoken in Myanmar and Bangladesh. Kato (2009:180) claims that it is already an implosive at the proto-Karenic stage. At least this phoneme does not date back to PTB. Hence, we set Type E independently. A potential source of its acquisition is language contact with Kra-Dai, Mon-Khmer, or Austroasian languages.

Even languages distributed in a small area display different types. For example, Selibu has two points: Shuimofang belongs to Type B3b, and Longwangbian to D2. The difference is due to the degree to which words of Tibetan origin are incorporated into the system. Tujia is also in the same situation. Tanxi belongs to Type B5, and Pojiao to D3. The former has a more complex system than the latter that reflects the sound change process.

We collected the data of around 710 points. Although PTB (and Old Chinese) includes a voiceless and voiced series in the plosives, almost all languages and varieties of Tibeto-Burman have a distinction of aspirated voiceless and non-aspirated voiceless series. Languages in Nepal often have a quadripartite system of aspirated voiceless, non-aspirated voiceless, voiced, and breathy voiced, and several languages in Myanmar have acquired an implosive which does not exist in PTB. These phenomena are suggestive of intense language contact influencing the sound system.

(SUZUKI Hiroyuki, EBIHARA Shiho, IWASA Kazue, KURABE Keita, and SHIRAI Satoko)



A: /t-d/ (/t-d-n/ only)
B1a: /th-t-d-nd-nt-nth-n-n̥/
B1b: /th-t-d-nd-nt-nth-n/
B2a: /th-t-d-nd-nth-n-n̥/
B2b: /th-t-d-nd-nth-n/
B3a: /th-t-d-nd-n-n̥/
B3b: /th-t-d-nd-n/
B4: /th-t-d-nth-n/
B5a: /th-t-d-n-n̥(?n)/
B5b: /th-t-d-n/
B6: /th-t-d/
Ca: /th-t-d-dh-n-n̥/
Cb: /th-t-d-t‘-n/
D1: /th-t-nd-n/
D2a: /th-t-n-n̥/
D2b: /th-t-n/
D3: /th-t/
E: others

Figure 13.8.1: Stop series in Tibeto-Burman.

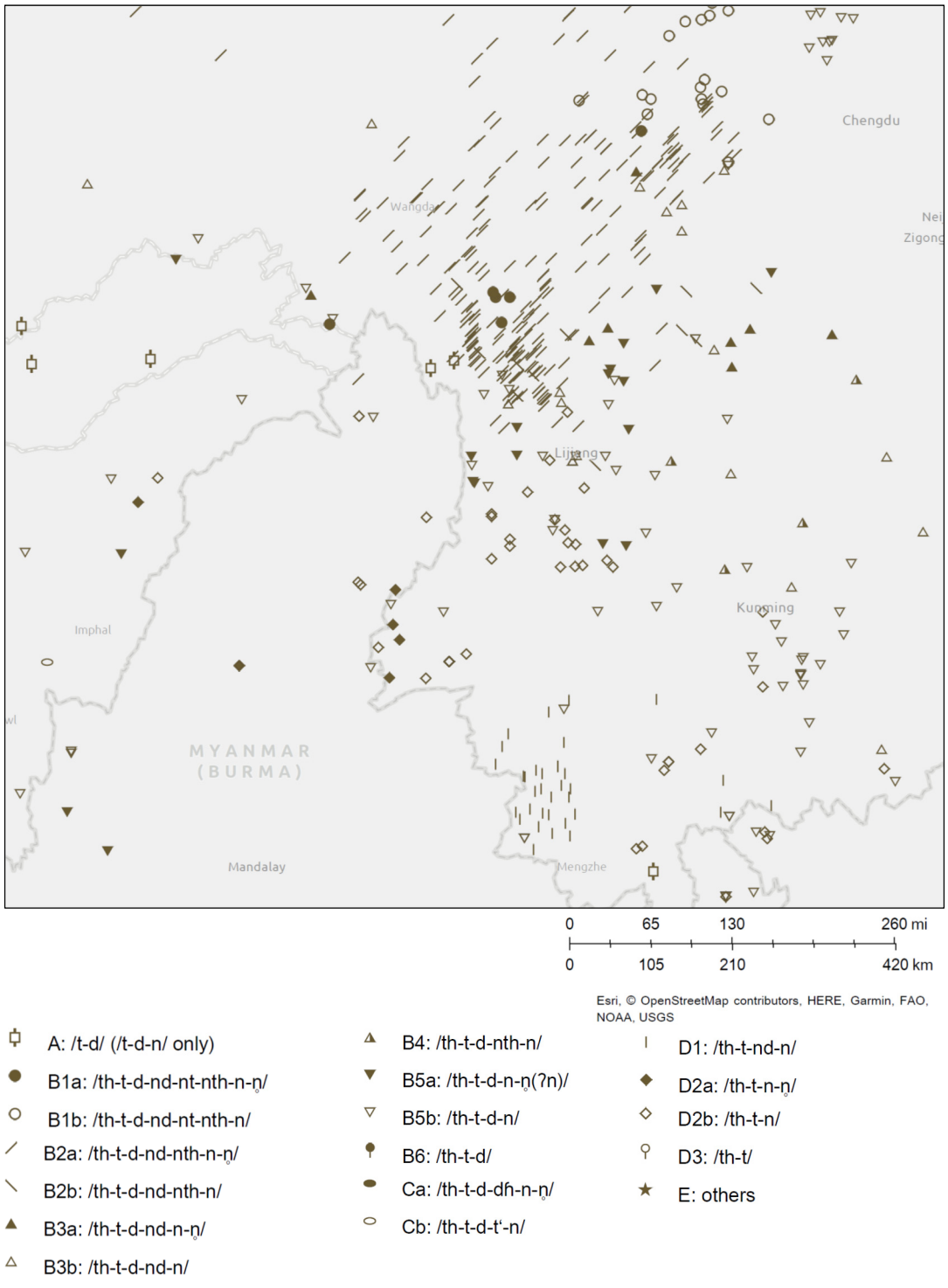


Figure 13.8.2: Stop series in Tibeto-Burman: An enlarged version.

Stop series in Austroasiatic

1. Classification

Figures 13.9.1 and 13.9.2 show types of (denti-)alveolar plosive consonant series in Austroasiatic (AA) languages. In languages with sesquisyllabic structures (C_1-C_2VC_3), the initial consonants of the major syllable (C_2) are taken into account. The types are classified into five large categories as follows ('/' stands for 'or' and '|' for 'and/or').

A. th-t-d/ɗ-n type
  A-1 th-t-d-n type
    th-t-d-n
    th-t-d-dh-n-(nh)
    th-t-d-n-nh-(ʔn)
    th-t-d-nd|nt|nth-n-(nh)
  A-2 (th)-t-ɗ-n type
    (th)-t-ɗ-n
    th-t-dh-ɗ-n
    th-t-ɗ-nd|nt|nth-n
    th-t-ɗ-(nd|nt|nth)-n-nh
    th-t-ɗ-nd|nt|nth-n-nh-ʔn
B. th-t-d-ɗ-n type
    th-t-d-ɗ-n-(nh)
C. th-t-n type
    th-t-n
    th-t-n-ʔn/ʔd
    th-t-nd|nt|nth-n
    th-t-(nd|nt|nth)-n-nh
    th-t-nd|nt|nth-n-nh-ʔn
D. t-d-n type
    t-d-n
E. t-n type
    t-n

2. Geographical distribution and interpretation

Proto AA contains the dental series *t-d-ɗ-n (Sidwell 2015), and most of the languages cited here contain /th/. Hence, the classification proposed above is based on the series: th-t-d-ɗ-n. Since the implosive sounds are widespread in Southeast Asia (Maddieson 2013) and phonologically distinguished from the normal plosives in proto AA, we think it important to distinguish the voiced plosive /d/ from the implosive /ɗ/, even when the distinction makes no sense phonologically. Here, we follow the description of each author, even though the identical target is sometimes described differently, such as in the case of [ɗ] as /d/ or [ʔd].

Type A-1 is quite common in the Katuic, Khasic, Khmuic, Mangic, Monic and some Vietic languages.

Type A-2 is common in the Bahnaric, Khmeric, Monic, Palaungic, Pearic, Vietic and a few Waic languages.

Type B is conservative in that it contains the complete proto AA series. The Koʰho (Southern Bahnaric), Mlabri (Khmuic) and Danaw (Palaungic) languages use this system.

Type C lacks the voiced and voiceless contrast, which is a typical case of the emergence of registral or tonal contrasts. Most languages of this type (Suai, Khmu, Lamet, Lai, Mường Danh, Kontoi Plang and Samtao) have registral or tonal contrasts, except for Lawa (Waic). However, Lawa has a full nasal series: th-t-nd-n-nh-ʔn.

Type D is typical in the Aslian and Munda languages.

Type E is only found in the Car Nicobarese language.

In addition to the stop series considered above, AA languages also possess prenasalized stops (nd-nt-nth), a voiceless nasal (nh) and a preglottalized nasal one (ʔn). Their geographical distribution is shown in Figure 13.9.1.

Prenasalised series are found in the Bahnaric, Khmuic, Palaungic, Bugan (Mangic), Nyah Kur (Monic), Lai (Vietic) and Lawa (Waic) languages.

Voiceless nasal stop is quite common in the Bahnaric, Katuic, Khmeric, Khmuic, Monic, Palaungic and Waic languages.

Preglottalized nasal stop is found in the Sedang (Northern Bahnaric), Eastern and Western Khmu (Khmuic) and Lawa (Waic) languages.

(SHIMIZU Masaaki, MINEGISHI Makoto)

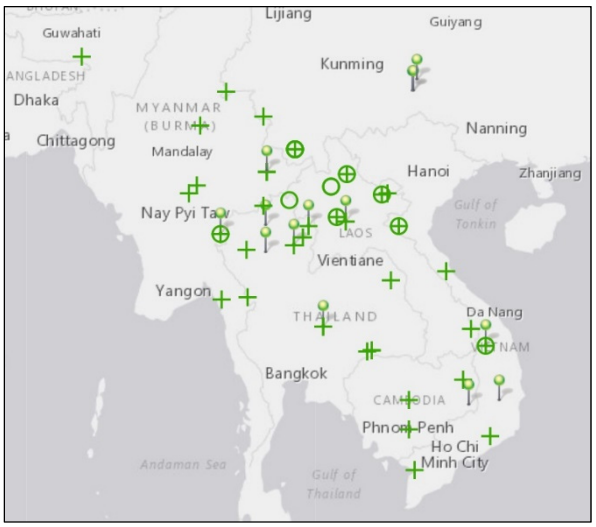


Figure 13.9.1: nd/nt/nth: nh:+ ʔn:○.

A-1.	th-t-d-n type	○
A-2.	(th)-t-ɗ-n type	×
B.	th-t-d-ɗ-n type	⊗
C.	th-t-n type	●
D.	t-d-n type	+
E.	t-n type	◇

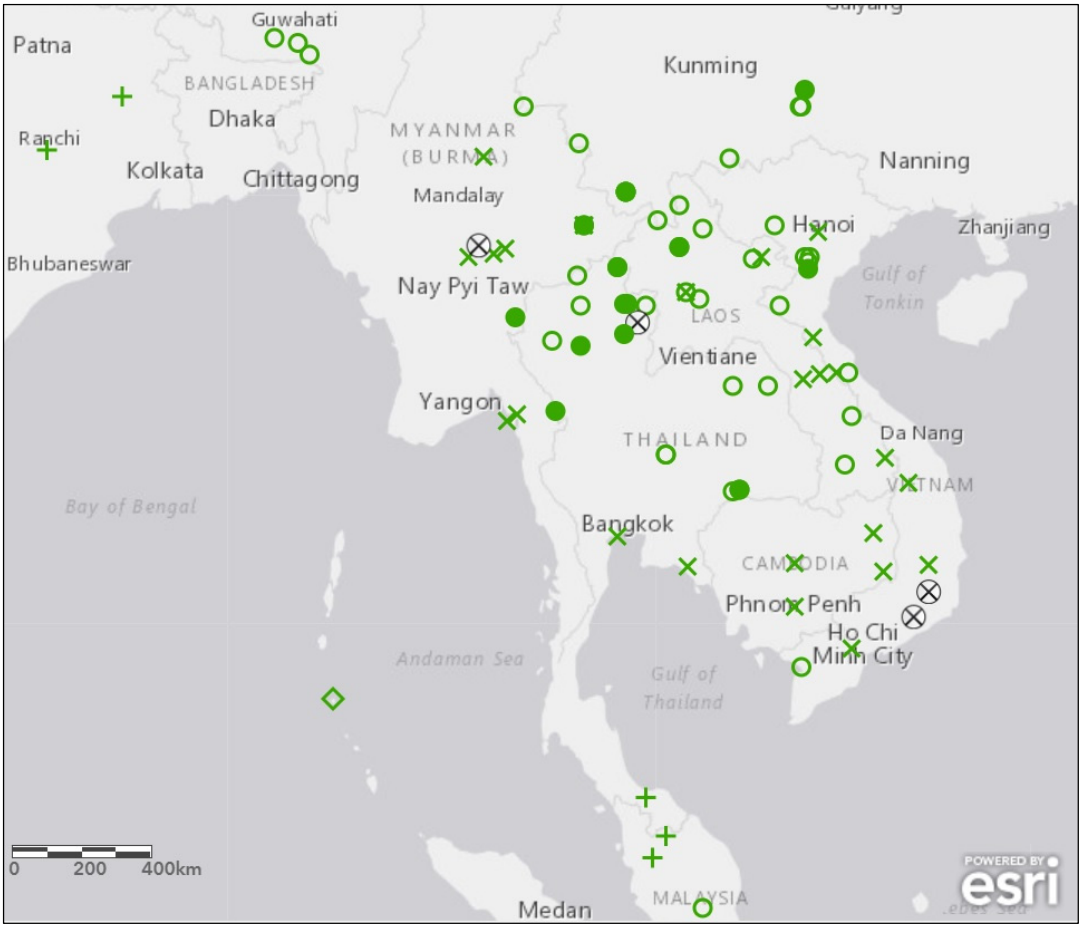


Figure 13.9.2: Stop series in Austroasiatic.

Stop series in Austronesian

1. Classification

Consonants in Austronesian languages do not differ very much if we look at dental stops and nasals.

A: t-n Languages with a voiceless stop and a nasal.

B: Languages with a voiceless and a voiced stop and a nasal.

B1: t-d-n

B2: t-ɖ-n, t-d-ɖ-n

With a retroflex voiced stop

C: t-d-dh-n Languages with a voiceless and a voiced stop, a nasal and other voiced consonant; an aspirated voiced stop or a voiced retroflex stop

D: Languages with (a) prenasalised stop(s)

D1: t-d-nd-n/t-d-ndr-n/t-d-nt-n

Languages with a voiceless and a voiced stop, a prenasalised stop, and a nasal

D2: t-d-nt-nd-n/t-d-ɗ-n-nt-nd-n

Languages with two prenasalised stops in addition to stops and a nasal

E: Others: languages with aspirated consonants.

t-th-d-n-nh/t-th-nt-d-n-nh/t-th-d-n-nh

2. Distribution

Austronesian languages most frequently exhibit Type B-1, in which a voiceless and a voiced stop and a nasal dental consonants (t-d-n) are found, and one language has a retroflex voiced stop /ɖ/ instead of /d/. A few languages lack a voiced stop (Type A). In Taiwan, the Philippines, and Sumatra, most languages fall into either Types A or B. Type C is only found in Madurese. There are no languages that exhibit a prenasalised stop.

Types D-1 and D-2, which have more than one prenasalised stop in a dental series are found in Sulawesi, Papua, and some Oceanic languages. More complicated inventories, such as t-th-d-n-nh, t-th-nt-d-n-nh, and t-th-d-n-nh, are found in some Oceanic languages. However, Type B remains the most frequent pattern in these areas, too.

Consonants in Proto-Austronesian (PAN) do not add up to a large number. The four vowels, /i, a, u, ə/ and four dephthongs /iw, ay, aw, uy/, are almost unequivocally posited for proto-Austronesian (Dyen 1953, Dahl 1981, Mills 1981, Blust 2009, Wolff 2010, among others). As for consonants, including semi-vowels, researchers may disagree. Blust (2009) posits 25 consonants, *p, *b, *m, *t, *d, *n *S, *C(ʦ), *l, *r, *R(/r/ or /ʁ/), *ɲ, *s(ç), *c(ʧ), *z(ʝ), *N(lʲ), *D(/ɖ/), *k, *g, *j(gʲ), *ŋ, *q, *h, *y(/j/), *w (Symbols in the brackets are suspected actual phones). Wolff (2010) reconstructs the following 19: *p, *b, *m, *t, *d, *s, *n, *ɬ, *l , *c, *j, *k , *g, *ŋ, *ɣ, *q, *h, *w, *y, and Ross (1995) posits the following 23 consonants: *p, *b, *m, *t, *d1(/d/), *d2(/ʣ/), *d3(/ɖ/), *C(ʦ), *n, *s, *S(/ʂ, ç/), *Z(/ʒ/), *L(l, ɬ/), *l(/ɭ, ɻ/), *r, *k, *g, *ŋ, *q, *ʁ, *h, *w, *y.

It is hard to determine the sets of consonants, but there are some consistent points. Place of articulation are bilabial, alveolar, alveolar retroflex, palatal, velar, uvular, and glottal. All the consonants are pulmonic. Manners of articulation are stop, nasal, fricative, lateral, and trill.

Overall, PAN is supposed to have a fairly simple phonemic system, so are the stop series.

(UTSUMI Atsuko)

- A type: t-n
- B-1 type: t-d-n
- B-2 type: t-ḍ-n, t-d-ḍ-n
- C type: t-d-dh-n
- D-1 type: t-d-nd-n/t-d-ndr-n/t-d-nt-n
- D-2 type: t-d-nt-nd-n/t-d-ḍ-n-nt-nd-n
- E type (others) t-th-d-n-nh/t-th-nt-d-n-nh/t-th-d-n-nh

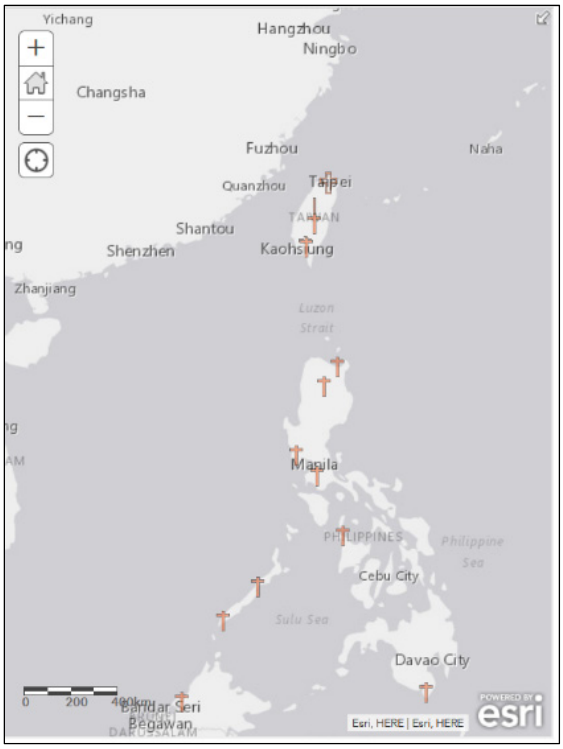


Figure 13.10.1: Stop series in Austronesian, Taiwan and Philippines.

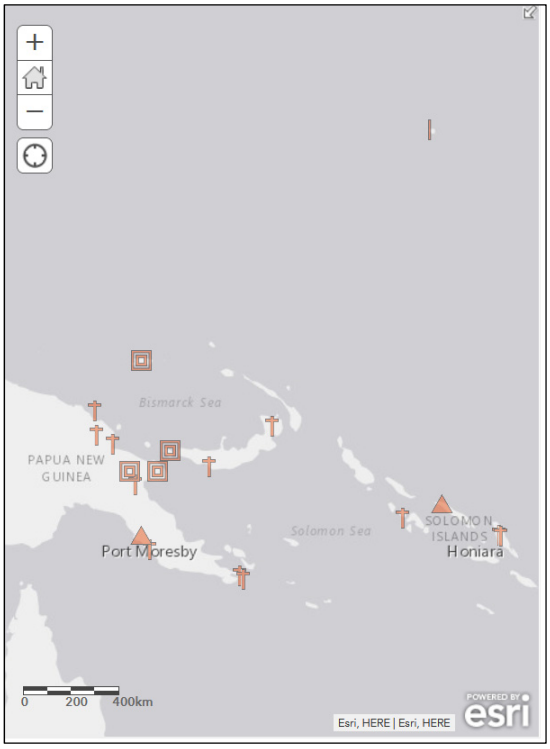


Figure 13.10.2: Stop series in Austronesian, Papua New Guinea and nearby islands.

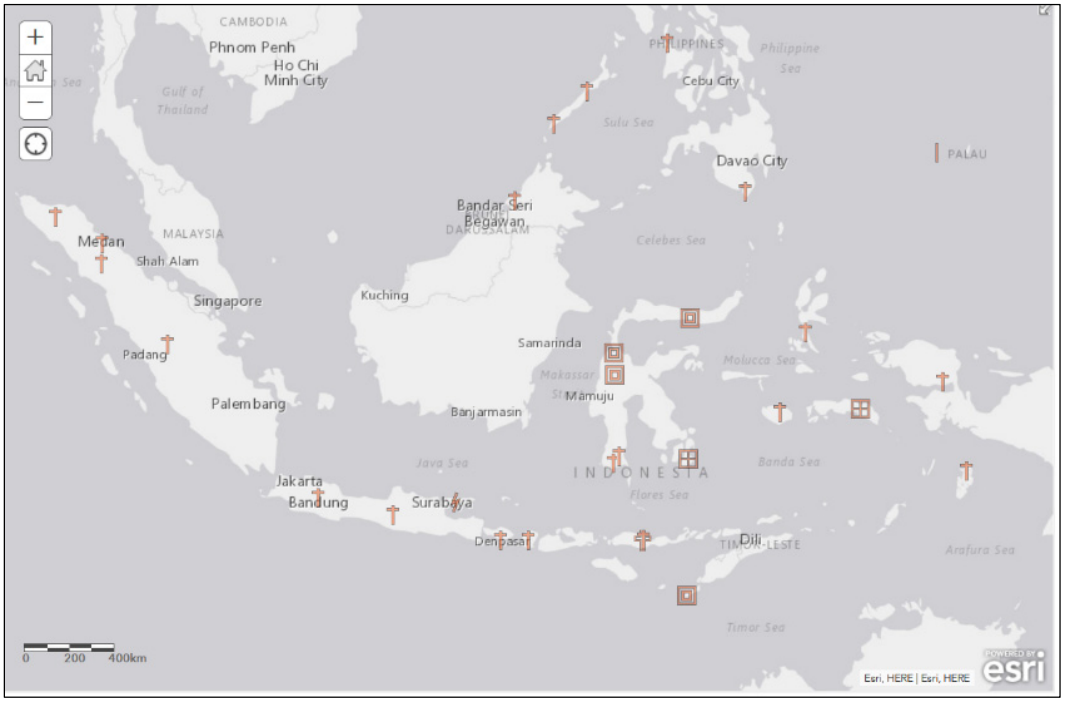


Figure 13.10.3: Stop series in Austronesian, Indonesia.

Stop series in Tungusic

1. Classification

Tungusic languages have almost less differences in phonetic inventories. For example, Evenki's inventory is as following: stops /p/ [p], /b/ [b], /t/ [t], /d/ [d], /k/ [k], /g/ [g], affricates /č/ [tɕ], /ǰ/ [dʑ], fricatives /s/ [s], /h/ [h], nasals /m/ [m], /n/ [n], /ŋ/ [ŋ] and others /l/ [l], /r/ [r], /v/ [w], /j/ [j].

In Tungusic languages only one type A is observed:

A t-d-n

2. Geographical distribution and interpretation

It is possible to say that all Tungusic languages have the distinctive features of [+/- voice] and [+/- nasal].

This type is also observed in other obstruents as [k]/[g]/[ŋ], but it may not be applicable to the labial plosive, as in Evenki words which begin with /p/ are relatively less than words with /b/ in the initial.

(MATSUMOTO Ryo)

● A. t-d-n type

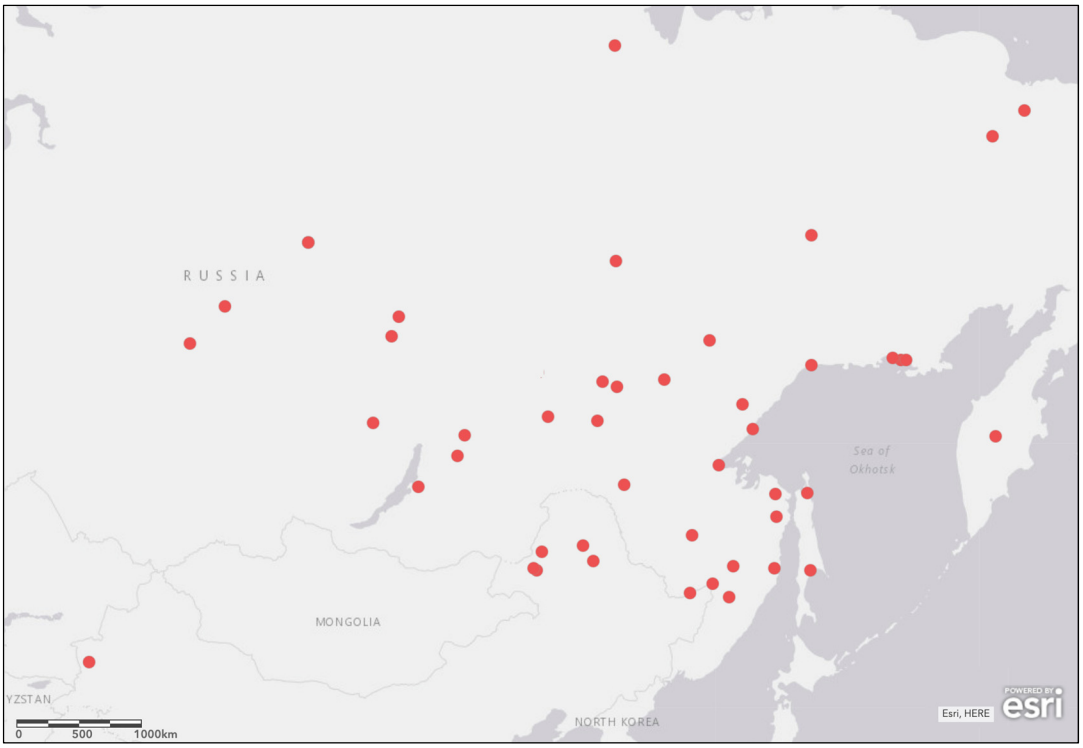


Figure 13.11.1: Stop series in Tungusic.

Stop series in Uralic

1. Classification

Uralic languages are classified into 3 groups, A1, A2 and B, as shown in Figure 13.12.1.

Type A1 has phonetically 3 series of alveolar plosive; voiceless, voiced and nasal, but in Type A2 and B voiceless and nasal are distinctive in the initial phoneme of a word. In Type A2 voiced plosive can appear only in the middle of a word mainly as the result of the morpho-phonemic alternation.

The phonetic form of the phoneme /d/ is different according to the language group. In Type A2, which includes Balto-Finn languages, it is pronounced by weak-voiced or half voiced [d̬].

2. Geographical distribution and interpretation

Type A2 is widely observed in Uralic languages. Especially most of the Finnic languages are belonging to this type with the consonant gradation (CG). For example, the Finnish CG of the alveolar series is shown in Table 1. It depends mainly on the morphological and phonetic conditions, which grade should be used.

Table 1. CG in Finnish

Strong Grade (SG)	Weak Grade (WG)
tt	t
t	d

For example:

(1) mai**t**o mai**d**o-ssa
milk.NOM *milk*.INESS

Sulkala and Karjalainen (1992: 366) also mention:

> */d/ is substituted by other phonemes in the dialects, and occurs only in word-medial position in native Finnish words, acting as the weak variant of /t/ in consonant gradation.*

In Permic languages, 3 series of stops as Type A1 are distinctive by [±voice] and [±nasal], as in Altaic languages observed. They reside next to Tatar and Bashkir, it is possible to expect that it is influenced from neighboring Turkic languages. On the other hand, Tatar and Bashkir are having contact also with Volga-Finnic languages (Mari and Mordvin) in the west, which are type A2 without the phoneme /d/.

In Ugric languages, except Hungarian which was moved to far west from the homologous Khanty and Mansi, it seems that they do not have the feature of [±voice]. Selkup in Samojed has the same stop series as Ugric, it may be because of the areal feature of languages of peoples along Yenisei called “Ostyak”.

(MATSUMOTO Ryo)

A1 t-d-n
A2 t-(d)-n
B t-n

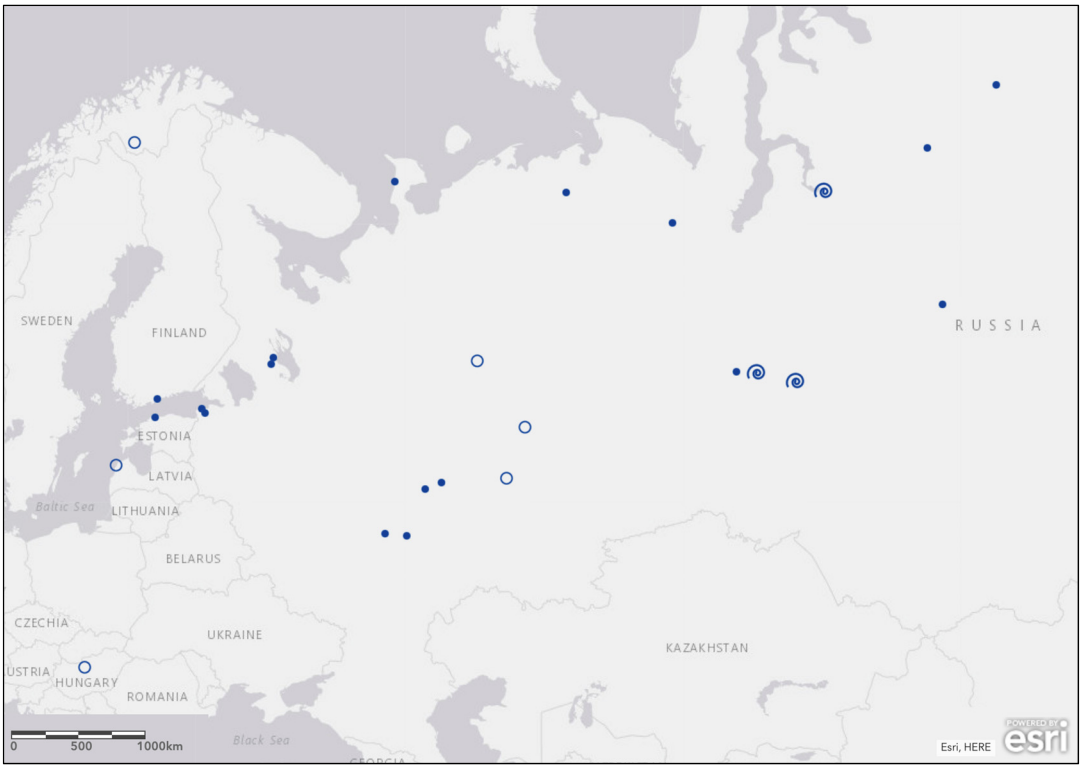


Figure 13.12.1: Stop series in Uralic.

Stop series in Mongolic and Turkic

1. Classification

At the phonological level, there are two types of initial dental stop series in Mongolic and Turkic:

t-n Chuvash (a Turkic language)

t-d-n Other languages

The *t-d-n* type includes two sub-categories on the phonetic level characterized by voice and aspiration.

All Mongolic languages belong to the */t/-/d/-/n/* type. Kalmyk in the lower Volga region, Buryad in southern Siberia and Moghol in Afghanistan have a voicing contrast between /t/ and /d/, while the other Mongolic languages including Oirad, which is closely related to Kalmyk, show a contrast in terms of aspiration between them:

A [t]-[d]-[n] (/t/ can be realized as a slightly aspirated [t^h])
Buryad, Kalmyk, Moghol

B [t^h]-[t]-[n]
Dagur (Butha), Dagur (Tacheng), Khamnigan, Bargu Buryad, Mongol (Chakhar, Khalkha, etc.), Oirad, Shira Yughur, Monguor, Baoan, Dongxiang, Kanjia

All Turkic languages except Chuvash belong to the */t/-/d/-/n/* type.

A [t]-[d]-[n] (/t/ can be realized as a slightly aspirated [t^h])
Turkish, Azeri, Gagauz, Turkmen, Tatar, Bashkir, Crimean Tatar, Kyrgyz, Kazakh (Kazakhstan), Noghay, Uzbek, Uighur, Sakha, Dolgan, Khakas, Shor, Chulym

B [t^h]-[t]-[n]
Kazakh (China), Sarïg Yughur, Salar, Tuvan (, Uighur)

Chuvash belongs to the *t-n* type. Voiced stops, however, appear in Russian loanwords. The realization of /t/ may vary to some extent depending on the environment.

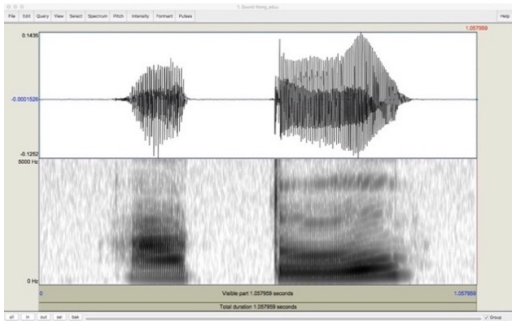
(For the languages and dialects for which clear phonetic descriptions are not available, the author made use of recordings of native speakers provided by institutions and individuals including those uploaded on the Internet as well as linguistic and learning materials on the market. If voicing was observed in /d/ in a sentence-initial position, the language or dialect was classified as a member of the type A group. As the amount of data the author could obtain was small and variations in terms of area and generation within a language or a dialect may exist, this is just a tentative classification.)

2. Geographical distribution and interpretation

Oral stops are mainly distinguished by voice in the western and northeastern languages as in Russian and by aspiration in southeastern ones as in Chinese.

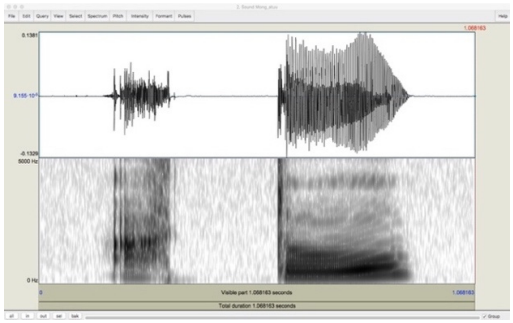
Development of preaspiration in some languages is reported (Karlsson and Svantesson 2012). The figures below show phonetic features of the Khalkha Mongolian intervocalic /t/ and /d/. The phonemes are both realized as a voiceless stop. The spectrograms show that /a/ is breathy voiced with partial devoicing before /t/ and modal voiced before /d/. (Creakiness observed at the beginning of /a/ in /atū/ in this utterance is just an

accompaniment of a low pitch and should be ignored in this discussion. Noise caused by breath before /a/ in /adū/ has nothing to do with the discussion, either.) We can also see the difference between the vowels from the waveforms. This situation can be interpreted as the contrast between preaspirated and non-preaspirated consonants, and the words can be transcribed as [a^{h}to:] and [ato:] respectively.



/adū/ ‘horse’

Figure 13.13.1: Preaspirated and plain voiceless stops in Khalkha Mongol (The waveforms and spectrograms were obtained using the Praat program developed by Paul Boersma and David Weenink.)



/atū/ ‘female fish’

(SAITÔ Yoshio)

Mongolic	t-d-n	t^{h}-t-n	
Turkic	t-d-n	t^{h}-t-n	t-n

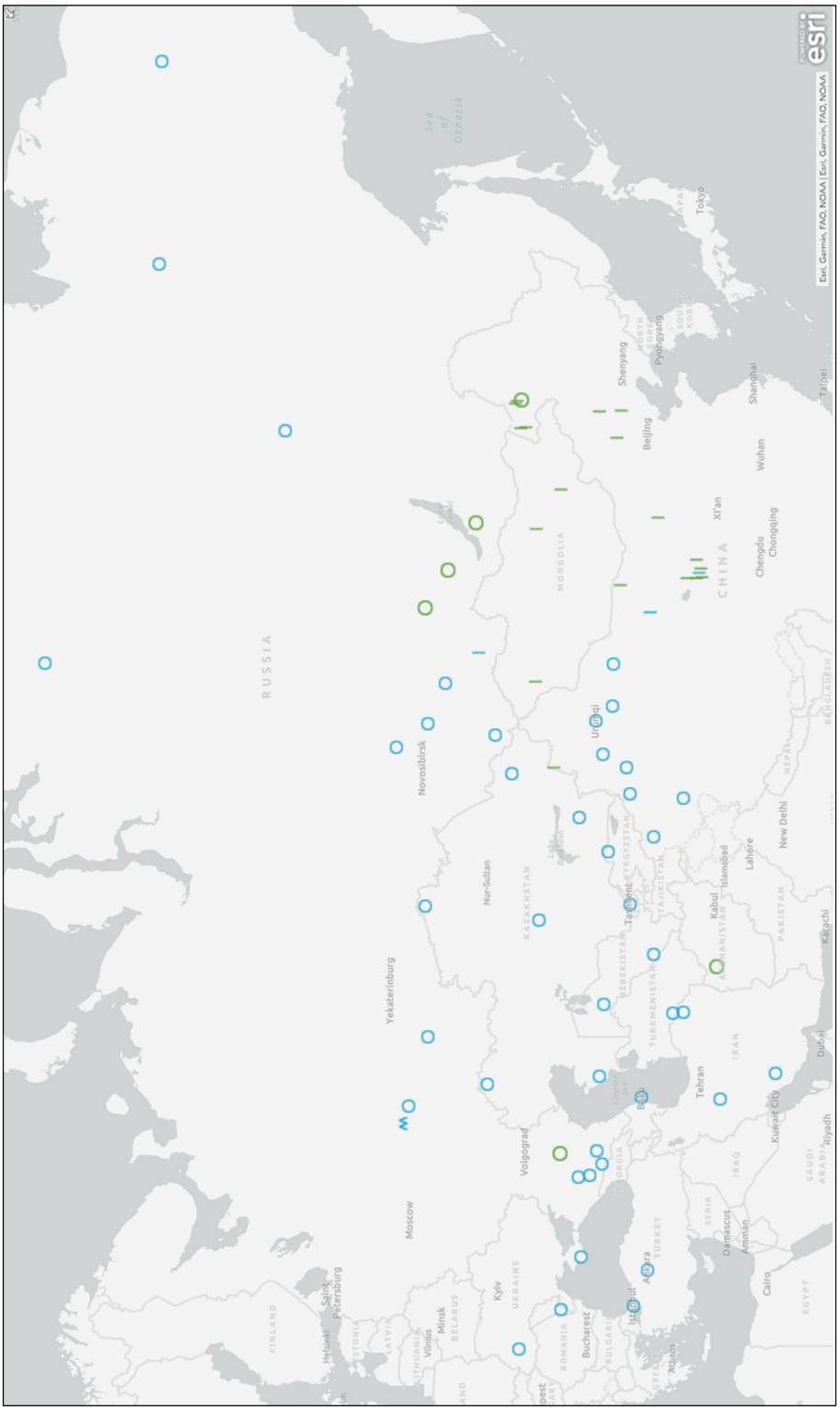


Figure 13.13.2: Stop series in Mongolic and Turkic.

Stop series in South Asia

1. Classification

Here, I discribe the languages of Indo-Aryan (hereinafter IA), several small language families/branches, and language isolates in South Asia. On Figure 13.14.1, the manners of articulation of the alveolar stop series are classified into ten types.

2. Geographical distribution and interpretation

As far as the entire Indian subcontinent is concerned, we can see that geographical rather than genealogical relationships have a stronger influence on the manner of articulation of consonants (Figure 13.14.1).

Historically, Sanskrit, an archaic language of IA, had five distinct alveolar stops <th-t-d-dh-n> (Cardona 2003); thus, it belongs to the type A classification of this paper. Even now, 21 (19 are IA) out of the 76 languages have the same five stops, that is, voiceless aspirated, voiceless nonaspirated, voiced nonaspirated, voiced aspirated plosives, and voiced nasals, just like Sanskrit. These languages are distributed over India (except the south), Bangladesh, and Nepal, and some northwest IA languages in northern Pakistan belong to this type.

Type B has lost the consonants of the voiced aspirated series and kept the four-way distinction. Languages belonging to this type are located in peripheral zones such as the Andaman Islands and an area from Indian-administrated Kashmir via northern Pakistan to northeastern Afghanistan. Genealogically, it includes the languages of the Andamanese family, Burushaski (isolate), and the northwest group of IA. Besides the inland languages other than Andamanese, Panjabi (both western and eastern dialects), which is located slightly to the south, also exhibits this four-way distinction. This language has lost the aspirated voiced plosives and exhibits distinctive tones instead.

Further along in type B, the distinction of aspiration has been lost even in voiceless stops, and the distinction has become three-way <t-d-n> in type C1 languages. In South Asia, this type consists of Onge (Andamanese) on Little Andaman Island, Pashayi (IA) and most Nuristanis in Afghanistan, and Chittagonian (IA) in Bangladesh. How has this language completely lost its aspiration distinction while being surrounded by type A languages is not clear. That Učida (1970) says that the fact that the language has a tonal system may be relevant to the loss.

In the Indian Ocean, there exist type C2 languages. This type is a subtype of C1 and has three languages: Sinhala, Dhivehi (both IA), and Vedda (a creole between pre-Vedda, a language isolate, and Sinhala). These languages have a three-way distinction, <t-d-n>, in the word-initial position, same as the C1 languages. However, the existence of a series of prenasalised stops <nd> in C2 languages is noteworthy. Prenasalised stops in such languages occur only in the onset of a word-internal syllable. They behave as single consonants and contrast with nasal + stop clusters (e.g. *ka.^{n}də* ‘tree trunk’ vs. *kan.də* ‘hill’ [Gair 2003: 779] in Sinhala).

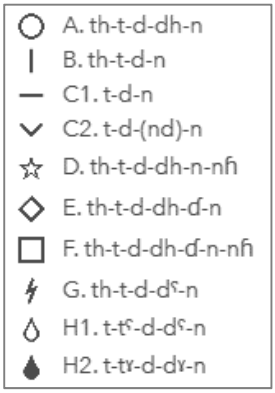
Marathi, Konkani, Vaagri Boli, and Saurashtra in central-to-southern India, Bhojpuri and Awadhi in northern India, and Torwali in northern Pakistan are all IA and type D languages, which have a series of aspirated nasals <nɦ> even word-initially, in addition to the series of type A languages.

Type E contains only Sindhi (IA), and type F has two languages, Saraiki and Marwari (both IA). These three languages are located across or near the borders of Pakistan and India. They gained a series of implosives, anew, and type F languages have voiced aspirated nasals, similar to type D languages. Type E has a six-way distinction, <th-t-d-dh-ɗ-n>, while type F has a seven-way, <th-t-d-dh-ɗ-n-nɦ>. Note that Sindhi and Saraiki have lost the distinction between dental and retroflex implosives; they are actually pronounced as the merged implosive as [ɖ] rather than [ɗ] in Sindhi, and vice versa in Saraiki, for example *ɖiṯho* 's/he saw' (Khubchandani 2003: 647) in Sindhi vs. *ɗiṯhimis* 'I saw it' (Bashir & Conners 2019: 220) in Saraiki.

Only Bishnupriya, which has the set <th-t-d-dʔ-n>, is classified as Type G in this study. This language has lost its voiced aspirated series, instead showing a series of 'voiced plosive with glottal closures' (Sinha 1981).

Outside South Asia, there are several IA languages in the west, see Map 2. All Romani dialects and Lomavren in and around Europe belong to type B, and the dialects of Domari in the Middle East are of type H1 <t-tˤ-d-dˤ-n> (Jerusalem) and H2 <t-tᵛ-d-dᵛ-n> (Aleppo). These pharyngealised or velarised dentals are the result of contact with Arabic, and are found mainly in words borrowed from Arabic, but also in some non-Arabic words, i.e., Indo-Aryan words. For example, Jerusalem *dˤandˤ* 'tooth' (Matras 2012: 43) and Aleppo *pēṭ* [peːtᵛ]'belly' (Herin 2012: 7) correspond to *dā̃t* and *pēṭ* in Hindi-Urdu respectively.

(YOSHIOKA Noboru)



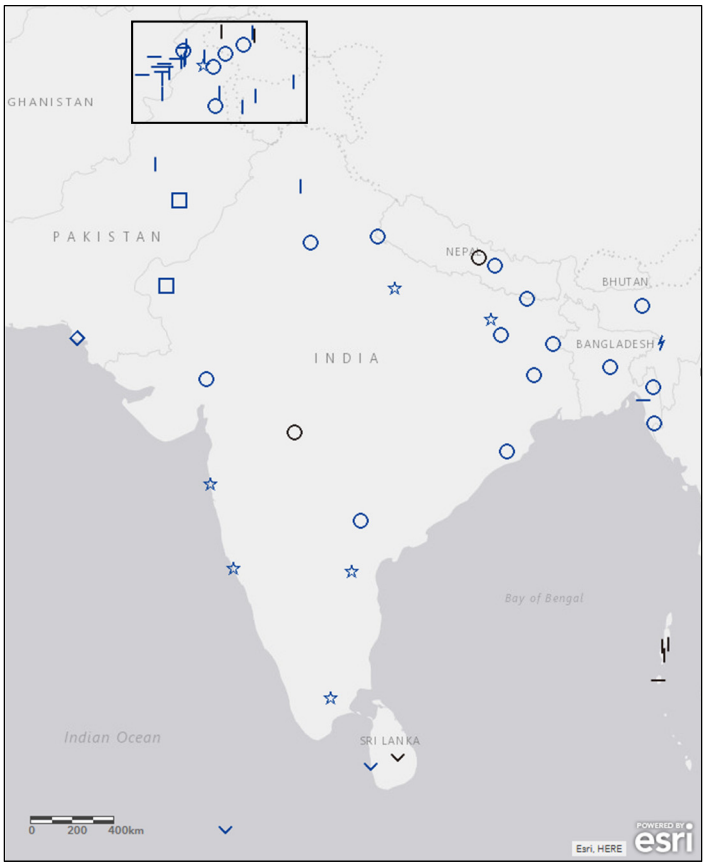


Figure 13.14.1: Types of stop series in Indo-Aryan, Nuristani (both in blue), Andamanese, and language isolates (those in black).

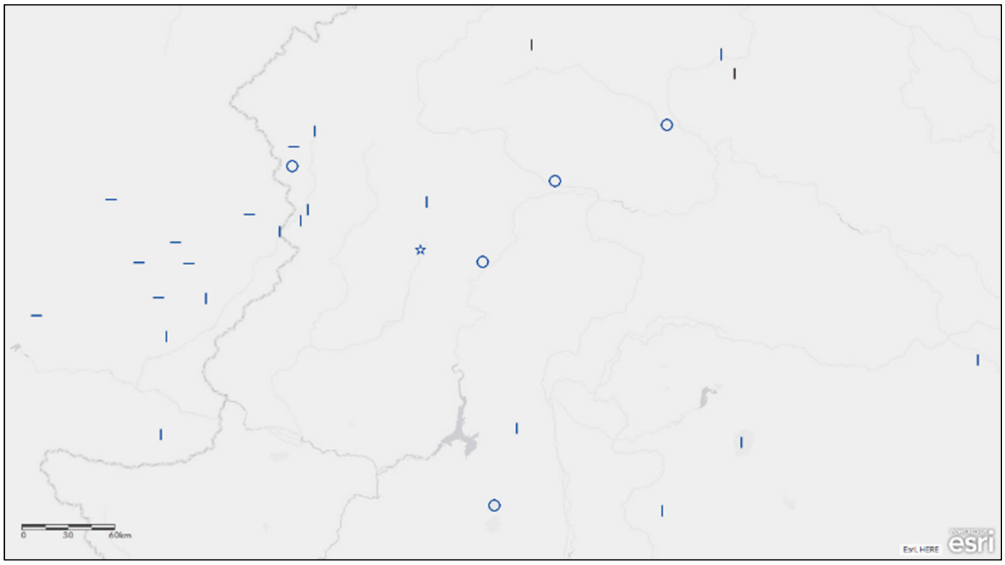


Figure 13.14.2: Types of stop series in Indo-Aryan, Nuristani (both in blue), Andamanese, and language isolates (those in black), enlarged.

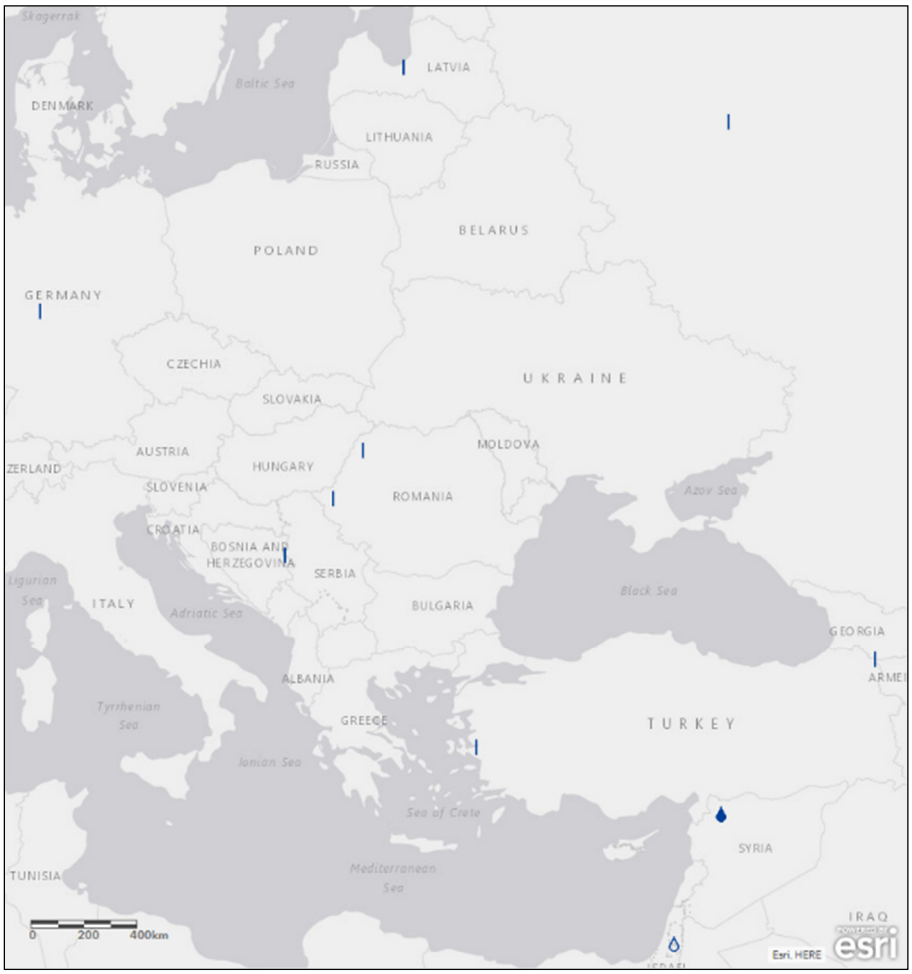


Figure 13.14.3: Types of Indo-Aryan languages outside of South Asia.

Stop series in Dravidian

1. Classification

In this map, stop series are classified as 3 large categories: *t-n* type, *t-d-n* type, and *t-d-dh-n* type.

A. *t-n* type
  t- [t-], n- [n-]
    -t- [-ð-], -n- [-n-]
    -tt- [-tt-] (< *-tt, <*-ntt)
    -nt- [-nd-] -nn- [-nn-]

B. *t-d-n* type
  t- [t-], d- [d-], n- [n-]
    -t- [-t-], -d- [-ð-]~[-d-], -n- [-n-]
    -tt- [-tt-], -dd- [-dd-], -nn- [-nn-]
    -nt- [-nt-](<*-ntt), -nd- [-nd-](<*-nt)

C. *t-d-dh-n* type
  C-1 *t-d-dh-n* type
  t [t^h], d [d], dh [$d^ɦ$]~[$t^ɦ$], n [n]
  C-2 *th-t-d-dh-n* type
  th [t^h], t[t], d [d], dh [$d^ɦ$], n [n]

2. Geographical distribution and interpretation

The Proto-Dravidian stop series are reconstructed as belonging to the *t-n* type, without phonemic distinctions between voiced/voiceless nor aspirated/non-aspirated stop consonants, as is the case with Old Tamil. Since Tamil Brāhmī script shed all the voiced or aspirated consonant characters for oral stops, Tamil orthography has never reintroduced a device to distinguish the voiced stops which appears to have become phonemic through lexical borrowing at least in the word initial position in most spoken Tamil dialects on the subcontinent. The stop series in Lankan Tamil dialects, on the other hand, are reported to have remained of this archaic type, such as in Jaffna variety recorded by S. Kuno (1958).

The *t-d-n* type is dominant elsewhere for the stop series in Dravidian. The reflex of the PDr. series as reconstructed above is observed in alternation between the initial voiceless and voiced stops as allomorphy in most languages.

Types which involve aspiration, i.e. C-1 and C-2 are attributed to the contact with Indo-Aryan languages. C-1 type is typically observed in the so-called educated speech of the languages with a long literary tradition, which incorporated a large amount of Sanskrit vocabulary and (except Tamil) its phonetic and phonological treatises. Voiceless aspirated stops are the less stable of the two aspirated series in this type, probably because voiceless stops in these languages were inherently aspirated. In order to maintain the distinction, borrowed voiceless aspirated stops tend to be either replaced by voiceless fricatives or characterized by an extra-long VOT which makes the following vowel as breathy as those following the voiced aspirated stops. The dental series in Telugu is known to have taken the latter course, ending up in a merger of the voiced and voiceless aspirated stops, as is shown on the map. Similar phonetic descriptions on some varieties of Kannada and Malayalam are found in literature but not represented on the map.

C-2 type is also found in tribal languages in north and central India. Only Kurukh and Naiki are shown on the map as this type, although there are reports of dialects of this

type in Gondi and Pengo. This type may suggest the extent of bilingualism with a Modern Indo-Aryan languages in the area.

Aspirated stops are not limited to borrowings and expressives in some languages. Kobayashi & Tirkey (2017: 34-35) discusses Kurukh spontaneous aspiration in medial positions in addition to the initial *kh* which is cognate to Malto *q*. Aspirated sounds and consonant clusters are reported to be distinct in Kurukh.

Bh. Krishnamurti (2003: 155) includes aspiration in Telugu and Old Kannada numerals in his evidence for the Dravidian laryngeal theory. (PDr. *CVHCV > Telugu C[H]VCV).

OTe. *padi* '10' *ēṃbhadi* '50'

OKa. *ombhattu* '9' *tombhattu* '90' cf. *hattu* < *pattu* '10'

Mod.Te. *padi* '10' *iravay* '20' *mupphay* '30' *nalabhay* '40' *ēbhay* '50' *aravay* '60' *ḍebbhay* '70' *enabhay* '80' *tombhay* '90'

Together with some more examples of Telugu numerals shown below, these may suggest a cluster origin of the Telugu aspirated stops. Inserted *-h-* appears below to block deletion of the preceding short vowel by a V-V sandhi which would result in a monomoraic allomorph:

pada-k-oṇḍu '11' *paṇ-ṇeṇḍu* '12' *pada-mūḍu* '13' *pad(h)-nālugu* '14' *padi-h-ēnu* '15' *pada-h-āru* '16' *padi-h-ēḍu* '17' *padd-h-enimidi* '18' *pan-dhommidi* '19'

(KODAMA Nozomi)

A. *t-n*
B. *t-d-n*
C-1 *t-d-dh-n*
C-2 *th-t-d-dh-n*

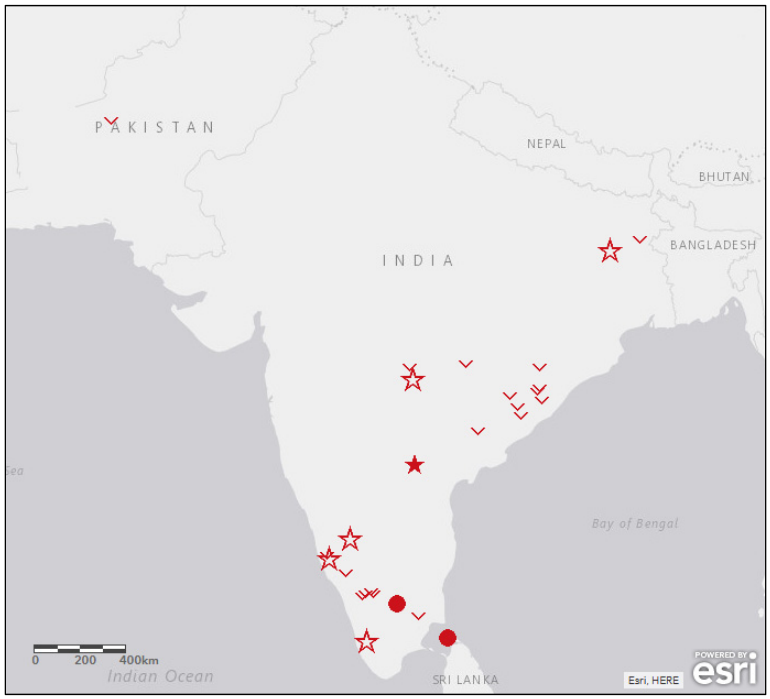


Figure 13.15.1: Stop series in Dravidian.

Stop series in Iranian and Armenian

1. Classification

Many modern Iranian languages have the same stop series as that of Proto-Iranian (PIr.), in which *t-*d-*n is reconstructed. However, some Iranian languages have developed new stop series, such as aspirated, ejective and pharyngealized/ velarized stops mainly due to language contacts. Although it is not an Iranian language, we also deal with Armenian here because it has strong relation with Iranian languages both linguistically and geographically.

In this map, stop consonant series are divided into following five large categories (type A through type E) with some subgroups.

Type A		t-d-n
Type B	B-1	th-d-n
	B-2	th-t-d-n
	B-3	th-t-dh-d-nh-n
Type C		th-t'-d-n
Type D	D-1	t-t$^{\varsigma}$-d-n
	D-2	t-t$^{\varsigma}$-d-d$^{\varsigma}$-n
Type E		th-t-t$^{\varsigma}$-d-n

2. Geographical distribution and interpretation

Type A is the commonest stop series in Iranian languages, especially in Eastern regions. Although type A is the same as the reconstructed PIr. stop series (*t-*d-*n), this does not straightforwardly correspond with type A (t-d-n) respectively. This type includes Persian, Tajik and Dari (Southwestern), Central Kurdish (Sorani) and some dialects of Balochi (Northwestern), Pashto, Yazglami, Shughni-Roshani group with Sarikoli, Ishkashimi-Snglechi, Wakhi and Munji-Yidgha (Southeastern), Yaghnobi (Northeastern) and Ormuri (controversial).

Type B is frequently observed around the Caspian Sea and Armenia. This group is divided further into three subgroups: B-1, which includes Tatic (Tati, Vafsi, Talysh), Caspian (Gilaki) and Gorani/Hawrami (Northwestern); B-2, which includes Northern Kurdish (Kurmanji, Northwestern), Eastern Armenian and Eastern Balochi (Northwestern); B-3, which includes Parachi (controversial).

The aspirated stops occur not only in loanwords. For example, according to Stilo (2019: 676), all voiceless stops except /ʔ/ are aspirated in Tatic and Caspian languages. Also, aspirated consonants are found in Northern Kurdish native words (see Haig and Öpengin 2018: 171). Haig and Öpengin (2018: 170) suppose that Kurdish aspirated phonemes are due to Armenian influence.

Parachi, spoken in Afghanistan, has many aspirated stops /th, dh, nh/, although their phonemic status remains to be determined. According to Kieffer (2009), aspiration is confined to initial and medial position, and occurs both in loanwords and Iranian words.

As for Eastern Balochi, the phonemic status of aspirated stop needs further research (see Korn 2005).

Type C is quite a unique series, which is found only in Ossetic (Iron and Digoron dialect), spoken in Caucasus, where ejective is quite common. There is a three-

way contrast in stops: aspirated voiceless, ejective and voiced. Ejectives occur mainly in loanwords, although they could occur in some inherited words from Proto-Iranian.

ex.) Iron. t'yssyn-/ t'st- 'to thrust' < *tund-s cf. OIA. tud- 'beat, hit' (Abaev 1979: 358)

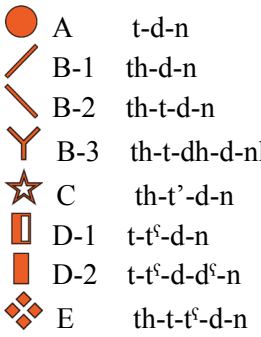
Type D has pharyngealized (or velarized) stops. This group is scattered around Arabian Peninsula, where Arabic is overwhelmingly dominant. It has two subgroups: D-1 and D-2. The former has one pharyngealized stop /tˤ/ whereas the latter has two (/tˤ/ and /dˤ/).

Type E shows features of both type B-1 and D-1, that is, it has both aspirated stop /th/ and pharyngealized one. Only Behdini dialect of Kurdish (Iraq) falls into this type. Interestingly, Type E is located between Type B and Type D-1 zone, which enables us to suppose that neighboring languages play an important role here too.

In conclusion, the stop series types in Iranian languages correlate roughly with their geographical distribution except for Sorani (Type A) and Parachi (Type B-3). This implies that language contact plays important role in Iranian stop series. In fact, many scholars point out that Iranian languages have attained new phonemes through the substrata or neighboring languages (Oranskij (1988: 41) for Eastern Iranian, Ėdel'man and Dodyxudoeva (2009) for Pamir languages, Haig and Öpengin (2018: 170) for Kurdish).

Aspiration in Parachi (especially /nh/) does not seem to be explained by neighboring language's influence because it is not contiguous to any other languages that have aspirated phonemes.

(IWASAKI Takamasa)



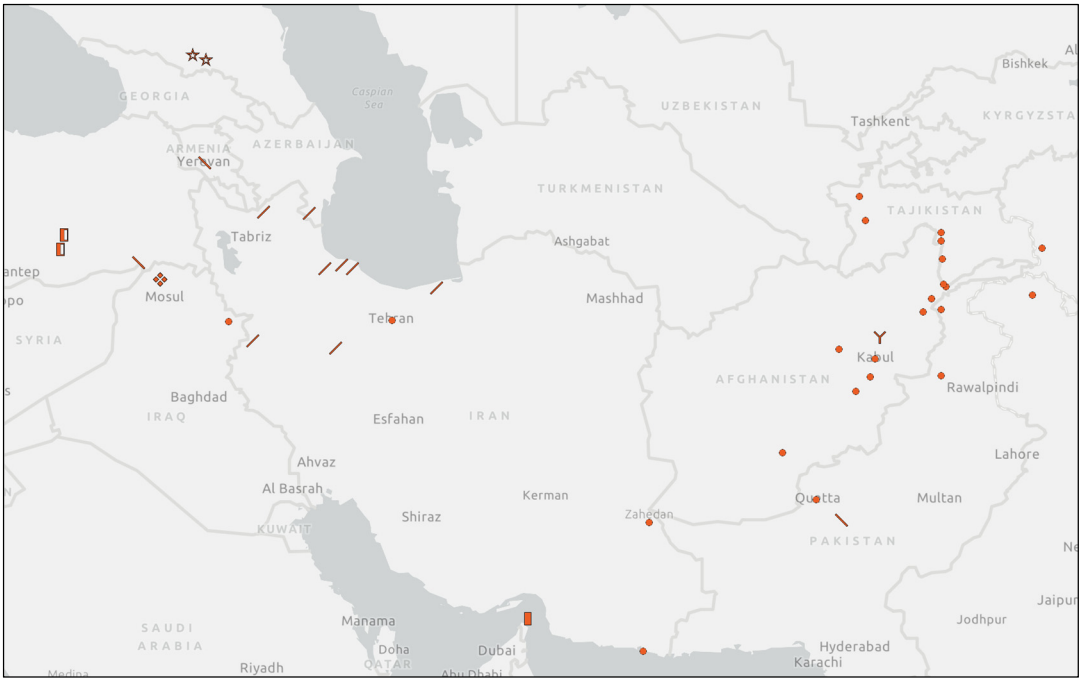


Figure 13.16.1: Stop series in Iranian and Armenian.

Stop series in Caucasian languages

1. Classification

The variation of the denti-alveolar stop series of three language groups in Caucasus, Kartvelian, Abkhazo-Adyghean, and Nakho-Daghestanian is classifies as follows:

A tripartite plosives + a nasal
- A1 t-t'-d-n
- A2 t^h-t'-d-n
- A3 t-t:-d-n

B quadripartite plosives + a nasal
- B1 t-t'-t:-d-n
- B2 t^h-t-t'-d-n
- B3 t^h-t'-t:-d-n

C t-t'-t:-t:'-d-n

Among various phonological interpretations, /t/ in Kartuli may include unaspirated and aspirated variants: [t, t^h], whereas Tschenkéli (1965:XXXII-XXXIII), Fähnrich (1993:18-19), and Kojima (2011:17) clearly mentions that /t/ (interpreted in Table 1) is aspirated. Aliroev (2004:18) describes /t/ in Chechen as an aspirated sound.

The phoneme described as /тт/ in the original documents has two explanations: 'intensive' (Alekseev 1999) and 'unaspirated' (Meylanoba and Sheykhov 1999, Ganieva 2011). In Table 1, they are interpreted as /t:/ and /t/, respectively; in the latter case, the original /т/ is consequently interpreted as /t^h/.

According to Desheriev (1959:12-15), the alveolar stop series in Khinalug is /t-t'-t:-t:'-d-n/. This description and that of Alekseev (1999) are mutually different, but the background of the difference is unidentified.

2. Geographical distribution and interpretation

Noteworthy differences in the components of the series are aspirated and geminate features. It is unclear whether a given language has an aspirated or unaspirated (non-ejective) feature in most cases (except for Type B2). In the Caucasian languages, the ejective feature is nearly pervasive, while the aspirated feature is regarded as a variant of an unaspirated voiceless sound and thus analysed as 'non-ejective', for example in the description by Kartuli by Dzidziguri and Chanishvili (1999:27). A similar phenomenon that a voiceless plosive is often realised as an aspirated sound is attested in several Iranian languages, as reported in Iwasaki (in this volume).

The distribution of Type B is limited to the southern part of the Daghestanian language area. Focusing on the distribution of Type A3, the geminate feature characterising A3 is probably influenced by languages of Type B.

The existence of the geminate ('intensive' or 'unaspirated') consonant is attested in several languages spoken in the region from South Daghestan to Azerbaijan. This sound is related to an unaspirated feature, which might form a contrast with an aspirated feature represented by non-intensive voiceless phoneme.

(SUZUKI Hiroyuki)

A tripartite plosives + a nasal

- ○ Al t-t’-d-n
- ● A2 t^h-t’-d-n
- ⊕ A3 t-tː-d-n

B quadripartite plosives + a nasal

- □ B1 t-t’-tː-d-n
- ■ B2 t^h-t-t’-d-n
- ◪ B3 t^h-t’-tː-d-n
- ▌ C t-t’-tː-tː’-d-n

Figure 13.17.1: Stop series in Caucasian languages.

Stop series in Semitic languages

1. Classification of stop series

The stop series are classified as follows.

A. *t-t'-d-n* series
  A-1. *t-t'-d-n* type
  A-2. *t-ṭ-d-n* type
B. *t-tˤ-d-n* series
  B-1. *t-tˤ-d-n* type
  B-2. *t-tˤ-d-n-dˤ* type
  B-3. *t-d-n-ɗ-dˤ* type
  B-4. *t-d-n* type
  B-5. *t-n* type

2. Geographical distribution and interpretation

A. *t-t'-d-n* type

Type A (*t-t'-d-n*) with an ejective is distributed in the modern period in the Ethiopian area and the southern Arabian peninsula. The Ethiopian Semitic languages such as Amharic (*t'amə* 'taste'), the official language of Ethiopia, Tigrinya (*t'əʃmə*) of the Christian language in Eritrea, and Tigre (*t'əʃma*), spoken by Muslims in the area, all exhibit this type. In addition, in the southern Arabian peninsula, the South Arabian languages including Jibbali (*t'ad* 'one'), Hobyot (*t'áat'*) in Oman, Mehri (*t'ād*) in Yemen, and Soqotri (*t'ad*) in the Soqotra island exhibit this type.

The ancient Semitic languages in these area such as Ge'ez, the classical language of Ethiopia, and the South Arabian epigraphic languages such as Sabaean in Yemen probably had the ejective *t'* as the emphatic *ṭ* because the modern varieties have it.

Type A-2 (*t-ṭ-d-n*) was distributed in the ancient Semitic languages in the Mesopotamia and Syria area, such as Akkadian (*ṭaːbu* 'good'), Ugaritic (*ṭaːbu* 'good'), and ancient Hebrew (*ṭoːb* 'good') These Semitic languages had the emphatic consonants *ṭ*, *ṣ*, and *ḳ*. These emphatics were probably ejectives rather than the pharyngalized *tˤ* (or uvularized, palatalized) as in Arabic because of the lack of a voiced counterpart (Nakano 1998: 15).

B. *t-tˤ-d-n* type

This type is widely distributed throughout the Arabic area, namely in the regions other than the Ethiopian area and the southern Arabian peninsula. In the Syrian region, where the North Semitic languages were spoken, the innovation of the pharyngalization of the ejective *t'* took place.

Thus, classical Syriac (*ṭaːb* 'good') and classical Arabic likely exhibited Type B-1 (*t-tˤ-d-n*). It is possible that the realization of *ṭ* of classical Arabic was a voiced [dˤ], but there is some discussions about this realization in Proto Arabic (Nakao 2018).

Modern Aramaic languages such as Syriac (*tˤaːb* 'good'), Mandaic, Ma'lula Aramaic (*tˤoːb* 'good') and Assyrian, spoken in Iraq and Syria (*tˤava* 'good), Arabic nomadic (Bedouin) dialects such as Iraqi (*tˤeːr* 'bird'), Arabian peninsula and Tunisian have a pharyngal *tˤ* as an emphatic but do not have a pharyngal *dˤ*. In addition, the reflex of *ḍ* in Arabic Bedouin dialects is the fricative *ðˤ*.

Type B-2 (*t-tˤ-d-n-dˤ*) is found in the urban dialects of Arabic such as Cairene in

Egypt, Damascine in Syria, and Maghrebi in Morocco, in which the interdental pharyngal fricative *ðˤ* and **d͡ʒˤ* or **d͡lˤ* (< Proto Semitic *ɬˤ* according to Lipiński 2001: 135) merged into *dˤ* (Cairene *dˤalma* < **ðˤalma* 'darkness'; *dˤarab* < **d͡ʒˤarab* 'to hit'). Thus, the system of the stop series has achieved symmetry in these dialects. Moreover, the dialects have developed the pharyngalized consonants *bˤ*, *mˤ*, and *zˤ* apart from the *rˤ* and *lˤ* that exist in classical Arabic.

Type B-3 (*t-d-n-ɗ-dˤ*), which is a variety of Type B-2, is found in Nigeria. In this dialect, the reflex of *ṭ* is a dental implosive emphatic [ɗˤ] and the reflex of *ḍ* of classical Arabic is *dˤ* (Owens 1993.) An implosive [ɗ] as the reflex of *ṭ* in Aswan, South Egypt is also reported (Schroepfer 2015).

Type B-4 (*t-d-n*) is found in the peripheral varieties of Arabic dialects such as Maltese, Ki-Nubi in Kenya and Bukhari in Uzbekistan, and modern Hebrew. In these varieties, the emphatic phonemes have merged into the non-emphatic counterpart, thus *tˤ* into *t* and *dˤ* into *d*; Ar. *tˤaːr* > Maltese *tar*. In the Bukhari dialect, the interdental fricatives *θ*, *ð*, and *ðˤ* merged into plane dental fricatives *s* and *z*. The speakers of modern Hebrew of European origin pronounce /ṭ/ as *t*, although the speakers of that of the Arabian origin pronounce it as *tˤ* as in Arabic.

Type B-5 (*t-n*) is found only in Cypriot Arabic. In this dialect, the opposition between voiced and unvoiced has disappeared in the stops, as has that between emphatic and non-emphatic.

(NAGATO Youichi)

▲ A-1. t-t'-d-n
△ A-2.t-ṭ-d-n
◐ B-1. t-tˤ-d-n
● B-2. t-tˤ-d-n-dˤ
★ B-3. t-d-n-ɗ-dˤ
✖ B-4. t-d-n
✚ B-5. t-n

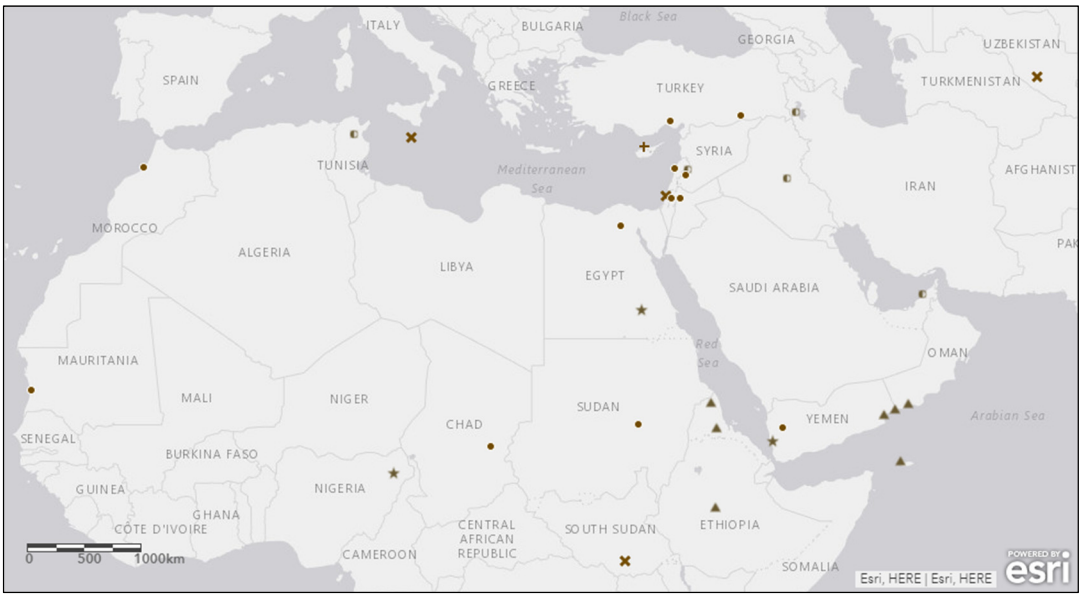


Figure 13.18.1: Stop series in Semitic.

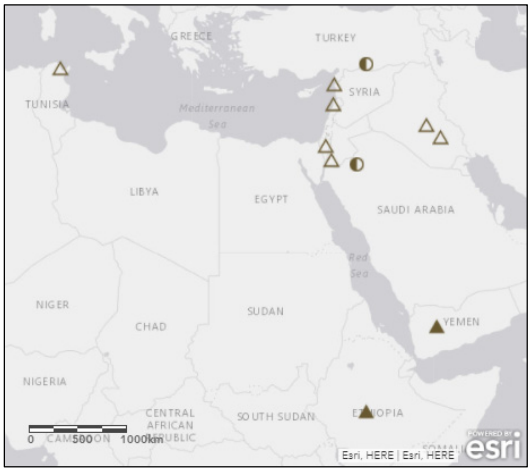


Figure 13.18.2: Stop series in Semitic (Old Semitic).

Stop series in Nilo-Saharan

1. Classification

On this map, the stop series is classified as consisting of 10 types, which could be consolidated into five main types:

A. *t-D-n* types (two-way laryngeal distinction)

A-1 *t-d-n* type (core type)

A-2 *t-ɗ-n* type (including /ɗ/ [d]~[ɗ])

A-3 *t-d-n-nd* type (A-1 plus a prenasalized stop)

A-4 *t-ɗ-n-nd* type (A-2 plus a prenasalized stop)

B. *t-n* type (no laryngeal distinction)

C. *T-d-ɗ-n* type (three-way laryngeal distinction)

C-1 *t-d-ɗ-n* type (core type)

C-2 *t'-d-ɗ-n* type (with an ejective stop)

C-3 *t-d-ɗ-n-nd* type (C-1 plus a prenasalized stop)

D. *t-t'-d-ɗ-n* type (four-way laryngeal distinction)

E. *th-t-t'-d-ɗ-n* type (five-way laryngeal distinction)

2. Geographical distribution and interpretation

Nilo-Saharan is a loosely defined group of African languages spoken between the domains of Afroasiatic and Niger-Congo language phyla. Although there is no consensus about the phylogenetic membership or the internal relationships, at least two large families have been established in the comparative linguistic debates: Central Sudanic (with its Western and diverse Eastern branches) and Eastern Sudanic (Nubian, Nara, Taman, Nyimang, Eastern Jebel, Temein, Daju, Surmic, and Nilotic). The largest of these groups is the Nilotic languages, with its Southern, Eastern and Western sub- branches. In addition, the following languages have appeared in the arena of "Nilo-Saharan" linguistics: Berta, Fur-Amdang, Mabang, Kuliak, Kunama and Saharan, as well as Koman, Gumuz, Songhay, Kadu and Shabo, whose Nilo-Saharan affiliations have been disputed (Dimmendaal 2020). In the following maps, at least one member of these groups is represented. To these we could add the two extinct languages not represented here: Meroitic, spoken in ancient Sudan, and "Mimi of Decorse," recorded in ca. 1900 in Chad.

There is as yet no accepted phylum-level sound correspondence, and the findings of previous studies cannot be taken for granted. To take an example, Ehret (2001) once proposed the proto-Nilo-Saharan stop series as **t-*t'-*d-*ɗ-*n-*nd*, analyzing then available Uduk (Koman; Type E) and typical Central Sudanic (Type C-3) data as the most archaic types. The membership of Uduk (or Koman in general) within Nilo-Saharan, however, is disputable, and it has been recently confirmed that the Koman languages have an additional phoneme *th* (Killian 2015; Otero 2019).

Type A-1 is the most prevalent type, represented by Korandje (Songhay) in Algeria, Nobiin (Nubian) in Egypt, and Datooga (Southern Nilotic) in Tanzania (although the *t* vs. *d* opposition in Datooga could be theoretically analyzed as /tt/ vs. /t/; see Hieda 2001).

Some phonemes and types exhibit obviously areal distributions. The

implosive *ɗ* (Types A-2, A-4, C, D, E) is frequent in so-called Sudanic belt, i.e., from the West African coasts to the southern and western fringes of the Ethiopian Highlands, which is often postulated as an areal feature of this region (Clements and Rialland 2008; Güldemann 2008). Ejective *t'* (Types C-2, D, E) is found almost exclusively among disputed Nilo-Saharan branches (Koman, Gumuz, and Shabo) spoken on the fringes of Ethiopian Highlands, with the notable exception of Ethiopian Berta, which uniquely attests C-2. The existence of the ejective series (and *ɗ*) is a feature shared with Afroasiatic languages of the same region (Omotic, Cushitic, and Ethio-Semitic) and often postulated as an areal feature (Crass & Meyer 2008). Central Sudanic Ngiti (Type D) actually has an implosive [ɗ̥], which is here integrated as a phonetic variant of *t'*. Type B is found only among Southern Nilotic languages in East Africa.

Many Nilo-Saharan branches attest a full or partial distinction of dental vs. alveolar (most Western Nilotic, Gaam (Eastern Jebel), Nyimang and Maba (Mabang)) or alveolar vs. retroflex (Eastern Nilotic Kakwa and some Central Sudanic) series. The point of articulation may be incoherent in some other languages, such as in Mamvu (Central Sudanic; Type A-1) and in Chamus and Ongamo (Eastern Nilotic; Type A-2). Phonetically, however, Mamvu has *ʈ-d-ɳ*, Chamus has *t̪-ɗ-n̪*, and Ongamo has *t̪-ɖ-n*. All Kadu languages have alveolar *ɗ* and *n* in addition to dental *t̪* (and *ɖ*) and retroflex *ʈ* (and *ɖ*) and are classified here as Type A-2. Similarly, Didinga (Surmic), with *t̪-ʈ-d-ɗ-n*, is classified as C-1. Central Sudanic Kresh and Bagirmi (*t-d-n-ʈ-ɖ-ɖ*) and Mangbetu (*t-d-n-ʈ'-ɖ-ɖ*), classified as Type A-1, and Southern Nilotic Pökoot (*t-n-ɖ*), classified as Type B, have retroflex implosive (and ejective). In addition, a few languages under Arabic and/or Berber influence, such as Northern Songhay and Sudanese Berta attest pharyngealized stops /tˤ/ (and /dˤ/).

(NAKAO Shuichiro)

A-1. *t-d-n*
A-2. *t-ɗ-n*
A-3. *t-d-n-nd*
A-4. *t-ɗ-n-nd*
B. *t-n*
C-1. *t-d-ɗ-n*
C-2. *t'-d-ɗ-n*
C-3. *t-d-ɗ-n-nd*
D. *t-t'-d-ɗ-n*
E. *th-t-t'-d-ɗ-n*

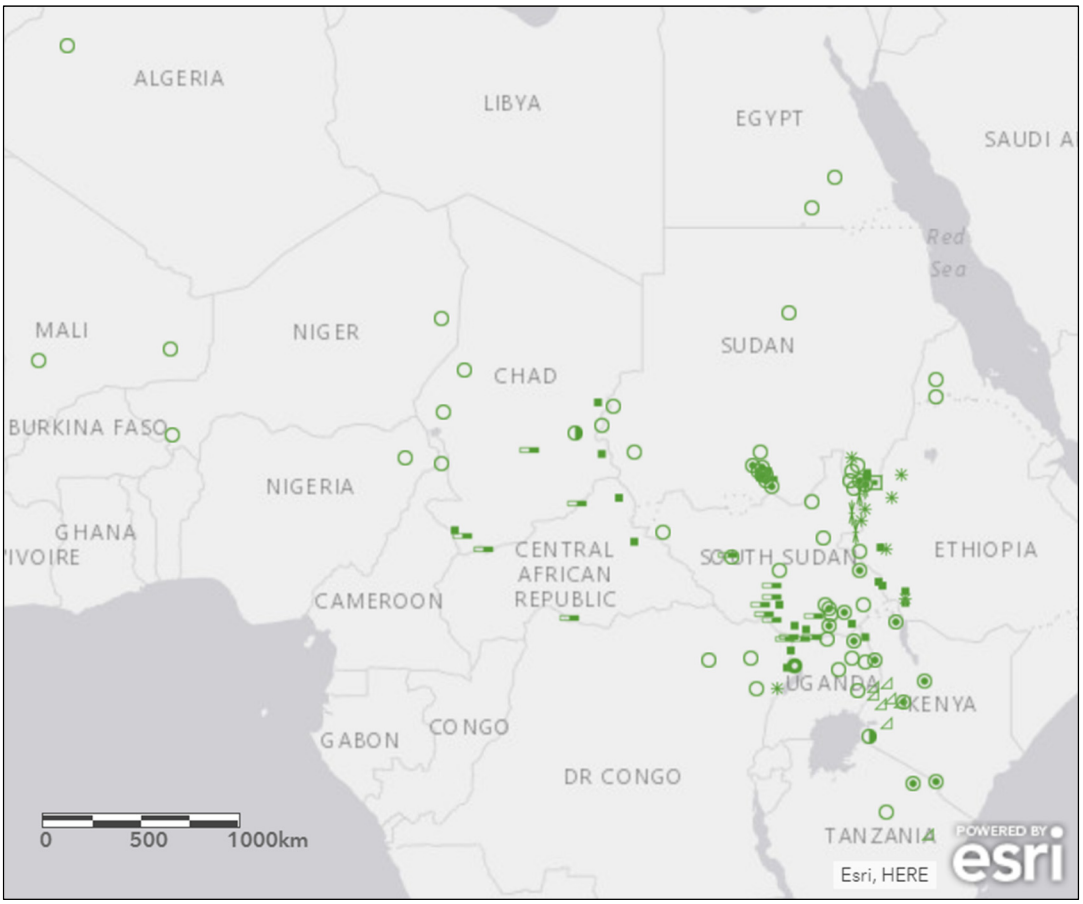


Figure 13.19.1: Stop series in Nilo-Saharan (in general).

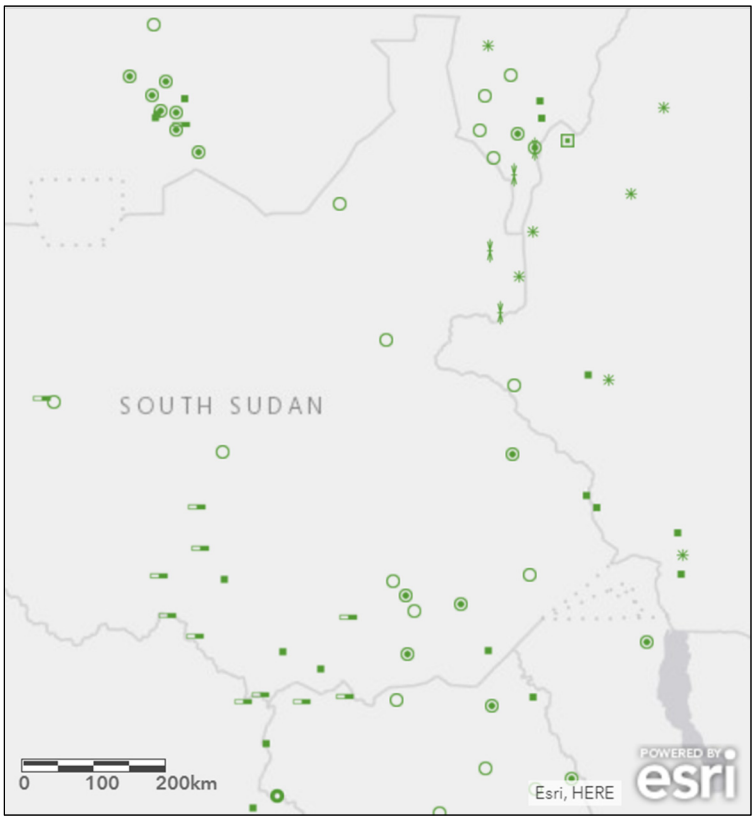


Figure 13.19.2: Stop series in Nilo-Saharan languages around South Sudan.

Stop series in Niger-Congo

1. Classification

The following is a list of articulatory types of stop consonant series attested in 85 sample languages from the following branches: 1. Kordofanian (2 languages), 2. Mande (5), and sub-branches of Atlantic-Congo including 3a. Atlantic (8), 3b. Ijoid (2), and 3c. Volta-Congo (68, including 25 Bantu languages spread over 13 of 15 zones of geographic classification proposed by Guthrie 1970:11−15). Systematic types are primarily classified by the number of distinctions, ranging from 2 to 7, and further subcategorised by the following features defining each subtype: [A̰] Aspiration (*th*), [A̤] Breathiness (*dh*), [NC̥] Voiceless Prenasal/ Nasal Cluster (*nt*), [NC̬] Voiced Prenasal/ NC (*nd*), [GC̥] Ejective (*t'*), and [GC̬] Implosive (*ɗ*). Type codes, consisting of the number of distinctions and feature tags, are provided in square brackets (e.g. [4-A̰-GC̥] for the 4-way distinction with aspirated and ejective consonants), and the number of attested languages of each pattern is shown in parentheses.

A. 2-way distinction
  A-1: [2] *t-d* (1)
  A-2: [2'] *t-n* (1) or *d-n* (1)
B. 3-way distinction
  B-1: [3] *t-d-n* (36)
  B-2: [3-A̰] *th-d-n* (4) or *th-t-n* (1)
  B-3: [3-NC̬] *t-nd-n* (3) or *t-d-nd* (1)
C. 4-way distinction
  C-1: [4] *t-d-nd-n* (5)
  C-2: [4-A̰] *th-t-d-n* (2)
  C-3: [4-A̰-GC̥] *th-t-t'-n* (1)
  C-4: [4-A̰-NC̬] *th-t-nd-n* (1)
  C-5: [4-GC̬] *t-ɗ-ɗ-n* (7)
  C-6: [4-GC̬-NC̬] *t-ɗ-nd-n* (1)
  C-7: [4-NC̥] *t-d-nt-n* (1)
  C-8: [4-NC̥-NC̬] *t-nd-nt-n* (1)
D. 5-way distinction
  D-1: [5] *t-d-ɗ-nd-n* (8)
  D-2: [5-A̰] *th-t-ɗ-nd-n* (1)
  D-3: [5-A̰-A̤-GC̥] *th-t'-dh-ndh-n* (2)
  D-4: [5-NC̥] *t-d-nd-nt-n* (5)
E. 6-way distinction
  E-1: [6] *t-d-ɗ-nd-nt-n* (1)
  E-2: [6-A̰-A̤] *th-t-f-d-dh-n* (1)
F. 7-way distinction
  F-1: [7-A̰-A̤] *th-t-dh-ɗ-nd-ndh-n* (1)

What is immediately suggested by these patterns is that feature [A̰] plays a distinctive role in subcategorisation of all types defined by the number of distinctions, except Type A, which itself can be divided into the 'voice contrast' type (A-1: [2] *t-d*) and the 'oral-nasal' type (A-2: *t-n* or *d-n*). The latter type, in turn, can be regarded as a basis for Type B-2: [3-A̰], which is configurated by adding [A̰] to Type A-2: [2']. On the other hand, the 'voice contrast' type serves as a basis for the 'canonical' type where the marked feature [A̰] is not relevant to the systematic configuration. A configurational hierarchy of the canonical types is formalised as follows: A-1 > +*n* > B-1 > +*nd/ɗ* > C-1/C-5 > +*ɗ/nd* > D-1 > +*nt* > E-1.

2. Geographical distribution and interpretation

A general tendency of geographical distribution is that the Benue-Congo (BC) sub-branch of the Volta-Congo (VC)

languages, especially Southern Bantu languages, shows more complexity than other languages spoken in the western part of the continent (for further discussions on phonological areas in Africa, see Clements & Rialland 2008).

Type A is attested in Kwa and Kru subbranches of VC as well as in Kordofanian.

Type B, which is the overwhelming majority of all types of distinctions, is spread widely throughout the continent but with a concentration in the west. It is also noted that the less marked canonical subtypes are densely distributed in the non-BC area, while [3-Ạ] seems to be typical in Kwa (VC), and [3-NÇ] is dominant in the Bantu area.

Type C consisting of the greatest number of subtypes still shows a regular geographic pattern; while all subtypes with [NÇ] are distributed in the Bantu area, those with the [GÇ] feature are well observed in non-Bantu BC and other VC subgroups such as Kwa and Ubangian.

Type D is also distributed in a principled way; the canonical type is exclusively found in North Volta-Congo as well as in non-VC languages such as Atlantic and Ijoid, while marked subtypes are predominantly distributed in the Bantu area.

While two Type E languages are sporadically found in Igboid (VC) and Atlantic, Type F, the most complex pattern in our sample, is found in the Southern Bantu zone, following the general tendency.

(SHINAGAWA Daisuke & KOMORI Junko)

A-1	[2]	*t-d*	C-6	[4-GÇ-NÇ]	*t-ɗ-nd-n*	
A-2	[2’]	*t-n* or *d-n*	C-7	[4-NÇ]	*t-d-nt-n*	
B-1	[3]	*t-d-n*	C-8	[4-NÇ-NÇ]	*t-nd-nt-n*	
B-2	[3-Ạ]	*th-d-n* or *th-t-n*	D-1	[5]	*t-d-ɗ-nd-n*	
B-3	[3-NÇ]	*t-nd-n* or *t-d-nd*	D-2	[5-Ạ]	*th-t-ɗ-nd-n*	
C-1	[4]	*t-d-nd-n*	D-3	[5-Ạ-Ạ-GÇ]	*th-t’-dh-ndh-n*	
C-2	[4-Ạ]	*th-t-d-n*	D-4	[5-NÇ]	*t-d-nd-nt-n*	
C-3	[4-Ạ-GÇ]	*th-t-t’-n*	E-1	[6]	*t-d-ɗ-nd-nt-n*	
C-4	[4-Ạ-NÇ]	*th-t-nd-n*	E-2	[6-Ạ-Ạ]	*th-t-f-d-dh-n*	
C-5	[4-GÇ]	*t-d-ɗ-n*	F-1	[7-Ạ-Ạ]	*th-t-dh-ɗ-nd-ndh-n*	

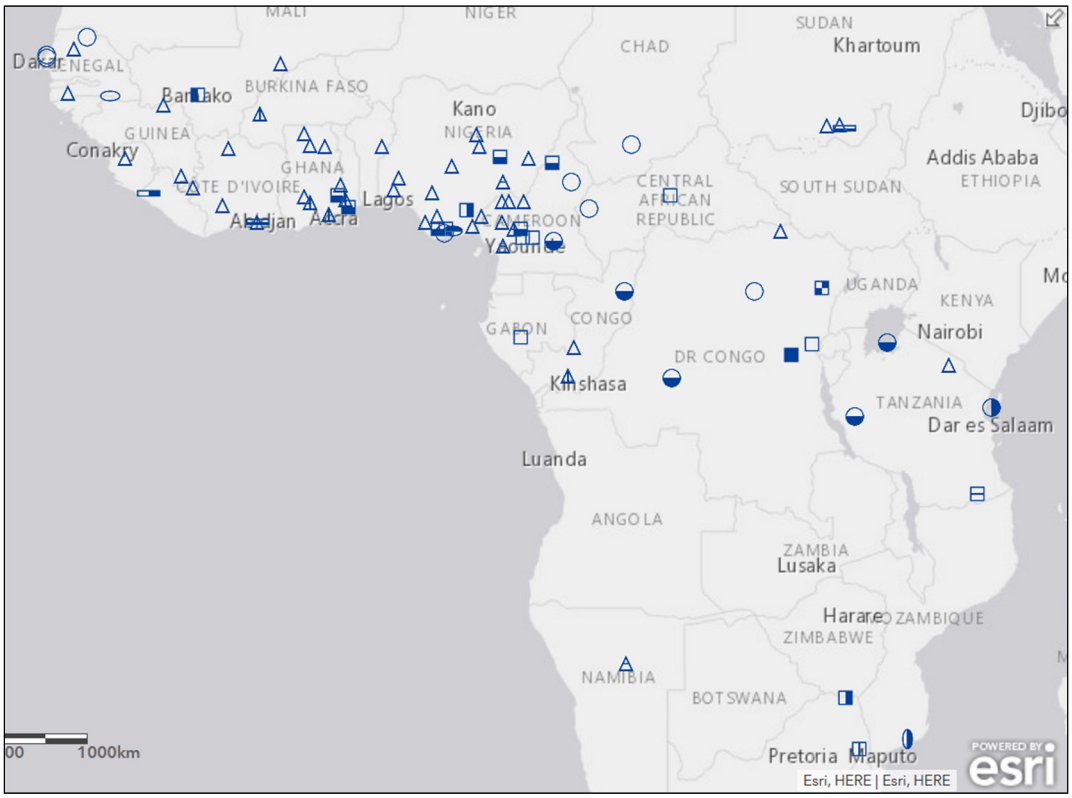


Figure 13.20.1: Geographical distribution of systematic types of stop series in Niger-Congo languages
*Geographical information for plotting the sample languages is based on Glottolog 4.3 (Hammarström et al. 2020).

Figure 13.20.2: Enlargement of the Western Coastal area

Coronal stop series in the Kalahari Basin area

1. Classification

Figure 13.21.1 shows the geographical distributions of selected coronal stop consonants in KBA languages. Classifications are made based on the series types, which are specified in terms of three laryngeal features, that is, voicing, aspiration, and ejection.

In the current sample, five click series types are attested, as illustrated with the relevant laryngeal features below (the click type is represented by the dental |).

A: |-g|-|ʰ-g|ʰ-|'-g|' [±voiced, ±aspirated, ±ejective]

B: |-g|-|ʰ-g|ʰ [±voiced, ±aspirated]

C: |-g|-|ʰ-|' [±voiced, ±aspirated, ±ejective]

D: |-g|-|ʰ [±voiced, ±aspirated]

E: |-|ʰ [±aspirated]

N.B.: There is a hierarchy: {|, |ʰ} > g| > {g|ʰ, |'} > g|'. (This implies [±aspirated] > [±voiced] > [±ejective].)

Non-click alveolar stops also show a parallel tendency to the click series with some disagreements, which yield two subtypes for series types C and D. Table 1 presents the series types of KBA coronal clicks and non-clicks, together with sample languages.

On Figure 1, series types of click and non-click stops are combined, and displayed as types A-E. The three language families in KBA, namely Tuu, Kx'a and Khoe-Kwadi, are marked with orange, brown and light blue symbols respectively.

Table 1: Series types of KBA coronal consonants (Gaps are indicated with ∅).

Type	Click	Non-click	Language
A	|-g|-|ʰ-g|ʰ-|'-g|'	t-d-tʰ-dʰ-t'-∅	• West !Xoon • East !Xoon
B	|-g|-|ʰ-g|ʰ	t-d-tʰ-dʰ	• Tsumkwe Jul'hoan • Heikkinen !Xuun W • Heikkinen !Xuun E
C1	|-g|-|ʰ-|'	t-d-tʰ-t'	• ǂHaba • Xade G|ui • G‖ana • Tshila
C2	|-g|-|ʰ-|'	∅-∅-∅-∅	• Khute G|ui • N!aqriaxe
D1	|-g|-|ʰ	∅-∅-∅	• N|uu
D2	|-g|-|ʰ	t-d-tʰ-t'	• Naro
E	|-|ʰ	t-∅	• Windhoek Khoekhoe

Note that, for reasons of space, other non-click coronal consonants, namely affricates, fricatives, nasals, and liquids, are not discussed in this article.

2. Geographical distribution and interpretation

As seen on Figure 13.21.1, the relationship between the geographical and genealogical distributions of stop series types is not straightforward. Type A is observed only in the Tuu family, type B only in the Kx'a family, and type E only in Namibian Khoekhoe of the Khoe-Kwadi family. In contrast, the other two types are shared by two families: type C by Khoe-Kwadi and Kx'a, and type D by Khoe-Kwadi and Tuu. The cross-family distribution of type C can apparently be explained in terms of language contact, but at this stage, it is still

unclear how the cross-genetic distributions of type D should be accounted for.

In addition, the distribution of click series types indicates that the ubiquitous laryngeal feature among KBA languages is [±aspirated] instead of globally unmarked [±voiced]. Historically, this can be explained as a result of the tonogenesis occurring in Khoekhoe (Haacke 1999). Khoekhoe is the only language (cluster) in KBA that contrasts four level tones and two simple (non-complex) stop series, that is, [–voiced, –aspirated] vs. [–voiced, +aspirated]. In contrast, all other sample languages contrast less than four level tones and at least three simple series, that is, [–voiced, –aspirated], [–voiced, +aspirated], and [+voiced, –aspirated]. This suggests that the pitch lowering associated with the voiced series was phonologized and the contrast in voicing was neutralized during a certain stage of Khoekhoe history. However, this is an over-simplified scenario and there are complicated details that we are not ready to present at this stage.

Unlike clicks, the non-click alveolar stops in the KBA languages involve many gaps, the detailed discussion of which is beyond the scope of this article. Comparative investigations have revealed that palatalization (/t d t^h t'/ > /c ɟ c^h c'/) yielded the gaps of the alveolar stops (type C2) in Khute G|ui, Khoe-Kwadi (Nakagawa 1998), and N!arqriaxe, Kx'a (Gerlach 2018). A similar sound shift probably involved the loss of the alveolar stops in N|uu (type D2) in the Tuu family.

(KIMURA Kimihiko,
NAKAGAWA Hirosi)

Tuu (orange)

A:	|-g|-|^h-g|^h-|'-g|'	t-d-t^h-d^h-t'-Ø
D1:	|-g|-|^h	Ø-Ø-Ø

Kx'a (brown)

B:	|-g|-|^h-g|^h	t-d-t^h-d^h
C2:	|-g|-|^h-|'	Ø-Ø-Ø-Ø

Khoe-Kwadi (light blue)

C1:	|-g|-|^h-|'	t-d-t^h-t'
C2:	|-g|-|^h-|'	Ø-Ø-Ø-Ø
D2:	|-g|-|^h	t-d-t^h-t'
E:	|-|^h	t-Ø

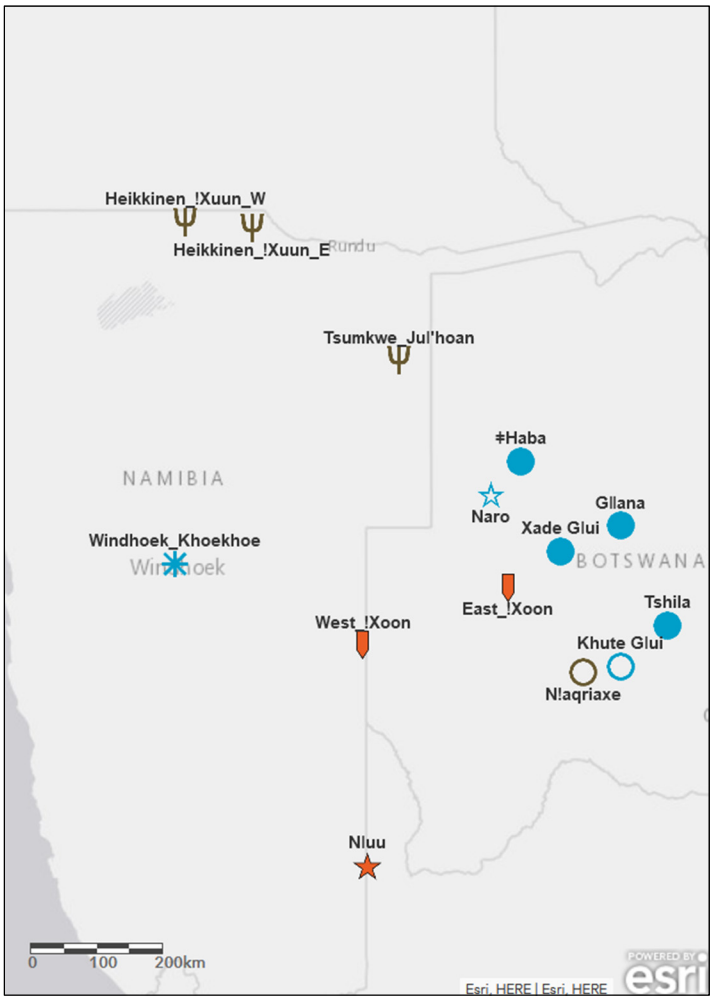


Figure 13.21.1: The geographical distribution of coronal series types.

References

Abaev, V. I. [Абаев, В. И.] (1973) *Историко-этимологический словарь осетинского языка том II. l-r*. Ленинград: Наука.

Abaev, V. I. [Абаев, В. И.] (1979) *Историко-этимологический словарь осетинского языка. том III. s-t'*. Ленинград: Наука.

Alekseev, M. E. [Алексеев, М. Е.] (1999) Хиналугский язык. In M. E. Алексеев (ред.) *Языки мира: Кавказские языки*, 460–469. Москва: Издательство Academia.

Alekseev, M. E. [Алексеев, М. Е.] (отв. ред) (1999) *Языки мира: Кавказские языки*. Москва: Издательство Academia.

Aliroev, Ibragim Yu. [Алироев, Ибрагим Ю.] (2004) *Самоучитель чеченского языка*. Москва: Издательство Academia.

Ball, Martin J., Sara. J. Howard and Kirk Miller (2018) Revisions to the extIPA chart. *Journal of the International Phonetic Association* 48(2): 155–164. doi: https://doi.org/10.1017/S0025100317000147

Bashir, Elena and Thomas J. Conners (2019) *A descriptive grammar of Hindko, Panjabi, and Saraiki*. Berlin: De Gruyter Mouton. doi: https://doi.org/10.1515/9781614512257

Batchelor, John (1938) *An Ainu-English-Japanese dictionary* 『アイヌ・英・和辭典』. 4th edition. Tokyo: Iwanami Syoten.

Baxter, William H. and Laurent Sagart (2014) *Old Chinese: A new reconstruction*. Oxford: Oxford University Press. doi: https://doi.org/10.1093/acprof:oso/9780199945375.001.0001

Berger, Hermann (1998) *Die Burushaski-Sprache von Hunza und Nager. Teil III: Wörterbuch*. Wiesbaden: Harrassowitz Verlag.

Blust, Robert A. (2009) *The Austronesian languages*. Canberra: Pacific Linguistics, Research School of Pacific and Asian Studies, Australian National University. Revised edition (2013) available at http://hdl.handle.net/1885/10191

Bradley, David (2011) Proto-Tibeto-Burman grain crops. *Rice* 4: 134–141. doi: https://doi.org/10.1007/s12284-011-9074-y

Burrow, T. and M. B. Emeneau (1984) *A Dravidian etymological dictionary*, 2nd ed. Oxford: Clarendon Press.

Cao, Zhiyun 曹志耘 (2008) *Hanyu fangyan dituji* 《汉语方言地图集》. Beijing: Shangwu Yinshuguan.

Cardona, George (2003) Sanskrit. In: George Cardona and Dhanesh Jain (eds.) *The Indo-Aryan languages*, 104–160. Oxford: Routledge. doi: https://doi.org/10.4324/9780203945315

Chao, Yuan Ren 赵元任 (1928) *Xiandai Wuyu de yanjiu* 《现代吴语的研究》. Guoli Qinghua daxue chuban shiwusuo.

Chao, Yuan Ren (1968) *A grammar of Spoken Chinese*. Berkeley: University of California Press.

Chiri, Mashiho 知里真志保 (1976[1953]) *Bunrui Ainugo jiten: Shokubutsu-hen* 『分類アイヌ語辞典：植物篇』. *Chiri Mashiho chosakushū bekkan* 1 『知里真志保著作集別巻1』. Tokyo: Heibonsha.

Clements, G. N. and Annie Rialland (2008) Africa as a phonological area. In: Bernd Heine and Derek Nurse (eds.) *A linguistic geography of Africa*, 36–85. Cambridge: Cambridge University Press. doi: https://doi.org/10.1017/CBO9780511486272.004

Cui, Rongchang 崔荣昌 (1996) *Sichuan jingnei de Xiang fangyan*《四川境内的湘方言》. Zhongyang Yanjiuyuan Lishi Yuyan Yanjiusuo.

Dahl, Otto Christian (1981) *Early phonetic and phonemic changes in Austronesian*. Oslo: Universitetsforlaget.

Desheriev, Yu. D. [Дешериев, Ю. Д.]. 1959. *Грамматика хиналугского языка*. Москва: Издательство Академии Наук СССР.

Diao, Xianmin, Guanqing Jia (2017) Origin and domestication of foxtail millet. In A. Doust and Xianmin Diao (eds) *Genetics and genomics of Setaria. Plant genetics and genomics: Crops and models*, vol 19, 61–72. Cham: Springer. https://doi.org/10.1007/978-3-319-45105-3_4

Dryer, Matthew S. and Martin Haspelmath (eds) (2013) *The world atlas of language structures online*. Jena: Max Planck Institute for Evolutionary Anthropology. URI: https://wals.info/

Duan, Haifeng 段海凤 and Xiaonong Zhu 朱晓农 (2018) Chaoxianyu de ruan yin fuyin: Cong yuyin shuju dao yinfa fanchou 朝鲜语的软音辅音：从语音数据到音法范畴. *Minzu Yuwen* 3: 13–25.

Duanmu, San (2016) *A theory of phonological features*. Oxford: Oxford University Press. doi: https://doi.org/10.1093/acprof:oso/9780199664962.001.0001

Dyen, Isidore (1953) *The Proto-Malayo-Polynesian laryngeals*. Baltimore: Linguistic Society of America. doi: https://doi.org/10.2307/j.ctt1x76d60

Dzidziguri, Sh. and N. Chanishvili [Дзидзигури, Ш. и Н. Чанишвили] (1999) Грузинский язык. In M. E. Алексеев (ред.) *Языки мира: Кавказские языки*, 20–52. Москва: Издательство Academia.

Evans, Jonathan (2006) Origins of vowel pharyngealization in Hongyan Qiang. *Linguistics of the Tibeto-Burman Area* 29(2): 95–126. doi: http://doi.org/10.15144/LTBA-29.2.95

Ėdel'man, D. I. and L. R. Dodyxudoeva (2009) The Pamir languages. In: Gernot Windfuhr (ed.) *The Iranian languages*, 773–786. London: Routledge. doi: https://doi.org/10.4324/9780203641736

Fähnrich, Heinz (1993[1986]) *Kurze Grammatik der georgischen Sprache*, 3. Auflage. Leipzig: Langenscheidt-Verlag Enzyklopädie.

Fukazawa, Mika (2021) 'Rice plant' in Ainu. In: Mitsuaki Endo et al. (eds) *Linguistic atlas of Asia*, 61. Tokyo: Hituzi Syobo.

Gair, James W. (2003) Sinhala. In: George Cardona and Dhanesh Jain (eds.) *The Indo-Aryan languages*, 766–817. Oxford: Routledge. doi: https://doi.org/10.4324/9780203945315

Ganieva, F. A. [Ганиева, Ф. А.]. 2011. *Диалекты и говоры самурского наречия лезгинского языка (мазинский и гутумский диалекты, мичахский и смугульский говоры)*. Махачкала: Институт языка и литературы и искусства им. Г. Цадасы.

Gong, Xun (2019) Uvulars and uvularization in Tangut phonology. *Language and Linguistics* 21(2): 175–212. doi: https://doi.org/10.1075/lali.00060.gon

Gruzdeva, E. Yu. [Груздева. Е. Ю.] (1997) Нивхский язык. In: *Языки мира. Палеоазиатские языки*, 139-154. Москва: Индрик.

Güldemann, Tom (2008) The Macro-Sudan belt: Towards identifying a linguistic area in Northern Sub-Saharan Africa. In: Bernd Heine and Derek Nurse (eds.) *A linguistic geography of Africa*, 151–185. Cambridge: Cambridge University Press. doi: https://doi.org/10.1017/CBO9780511486272.006

Guthrie, Malcolm (1970) *Comparative Bantu: An introduction to the comparative linguistics and prehistory of the Bantu languages*. Vol. 3. London: Gregg Press.

Haig, Geoffrey and Ergin Öpengin (2018) Kurmanji Kurdish in Turkey: Structure, variety and status. In Bulut Christiane (ed.) *Linguistic minority in Turkey and Turkic-speaking minority periphery*, 157–229. Wiesbaden: Harrassowitz.

Hattori, Shirō 服部四郎 (1964) *Ainugo hōgen jiten* 『アイヌ語方言辞典』. Tokyo: Iwanami shoten.

Hattori, Shirō 服部四郎 and Mashiho Chiri 知里真志保 (1960) Ainugo shohōgen no kisogoi tōkeigakuteki kenkyū アイヌ語諸方言の基礎語彙統計学的研究. *The Japanese Journal of Ethnology* 24(4): 307–342. doi: https://doi.org/10.14890/minkennewseries.24.4_307

Herin, Bruno (2012) The Domari language of Aleppo (Syria). *Linguistic Discovery* 10(2): 1–52. doi: http://doi.org/10.1349/PS1.1537-0852.A.412

Hieda, Osamu 稗田乃 (2001) Datoogago no shiin taikei saikoo ダトーガ語の子音体系再考. *Suwahiri & Afurika Kenkyuu* 11: 102–117. doi: https://doi.org/10.18910/71091

Huang, Jingui [黄金贵] (2016) *Gudai wenhua ciyi jilei biankao (xin yi ban)* 《古代文化词义集类辨考（新一版）》. Beijing: The Commercial Press.

Inenlikej, P. I. [Инэнликэй П. И.] (2006) *Словарь чукотско-русский и русско-чукотский*. Санкт-Петербург: Дрофа.

Iosad, Pavel (2022) Den historiske utviklinga til preaspirasjon i samiske språk. *Nordlyd* 46(1): 75–101. doi: https://doi.org/10.7557/12.6350

IPA (1999) *Handbook of the International Phonetic Association*. Cambridge: Cambridge University Press.

Kagaya, Ryohei (1974) A fiberscopic and acoustic study of the Korean stops, affricates and fricatives. *Journal of Phonetics* 2: 161–180. doi: https://doi.org/10.1016/S0095-4470(19)31191-X

Karlsson, A. and Jan-Olof Svantesson (2012) Aspiration of stops in Altaic languages: An acoustic study. *Altai Hakpo* 22: 205–222. URI: http://hosting01.snu.ac.kr/~altai/askv001/index.php?document_srl=1519

Kato, Atsuhiko (2009) A basic vocabulary of Htoklibang Pwo Karen with Hpa-an, Kyonbyaw, and Proto-Pwo Karen forms. *Asian and African Languages and Linguistics (AALL)* 4: 169–218. URI: http://hdl.handle.net/10108/61391

Kawamoto, Kunie 川本邦衛 (2011) *Vietnamese-Japanese dictionary* 『詳解ベトナム語辞典』. Tokyo: Taishukan.

Kayano, Shigeru 萱野茂 (2002) *Kayano Shigeru no Ainugo jiten: Zōho-ban* 『萱野茂のアイヌ語辞典：増補版』. Tokyo: Sanseidō.

Kieffer, Charles M. (2009) Parachi. In: Gernot Windfuhr (ed.) *The Iranian languages*, 693–720. London: Routledge. doi: https://doi.org/10.4324/9780203641736

Killian, Don (2015) Aspects of Uduk morphology and phonology. Doctoral dissertation, University of Helsinki. URI: http://hdl.handle.net/10138/229495

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.0000007344.43938.4c

Klimov, G. A. [Климов, Г. А.] (1994) *Einführung in die kaukasische Sprachwissenschaft* (Aus dem Russischen übersetzt und bearbeitet von Jost Gippert; *Введение в кавказское языкознание*, 1986, Наука). Hamburg: Helmut Buske Verlag.

Kojima, Yasuhiro 児島康宏 (2011) *Nyuuekusupuresu Guruziago*『ニューエクスプレス グルジア語』. Tokyo: Hakusuisya.

Kobayashi, Masato and Bablu Tirkey (2017) *The Kurux language: Grammar, texts and lexicon*. Leiden: Brill. doi: https://doi.org/10.1163/9789004347663

Konow, Sten (1913) *Bashgali dictionary: an analysis of Colonel J. Davidson's notes on the Bashgali language*. Calcutta: Asiatic Society. (Reprinted in 1986, Delhi: Gian Publishing House.)

Korhonen, Mikko (1981) *Johdatus lapin kielen historiaan*. Helsinki: Suomalaisen Kirjallisuuden Seura.

Korn, Agnes (2005) *Towards a historical grammar of Balochi: Studies in Balochi historical phonology and vocabulary*. Wiesbaden: Ludwig Reichert.

Krishnamurti, Bhadriraju (2003) *Dravidian languages*. Cambridge: Cambridge University Press. doi: https://doi.org/10.1017/CBO9780511486876

Kuno, Susumu (1958) Phonemic structure of Colloquial Tamil. *Gengo Kenkyu* 34: 16–42. doi: https://doi.org/10.11435/gengo1939.1958.34_41

Kurebito, Megumi 呉人惠 (2009) *Koryak gengo-minzokushi*『コリャーク言語民族誌』. Sapporo: Hokkaido University Press.

Kurebito, Megumi, Tokusu Kurebito, Yukari Nagayama, Chikako Ono, Mitsuhiro Yazu (2001) *Comparative basic vocabulary of the Chukchee-Kamchatkan language family:1*. (Megumi Kurebito, ed.) Suita: Osaka-Gakuin University.

Ladefoged, Peter and Ian Maddieson (1996) *The sounds of the World's languages*. Hoboken: Blackwell Publishers.

Lee, Ki-Moon and S. Robert Ramsey (2011) *A history of the Korean language*. Cambridge: Cambridge University Press. doi: https://doi.org/10.1017/CBO9780511974045

Li, Fang-Kuei (1977) *A handbook of comparative Tai*. Honolulu: The University Press of Hawai'i.

Lipiński, Edward (2001) *Semitic languages: Outline of a comparative grammar* (2nd edition). Leuven: Peeters.

Long, Haiyan 龙海燕 (2008) *Dongkou Gan fangyan yuyin yanjiu*《洞口赣方言语音研究》. Beijing: Minzu Chubanshe.

Lu, Houyuan, Jianping Zhang, Kam-biu Liu, Naiqin Wu, et al. (2009) Earliest domestication of common millet (*Panicum miliaceum*) in East Asia extended to 10,000 years ago. *Proceedings of the National Academy of Sciences* 106(18): 7367–7372. doi: https://doi.org/10.1073/pnas.0900158106

Lu, Jifang 卢继芳 (2007) *Ganyu Changdu pian fangyan yuyin yanjiu*《赣语昌都片方言语音研究》. Beijing: Shangwu Yinshuguan.

Machida, Kazuhiko 町田和彦 (ed.) (2001) *Kareinaru Indokei mozi*『華麗なるインド系文字』. Tokyo: Hakusuisya.

Maddieson, Ian (1984) *Patterns of sounds*. Cambridge: Cambridge University Press. doi: https://doi.org/10.1017/CBO9780511753459

Maddieson, Ian (2013) Consonant Inventories, WALS online. https://wals.info/ [accessed October 2020]

Matisoff, James A. (1975) Rhinoglottophilia: The mysterious connection between nasality and glottality. In Ch. A. Ferguson, L. M. Hyman and J. Ohala (eds.). *Nasálfest: Papers from a Symposium on Nasals and Nasalization*, 265–287. Stanford: Stanford University.

Matisoff, James A. (2003) *Handbook of Proto-Tibeto-Burman: System and philosophy of Sino-Tibetan reconstruction.* Berkeley, Los Angeles and London: University of California Press. URI: https://escholarship.org/uc/item/19d79619

Matisoff, James A. (2015) *The Sino-Tibetan etymological dictionary and thesaurus*. Berkeley: The Regents of the University of California.

Matras, Yaron (2012) *A grammar of Domari*. Berlin: De Gruyter Mouton. doi: https://doi.org/10.1515/9783110291421

Meylanova, U. A. and E. M. Sheykhov [Мейланова, У. А. и Э. М. Шейхов] (1999) Лезгинский язык. In M. E. Алексеев (ред.) *Языки мира: Кавказские языки*, 373-385. Москва: Издательство Academia.

Mills, Roger Frederick (1981) Additional addenda. In: Robert A. Blust (ed.) *Historical linguistics in Indonesia*, Part I, 59–82. Jakarta: Badan Penyelenggara Seri Nusa, Universitas Atma Jaya Jakarta.

Mon-Khmer Etymological Dictionary. http://sealang.net/monkhmer/

Moshiogusa 藻汐草 = Uehara, Kumajirō 上原熊次郎 (1792) Moshiogusa 藻汐草. https://www.wul.waseda.ac.jp/kotenseki/html/ho02/ho02_05038/index.html [accessed January 2023]

Munda Etymological Dictionary. http://sealang.net/munda/dictionary/

Murayama, Shichirō 村山七郎 (1971) *Kitachishima Ainugo*『北千島アイヌ語』, 134–244. Tokyo: Yoshikawa Kobunkan.

Nakano, Akio 中野曉雄 (1998) A sketch of phonetic and phonological systems of Afro-Asiatic languages: Consonantal features アフロ・アジア語の音声・音韻ー子音音素論. *Journal of the Phonetic Society of Japan* 2(3): 9–30. doi: https://doi.org/10.24467/onseikenkyu.2.3_9

Nakao, Shuichiro 仲尾周一郎 (2018) Shūen Arabiago ni okeru kōtōkaon 周縁アラビア語における喉頭化音. *Journal of Arabic and Islamic Studies* 16: 71–92.

Nickel, Klaus Peter (1994[2]) *Samisk grammatikk*. Guovdageaidnu: Davvi Girji OS.

Nielsen, Konrad (1979[1926-1929]) *Lærebok i lappisk (samisk)* I-III. Oslo: Universitetsforlaget.

Ogura, Shinpei 小倉進平 (1944) *Chōsengo hōgen no kenkyū*, 2 vols.『朝鮮語方言の研究』. Tokyo: Iwanami Shoten.

Ono, Chikako 小野智香子 (2020) *Itelmen-go bunpo. Dousi keitairon wo chuusin ni*『イテリメン語文法：動詞形態論を中心に』. Sapporo: Hokkai-Gakuen University Press.

Oranskij I. M. [Оранский, И. М.] (1988) *Введение в иранскую филологию.* Москва: Наука.

Osada, Toshiki 長田俊樹 (1995) *Munda zin no nookoo bunka to syokuzi bunka---Indo bunka-inasaku bunka-syooyoozyu bunka---*『ムンダ人の農耕文化と食事文化—インド文化・稲作文化・照葉樹林文化—』. Kyoto: International Research Center for Japanese Studies.

Otero, Manuel Alejandro (2019) A historical reconstruction of the Koman language family. Doctoral dissertation, University of Oregon. URI: https://scholarsbank.uoregon.edu/xmlui/handle/1794/25239

Owens, Jonathan (1993) Nigerian Arabic in comparative perspective. *Sprache und Geschichte in Afrika* 14: 85–176.

Pittayaporn, Pittayawat (2009) The phonology of Proto-Tai. Ph.D. dissertation, Cornell University. URI: https://hdl.handle.net/1813/13855

Rastorgueva, V. S. [Расторгуева В. С.] and D. I. Ėdel'man [Д. И. Эдельман] (2007) *Этимологический словарь иранских языков. том 3. f-h*. Москва: Восточная литература РАН.

Ratliff, Martha (2010) *Hmong-Mien language history*. Canberra: Pacific Linguistics. doi: https://doi.org/10.15144/PL-613

Ross, Malcom (1995) Introduction to the *Comparative Austronesian dictuionary*. In Darrell T. Tryon (ed) *Comparative Austronesian dictionary*. Part 1: Fascicle 1. Berlin, New York: Mouton de Gruyter. doi: https://doi.org/10.1515/9783110884012.1.1

Sakamoto, Yasuyuki 坂本恭章 (2001) *Cambodian-Japanese Dictionary*『カンボジア語辞典』. Tokyo: Research Institute for the Study of Asian and African Languages and Cultures. URI: http://hdl.handle.net/10108/82382 ; http://hdl.handle.net/10108/82383 ; http://hdl.handle.net/10108/82384

Sawai, Harumi 澤井春美 (1989) *Sawai Tomeno Tokachi Hombetsu bunrui Ainugo jiten* 『沢井トメノ十勝本別アイヌ語分類辞典』. Hombetsu: Hombetsu Board of Education Hokkaido Government Office 本別町教育委員会.

Schroepfer, Jason (2015) Ethnic variation of */tˤ/ in Aswan Arabic. *Selected Papers from New Ways of Analyzing Variation* 22(2): 151–160. URI: https://repository.upenn.edu/pwpl/vol22/iss2/17

Shiraishi, Hidetoshi (2010) *Topics in Nivkh phonology. A description and analysis of the phonological system of Nivkh*. Chișinău: VDM Verlag Dr, Mueller.

Shiraishi, Hidetoshi 白石英才 and Itsuji Tangiku 丹菊逸治 (2015) Nivkh-go Amur hougen no kisogoi 3 ニヴフ語アムール方言の基礎語彙3. *Hoppou Gengo Kenkyu* 5: 215–226. URI: http://hdl.handle.net/2115/58361

Sidwell, Paul and Felix Rau (2015) Austroasiatic comparative-historical reconstruction: an overview. In Mathias Jenny and Paul Sidwell (eds) *The handbook of Austroasiatic languages*, 221–363. Leiden, Boston: Brill. doi: https://doi.org/10.1163/9789004283572_005

Silva, David J. (2006) Acoustic evidence for the emergence of tonal contrast in contemporary Korean. *Phonology* 23(2): 287–308. doi: https://doi.org/10.1017/S0952675706000911

Sinha, Kali Prasad (1981) *The Bishnupriya Manipuri language*. Calcutta: Firma KLM.

Skorik, P. Ya. [Скорик П. Я.] (1961) *Грамматика чукотского языка. Часть первая. Фонетика и морфология именных частей речи*. Москва: Издательство академии наук СССР. М-Л.

STEDT = The Sino-Tibetan etymological dictionary and thesaurus. Database. Online: http://stedt.berkeley.edu/search/

Stilo, Donald (2019) The Caspian region and south Azerbaijan: Caspian and Tatic. In Geoffrey Haig and Geoffrey Khan (eds.) *The languages and linguistics of western Asia: An areal perspective*, 659–823. Berlin: De Gruyter Mouton. doi: https://doi.org/10.1515/9783110421682-019

Sulkala, Helena and Merja Karjalainen (1992) *Finnish*. London and New York: Routledge.

Suzuki, Hiroyuki 鈴木博之 (2011a) Gagatang Zangyu de yanhua yuanyin ji qi laiyuan 嘎嘎塘藏語的咽化元音與其來源. *Language and Linguistics* 12(2): 477–500.

Suzuki, Hiroyuki 鈴木博之 (2011b) Tibettogo syohoogen ni mitomerareru kootoo tokutyoo no tayoosei チベット語諸方言に認められる喉頭特徴の多様性. *Journal of the Phonetic Society of Japan* 15(2): 52–60. doi: https://doi.org/10.24467/onseikenkyu.15.2_52

Suzuki, Hiroyuki 铃木博之 (2013) Yunnan Weixi Zangyu de r jieyin yuyin yanbian: Jiantan "erhua" yu "jinhou" zhi jiaocha guanxi 云南维西藏语的r介音语音演变：兼谈"儿化"与"紧喉"之交叉关系. *Dongfang Yuyanxue* 13: 20–35.

Suzuki, Hiroyuki 鈴木博之 (2015) h̃ to ɦ̃: Bikuu kyoomei wo tomonau seimon masatuon ni kansuru oboegaki h̃とɦ̃：鼻腔共鳴を伴う声門摩擦音に関する覚え書き. *Journal of RIHN Descriptive Linguistic Group* 7: 141–149. URI: http://id.nii.ac.jp/1422/00000875/

Tamura, Suzuko 田村すず子 (1996) *Ainugo Saru hōgen jiten*『アイヌ語沙流方言辞典』. Tokyo: Sōfūkan 草風館.

Tamura, Suzuko (2000) *The Ainu language*. Tokyo: Sanseido.

Tomita, Takejiro (ed.) 冨田竹二郎（編）(1997) *Tai-Nichi dai jiten*『タイ日大辞典』. Tokyo: Mekhong.

Tournadre, Nicolas and Hiroyuki Suzuki (2022) *The Tibetic languages: An introduction to the family of languages derived from Old Tibetan* (with the collaboration of Xavier Becker and Alain Brucelle for the cartography). Villejuif: LACITO Publications (CNRS).

Tryon, Darrell T. (eds.) (1995) *Comparative Austronesian dictionary*. Berlin and New York: Mouton de Gruyter. doi: https://doi.org/10.1515/9783110884012

Tschenkéli, Kita (1965-1974) *Georgisch-Deutsches Wörterbuch* I-III. Nach dem Tode des Verfatters fortgeführt von Yolanda Marchev unter Mitwirkung von Lea Flury, Ruth Neukomm und Victor Nosadzé. Zürich: Amirani-Verlag.

Turner, Ralph Lilley (1966[1985]) *A comparative dictionary of the Indo-Aryan languages*. London: Oxford University Press.

Učida, Norihiko (1970) *Der Bengali-Dialekt von Chittagong: Grammatik, Texte, Wörterbuch*. Wiesbaden: Otto Harrassowitz.

Wang, Fengyang 王凤阳 (2011) *Gu ci bian (zengding ben)*《古辞辨（增订本）》. Beijing: Zhonghua Book Company.

Watanabe, Hitoshi 渡辺仁 et al. (1984) *Ainu minzoku bunkazai chōsa hōkokusho (Ainu minzoku chōsa III Shizunai chihō)* アイヌ民俗文化財調査報告書（アイヌ民俗調査III 静内地方）. Sapporo: Hokkaido Government Office 北海道教育委員会.

Watanabe, Hitoshi 渡辺仁 et al. (1991) *Ainu minzoku bunkazai chōsa hōkokusho (Ainu minzoku chōsa X Chitose)* アイヌ民俗文化財調査報告書（アイヌ民俗調査X 千歳）. Sapporo: Hokkaido Government Office 北海道教育委員会.

Watanabe, Hitoshi 渡辺仁 et al. (1995) *Ainu minzoku bunkazai chōsa hōkokusho* アイヌ民俗文化財調査報告書（アイヌ民俗調査XIV 補足調査1）. Sapporo: Hokkaido Government Office 北海道教育委員会.

Wheat Flour Institute (1976) *From wheat to flour: The story of man in a grain of wheat*. Washington, D.C.: Wheat Flour Institute.

Wolff, John U. (2010) *Proto-Austronesian phonology with glossary*. Southeast Asia Program Publications. Vol. 1. Ithaca, New York: Cornell University. doi: https://doi.org/10.7591/9781501735981

Yoneda, Yūko 米田優子 (1995) Problems in the use of oral tradition in research of Ainu agricultural history: An examination of legends about the origins of grains アイヌ農耕史研究にみられる伝承資料利用の問題点－穀物の起源説話に関する検討を中心に－. *Bulletin of the Hokkaido Ainu Culture Research Center* 1: 1–25. Online: https://ainu-center.hm.pref.hokkaido.lg.jp/kankou/kiyou/pdf/kiyou01-01.pdf

Zhang, Huiying 张惠英 (2009) *Chongming fangyan yanjiu*《崇明方言研究》. Beijing: Zhongguo Shehui Kexue Chubanshe.

Zhu, Xiaonong 朱晓农 (2010) *Yuyinxue*《语音学》. Beijing: Shangwu Yinshuguan.

Zhukova, A. N. [Жукова А. Н.] (1967) *Русско-корякский словарь*. Москва: Сов. Энциклопедия.

Linguistic atlas of Asia and Africa
Cumulative index of chapters