Table of Contents

Mitsuaki ENDO, Goals of the Project “Studies in Asian Geolinguistics” 2015-2017  .................. 1

“Sun” in Asian Languages

Ryo MATSUMOTO, “Sun” in Tungusic and Uralic ................................................................. 5
Yoshio SAITO, “Sun” in Mongolic and Turkic ................................................................. 6
Mika FUKAZAWA, “Sun” in Ainu ................................................................. 8
Takashi UEYA and Kenji YAGI, “Sun” in Sinitic languages ............................................. 10
Yoshihisa TAGUCHI, “Sun” in Hmong-Mien ................................................................. 12
Satoko SHIRAI, Keita KURABE, Kazue IWASA, Hiroyuki SUZUKI, and Shiho EBIHARA,
  Sun: Tibeto-Burman ................................................................. 14
Mitsuaki ENDO, “Sun” in Tai-Kadai ................................................................. 18
Mika KONDO, “Sun” in Austroasiatic ................................................................. 20
Noboru YOSHIOKA, Sun: South Asia ................................................................. 22
Yoichi NAGATO, The Sun: Arabic Languages ................................................................. 24
References ................................................................. 26

Articles

Shinsuke KISHIE, Yukichi SHIMIZU, and Yukako SAKOGUCHI, Dialectal Forms Associated with
  the Word Taiyō (Sun) in Japanese ................................................................. 37
Mika FUKAZAWA, Geographical distribution of ‘daytime’ in Ainu .................................... 44
Rei FUKUI, The sun in Korean ................................................................. 55
Liying QIN and Hiroyuki SUZUKI, Chasing a Cat from the Mekong to the Salween: A Geolinguistic
  Description of ‘Cat’ in Trung and Khams Tibetan in North-western Yunnan ..................... 61
Dawa Drolma and Hiroyuki SUZUKI, Preliminary Report on the Darmdo Minyag Linguistic Area,
  with a Geolinguistic Description of Terms for ‘Sun’ ................................................................. 72
Hiroyuki SUZUKI, A Geolinguistic Description of Terms for ‘Sun’ in Tibetic Languages of the
  Eastern Tibetosphere ................................................................. 79
Makoto MINEGISHI, Toshiki OSADA, Nathan BADENOCH, Masaaki SHIMIZU, Atsushi YAMADA,
  and Yuma ITO, A Survey of Recent Austroasiatic Studies ................................................................. 86
Goals of the Project “Studies in Asian Geolinguistics” 2015-2017

Mitsuaki Endo
College of Economics, Aoyama Gakuin University

1 Introduction

The goal of this study is to publish The Linguistic Atlas of Asia, Vol. 1 in the academic year 2018. It comprises 9 linguistic properties: 1) the Sun, 2) rice, 3) milk, 4) wind, 5) iron, 6) means to count noun, 7) tonal phenomena including tone, pitch accent and stress, 8) "it rains", and 9) types of manner of articulation of consonant. Research meetings will be held thrice a year, 9 times in total. Each meeting will be devoted to one theme above. It mainly deals with vocabulary, but phonology, morphology, and syntax are also included. Regarding the density, over 1,000 locations of all language families in Asia will be included. Furthermore, the geographical distribution will be interpreted from the viewpoint of linguistic history, language contact, and the transmission of matters as well as the migration of human groups.

The tool for mapping is Arc GIS online. Thanks to the great efforts by Mr. Yukichi Shimizu, an additional set of symbols suitable for geolinguistics first designed by Ms. Yuko Okada is open to the public with colorful symbols. Language families are distinguished by colors in order to avoid confusion between one another.

In the following chapters, some details of the project are described.

2 Discussion of linguistic properties

2.1 The Sun

The first item of the Linguistic Atlas of Asia is "the Sun". The reason is that a natural phenomenon and easy to understand. Many linguistic atlases begin with this word, including Atlas Linguarum Europae (Vol. I, fasc. 1, Assen: Van Gorcum, 1983, and following volumes) which we intend to connect with in the future, Zhiyun Cao (ed.) Linguistic Atlas of Chinese Dialects (Beijing: the Commersial Press, 2008), and among others, Ray Iwata (ed.) The Interpretative Maps of Chinese Dialects (Vol.1, Tokyo: Hakuteisha, 2009; Vol. 2, Tokyo: Kobun Shuppan, 2012). The last one is especially important for us, since this project is a follow-up and expansion of the activities on Chinese geolinguistics lead by Professor Ray Iwata from 1989 on. Linguistic Atlas of Asia will be compiled according to this model.

An important difference is that each Asian language family is discussed in two pages, and an overview of all of Asia is also given for each item.

Such a basic vocabulary may show no lexical varieties in a particular language, like Ainu or Korean, and so on. Still phonological differences can be traced, and even a uniform distribution is meaningful in a wider context surrounding that language.

2.2 Rice

The importance of rice in daily life is obvious in Southeast and East Asia. In North Asia, rice isn't grown, and therefore no word exists at all in several languages. Again, this lack is also an interesting fact, and its geographical distribution deserves to be traced. The spreading of rice and its culture along with the words which denote them should be closely related. Collaboration with the specialists of rice is also desirable. This item is chosen as a specimen of agricultural vocabulary.
2.3 Milk
Stock raising is also a basic industry especially in Central Asia. Milk is chosen as a representative word in this area. One and the same word is used to denote "milk" and "breast (of mother)" in some languages like Japanese and Chinese. It is a natural semantic derivation process, similar to "the Sun" and "day". Such a phenomenon will also be dealt with.

2.4 Wind
See Saito (2013) and Endo (2014). Although this is a basic vocabulary item, similar forms are wide spread among languages in Eastern Asia. Further scrutiny across the whole Asia might be interesting.

2.5 Iron
Production of iron uses coal and high temperature technology; hence it was a great innovation dating around 2 thousand years ago. Its invention contributed to the rise of production using solid machines not only in manufacturing industry but also in agriculture. Chang (1972) reconstructed the proto Sino-Tibetan form for 'iron' as *qhleks. Hlek in modern Siamese is a reflection of this form, and tiê in Chinese as well as tetsu in Japanese have the same origin despite the difference in appearance.

2.6 Means to Count Noun
Classifiers are used to count noun in Southeast and East Asian languages. For examples, "these three books" in Chinese is expressed as follows: zhè (this) sān (three) běn (classifier) shū (book). Word order of these items differs according to various languages in Southeast and East Asia. In other areas in Asia, no classifiers are needed. Still, the simpleness of these languages is also significant to draw a map.

2.7 Tonal Phenomena
Tone is a shared feature among Southeast and East Asian languages. Synchronously and diachronically speaking, there are abundant variations. In addition, there are pitch accent languages like Japanese and some dialects of Korean, and stress accent languages in Altaic. A system with a fixed position stress also occurs in a wider geographical context. This is a very complicated feature, so about double the space for this topic is considered in the book.

2.8 It rains
In Chinese, the word order of "it rains" is VS, although the basic word order is SVO. In Tai-Kadai, SV is dominant, but a VS type is also observed in neighbouring languages with Chinese. In the oracle-bone inscription age in Chinese, "rain" can serve as a noun as well as a verb by itself as in English. In Nivkh, there is no word to express "rain."

2.9 Types of manner of articulation of consonant
Voicedness and aspiration are dominant distinctions among Asian languages. Prenasalised consonants are also important issues for many languages.

3 Covered Language Families and Languages
All language families and isolated languages in Asia should be included in this Atlas. Fortunately, almost all branches have been covered by specialists:

- Paleo-Aisiatic: Tokusu Kurebito
- Nivkh: Hidetoshi Shiraishi
- Ainu: Mika Fukazawa
- Uralic: Ryo Matsumoto
- Tungusic: Ryo Matsumoto
- Mongolic: Yoshio Saito
4 Measures to Spread the Research Results to the Public

The last meeting in February or March 2018 will be held jointly as the 4th International Conference of Asian Geolinguistics at ILCAA, TUFS. At the same time, a symposium regarding the relationship and formation processes between language distribution and cultural and natural geography will be held with geneticists, archaeologists, historians, and geographers focusing on Asia.

Nine issues of e-publication *Studies in Asian Geolinguistics* in English are open to the public in the world through the web site. The contents consist of the draft version of each item to be included in the Atlas, more detailed discussions and supplementary materials, as well as original papers by members. A printed version including 3 issues each will be circulated annually.

Longitude and latitude data of each place, lists of source materials, and raw linguistic data will be open to the public on the internet after the end of the project in order to promote future studies by broader range of researchers.

Linguistic Atlas of Asia, Vol. 1 will be published hopefully by ILCAA, TUFS, if the publication proposal is accepted. The maps will be open online through Arc GIS Online.

5 Prospective Research Results

1) Orthodox research with classic models has been started in the areas where geolinguistics are still not developed. Further developments are promoted more systematically in the areas where dialect geography has already begun.

2) Descriptions on dialects of each language family are overviewed. The last decades witnessed a huge progress in many Asian countries, and this should be traced thoroughly. Professor Makoto Minegishi tried to do a model survey on Austroasiatic languages this time. As a result, a lack of materials was revealed, and the necessity of further descriptive studies on these dialects was clarified.

3) Comparative linguistics of each language family is also essential for geolinguistic studies; otherwise it is impossible to identify whether some forms are cognate or not. Moreover, exceptional changes due to folk eymology or analogy and so on can be detected only using a strict comparative method. Philological studies are also useful to interpret geographical distribution of linguistic properties. Reversely, linguistic maps can provide positive evidence to determine the originating area of a linguistic feature in an old document.

4) Maps based on a super-macro perspective among languages across the whole of Asia with micro and dense places enable us to find out changing and spreading mechanisms inside a language or a language family as well as facts with concrete instances caused by language contacts. Natural, easy-to-emerge or the opposite changing processes will be detected with a vast range of examples with over thousands of places, and universal patterns will be extracted empirically.
5) Linguistic facts are compared with extra-linguistic factors such as migration history, spreading processes of things, transportation and land shapes. Interdisciplinary and synthetic studies on human groups are expected based on geographical distribution of particular linguistic features.

6 Prospective Biproducts

1) Synergy effects are expected from the close collaboration among domestic specialists of various languages in Asia. They will be cover all the generations in the country.

2) Closer networks will be formed and strengthened with the researchers of geolinguistics, dialectology and historical linguistics in Asia as well as other areas in the world. We can stimulate the autonomous geolinguistic studies in the other countries in Asia. A consortium of geolinguistic studies in Asia is also expected to be formed in the near future, or it is already in the process of formation in the form of the International Conference of Asian Geolinguistics first held in Japan in 2012.

3) Manuals and guides to draw linguistic maps using Arc GIS Online is provided to the public. It will help spread the method of drawing linguistic maps anywhere in the world. Moreover, after providing the longitude and latitude information of places in the Linguistic Atlas of Asia, as well as lists of source materials, it will be far easier for anyone to draw linguistic maps in Asia.

References


Acknowledgements

I appreciate Dr. Donna Erickson for checking my English.
"Sun" in Tungusic and Uralic

**Tungusic**

Many minority peoples inhabit Siberia. Tungusic languages have three words for *sun*: *dilača*, *šigun* and *jūlten*. Evenki dialects have all these three forms, while Ewen has one form *jūlten*, and the other tungusic languages have one form *šigun*. The word *jūlten* is derived from the root *jū-* ~ *ǰū-~* *ňū-* “to go out, to appear, to rise”, and the meaning is “the place from where the sun rises”. From the majority distribution and the word formation, one could argue that the form *šigun* dates back to Proto-Tugusic.

Map 1 shows the overlapping distribution of the three forms. Some dialects use two forms concurrently. According to the Map 1, in the central area of Evenki (the south-west Yakutia) only *dilača* is found, while *šigun* is located in the peripheral areas. On ther other hand, in Evenki the form *jūlten* is found only in the east, where Evenki has contact with Ewen. From the point of view of the theory of language form propagation, the form *šigun* could be older than the others.

There remains the etymological problem of the word *dilača*, as Kazama (2003: 46) referred to it as “a rather unique vocabulary item”. According to Tsintsius (1975: vol.1, 206), the cognates of *dilača* are found also in Solon *dečevā* and Negidal *dulana*, which mean “sun” in both. However its origin is unknown. I propose here one possibility that they might have the same root with *dil* “head” as a sequel to the personification of sun, as well in Baltic-Finnic: Fin. *päivä* “day, sun”, *pää* “head”.

**Uralic**

In Uralic there seem to be several words for *sun*, but phonologically it could be divided in three: the vowel of the first syllable is A) back vowel, B) front vowel, or C) the first consonant is labial. Its geographical distribution is also clearly divided in three in same way. As for the first consonant, type A and B have a common character – velar obstruent, but the different change – palatalization, sibilantization – had happened due to the following vowel. Type A is found in Samoyedic, where velar sound is retained. Forest Nenets *dālha* and Selkup *cēl* are the exceptional form, but these words are the cognate to “day”; ex. Tundra Nenets *jall’a* (cf. *jālā* “Licht, Tag, Sonne”, *kājā* “Sonne”: Janhunen 1977). On the other hand type B is found around the Ural Mountains, and the affected consonants is *č*, *š*, *sʲ* or *x*. Type C is found only in the limited Baltic-Finnic area. In this area the forms are very similar each other, *p*-initial words possibly appeared after the split from the other Uralic languages.

The exceptional forms in Uralic are found in Finnish *aurinko*, and in Hungarian *nap*. There words’ origins are still unclear (In Finnish *päivä* has the meaning “sun”, but mainly used for “day”).

**“sun” and “day”**

As already mentioned above, sometimes in some languages the words for “sun” and “day” are not distinguished clearly. Tungusic and Samojed languages consistently have different words: Evenki *dilača* “sun”, *injni* “day” : Tundra Nenets *xajer?* “sun”, *jall’a* “day”.

(Ryo Matsumoto)

---

**Map 1 “Sun” in Tungusic**
- (red) *dilača*
- (blue) *šigun*
- (green) *jūlten*.

**Map 2 “Sun” in Uralic**
- (yellow) A *k + back vowel*
- (Blue) B *k + front vowel*
- (red) C labial *p-initial*
- (black) others
“Sun” in Mongolic and Turkic

The Mongolic and Turkic languages have been in continuous contact with each other over long periods of time and thus share a certain amount of vocabulary such as tämür/tämir “iron,” bütün/bütün “whole, entire,” bajan/baj “rich,” and so on. But, they use completely different words for “the sun.”

1. Mongolic

The word forms representing “the sun” in Mongolic derived from naran. Some languages lost the word-final nasal in the [−human] words in the nominative case while other languages preserved it. In terms of this characteristic, the forms can be classified into the following two categories:

A) Forms retaining the word-final nasal in the nominative: naran, narn, narn, nar.
   E.g. Buriad, Dongxiang, Bonan, Moghol, Shira Yughur, Kalmyk.

B) Forms without the word-final nasal in the nominative: nara, naro, nar.
   E.g. Khalkha, Ordos, Dagur, Monguor.

The map shows that the forms retaining the word-final nasal appear in marginal zones of the Mongolic-speaking area. This distribution suggests that they are older than those with no word-final nasal, and agrees with the data from written documents. (The Kalmyk language in the lower Volga region should be treated separately because its speakers migrated there from the present-day Xinjiang region in the 17th century.)

2. Turkic

The word forms representing “the sun” in Turkic can be classified into the following two types:

A) kün-type

B) küneş-type, including the form kujaš, etc.

Voicing of the initial consonant has taken place in the Southwestern group, which includes Turkish, Azeri, Turkmen, etc., producing the forms gün and güneş. Vowel changes have also taken place in some languages. (E.g. Khalaj has both kün and kın.)

The form kün, which originally meant “the sun,” has extended its meaning to refer also to “day.” This phenomenon is “common in all periods and all languages.” (Clauson 1972: 725)

Chulym developed a compound word kün karayi with the word karak “eye.”

The form kujaš, which meant “sunny place” or “sun heat,” existed in Old Uighur. It had developed from *kuñaš, which is preserved in modern Sakha and Dolgan as kuñas “warm weather.” (Tenishev 2001: 65, Erdal 2004: 72)

The Chuvash form xavel corresponds to the form küneş, etc. in the other Turkic languages. (The key sound correspondence is ʃ : s)

The type B word is used (almost?) exclusively for “the sun” in Turkish, Crimean Tatar, and Chuvash, while the type A form is used for both “the sun” and “day” in the other languages.

In Turkish, the word for “the sun” is güneş. The word gün is used for “day,” and is only used in the meaning “the sun” in such fixed phrases as gün ağlı “halo around the sun,” güne bakan “sunflower” (bak- “to face”), gün batması “sunset” (bat- “to sink”), gün doğ- “(of the sun) to rise” (doğ- “to be born”), etc.

In Karaim, kujaš/kujaš is used for “the sun,” while kün is used both for “the sun” and “day.” kujaš anıktanred’ batma “the sun has already set,” kujaš balkuvu “sunshine,” kujaš sagatlar “sundial”; kün batşi “sunset,” birsi kün “the day before yesterday.”

In Tuvan, xün is used both for “the sun” and for “day,” and the two meanings are mainly distinguished by context as they are in most other Turkic languages. E.g. xün ünüp kelgen “The sun rose,” töriüttingen xün “birthday.” But in Tuvan there are some cases where the two meanings, for example as in ol xün “that sun / that day,” are distinguished by difference in pitch and by the presence and absence of pause between the two elements. (A. Syuryun, personal communication)

In addition to the above two types of words, there is a word jašik (< jašu- “bright”), which was used for “the sun” in Karakhanid in the 11th century and survives in only a few modern languages with its original meaning, e.g. Turkish işik “bright.”

We can interpret the modern distribution of the Turkic words as the result of a limited spread of the use of the type B word for “the sun” from the Black Sea region.

Keywords: day, sunny place, sun heat, warm weather, bright, eye

(Yoshio Saitô)
“sun” in Mongolic

“sun” in Turkic
“Sun” in Ainu

In Ainu, the term cúp (c=[ʣ, ʤ, ʦ, ʧ]) stands for both “sun” and “moon.” It is distributed over three large dialectal groups of Hokkaido, Sakhalin and northern Kuriles, where “●” is denoted in Map 1. The form of cúp in Sakhalin, “●”, originates from cúp in Hokkaido because the most of the Sakhalin dialects have a special phonemic variant /-h/ [h], which is substituted for the coda /-p, -t, -k, -r/ [p, t, k, r] (cf. Chiri 1942: 471–472; Tamura 2000: 20). When specifically referring to the sun, we have to add the word for “daytime” before cup. Then, the terms for “sun” can be classified into 3 types: the A) toono type, B) sirpeker type and C) tôkap type.

The lexical forms of Type A are distributed over Sakhalin and the northernmost part of Hokkaido and the modifier tôno/toono would be the oldest form for “daytime.” The monosyllabic forms in Type A and C, too and tô, mean “day (24 hours)” in Ainu, since the long vowel oo [o:] in Sakhalin is phonetically equivalent to the pitch accent ó in Hokkaido. The terms tôno and toono would be interpreted as “the middle of the day” if the following morpheme no was related to the locative noun nósí “the middle of.”

In Type B, sirpeker originally means “day breaks; it gets light,” consisting of the prefix sir- “sight; view” and the intransitive verb pekér “to be light.” The dialects of eastern Hokkaido and northern Kuriles show this type. In some dialects, the final consonant /-t/ is changed to /-t/ before the consonant /c/. The southernmost dialect of Hokkaido, Samani, uses the short form peker.

Western Hokkaido dialects are categorized into Type C. In C-1, the term for “daytime” tôkap consists of tô “day” and káp “skin” (uncertain). In C-2, the coda /-p/ [p] of káp is changed to /-m/ [m], possibly by analogy with the meaning of kám “meat.” The history of this type is unclear. The word tôkap for “daytime” is considered as a homonym for “older woman’s breast” or a polyseme (for more details see Fukazawa 2015).

In addition to the three types above, the rikóma type is often used as a reverential form in a specific context. When people pray to the god of the sun, and sometimes the god of the moon, they are called rikoma kamuy in the Shizunai dialect, and rikómá tonpi in the Yakumo dialect. In Raichisika, one of the Sakhalin dialects, rikomah (≪ rikómap, as in Biratori) is recorded as a special term that older people use. In the text of yūkar, the heroic epics in Ainu, rikóma cup is used for “the sun” in expressions like “He has a face like the sun.” The modifier rikóma is interpretable as rik “sky,” omá “to put O₁ (in) O₂,” and then rikoma kamuy means “the god in the sky.”

The head of the noun phrases for “sun” shows the following variations: a) cúp (kamuy)” (the gods of) the sun and/or the moon,” b) tônpi “the light (of the star),” and c) -p, the classifier for “thing,” which is seen in rikómap and rikomah. A standard form is a) cúp, and the following kamuy (y=][j]), “the god,” is optional except for Type A. In the Ainu culture, the sun and the moon are regarded as important gods; however, unlike the gods of the mountain, village, water etc., these gods are not necessarily worshipped in a visible manner (Uchiyama 2008). Their gender and relationship differ by region: they are either a married couple or siblings (Kitahara 2014: 14; Watanabe et al. 1992: 72–73).

The form of b) tonpi occurs in limited areas. In the dialect of Tarantomari, Sakhalin, cup is replaced by tônpi, which stands for “sun,” “moon” and even for “month,” “●” in Map 1. The lexical form of tônpi could be analyzed into tom “to sparkle; to shine” and -pi (< pe) classifier of “thing” (uncertain) (cf. Tamura 1996: 580).

Since at least the middle of the nineteenth century, the terms tô and too, “day,” have been metonymically used for the meaning of “day of the month,” and cúp, for “month,” in the same way as -kai-nichi ❈ stands for “day of the month” and -gatsu ❈, for “month” in Japanese.

Keywords: daytime, the moon, the god of the sun

(Mika Fukazawa)
“Sun” in Ainu

A. toono type
   A-1 toono cup type
   A-2 toono tonpi

B. sírpeker type
   B-1 sírpeker cup (kamuy) type
   B-2 peker cup kamuy

C. tókap type
   C-1 tókap cup (kamuy)
   C-2 tókam cup (kamuy)
"Sun" in Sinitic languages

1. Classification of word forms

With respect to the Chinese language atlas, there are two main previous studies, Cao2008 and Iwata2009. These atlases both include maps of Sun-related terms. The maps of Sun-terms contained in Iwata2009 are made by Takashi Matsue. The distribution of the main forms of terms for "the sun" in China has already been explained in Matsue2009. This map was made to emphasize the other forms of Sun-terms.

In this map, word forms representing "the sun" in Chinese are classified into 9 categories.

A. ri 日 type
   A1 ri 日[zit]
   A2 ritou 日頭[Z0, t'ou], retou 熱頭[nie? t'eu], yuetou 月頭[nue t'eu]

B. taiyang 太陽 type
   太陽[t'ai ian], taiying 太影[t'ie in]

C. Bw- type
   C1 yangpo 陽婆[ia p'ei] [ie p'ei]
   C2 tayiyang bao 太陽包[t'ai ian pau]
   C3 ritou pu 日頭婆[ni? dav bu] [nie? do va], tayiyang bo 太陽婆[t'a ji vai]
   C4 pusha ye ye 菩薩㖶婆[p'u sa i ie]
   C5 ritou pusa 日頭菩薩[ni? dav bu sw?], tayiyang pusa 太陽菩薩[t'a fia in bu sa?]

D. ye 爺 ("grandfather") type
   爺[ie] 爺[ie ie], laoye 老爺[lau ior], ritou ye 日頭 爺[ri ou i]

E. wo 烏 "burrow" type
   ye woor 爺窩兒[ia uor], yangpo wozi 陽婆窩子[iaan p'x iui uo sa?]

F. yan 眼 "eye" type
   ritouyan 日頭眼[Z0, t'ou niir], tayiyang yan 太陽眼[t'a ian niir], ye ye yan 爺眼眼[ia ia niir]

G. Kw- type ("sun" + /k/-, /k'/-) type
   G1 [kong] "male, grandfather" ritou gong 日頭公 [zik t'au kon], tayiyang gonggong 太陽公公[t'a fia on kon]
   G2 [kön] "hole" ritou kong 日頭孔[nia? tui k'ou] (type A2)
   G3 [ku] "cow" ritou gū 日頭咕[ni t'ak u] etc.
   G4 [k'ou] "burrow" ritou ke 日頭窟[ni div k'ou]

H. Taiyang di 太陽帝 ("emperor") type
   tiandi 天地[t'ai ian tier], taimian 太陽帝[t'ai ian ti]

I. others: nair 奶兒[ner], nair nair 奶兒奶兒[ner ner]

2. Geographical distribution and interpretation

To put it shortly, word forms representing "the sun" in Chinese can classified into two main categories - ri 日 type(Type A) and taiyang 太陽 type(Type B). It is estimated that the oldest form may be the monosyllabic form ri 日(A1). But its distribution is limited to a certain part of southeast China. The main form of ri 日 type is "ri 日+ou 頭(suffix)"(A2). It includes retou 熱頭("hot"+suffix) and yuetou 月頭 ("moon"+suffix). Both of them are forms derived from ritou 日頭. (Matsue2009:64-65)

ri 日 type and taiyang 太陽 type are coexistent in many areas. As an overall tendency, ri 日 type is seen throughout all of China. On the other hand, taiyang 太陽 type is very rare in southern China.

Bw- type (Type C) is the form including /bw/, /bw/ is Bilabial Plosive + Rounded vowel. It tends to change to unvoiced-aspirated consonants or unvoiced aspirated consonants /pʰw-//pw-/ derived from voiced consonants. Yangpo 陽婆(C1) is seen in Shaxi area. /bw/ in Shaxi represents "old woman (Po 婆). Ritou pu 日頭婆(C3) and ritou pusa 日頭菩薩(C5) are distributed in two separate areas, Zhejiang and Sichuan. In these areas /bw/ represents the "Buddhist saint (Pusa 菩薩)". In southern dialect, jinri 今日 "today" is called jinbu 今晡[teen pu]. It may suggest that they are related to each other somehow. The form ritou bao 日頭寶(C2) is distributed in Sichuan and Guangdong.

In type D, Ye 爺 is the morpheme meaning the kinship term "grandfather". ye 爺 can be seen as the form that omitted the "sun" morpheme of 日頭爺 (Matshue2009). However, the first syllable of C1 yangpo 陽婆 has largely lost its original nasal coda -ŋ, then the pronunciation of yang 陽 and ye 爺 can be identical in these areas.

wo 窩(type E), yan 眼(type F), kong 孔(G2), ke 窟(G4) all have forms which are related with "hole", therefore folk-etymology that looks upon "sun" as "hole in the sky" probably influenced the formation of these types. ritou kong 日頭孔(G2) shows a scatter distribution in Anhui, Zhejiang, Shaxi, Hebei. Ritou kong 日頭孔(G1) is densely distributed in Anhui and Zhejiang. Kong 孔 (k'ong) represents "hole". In its neighboring areas, ritou gong 日頭公(G3) is distributed rather broadly. Gong 公 means "male, grandfather ". We can hypothesize that the form ritou kong 日頭孔 was derived from ritou gong 日頭公. (/kong/>/k'ong/). (Takashi Ueya, Kenji Yagi)
"Sun" in Sinitic

A1 日zhī/nǐ?
A2 日類jet d'ěu/ze tū
B 太陽tāi liān
C1 陽塞iā prā/pra prā
C2 太陽包tāi liān pau
C3 日頭看niā dão bù/太陽佛tāi jī vāi
C4 聲嘈jīnghuáng sì le le
C5 日頭若hēi? dào bù sè?
/太陽音随tāi liān phu sa

D 聲heī/老鸽鷄lǎo gūié
E 鳥nuò鴿鷄lǎo 乌乌乌 ēr ēr ēr
F 日頭眼te tāi/太陽眼tāi tāi
G1 日頭公zhī tōun kǒnɡ/太陽公tāi liān kǒnɡ kǒnɡ
G2 日頭kǔ? diwū k'ǒnɡ
G3 日頭看tāe kū
G4 日頭音ni dǐ k'ǔ
H 太陽帝tāi liān tī/天地tī tī tī tier
I others
“Sun” in Hmong-Mien

1. Classification of word forms
Most languages belonging to Hmong-Mien exhibit a word-form that has a root with a nasal onset. Based on the comparative work in Hmong-Mien linguistics, this root goes back to a Proto-Hmong-Mien root that has a voiceless nasal onset and Tone 1. We represent all the forms cognate with this proto-form as setQuery, and call the languages that have setQuery as the root for “the sun” Type A languages (the number on each word-form represent historical tone). In these languages, the form setQuery is also used as the word-form for “day”.

The Type A languages are divided into several subgroups. First, the most numerous subgroup A-1 uses setQuery as a monosyllabic word-form. Other languages in this group can be divided according to the element that is added to the root. Subgroups A-2 to A-19 are languages where the additional element is placed before the root, and subgroups A-20 and A-21 are languages where the additional element is placed after the root.

A-1: setQuery as a monosyllabic word
The element before the root in subgroups A-2 to A-19 is as follows.
- A-3: qa7/8
- A-4:  qed/ qa1
- A-5: ha8
- A-6: ci5
- A-7: khu4
- A-8: qhat3 “hole”
- A-9: ps7
- A-10: qa0
- A-11 po3 “father”
- A-12: vei3 “wife”
- A-13: aɔ 51
- A-14: a2/ ja2 qa1d “moon” (first two syllables)
- A-15: nan1
- A-16: la2
- A-17: nu2
- A-18: mvo1
- A-19: a1/4/7

The element after the root in subgroups A-20 and A-21 is as follows.
- A-20: ka3
- A-21: ntor2 “sky”

The remaining forms are divided into four types.

B: root is nan or ntor6

C: tau2 ŋkου1
D: qa3/8 ha1
E: ŋa4/6/8 tau2 or ni2 tau2

2. Geographical distribution and interpretation
Type A-1, which is monosyllabic, has the widest geographical distribution. Since this form also exhibits the widest distribution in terms of the phylogenetic tree of the Hmong-Mien languages, it can be interpreted as preservation of the oldest state among the A type forms. The other forms with some additional element must have been formed in each language/dialect. Type A-2, which has a bilabial element before the root, is concentrated in Guangxi and Vietnam, but this is a feature of two branches of Mienic, Iu-mien and Kim-mun. The root of the Type A forms is considered as a loan from some Tibeto-Burman language (Benedict 1987, Ratliff 2010:235).

The other types than Type A indicate a limited distribution. The etymology of Type C and Type D is unclear. The forms of Type E seem to be a loan from Chinese “日” or “阳”. What might attract interest is Type B. It shows periphrastic distribution both in terms of geography and phylogeny. The Type B forms (nan and ntor) are distributed in one of the North Hmong languages in Hunan and Jiong-nai in Guangxi (based on Rafliff 2010, the proto-form for these forms can be reconstructed as *ndang). Furthermore, the reflex of this proto-form has a meaning “sunshine” in some languages including one of the other North Hmong (Xiang 1992) and Pa-na, both of which have a Type A form for “the sun”. This fact suggests two possible stories. Based on the assumption that the Type A forms derive from a loanword, the Type B forms might be the retention of the original Hmong-Mien form for “sun”. After borrowing the root of Type A, the Type B forms shifted the meaning to “sunshine” through a metonymic semantic change (or narrowed the original meaning). The other story is that the original word for “sunshine” has acquired a meaning of “sun” in the Type B languages. The issue is still to be settled.
"Sun" in Hmong-Mien

A-1: ŋV¹ as a monosyllabic word
A-15: nャ

Polysyllabic word with an element before root:
A-2: pu²/ bu²/ pa²/ pa³/ ba²/ pe³
A-16: la³
A-17: nui²
A-18: mwα¹
A-19: a¹/³/⁷

A-3: qha⁷/⁸
A-4: ʔe³/ʔa¹
A-5: ha⁸

Polysyllabic word with an element after the root:
A-6: ci³
A-10: qo⁸
A-11: po³ “father”
A-12: vei³ “wife”
A-13: a⁶ ³¹
A-14: a² la³ qa³⁴ “moon”
1. Classification of word forms

In Tibeto-Burman (TB), the most common word form of “sun” is with an initial nasal (n/q/ŋ). However, this type originates from two different proto-forms: Proto Tibeto-Burman (PTB) *nay or *s-nay ‘sun/day/dwell’ and *g-nam ‘sun/sky’ (Matisoff 2013, STEDT). In the maps below, we distinguish these two types: Ax) N- <*>nay type and A0) N- <*>g-nam type. (Although Ax can be further classified into a number of subtypes, they are not distinguished on the map). Additionally, more than six other types of stem are found: B) b/p- type, C) m- type, D) ts(h)/s/dz/tcV [-front] type, E) g/q/dz/tcV [+front] type, F) l/- type. Some types of stem such as F (l/-) are found only in compounds. The forms compounded with other stems are distinguished on the maps (e.g., ne21ge21 : Ax + E). Forms are classified independently of affixes, although some forms do contain an affix/affixes. (e.g., ȵi-ma : Ax type). Examples of each type follow:

Ax. N- <*>nay type

A-1: n- type. ne, ȵi, nī, nō, ning, nay, etc.

A-2: ŋ/ŋ- type. ŋi-ma, ŋ-ŋa-ma, ŋim, ŋā, ŋ3i; ŋ-ŋa-mō, ŋin-ga-ma, etc.

A0. N- <*>g-nam type

nam31, nam53, nam31, nam53+nam53, yna55.

B. b/p- type

pēi, paF, jbo24, yba24, ŋmbyi, ho222bō44, etc.

C. m- type

mũ̃mē, mgūŋ, mĩs31, a55-mq21-yu55, etc.

D. ts(h)/s/dz/tcV [-front] type

a33-thsu55, jan, k3-jam, etc.

E. g/q/dz/tcV [+front] type

gĩ33, dzǐ32, tị̃̃22, dzị̃̃32-mo21, etc.

In addition, various types of compounds are found, such as:

Ax + C type

nĩ55m33, n53ma33, naŋ55ma33, etc.

C + Ax type

mu21ŋi33, m21ŋi33, a55-mu21ŋi33-ma33 , etc.

Ax + E type

ne21ge21, ni44gi21-mo33, ŋi21ndzi21, etc.

A0 + F type (F: l/- type)

nam31ón55, nam55+nam31ŋ55, etc.

C + D type

mași, meswód, mĩs55tısı55, etc.

F + E type (F: l/- type)

li55tısı33, lo21tısı33-ma33, lo21gu33, etc.

2. Geographical distribution and interpretation

It is certain that Ax (N- <*>nay type) is the oldest of the types. It can be traced back to Proto Sino-Tibetan *ŋyi (Coblin 1986, STEDT). From the geolinguistic point of view, Ax is the most commonly and broadly found type in all of the southern, western and northern parts of the whole TB area. Most of the spots are of Burmese and Tibetan dialects, since they preserve reflexes inherited from Written Burmese ne and Old Tibetan (géni ma (ma is a suffix), respectively. However, the Ax type is also found in other languages: Akha naŋ, Haka Lai nīi, Khir ŋi55wa33, Nusu ŋi35a55, etc. Moreover, all of these spots are located in the neighboring or inner areas of the Burmese and Tibetan speaking regions.

Compound forms involving the Ax type, such as Ax + E, C + Ax, Ax + C, are also commonly found—more commonly than the other stems. They are found in Loloish languages that are spoken in the central and southeastern parts of the TB area. Ax + C and C + Ax show ABA distribution; therefore, Ax + C is considered older than C + Ax. Ax + E is broadly found in the eastern area of Loloish.

A0 type (N- <*>g-nam) is concentrated in the Nungic languages in the central part of the TB area, e.g., Rawang nam31, Trung num53, etc. Compound types such as Trung nam31on55 (A0 + F) are also found in the same area, with some exceptions such as gSerpo Tibetan na’tsa and sTau namtsa (both A0 + D) that are found in the northeastern and northerly-central parts of the TB area. Interestingly, some Tibetan dialects spoken in the immediate area to Trung also have compound types with A0: e.g., Jeshu Tibetan ‘n3 la. A0 type and its compounds are apparently new, because its distribution is narrow and concentrated in the central area.

B type (b/p-) is the second-most commonly found. Geographically, it is found both in the southerly-central area and the northerly-central area. In the southerly-central area, Northern Burmish languages such as Lashi pēi and Zaiwa bui, have this type. Burling (1967) provides a Proto-Burmish form *pweį ‘sun’. In the northerly-central area, Qiangic languages such as Nyagrong Minyag jbo24 and Stau yba24 have B type. Compound types involving B are not so common.

C type (m-) is not so common as A and B but is found in both the southern and northern areas: in
Karenic languages such as Hpa-an Pwo mѧթ mé in the south, Loloish languages such as Lalu Yi ъ55m֌ 21ћѧ55 in the southerly-center, and Qiangic languages such as Mawo Qiang њ55mɨθ55 in the southwestern area and C + D type, such as Northern Qiang məsi, that are found in the central area. In the Sal group of languages (Jinghpaw, Cak, and Kadu, etc.) this form originates from the common root *sal (Burling 1983), and Matisoff (2013) points out that it is derived from PTB *tsyar ‘sunshine.’ Interestingly, the compound types (D + C and C + D) and the plain D type form an ABA distribution. This may suggest that C type is older than D, and in the older stage D had to be compounded with C to mean “sun.”

E type (g/t/h/s/dz/tcV [-front]) is scattered in the central area: Southern Yi A55 tshu21, Jinghpaw jan, etc. Compound types involving D are of the D + C type, such as Cak comiʔ, that are found in the southwestern area and C + D type, such as Northern Qiang məsi, that are found in the central area. In the Sal group of languages (Jinghpaw, Cak, and Kadu, etc.) this form originates from the common root *sal (Burling 1983), and Matisoff (2013) points out that it is derived from PTB *tsyar ‘sunshine.’ Interestingly, the compound types (D + C and C + D) and the plain D type form an ABA distribution. This may suggest that C type is older than D, and in the older stage D had to be compounded with C to mean “sun.”

E type (g/t/h/s/dz/tcV [-front]) is found in the central and southeastern areas. All the languages that have plain E type are Loloish, e.g., Axi Yi tɕi22. The distribution is rather narrow; thus, we can conclude that this type is newer than Ax, B, C, and D. The proto form of E is unknown. Since Written Northern Yi has gwsi33, we consider that the older form of this type is gV and affricate initials result from palatalization by a front vowel. E is more commonly found as a part of the compound type Ax + E than by itself. Ax + E type shows more peripheral and broad distribution than plain E. Moreover, as we mentioned above, Ax is considered older than other types. This may suggest that E was first used as part of a compound, and later became an independent word meaning “sun.” Another compound type, F (l-) + E, is found only in the southern parts of the Ax + E area. F (l-) + E is newer than Ax + E.

F type (l/-) is not found as an independent word for “sun,” but it is found in several types of compound: Ax + F type (e.g., Trung nam31loŋ55), which is found in the central part, F + E type (e.g., Axi Yi li55ti33) in the southeastern part, and F + C type (e.g., Bwe lomu) in the southern part. One possible origin of the F-type stem is PTB *s-la(m/p) “dry (by fire / sun)” (STEDT). The Ax + F type is also found in some Tibetan dialects spoken in the central part of the TB area (or the southeastern end of the Tibetan area), e.g., Jesha Tibetan ⁿn5 лa. Suzuki (this volume) points out that such compounds literally mean “heaven-deity” (Written Tibetan gnam lha). That is, F-type stems here probably do not originate from PTB *s-la(m/p) “dry (by fire / sun)” but do originate from a word that means “deity.” The etymology of the F type is a subject for future investigation, and we tentatively do not distinguish these possible etymologies on the maps. The distribution of the F type suggests that each type of compound has been newly developed in each area.

### 3. Conclusion

In the present work, we collected Tibeto-Burman data of “sun” from 354 languages and dialects, and found more than seven stems that mean “sun.” On the maps, we distinguished six types of plain form and seven types of compound form. We also analyzed its geographical distribution and tentatively conclude the chronological order of plain forms as follows:

[Older] >>>>>>> [Newer]
Ax > B/C > D > E/A0

Interestingly, some compound types are considered older than plain types from the geolinguistic viewpoint: for example, compound types with D (D + C and C + D) are apparently older than the plain D type, while Ax + E seems older than E. This may suggest that some stems such as D and E cannot mean “sun” by themselves at the older stage.

**Keywords:** Tibeto-Burman, geolinguistics, compound, chronological order, “sun” and “eye”

(Satoko Shirai, K. Kurabe, K. Iwasa, H. Suzuki and S. Ebihara)
Map 1: “Sun” in Tibeto-Burman: The whole area
"Sun" in Tai-Kadai

The word forms representing "the sun" in Tai-Kadai can be classified into A) van type, B) ta van type, C) tang ngon type, and D) others.

Logically speaking, the monosyllabic form van meaning "the sun" should be the oldest form. It has the meaning "day" in many other dialects, which is more abstract and consequently should be due to a secondary derivation "sun" > "day (time period while the sun shines)". This type A is seen in the peripheral area, Li language in Hainan Island, Shan languages in Myanmar and Dehong, Yunnan, and the Surat dialect in the southern part of Thailand. In Surat, "the sun" is tone 1 and "day" is tone 4 of the same form "wan", thus the differentiation between two meanings is made by tone change in this dialect.

Otherwise, it was the original meaning "the sun" which developed a disyllabic form in order to distinguish with "day": ta van, or "eye" + "day", meaning "the eye of day", since an attributive is placed after the head in Tai-Kadai. Indonesian matahari "the sun" consists of the same word formation mata "eye" + hari "day". Similar calque phenomena are observed as well in Austroasiatic (for example, Aslian).

Since iambic is the dominant rhythm pattern in Tai-Kadai, the first syllable taa is shortened in several dialects, for example, tawan in Siamese. As a result, it became a sesquisyllabic word, and the etymological meaning of first syllable faded away.

First syllables appear to be aspirated forms in the western part of Guangxi as well as Hainan. The aspiration was caused by medial -r- according to Li (1977:119) who reconstructed its proto-Tai form as *traa, or *praa, taking the Saek form into consideration. Nishida (2000:95) supposed a different process from Austronesian *mata’ > (*m-)tha > tha~ha. The form starting with ha is seen in Vietnam and beyond. Doan (1996:104) drew a precise map of these forms in northeastern Vietnam.

On the other hand, the form ngon for "day" is a cognate word with van, for which Li (1977: 239,240,271) reconstructed *ŋwaŋ. In Map 1, "×" denotes the places where the onset of "day" starts from ŋ-, while "|" denotes the v- type.

Map 2 shows the geographical distribution of the first syllable ending with a coda, "○", while "." without coda -ŋ (the northern part shows no information compared to map 1, since the data for Buyi were available only for "day"). As seen from Map 3, which synthesized both maps 1 and 2, the coda -ŋ of the first syllable emerged mainly in the area where the second syllable starts with ŋ-. The dominant form in this area is tang ngon which means "lamp of day", as Chen (2011:170) pointed out. This is an example of paronymic attraction, where the new forms came about by folk etymology.

In Siamese, ดวงอาทิตย์ duang aathít or พระอาทิตย์ phrá aathít from Sanskrit ādityā, "a solar deity", are also used.

(Mitsuaki Endo)
A. "van" type: monosyllabic word denoting "sun" and "day"  
B. "ta van" type: "eye" + "day"  
   B-1: first syllable is long  
   B-2: first syllable is short  
   B-3: first syllable is aspirated  
   B-4: first syllable is "ha"  
   B-5: first syllable is "tsha"  
C. "tang ngon" type: "lamp" + "day"  
   C-1: first syllable is unaspirated  
   C-2: first syllable is aspirated  
   C-3: first syllable is "teang"  
   C-4: first syllable is "kjang" (meaning "time")  
D. Others  
   tai ngon  
   la:k van  
   nda van  
   thak ngn  
   thu rua~
“Sun” in Austroasiatic

Most word forms representing “sun” in Austroasiatic have other meanings, such as “day,” “sky” or “god,” or consist of elements denoting these meanings. They can be classified as follows:

I. Word forms including an element meaning “day”

A) ŋaj type
   A-1: ŋaj (ŋaj, ʔŋaj, ʔŋaj, ʔŋaj, ʔŋaj, ʔŋaj, ʔŋaj, ʔŋaj)
   A-2: ŋaj + ɡaj (ŋaj ɡaj, ʔŋaj ɡaj, ʔŋaj ɡaj, ʔŋaj ɡaj, ʔŋaj ɡaj)
   A-3: mat “eye” + ɡaj (mat ɡaj, mat ɡaj, mat ɡaj, mat ɡaj)
   A-4: Others

   leːŋ “sky” + ɡaj (leːŋ ɡaj, leːŋ ɡaj, leːŋ ɡaj, leːŋ ɡaj, leːŋ ɡaj, leːŋ ɡaj, leːŋ ɡaj, leːŋ ɡaj)

   The forms of this type are seen most widely in Austroasiatic; therefore, it is quite likely that this is the oldest type. There are monosyllabic forms, disyllabic forms, and compound forms. Most of them have the initial consonant ɡ-. However, the first syllable of it is the honorific term only “day” but also “sky” and “time.”

B) nar/dar type
   B-1: nar/dar (nar, naːr, dar)
   B-2: mat + nar/dar (madal)

C) ɡtɔʔ type
   C-1: ɡtɔʔ (kit ɡtɔʔ, kitkat ɡtɔʔ)
   C-2: mat + ɡtɔʔ (kit ɡtɔʔ, mat ɡtɔʔ, mat ɡtɔʔ)

   In Aslian, ɡtɔʔ is the word representing not only “day” but also “sky” and “sun.”

D) Others
   mat + ji:s (mat jis, mat jis), mat + ʔareʔ (mat ʔareʔ, meːŋ (meːŋ, meːŋ)

II. Word forms including an element meaning “sky, heaven”

E) leːŋ type
   E-1: leːŋ (harbaŋ, caroŋ…)
   E-2: mat + leːŋ (mat leːŋ, mat leːŋ, mat leːŋ, mat leːŋ)

F) loːj type
   mat + loːj (mat loːj, mat loːj, mat loːj, mat loːj

III. Word forms including an element meaning “god”

G) preah type
   G-1: preah + athit (preah athit, preah athit)

   The second syllables of these forms are related to Aditya “the sun god.”

G-2: mat + preah (mat p’rah)

IV. Others

gehən (qəhən), tawe (tawe, tawe, tawe, tawe: originated from ta van in Tai-Kadai), oyəŋ

Many compound forms have the first element mat, meaning “eye.” A combination of “eye” + “day” meaning “sun” can be observed in other language families (e.g., matahari in Indonesian). Other than “day,” the second element of compound forms beginning with mat can mean “sky” or “god.” However, the element mat is not used independently to denote “sun.” The presence or absence of the first element mat is shown in Figure 2, which shows that compound forms including mat are used in Eastern and Southern Mon–Khmer.

In Vietnamese, the form mat ɗːj “eye of the sky” is used. Besides this form, there is also the form oyəŋ ɗːj. The first syllable of it is the honorific term meaning “grandfather.”

**Keywords:** day, sky, god, eye

(KONDO Mika)

---

**Figure 1. Rhymes of the ŋaj type forms**

- :j, -ie~e~ɛ, -aj/-j,
- ɛj, -aj,
- ik, -wa

**Figure 2. Presence or Absence of First Element mat**

- presence, abs: absence
I. Word forms including an element meaning “day”

A) ŋaj type
   A-1: ŋaj
   A-2: ŋaj + ŋaj
   A-3: mat + ŋaj
   A-4: Others

B) nar/dar type
   B-1: nar/dar
   B-2: mat + nar/dar

C) ktsʔ type
   C-1: ktsʔ
   C-2: mat + ktsʔ
   D) Others

II. Word forms including an element meaning “sky”

E) leːy type
   E-1: leːy
   E-2: mat + leːy

F) lə:j type
   mat + lə:j

III. Word forms including an element meaning “god”

G) preah type
   G-1: preah + athit
   G-2: mat + preah
Sun: South Asia (IE (Indic, Iranian, Nuristani), Dravidian, Andamanese, Nihali, Burushaski)

1. Classification of word forms

In this map, there are eight major categories of word form: sūrya, āśāb, divākara, ravi, mithra, nēśara, vēlā, and porutu, and five minor categories.

A. sūrya: sūrya, sūryu, sūryo, sūri, suurī, sūuri, sūrē, sirē, sūr, sūrō, sūrya, sūrajā, sūraj, sūrī, sūrzō, sūj, su, sūryan, sūran, cūriya, suryya, suryyaya, sūryuddu

B. āśāb: āśāb, āśābu, oftob

C. divākara: divākara, dabākara, ditakara, dinayara

D. ravi: ravi, rav, rabi, rōbi

E. mithra: mihir, mier, mira, mer

F. nēśara: nēsar(a), nāyiku, nōś, dē

G. vēlā: vēla, vēle, vēra, vēda, verra, bēru, bēri

H. porutu: porutu, portu, port, portu, pord, podd, poddu, prodhu, pod

I. hiru: hiru, iru, ira, yir, irapojja

J. garm: ghām, yarma, gōmōej

K. din: dī, deō

L. xwar: xoršīd, nwar, nmar, lmar

M. nēra: ne-ra, nēram

N. others: nóo, yöor, roć, sa, bōdō-da, pute, die, diu, ēkē, pi, s, dūmba, devta

2. Geographical distribution and interpretation

The lexical forms representing the sun can be classified into A) sūrya type, B) āśāb, C) divākara, D) ravi, E) mithra, F) nēśara, G) vēlā, H) porutu, I) hiru, J) garm, K) din, L) xwar, M) nēra, and N) others.

The most major type is sūrya, which can be verified even in Ardda-Māgadhi, a kind of Prakrits, sūr sūr, sūria sūrī, sūrya sūrī, and in Sanskrit sūrya sūrye ‘sun’. Forms of this type are observed throughout South Asia, from the northern part of Pakistan to Sri Lanka, from Aryan languages, to Nuristani and even Dravidian languages, but not in Iranian languages. Some of the forms include /dāz/, /dā/, /j/, or /z/ sounds, which have been regularly changed from y /j/.

Āśāb forms are located in the western Islamic area. The forms of this type are derived from the Persian word āśāb ‘the sun, sunshine’.

The next divākara type can be seen in middle-north India and Sri Lanka. It is certain that this type of form is derived from the original form divasakāla and its meaning is ‘daytime’ (divasā ‘heaven, day’ + kālā ‘time’).

The Ravi type is observed in peripheral areas. Assamese languages are located in the north-east end of South Asia, Kashmiri in the north-west end, Konkani facing onto the Arabian Sea, and Oriya fronting the Bay of Bengal. Sanskrit also has the word ravi रवि ‘sun, mountain’. The difference in meaning between sūrya and ravi is uncertain there.

Mithra is the name of a god, who originally presided over promises, but later assimilated the characteristics of the sun god. In Pahlavi (Middle Persian), the form was mihr मिह्र meaning ‘sun’. This type is mainly seen in Iranian languages, whereas Bengali also has the form mihir.

Nēśara originally meant ‘sun’ in Prakrit but the forms of this group are widely distributed just among Dravidian languages, including the Brahui spoken in Pakistan and Afghanistan.

Velā वेला means ‘sun’ in Sanskrit. This type is also employed by Dravidian languages, but not by Aryan.

Porutu seems to have meant ‘time’ originally. This type is only observed in Dravidian languages.

Hiru type forms are distributed in the southern islands, i.e. Sri Lanka and the Maldives. The original form and meaning of this type are not clear. The Wakhi form yir may be classified into this type, but it is geographically the most separate from the other forms.

Nepali (Aryan) ghām, Pashto (Iranian) yarma, and Nihali (language isolate) gōmōej are clearly derived from Sanskrit gharma गर्म ‘hot’.

Gujar and Hindko in the northwest part of India and in Northern Pakistan employ dī and deō, respectively. These forms may be cognate with the word din ‘daytime’ in Hindi-Urdu.

Three Pashto forms nmar, nwar, and lmar and xor- of the Persian word xoršīd are cognate with the Pahlavi (Middle Persian) word xwar xor- ‘sun’.

Nēra originally meant ‘time’. This type can only be observed in Southern Dravidian languages.

(YOSHIOKA Noboru)
Map 1. ‘Sun’ in South Asia
The Sun: Arabic languages

1. Classification of word forms

In this map, word forms are classified as 5 large categories according to the root: \( f-m-s \), \( f-m-f \), \( s-m-s \), \( s-m-f \), and others. Many of the words of Semitic languages consist of three consonants, and vowels are allocated to the root on a different level.

A. \( f-m-s \) root type (شُمس) (●)

\( \text{fams, fams, fim}s \) in the singular form.

\( \text{famsa, fime:se} \) in the singular form with the feminine ending (-\( a \) or -\( e \) ـ).

\( \text{fims, fimu:s} \) in the plural form.

\( \text{hams} \).

B. \( f-m-f \) root type (شُمس) (○)

\( \text{famf}, \text{famf}, \text{femf}, \text{fimf}. \)

C. \( s-m-s \) root type (سعَد) (+)

\( \text{sams}, \text{sim}s, \text{sim}s, \text{semsi}. \)

D. \( s-m-f \) root type (سعَد) (×)

\( \text{simf}. \)

E. others

\( \text{harray} [\text{hara}ja], \text{harray} [\text{hara}j], \text{har(r)ā} [\text{hara}ha] \)

\( \text{nahar} [\text{naha}s] \)

2. Geographical distribution and interpretation

A. The oldest form of “the sun” in Arabic is \( \text{fams} \) and the root of it is \( f-m-s \). This form is used in Classical Arabic and is widely distributed throughout the Arabic-speaking areas. \( \text{fams} \) is used also in the less investigated areas. And even in the areas where other forms are dominant, the prestigious form \( \text{fams} \) is frequently used in formal speech.

\( \text{hams} \) is used by some tribes in South-west Saudi Arabia. These dialects have undergone a phonological change *\( f > h \), for example *\( \text{farib} > \text{harb} \). This change is exceptional in Arabic dialects (Behnstedt 2011: 403).

B. The root \( f-m-f \) is a result of assimilation of the third radical \( s \) to the first radical \( f \). In Maltese \( \text{femf} \) (orthographically ‘xemx’) is the form of the official language, Maltese. Thus the root is dominant in Morocco while in the other areas such as Sudan and a part of west Yemen and the Upper Egypt both \( f-m-f \) and \( f-m-s \) forms are used.

C. The root \( s-m-s \) is a result of assimilation of the first radical \( f \) to the third radical \( s \). This form is found in a wide area of North Africa including Egypt, Mauritania and Mali, and some parts of Yemen but only sporadically in dominantly \( f-m-s \) areas.

D. The \( s-m-f \) is a result of metathesis of the first radical and the third radical of the original \( f-m-s \). This form is found in Algeria and Tunisia but only sporadically and there is no \( s-m-f \) dominant area.

E. There are other words for “the sun”. The words found in Chad and Nigeria are \( \text{harrāya}, \text{harrāy}, \text{har(r)ā} \) with the common root of \( \text{harr} “\text{hot}” \) of Classical Arabic. \( \text{naha:r} \) used in Anatolia originally meant “day” in Classical Arabic.

The vowel pattern added to the roots depends on the rule of each dialect. \( \text{fams}, \text{famf}, \text{simf} \) can be found in Maghreb dialects (the North Africa except Egypt) as there are only two short vowels \( \text{a} (< *\text{a} \text{and} *\text{i}) \) and \( u \) in these dialects. The vowel \( i \) in \( \text{fims} \) (Oman, Daftar), \( \text{fimf} \) (Dakhla, Kharqa oasis) and \( \text{simf} \) is chosen probably because of non-emphatic environment. Maltese \( \text{e} \) in \( \text{femf} \) (orthographically ‘xemx’) was originally \( *\text{a} \) in Arabic. The Arabic \( *\text{a} \) was divided into \( a \) and \( e \): in the neighbourhood of an emphatic consonants \( (*\text{f}, *\text{d}, *\text{s}, *\text{t}) \) or pharyngeal consonants \( (*\text{s}, *\text{ḥ}) \) an open vowel \( a \) was retained, and in other environments the vowel was changed to \( e \). Some dialects, such as Baghdad (famis), Babylon (femes), Irbid Jordan (fames), have an epenthetic vowel between \( m \) and \( s \).

The final vowels in Anatolian \( \text{famse} “\text{Morning sun}” \) and Uzbekistan \( \text{famsa} \) are the feminine endings.

Diminutive form (*\( \text{fimaysa} \)) are also found: \( \text{fime:se} \) (Palestinian) and \( \text{sime:sa} \) (Bahriyya). The plural form \( \text{fimu:s} \) is also found in Yemen. \( \text{fimu:s} \) is a variant of it.

The roots \( f-m-f \) or \( s-m-s \) as results of assimilation are interesting because of the fact that, in Semitic languages, a general principle in root-formation states that the first and the third radicals may not be identical in principle. In Morocco, besides \( \text{famf}, \) there is also a root word \( *\text{zaw} ‘a pair’ > ʒʒɛs ‘two’ \). Thus Moroccan dialects seem to have a weak constraint against the identical radical rule.

However, in a number of Semitic languages the roots of “the sun” are the first- third radical identical root: \( f-m-f \) (e.g. \( \text{femef} \) in Hebrew, \( \text{fimfa} \) in Syriac, \( \text{famfu} \) in Akkadian, \( \text{fapu} (< f-m-f) \) in Ugaritic. This is result of the merger of \( *\text{th} \) into \( *\text{s} [J] \) of Proto-Semitic. Only a few Semitic languages have the first-third not identical root: Arabic \( f-m-s \) and South Arabian languages such as Sabaean \( s-m-f \).

\( \text{Keywords:} \) assimilation, merger, root

(Youichi Nagato)
References (the other languages will be included in the following issues)

Ainu


Arabic


Omar, Margaret K. (), Saudi Arabic Urban Hijazi dialect Basic Course. Foreign Service Institute: CT.


Yoda, Sumikazu (伊藤純和) (2012), 『アラビア語エルサレム方言文法研究』. 渓水社: 濱松。

Sinitic languages
《Changningxianzhi》 bianzuan weiyuanhui bangongshi [常宁县志编纂委员会办公室]. 1990. Changningxian fangyanzhi [常宁县方言志]. [].
Cui, Rong [崔容]. 2004. Taiyuan bei jiaou fangyan yanjiu [太原北郊区方言研究]. Shanxi renmin chubanshe [山西人民出版社].
Daxi, Bozi [大西博子]. 1999. Xiaoshan fangyan yanjiu [萧山方言研究]. Haowenchuban [好文出版社].


Zeng, Yumei[曾毓美]. *Hunan Jianghua Zhaihanhua yanjiu*[湖南江华寨话研究]. Hunan shifan daxue chubanshe[湖南师范大学出版社].


Jiangsu jiaoyu chubanshe[江苏教育出版社].


Tibeto-Burmese


Hanson, Ola. 1906. A Dictionary of the Kachin Language. Rangoon: Baptist Mission Press. (Jinghpaw (JA))


Dialectal Forms Associated with the Word Taiyō (Sun) in Japanese

Shinsuke Kishie¹
Yukichi Shimizu¹
Yukako Sakoguchi¹

Keywords: Japanese Dialect, Ryukyuan dialects, Geographical Distribution, Chinese Origin, Honorific.

1. Introduction

There is a variety of dialectal word-forms associated with the term Taiyō (sun) in Japanese. They can roughly be classified into four groups: word-forms of Japanese origin, word-forms of Chinese origin, indigenous word-forms of Ryukyu dialect as well as the word-forms of sun with affixation of honorific formatives. As for Ryukyu dialect, we discuss whether or not this has connection with the word form in southern language.

We herewith will put forward a geographical distribution of the dialectal word-forms associated with the word Taiyō (sun) along with a description of their constituents followed by a discussion.

2. Distribution of Dialectal Word-forms Associated with the term Taiyō (sun) in Japanese


The outcome of LAJ has been published as a simplified dialect map in the Shougaku-tosho (1991) and compiled version of it by Sato (2002) as well as the Nihon-Hogen-Daijiten (Japanese Dialect Dictionary). Dialect map-1 depicts a dialect distribution based on the result published in the Gendai-Nihongo-Hogen-Daijiten. Though this dialect map has been drawn based on the data from very few locations, it significantly coincides with the result presented in the dialect map 251 of LAJ Vol. 6.

According to dialect map-1, the dialectal forms representing the word Taiyō (sun) can be grouped into several lineages such as hī [çi:], hidon [çiidoN], ohī-sama [oçi:sama]/ohī-san [oći:saN], otentō-sama [otento:sama]/otentō-san [otento:saN], kon'nichisan [konnitʃisaN], nichirinsama [niʃirinsama], nittensama [nittensama] and tida [tida].

These dialectal forms can be characterized with four following features: abundance of word-forms of Japanese origin, abundance of word-forms of Chinese origin, idiosyncratic word-forms in Ryukyu dialect of Okinawa and abundance of dialectal forms with the honorific affixes.

¹ The Tokushima University
Firstly the word-forms having Japanese origin are realized with the dialectal forms, all of which are constituted of formative *hi* [çi], for example *ohi-sama* [oçi:sama], *ohi-san* [oçi:saN] and *hidon* [çidoN] etc. Except these word-forms, all other are the word-forms of Chinese origin which includes shared word form *taiyō* [taijo:]. Accordingly this is now evident that *otentō-sama* [otentoːsama], *nichirin-san* [nitʃirinsaN], *nittensan* [nittensaN] and *kon'nichi-san* [konnitʃisaN] have respectively derived from the words of *tentō* [tento:], *nichirin* [nitʃiriN], *nitten* [nitteN], *konniichi* [konnitʃi], which are of Chinese origin.

In highlighting the countrywide geographic distribution of this dialect map, Sato (2002) along with some other dialectologists have pointed out that *otentō-sama* [otentoːsama] /
otentōsan [otento:saN] and ohīsama [oçi:sama] / ohīsan [oçi:saN] are respectively distributed mostly in the center and the periphery of the Archipelago. From this distribution of dialectal forms, it is can now be inferred that the word-forms of ohīsama [oçi:sama] / ohīsan [oçi:saN] of Japanese origin are older than the word-forms of otentōsama [otento:sama] / otentōsan [otento:saN] of Chinese origin. In connection with this inference, we can now fervently expect the studies on the dialectal distribution of word-forms derived from the word Taiyō (sun) of Chinese origin in other Asian languages.

2. On Etymology of the Word-form tida

Tida [tida] is a word-form that only can be found in the Ryukyuan dialect. There are several theories in connection with etymology of word-form tida [tida] including the theory of southern language lineage. The affiliation with the southern languages was first postulated by Izuru Shinmura in the early year of Showa era. Shinmura (1927) hypothesized that tida [tida] is a word-form of Amis language of Taiwan which is associated with chidaru [tsidar] that represent meaning of Taiyō. This is the first hypothesis to associate the tida [tida] with theory of southern language lineage. Subsequently Ando (1935) showed etymological connection of chidaru [tsidar] of Amis language as hypothesized by Shinmura (1927) with word-forms todaru [todoru] and chidaru [tʃidar] that appeared in the Kojiki as well as the word-form sinar [sinar] that is used to mean light in Malay language. Tida [tida] being a word-form of archaic Yamato language is considered to share the origin with southern language. Thus he hypothesized that tida [tida] holds etymological connection both with archaic Yamato language and southern languages. According to Mamiya (2014), Yoshitada Nakahara is the first to mention the connection of tida with teru. Nakahara in his thesis Omoro Shinyaku (1957) demonstrated that tida has derived from teru with the analogy that teriya became tera which changed to teda due to the confusion between the sound of syllable of ra-row to that of da-row.

Sun in Ryukyu Dialects

Zusetsu Ryukyugo Jiten, Kinkeisha

Dialect map- 2-1 Nakamoto(1981)
On the other hand, Kamei (1973), Uemura (1963) and Mamiya (2014) have demonstrated that tida [tida] has derived from the word tendō [tendo:] of Chinese origin. If tida [tida] is taken to be derived from the word-form of the tendō [tendo:], then it can be regarded to share the origin with word otento-sama [otento:sama] of mainland dialect. According to Mamiya (2014) if tida [tida] is taken to be word-form derived from tendō [tendo:], then both Yamato language and Ryukyuan language essentially share not only the word-forms having origin in hiç[çi:] and tendō [tendo:] of Japanese language, but also the suffixational honorific formatives like ohisama [oçi:sama] and otento:sama [otento:sama].

Though there have been several hypotheses with regard to the etymological origin of the word-form tida, Nakamoto (1981), Hokama (1981) and Mamiya (2014) demonstrated that none of them is decisive. Of these hypotheses, the hypothesis of showing connection of tida [tida] with tendo [tendo:] is considered to be irrefutable at this time. This hypothesis endorsed in particular by a number of dialectologists in Japan.

Nakamoto (1981), as shown in dialect map-2, presented a distribution of dialectal word-forms for the sun in Ryukyuan language showing its etymological connection in the following 6 probable lineages. These word-forms: tidan [tidaN], tida [tida], tiida [tii:da], tira [tira], chida [tfida] and shira [jira]. shira [jira] (or shina [jina]), as described in the above, is thought to have connection with Amis language and Malay language, while Nakamoto (1981) hypothesized to have derived they from tida [tida] through morphological change, both of which contrast Murayama’s (1976) claim, for which it cannot directly be connected with the southern languages. Though distribution of dialects in the dialect map-2 shows that shira [jira] and shina [jina] are seen in some locations of Main Island in Okinawa and Hateruma, it cannot be concluded that they bear etymological origin to the word-form tida [tida]. One of the reason for endorsing this hypothesis lies with presence of extensive word-forms having lineage to the word tida [tida] distributed throughout the Ryukyu Island for which derivation of shira [jira]
and shina[ʃina] etc. can be described as change analogous to change from [t] to [ʃ] and that from [d] to [r].

According to Mamiya (2014) it is alarming to show connection of tida [tida] excessively with chidar (sun) of Amis language and sinar (sunshine) of Malay language. Though these resemble with the word-form and meaning of southern languages such comparison should be avoided for not having evidence in favor of the familial relationship between Japanese and these southern languages.

3. On the Dialectal Forms and Suffixational Titles Associated with Word Taiyō (sun)

In Japan, the sun, like God or Buddha, has been worshiped and embraced for its reverence since ancient time. This custom has not been focused only on the sun, but on other similar things. Therefore the moon and the lightning as well as large trees, large stones and natural materials had also been object of worship. There are several dialectal forms for the Taiyō (sun), which include word-forms with the prefixation of ‘o (go)’ representing respect and the suffixation of “san (sama)” representing courtesy title as well as the word-forms with the suffixation of formative don derived from dono (tono) in the dialect of southern Kyushu region. These dialectal forms are distributed throughout the whole region of Honshu. There are also word-forms such as o-tentou-sama [otentoːsama] and o-hi-san [oçiːsaN] which carry both of prefix and suffix. Yet there is a small number of word-forms as to tayō which carry honorific affixes, for example taiyo-sama [taijoːsama] and taiyo-san [taijoːsaN]. Still there is a word-form like tidaganashi [tidaganaʃi] that carries a courtesy title ganashi [ganaʃi] for the sun in the several Ryukyuan dialects, hence shares a common characteristic with the Mainland dialect that both of them carry formative of courtesy title.

The word-forms containing the honorific formatives show an idiosyncratic geographical distribution. Firstly the word-forms carrying an honorific formative, e.g. Nichirin-san [nitʃirinsaN], Nitten-san [nittensaN] and Kon’nichi-san [konnitʃisaN] are widely distributed in western Japan contrary to Eastern Japan where it is scarcely distributed.

Distribution maps which have been created focusing on the regional differences with regard to honorific suffixes with the data from Teruo Hirayama et al. (1992) and LAJ Vol.6 are shown in the dialect maps 3-1 and 3-2. Although there remains a difference in the number of locations surveyed, the two dialect maps show the identical results in comparison. However it can be noted that the distributions of sama and san show contrast along the line of east-west opposition. While sama is distributed in Eastern Japan, san is distributed in western Japan representing an opposition between these two linguistic situations. In strict sense, however, the distribution of Sama-San opposition can said to be an A-B-A type arrangement rather than an arrangement of East-West opposition, since sama is heavily concentrated in the regions like western Midland and southern Shikoku as well as Kyushu showing a distribution of old word-form in the periphery. Though san cannot be said to derive from sama by change, the dialect map 3-2 shows a dialect situation that san is distributed from central Kansai district towards the different regions of western Japan. This change, however, can be said to reflect prompted by custom of deep affection as well as respect towards the sun in the Kansai region.

Therefore it would be interesting to know if there is any language or dialect in Asia, like Japanese, which contains the honorific word-forms for the sun, the moon and the thunder for having custom of worshipping, and faith and reverence towards the sun, the moon and the thunder.
Figure 3-1. Sama/San in Japanese

Figure 3-2. Sama/San in Japanese
4. Conclusion

Though there are dialectal word-forms of Japanese origin *hi* and numerous dialectal word-forms of Chinese origin representing the *Taiyō* (sun) in Japanese, it can be recognized form the dialectal situation of geographical distribution that the dialectal forms associated with *hi* is the oldest of all in the Japanese archipelago. In addition there are numerous dialectal-forms of Chinese origin including the *tentō*, they tend to be distributed towards western Japan.

Though the dialectal word-forms associated with *tida* and *tida* distributed in different regions of Ryukyu had been hypothesized to be connected with Austronesian languages by conducting comparative study in the early period of Showa era, there are some dialectologists who think *tendō* to bear connection with the Inland dialect. Therefore it would be premature to conclude on the forms of two different languages having no phyllogenetic relationship to share same lineage with the analogy that they match in meaning and forms.

It is remarkable to see that word-forms associated with *taiyō* (sun) in Japan characteristically bear abundant honorific formative in the dialects distributed throughout the country. The distribution of word-form *sama* and that of *san* show difference by regions given that *san* has changed to *sama* which likely dispersed from central Japan toward the peripheral regions.

We therefore expect the study aimed to reveal the phenomena with regard word-forms denoting the sun in other languages of Asia as we revealed the phenomena of dispersion of dialectal forms originated from word of Chinese origin and that of presence of word-forms affixed by honorific formatives for the sun being object of reverence in Japanese.

Acknowledgements:

We would like to give special thanks to Mr.Yasuo Kumagai and Mr.Takuichiro Onishi for giving us a chance to use Linguistic Atlas of Japan Database.

Reference:

Shinmura Izuru(1927) : Nihonjin to Nanyo, Toho Gengoshi Ronso, Iwanami Shoten.
Ando Masatsugu(1935) : Koten to Kogo, Sanseido.
Kamei Takashi(1973) : Tida no Gogen, Nihongo Keitoron heno Michi, Yoshikawakobunkan.
Murayama Shichiro(1979) : Nihongo no Tanjo, Chikuma Shobo.
Kamimura Takaji (1963) : Ryōkyu Hōgen no Taiyō wo Imisuru Go ni tuite: Kagoshimadaigaku Bunka Hōkoku.
Geographical distribution of ‘daytime’ in Ainu

Mika Fukazawa
Graduate School of Humanities and Social Sciences, Chiba University

Abstract
This paper attempts to describe the geographical distribution of ‘daytime’ in Ainu. To demonstrate the origin and history of the words for ‘daytime,’ we researched the compound nouns for ‘the sun’ and the expressions for time divisions of the day: ‘morning,’ ‘noon,’ ‘afternoon,’ ‘evening’ and so on. As a result, we classified the words for ‘daytime’ into three types, and suggested that the first toono type, which has the root tó/too ‘day (24 hours),’ would be older than the other two. The second sirpeker type etymologically comes from the meaning of ‘day breaks; it gets light,’ and is distributed over the eastern Hokkaido and Kuril Islands. The third tôkap type shows the different meanings of ‘(older woman’s) breast’ and ‘daytime’ and it is uncertain how the homonymy or polysemy occurred. However, the term tôkap for ‘breast’ seems to limit the use and meaning to avoid the homonymic clash. If the sense of ‘daytime’ was derived from ‘breast,’ the semantic shift would be related to some “sociocultural change” in Ainu.

1 Introduction
This article describes the geographical distribution of ‘daytime’ in Ainu and sheds light on the historical process of it. The Ainu language1 is one of the indigenous languages in Japan, spoken throughout Hokkaido, Sakhalin, and the Kuril Islands. Its genetic relationship with Japanese and any other languages has not been demonstrated, namely, it is a language isolate. It has five vowels /i, e, a, o, u/ and eleven consonants /p, t, k, c, s, m, n, r, w, y, h, (’)/.

The first, in the next section, is to show the words for ‘night’ and ‘daytime,’ and the compound nouns ‘the moon’ and ‘the sun.’ After the classification of these words, we will survey the words that seem to share the same root as the words for ‘daytime,’ and the expressions for the time divisions of the day with the focus on the time of ‘noon.’ Finally, we will suggest the origin and history of the words for ‘daytime,’ through the geographical distributions and the etymological analysis of them.

2 Large bright stars in the daytime and the night
In this section, we will see the distributions of the words for ‘daytime’ and ‘night,’ and the compound nouns for ‘the sun’ and ‘the moon.’ In Ainu, the term cup stands for both ‘the sun’ and ‘the moon.’ When we distinguish them individually, the former is expressed as “the daytime cup” and the latter, as “the night cup.” It should be noted that the terms for ‘the sun’ show dialectal variation.

2.1 ‘Night’ and ‘the moon’
Before we will introduce the geographical distributions of ‘night’ and ‘the moon,’ we have to note that the Sakhalin and Hokkaido dialects show the typical differences in the phonological and/or phonetic structures (see Chiri 1942, Tamura 2000 etc.), as follows:

(a) Accent: the majority of Hokkaido dialects have a pitch accent, and instead, in the Sakhalin dialects, vowel length is distinctive. A few dialects in Hokkaido have no accent.

---

1 The language is a polysynthetic SV/AOV, Basically head-marking type and allows ‘pro-drop,’ which is compensated with verbal agreement (singular/plural forms and personal affixes/clitics).
(b) Phonetic correspondence: the codas /-p, -t, -k/ in Hokkaido have historically changed to /-h/ in most Sakhalin dialects, and /-r/ has changed to /-h/ or /-rV/ in Sakhalin.

The words for ‘night’ are künne in Hokkaido and kunne in Sakhalin\(^2\) (Hattori and Chiri 1960). In northern Kuriles, Krascheninnikov’s “Vocabularium Latino-Curilice”\(^3\) takes the form sirkunne ‘night,’ consisting of sir- ‘view’ and kunne ‘to be dark/black.’ Hattori (1999 [1959]: 152) suggested that kunne originated from √kur + the copula ne, and the following words were based on the root: kur ‘shadow,’ niskur ‘cloud’ and ekurok ‘to be dark.’ If so, the other terms in the Kuril dialects could also be constructed with √kur: sirikurako (Krascheninnikov) for ‘to be dark,’ sirkurka (Dybowski 1892) and sirekorak (Torii 1903) for ‘night.’ The words for ‘the moon’ show a similar distribution to the ones for ‘night.’ Along the rules of (a) and (b) above, the terms for ‘the moon’ are künne cup or kunne cup in the Hokkaido dialects, and have turned into kunne cuh in most Sakhalin dialects.

2.2 ‘Daytime’ and ‘the sun’

The distributions of the words for ‘daytime’ and ‘the sun’ are different from the ones for ‘night’ and ‘the moon.’ The lexical forms are classified into three types, as in Table 1, and plotted out on three dialectal areas respectively: A) Sakhalin, B) eastern Hokkaido and northern Kuriles, C) western Hokkaido. This distributional pattern is not reported in Nakagawa (1996), though it can be realized if “Eastern-Western Type” and “Sakhalin Type” are merged into one type.

<table>
<thead>
<tr>
<th></th>
<th>‘daytime’</th>
<th>‘the sun’</th>
</tr>
</thead>
<tbody>
<tr>
<td>A) toono type</td>
<td>toono type</td>
<td>toono cup type</td>
</tr>
<tr>
<td>A-1</td>
<td>toono</td>
<td>toono cup type</td>
</tr>
<tr>
<td>B) sirpeker type</td>
<td>B-1 sırpeker/sırpeker</td>
<td>sirpeker cup (kamuy) type</td>
</tr>
<tr>
<td>B-2</td>
<td>—</td>
<td>peker cup kamuy</td>
</tr>
<tr>
<td>C) tókap type</td>
<td>C-1 tókap/tókap</td>
<td>tókap cup (kamuy) type</td>
</tr>
<tr>
<td>C-2</td>
<td>tókap</td>
<td>tókap cup (kamuy)</td>
</tr>
<tr>
<td>C-3</td>
<td>tóykapne</td>
<td>—</td>
</tr>
</tbody>
</table>

Table 1: ‘daytime’ and ‘the sun’

A) *toono* type

Type A is observed in Sakhalin and the northernmost part of Hokkaido, Sooya. The long vowel oo [oː] in Sakhalin is phonetically equivalent to a pitch accent ó in Hokkaido. The monosyllabic form *too* means ‘day,’ although it is uncertain what the following morpheme *no* is. I will discuss how it might be related to the locative noun *nóski* ‘the middle of,’ and how this type is older than Type B and C, in section 3.

B) *sirpeker* type

The basic term *sirpeker* means ‘day breaks; it gets light,’ which consists of the prefix sır- ‘sight; view’ and the intransitive verb peker ‘to be light.’ In the Hokkaido dialects of Obihiro and Bihoro, the final consonant /r/ [ɾ] is changed to /t/ [t] before the consonant /c/ [ʣ, ʤ, ʦ, ʧ]. Since this type is found also in the Kuril Islands, it may be older than Type C.

---

\(^2\) In addition to this, Hattori and Chiri (1960) recorded the word for ‘the moon’ as ku**nne cuh** in Ochiho, Sakhalin. Sometimes, maybe idiolectally, [n] is pronounced as /n/, but not [n] in Ainu.

\(^3\) The year of completion is unknown.
C) tókap type

This type has three subtypes: tókap/tokap, tókam and tóykapne. In most Hokkaido Ainu, “[i]f the first syllable is open, the accentual nucleus is on the second syllable” (Tamura 2000: 21), but “[f]or compound nouns and derived words, because the first element’s accent takes priority, exceptions may arise,” i.e., rékor ‘to have a name’ from re ‘name’ + kor ‘to have’ (ibid.: 23). Sato (2015: 2) reexamined the accentuation rules of compounds and exhibited the strong tendency in the case that the former element is CV. This would be the reason why the words of type C have exceptional accents on the first syllables.

In C-1, the former CV element, tó, of tókap represents ‘day (24 hours)’ and/or ‘breast,’ and the following káp may come from ‘skin.’ See a more detailed discussion in section 3. In C-2, the coda /m/ of kám would be changed from /p/ of káp in C-1, which may be phonetically attracted by analogy with the meaning of kám ‘meat.’ The word tóykapne of C-3, uncertainly from tóykap ‘daytime’ + the copula né, is in Sooya near the Sakhalin Island. The element tóy of tóykap could be affected by a long vowel, i.e., too [to:] ‘day’ in the Sakhalin dialects.

Map 1: ‘daytime’
3.1. ‘Noon’

Map 3 shows the distributions of the words for ‘noon’ in Ainu. The terms tónonoski and tónanoski are older than any other lexical forms. The word tónonoski can be decomposed into tóno ‘daytime’ and nòski ‘the middle of.’ The element tóno would be the same origin as the word for ‘daytime,’ toono, in Sakhalin. The word tónoski can also be decomposed into tó ‘day’ and nòski ‘the middle of,’ but it could be simply interpreted as the contracted form of tónonoski. The dialects of Yakumo and Horobetsu have both types of the words, tónonoski and tónoski.
The words for ‘noon’ include a complicated problem—What time is ‘noon’? For example, the word *tókes* means ‘noon’ in the south-eastern dialects of Hokkaido, while it means ‘afternoon’ and/or ‘evening’ in the other regions.

In Obihiro (Sawai and Tamura 2005: 339; Hattori (ed.) 1964:92, 252):

- *tókes pákno*  
  ‘morning’ (lit. before the noon)
- *tókes*  
  ‘noon’

4 In northern Kuriles, the other lexical forms for ‘noon’ are also reported: *toanonoschi/toanonoski* and *dohnonóský/tononoski* (Torii 1903; Krascheninnikov’s Vocabularium Latino-Curilice; Klaproth 1823).
tôkes’îpé  ‘lunch’ (lit. a noon meal)  
tôkes oráno; tôkes wano  ‘afternoon’ (lit. from the noon)  
oúman  ‘evening’

In Shizunai (Watanabe et al. 1984: 100-101):
tokap etok  ‘morning’ (lit. before the noon)  
(unknown)  ‘noon’  
tokap’îpe  ‘lunch’ (lit. a noon meal)  
tôkes  ‘afternoon’  
onúman  ‘evening’

(unknown)  ‘morning’  
tôkam; tônoski  ‘noon’  
tôkam’îpe; tônoski’îpe  ‘lunch’ (lit. a noon meal)  
tôkes; onúman  ‘afternoon; evening; about 3:00 p.m. - 6:00 p.m.’

In Bihoro (Watanabe et al. 1986: 104; Hattori (eds.) 1964: 92, 252):
tonanoski etok  ‘morning’  
(lit. before the noon)  
tonanoski  ‘noon’  
tonanoski’îpe  ‘lunch’  
tonanoski ipe okaketa  ‘afternoon’  
(lit. after the lunch)  
tonanoski wano  ‘afternoon’  
(lit. from the noon)  
tôkes; onúman  ‘evening’

In Raichishika, Sakhalin (Hattori (eds.) 1964: 252):
(unknown)  ‘morning’  
toonoske; toonoski  ‘noon’  
toonoski’îpe  ‘lunch’  
(lit. a noon meal)  
tookes  ‘afternoon; about 4:00 p.m. - 5:00 p.m.’  
onúman  ‘evening; after the time of tookes’

In the 19th centuries, some written materials are made by a translator between Japanese and Ainu, so that the time expressions in the Japanese-Ainu glossaries are along “unequal hour§,” which Japan has used until the Edo period.

In Notoya (1868)’s the Japanese-Ainu glossary:
sirpeker  [Siripesker]  辰時／朝五ツ {about 7:00 a.m. - 9:00 a.m.}  
tôkap’etoko  [トウカフイトコ]  巳ノ時／昼四ツ {about 9:00 a.m. - 11:00 a.m.}  
tôkap  [トウガフ]  午ノ時／昼九ツ {noon; about 11:00 a.m. - 1:00 p.m.}  
tôkes  [トウケシ]  未ノ時／昼八ツ {about 1:00 p.m. - 3:00 p.m.}  
onuma(n)  [ヲヌマ]  申ノ時／昼七ツ {evening; about 3:00 p.m. - 5:00 p.m.}

§ It is the twelfth period of the daytime (from sunrise to sunset), and the night (from sunset to sunrise). The hours are “unequal” because the length of daytime and night varies according to the seasons and the geographical latitude.
Undoubtedly, the original meaning of the word tókes must be ‘afternoon’ and/or ‘evening’ because it can be analysed into tó ‘day’ and the locative noun kés ‘the end of.’ The meaning of ‘noon’ by semantic shift occurred in the limited geographical area, as shown in Map 5.

Also in English (c.f. OED 1989: 508), the word noon meant “[t]he ninth hour of the day, reckoned from sunrise according to the Roman method, or about three o’clock in the afternoon,” coming from the Latin word of nona (hora). It was the name of the time to pray and prayers. By the 14th century, the sense has been shifted to ‘twelve o’clock in the day’ and fixed, because the time for praying (and eating a meal)⁶ has been changed from 3. p.m. to 12 p.m..

Map 5: ‘noon’ and ‘lunch’

- tókapnoski ‘noon’ : Biratori, Chitose, Mukawa
- tokap’ipe ‘lunch’ : Yakumo, Horobetsu, Biratori, Mukawa
- tokamnokski ‘noon’ : Nayoro
- tókam ‘noon’ : Asahikawa
- tókam’ipe ‘lunch’ : Aasahikawa
- tókes ‘noon’ : Obihiro
- tokes ‘noon’ : Samani
- tókes’ipe ‘lunch’ : Nayoro, Obihiro, Hombetsu

According to the report of Obihiro, where the word for ‘noon’ is tókes, the native speaker said “I do not call the sun tókes cup” (Sawai and Tamura 2005: 272). In contrast, the words tókapnoski for ‘noon’ and tókap’ipé for ‘lunch’ are used in the same place there is tókap for ‘daytime’ (see also Map 1).

⁶ One of the semantic shifts due to “sociocultural change” (Blank 1999) is the names for meals in European languages (Blank 1999, González 1993) and Jamaica (Hock 1986). For example, in French, “the binary system of the Middle Ages thus developed through a ternary to a fourpartite structure shifting the words along the temporal contiguity of meals”; until the 16th century, people used to have a lighter meal, souper, in the afternoon, but souper now serves to designate a late-evening meal in the 19th/20th century (Blank 1999:73-74).
3.2. ‘Breast’—The same lexical form as ‘daytime’

Here, let us consider the words for ‘breast’ on Map 6, which take the same lexical forms as the ones for ‘daytime.’ The monosyllabic forms to/too must be oldest for ‘breast,’ and the words tótto/totto are the reduplicated form of it. The type of tókap, coming from the compound tô ‘breast’ + káp ‘skin,’ is distributed over Hokkaido. In Asahikawa, Nayoro, and Horobetsu, where some lexical forms are overlapped with each other, there are various forms with different meanings and uses: tô means ‘breast,’ tótto is used in baby talk, and tôkap stands for ‘older woman’s breast,’ vis-a-vis tôkap/tókam for ‘daytime.’

Then, why are the words for ‘breast’ and ‘daytime’ taking the same form? There is no answer for now; however, the sense of ‘breast’ seems to have been not used or limited to the meaning of ‘older woman’s breast,’ and sometimes the form itself was shortened as káp/kap (lit. skin). It may be caused by a “homonymic clash.” Otherwise, the semantic shift from ‘breast’ to ‘daytime’ would be motivated by the metonymy, i.e., the custom of breast-feeding for lunch, and some kinds of “sociocultural change” as we saw in section 3.1. above.

Map 6: ‘breast’

- to, -ho : Horobetsu, Nayoro, Asahikawa
- too, -ho : Raichishika
- tótto, (-ho) : Yakumo, Horobetsu, Biratori, Asahikawa, Nayoro, Sooya, Chitose
- toto : Samani, Shizunai
- tôkap : Yakumo, Horobetsu, Obihiro, Asahikawa, Nayoro, (Hombetsu)

7 In the 1st meeting, academic year 2015 of Studies in Asian Geolinguistics at AA Institute (October, 3rd, 2015), Prof. Chitsuko Fukushima and Prof. Kazue Iwasa told me how to consider the homonymic clash and semantic change. I still could not have a clear answer, but their suggestions inspired me to reconsider the history of tôkap/tókam.
4 Conclusion

The following table summarizes the types of the lexical forms for ‘daytime,’ ‘sun’ and ‘noon.’

<table>
<thead>
<tr>
<th>A) the toono type</th>
<th>‘daytime’</th>
<th>‘the sun’</th>
<th>‘noon’</th>
</tr>
</thead>
<tbody>
<tr>
<td>A-1 toono</td>
<td>toono</td>
<td>toono cup type</td>
<td>tónonoski type</td>
</tr>
<tr>
<td>A-2</td>
<td></td>
<td></td>
<td>tónoski type</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>B) the sirpeker type</th>
<th>‘daytime’</th>
<th>‘the sun’</th>
<th>‘noon’</th>
</tr>
</thead>
<tbody>
<tr>
<td>B-1 sirpeker/sirpeker</td>
<td>sirpeker cup (kamuy) type</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B-2</td>
<td></td>
<td>peker cup kamuy</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>C) the tókap type (related to ‘breast’?)</th>
<th>‘daytime’</th>
<th>‘the sun’</th>
<th>‘noon’</th>
</tr>
</thead>
<tbody>
<tr>
<td>C-1 tókap/tókap</td>
<td>tókap cup (kamuy)</td>
<td>tókapnoski</td>
<td></td>
</tr>
<tr>
<td>C-2</td>
<td>tókam</td>
<td>tókam cup (kamuy)</td>
<td>tókamnoski</td>
</tr>
<tr>
<td>C-3</td>
<td>tóykapne</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>D) the tókes type (originating from ‘afternoon’)</th>
<th>‘daytime’</th>
<th>‘the sun’</th>
<th>‘noon’</th>
</tr>
</thead>
<tbody>
<tr>
<td>D-1</td>
<td></td>
<td></td>
<td>tókes/tókes</td>
</tr>
</tbody>
</table>

Table 2: ‘daytime,’ ‘sun’ and ‘noon’

The noun root tó/too means ‘day (24 hours)’ in Type A, C and D. We suggested the derived noun tóno/toono would be the oldest form for ‘daytime,’ and the words for ‘noon,’ i.e., tónonoski and tónoski, are constructed with it and the locative noun nóski ‘the middle of.’ These words are widespread, and this is the evidence that A) the toono type is oldest. B) the sirpeker type is not used for ‘noon,’ but must be older than the types of tókap and tókes. The word sirpeker originally has the meaning of “day breaks; it gets light,” and this type can be also seen in northern Kuriles.

The words for ‘breast’ also have tó as the noun root, and might be the source of C) tókap type. They are distributed over Hokkaido, and the places using the term tókap for ‘breast’ and ‘daytime’ are overlapped with each other. However, the term tókap for ‘breast’ appears to be limited for use and meaning due to homonymic clash. Among the lexical forms for ‘noon,’ D) the tókes type is newest, as a result of the semantic shift from ‘afternoon.’

Appendix

List of Used Data:

<table>
<thead>
<tr>
<th>Dialects</th>
<th>Materials</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yakumo/ Horobetsu/ Nayoro/ Sooya/ Raichishika</td>
<td>Hattori and Chiri (1960) and Hattori (ed.) (1964)</td>
</tr>
<tr>
<td>Oshamambe/ Nukibetsu/ Niikappu/ Kushiro/ Ochiho/ Tarantomari/ Maoka/ Shiraura/ Nairo (Sakhalin)</td>
<td>Hattori and Chiri (1960)</td>
</tr>
<tr>
<td>Biratori, Fukumitsu</td>
<td>Hattori and Chiri (1960), Hattori (ed.) (1964) and Tamura (1996)</td>
</tr>
<tr>
<td>Samani</td>
<td>Hattori and Chiri (1960) and Watanabe et al. (1985)</td>
</tr>
<tr>
<td>Obihiro</td>
<td>Hattori and Chiri (1960) and Sawai &amp; Tamura (2005)</td>
</tr>
<tr>
<td>Bihoro</td>
<td>Hattori and Chiri (1960), Hattori (ed.) (1964)</td>
</tr>
<tr>
<td>Area</td>
<td>References</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>---------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Chitose</td>
<td>Nakagawa (1995) and Watanabe et al. (1994)</td>
</tr>
<tr>
<td>Shizunai</td>
<td>Okuda (1999) and Watanabe et al. (1984)</td>
</tr>
<tr>
<td>Mukawa</td>
<td>Nakagawa (ed.) (2014)</td>
</tr>
<tr>
<td>Hombetsu</td>
<td>Sawai (2006) and Watanabe et al. (1987)</td>
</tr>
<tr>
<td>Shumushu (Northern Kuril)</td>
<td>Torii (1903) and Murayama (1971)</td>
</tr>
</tbody>
</table>

The Atlas of Ainu Dialects

**Hokkaido:**

**Sakhalin:**

**Kuril Islands:**
25. Shumshu

**References**


The sun in Korean

Rei Fukui

Department of Korean Studies, Graduate School of Humanities and Sociology, The University of Tokyo

Abstract

This paper discusses dialectal and derivational variation, history and etymology for words meaning the ‘sun’ in Korean. Native and Sino-Korean words are taken into consideration. Also the native word for the sun has a polysemic relationship with the temporal meaning ‘year’, which is peculiar to this language.

1 Introduction

In modern standard Korean two words are used to refer to the sun: one is the native word /hɛ/ (해) and the other a Sino-Korean word /tʰɛjan/ (태양, 太陽). Dialectal variation, history and etymology for these words will be discussed in this paper.

2 /hɛ/

This native word for the sun shows virtually no dialectal variation. In some dialects the vowel /ɛ/ is pronounced slightly different from the standard [ɛ], but aside from such minor phonetic details we have virtually no dialectal variation. Therefore drawing a map for this lexical item is almost of no use so far as the Korean Peninsula alone is concerned.

In Ogura (1944, 1st vol.: 3) the description for this item is as simple as the following:

(1) [hɛ]
(2) [pit-tʃʰi]
(The sun: (1) [hɛ] in general. (2) [pit-tʃʰi] HamNam Kapsan (slang used by wild Ginseng hunters.)

Here ‘in general’ means that the form in (1) is generally used throughout the Korean Peninsula. The only exception to this is the form shown in (2), which is apparently derived from the word /pich/ (‘light’) by adding the suffix /-i/. But as noted above this is a kind of slang used by wild Ginseng hunters often referred as ‘Simmani’,1 so that this form is by no means a normal variety of this word used in everyday life in certain fixed locations.

As to the quality of the vowel /ɛ/, two kinds of additional explanation is necessary. First, some dialects, typically the Kyeongsangdo dialects, are known to have lost the distinction between the vowel /e/ and /ɛ/, and the vowel in question is pronounced somewhere between the two vowels.

Secondly, Ogura Shinpei gives a slightly different description of this vowel for Cheju dialects (Ogura (1931a, 1931b, 1944: 2nd vol.)).

The origin of the vowel /ɛ/ in standard Korean goes back to one of the two diphthongs /ai/ and /ʌi/ in Middle Korean (abbreviated as MK hereafter). The Cheju dialects are the only ones that are (or, had been) preserving the distinction of these two diphthongs in the forms of two different monophthongs,

1 In some cases they are known to use borrowings from Tungusic languages but this case clearly shows a Korean origin.
while in other dialects the distinction is completely lost and both have merged into /ɛ/. What follows is his description of the vowel corresponding to the MK diphthong /æi/ ( jihadists) in this dialect (Ogura 1931b: 146).

His transcription of the vowel ð (umlaut italic o) is somewhat unusual (using italic is essential here but easy to be overlooked). In the framework of the current IPA system the vowel symbol that comes most closely to such a vowel would be [æ], front, half-open, rounded vowel.

Similar explanation is also found in Ogura (1931a: 32-33, also 1944: 2nd vol.).

As we have seen above, he gives us no further description than ‘[ɛ] in general’ in Ogura (1944: 1st vol.: 3). It may be the case that minor phonetic details are neglected in this expression.

Now let us turn our attention to the works on dialect survey made in Korea. Unfortunately, however, the word for the ‘sun’ is not included in the dialect dictionary compiled by Choi (1978) and in the linguistic atlas compiled by Lee et al. (2008). A large-scale dialect survey had been done in Korean in the 1980s and the results of which have been published in Han’guk Chŏngsin Munhwa Yŏn’guwón (ed.) (1987-1995) in 9 vols., but the item for the ‘sun’ is not included either. The lack of this entry in these works seems to be the result of the fact that there is very little variety for this word in Korean.

There is however another native word which can be used instead of /he/ ( 헤). It is formed by attaching the honorific suffix ‘-nim’ to it: /hennim/ ( 헤 달) or /hen(n)im/ ( 헤 달). This formation is quite similar to that of the Japanese word ‘Ohisama’ formed by attaching the honorific suffix ‘-sama’ and a prefix ‘o-’ to ‘hi’. But again we don’t have any data for these words in order to draw a linguistic map and, if any, it would be difficult to do so because in most cases /he/ and /he(n)nim/ can be used by the same speaker and are in a kind of stylistic variation, not a regional one.

Finally I would like to point out a cross-linguistically interesting question. The word /he/ has a homonym with the meaning of ‘year’. Synchronically it is possible to treat them as homonyms but there is a clear semantic relationship so that it is also possible to treat them as a single lexical item having two different meanings. As to the relationship between the sun and temporal notions, it is interesting to note

---

2 Although we find two forms for this item in modern Korean dictionaries, the actual pronunciation is invariably the latter /hennim/ in my experience.
that in many languages including Japanese and Chinese, it is related to ‘day’, but in Korean it is ‘year’. I have no knowledge of other languages of the latter type, but this poses cross-linguistically an interesting question. As compared to the sun, the moon seem to be more closely related to the temporal notion ‘month’ in much wider range of languages including Korean.

2.1 The history of the word /hʌi/

In Middle Korean the word corresponding to the modern /hʌi/ was /ʌi/ (abbreviation as H below). This word too had another meaning ‘year’ in MK, and the forms corresponding to these two meanings are completely homophonous including the tone. A few examples containing this word are shown below:

힌 므지개 현에 배니이나 <1447 Yongbi och’on ka (龍飛御天歌) 50>

“a white rainbow is penetrating the sun.”

禄은 아오미니 환 東鄉が 未し면 아오미오 <1459 Wörin sŏkpo (月印釋譜) 2:50a>

“(the character)禄 means morning, if the sun is in the east, it is morning...”

If we go back to the times before the Hangul script was invented, we have two documents containing the Korean word for the sun. One is the Cháoxiāngyuàn yìyù (朝鮮館訳語), compiled in the beginning of the 15th century, in which the item sun (日) is recorded as follows:

日害孳 (cf. MK_https 일 zil)

The Korean part of this item consists of two characters, of which the first character ‘害’ corresponds to the native Korean word and second character ‘忍’ the Sino-Korean reading of the character ‘日’. The sound of the character ‘害’ is quite similar to the MK form of the word.3

Another document is the Jīlin lèishi (鴨林類事), recorded in the early 12th century, in which the description for this item is complicated and apparently some kind of errors must be included.

方言天曰漢木奈
日曰姪
月曰契 [黑隄切]
雲曰屈林
風曰孛纏
……

(cf. 姍 héng 胡登切，姪 dá 当割切
契 qí 話計切，祎 xi 胡計切)

(Right: a part of the text Jīlin lèishi (鴨林類事)
included in the Shùnzhì (順治) 4 (1647) edition of
the Shuofu (說郛), vol. 55)

3 The reason for selecting this character may have something to do with the fact the this character has the qusheng (去声) and the tone of the word ‘히’ (hʌi) is H.
The item ‘sun (᪥)’ appears as the second entry of the list, followed by the third entry ‘moon (᭶)’. But the two entries seem to be wrongly interchanged. Chin T’ae-ja (陳泰夏, 1974: 250-251) explains this as follows:

(The two characters ‘姮’ and ‘契’ seem to have been wrongly interchanged. If we take ‘姮’ as ‘祲’ part of it being omitted, then fanqie of this character being ‘胡計’, its reading matches well with MK [hei]. And if we take ‘姮’ as a miswritten ‘姐’, then its reading matches well with MK [tel]. ....... Accordingly we can assume that the note ‘黑陰切’ in fact represent the sound for [hei] and must have been placed under the entry for ‘日 (the sun)’. Translation mine.)

It is very complicated but this interpretation is a plausible one and many others seem to agree with this (for example, Kang, Sin-hang (1980)). Therefore it can be concluded that this word must have had almost the same phonetic shape as the Mk form in the 12th century.

If we go back further to the Silla period we don’t have any example of this word used in the sense of the sun in Hyangga (郳歌), although we have an example used in the sense of ‘year’ in ‘龜竹旨郳歌’ but in this case the lexical meaning was expressed using the Sino-Korean ‘年’ so that its pronunciation is unknown.

To sum up, the word shape of this word was /hʌi/ with a high tone in MK, and it must have sounded very similar to that in the 12th century, but we don’t have data older than that.

2.2 Etymology of /hɛ/

As to the etymology of this word, there have been several theories advanced among which the following one, found in an etymological dictionary, calls for attention.


(he (the sun) (n.) tʰɛjaŋ. [Etymology: √ hʌi- [white]. Change: hʌi (Yongka 7:1) > hɛ])

As is shown here, one form of the adjective with the meaning of ‘white’ was ‘hʌi-’ in MK, but it had an alternate form ‘hii-’ so that I am not sure about this etymology.

3 Sino-Korean word /tʰɛjaŋ/ (太陽)

This Sino-Korean word must have been known for hundreds of years before the Hangul script was invented in the 15th century. Its usage can be attested in many kinds of Hanmun (漢文) documents. But they are a kind of Koreanized Chinese writings so that it is difficult to judge how much this particular word was used as part of the daily vocabulary at the time.

Aside from documents written in Hanmun, its usage in native Korean sentences written in the Hangul script is very scanty until late 19th century. The following example is the oldest one so far I have found and is a unique example from 15 through 16 centuries.

변화 희(음가 7:1) > 헴

“The brightness follows the sun and the darkness follows the black moon”
But this usage of ‘太陽’ is a translation of a passage from sutra written in Chinese (“明從太陽暗隨黑月”), so that it is difficult to decide whether or not this word was used in daily life in the 15th century. Although other sporadic usages are found from 17 through 18 centuries (but less than 5 examples in total), it is as late as the late 19th century that we see plenty of examples of this word, many of the early examples being from documents on Christianity like the following:

예수 엽골에 광칙 발효 아 태양 광황시고 <1865 主年瞻禮廣益 9b>
jeisu erkor-ei kwaneʌli parhʌ-ja tʰaiʌŋ kʌlʌ-si-ko
“Jesus’s face shone like the sun …”

비유컨대 거울이 태양의 빛을 빛을 때태양의 태와 빛치 나타나 <1883 真教切要 26b>
piju-keʌntai ke’ur-i tʰaiʌŋ-i piʌb-ʌr pat-imai tʰaiʌŋ-i tʰjei-wa piʌb-i natʰana
“it is as if a mirror reflects the sun and the sun’s body and the light appear ...”

Therefore it is likely that although this Sino-Korean word have been known for many centuries but it was in the late 19th century that this word began to be used as part of the daily vocabulary probably under the influence of the introduction of western culture.

4 Conclusion

The native Korean word for the sun has very little dialectal variation and has been steadily used at least from the 12th century. This word is also used with the meaning of ‘year’, contrary to many other languages in which this word has closer relationship with the temporal meaning ‘day’. So far I know of no other language which has a semantic relationship between the sun and year, so that the existence of such a language poses a cross-linguistically interesting question. Also a personified honorific variety of this word can be made by attaching the suffix /-nim/, and this too can be cross-linguistically interesting, since similar phenomena are found in other languages such as Japanese and Chinese. The etymology of this word is unclear but it might be related to color terms.

Another Sino-Korean word /tʰjʌŋ/ seems to have been known for a long time but it is until the late 19th century that this word began to be used in the daily life.

Based on these observations I would like to propose the following list of cross-linguistic questions:

(1) Existence of personified varieties.
(2) Existence of a temporal meaning (day, year).
(3) Relationship with color terms (white, yellow, red, etc.)

Acknowledgements

I would like to thank Professor Mitsuaki Endo for inviting me to this conference. Otherwise I could not have noticed the rarity and importance of the relationship between the ‘sun’ and ‘year’ found in this language.

References


Chasing a Cat from the Mekong to the Salween: A Geolinguistic Description of ‘Cat’ in Trung and Khams Tibetan in North-western Yunnan

Liying Qin\textsuperscript{a} and Hiroyuki Suzuki\textsuperscript{b}

\textsuperscript{a}School for the Study of the Minorities, Yunnan University / Kunming University of Science & Technology
\textsuperscript{b}IKOS, University of Oslo / National Museum of Ethnology

Abstract

This brief essay provides a linguistic map that describes the lexical distribution and spread of the word form ‘cat’ which includes a nasal initial (/n/ or /n/) of Khams Tibetan (sDerong-nJol Kham) and Trung, spoken in Deqin, Weixi, and Gongshan counties located in north-western Yunnan. Through a geolinguistic analysis, it concludes that the form for ‘cat’, with a nasal initial attested both in Trung and Khams Tibetan, is not an inherited word, but possibly a loan from a certain language spoken alongside the Nujiang River.

1 Linguistic overview of Trung and Khams Tibetan in Gongshan County

This article uses a geolinguistic methodology to examine the distribution of the word form for ‘cat’ in Trung and sDerong-nJol Khams Tibetan spoken in the three counties Gongshan, Deqin, and Weixi, located in north-western Yunnan. Although both languages have a similar word form regarding ‘cat’, due to the lack of information on their geographical distribution, it has been difficult so far to discuss the mutual relationship between these two languages. The article provides a preliminary geolinguistic analysis of the issue using first-hand data.

The focus of this essay is mainly the languages of Gongshan County. We will thus present an overview of two languages of Gongshan: Trung and Khams Tibetan, including geographical distribution, language situation, dialectal difference, and phonological system. Regarding Trung, since few references explain its dialectal differences, we provide a detailed description below.

1.1 Trung

The Trung people are one of the cross-border nationalities with small population in China. They are distributed in the Gongshan Dulong and Nu Autonomous County of the Nujiang Lisu Autonomous Prefecture (\textit{Dulongzu Jianshi} 1986:1). Most Trung people live along the banks of the Dulongjiang River, and Xiaochala Mountain. Bingzhongluo Township along the Nujiang River is also a Trung settlement. A few Trung people are scattered in Qile Village, Weixi County, in Yunnan Province, and Chawalong Township, Chayu County in the Tibetan Autonomous Region (TAR). Within Myanmar, people who speak the Trung language (possibly up to 100,000 people) live in northern Kachin State (LaPolla 2003). According to the 2000 census, China has 7426 Trung people. The origin and migration of Dung people has been discussed by \textit{Nujiang Juizhi} (1998), Yang and Li (2010), Gao (2009), Wang (2011), He and He (2007), Sun (2013), Liu (2009), and so on.

Trung is a Tibeto-Burman language. The geolinguistic location of Trung is similar to its geographical distribution: in the east it is close to the Loloish languages, in the south it is adjacent to Rawang in Burma, in the west there are a multitude of languages in the Himalayan Massif, and in the north it adjoins the Tibetan language area (Huang 1997). The affiliations of the language are still unclear. Scholars tentatively put it under the Jingpo branch of Tibeto-Burman. Sun Hongkai (1983) classified the
Trung language in China into two dialects: Dulong River (Dulonghe) dialect of Trung and the Nujiang dialect of Trung. The latter is spoken by the Nu people in Gongshan who call themselves “Nung”. However, according to local people’s opinions and the first author’s research on Trung vocabulary, the diversity between vernaculars in Dulong River (the vernaculars in the upper, middle and lower reaches of Dulong River) is much bigger than that between the Dulong River dialect and the Nujiang dialect. Therefore, that classification does not reflect the internal divergence of Trung along the Dulong River. In fact, according to the specific situations of the Trung language in the upper, middle, and lower reaches in Dulong River, the Trung language can be classified into four dialects based on the areas where it is spoken:

<table>
<thead>
<tr>
<th>Based on the administrative names</th>
<th>Local simplified name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dizhengdang &amp; Longyuan</td>
<td>Yixiang ‘first township’ &amp; Longyuan</td>
</tr>
<tr>
<td>Kongdang</td>
<td>Sanxiang ‘third township’</td>
</tr>
<tr>
<td>Maku</td>
<td>Sixiang ‘fourth township’</td>
</tr>
<tr>
<td>Nujiang dialect (according to Sun 1983)</td>
<td>Shuangla</td>
</tr>
</tbody>
</table>

Local names are generally used by the Trung people and also appear in LaPolla (2000). However, they are nonexistent as administrative toponyms. And although some names include the word ‘xiang’, normally translated as ‘township’ in English, in this case it designates an administrative village level. Therefore, in this article we consistently use the administrative names.

Except tone differences, the Dizhengdang dialect (especially the Trung in Longyuan Village) is very close to the Nujiang dialect of Trung. In Randy J. LaPolla’s discussion (1997), the Nujiang dialect of Trung seems to be the same as the Kwinpang dialect spoken in Myanmar, which should thus be considered a dialect of Trung. However, Nung people claim that their Nung language is different from Trung, even though they know Nung is similar to Trung. See also the description of 1.2.

Since these places inhabited by Trung people are hard to get to and people there rarely make contact with the outside world, the proportion of monolinguals is rather high. The transmission of Trung is natural, and it is the major medium of communication for local people. Chinese and/or Lisu languages are Trung people’s second language. In addition, due to the close contact with the Tibetans and Lisu people in old days, the Trung language in the upper reaches of Dulongjiang River and in the lower reaches is respectively influenced by Tibetan language and Lisu language. Of loan words, 80% are Chinese loans, 10% are Tibetan, and 5% are Yi (LaPolla 1987). According to UNESCO’s nine criteria and Chinese experts’ 6-scale criteria used to test language vitality, Trung is at the 2nd level (vital or still active) (Sun 2006). Moreover, according to Daniel Nettle’s calculating standard (1999), the existence and maintenance of Trung in the Dulong River area is still at a safe stage.

The following are some words extracted from the 400 core words which can reflect the divergence of the dialects (vernaculars) between the upper (Dizhengdang & Longyuan hamlets), middle (Kongdang Hamlet) and lower (Maku Hamlet) reaches of Dulong River.

Through interviews we learned that the Trung language in the Longyuan hamlet (upper reaches of Dulong River) differs greatly from other places’ Trung language, and is much more similar to the Shuangla vernacular of Trung ‘Nujiang dialect’.

<table>
<thead>
<tr>
<th>Kongdang</th>
<th>Buer-Longyuan</th>
<th>Nung/Shuangla</th>
<th>Xiaoachala</th>
<th>Maku</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>na³³</td>
<td>a³⁵ go³⁵</td>
<td>gu³³</td>
<td>na³³</td>
<td>na³³</td>
<td>I</td>
</tr>
<tr>
<td>a³⁴ me³³</td>
<td>a³⁴ ma³³</td>
<td>a³⁴ me³³</td>
<td>a³⁴ me³³</td>
<td>a³⁴ me³³</td>
<td>mother</td>
</tr>
<tr>
<td>a³⁴ pe³³</td>
<td>a³⁴ pe³³ (B)²</td>
<td>a³⁴ pa³³</td>
<td>a³⁴ pe³³</td>
<td>a³⁴ pe³³</td>
<td>father</td>
</tr>
</tbody>
</table>

1 Nung people who live along the lower reaches of the Nujiang River call themselves /nuŋ³¹ teŋ³⁵/.
2 This form is similar to Rawang spoken in Myanmar.
In addition, there is a large amount of synonyms in the upper and middle reaches of the Dulong River. Also, there exists phonetic correspondence between these words. For example, the consonant /p/ in the 3rd Township corresponds to /bx/ and /b/ in the 4th Township, the same as the correspondence of /b/ to /z/; the vowels /u, uu/ in the 3rd Township corresponds to /i/ in the 4th Township; the consonant /c/ corresponds to /k/; /m/ corresponds to /n/ and /ŋ/.

The phonological inventory of Trung (vernacular of Kongdang3) is as follows:

**Table 3: Consonantism of Kongdang Trung.**

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
</tr>
</thead>
<tbody>
<tr>
<td>plosive</td>
<td>voiceless</td>
<td>p</td>
<td>t</td>
<td>c</td>
<td>k</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>voiced</td>
<td>b</td>
<td>d</td>
<td>j</td>
<td>g</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>affricate</td>
<td>voiceless</td>
<td>ts</td>
<td>te</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>voiced</td>
<td>dz</td>
<td>dz</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fricative</td>
<td>voiceless</td>
<td>s</td>
<td>e</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>voiced</td>
<td>z</td>
<td>z</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>nasal</td>
<td>voiced</td>
<td>m</td>
<td>n</td>
<td>n̄</td>
<td>j</td>
<td>j</td>
<td></td>
</tr>
<tr>
<td>liquid</td>
<td>voiced</td>
<td>l</td>
<td>r</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>semi-vowel</td>
<td>voiced</td>
<td>w</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>j</td>
</tr>
</tbody>
</table>


The Trung language has 28 consonants and 14 consonant clusters: /pl, bl, ml, kl, gl, pu, bu, mu, ku, gi, xi, m2, n2, n̄2/. The consonants /p, t, k, ?, m, n, n̄, l, r/ often occur in final position.

**Table 4: Vocalism of Kongdang Trung.**

```
i  uu
|e     |
a
```

3 B=the form of Buer; /iŋ/5 for Longyuan

4 B=the form of Buer; /nəŋ/5 for Longyuan

5 B=the form of Buer; /taŋ/ for Longyuan

6 B=the form of Buer; /kəŋ/ for Longyuan

7 The current township government of Dulong River is located at Kongdang, therefore, Trung people who live there take the Kongdang dialect as the lingua franca of Trung and the “Pinyin Plan of Trung” is made upon it.
The vowel length (short/long) is distinctive. Nine diphthongs are attested: /ai, si, ui, uii, ua, a:i, o:i, u:i, u:ii/. Tones of Kongdang Trung:
A three-way distinction in word tone: high level [55], falling [53], low falling tone [51].

1.2 Khams Tibetan
The Khams Tibetan variety spoken in Gongshan County, e.g. Bodgrong Tibetan (Suzuki 2014ac), belongs to the sDerong-nJol dialect group. Bodgrong Tibetan is spoken by Tibetans and Nu-nationality people living in the central area of Bingzhongluo [Bod-grong] Township. This township adjoins Chawalong [Tsha-ba-rong] Township, Chayu County of the TAR and Yunling [Lung-gling] and Yanmen townships of Deqin [Jol] County, Diqing [bDe-chen] Prefecture, both of which belong to the Tibetan cultural area. In Nujiang, Tibetan dialects are found only in Bingzhongluo and Bangdang townships, and they are a minority language in this area, where Lisu, Nung (a.k.a. Anu, regarded as a dialect of Trung), and Chinese are also spoken. And whereas Lisu has played a role as lingua franca, this role is currently being replaced by Chinese. Dialectal divergence within the two villages is to some extent attested. There are at least three varieties: Bodgrong (Bingzhongluo [Bod-grong] - ‘luo’ is a Lisu word designating ‘place’), Chunathang (Qiunatong [Chu-nag-thang]), and Dimalo (Dimaluo).

According to local oral tradition, the Tibetans living in Nujiang have migrated from gYanggril (Yongzhi [Glang-sgril], Yunling) and Tshedrug (Cizhong [Tsho-drug], Yanmen) villages in the present Deqin County several generations and around 200 years ago. On the other hand, no specific relation between Bodgrong and Tshawarong (Chawalong [Tsha-ba-rong]) has been attested.

According to native speakers in Ridang Hamlet, Bingzhongluo, it used to be ordinary that they are multilingual of Khams, Lisu, Yunnanese (a variety of Southwestern Mandarin), Nung, and Trung. Still now, most of them are trilingual of Khams, Lisu, and Yunnanese. A noteworthy thing is that the language data described in Suzuki (2014c) was obtained from Nu nationality people. Not a few Tibetan-speakers in Bingzhongluo are officially registered as Nu nationality. Hence they do have to some extent competence of Nung, which may influence the formation of the local Khams vernacular. They also consider the Nung language to be different from Trung but quite similar to it. They have a frequent contact with Trung people living in Xiaochala (see 1.1).

The phonological inventory of Bodgrong Tibetan (vernacular of Rithang) is as follows: 9

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
</tr>
</thead>
<tbody>
<tr>
<td>plosive</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>aspirated</td>
<td>pʰ</td>
<td>tʰ</td>
<td>kʰ</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>non-aspirated</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>voiced</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>affricate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>aspirated</td>
<td>tsʰ</td>
<td>teʰ</td>
<td>cʰ</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>non-aspirated</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>voiced</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fricative</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>aspirated</td>
<td>sʰ</td>
<td>eʰ</td>
<td>xʰ</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>non-aspirated</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>voiced</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>nasal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>voiced</td>
<td>m</td>
<td>n</td>
<td>ŋ</td>
<td>ŋ</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>voiceless</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>liquid</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>voiced</td>
<td>l</td>
<td>r</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>voiceless</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>semi-vowel</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>voiced</td>
<td>w</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

8 Each name is given with a Written Tibetan (WrT) form in square brackets. This essay consistently uses the pinyin name for each vernacular.
9 See Suzuki (2014c) for a detailed description.
Table 6: Vocalism of Bodgrong Tibetan.

<table>
<thead>
<tr>
<th>i</th>
<th>u</th>
<th>uu</th>
</tr>
</thead>
<tbody>
<tr>
<td>e</td>
<td>o</td>
<td>o</td>
</tr>
<tr>
<td>a</td>
<td>a</td>
<td>a</td>
</tr>
</tbody>
</table>

Tones of Bodgrong Tibetan
A four-way distinction in word tone. The following phonemic signs will be used at the beginning of a word:


2 Data and phonetic description of the word form

2.1 Research sites
We have collected data from 42 locations, of which 7 varieties are Trung, 2 Nung, and 33 Khams Tibetan affiliated with the sDerong-nJol dialectal group. The data reflected in the linguistic maps are limited to our first-hand materials for the sake of the consistency of the phonetic description. The list of vernaculars is following:

Table 7: Research sites.

<table>
<thead>
<tr>
<th>Township, County</th>
<th>Hamlet (language name)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dulongjiang, Gongshan</td>
<td>Buer, Dizhengdang, Longyuan, Kongdang, Mabilidang, Bapo, Maku (Trung)</td>
</tr>
<tr>
<td>Bingzhongluo, Gongshan</td>
<td>Xiaochala (Trung)</td>
</tr>
<tr>
<td></td>
<td>Shuangla, Gongka (Nung)</td>
</tr>
<tr>
<td></td>
<td>Ridang [Ri-thang] (Khams Tibetan)</td>
</tr>
<tr>
<td>Bangdang, Gongshan</td>
<td>Dimaluo (Khams Tibetan)</td>
</tr>
<tr>
<td>Badi, Weixi</td>
<td>Jieyi [sBrul-yul], Luotong [Lo-thang] (Khams Tibetan)</td>
</tr>
<tr>
<td>Yanmen, Deqin</td>
<td>Badong [dPa’-gdong], Cizhong [Tsho-drug], Siga [Sa-dkar], Gongniang [sGo-gnyan],</td>
</tr>
<tr>
<td></td>
<td>Chunduole [Chu-mdo-log], Nitong [sNyin-thang], Guzha [sGo-grags], Yeka [Yar-kha],</td>
</tr>
<tr>
<td></td>
<td>Muda [Mo-rtags], (Khams Tibetan)</td>
</tr>
<tr>
<td>Yunling, Deqin</td>
<td>Yongzhi-2 [gLang-sgril], Yongzhi-3 [gLang-sgril], Chalitong [Tsha-re-thang],</td>
</tr>
<tr>
<td></td>
<td>Chaliding [Tsha-re-steng], Hongpo [dNgul-phung], Jiunongding [ICang-nang-steng],</td>
</tr>
<tr>
<td></td>
<td>Balida [Ba-ri-steng], Guonian [sGo-nyan], Jibi [ICags-spel], Yubeng [gLegs-sbam], Xidang [Shar-thang], Mingyong [Me-long] (Khams Tibetan)</td>
</tr>
<tr>
<td>Shengping, Deqin</td>
<td>Adunzi, Wunongding [mGo-nang-steng], Niangyi [Nyang-yas], Gongda [kKang-rtags],</td>
</tr>
<tr>
<td></td>
<td>Zhiren [Bri-zhing], (Khams Tibetan)</td>
</tr>
<tr>
<td>Foshan, Deqin</td>
<td>Foshan, Jiangpo [ICang-phud] (Khams Tibetan)</td>
</tr>
</tbody>
</table>

Other than these, a variety of Trung, Lula, spoken in the Lula hamlet near Kongdang, was also recorded, but it is not included the linguistic map because its data awaits confirmation.
2.2 Phonetic description of the word form

The following table 8 is a phonetic description (segmental part only\textsuperscript{10}) of the word form of ‘cat’ in each variety. The tonal description is uniformly omitted in order to provide a classification of word forms.

Table 8: Phonetic description of ‘cat’.

<table>
<thead>
<tr>
<th>Type</th>
<th>Segmental form</th>
<th>Distribution of hamlets (Language name in parentheses)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A-1</td>
<td>ɲa me / ɲa mje</td>
<td>Shuangla (Nung); Chaliding, Chalitong, Muda, Yeka, Nitong, Guzha, Gongnianq, Siga, Chunduole, Cizhong, Badong, Jieyi, Luoyi, Ridang, Dimaluo (Khams Tibetan)</td>
</tr>
<tr>
<td>A-2</td>
<td>na me / na me</td>
<td>Longyuan, Kongdang, Mabilidang, Bapo*, Xiaoachala* (Trung),</td>
</tr>
<tr>
<td>B</td>
<td>o li / a li</td>
<td>Buer, Dizhengdang, Bapo*, Maku, Xiaoachala* (Trung); Gongka (Nung)</td>
</tr>
<tr>
<td>C-1</td>
<td>li la</td>
<td>Shengping, Wunongding, Niangyi, Gongda, Zhiren, Foshan, Jiangpo, Mingyong, Xidang, Yubeng, Yongzhi-2* (Khams Tibetan)</td>
</tr>
<tr>
<td>C-2</td>
<td>lu lu / lu lu</td>
<td>Jiabi, Balida, Guonian, Jiunongding, Hongpo, Yongzhi-2*, Yongzhi-3 (Khams Tibetan)</td>
</tr>
</tbody>
</table>

The data points with an asterisk means that they use two forms. They do not designate different semantic categories of ‘cat’ (e.g. ‘domestic cat’ and ‘wild cat’), but mainly reflect a difference of the speakers’ generation. The case of Bapo and Xiaoachala is that the B form (/o li/) is principally used by elder people, and the A-2 form, by younger and middle-aged speakers. In the data from Kongdang, which also uses the A-2 and B forms, the B form is mostly spoken by people from Dizhengdang or Maku who work or do business in Kongdang. This is the reason why we do not regard the Kongdang dialect as a dialect with two forms in the map; the variation is due to sociolinguistic factors.\textsuperscript{11}

The case of Yongzhi-2 is, on the other hand, unclear in terms of the use of each form (C-1 and C-2). Noticing the Yongzhi-3 dialect, spoken in the village next to Yongzhi-2, situated in an area above it, we can see that only the C-2 form is used. It may imply that the C-2 form is a more localised word. We will see the geographical distribution in greater detail in the next section.

3 Map and analysis

Based on the phonetic description in 2.2, we will draw a linguistic map and describe a geographical distribution and its features.

3.1 Linguistic map

Map 1 is an overall distribution of the word form of ‘cat’ in Trung and Khams Tibetan spoken in Northwestern Yunnan, designed with the online Geocoding mapping method provided by the site: http://ktgis.net/gcode/lonlatmapping.html.

\textsuperscript{10} Suprasegmental phenomena are hereby not considered.

\textsuperscript{11} Whether a linguistic map reflects sociolinguistic variations or not depends on the purpose of a geolinguistic analysis. In this article, same as in most cases of geolinguistic analysis, sociolinguistic differences are not dealt with if the current sociolinguistic situation is evident. See also Suzuki (2015) and Suzuki and Sonam Wangmo (2016).
3.2 Analysis

The word forms for ‘cat’ are principally divided into two categories as displayed with a ‘pin’-type (forms including a nasal initial; A-1 and A-2) and a ‘circle’-type (forms including a lateral initial; B, C-1, and C-2) on the map. However, the green ‘pin’ indicates dialects with two word forms. These dialects are of Trung only, and each of Kham Tibetan dialects has one form. Seen from a distribution of each word form, the A-1 and A-2 forms are attested in the central area of the map. In Trung, in both the northern and southern edges, the B-type is employed, where the word form for ‘cat’ demonstrates the so-called ‘ABA distribution’ in the geolinguistic method (See Iwata 2010). It implies that the A-1 and A-2 forms are newly developed or acquired ones, and that the form distributed in its periphery is thus usually analysed as a more archaic one. In addition, the main pathway just exists from the riverside of Nujiang to the central area of Dulongjiang Township, which can also be regarded as a factor of language
change. On the other hand, the data of sDerong-nJol Khams does not show the ABA distribution for ‘cat’; however, its language area is already at the southernmost tip, beyond which no Tibetic languages, are spoken.\textsuperscript{12} Comparing the map with a wider perspective regarding ‘cat’ (see Suzuki 2014b), the limited distribution of the A-2 form to the area demonstrated in Map 1 implies a loan from a surrounding language.

Then, we describe the case of two Trung dialects (vernaculars) which possess two forms for ‘cat’ in more detail. The first author’s research has found that the vernaculars spoken from Longyuan to Bapo along Dulong Rivier can use two forms, in which the /na me/-type (A-2) is much more frequently used than the /lo/-type (B). According to the information provided by a native of Mabilidang, the /lo/-type (B) existed in the speech of elder people, and, perhaps, current elders do not use it but the /na me/-type (A-2) instead. This description makes it more interesting that the vernacular of Xiaochala, spoken by Trung people who lived on the mountain along Nujiang River, also has two forms for ‘cat’. The Trung-speakers in this village are descendants of immigrants from Kongdang Township in 1950s. People lived isolated from the rest of the Trung-speaking area as a language island; hence they maintain the language situation at the moment that their ancestors’ dialect, in which the two forms for ‘cat’ might already exist. Notably, the form for ‘cat’ in Nung (the vernacular of Shuangla) is the /na me/-type (A-1), not the /na me/-type (A-2). This sound form, A-1, is common to Khams (Bodgrong Tibetan), and this situation suggests that Nung has to some extent been influenced by Khams, which is discussed in articles about Nung’s origin and migratory route (Liu 2009, Yang 2010, Wang 2011, Sun 2013). Meanwhile, it is also noticeable that some Nung-speakers also speak Bodgrong Tibetan (see 1.2). Another Nung form (the vernacular of Quinatong) for ‘cat’ (B) is also interesting from the viewpoint of geographical distribution. Because of the lack of data of the Chunathang dialect of Bodgrong Khams, it cannot be asserted that Quinatong Nung and Chunagthang Khams have the same form, which would indicate a mutual influence between them.

Next, we examine the geographical distribution of the form for ‘cat’ in sDerong-nJol Khams in more detail. As Suzuki (2014b) shows, the distribution of the A forms is limited in the area demonstrated in Map 1, whereas the B and C forms are attested to everywhere in the eastern Tibetosphere. Suzuki (2014b) does distinguish C-1 from C-2 on the maps present in that article, neither does it explain whether they two are cognates or not. As far as the geographical distribution, each form of the C type to some extent has its own field, and in Map 1, the C-2 form exhibits much limited distribution. On the other hand, the C-1 form is not only distributed in a wider range than C-2, but also occurs further north, in such regions as Batang County (Suzuki 2014b). The question regarding the distribution of the C forms within Map 1 should be an independent appearance of the C-1 form in the Yongzhi dialect. As the role and position of Yongzhi is crucial for this article, a detailed analysis is provided later. In the area of Map 1, the B form does not appear within Khams Tibetan dialects, however, the dialects belonging to the Sens-kyi-nyila group, mainly spoken in Map 1’s eastern neighbour Shangri-La Municipality, use the B form.\textsuperscript{13} However, Trung does not have any occasion to contact dialects of the Sens-kyi-nyila group, the B-form attested to in these two languages has independently been developed from each other. The A form is just peripheral, but to some extent has a certain geographical continuity of distribution. In addition, it is shared by the dialects spoken alongside Lancangjiang and Nujiang.

As introduced in 1.2, Khams Tibetan spoken in Gongshan County (Nujiang) seems to be genetically related to that spoken in Deqin County (along Lancangjiang). There are at least three main pathways to connect the one to the other: Quinatong-Yongzhi (3 to 4 days’ walk), Dimaluo-Yongzhi (3 days’ walk\textsuperscript{14}), and Dimaluo-Cizhong (1 to 1.5 day’s walk). All of these routes are used by local people, especially local Tibetans. Hence, we can easily find numerous dialectal similarities and shared lexical items between Khams varieties along the Lancangjiang and Nujiang. Then, a question emerges: why does the Yongzhi dialect use the C-1 type regardless of its geographical position and appear to have a closer relationship

\textsuperscript{12} The next place where a Tibetic language is spoken is Gagatang, located near to the administrative centre of Weixi County (see the maps of Suzuki 2014b, 2015). Lisu is a dominant language between these two areas.

\textsuperscript{13} It appears most frequently in the eastern Tibetosphere from the viewpoint of geographical distribution.

\textsuperscript{14} At present a motorway is being constructed between these two hamlets, and it will become the first motorway between Deqin and Gongshan counties.
with Bodgrong Tibetan? Unfortunately, we have no sufficient evidence to speculate it. It is certain that
the Yongzhi dialect is considered by local Tibetans outside Yongzhi Hamlet as a somewhat divergent
speech form.\(^{15}\) Indeed, the distribution of the word form for ‘cat’ in Yongzhi seems curious seen from a
geolinguistic view. However, if the A forms are not an inherited word but a loan, an explanation would
be simple: the Yongzhi dialect maintains an older form, which has now replaced by elsewhere. If this
hypothesis is true, the A forms came from the Nuijiang area, and entered the sDerong-nJol-spoken area
alongside Lancangjiang through the three paths. It implies that Bodgrong Tibetan, which is derived
from the Yongzhi dialect and Cizhong dialect (see 1.2), could have played a role as a donor language
regarding the A forms for ‘cat’, which has not originally been attested in these dialects. Thinking of this
point, we suppose that Bodgrong Tibetan also borrowed the A forms from a non-Tibetic language at a
certain period, which should be after the Tibetans’ immigration to Bingzhonghuo, i.e. 200 years ago, at
earliest.

As mentioned earlier, the A forms attested in Trung might be a loan, which possibly originates from
non-Trung languages spoken alongside Nuijiang. So are those in Khams. In consequence, the A forms
(A-1 and A-2) are likely to have a donor language which is neither Khams Tibetan nor Trung. The terms
for ‘cat’ in surrounding languages which are not displayed on Map 1 are as follows:

Lisu: /a nia/\(^{16}\) (Gongshan dialect), /a\(^{31}\) ni\(^{33}\) za\(^{33}\)/ (Lushui dialect\(^{17}\)), /xwa\(^{55}\) lž\(^{22}/18\) (Weixi-Kangpu
dialect)

Nusu: /mu\(^{35}\) ni\(^{31}/19\)

Anong: /mu\(^{35}\) ni\(^{31}/20\)

With the examples displayed above, it is difficult to claim to which language(s) Trung and Khams
Tibetan are related. However, Lahu Xi, a Loloish language spoken in Xishuangbanna and the area south
to it, also has /na\(^{33}\) mi\(^{33}\)/ for ‘cat’ (YS59 1998). Thus, the forms attested in Trung and Khams might have
a donor language which is close to Loloish in the point of word form of ‘cat’.

Nevertheless, questions still remain: (1) How have Trung Khams Tibetan dialects acquired different
initials (/nj or /n/) for the word ‘cat’? (2) Is the B form in Trung an inherited word or a loan? For the first
question, the linguistic map tells us nothing. The consonantism in each language taken consideration.
We cannot directly know the background of this sound correspondence. There is still a possibility to
consider that this word originated as an onomatopoeia. A way to solve this problem would be
hamlet-to-hamlet research (Dawa Drolma & Suzuki 2015) of whichever languages are spoken in a given
area, which must be Gongshan County here.\(^{21}\) For the second question, we must examine the word form
for ‘cat’ spoken in Chayu, TAR, north of the Trung-spoken area, which may establish connection
between the A and B forms.

4 Conclusion

This article discussed the word form of ‘cat’ in Trung and sDerong-nJol Khams Tibetan dialects spoken
in north-western Yunnan from a geolinguistic view, with an introduction to Trung and Bodgrong
Tibetan. The linguistic map of ‘cat’ shows that the forms such as /na me/ or /na me/ are an newly
acquired forms from an unspecified donor language, which must be or have been spoken alongside the

\(^{15}\) The reason why the Yongzhi dialect is well known to outsiders is because Yongzhi Hamlet is a starting point of
circumambulation of Kha-ba dKar-po, one of the great sacred mountains in the Tibetan cultural area, and many
pilgrims come through here. Meanwhile, Tibetans in Yongzhi also generally know that surrounding dialects of
sDerong-nJol Khams uses the A-1 form for ‘cat’, which is different from theirs. This fact should be considered,
however, we lack data enough to analyse linguistic features.

\(^{16}\) A suprasegmental description is omitted in the forms of the Gongshan dialect.

\(^{17}\) From YS59 (1998).

\(^{18}\) This form is similar to the Naxi form (YS59 1998).

\(^{19}\) From Sun & Liu (2009).

\(^{20}\) From Sun & Liu (2009).

\(^{21}\) A similar issue regarding the nasal initial is attested in Khams Tibetan spoken alongside Lancangjiang. See
Suzuki (2009).
Nujiang. Both Trung and sDerong-nJol Khams have a word form including the /l/-sound (e.g. /s li/ and /li la/), however, they two are unlikely to possess the same origin. But other Khams Tibetan dialects do have a /s li/-like form, which thus implies the Trung word for ‘cat’ is related to the Tibetan spoken form. The lack of data from the TAR (especially Chayu County) is a crucial problem at present, but when this is resolved, we will obtain a more reliable interpretation regarding the distribution of the word form for ‘cat’.

Acknowledgements

Our thanks go to the friends from north-western Yunnan who provided language data for the authors. We should also like to thank Gerald Roche for checking our English and contributing insightful comments.

A part of the fieldwork for collecting data of Khams Tibetan was funded by Grants-in-Aid for Scientific Research of Japan Society for the Promotion of Science: “Linguistic Substratum in Tibet” headed by Yasuhiko Nagano, No. 16102001; “Dialectological Study of the Tibetan Minority Languages in the Tibetan Cultural Area in West Sichuan” headed by the present author; “International Joint Survey of rGyalrongic Languages” headed by Yasuhiko Nagano, No. 21251007; “Study on the Dialectal Development of Tibetan Spoken in Yunnan, China, through a Description of the Linguistic Diversity” headed by the present author, No. 25770167.

References


Preliminary Report on the Darmdo Minyag Linguistic Area, with a Geolinguistic Description of Terms for ‘Sun’

Dawa Drolma\textsuperscript{a} and Hiroyuki Suzuki\textsuperscript{b}

\textsuperscript{a}University of Virginia / Minzu University of China

\textsuperscript{b}IKOS, University of Oslo / National Museum of Ethnology

Abstract

This brief essay attempts to describe a dialectal difference of the word ‘sun’ attested in Darmdo Minyag, generally known as the western dialectal group of Minyag, by demonstrating the linguistic area of this language. Forty-four vernacular forms from the Minyagrong Valley are recorded here, of which 34 are counted as varieties of Darmdo Minyag, and the rest, Kham Tibetan. The lexical variety among terms for ‘sun’ in Darmdo Minyag is small, especially in terms of tonal variation.

1 Introduction

This essay aims to provide an overview of the present Darmdo Minyag linguistic area, and attempts to describe a dialectal difference of the word ‘sun’ attested in this language. Darmdo Minyag is generally known as the western dialectal group of Minyag,\textsuperscript{1} a member of the Qiangic languages, mainly spoken inside the valley called Minyagrong,\textsuperscript{2} between Jiagenba Village of Kangding Municipality and Tanggu Village of Jiulong County, both in Ganzi Tibetan Autonomous Prefecture, Sichuan Province, China.\textsuperscript{3} Multiple scholars have already described this language in several articles, and wordlists have already been published, including Huang (1985, 2007b), Sun (1983), Ikeda (1998, 2002, 2006, 2007), ZYC (1991), and TBL (1992). However, a dialectological study of Darmdo Minyag has not been conducted so far, and hence we cannot assess how large the dialectal difference is within this language. According to the first author’s brief research, native speakers of Darmdo Minyag consider the dialectal difference to not be particularly evident. However, to some extent, difference is perceived between the varieties of Kangding and Jiulong, though not to the extent that mutual intelligibility is effected.

Darmdo Minyag is estimated to have 10,000 speakers.\textsuperscript{4} The area where Darmdo Minyag is spoken is gradually decreasing, as we can see in the description of Huang (2007a). Meanwhile, the area where Kham Tibetan is spoken is growing.\textsuperscript{5} Speakers of Darmdo Minyag are generally bilingual in the local

\begin{enumerate}
\item The eastern counterpart of Darmdo Minyag is henceforth called Shimian Minyag, which is spoken mainly in Shimian County, east of Mt. Minyag Gongkar. There are two main reasons why we propose to distinguish Darmdo Minyag from Shimian Minyag: firstly, there are rare occasions for communication between the two populations, and secondly, there is low intelligibility between the two varieties according to the description by Sun (1983); see also Ikeda (2003). At present, there is no contact between the speakers of the two Minyag languages, which also display large linguistic differences.
\item The Minyag area is principally divided into two parts: Minyaggang and Minyagrong. The name Minyag Rabgang is also used in traditional Tibetan geography to refer to both areas together (cf. Karma rGyal-mtshan 2002:438). See also Suzuki and Sonam Wangmo (2014) and Dawa Zhuoma (2014).
\item Additionally, there are a small number of Darmdo Minyag speakers living in Zhusang Township, Yajiang County, to the west of Kangding Municipality.
\item This number is also mentioned in Sun et al. (2007:905), which, however, includes Darmdo and Shimian Minyag. Shimian Minyag is estimated to have 3,000 speakers (p.c. with Yin Weibin 2015).
\item At present, Darmdo Minyag seems to have no native speakers in Minyaggang. According to oral histories
\end{enumerate}
Khams Tibetan dialect, which we refer to as Minyag Rabgang Khams. The area where Darmdo Minyag is spoken is also adjacent to another region, where a different variety of Tibetan, Muli-nDappa, is spoken (Suzuki 2014). The latter is spoken by Darmdo Minyag-speakers living in the west of Pusharong Township.

We have collected data on 44 varieties spoken in the Minyagrong Valley through fieldwork conducted in Kangding. Most of the people who provided the data were in their 20s. The data reflected in the linguistic maps are limited to our first-hand materials for the sake of the consistency of the phonetic description. The list of varieties is:

Table 1: Research sites.

<table>
<thead>
<tr>
<th>Town/township</th>
<th>Hamlet (vernacular name)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pengbu (Phung-po-gshis or Bon-po-gshis)</td>
<td>Jiagenqiao (Zam-pa-kha), Xishaka (Shing-zam-kha), Mugu ('Bor-khul), Ritou (Ri-thog), Geerdi (Kun-legs-sde), Mudu (Mun-gtub), Mada (dMar-sde), Duorang (rDo-ra), Rioiu (Ri-'og), Jiangde (Cag-'dul), Naze (Lha-btsan), Tiku (mThul-lang), Mase (Ming-gser), Nadi (Lha-brtsse-gshis)</td>
</tr>
<tr>
<td>Shade (Sa-bde)</td>
<td>Shade (Sa-bde), Shenggu (gSer-'go), Wayue (Wa-yod), Ebarong-1 ('Go-pa-rang), Ebarong-2 ('Go-pa-rong), Laha (La-o), Chijiixi-Baiyi (Khrod-rgyu-gshis stod), Chijiixi-Wuyi (Khrod-rgyu-gshis smad)</td>
</tr>
<tr>
<td>Gonggashang (kLu-pa)</td>
<td>Liuba (kLu-pa), Sewurong-1 (Se'u-rong), Sewurong-2 (Se'u-rong), Shangchenzhi (Khri-'dzin stod), Xiachengzhi (Khri-'dzin smad), Shangmuju (Mun-rgya stod), Xiamuju (Mun-rgya smad), Yulongxi (gYang-legs-gshis), Chimei (Tsher-ma)</td>
</tr>
<tr>
<td>Pusharong (dPa-sreg-rong)</td>
<td>Yidai (Jig-rgen), Binggu (dPe-rong), Changcaoping (rTswa-ring), Kuxiro (Khru-shi-rong), Huoshan (Me-ri), Pusharong-1 (dPa-sreg-rong), Pusharong-2 (dPa-sreg-rong)</td>
</tr>
<tr>
<td>Jiju (Ce-cus)</td>
<td>Jiju (Ce-cus), Matic (Ma-mo), Geba (Dar-sha-gting), Songyu (gSer-gzhong), Caiyu (Tsha-yul)</td>
</tr>
<tr>
<td>Tanggu (Thang-mgo)</td>
<td>Tanggu (Thang-mgo)</td>
</tr>
</tbody>
</table>

Regarding the Lazexi dialect, we recorded varieties from three different generations (in their 70s, 40s, and 20s), in which the form used by people in their 40s will be used for the linguistic maps here.

We have had information that Darmdo Minyag was also spoken in Jiagenba Village, just north of Pengbu Village, around 30 years ago, however, our preliminary investigation of the village did not locate any speakers. Hence, Jiagenba was determined to be beyond the scope of our current research.

Through the present research, we have added detail to our knowledge of the distribution of Darmdo Minyag. The whole area listed in table 1 is generally known as the Darmdo Minyag linguistic area by people living in Minyaggang, non-local officials, and non-local scholars, however, in several of the above-mentioned hamlets, the inhabitants speak Tibetan languages (Minyag Rabgang Khams and Muli-nDappa Khams), suggesting that the Tibetan languages may have already replaced Darmdo

narrated by elders living in Minyaggang (Xinduqiao Town and Waze Township), there might have been Darmdo Minyag speakers in Minyaggang in the past. See also the description of Huang (2007a).

6 See Suzuki (2007) and Suzuki and Sonam Wangmo (2014) for detailed information regarding Minyag Rabgang Khams. However, meanwhile, some vernaculars of Minyag Rabgang Khams face language endangerment because of various sociolinguistic factors (Suzuki and Sonam Wangmo 2015).

7 We were unable to find speakers from Zhushang Village.

8 Each name is given with a Written Tibetan (WrT) form in parentheses. This essay consistently uses the pinyin name for each vernacular.

9 Also called Lazexi, which we use in the essay.

10 See Suzuki (2014) for a detailed classification of Khams Tibetan. There are many clear differences between Minyag Rabgang Khams and Muli-nDappa Khams in terms of phonology and morphology.
Minyag. At present, only a few people are monolingual in Darmdo Minyag, and the majority of Darmdo Minyag speakers are bilingual in Minyag Rabgang Kham. In Pusharong Village, the use of Muli-nDappa Kham is also attested. At present, the number of trilingual speakers of Darmdo Minyag, Minyag Rabgang Kham, and Chinese (Sichuanese, Southwestern Mandarin) is increasing, with an expansion of the use of Chinese in various social contexts, such as media and schooling.

Map 1 shows the overall distribution of languages spoken in the Minyagrong Valley, designed with the online Geocoding mapping method provided by the site: http://ktgis.net/gcode/lonlatmapping.html.

The blue dots show the distribution of communities where Darmdo Minyag is spoken, reflecting the present-day geographical range of this language.

As an additional remark, although Yang (1994) reports the existence of the ‘Zhaba’ language (including nDrapa and Choyu; cf. Ikeda 2003:97-101) in Jiju Village, it was not found there in our current research.
2 Phonetic description of the word form

This section provides a phonetic description of the word form ‘sun’ in each variety, following the main objective in the present volume of Studies in Asian Geolinguistics. Since we have not conducted an exhaustive survey of the phonology of each variety,\footnote{Except for the Lazexi dialect, which was described and analysed by the second author. A part of the analysis is provided in Suzuki (2011).} the description is highly phonetic. So is the tone; the given 5-grade value of each syllable is based on our phonetic observation and may be subject to change in forthcoming phonological analyses. A list of the phonetic forms attested in the vernaculars of the Minyagrong Valley that we recorded is:

<table>
<thead>
<tr>
<th>Type</th>
<th>Phonetic form</th>
<th>Distribution of hamlets (village name in parentheses)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A-1</td>
<td>m\textsuperscript{35}</td>
<td>Geerdli, Mudu, Mada, Duorang, Riwu, Jiangde, Naze, Tiku, Mase, Lazexi (Pengbuxi)</td>
</tr>
<tr>
<td>A-2</td>
<td>m\textsuperscript{55}</td>
<td>Shade, Shenggu, Wayue, Ebarong-1, Ebarong-2, Laha, Chijixi-Bayi, Chijixi-Wuyi (Shade); Sewurong-1, Sewurong-2 (Gonggashan)</td>
</tr>
<tr>
<td>B-1</td>
<td>n\textsuperscript{35}</td>
<td>Liuba, Shangchengzi, Xiachengzi, Shangmuju, Xiamuju, Yulongxi, Chimei (Gonggashan)</td>
</tr>
<tr>
<td>B-2</td>
<td>n\textsuperscript{55}</td>
<td>Yidai, Binggu, Changcaoping, Kuxi, Pusharong-1, Pusharong-2 (Pusharong)</td>
</tr>
<tr>
<td>C</td>
<td>n\textsuperscript{55}</td>
<td>Tanggu (Tanggu)</td>
</tr>
<tr>
<td>MR</td>
<td>n\textsuperscript{4}ma\textsuperscript{4}</td>
<td>Jiagenqiao, Xishaka, Mugu, Ritou (Pengbuxi)</td>
</tr>
<tr>
<td>MD</td>
<td>n\textsuperscript{24}</td>
<td>Huoshan (Pusharong); Jiju, Mati, Geba, Songyu, Caiyu (Jiju)</td>
</tr>
</tbody>
</table>

The first five forms (A-1, A-2, B-1, B-2, and C) are regarded as descendants of a common Darmdo Minyang etymon. MR is a form of Minyang Rabgang Kham, while MD is a form of Muli-nDappa Kham. Additionally, we collected the form of the Lazexi dialect of ‘sun’ from three different generations: from people in their 70s, 40s, and 20s. The word from the eldest generation is [m\textsuperscript{35}], with a slight pharyngealisation of the vowel, which is not attested in table 2. That of the other two generations is [m\textsuperscript{32}], the A-1 form.

Regarding the description of Darmdo Minyang in previous works, Huang (2007b) describes /n\textsuperscript{24}/\textsuperscript{12} for ‘sun’ (in the Muju dialect spoken in Gonggashan), which is close to A-2 in terms of a clear appearance of the ‘tense’ vowel, as well as B in terms of the vowel tongue position. Ikeda (2006:110) describes /n\textsuperscript{55}/ for ‘sun’ (in the Shenggu dialect spoken in Shade) and /n\textsuperscript{5}/ (in the Tanggu dialect spoken in Tanggu). The former is close to B-1 in terms of its lack of the ‘tense’ feature of the vowel, however, the Shenggu dialect in our data clearly has the ‘tense’ feature. It may imply the existence of difference among generations of speakers, however, more detailed investigations are needed. Thub-bstan dGe-legs et al. eds. (2008) give two forms for ‘sun’ (nii and nii ‘bus lus lus’ in Tibetan script), however, we have not found the latter form. The form of the four hamlets of Pengbuxi corresponds to WrT nyi\textsuperscript{13} ma, and the phonetic form is close to the form attested in the Minyang Rabgang Khams (a disyllabic form corresponding to WrT), whereas the form of the five hamlets of Pusharong and Jiju corresponds to that of the Muli-nDappa Khams (a monosyllabic form corresponding to WrT).\footnote{See Suzuki (2015) for a detailed description of the word form ‘sun’ in Tibetic languages in the eastern Tibetosphere.} See section 3.2 for details.

\footnotetext[11]{The underlining of the vowel designates a ‘tense’ vowel (Chn. \textit{jin yuan\textsuperscript{11}yin}). The ‘tense’ vowel in Darmdo Minyang is so problematic that Suzuki (2011) attempted to elucidate the basic vocalic characteristics of the Phungposhis dialect, which said: the ‘tense’ vowels are basically \textit{pharyngealised} or simply more back vowels to their counterpart caused by retraction of the tongue. /\textsuperscript{2}/ in Huang (2007b) basically corresponds to /\textsuperscript{3}/ in our description here.}
3 Map and analysis

Based on the phonetic description in section 2, here, we draw a linguistic map and describe the geographical distribution of forms of the term 'sun' in Darmdo Minyag and discuss the features of this distribution.

3.1 Linguistic map

The following map is designed with the Arc GIS online.

Map 2: Overall distribution of Darmdo Minyag ‘sun’.
3.2 Analysis

The word forms with a /n/-initial (MR and MD), similar to that of Khams Tibetan, are attested in the north of Pengbuxi Village and in the west of the Pusharong Valley. They are a Tibetic etymon. The vernaculars using them are not Darmdo Minyag but Khams Tibetan (Minyag Rabgang Khams and Muli-nDappa Khams respectively), and their distribution is already inside the Minyagrong Valley and its tributaries. On the other hand, the word form with a /n/-initial is regarded as Darmdo Minyag in origin. This type has differences regarding the vowel quality, e.g. whether it has a ‘tense’ (mainly pharyngealised here) feature or not, and whether the tongue position is low, mid-central, or mid-back. From an historical viewpoint, the ‘tense’ feature realised as a pharyngealisation is gradually being lost, hence the form A-2 may be the most archaic form, and the form with a mid-back vowel (ɤ), A-1, is the second-most archaic one. Both of them are distributed in the northern part of the Minyagrong Valley, which adjoins the area of Minyag Rabgang Khams. We explore possible reasons for this below.

The distribution of the two above-mentioned archaic forms is, however, a bit peculiar from the viewpoint of dialectology. The main road goes through Pengbuxi and Shade, where there might be the most occasions for language or dialect contact, which could easily trigger phonetic changes. However, the present data reflected in map 2 shows a different pattern. In this case, we rather assume that the varieties spoken in the central area of Minyagrong—Sewurong, just south of the present Shade Village (see Yudru Tsomu 2009)—maintain archaic traits. Investigations regarding other word forms are needed.

4 Conclusion

This article provided a linguistic map displaying the current Darmdo Minyag linguistic area within the Minyagrong Valley and its tributaries, and presented a preliminary geolinguistic analysis of Darmdo Minyag, taking the word ‘sun’ as an example. The analysis has shown that a great lexical difference is not attested regarding this word form, and that the vernaculars spoken in the northern area of the Minyagrong Valley maintain a more archaic form.

Acknowledgements

Our thanks go to the friends from Minyagrong who provided language data for the authors, especially to rGya-ma, rGya-mtsho, and rGyal-dkar. We should also like to thank Gerald Roche for checking our English and contributing insightful comments. A part of the fieldwork for collecting data of Darmdo Minyag was funded by a Grant-in-Aid for Scientific Research of Japan Society for the Promotion of Science: “Linguistic Substratum in Tibet” headed by Yasuhiko Nagano, No. 16102001.

References


Suzuki, H. and Sonam Wangmo. 2014. Language evolution and vitality of Lhagang Tibetan, a Tibetan language as a minority in Minyag Rabgang. Paper presented at the workshop on the linguistic minorities of the Chinese Tibetosphere (Uppsala) [A revised version is to be published in International Journal of the Sociology of Languages, 2016].


A Geolinguistic Description of Terms for ‘Sun’ in Tibetic Languages of the Eastern Tibetosphere

Hiroyuki Suzuki

IKOS, University of Oslo / National Museum of Ethnology

Abstract

This article attempts to describe a dialectal difference of the word ‘sun’ attested in around 230 dialects of the eastern Tibetic languages. The word ‘sun’ generally corresponds to WrT nyi ma, however, a small number of vernaculars spoken in the southernmost area of the eastern Tibetic languages (Yunnan) use WrT gnam lha form for ‘sun’, and one dialect from Sharkhog uses a form /na ʰtsa/, of unclear origin.

1 Introduction

This article provides a detailed description of the geolinguistic analysis of the word forms for ‘sun’ in the Tibetic languages spoken in the eastern Tibetosphere, which Shirai et al. (2015) did not describe in detail due to their focus on the whole of the Tibeto-Burman linguistic area. The geographical scope of the eastern Tibetosphere in this essay principally includes Khams and Amdo in the traditional Tibetan geography, which basically corresponds to China’s Qinghai, Gansu, Sichuan, and Yunnan provinces as well as a part of Chamdo District of Tibet Autonomous Region (TAR). One dialect from Myanmar is also included, however, several data points from Khysungpo (North-eastern part of TAR) and Yulshul (Southern part of Qinghai) as well as Minyagrong (see Dawa Drolma & Suzuki 2015 for a detail) were omitted because of practical reasons.

The data used to create the linguistics maps at the end of this paper only includes first-hand materials collected by the author from 2003 to 2015. Because of this, as well as because of time constraints on the part of the author, the data points are not equally distributed within this area, and the points on the map only reflect the current research situation. The present map contains 228 points.

The linguistic maps reflect so-called ‘regiolects’, i.e. dialects with regional differences. Sociolects, which certainly exist in the given area,1 are not dealt with in this essay.

2 Classification of word forms

This section provides a classification of word forms of ‘sun’ based on the phonetic differences. There are three large categories: the nyi ma type,2 the gnam lha type, and the /na ʰtsa/ type. The first type includes numerous types of phonetic realisations. The classification proposed in the article is as follows:3

A. nyi ma type
   A-1: disyllabic form as /ŋV mV/
      [ŋ ma], [ŋə ma], [ŋi mə], [ŋə mə], [ŋə mo], [ŋə mə], [ŋə mō], [ŋi mā], etc.
   A-2: disyllabic form with /n/ initial
      [nə ma]
   A-3: disyllabic form as /ŋV wV/

1 Lhagang Tibetan, for example. Cf. Suzuki & Sonam Wangmo (2015).
2 Each form of Written Tibetan (henceforth WrT) is given in italics, transliterated based on the Wylie system.
3 A suprasegmental description is uniformly omitted.
A-4: monosyllabic form with /ŋ/ initial
[näː], [nəː], [nɵː], [nũː wã], [nũ wɔ]
A-5: monosyllabic form with /ŋ/ initial
[naː], [nəː]
B. *gnam lha* type
[näː ʃa]
C. /na ʰtsa/ type
[nə ʰtsa]

Depending on the purpose for drawing a linguistic map, the subclassification (A-1 to A-5) above can be simplified and three major groups (A, B, C) are principally concerned, as reflected in Shirai et al. (2015). This article does not apply this simplification.

3 Geographical distribution and interpretation

The lexical forms representing the ‘sun’ can be classified into: A) the *nyi ma* type, B) the *gnam lha* type, and C) the /na ʰtsa/ type. Type A is far more frequent than the other two types. It is a common form found in the Tibetan languages and is therefore observed throughout the eastern Tibetosphere, as can be seen in Map 1. Type B (*gnam lha*; lit. heaven-deity⁴) and type C are in the minority, and are merely attested in an extremely small area of the southernmost and easternmost part respectively of the eastern Tibetosphere.

Type A includes various phonetic realisations, most of which, however, follow the phonological change of each dialect. The disyllabic form /ŋ mV/ (A-1) is the form which directly corresponds to WrT *nyi ma*, and its present distribution is the largest throughout the area. The distribution of monosyllabic forms (A-4 and A-5) is limited in the easternmost and southeastern areas. These areas are located on the border between the Tibetosphere and Sinosphere. Another disyllabic form /ŋ wV/ (A-3), which is analysed as a transitional form from A-1 to A-4, is attested to in the area close to the distribution field of A-4.⁵ However, this geographical position of dialects (i.e. the border zone between the Tibetosphere and Sinosphere) with a monosyllabic form of ‘sun’, and the monosyllabification is not necessarily directly related to each other. For example, the dialects of Rongbrag and Minyag Rabgang do not have a monosyllabic form in spite of their proximity to the Sinosphere. Another noteworthy feature of type A is the existence of the /ŋ/-initial form (A-2 and A-5). The reason why this form is related to type A (an etymon of WrT *nyi ma*) is because a regular sound correspondence of WrT *ny* with /ŋ/ is attested in several dialects, such as Sangdam (Suzuki 2012a) and Bragkhoglung (Suzuki 2012b).⁶

Map 2 is an enlarged version of the southern part of Map 1, in which the distribution of types A-3 and A-4 are analysed. Looking at the border between A-1 and A-3, A-4 which can be seen in the northeastern area of Map 2, it seems that there is an ‘isogloss’ formed by these two areas. However, there actually exists another language called Darmo Minyag between the Muli-nDappa and Minyag Rabgang dialectal groups (Dawa Drolma & Suzuki 2015). In the southern area of Map 2, the geographical distribution of the forms for ‘sun’ are complicated and distinctions are attested to within dialectal groups such as Sem-kyi-nyila and sDerong-nJol. In this area, the difference of the word form can be regarded as a sound change which independently occurred. Regarding the case of the Sem-kyi-nyila group, it may be because of influence from the neighbouring dialects of Muli-nDappa.

---

⁴ This word formation, including an expression of deity, may imply a background of sun worship. However, no other evidence has been attested to which suggests that the dialectal area (southernmost rGyalthang) has ever had such a religious practice.

⁵ Regarding the fusion of disyllabic words in Thewo Tibetan, see Rig-‘dzin dBang-mo (2013).

⁶ This phonological correspondence taken into consideration, type C, containing /n/ at a word-initial position, is not regarded as a form derived from WrT *nyi*, because the dialect (gSerpo) does not display this sound correspondence as a phonological rule.
which have lead to their use of the A-3 and A-4 forms. As for the dialects spoken alongside the Lancang River, three word forms A-1, A-3, and A-4 are used. In both the Northern and Southern tips of this area the form A-1 is used, hence the A-3 and A-4 forms may have emerged in the central position of this area, for similar distribution of variations regarding certain word forms are also reported (Suzuki forthcoming).

Map 3 is an enlarged version of the north-eastern part of Map 1. The dialects displayed in Map 3 consist of various genetically different language groups (Suzuki 2009, 2015). Within this region, we can note that the Sharkhog and Khodpokhag varieties include A-1, A-3, and A-4 types. The former language’s form for ‘sun’ can be analysed as a transitional form from A-1 to A-3, for there are no monosyllabified forms (A-4 type) attested there. The latter language, on the other hand, mainly has A-4 type, which is a form created by the a coalescence of two syllables. The monosyllable forms (A-4 and A-5) are also attested in Cone Tibetan and in one dialect from Diebu County which belongs to the Thewo-stod group. These two areas are connected to each other with a mountain path, hence this phenomenon may be analysed as an areal feature brought about by frequent language contact.

Old Tibetan has another orthographic form for ‘sun’ gnyi ma, which, however, does not have any attested cognates in the eastern Tibetosphere. In addition, Literary Tibetan has huge amounts of expressions meaning ‘sun’ such as kun gsal (lit. all-shining) and jeg rten dbang po (lit. world-lord), but none of them are used as dialectal forms for referring to the sun. See appendix for a list of literary word forms.

4 Conclusion

The word form of ‘sun’ in the eastern Tibet languages mainly corresponds to WrT nyi ma, and its geographical distribution is the largest among the three attested lexical items. The lexical variation of morphemes is therefore not rich; the other two forms are: WrT gnam lha and /na ‘tsa/, both of which are used in isolation or in a limited geographical area.

Acknowledgements

I should like to express my gratitude to my Tibetan friends who helped me and taught me their mother tongue. My thanks also go to Abe Powell for polishing my English and contributing insightful comments. A part of the discussion was ameliorated based on that in the first meeting of the Project “Studies in Asian Geolinguistics” (3-4. October 2015). Field research was funded by four Grants-in-Aid for Scientific Research from the Japan Society for the Promotion of Science [JSPPS]: “Linguistic Substratum in Tibet” (headed by Yasuhiko Nagano, No. 16102001), “Dialectological Study of the Tibetan Minority Languages in the Tibetan Cultural Area in West Sichuan” (headed by the present author), “International Field Survey of the rGyalrongic Languages” (headed by Yasuhiko Nagano, No. 21251007), and “Study on the Dialectal Development of Tibetan Spoken in Yunnan, China, through a Description of the Linguistic Diversity” (headed by the present author, No. 25770167), as well as and a private financial support of Tibetan Studies Committee of Yunnan Ethnology Association headed by Xu Jianhua.

Appendix: List of Tibetan literary words for ‘sun’


References


Maps 7

Map 1: Overall distribution of word forms

Legend:
A WrT nyi ma type
- [nĩ ma], [ŋo ma], [ŋo ma], [ŋo mo], etc.
- [nɔ ma]
- [ŋo wɔ], [ŋo wɔ], [nɕ wɔ], [ŋi wɔ], etc.
- [ŋː], [ŋː], [ŋːː], [ŋːː], [ŋːː]
- [ʰaː], [ŋːː]
B WrT gnam lha type
- [nː ŋa]
C /na ‘tsa/ type
- /na ‘tsa/

7 The linguistic maps here were designed with ArcGIS online (Map 1) as well as with the online Geocoding mapping method provided by the site: http://ktegis.net/gcode/lonlatmapping.html. (Maps 2 and 3)
Map 2: Distribution of word forms in the southeastern Khams region

Legend
- [ŋi ma], [ŋə ma], [ŋə ma], [ŋə mo], etc.
- [ŋə ma]
- [ŋə wū], [ŋi wū], [ŋi wō]
- [ŋu:], [ŋu:], [ŋu:]
- [ŋu:]
Map 3: Distribution of word forms in the easternmost Amdo region

Legend
- [ŋi ma], [ŋə ma], [ŋə ma], [ŋə mo], etc.
- [naː], [njaː]
- [ŋə wà], [ŋi wà], [ŋi wɔ]
- [ŋuː], [ŋuː], [ŋaː], [ŋaː]
- [na 'tsa]
A Survey of Recent Austroasiatic Studies

Makoto MINEGISHI (ILCAA, Tokyo University of Foreign Studies)
Toshiki OSADA (Research Institute for Humanity and Nature)
Nathan BADENOCH (Kyoto University)
Masaaki SHIMIZU (Osaka University)
Atsushi YAMADA (Japan Health Care College)
Yuma ITO (Kyoto University)

1. Introduction
This paper surveys studies in Austroasiatic languages (hereafter abbreviated as AAL) in order to identify the data used in the Asian geolinguistics project. First, we will focus on the data published after around 1990. Second, we will introduce some of the studies on AAL done by the scholars in Japan. Since the number of Japanese AAL researchers is very limited, we will refer not only recent works, but also some important ones done in 1960’s and 70’s, such as those by Sakamoto. Third, authors of the present paper currently working in different branches of AAL will appraise works by offering relevant bibliographies, and commenting on recent progress made in their subgroups of interest.

Given the broad geographic distribution and sheer number of languages in the Austroasiatic phylum, this review is necessarily selective. From the rich body of literature that continues to emerge, we have chosen the most fundamental and relevant work to this project. Needless to say, an exhaustive review of AAL would be a much more significant work of research, and the authors designed the present survey with the objectives and scope of the project in mind.

2. Typological Characteristics of AAL
AAL is comprised of the Mon-Khmer languages in Southern China and mainland Southeast Asia, the Munda languages in India, and the Nicobarese languages in the Andaman Sea. Most AAL in Southeast Asia are spoken in ethnic minority communities except Khmer (Cambodian) and Vietnamese, the official languages of Cambodia and Vietnam respectively. Although some languages in India have large populations, such as Santali at 6.5 million, Mundari 1.5 at million, Ho 1 at million, and Khasi 1 at million according to Osada & Onishi (eds) (2012), others are mostly spoken by minorities, and some of them are endangered.

The notion of an Austroasiatic language family was advocated by Rev. Wilhelm Schmidt (1868-1954), SVD (Societas Verbi Divini) in the beginning of twentieth century. He proposed the idea of the Austroasiatic language family comprising the Mon-Khmer languages and Munda languages, and further put forward the idea of the Austric superfamily including Austroasiatic and Austronesian languages. It should be noted, however, that the term ‘Austric’ is commonly understood in India as equivalent to Austroasiatic.

Typologically, verb-medial (Subject-Verb-Object) Mon-Khmer languages in the east contrast with verb-final (Subject-Object-Verb) Munda languages in the west. In between are verb-initial (Verb-Object-Subject) Nicobarese languages in Andaman Sea. Morphologically, while languages in Southeast Asia are basically isolating, those in South Asia have complex derivations with prefix, infix, and suffix. Lexically, not many words in AAL are attested as cognates.

3. Recent Researches in AAL
Since Schmidt’s propagation of AAL, comparative studies of the phylum did not show much progress until the beginning of the twenty-first century. The high level of language diversity and broad geographic distribution of AAL, especially minority languages in dispersed areas, caused difficulties for linguists working in one language in extending their interest to another one. Knowledge in Khmer, for example,
would not help linguists study Mon since these two languages differ both in grammatically and in lexically. Most linguists in the area therefore have focused on a few languages in a subgroup or two trying to complete their descriptive studies. Dictionaries and grammars were thus published concerning major languages, such as Khmer, Vietnamese, Mon, and Santali.

Another obstacle for development in research was the lack of communication among linguists. Those working in Southeast Asia or China have not had frequent contact with those working in South Asia. Situation has drastically changed, however, with resumption of the International Conference of Austroasiatic Linguistics (ICAAL). The third ICAAL was held in 2007 at Deccan College, Pune, after a long interim – the first was held in 1973, and the second in 1978. Now ICAAL is held almost regularly: sixth in 2015, and seventh to be held in the next two years.

A major output from ICAAL is Jenny & Sidwell (eds) (2015), *the Handbook of Austroasiatic Languages*. It is the first reference work of AAL, which is now regarded as an essential and comprehensive guide for AAL studies. Part I of the book comprises a typological overview, genetic classification, comparative reconstruction. Part 2 provides grammatical sketches of subgroups Aslian, Monic, Pearie, Khmeric, Bahnaric, Katuic, Vietic, Khmuic, Mangie, Palaungic, Khasian, and Nicobarese.

Thai researchers, primarily at Mahidol University, have been conducting research on the Khmu and Palaung languages. These studies, importantly, have been regional projects, not limited to varieties spoken in Thailand. Work on Khmu included a multi-country survey of dialects, producing not only lexical resources, but a descriptive grammar, extensive treatment of tonogenesis, and conversational materials. These materials complement the body of works by Svantesson, Lindell and Kam Raw, covering linguistic and folklore studies of the 1970s and 1980s. With regards to Palaung, researchers at Mahidol have recently completed a project to document the diversity of the Palaung languages, covering varieties spoken in Myanmar, China and Thailand. The Journal of Language and Culture, published by Mahidol, issued a special volume on Palaung in 2009. The depth of Palaung linguistic diversity has only started to be understood, but these recent materials move forward the important historical work of Ostapirat (2009) and earlier linguists, such as Mitani’s 1977 reconstruction of Proto-Palaung.

In terms of cross-cutting typological issues, researchers of AAL have been engaged in exploration and discussion of the processes of tonogenesis and registrogenesis. The influence of these suprasegmental dynamics on the trajectories of change that are seen in the AAL family is significant. The well-known cases of register in Khmer and Mon have been followed by a growing body of research in Katuic, Bahnaric and Pearic. While the process of devoicing and register and in some cases tone development has been observed widely across the family, other mechanisms of tonogenesis have been described by Svantesson for the Angkuic branch of Palaungic languages. The question of register in the history of the AAL phylum is yet to be resolved.

4. Resources for AAL Geolinguistics

4.1 Online resources

In what follows, we will review the sources of data to be used in the study of AAL geolinguistics.

First of all, information provided by SEALANG projects on the following web site will provide a critical resource for the Asian geolinguistics project.

http://sealang.net/

The site provides information of Southeast Asian languages with the following online library collecting materials, including dictionaries.

http://sealang.net/library/

Among the services provided by SEALANG project, the most convenient resource for our research is their “Etymological Dictionary, built to support work in comparative and historical linguistics, and a companion Languages Database devoted to preservation and sharing of language and lexical resources. [cited from the site]”

http://sealang.net/monkhmer/dictionary/
The site above provides a database based on published material, where we can retrieve words with bibliographical information about the original first-hand data.

Some other web sites provide online dictionaries. For example, the following site at ILCAA provides word search in the Santal Dictionary by P.O. Bodding (1868-1936). This is part of the collaboratively work on Santali by M. Minegishi, Jun Takashima at ILCAA, Tokyo, and Ganesh Murmu at Ranchi University, India.¹

http://www.aa.tufs.ac.jp/~mmine/india/Bodding2k/index.html

Concerning languages in South Asia, Christian-Albrechts-Universität zu Kiel (Kiel University) provides information on “seldom studied and endangered South Asian languages”.

http://www.isfas.uni-kiel.de/de/linguistik/forschung/projekte/southasiabibliography/bibliography/austroasiatic/austroasiatic

4.2 Dictionaries

Since the online dictionaries and materials available on the web sites above are digital reprint versions of paper publications, it is recommended to refer to the original dictionaries and grammar books if at all possible.

Although dictionaries on major languages such as Khmer, Vietnamese, and Santali are available, those of minority languages are not numerous. Among such dictionaries, the following have appeared only recently:


Moreover, linguists in Vietnam have been steadily producing a valuable collection of dictionaries of Vietic, Katuic and Bahnaric language spoken in that country, all of which are glossed in Vietnamese and are difficult to obtain outside of Vietnam.

5. AAL Education and Studies in Japan

Among AAL, Vietnamese is taught and researched in several universities in Japan, such as Osaka University, Kanda University of International Studies, and Tokyo University of Foreign Studies (TUFS). Khmer is taught in TUFS only. These universities have regular staff teaching undergraduate and graduate students. Among them are Prof. Kenji Tomita (grammar) and Masaaki Shimizu (phonology and syntax) at Osaka University and Atsushi Kasuga (phonetics and syntax) at Kanda University concerning Vietnamese. Hiromi Ueda and Tomoko Okada at TUFS work on Khmer syntax.

¹ Based on the examination of headwords of the Santal dictionary, Minegishi et al. (2010), (2011) claim that Bodding’s distinction between narrow and open vowels, i.e., “e” vs. “ë”, and “o” vs. “ö” is not a phonemic one, thus, the Santali vowel system has only six vowels, rather than eight that Bodding claimed.
5.1 Prof. Sakamoto’s works on Khmer and Mon

Yasuyuki Sakamoto (坂本 晃), Prof. Emeritus of TUFS, is one of the linguists who initiated the study of Khmer and Mon languages in Japan. He made some important studies on phonologies of modern standard Khmer [坂本 (1968)] and old Khmer (for example Sakamoto (1970), (1971), (1974)). He was the first professor in Khmer studies at Tokyo University of Foreign Studies, to offer a regular course in Cambodian since 1992. He published grammar and dictionaries of Khmer, and dictionaries of Mon as follows:


He also published the following Mon dictionaries based on his field work done in Bangkok:


5.2 Khmer phonology in Japan

Along with his phonetic description and phonological analysis of modern Khmer based on his fieldwork in Phnom Penh, Sakamoto made phonemic analysis of old Khmer based on inscriptions. From his analysis of distribution of vowel symbols in the inscriptions, he induced the original vowel system. Sakamoto (1974), for example, distinguishes three ways in usage of vowel symbols <a>, <aa>, or <o>. The first is to write a vowel in a word with either <a> or <aa>, but never with <o>. The second is to write a vowel with either <a> or <o>, but never with <aa>. The third is to write a vowel always with <a>. By examining the modern pronunciation of the words written in these three types, he reconstructs two phonemes: vowel */A/ written as <a> or <aa>, and vowel */O/, written as <a> or <o>. The one always written with <a> can be identified as either */A/ or */O/. See Sakamoto (1970), (1971) for the case of other vowels.

Following Sakamoto’s phonemic description of modern standard Khmer, 峰岸 [Minegishi] (1985) showed that there are dialectal differences in phonemic system in Khmer, especially in vowels. Minegishi (1986) gives the phonemic system of the local dialect in Takeo province near Phnom Penh, which differs significantly from that of standard Khmer.

5.3 Work on Palaungic languages [reviewed by Badenoch]

Yasuyuki Mitani (三谷 恭之) made early contributions to the study of the Waic languages in the 1960s and 70s, particularly with analysis of the Lawa and Khamet languages spoken in Thailand. Moreover, Mitani’s reconstruction of Palaungic is still respected as the most solid scholarship on the topic. In addition to the analytical material presented by Mitani that drew on his own field research, an important contribution of his was to bridge back to the earlier days of field surveying conducted by scholars such as Luce. Very recently, Cheeseman et al. (2015) have published “Palaungic Linguistic Bibliography with Selected Annotations”, now probably to be considered to reflect the state of the art of Palaungic literature.

5.4 Current AAL investigators

In addition to the linguists given above, the following are some of the active linguists in Japan
investigating AAL.

Naomitsu Mikami at the Keio Institute of Cultural and Linguistic Studies, Keio University works on Vietnamese, Thai, and other languages in mainland Southeast Asia. Toshiki OSADA has conducted field work on Munda languages in Bihar and Jharkhand states in India. He is specialized in Mundari language and has published Mundari grammars, textbooks and academic papers. Atsushi Yamada works on Waic in China. Yuma Ito has conducted fieldwork on Mlabri, endangered language in Thailand.

6. State of the art of AAL studies

In what follows, the authors of the present paper currently working in different fields, offer the bibliographies, and comment on the recent progress in the area of their concern. We will hereafter refer to the materials written in English, Japanese or Chinese only.

6.1. Comparative Studies [by Osada]

[General Comments]
Along with the recent steady progress in descriptive studies of AAL, reconstruction of the language family has made great progress. Paul Sidwell works most actively in this area.

[Recent publications]
Sidwell, P. 2005. The Katuic Languages: Classification, Reconstruction and Comparative Lexicon. Lincom Europa, Munich, Germany
Sidwell, P. 2009. Classifying the Austroasiatic languages: History and state of the art, Lincom Europa, Munich Germany

6.2. Paleontology [by Osada]

AAL paleontology has only very recently started to open new perspectives on the field:

6.3. Studies in China [by Yamada]

[Recent publications]
The following are publications concerning languages of Mon-Khmer subgroups in China.

[Austroasiatic in China]
刘岩 (2006)『孟高棉语声调研究』中央民族大学出版社.
颜其香・周植志 (1995)『中国孟高棉语族语言与南亚语系』 中央民族大学出版社.

[Waic]
王敏鹏・张化鹏・肖玉芬 编 (1994)『佤语研究』云南民族出版社.
王敏鹏 编 (1992)『佤语熟语汇释』云南民族出版社.


YAMADA, Atsushi (2007) Parauk Wa Folktales. 佤族（巴饶克）的民间故事, ILCAA.

[Angkuic]
陈国庆 (2005)『克蔑语研究』民族出版社.

[Bit–Khang]
刀洁 (2007)『蓝茂语研究』民族出版社.

[Khmuic]
陈国庆 (2002)『克木语研究』民族出版社.

[Pakanic]
高永奇 (2003)『莽语研究』民族出版社.
李旭敏 (1999)『佧语研究』中央民族大学出版社.
李云兵 (2005)『布頑语研究』民族出版社.

【General Comments】
Elucidation of little known languages, such as those of the Palaungic and Khmuic branches shows progress. Among them investigation of Waic languages has progressed in terms of both quality and quantity. Investigators focus on historical studies such as historical phonology and comparative studies. Studies in typology and sociolinguistics are expected to increase.
6.4. Studies in Munda languages
As mentioned above, Osada has made a significant contribution to the study of the Munda languages, including multiple articles in the important book *The Munda Languages*, edited by Anderson and released in 2008. Among these are “Mundari”, as a single author chapter, and the co-authored “Ho and Kherwarian languages” with Anderson and Harrison. Moreover, Kobayashi co-authored a chapter on Kera Mundari. A language learning text, *Lessons in Mundari*, has been released in pre-publication form, authored by Osada et al.

6.5. Studies in Khmuic languages [summarized by Ito]

**[Recent publications]**

**[General Comments]**
Comparative studies have shown noticeable progress. Some minor languages in the group have been investigated, such as Mlabri (by Ito), and Bit, Ksingmul and Phong (by Badenoch). Description of other members of this branch is necessary.

6.6 AAL in the Japanese grey literature [summarized by Badenoch]
Over the past two decades, Japanese field-based projects have produced a number of short, but valuable word lists for languages that remain largely undescribed. For example, publications from ILCAA have introduced data on Then (Khmuic), Khabit (Khmuic-Palaungic?) and Khmu (Khmuic), all presented by Kosaka (2001, 1999, 1999). Additionally, a Khang wordlist from Vietnam was produced by Ueda ed (2003), as well as Mlabri by Sakamoto ed (2005), both as part of the Endangered Language of the Pacific Rim project. Moreover, Palaungic wordlists have been released through ILCAA: The Palaung Language (Shintani 2008) and the Riang Language (Shintani 2014).

References


