

Studies in Asian Geolinguistics, Monograph Series No. 2

Proceedings of the Workshop “Geolinguistic Method and Southeast Asian Linguistics”



ILCAA Joint Research Project 2015 - 2017 “Studies in Asian Geolinguistics”

Research Institute for Languages and Cultures of Asia and Africa

Tokyo University of Foreign Studies

Studies in Asian Geolinguistics, Monograph Series No. 2

*Proceedings of the Workshop “Geolinguistic Method and
Southeast Asian Linguistics”*

11–13 May 2017
Padang, West Sumatra, Indonesia

First published 2017

Edited by Hiroyuki SUZUKI & Mitsuaki ENDO

Copyright © 2017 Individual Contributors

ISBN 987-4-86337-268-9

This is an electronic publication in PDF format.
This publication is offered under Creative Commons International License
Attribution–NonCommercial–NoDerivatives 4.0.



<http://creativecommons.org/licenses/by-nc-nd/4.0/>

Published by
 Research Institute for
Languages and Cultures of
Asia and Africa (ILCAA)
Tokyo University of Foreign Studies
3-11-1, Asahi-cho, Fuchu-shi,
183-8534, Tokyo, JAPAN
<https://publication.aa-ken.jp/>

Table of contents

Workshop Abstract and Programme	1
Introduction	
Mitsuaki ENDO	
<i>Asian geolinguistics: So far, and beyond</i>	2
Chitsuko FUKUSHIMA	
<i>Methods and principles of geolinguistics</i>	10
Mika FUKAZAWA	
<i>A manual for drawing geolinguistic maps with ArcGIS Online</i>	16
Case studies	
Satoko SHIRAI	
<i>A geolinguistic approach to Tibeto-Burman vocabulary:</i>	
<i>A case study of ‘sun’ and a future perspective</i>	34
Masaaki SHIMIZU & Mika KONDO	
<i>The distribution of diphthongs in Vietnamese dialects</i>	51
Kazue IWASA	
<i>Mapping the Yi characters</i>	
<i>– an attempt to explore a route of their historical change and origin –</i>	59
Appendix	
Hiroyuki SUZUKI	
<i>How a geolinguistic approach can contribute to a study of Tibetic languages:</i>	
<i>Examining the cases of ‘wind’ and ‘rice’</i>	68

Cover photo: Members of a traditional dance troupe of Minangkabau in Padang, at the occasion of SEALS 27, taken by Mika FUKAZAWA on 11 May 2017

Workshop Abstract and Programme

Workshop "Geolinguistic Method and Southeast Asian Linguistics"
The Twenty-Seventh Meeting of the Southeast Asian Linguistics Society (SEALS 27)
11-13 May 2017
Padang, West Sumatra, Indonesia

This workshop aims at applying the geolinguistic method to the study of Southeast Asian languages. This discipline, which is also called word geography, dialect geography, linguistic geography, or areal linguistics, has a long tradition of over hundred years in European and Japanese dialectology.

There are two sessions: Introduction and case studies. Firstly, we will overview previous geolinguistic studies on Asian languages, and survey future tasks. Next, geolinguistic methods and principles will be outlined based on the tradition of European and Japanese dialect geography. Further, how to draw dialect maps using concrete language data will be discussed. This practical part leads participants to draw maps by themselves according to the manual which will be distributed.

The second half will consist of case studies, dealing with vocabulary, phonology, as well as script. The first presentation will deal with Tibeto-Burman vocabulary using the geolinguistic approach. The second will be a case study on phonology based on field investigations of Vietnamese dialects. Lastly, the geolinguistic method will be applied to the study of the Yi script, which is known to show great diversity in terms of geographical distribution.

Date: 11 May 2017

Venue: Kyriad Hotel Bumiminang, Padang, West Sumatra, Indonesia

1. Introduction 13:30-15:15

1.1. Theoretical part

Asian geolinguistics: So far, and beyond, Mitsuaki ENDO (Aoyama Gakuin U.)

Methods and principles of geolinguistics, Chitsuko FUKUSHIMA (U. of Niigata Prefecture)

1.2. Practical part

A manual for drawing geolinguistic maps with ArcGIS Online, Mika FUKAZAWA (Chiba U.)

2. Case studies 15:45-17:30

2.1. *A geolinguistic approach to Tibeto-Burman vocabulary*, Satoko SHIRAI (Tsukuba U.)

2.2. *The distribution of diphthongs in Vietnamese dialects*, Masaaki SHIMIZU (Osaka U.) & Mika KONDO (Kyoto U.)

2.3. *Mapping the Yi characters – an attempt to explore a route of their historical change and origin –*, Kazue IWASA (Kyoto U.)

References

- Endo, Mitsuaki. ed. (2016). *Studies in Asian Geolinguistics I, II, III*. ILCAA ELECTRONIC PUBLICATIONS. Tokyo: Research Institute for Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies. <https://publication.aa-ken.jp/>
- Endo, Mitsuaki. ed. (2016). *Papers from the Third International Conference on Asian Geolinguistics*. ILCAA ELECTRONIC PUBLICATIONS. Tokyo: Research Institute for Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies. <https://publication.aa-ken.jp/>
- Kondo, Mika. (2013). *Vietnamese Dialect Maps on Vocabulary*. Asian Geolinguistic Society of Japan, Monograph Series No.1. <https://agsj.jimdo.com/monograph/>

Asian geolinguistics: So far, and beyond

Mitsuaki Endo

Faculty of Economics, Aoyama Gakuin University

Abstract











This article describes the outline of the Joint Research Project “Studies in Asian Geolinguistics” at the AA Institute, Japan 2015-2018. It consists of (1) Outline of the project, (2) Goals of the project, (3) Colouring: Resolution for mapping the area where many language families overlap, (4) Synthetic map in East and Southeast Asia: The sun, and (5) Future activities.

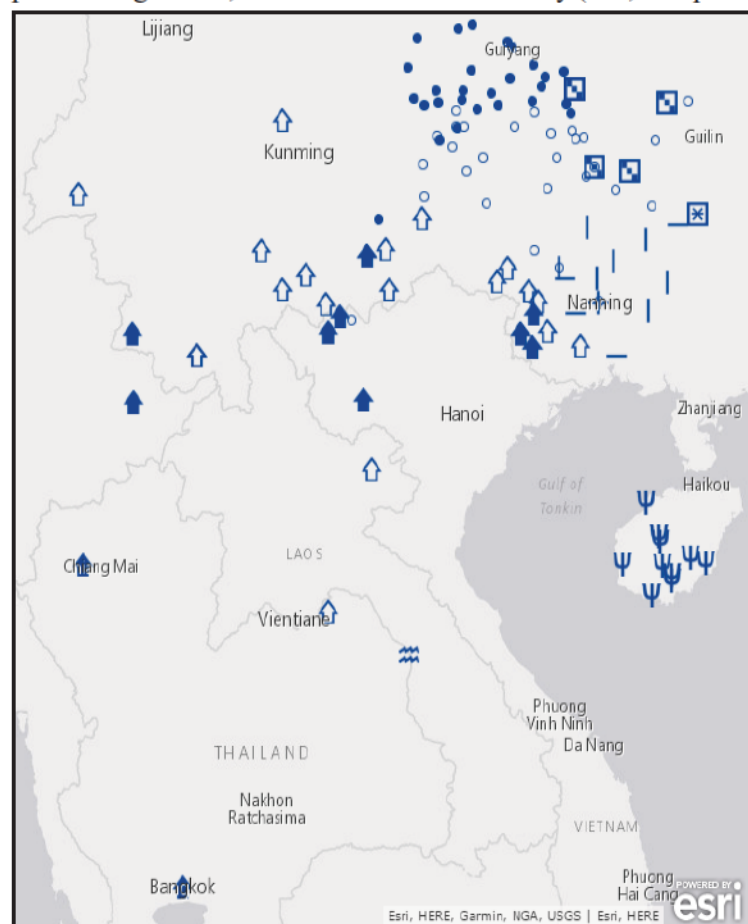
1 Outline of the project

The Joint Research Project “Studies in Asian Geolinguistics” has been in progress since April 2015 at The Research Institute for Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies (see <http://www.aa.tufts.ac.jp/en/projects/jrp/jrp210>).

The goal of this project is during the academic year 2018 to hopefully complete compilation of *The Linguistic Atlas of Asia*, Vol. 1, which is to cover all language groups throughout Asia. It will be comprised of 8 items with linguistic maps covering over 1,000 locations for vocabulary (sun, rice plant, milk, wind, iron), phonology (tone and accent), and syntax (how to count nouns, “it rains”).

The results of the first year 2015 are already open to the public in pdf files at the site of the AA Institute: <https://publication.aa-ken.jp/> Mitsuaki Endo (ed.), *Studies in Asian Geolinguistics* I (Sun), II (Rice plant), III (Milk), and Mitsuaki Endo (ed.), *Papers from the Third International Conference on Asian Geolinguistics*.

- | | | |
|---|-----|----------------------|
|  | A1 | lek with higher tone |
|  | A2 | lek with lower tone |
|  | B | khjak |
|  | C | khət |
|  | D-1 | thi:t |
|  | D-2 | ti:t |
|  | E-1 | va2 |
|  | E-2 | fa2 |
|  | F | go:i |
|  | G | maa |



Map 1 Iron in Tai-Kadai (Endo, forthcoming)

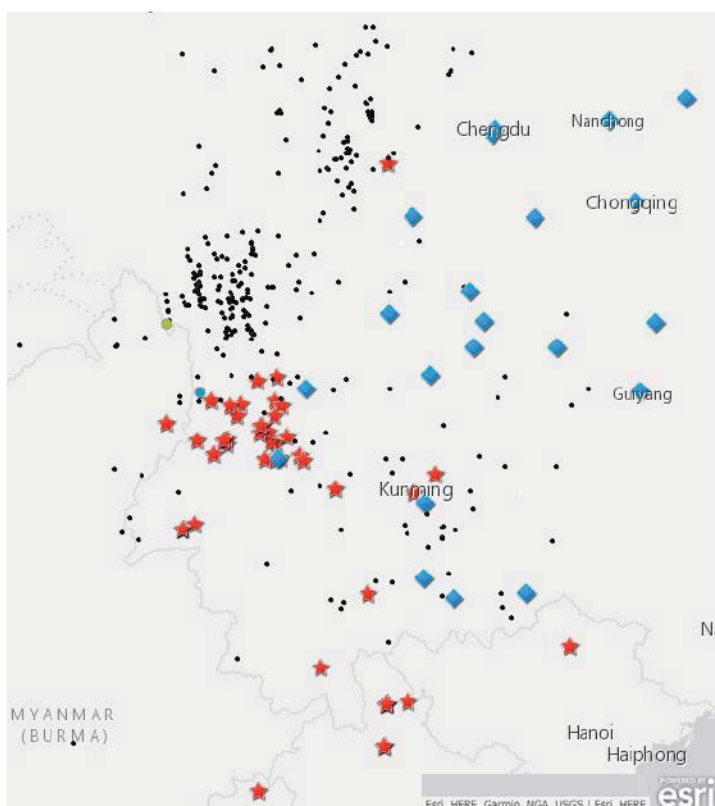
2 Goals of the project

2.1 First goal: Orthodox dialect geography inside a linguistic family

Our first goal is drawing classic model maps and giving diachronic interpretations for them according to the principles and traditions of the orthodox dialect geography.

Map 1 is a sample map which shows the geographical distribution of "iron" in Tai-Kadai (Endo, forthcoming). In the map, the arrow symbols denote the type *lek* (A-1, 2) as seen in Siamese. These types spread in the South-Western and Central Tai area, and correspond to the Old Chinese form. In the western Zhuang, there are *thi:t* types (D-1, 2) which reflect the Middle Chinese form, or are even possibly borrowed from Yue, Hakka, or Pinghua in that region at a relatively recent time. In the Buyi language, *va* or *fa* types (E-1, 2) are found. For further discussion, please see Endo (forthcoming).

2.2 Second goal: Chasing phenomena brought about by language contacts



Map 2 is cited from Suzuki (2016). The blue square denotes *ku* 穀 in Southwestern Mandarin of Chinese, the red star denotes Tibeto-Burman **kuk*. This relationship was already pointed out in STEDT. This fact implies that *ku* 穀 had a coda -k when this word was borrowed in TB. Our second goal is chasing such phenomena brought about by language contacts beyond language families.

Moreover, there are other goals, for example, extracting universal rules of diachronic changes from Asian languages, using maps to examine the relationship between linguistic properties and the external world, and so on.

Map 2 Rice plant in Yunnan province: TB and Sinitic

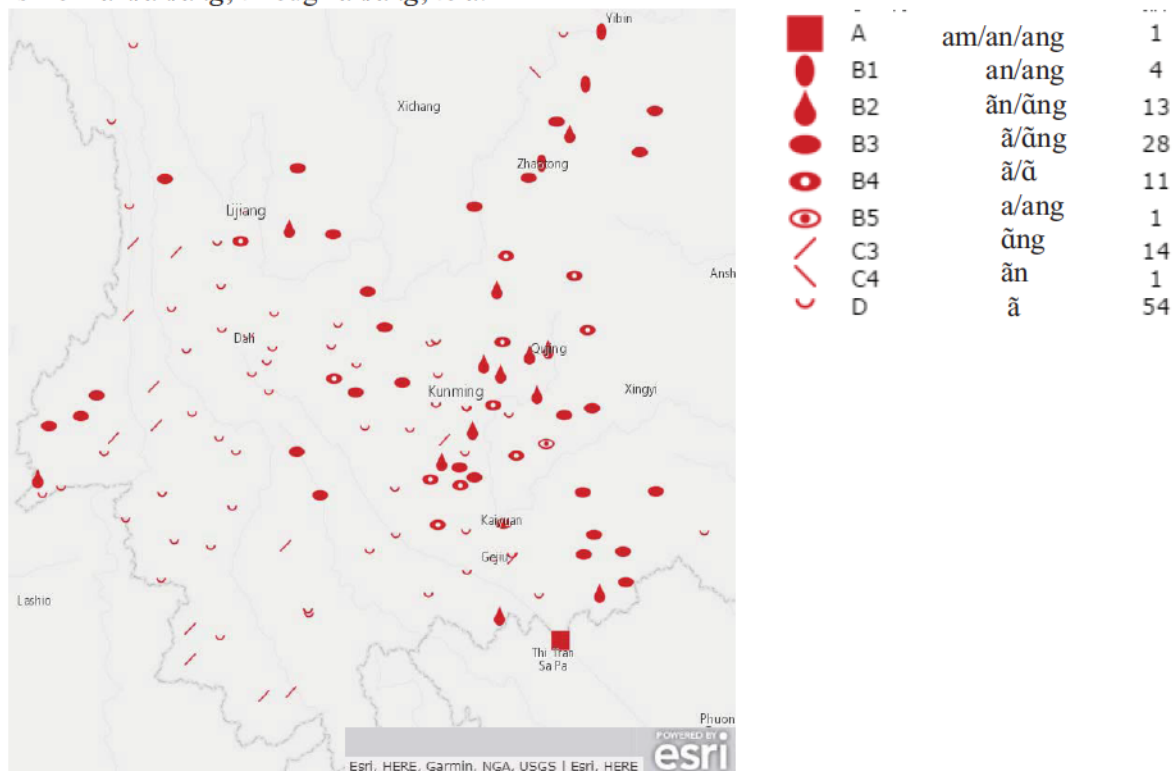
3 Colouring: Resolution for mapping the area where many language families overlap

There are three areas where many language families overlap in Asia, e.g. Northeastern Siberia, India, and mainland Southeast Asia. In order to resolve the technical problem to distinguish these language families on one and the same map, different colours are used for each language family. Here we see a concrete example in Yunnan.

3.1 Chinese

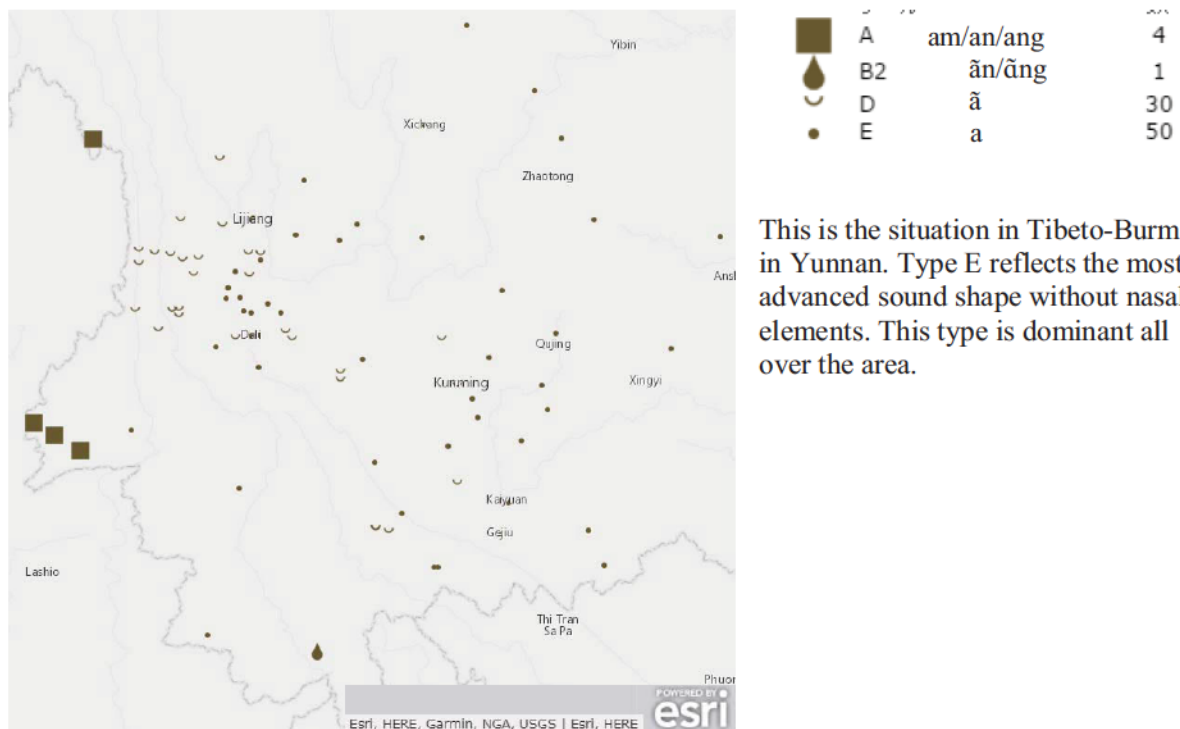
Map 3 shows the nasal coda types among Chinese dialects in Yunnan. Type A is the three-fold contrast of nasal codas -m/-n/-ng. As a matter of fact, this dialect is a variety of the Yue dialect. Their ancestors migrated from Guangdong Province in the early 20th Century in order to construct the railway between Yunnan and Vietnam. Types B1 to B5 are the two-fold contrast types. These types spread

across two thirds of Yunnan Province. Types C3, C4, and D have no contrast of nasal coda, but they still have nasal elements. These types occupy the western part of Yunnan. The chronological order of change is from *am/an/ang*, through *an/ang*, to *ã*.



Map 3 Nasal coda types among Chinese dialects in Yunnan

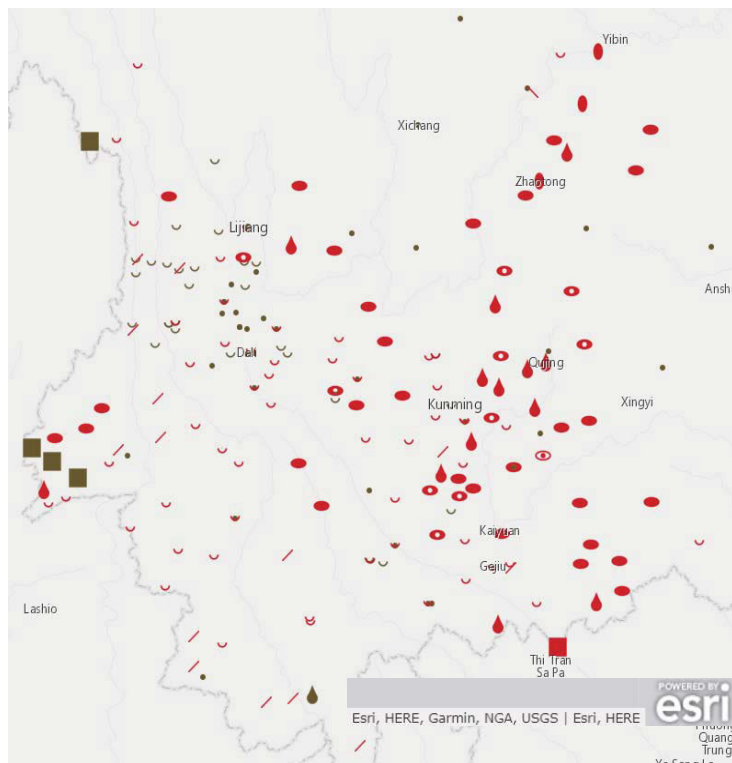
3.2 Tibeto-Burman



This is the situation in Tibeto-Burman in Yunnan. Type E reflects the most advanced sound shape without nasal elements. This type is dominant all over the area.

Map 4 Nasal coda systems after -a vowel in Tibeto-Burman

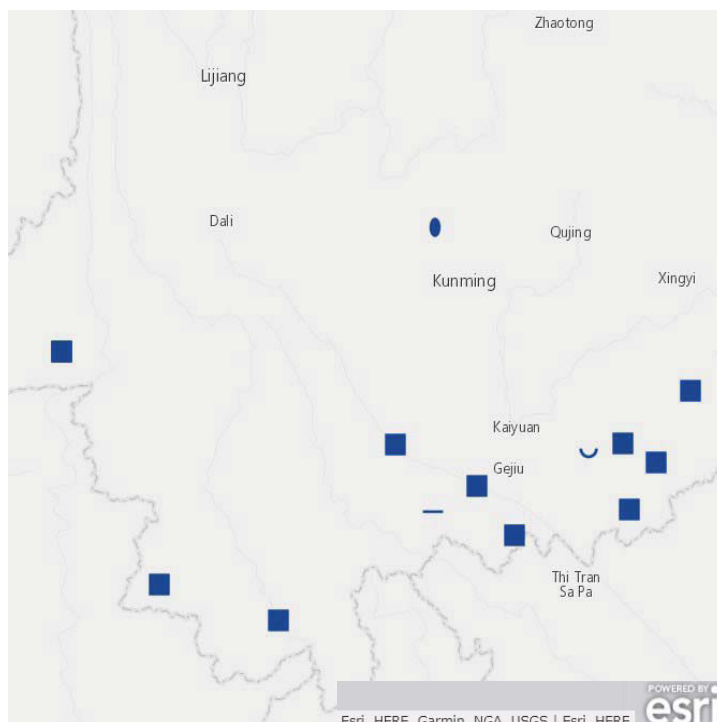
3.3 Overlapped map of Sinitic and Tibeto-Burman



A parallelism between both languages is observed in the Lijiang-Dali area. However, TB is more advanced than Sinitic in the other area.

Map 5 Overlapped map of Sinitic and Tibeto-Burman

3.4 Tai-Kadai

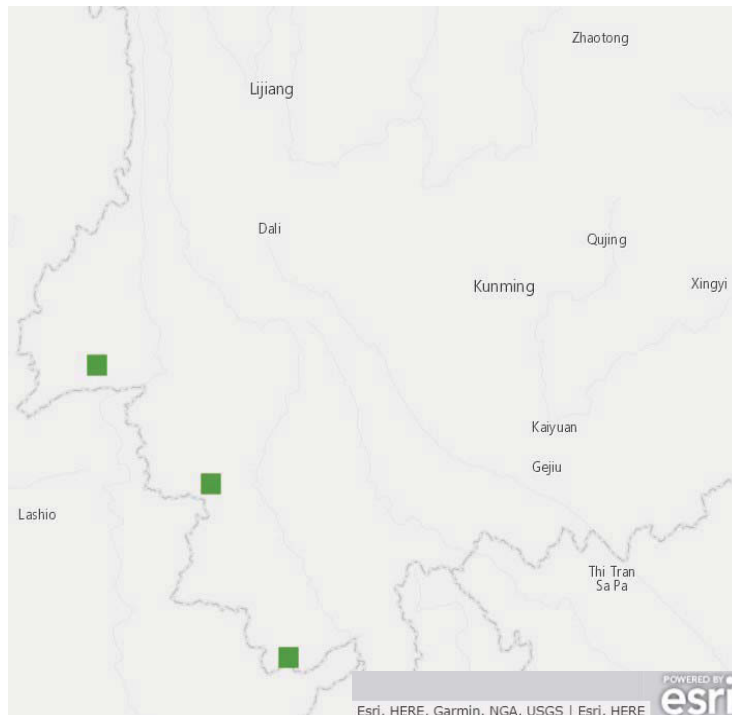


■	A	am/an/ang	11
●	B1	an/ang	1
—	C1	ang	1
⌒	D	ã	1

As for Tai-Kadai, Dai and Zhuang are spoken in Yunnan. The majority keep the three-fold distinction. However, there are some minor dialects which changed into two- or one-fold distinction types.

Map 6 Nasal coda types in Tai-Kadai in Yunnan

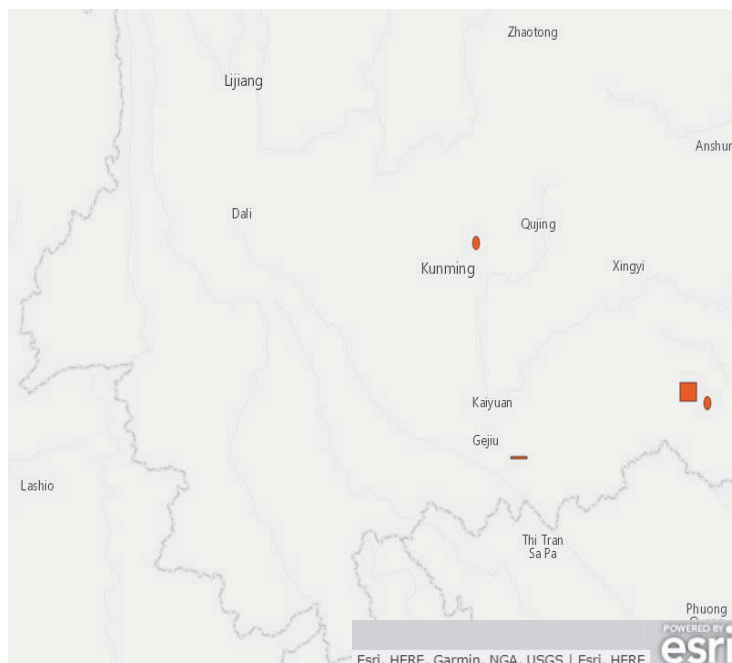
3.5 Austroasiatic






Wa is spoken in Yunnan. These three dialects belong to the three-fold type.

Map 7 Nasal coda types in Austroasiatic in Yunnan

3.6 Hmong-Mien

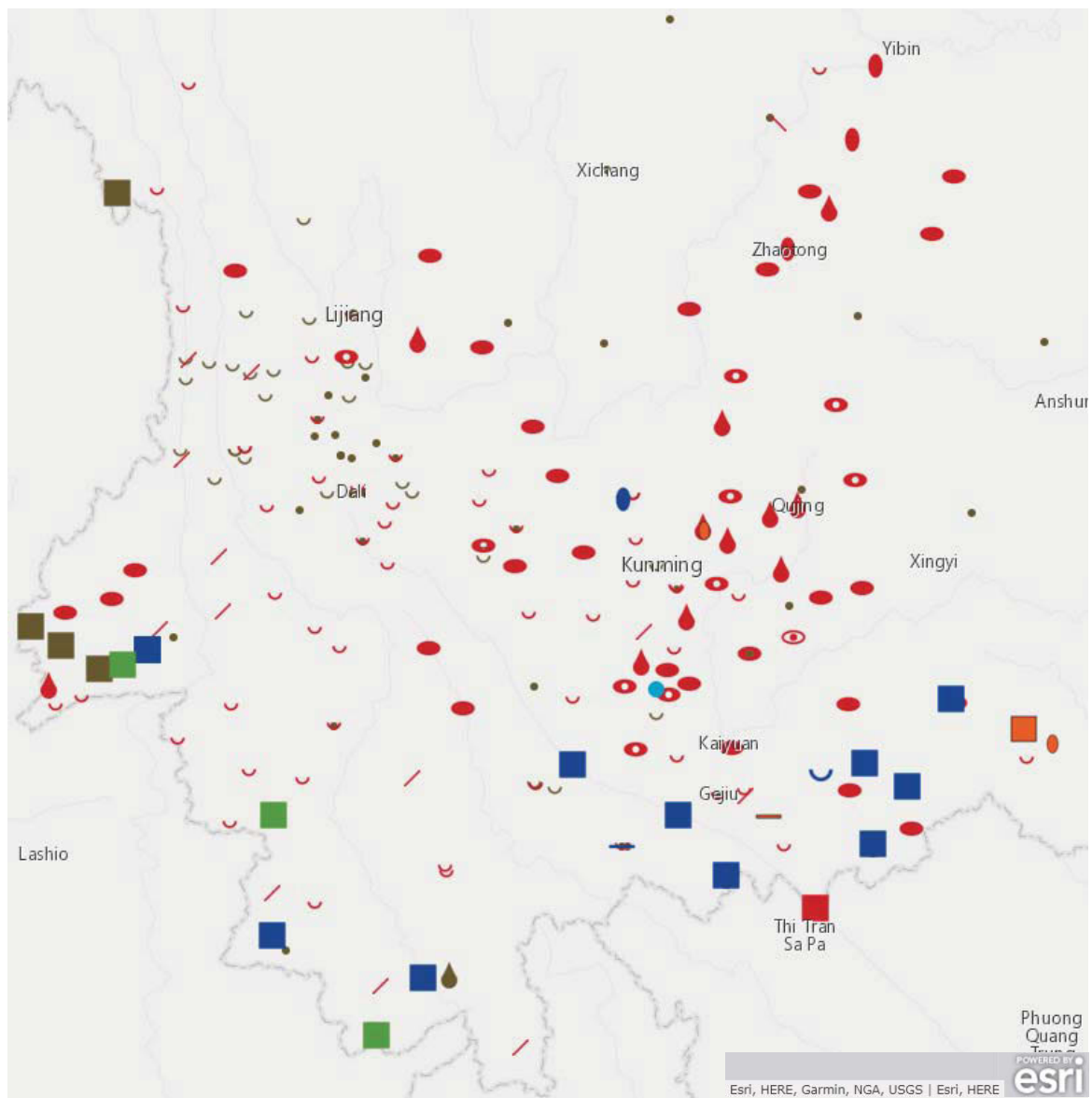


	A	am/an/ang	1
	B1	an/ang	2
	C1	ang	1

Hmong-Mien is also spoken in Yunnan. There are three types.

Map 8 Nasal coda types in Hmong-Mien in Yunnan

3.7 Synthetic map of Yunnan

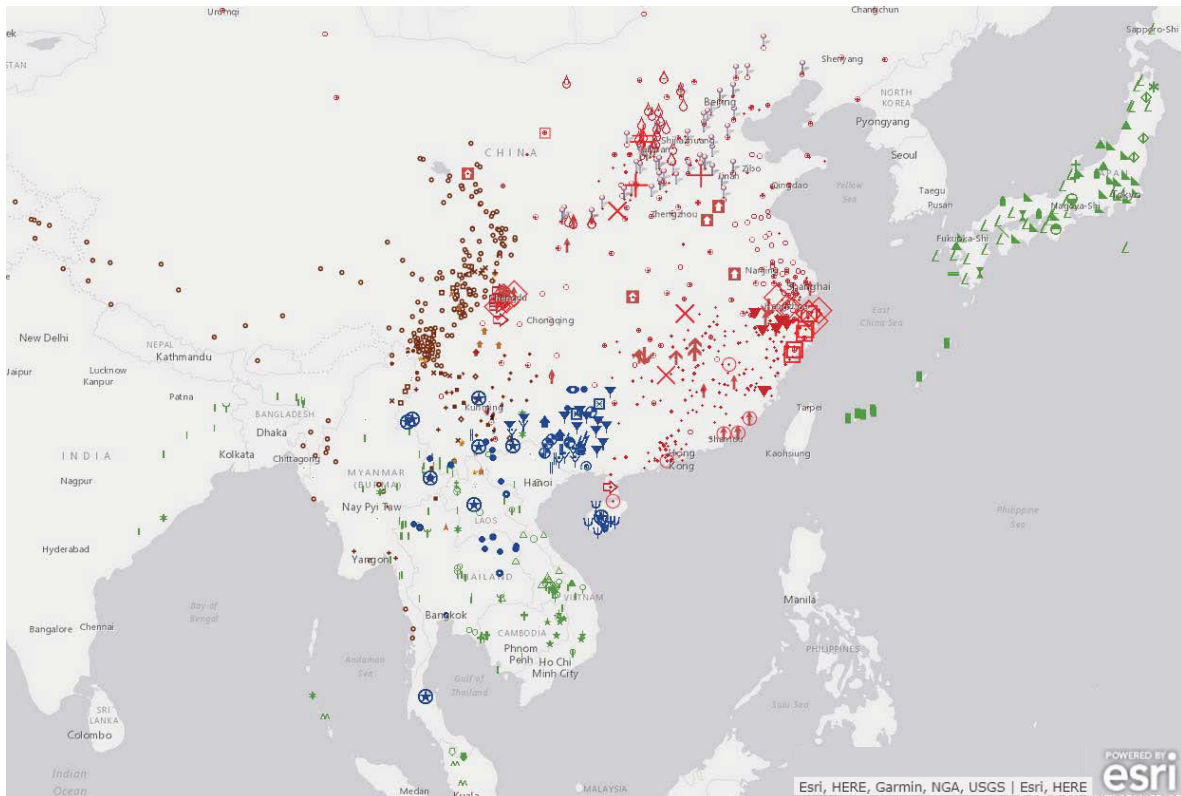


Map 9 Synthetic map of Yunnan

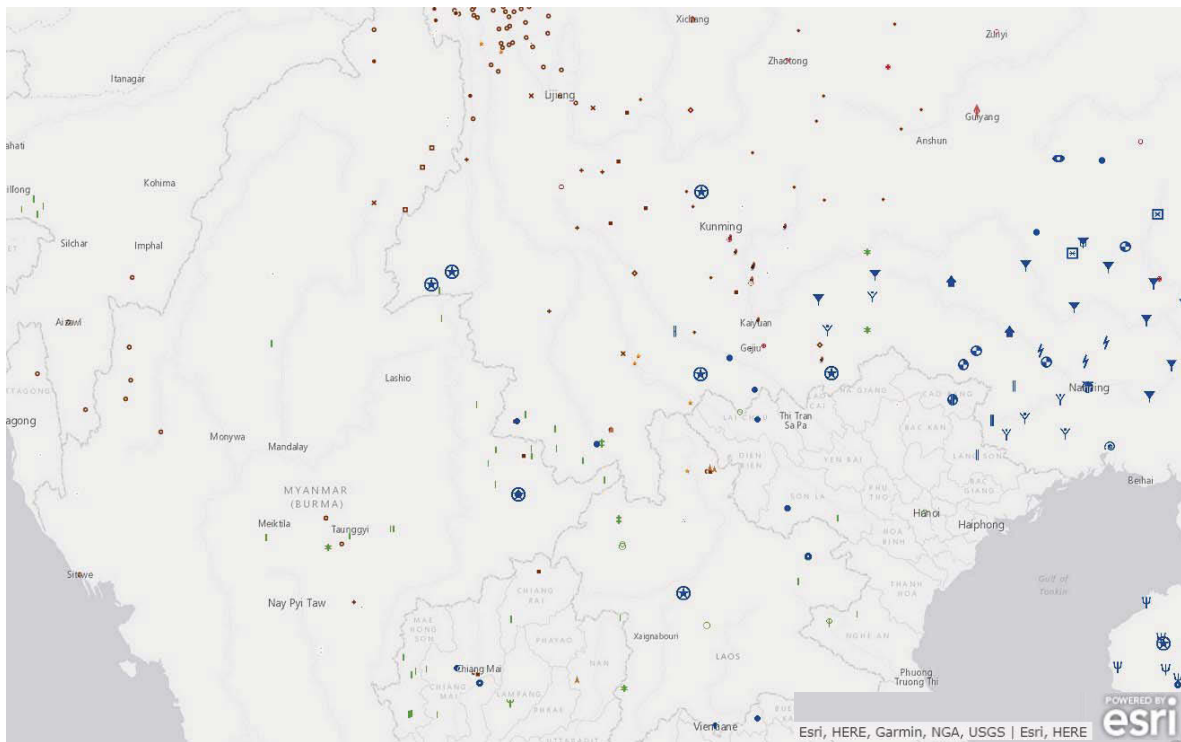
This is the synthetic map of all language families in Yunnan.

4 Synthetic map in East and Southeast Asia: The sun

Here is a sample of the synthetic map of a wider area, showing the case of "the sun" (see *Studies in Asian Geolinguistics*, I and II).



Map 10 Synthetic map in East and Southeast Asia: The sun



Map 11 Enlarged view of SE Asia

5 Future activities

5.1 International Conference on Asian Geolinguistics

A common place for scholars in this area is the International Conference on Asian Geolinguistics. The first meeting was held in Japan in 2012, the second in Thailand in 2014, and the third in Cambodia in 2016. The next meeting, the 4th conference, will be held in Indonesia in 2018. We warmly welcome the participation of many scholars. Those who are interested in this conference, please contact Mitsuaki Endo. As for the previous conferences, see <https://agsj.jimdo.com/>.

5.2 Two meetings of the project “Studies in Asian Geolinguistics” at the AA Institute in Tokyo in 2017

Moreover, two meetings of the project “Studies in Asian Geolinguistics” will be held at the AA Institute in Tokyo in 2017. The first meeting will be held on 5-6 August 2017. The theme will be “Tone and Accent in Asia and Africa”. This date was fixed for the convenience of the participants of the Sixteenth International Conference on Methods in Dialectology, which will be held on 7-11 August at the National Institute for Japanese Language and Linguistics in Tokyo. See <http://pj.ninjal.ac.jp/Methodsxxvi/>.

The second meeting will be held in November or December 2017; the theme will be “It rains”.

Both meetings will be open to the public. For those who are interested in participating, please check the site of the AA Institute: <http://www.aa.tufs.ac.jp/en/projects>, or directly contact Mitsuaki Endo.

Acknowledgements

This work was supported by KAKENHI 15K02525. I appreciate Dr. Donna Erickson for checking my English.

References

- Endo, M. forthcoming. Iron in Tai-Kadai, *Studies in Asian Geolinguistics*, V.
Suzuki, Hiroyuki. 2016. Rice (plant) in Asia, *Studies in Asian Geolinguistics*, III:1-3.

Methods and Principles of Geolinguistics

Chitsuko Fukushima

Department of International Studies and Regional Development, University of Niigata Prefecture

Abstract

Geolinguistics is a method of historical linguistics, so its aim is to reveal the history of language. In this presentation, its history, methods, and principles are introduced to invite scholars to Southeast Asian geolinguistics. By combining linguistic maps from different languages, not only the new discovery but also the theoretical reorganization might result.

1 Introduction

Geolinguistics (or linguistic geography) is a method of historical linguistics, so its aim is to reveal the history of language (Sibata 1969:11). The discipline has prospered in Europe, U.S., Japan and other countries. In this presentation, its history, methods, and principles are introduced to invite scholars to Southeast Asian geolinguistics.

2 Geolinguistics in Europe

Geolinguistics in Europe is characterized by two contrastive atlases, both of which were published about a century ago. One is *Deutscher Sprachatlas*. Georg Wenker's first work was published in 1878 and developed into the *Sprachatlas des deutschen Reichs*. Wenker started his survey, expecting to find a clear dialect boundary as an evidence of phonetic law, but isoglosses did not coincide. The atlas was based on postal survey, and colorful symbols were plotted at localities to make a linguistic map. The atlas is now digitalized and open to the public at a website called *DiWA (The Digital Wenker Atlas)*¹ (Herrgen 2007). The other is *Atlas linguistique de la France* (1902-1910) by Jules Gilliéron. The atlas was based on the surveys by interview, and word forms were printed at localities in the map. Gilliéron used the data to write papers to interpret the history of words found in the maps. Later Hans Goebel used the data in his study of dialectometry², in which statistical results were displayed in the maps (Goebel 2007).

A unique aspect of geolinguistics in Europe is that *Atlas Linguarum Europae*, namely *European Linguistic Atlas*, has been published (Vol. I, fasc. 1, Assen: Van Gorcum, 1983, and following volumes)³. A common questionnaire was used to make the atlas which covers the whole Europe. The project is still going on. In addition, many European countries now have national or regional atlases⁴. Some surveys are combined with sociolinguistics.

3 Geolinguistics in Japan

The first geolinguistic works in Japan are two atlases published in 1905 and 1906: *Phonetic Dialect Atlas* and *Grammatical Dialect Atlas* (Onishi 2007). A government-appointed language research commission made the maps using the data collected by local governments. The color-hatched maps showed the dialect boundary between Eastern dialects and Western dialects, which almost coincided with the prefectural border, the Japan Alps.

¹ DiWA <http://www.diwa.info>, Regional Language <https://regionalsprache.de/>

² Dialectometry Project - Salzburg <http://www.dialectometry.com>

³ *Atlas Linguarum Europae* <http://www.lingv.ro>

⁴ Many presentations on such works have been given at Methods in Dialectology conferences and International Society for Dialectology and Geolinguistics (SIDG) conferences.

Later a folklorist-dialectologist Kunio Yanagita made a map of dialectal words meaning “snail”, using the data from local dialect dictionaries. Based on the map, he argued that new words spread from the center of the country so that the older words remain at surrounding, distant areas. His theory of language propagation has something in common to the so-called wave theory. It is said that Yanagita got some hint from a German book on agricultural economics.

After the World War II, a Belgian Catholic priest, Willem A. Grootaers, introduced the methods of the European-style (specifically Dutch) geolinguistics to Japan. Inspired by him, Takesi Sibata started the survey for *Linguistic Atlas of Japan* (LAJ) (1966-1974) as a project of National Language Research Institute (NLRI) (the present NINJAL) and the linguistic survey in Itoigawa with Grootaers, Munemasa Tokugawa, and Yoshio Mase. The latter survey, conducted in a local city Itoigawa, produced many academic papers and developed new methods such as a survey at all hamlets in the area, glottogram (graph which crosses age and geographical factors), and a survey of all residents at a hamlet. Sibata wrote a book entitled *Methods in Linguistic Geography* (1969) and proposed the principles to interpret linguistic maps (Fukushima 2012). He believed that linguistic maps must be drawn to clearly reflect the researcher’s interpretation.

Thanks to these efforts, geolinguistics in Japan started to thrive in 1970s, and hundreds of surveys of different sizes have been conducted throughout Japan (Onishi 2016). As additional nation-wide linguistic atlases, *Grammar Atlas of Japanese Dialects* (GAJ) was published in 1989-2006, and *New Linguistic Atlas of Japan* (NLJ) (based on the FPJD survey) was published in 2016.

4 Methods of Geolinguistics

4.1 Survey Design

The geolinguistic data should be collected with the same criteria. For example, responses should be obtained from native informants from the same generation using the same questionnaire. Thus when we design a survey, we need to consider the following points: size and location of the area (whole country, region, prefecture, county, or city, etc.), character of informants (age, gender, etc.), localities (all hamlets in the area, localities in equal distance, etc.), representativeness (one informant, several informants, informants chosen by random sampling, etc.), focused type of data (vocabulary, phonology, morphology, syntax, folklore, etc.), method of data collection (interview, postal survey, tape-recorded conversation, dialect dictionaries or documents, etc.).

4.2 Data Analysis and Map Drawing

The way of data analysis used to be so simple that we made survey reports and data cards manually and then made linguistic maps using stamps or seals. However, we have been using personal computers for geolinguistics since 1980s. Now the data are digitalized and stored as documents or database files. There used to be no software to draw a linguistic map in early days, so original programs had to be created by the researchers: for example, GLAPS, SEAL⁵, EGL, etc. were made in Japan. Then, general purpose programs such as Word, Excel, Acrobat or Illustrator were used for map-drawing. Now the GIS (geographic information system) software, such as Mandara, ArcGIS, SIS, etc. is available. In the GIS system, the latitude/longitude information is combined with the linguistic data to make a file. GIS is useful to superimpose or combine different maps and draw maps in various styles as shown below. Nowadays the publication of data as well as maps is advised because this makes it possible for other researchers to reanalyze the data: for example, the GAJ and FPJD data are published in Japanese geolinguistics⁶.

5 Principles of Geolinguistics

⁵ A Room for Linguistic Geography (SEAL download) <http://www.unii.ac.jp/~chitsuko/english/>

⁶ Publication of Data and Programs (GAJ and FPJD) http://www2.ninjal.ac.jp/hogen/dp/dp_index.html

As stated in 3, linguistic maps should be drawn to reflect the researcher's interpretation. Then how can we interpret the distribution in a linguistic map? Two important principles were proposed by Sibata (1969). One is "the principle of adjacent distributions". As suggested by the wave theory, language spreads as if it crawls on the ground (or the wave spreads when a stone is thrown into a pond). Thus, if word forms a, b, and c are used in this order at adjacent localities A, B, and C respectively, then the chronological order of these forms must be either $a > b > c$ or $c > b > a$; neither $a > c > b$, $b > a > c$, $b > c > a$, nor $c > a > b$. Another is "the principle of surrounding distributions". Language tends to spread from the center of the area to the surrounding localities. Thus, if word forms a, b, a are used in this order at adjacent localities A, B, and C respectively and B is the center of the area, then the chronological order of these forms at the center must be $a > b$. When these principles are combined, it is possible for new words to jump to the center of a local area and spread to the surrounding area as flames leap to a distant house. In this case, language does not always spread like creeping on the ground. Also, we should examine if independent changes are not possible at distant localities where the same linguistic features are used.

Based on these principles, the dialectal distribution of a map can be interpreted. See Figure 1 map of "cat" in Tokunoshima, Amami, Japan (Sibata et al. 1977). The center of this island is located on the southeastern part. The chronological order of forms at the center of this area is considered to be *maju* > *mjau* > *njau*. We can see language spread from the southeastern center but also leapt to some points.

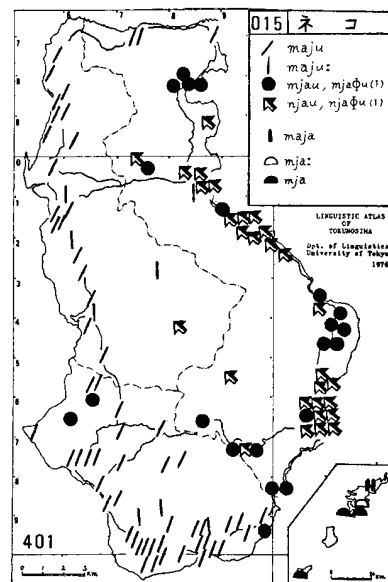


Figure 1 "cat" in Linguistic Atlas of Tokunoshima

6 Information Technology and Geolinguistic Data

When you map using personal computers, especially making use of the GIS system, it is easy to integrate, compare, or superimpose linguistic maps. You do these for different purposes.

6.1 Integration of linguistic maps

When you make linguistic maps from a linguistic survey, you can find similar distributions. These are explained by internal reasons or external reasons. Phonological maps and morphological maps tend to show similar distributions because the variation is supported by the phonological system or morphological system. This is considered as internal reasons. On the other hand, lexical maps tend to vary and are influenced by the topographical conditions. For example, language may stop spreading facing the high mountains, or language spread along the highway connecting big cities. This is considered as external reasons. In either case, you can integrate similar distributions to show the actual situation of dialect continuum. See Figure 2 and 3, maps from the southwestern Izumo linguistic survey: integrated morphological maps and lexical maps. These maps are related with the variation of the old *-au* forms and drawn with SEAL developed by the author (Fukushima 2010).

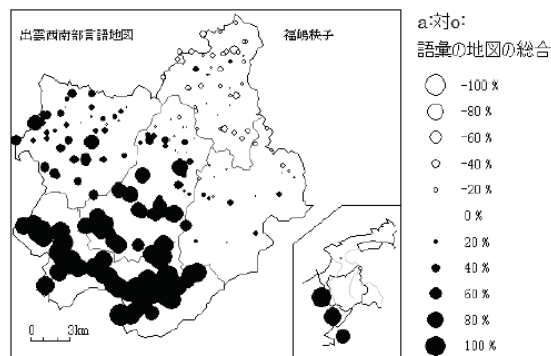


Figure 2 Integrated morphological maps

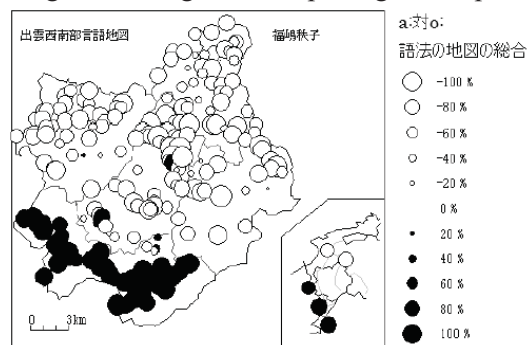


Figure 3 Integrated lexical maps

You can add up the number of word forms with a particular function used at a locality, make a table and draw small graphs at each locality. Figure 4 maps the frequency of a group of quasi-nominal particles originating from *-gan* (Fukushima in press). The map was drawn with the GIS soft SIS.

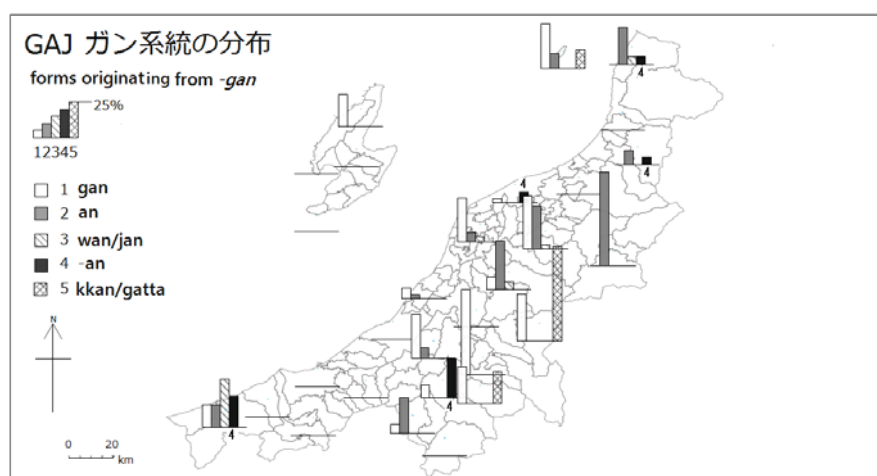
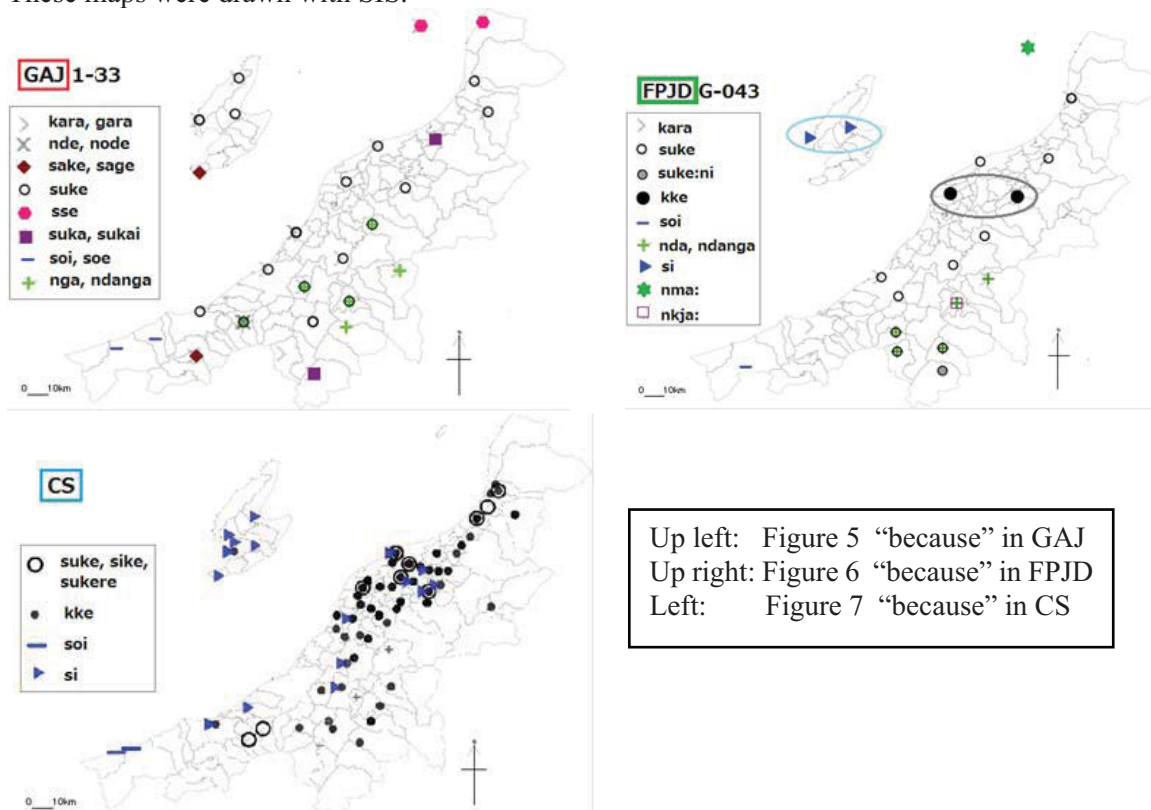


Figure 4 Quasi-nominal particles originating from *-gan*

6.2 Comparison of linguistic maps

You can compare linguistic maps from different surveys to trace linguistic changes. The author studied Niigata dialects situated on the border of the western dialects and eastern dialects of the Japanese language. Two national survey data of elderly informants, GAJ and FPJD were compared with the data of contemporary college students, CS [College Students] (Fukushima 2016). See Figure 5, 6 and 7 “because”. The oldest map of GAJ shows distributions of *suke* and its variants. The recent map of FPJD shows patches of new words *kke* and *si*. The newest map of CS has wider distributions of the two new forms. The form *kke* changed from *suke*, and *si* is a dialectal form which spread from western dialects. These maps were drawn with SIS.



Up left: Figure 5 “because” in GAJ
Up right: Figure 6 “because” in FPJD
Left: Figure 7 “because” in CS

6.3 Superimposition of linguistic maps

Yet another important function of the GIS software is to superimpose linguistic maps. You can easily confirm the change (or no change) by doing this (Fukushima 2016). See Figure 8 “bought” a superimposed map of GAJ, FPJD and CS. The distribution of western dialect form *ko:ta* from three surveys is superimposed on a map. This map was drawn with SIS.

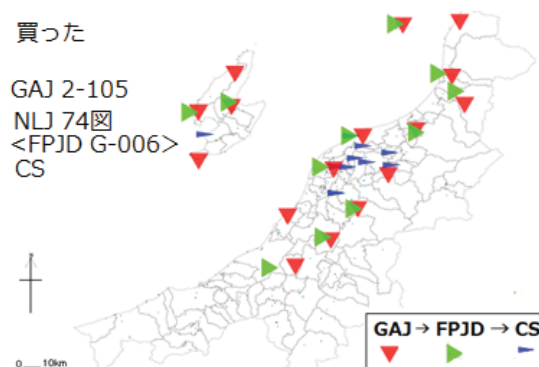


Figure 8 “bought” a superimposed map of GAJ, FPJD and CS.

This method can be used to combine linguistic maps from different languages. As Mitsuaki Endo has already presented, we are now involved in the making of *The Linguistic Atlas of Asia* (Endo 2016)⁷. The experts of each language (or language family) are making linguistic maps using ArcGIS online⁸. The basic color of each language (or language family) is decided, so if we only superimpose these maps online, then the map for the prospective *Linguistic Atlas of Asia* is ready. You can learn how you do this in the next presentation by Mika Fukazawa.

7 Conclusion

Following the long tradition of geolinguistics, we have been working to make *The Linguistic Atlas of Asia*. This has been made possible thanks to the cooperation among many scholars, the aggregated research data, and the development of technology. By reading these linguistic maps from different languages, we could find trends of language changes and connections between some languages. We found common expressions considered as the example of calques. For example, as the names for the Sun, expressions similar to “the eye of the sky” were shared by languages in the region from the southern China to Indochina, and to Indonesia (Fukushima 2016b). We are making lexical, phonological, and grammatical maps. Not only the new discovery but also the theoretical reorganization might result from this project.

Acknowledgements

This paper is part of the outcomes of the collaborative research project “Field-Research Project for Analyzing the Formation Process of Japanese Dialects”, carried out at the National Institute for Japanese Language and Linguistics (NINJAL), and the GAJ and FPJD data published by NINJAL have been utilized. This study was supported by JSPS Grant-in-Aid for Scientific Research (A) 23242024 and (C) 16K02688. I feel much grateful to those institutions. Thanks also go to Research Institute for Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies for offering the grant for ILCAA Joint Research Project: Studies in Asian Geolinguistics.

⁷ Asian Geolinguistics <https://agsj.jimdo.com>

⁸ ArcGIS online <https://www.arcgis.com/home/>

References

- Endo, Mitsuaki. 2016. Goals of the Project “Studies in Asian Geolinguistics” 2015-2017. *Studies in Asian Geolinguistics* I: 1-4.
- Fukushima, Chitsuko. 2010. Bumpu o doo yomuka [How do you interpret distributions?]. In: Zendo Uwano (supervision), *Nihongokenkyuu no 12 shoo [Twelve Chapters on Japanese Linguistics]*: 35-47. Meiji Shoin, Tokyo, Japan.
- Fukushima, Chitsuko. 2012. First Dialectologists: Takesi Sibata. *Dialectologia*, 8 (2012):175-181.
- Fukushima, Chitsuko. 2016a. Tracing real and apparent time language changes by comparing linguistic maps. In: Marie-Hélène Côté, Remco Knooihuizen and John Nerbonne (Eds.), *The future of dialects: Selected papers from Methods in Dialectology XV*: 363-376. Language Science Press, Berlin, Germany.
- Fukushima, Chitsuko. 2016b. Sun in Asia. *Studies in Asian Geolinguistics* II: 1-2.
- Fukushima, Chitsuko. In press. Hoogen bumpu no soogoo to hikaku kara miru hoogen no chiikisa to henka [Dialectal variation and change from the viewpoint of integration and comparison of dialectal distributions]. *Hoogen no kenkyuu [Studies in Dialects]* 3.
- Goebl, Hans. 2007. Dialectometry: Theoretical Prerequisites, Practical Problems, and Concrete Applications (mainly with examples drawn from the ‘Atlas Linguistique de la France’, 1902-1910). In: *Geolinguistics around the World: Proceedings of the 14th NIJL International Symposium*: 65-74. NINJAL, Tokyo, Japan.
- Herrgen, Joachim. 2007. Dialectology - Digital and Interactive: The Digital Wenker Atlas - DiWA. In: *Geolinguistics around the World: Proceedings of the 14th NIJL International Symposium*: 29-34. NINJAL, Tokyo, Japan.
- Language Research Commission (ed.). 1905. *On'in bunpu-zu [Phonetic Dialect Atlas]*. Tokyo, Japan.
- Language Research Commission] (ed.). 1906. *Kogo-ho bunpu-zu [Grammatical Dialect Atlas]*. Tokyo, Japan.
- National Language Research Institute (ed.). 1966-1974. *Nihon gengo chizu [Linguistic Atlas of Japan]*. 6 Vols. National Printing Bureau, Tokyo, Japan.
- National Language Research Institute (ed.). 1989-2006. *Hoogen bumpoo zenkoku chizu [Grammar Atlas of Japanese Dialects]*. 6 Vols. National Printing Bureau, Tokyo, Japan.
- National Institute for Japanese Language (NINJAL) (ed.). 2007. *Geolinguistics around the World: Proceedings of the 14th NIJL International Symposium*. National Institute for Japanese Language, Tokyo, Japan.
- Onishi, Takuichiro. 2007. Mapping Japanese Dialects. In: *Geolinguistics around the World: Proceedings of the 14th NIJL International Symposium*: 1-6. NINJAL, Tokyo, Japan.
- Onishi, Takuichiro. 2016. Hoogen chirigaku no kenkyuu dookoo [Research Trends in Japanese Geolinguistics]. *Hoogen no kenkyuu [Studies in Dialects]* 2: 83-97.
- Onishi, Takuichiro (ed.). 2016. *Shin nihon gengo chizu [New Linguistic Atlas of Japan]*. Asakura Shoten, Tokyo, Japan.
- Sibata, Takesi. 1969. *Gengo chirigaku no hoohoo [Methods in Linguistic Geography]*. Chikuma Shobo, Tokyo, Japan.
- Sibata, Takesi et al. 1977. *Amami Tokunoshima no kotoba: bumpu kara rekishi e [Language in Tokunoshima, Amami: from Distributions to History]*. Akiyama Shoten, Tokyo, Japan.
- Yanagita, Kunio. 1930. *Kagyuukoo [On the Lexicon of Snails]*. Tookoo Shoin. Tokyo, Japan.

A Manual for Drawing Geolinguistic Maps with ArcGIS Online

Mika Fukazawa
Sapporo Gakuin University, JSPS postdoctoral research fellow

Abstract

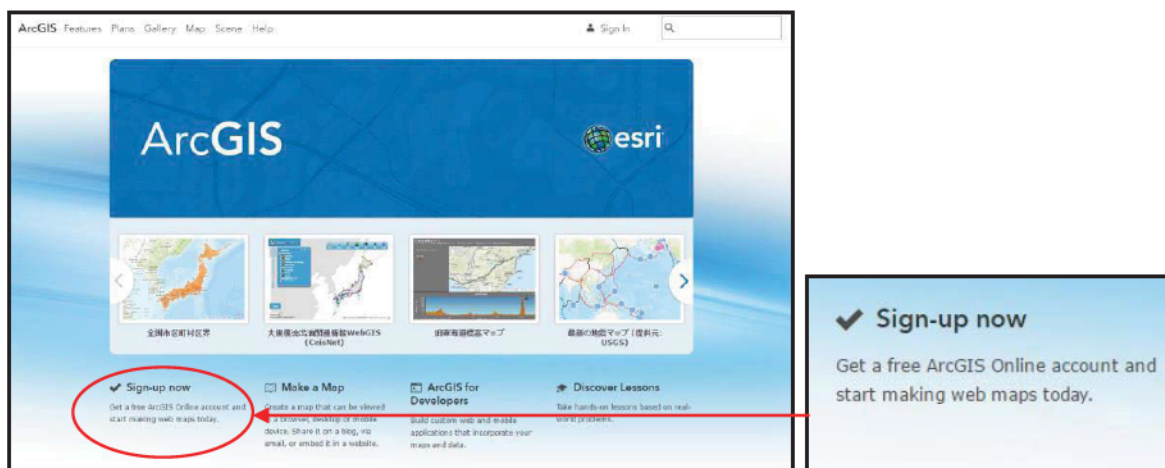
This is a manual of ArcGIS Online, hosted by Esri, for drawing geolinguistic maps for a geolinguistic study. In this paper, I attempted to provide a simple explanation of how to make a map specifically for geolinguistics. Sections 2, 3 and 4 are the basic manual, referring to the “ArcGIS Online Help” page and Endo’s (2016) “ArcGIS Online による方言地図作成法、βバージョン 1.0 [How to make geolinguistical maps using ArcGIS Online], β ver. 1.0.” In Section 5, I illustrated two geolinguistic maps of the words for “iron” in Ainu and Japanese, which I made with ArcGIS Online.

1 Introduction

ArcGIS Online is "a scalable and secure software-as-a-service" hosted by Esri. You can access it from the address <https://www.arcgis.com/home/index.html>. Section 2 shows how to create an ArcGIS public account, and Section 3 describes how to make a map with ArcGIS Online. In Section 4, I provide some tips for making a beautiful and technical map. Section 5 includes two geolinguistic maps of the words for “iron” in Ainu and Japanese, made by ArcGIS Online. These maps are also used for the pictures in this manual.

2 Create an ArcGIS Public Account and Sign in¹

1. First of all, you need to create your public account on ArcGIS online. ArcGIS public accounts are available for individuals. Click “Sign up now.”



(<https://www.arcgis.com/home/index.html>)

¹ Section 2 is based on the “Create a Public Account” page
<http://doc.arcgis.com/en/arcgis-online/reference/create-account.htm> (accessed 2017-5-3)

2. Then, click “Create a Public Account.”

The screenshot shows the ArcGIS Sign In page. On the left, under the heading "Need an ArcGIS Public Account?", there is a button labeled "CREATE A PUBLIC ACCOUNT" which is circled in red. A red arrow points from this button to a blue callout box on the right that contains the text "CREATE A PUBLIC ACCOUNT". The page also features a "Sign In" section with fields for Username and Password, and options to sign in with Facebook or Google+.

(<https://www.arcgis.com/home/signin.html>)

3. To create a public account, you can use your Facebook or Google credentials. Otherwise, click “Enter Your Information” at the bottom.

The screenshot shows the "Create Your ArcGIS Public Account" page. It offers three options: "USING FACEBOOK", "USING GOOGLE+", and "ENTER YOUR INFORMATION". The "ENTER YOUR INFORMATION" button is circled in red, and a red arrow points from it to a blue callout box labeled "ENTER YOUR INFORMATION". Below these options, there is a note: "If you have an Esri Account then you already have an ArcGIS Public Account and you can just sign in."

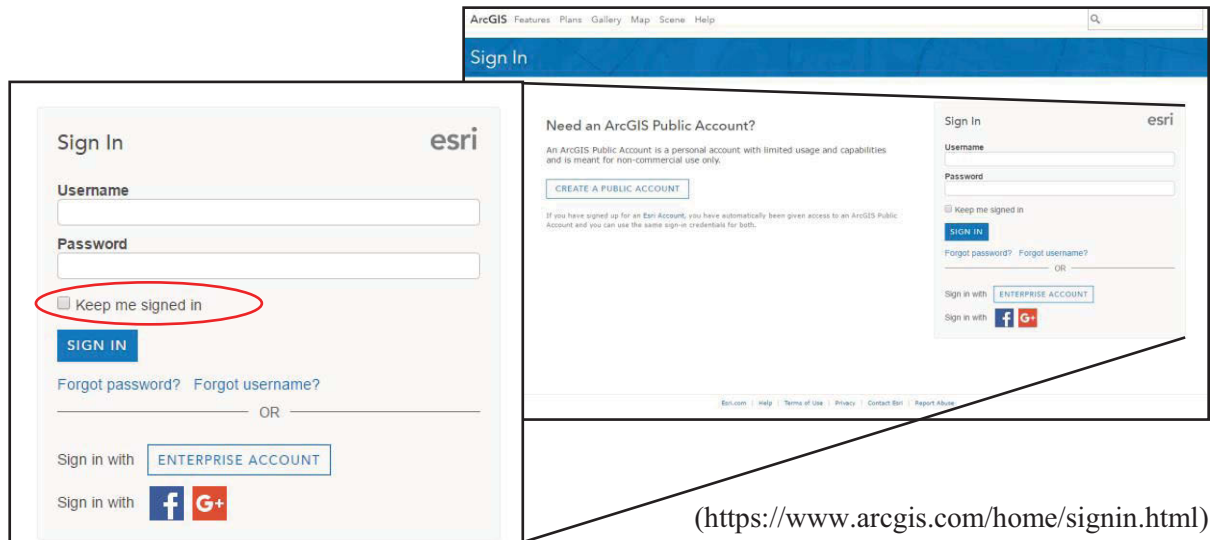
4. Enter a username,² password³ and so on, and finally, click “Create My Account.”

The screenshot shows the registration form on the "Create Your ArcGIS Public Account" page. The form includes fields for Username, Password, Confirm Password, First Name, Last Name, E-mail, Confirm E-mail, Security Question, and Answer. At the bottom, there is a checkbox for "I accept and agree to be legally bound by the Esri ArcGIS Online Terms of Use and Privacy Policy." and a button labeled "CREATE MY ACCOUNT" which is circled in red. A red arrow points from this button to a blue callout box labeled "CREATE MY ACCOUNT".

² Your username has to contain 6 to 128 alphanumeric characters, and the following special characters can also be used: . (dot), _ (underscore), and @ (at sign).

³ Your password cannot be the same as your username. Use at least 8 characters, including at least 1 letter and 1 number.

- If you check the box “Keep Me Signed in” on the “Sign in” page, you can be automatically signed in to ArcGIS Online.



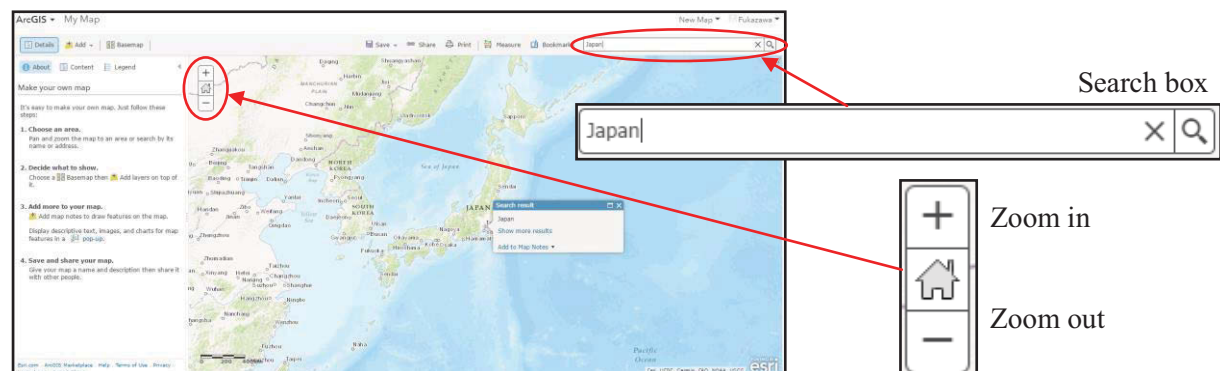
3 Make a Map

3.1 Choose a Basemap⁴

- After you are signed in, click “Make a Map.”



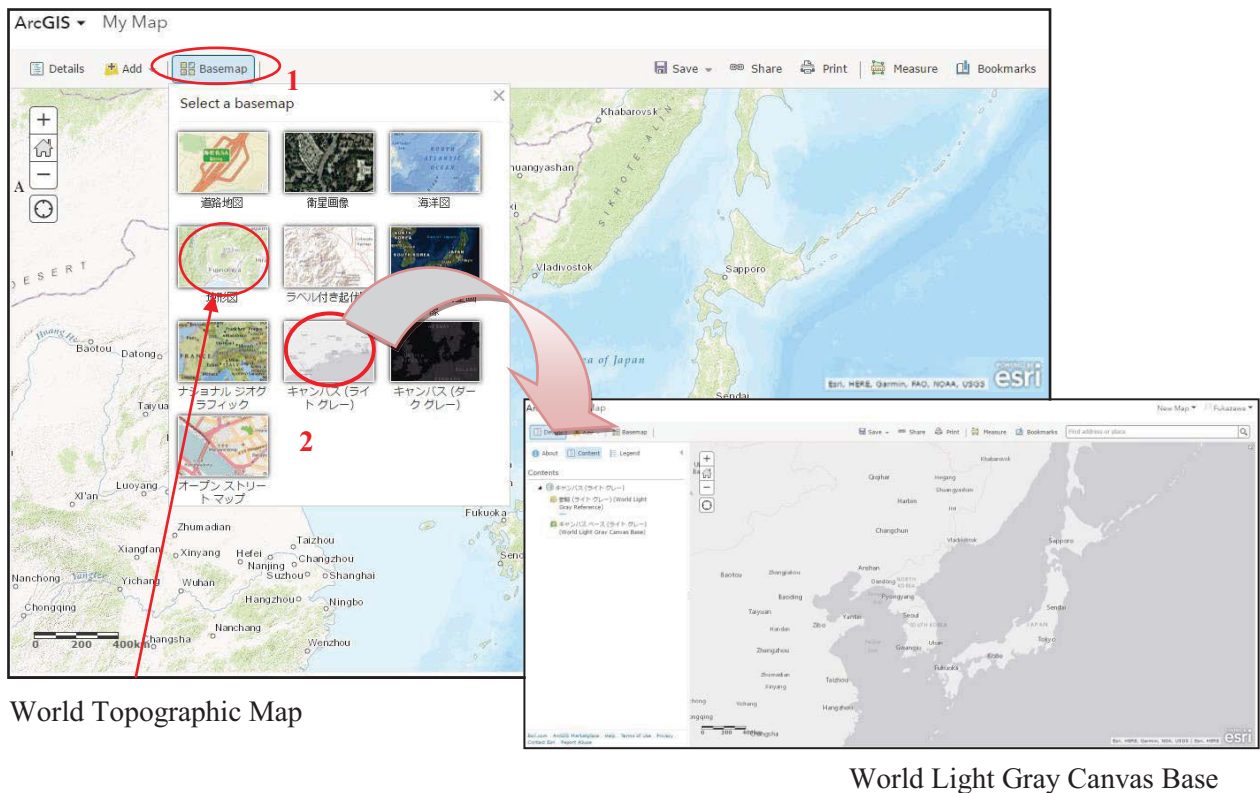
- Pan and zoom the map to the area you want to display. You can also focus on the area by entering the place name or address in the search box.



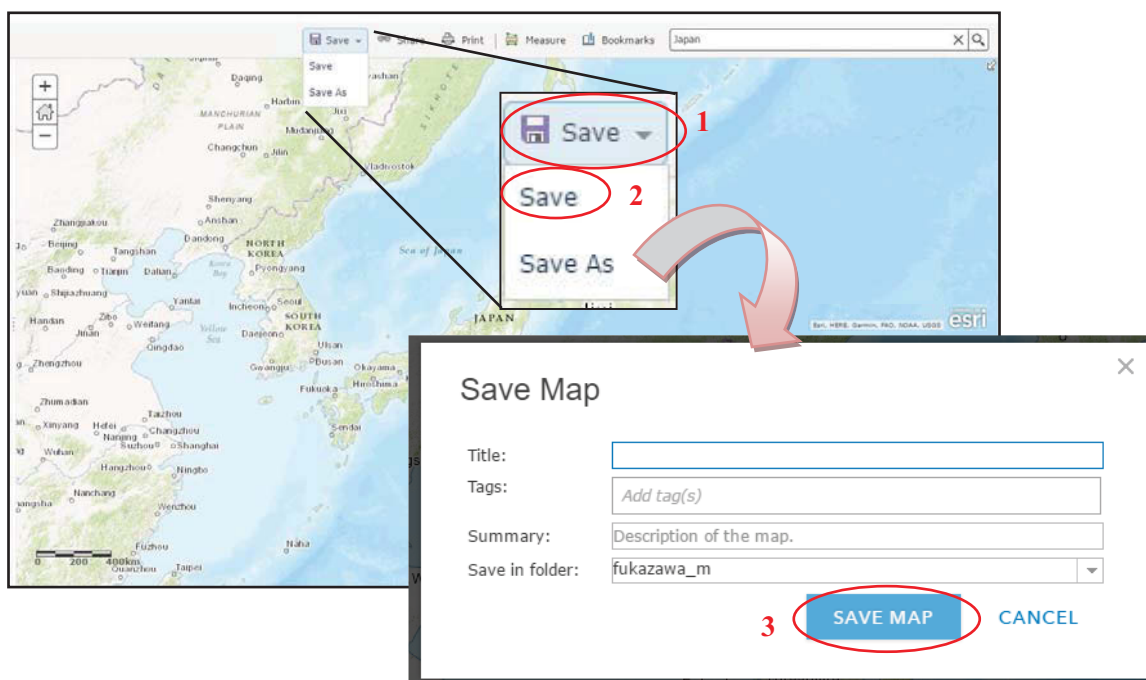
⁴ Section 3.1 is based on the “Choose Basemap” page

<http://doc.arcgis.com/en/arcgis-online/create-maps/choose-basemap.htm> (accessed 2017-5-3)

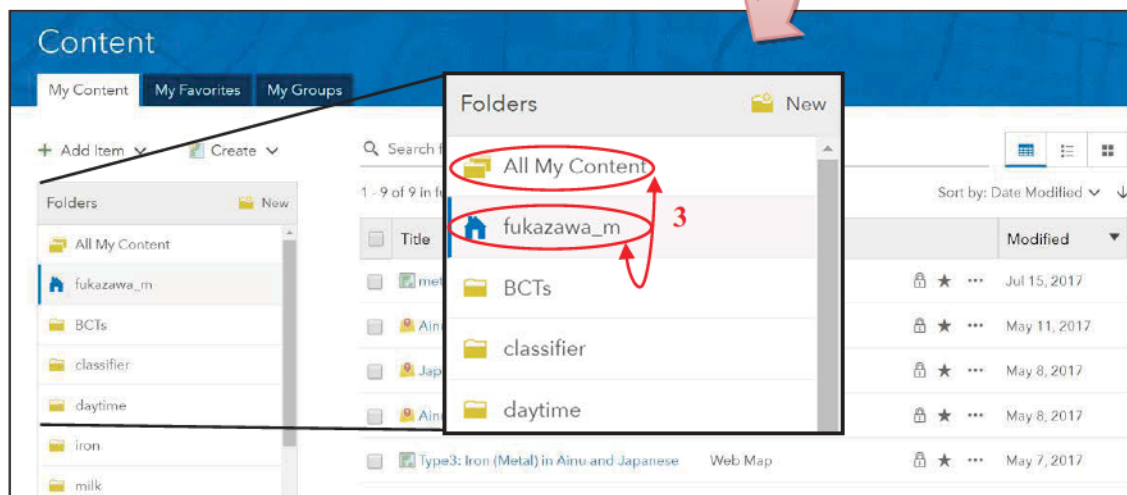
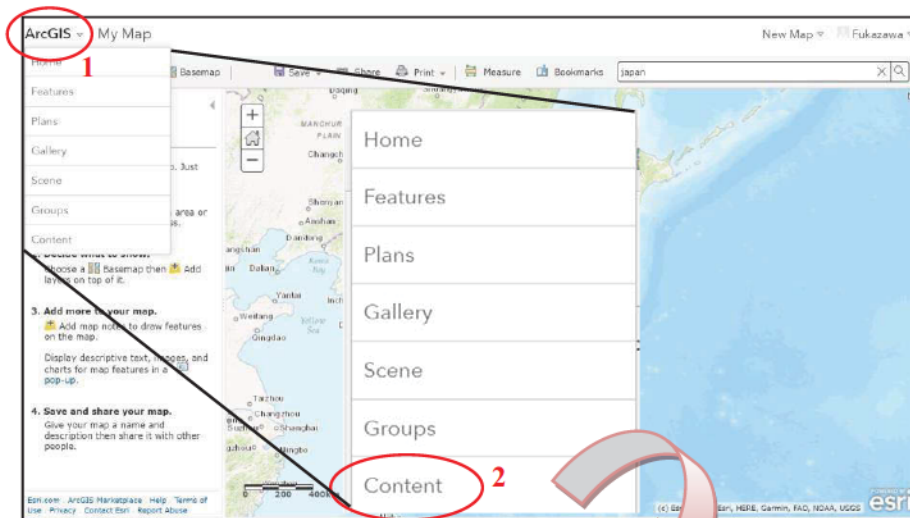
- When you click “Basemap,” ten thumbnails of basemaps are shown. The “World Topographic Map” and “World Light Gray Canvas Base” basemaps would be useful for geolinguistic maps. Select your favorite basemap, and click it.



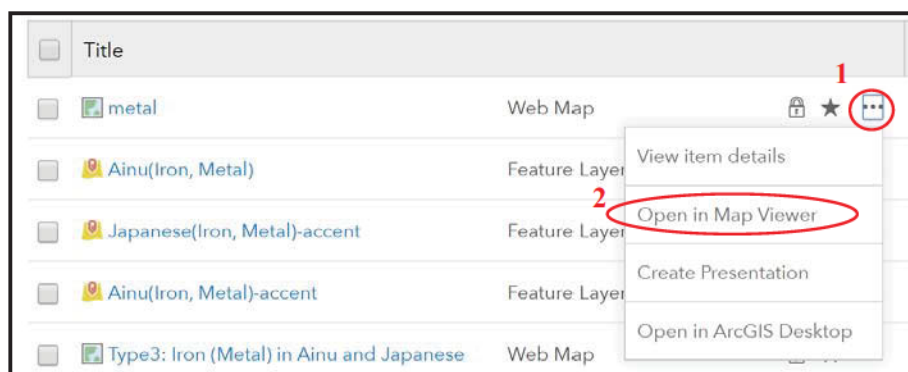
- The first time you save your new map, click “Save” and “Save.” Then, the “Save Map” page opens. Enter the title, tags and summary, and select a folder. Finally, click “Save Map.”



- The map has now been added to the “Content” folder. The next time, click “Content,” and choose “All My Content” or the folder name which you saved the map in.



- Select the map you want to open, and then click the icon “...” and “Open in Map Viewer.” Then, the map will directly open.



3.2 Prepare the Geographical Data

Prepare the geographical data for making a map. An Excel worksheet or spreadsheet is easy to use for the geographical data for ArcGIS.

1. Type the entries in the cells of the first row, as shown in Table 1:

	A	B	C	D	E	F	G	H
1	Latitude	Longitude	Place	Address	Word	Type1	Type2	Type3

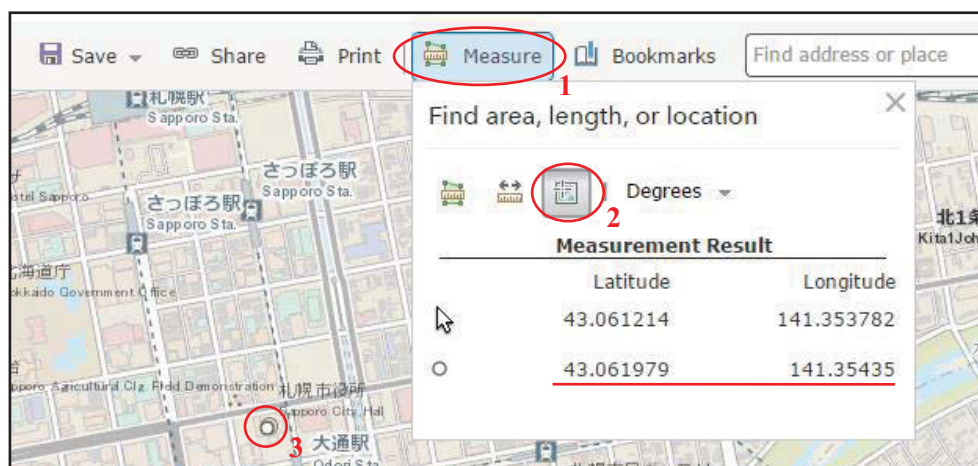
Table 1: Entries of Data in an Excel Worksheet

2. Enter the latitude, longitude, place (name), address, and word forms in columns A to E. The data must contain the location information, using the latitude-longitude data from the World Geodetic System (WGS). You can acquire it from Google Maps (<https://maps.google.com/>), ArcGIS and so on.

- Google Maps: Right-click on the map of the place for which you want to acquire data, and then click “What’s here?”



- ArcGIS: Click “Measure” and the “Location” icon. Search the map for place you want to investigate, and then click on it.



3. If one place has two word forms, insert a row for the second word. Copy and paste the location information of the place (latitude, longitude, place, address), and type the second word form in the cell of the inserted row.

- Enter the cells in column F; for example, the words for “iron” in Japanese are classified into two types: T (*tetsu* type) and K (*kane* type). In that case, T or K would be entered into the cells, as in Table 2. When the words are classified by several criteria, insert columns for the classifications, as with columns G and H in Table 2.

	A	B	C	D	E	F	G	H
1	Latitude	Longitude	Place	Address	Word	Type1	Type2	Type3
2	43.062035	141.3544	Hokkaido	北海道札幌市	tetsu	T		tetsu
3	45.30302	141.0477	Rebun	北海道礼文町	tedzi	T		tedzi
4	40.822072	140.747	Aomori	青森県青森市	tedzi	T		tedzi
5	40.822072	140.747	Aomori	青森県青森市	kane	K	2	kane
6	40.603124	140.46392	Hirosaki	青森県弘前市	tedzi	T		tedzi
7	40.603124	140.46392	Hirosaki	青森県弘前市	kane	K	2	kane

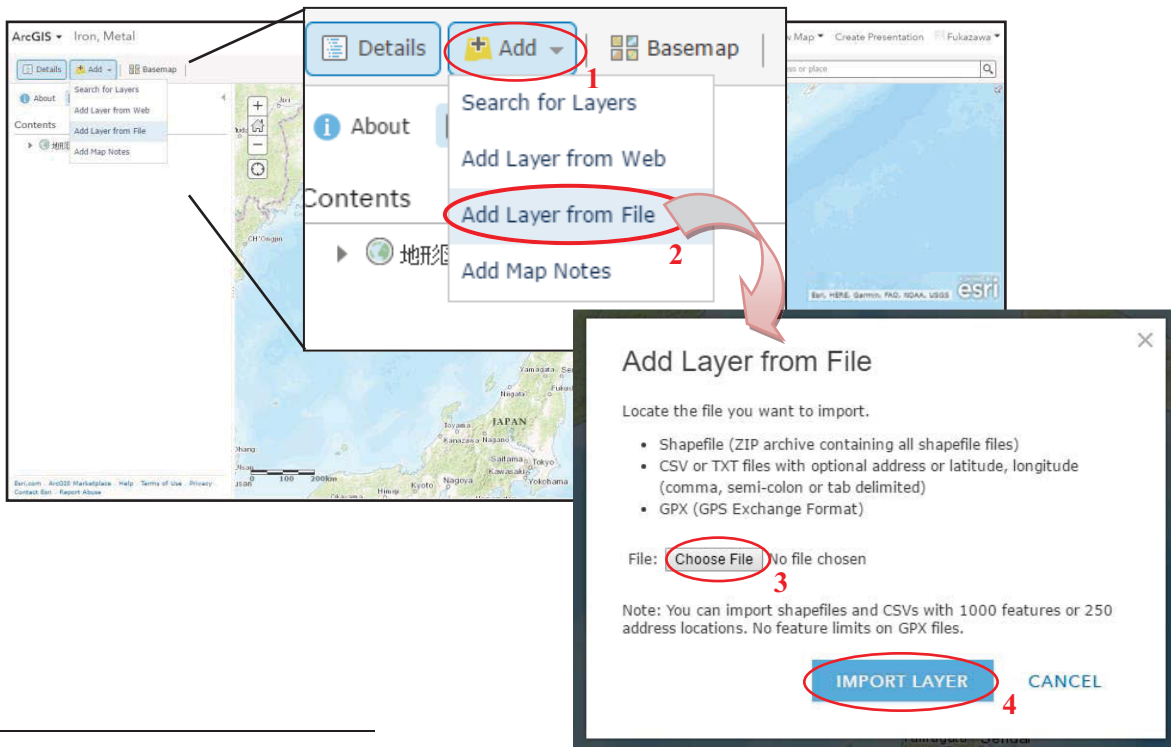
Table 2: Classification of Words into Types

- Save the workbook in the Excel file format for preservation and in the Unicode text file format (.txt) for ArcGIS.

3.3 Add Layers⁵

After choosing your basemap, you can display your text file of the geographical data, which was made in Section 3.2. In ArcGIS, the contents of a map are called “Layers.”


- Click “Add” and “Add Layer from File.”
- Choose your geographical data file from your computer, then click “Import Layer.”

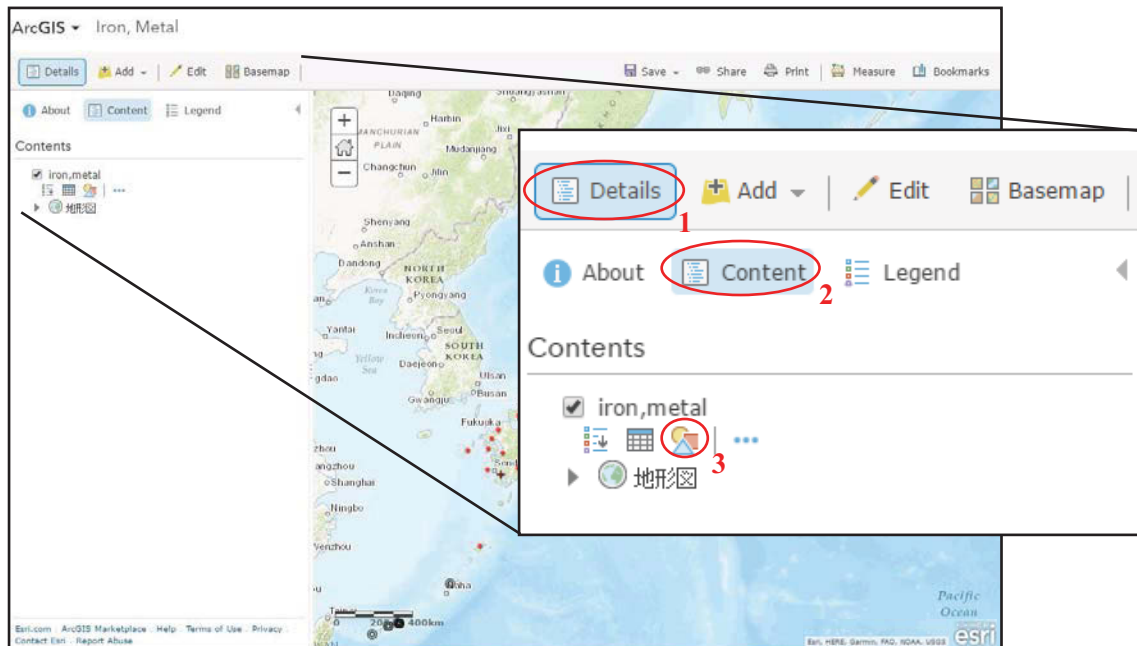



⁵ Section 3.3 is based on the “Add Layers” page

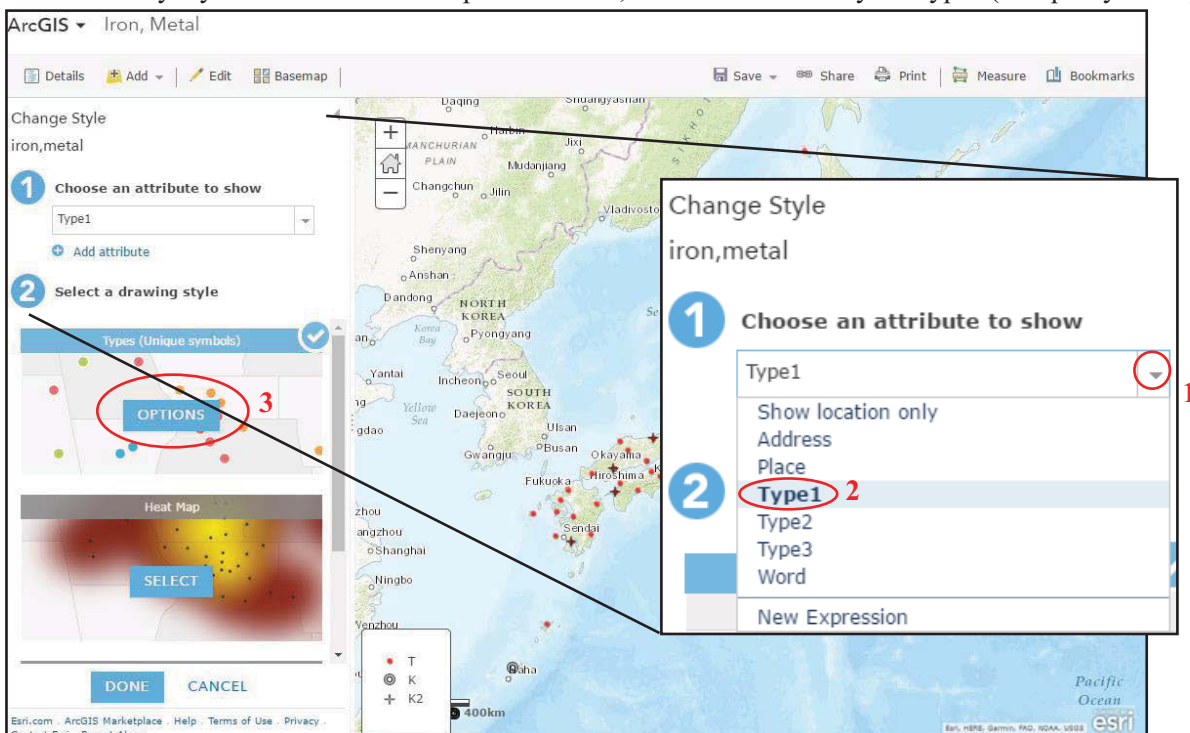
<http://doc.arcgis.com/en/arcgis-online/create-maps/add-layers.htm> (accessed 2017-5-3)

3.4 Change Style⁶

1. After you click “Import Layer,” the page for changing style opens directly. Alternatively, click “Details,” “Content” and the icon  in the list of contents of the layer you want to change.



2. In the drop-down menu, choose the attribute you want to display in your layer.
3. Select a drawing style, and click “Select.” Then, the check mark  appears on the upper right of the style you selected. Click “Options.” Here, I recommend the style “Types (Unique symbols).”



⁶ Section 3.4 is based on the “Change Style” page

<http://doc.arcgis.com/en/arcgis-online/create-maps/change-style.htm> (accessed 2017-5-3)

- Click the symbol to edit. On the editorial page, you can change the shape, color and size of the symbol you selected. Choose them and click “OK” and “Done.”

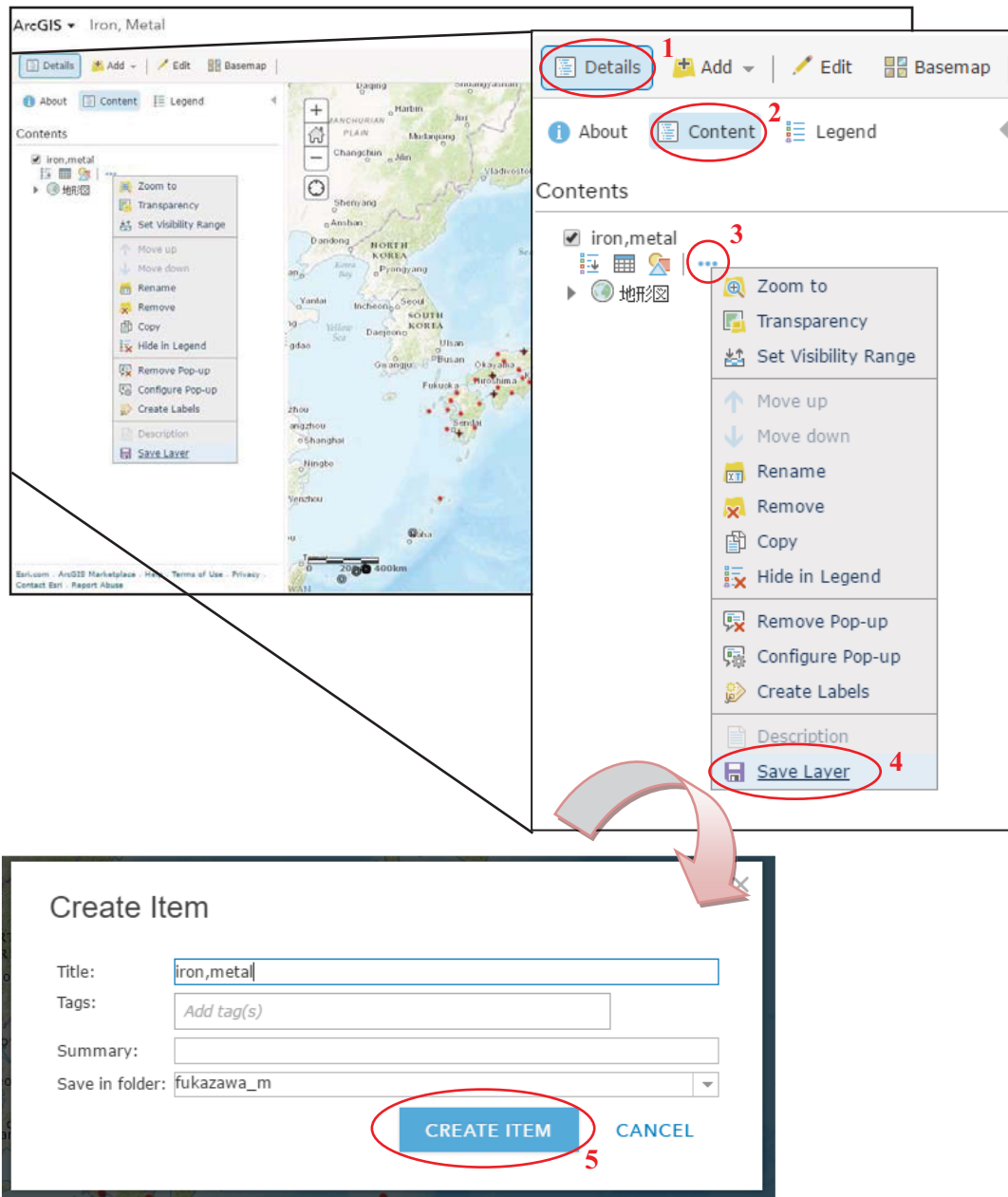
The following table represents the data shown in the 'Type1' symbol editor:

LABEL	COUNT
T	65
K	16
K2	7
Other	0

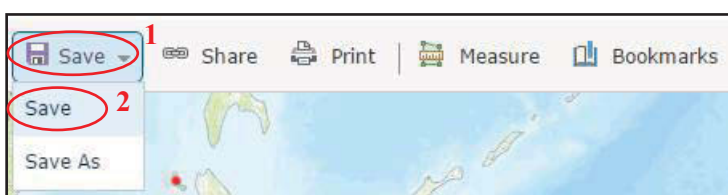
The 'Change Style' panel shows the following options:

- 1 Choose an attribute to show:** Type1
- 2 Select a drawing style:** Heat Map, Location (Single symbol)
- 6 DONE** (highlighted)

- To save the layer, click “Details” and “Content.” In the icons of the layer you want to save, click the “More Options” (**...**) button and “Save Layer.”
- If you are saving the layer for the first time, type the title, tags and summary, then choose a folder and click “Create Item.” The next time, this page will be skipped.



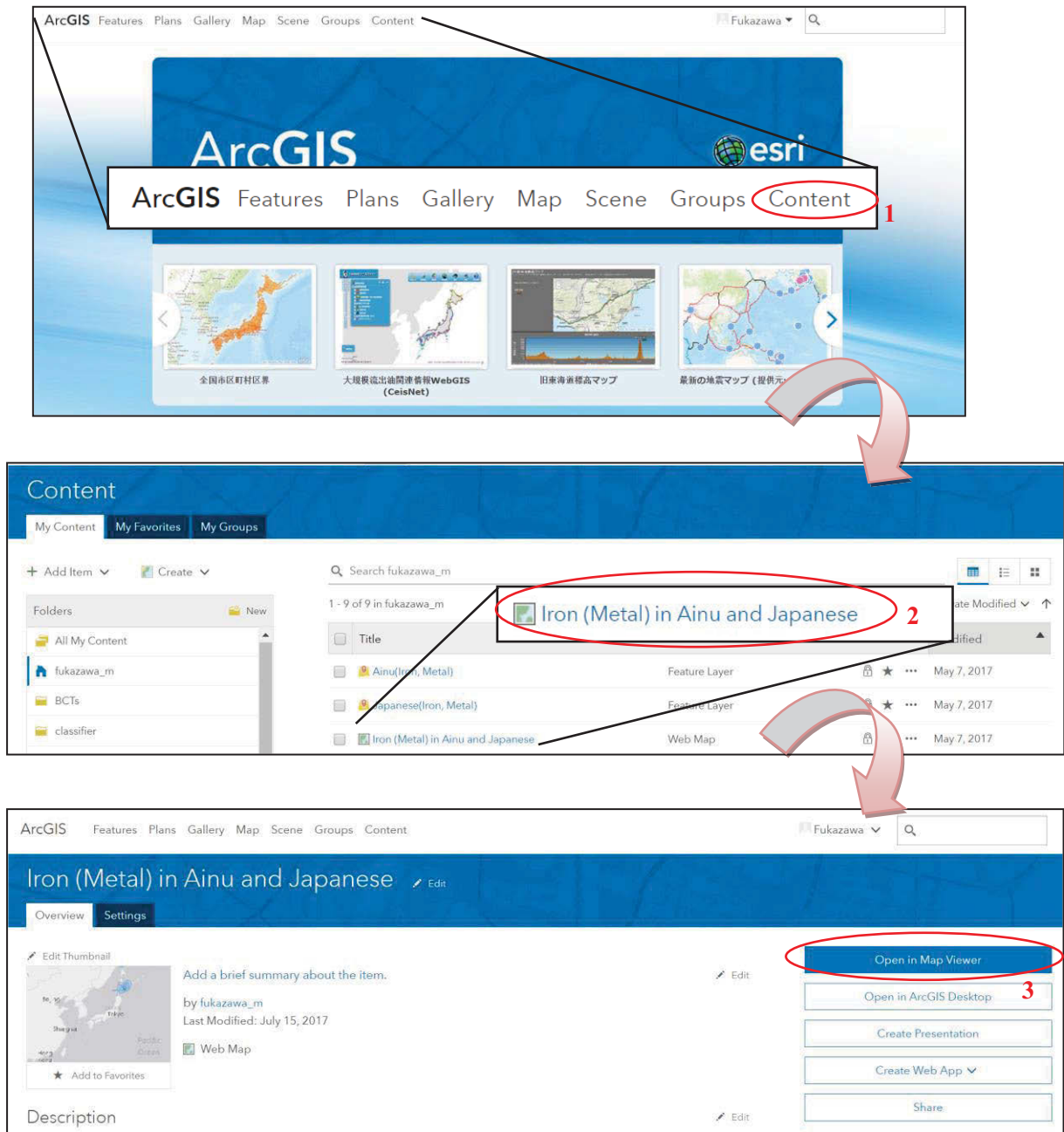
- Finally, click “Save” and “Save” to save the styling changes to the map (see also 3.1). If you copy the original map, click “Save As.” Modify the title, etc., and click “Save Map.”



4 Other Tips

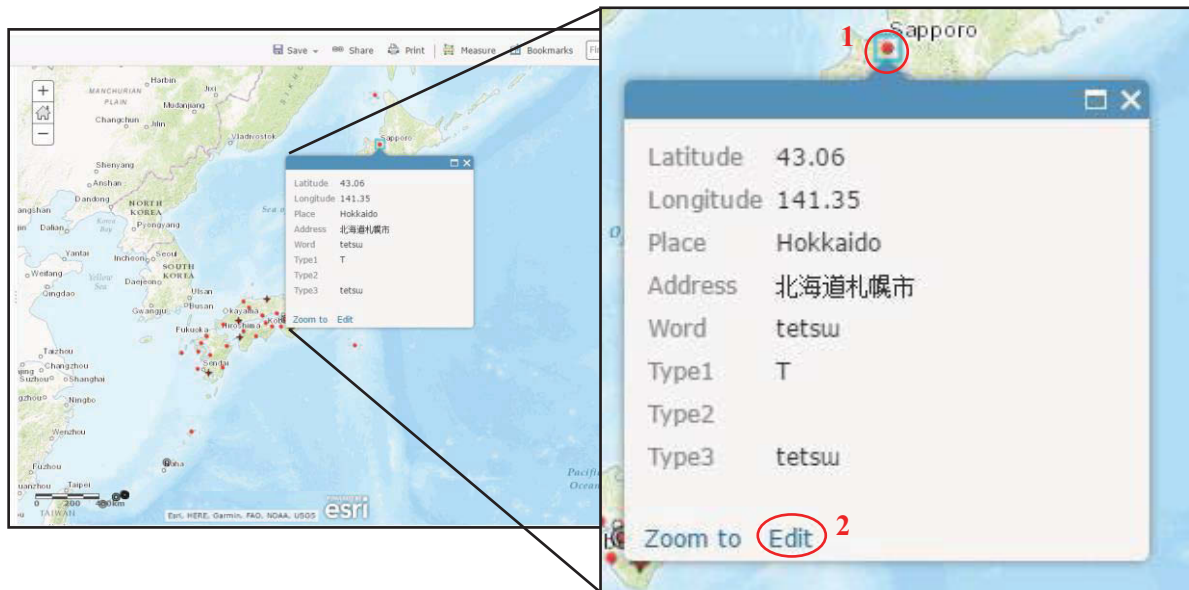
4.1 Using ArcGIS after the first time

1. After the first time you use ArcGIS, sign in and click “Content.” Then, you can go directly to the “Content” page.
2. If you click the name of the map or layer, you can get the information about it.
3. When you want to jump to the map or layer, click “Open in Map Viewer.”



4.2 Edit Features⁷

1. Click the feature on the map that you want to edit. The pop-up window will open. Click “Edit.”



2. The next pop-up window will open. Change the information of the feature you made in the Excel worksheet previously: latitude-longitude, place, address, word and types. Click “Close.” If you want to delete the feature, click “Delete.”

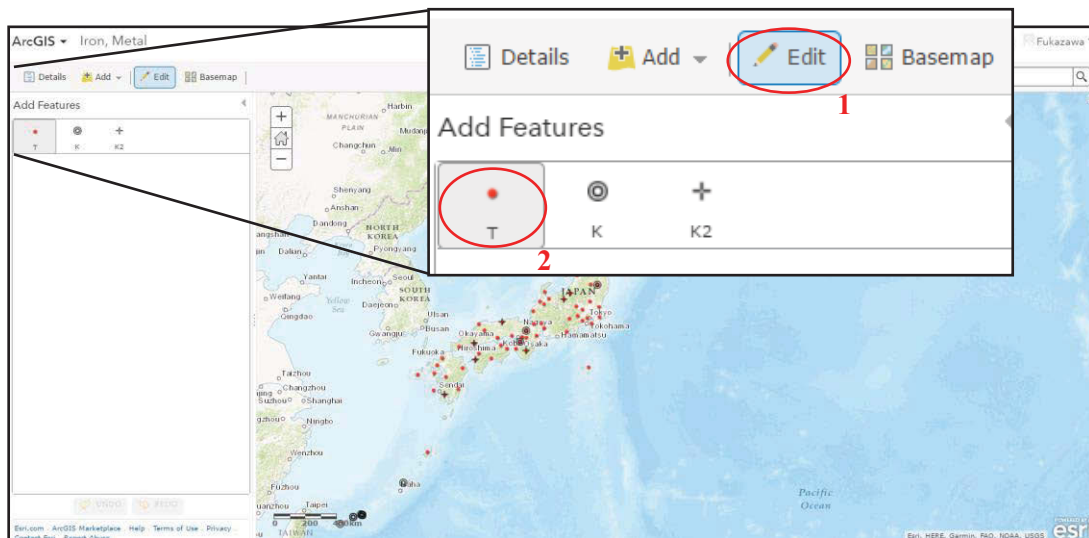


⁷ Sections 4.2 and 4.3 are based on the “Edit Features” page

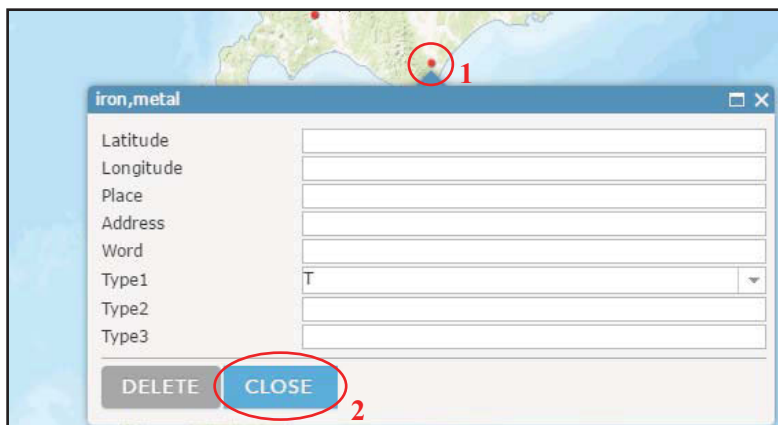
<http://doc.arcgis.com/en/arcgis-online/create-maps/edit-features.htm> (accessed 2017-5-3)

4.3 Add Features⁸

1. Click “Edit” and the symbol of the feature you want to add to the map.



2. Click the location where you want to add the feature to the map. Its pop-up will open. Type the information and click “Close.”



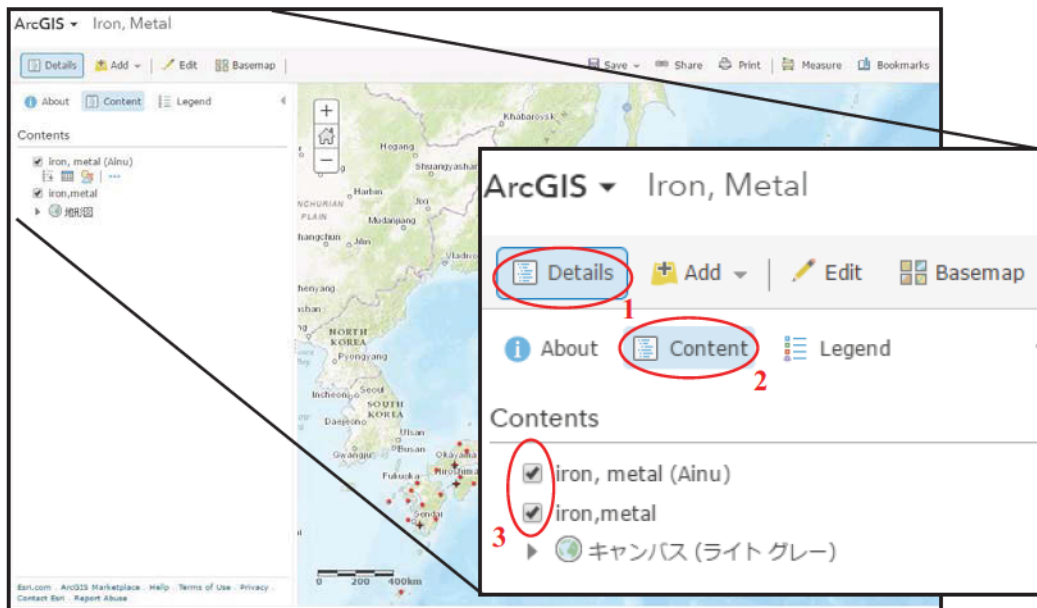
3. Do not forget to save the layer and the map (see Section 3).

4.4 Superpose layers

1. You can superpose several layers on a map. Add some layers, change the styles and save them, as we saw in Section 3.
2. Then, click “Details” and “Content,” and put a check mark on the layers you want to display.
3. Click “Save” to save the map.

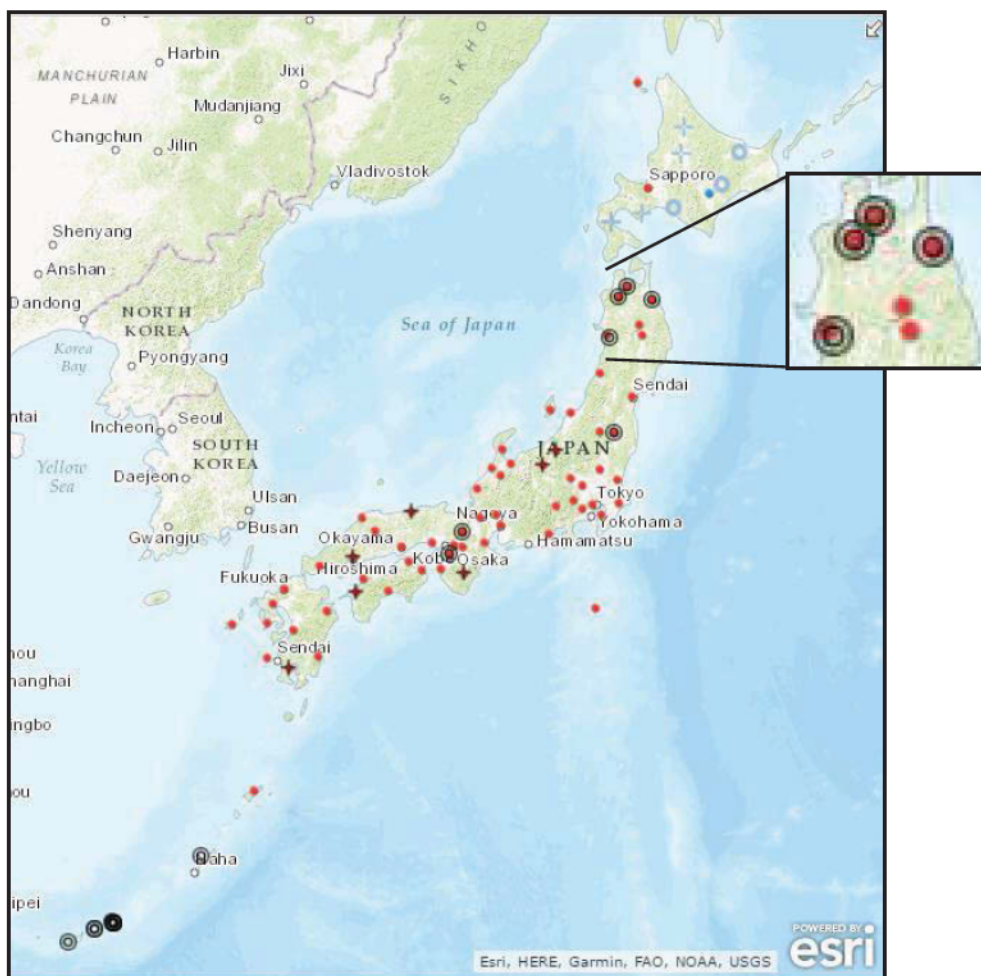
⁸ Section 4.2 is based on the “Edit Features” page

<http://doc.arcgis.com/en/arcgis-online/create-maps/edit-features.htm> (accessed 2017-5-3)

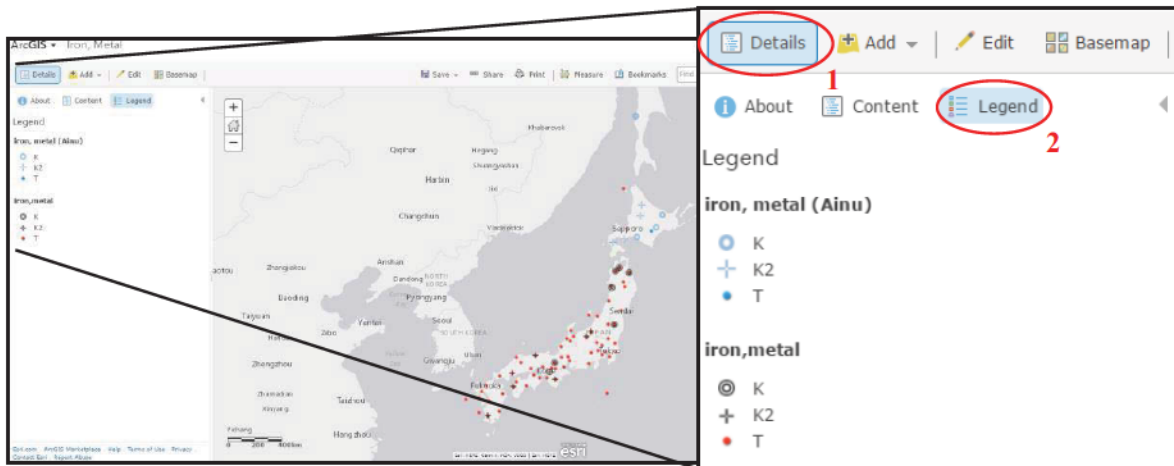


4.5 Tips for Visibility

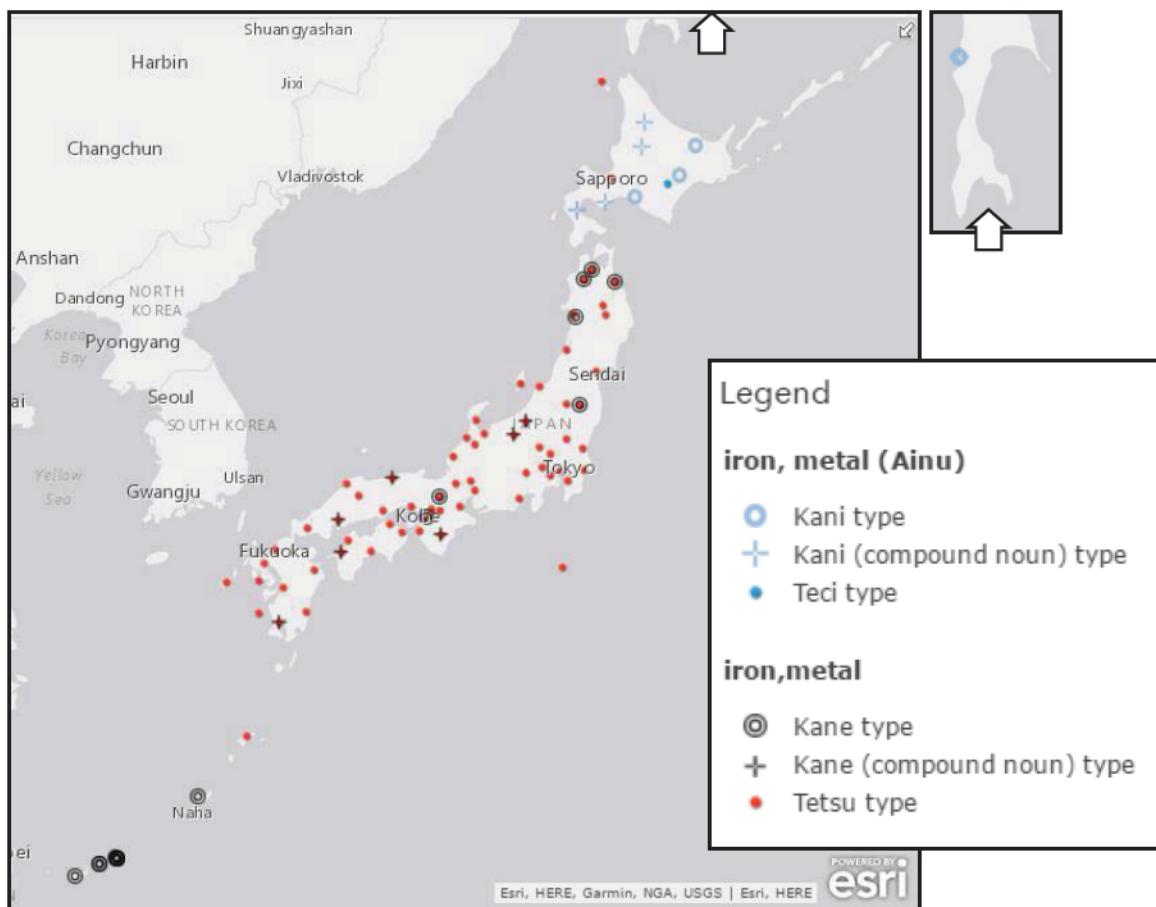
1. If there are two or three forms in one place, improve the symbols (e.g., ◎ and ●, + and ●, or |, ◎ and ●).



2. It is better to mark symbols in different languages by different colors. For example, blue is used for Ainu, but red and black are used for Japanese.
3. Assign the same symbols to the same forms, categories and loan words. Here, the map shows ● for K (*kani* and *kane*) type, + for K2 (*kani* and *kane* (compound noun)) type and ◎ for T (*teci* and *tetsu*) type.
4. If the symbols are not visible, change the basemap (see Section 3.1).
5. The legend shows the correspondence between the symbols and types (categories) you assigned. Click “Details” and “Legend.”



1 Example: The Map of Iron (Metal) in Japanese and Ainu



Map 1: Iron (Metal) in Ainu and Japanese

Now, take a look at the examples of the geolinguistic maps made by ArcGIS. Map 1 shows the words for iron (metal) in Ainu⁹ and Japanese. Fukazawa (2017) suggested the map of the words for “iron” in Ainu, as well as its description. Here, I added the layer of Japanese to the one of Ainu and modified the symbols of the original features of Ainu.

In Ainu, the words for “iron” are classified into the following three types (Hattori ed. 1964):

Kani type:

káni, káne, kani, kaani

Kani (compound noun) type:

yay'an kaani, yayán káni, yayán kani and *yayán káne* for “general metal”;

sínokane for “real metal”;

sirár kani for “rock metal”

Teci type:

tecí

In Japanese, the words for “iron” are categorized into the following three types:

Kane type:

kanę, kane, kani, hani

Kane (compound noun) type:

kanamono, kanamon, kanekenomono for “metallic thing”;

kanake for “metal contained in water”;

ffūkani, ffukani, kuroŋane for “black metal”;

tetsüŋane for “iron metal”

Tetsu type:

tetsu, tędzĩ, tędzũ, tetsu, tętsĩ, teT, tįtsĩ, etc.

The *kani* type in Ainu literally means “metal,” and it was borrowed from the Japanese *kane*, 金 for “metal,” and with it the words for certain kinds of metals, such as gold, silver, and so on, were created. Fukazawa (2017) also mentioned, “Around the 9th or 10th century, when Hokkaido was divided into the Okhotsk culture and the Satsumon culture, people were refining iron, which was taken from a trade with Japanese. Although Sakhalin Ainu had the skill of smithery, Hokkaido Ainu lost it because Japanese iron products were imported, Mamiya (1810) reported.”

In the *kani* type in Ainu, the words for “metal” are applied to the meaning of “iron.” The vowels /e/ and /i/ in the second syllable of *kane* and *kani* would be influenced by the Tohoku dialects in northeastern mainland Japan; the word *kanę* exists in the Aomori, Akita and Fukushima dialects (Hirayama 1993: 3376). Ainu has a five-vowel system /i, e, a, o, u/, which is the same as the most Japanese dialects. However, Okinawa dialects in Japan also use the form *kani* (Hirayama 1993: 3377) because they have a three-vowel system /i, a, u/.

The *tecí* [tetei] type is the loan word from the Japanese *tetu* 鉄 [tetsu]. Ainu does not possess the sound [tsu], and, moreover, the Hokkaido and Tohoku dialects in Japanese have the sound form [tędzĩ] for iron, using a close central unrounded vowel [ĩ]. Japanese *tu* [tsu] was usually borrowed in Ainu *ci* [tei]: the word for “cabbage,” *káypeci* [kajpeei], was borrowed from *kaibetu* [kaibetsu] in the Hokkaido and Akita dialects of Japanese, and the word for “New Year,” *sónkaci* [sonkaci], came from *syoogatu* [jo:gatsu] in Japanese.

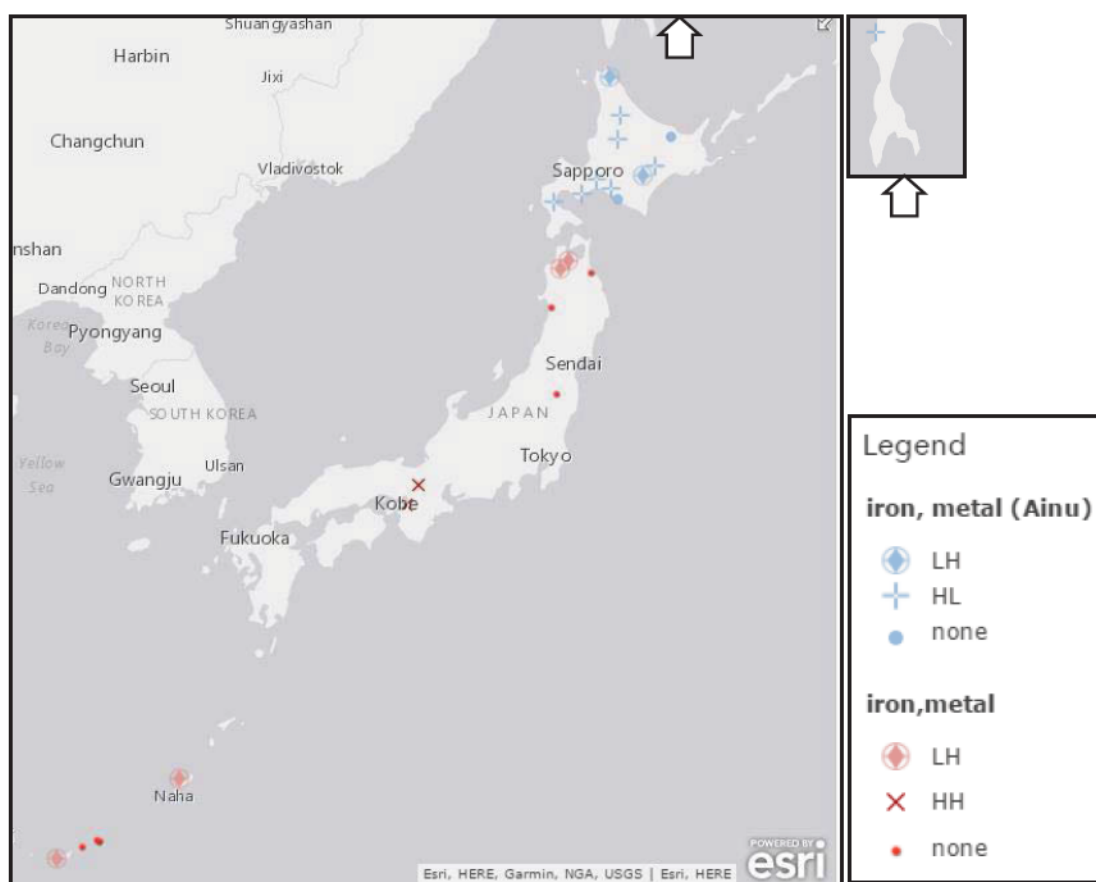
Map 2 below shows differences in accent among the dialects in Ainu and Japanese. In most of the dialects in Ainu, the words for “metal” are accented on the first syllable, i.e., *káni*, *káne*, and *kaani*¹⁰.

⁹ Ainu is an indigenous language in Hokkaido, southern Sakhalin Island and the Kuril Islands.

¹⁰ “The word *kaani* of the Sakhalin dialects would be borrowed from the word *káni* of the Hokkaido dialects,

According to the basic accent in the Hokkaido Ainu dialects, a high pitch falls on the second syllable, CVCV́. Also, in the Aomori and Hirosaki dialects (Japanese) of northeastern Japan, the word for “metal” has a high pitch on the second syllable, [kanɛ] LH¹¹. However, as suggested in Fukazawa (2017), the word accented on the second syllable, i.e., *kaní*, was not influenced by these dialects. Only the basic accent rule was adopted in the Hokkaido Ainu dialects. Therefore, it would be newer than the ones accented on the first syllable. On Map 2, the LH type shows local distribution. Vovin (1993: 73-74, 98) reconstructed the proto-Ainu form **kaani* HHH “metal,” which comes from the Old Japanese word *kane* HH “metal. In the Kyoto dialect of Japanese, the accent-less *kane* HH may be borrowed in Ainu with *káni* HL (Chiri 1956[2004]: 156, Fukazawa 2017). In Ainu, the unaccented form is only distributed in the unaccented dialects, but in Japanese, it is distributed also in the Tohoku and Okinawa dialects, which have an accent system.

Map 2, in terms of the accentual patterns of iron (metal) in Ainu and Japanese, presents more differences among the types of *kani* and *kane* in Ainu and Japanese. All of them are displayed as © on Map 1. Like these maps, we can draw several maps with the same data. In order to make a beautiful map and give historical interpretation to the map, you need to change the criteria by which the word forms (sound forms) are classified.



Map 2: Accentual Patterns of Iron (Metal) in Ainu and Japanese

following the correlation between a long vowel *aa* in Sakhalin and a high pitch *á* in Hokkaido.” (Fukazawa 2017)

¹¹ Okinawa and Hatoma dialects in southwestern Japan also have a LH accent in the words for “iron (metal),” *hani* and *kani*.

Acknowledgements

I would like to offer my special thanks to Prof. Mitsuaki Endo, who showed me his Japanese manual of ArcGIS (Endo 2016). It was a great contribution to my draft of this paper. This work was supported by JSPS KAKENHI Grant Number 16J01055.

References

- Chiri, Mashiho. 1956 [2004]. *The Ainu Language Primer* (アイヌ語入門). Sapporo: Hokkaido Publication Project Center.
- Endo, Mitsuaki. 2016. *How to make Geolinguistical maps using ArcGIS Online*, β ver. 1.0 (ArcGIS Online による方言地図作成法、 β バージョン1.0). private edition.
- Fukazawa, Mika. 2017 [to be submitted]. "Iron" in Ainu, Endo, Mitsuaki (ed.) *Studies in Asian Geolinguistics*, 5. Tokyo: Research Institute for Languages and Cultures of Asia and Africa (ILCAA), Tokyo University of Foreign Studies.
- Hattori, Shirô (ed.). 1964. *Ainu Dialect Dictionary* (アイヌ語方言辞典), Tokyo: Iwanami Shoten.
- Hirayama, Teruo. 1993. *Dictionary of Japanese Dialects* (現代日本語方言大辞典), 4. Tokyo: Meiji Shoin.
- Mamiya, Rinzô and Murakami, Teisuke. 1810. *Hokuibunkaiyowa* (北夷分界余話), 6, In the Safekeeping of National Archives of Japan (National Archives of Japan Digital Archive: URL: <https://www.digital.archives.go.jp/>).
- Nakagawa, Hiroshi. 1995. *The Ainu-Japanese Dictionary: Chitose Dialect* (アイヌ語千歳方言辞典). Tokyo: Sofukan.
- Nakagawa, Hiroshi. (ed.). 2014. *The Japanese-Ainu Dictionary: Mukawa Dialect* (アイヌ語鶴川方言日本語・アイヌ語辞典). Chiba: Chiba University. (Available online at <http://cas-chiba.net/Ainu-archives/>.)
- Okuda, Osami 1999. *Ainu Shizunai Dialect Lexicon in Context (with CD-ROM)* (アイヌ語静内方言文脈つき語彙集 (CD-ROMつき)). Ebetsu: Sapporo Gakuin University.
- Sawai, Harumi. 2006. *Ms. Tomeno Sawai's Basic Vocabulary of the Tokachi Dialect of Ainu* (アイヌ語十勝方言の基礎語彙集:本別町・沢井トメノのアイヌ語). Sapporo: The Hokkaido Ainu Culture Research Center.
- Tamura, Suzuko. 1996. *The Ainu-Japanese Dictionary: Saru Dialect* (アイヌ語沙流方言辞典). Tokyo: Sofukan.
- Vovin, Alexander V.. 1993. *A Reconstruction of Proto-Ainu*. Leiden: E. J. Brill.

[Website]

ArcGIS Online: <https://www.arcgis.com/home/index.html>

Google Maps: <https://maps.google.com/>

A geolinguistic approach to Tibeto-Burman vocabulary:

A case study of ‘sun’ and a future perspective

Satoko Shirai

Japan Society for the Promotion of Science/University of Tsukuba

Abstract

In this paper, I will introduce a geolinguistic investigation into Tibeto-Burman vocabulary, which was conducted as part of the Asian Geolinguistic Project. After surveying the background, I will examine the distribution of words for ‘sun’ in Tibeto-Burman from the geolinguistic perspective. I will also review our past attempt at the project and propose a cross-cutting analysis and perspectives for future research.

1 Introduction

This paper provides a case study of the geolinguistic investigation into Tibeto-Burman (TB¹) vocabulary that was conducted as part of the Asian Geolinguistic Project at the Research Institute for Languages and Cultures of Asia and Africa. During this project, we found that many different roots are used for each item in contemporary TB varieties. Consequently, our research mainly focuses on the distribution and relative time-depth of different roots—not on the phonetic variety of one or a few roots.

Section 2 will survey the background to the geolinguistic approach to TB vocabulary, including our role within the Asian Geolinguistic Project. In Section 3, I will re-examine TB words for ‘sun’ from the geolinguistic perspective with additional data. Section 4 will discuss the results of our works on TB vocabularies, including ‘rice plant,’ ‘milk,’ ‘wind,’ and ‘iron.’ Section 5 will summarize the study.

2 Background to the study

Based on previous research on the TB language family, the importance of the geolinguistic viewpoint has been emphasized repeatedly; for example, Matisoff (1990: 113) proposed the notion of the ‘Sinosphere’ and ‘Indosphere’ to clarify the linguistic/cultural influence in Southeast Asia. He introduced these terms as significant notions in the historical research of TB (Matisoff 1991: 485-486). Dryer (2008, 2017) also employs a geolinguistic viewpoint to investigate TB word order. However, as far as I know, the Asian Geolinguistic Project is the first attempt at creating a comprehensive language map with which to conduct a geolinguistic investigation of TB vocabulary.

The TB project team began with five scholars: Shiho Ebihara, Kazue Iwasa, Keita Kurabe, Hiroyuki Suzuki, and myself, and we were later joined by Ikuko Matsuse. We have been compiling and analyzing data from hundreds of TB varieties. For word forms of ‘sun,’ Shirai et al. (2016) compiled data from 354 varieties²; for ‘rice plant,’ Suzuki et al. (2016) compiled 426 varieties; for ‘milk,’ Ebihara et al. (2016) compiled 533; for ‘wind,’ Iwasa et al. (forthcoming) compiled 560, for ‘iron,’ Kurabe et al. (forthcoming-a) compiled 516, and for ways to count nouns, Kurabe et al. (forthcoming-b) compiled 491. However, varieties from some important language groups and areas such as Tamangish, Kiranti, North Assam, and Northern Naga are missing from four of our lists: ‘sun,’ ‘rice plant,’ ‘milk,’ and ‘wind.’

¹ Abbreviations: OC - Old Chinese; PLB - Proto Lolo-Burmese; PTB - Proto Tibeto-Burman; STEDT - *Sino-Tibetan Etymological Dictionary and Thesaurus* (database); TB - Tibeto-Burman.

² The number includes the multiple forms found in a certain variety of a language/dialect.

3 ‘Sun’ in TB: revisited

In this section, I will re-examine the geographic distribution of TB words for ‘sun.’ Shirai et al. (2016) were the first to attempt this theme; however, they lack data from some areas and groups—especially from Nepal, Bhutan, North Assam, and Northern Naga. This paper is based on enlarged data: 373 varieties of TB, including the abovementioned areas.³

3.1 Classification of word forms

The TB languages in our list show a variety of word forms for ‘sun.’ They consist of a single stem, a stem plus an affix/affixes, two or more compounded stems, or compounded stems plus an affix. First, I classify them into stem types and compound types. Affixes are ignored in the classification, although some forms do contain an affix/affixes. For example, the Xunhua Tibetan *ṣama* and Tshobdun rGyalrong *ta’ṣi* contain a suffix (*-ma*) and prefix (*ta-*) respectively, but both are classified into one of the stem types (the **nəy* type).

3.1.1 Stem types

In this paper, I classify the stem types into eight types: [1] **nəy*, [2] B-, [3] M-, [4] **ḡnam*, [5] GI-, [6] J-, [7] TS-, and [8] **riṣ*.⁴

[1] The **nəy* type is the most widely found. The etymology is PTB **(s-)nəy* ‘sun / day / dwell’ (Matisoff 2003, STEDT #85). This root is recognized as cognate with the Chinese 日 (OC *ṣjět*) (Matisoff 2003: 201, 464), and Coblin (1986) provides the Proto-Sino-Tibetan root **nyi?*. Word forms of this type contain an initial nasal (*n/ṣ/ṣ*). Examples of such forms are *ne*, *ní*, *ṣ³⁵*, *ṣ⁵ naṣ*, *niṣ*, *nyim*, *ṣā*, *ṣama*, *ṣima*, *ṣe⁵⁵ mi⁵³*, *kəṣi*, *ta’ṣi*, etc.

[2] The B- type is the second-most common among the eight stem types. Forms with initial bilabial stops (*p/b*) are tentatively classified as this type. Burling (1967) provides a Proto-Burmish form **pwei^{1a}* ‘sun’ and Matisoff (1978) PLB **mba³* ‘shine.’ However, the etymology of other groups is not clear. A possible etymology is PTB **(p/b/h)(w/y)a(:)(r/l/n/t)* ‘fire / burn / shine / bright / light / white’ (STEDT #6713). In this paper, I do not further examine the detailed etymology. Examples of such forms are *buí*, *péi*, *pu⁵⁵*, *ṣbā*, *ṣbṣi⁵³*, *ṣmbyi*, *ṣə²² bə⁴⁴*, etc.

[3] The M- type tentatively involves words with initial *m*. Their common etymology is unclear, though we can determine the etymologies of two groups: Proto-Loloish **mo²* ‘sun’ (Bradley 1979) and Proto-Karen **mi^B* ‘sun’ (STEDT). A possible common etymology is PTB **r-məw* > PLB **məw²* ‘sky / heavens / clouds’ (STEDT #2473). Examples of such forms are *mi³¹*, *mun*, *məwṣ*, *mūmé*, *a⁵⁵ ṣ²¹ yu⁵⁵*, etc.

[4] The **ḡnam* type of stems are characterized by both initial *n-* and final *-m*.⁶ The etymology is PTB **ḡ-nam* ‘sun / sky’ (STEDT #2484). Ogenort (2011: 268) provides the Proto-Kiranti form **ṣnam* ‘sun.’ Examples of such forms are: *nam³¹*, *nem³¹*, *nəm⁵³*, *num⁵³*, *ḡnam*, *nam⁵³ nəm⁵³*, etc.

[5] The GI- type of stems has initial *g/ṣ/dz/ṣ* followed by a [+front] vowel. Shirai et al. (2016) concluded that the older form of this type had initial *g-*, because of the Written Northern Yi form: *ḡu³³*. The etymology is unclear, but one related proto-form may be PTB **ḡ(l)(i/u)m* > PLB *ḡim¹* ‘set (of the

³ The data of Newar varieties were added by Ikuko Matsuse. The Meche item was provided by Kazuyuki Kiryu and inputted by I. Matsuse. The other addition was made by me.

⁴ They corresponds to the classification in Shirai et al. (2016) as follows: [1] A-1, [2] B, [3] C, [4] A-2, [5] E, [6] a new classification (ramified from D) and [7] D. [8] was not mentioned in Shirai et al. (2016).

⁵ The numbers in italics that appear as superscripts to examples indicate tones, as do the alphanumerical superscripts such as those next to the word *pwei^{1a}*.

⁶ Some varieties of the rGyalrongic group have a disyllabic form *kV n/ṣ/ṣ*- (e.g., Situ rGyalrong *kəṣi*) or initial cluster *ṣn-* (e.g. Khroskyabs/Lavrung *ṣnə⁵⁵*). It is difficult to ascertain whether they are formed of a **nəy*-type stem with a prefix or a **ḡnam*-type stem that has become disyllabic. In this paper, I follow the analysis of STEDT that supports the former possibility.

sun)’ (STEDT #6212, Matisoff 2003). It may have undergone the semantic change through compounding (see Section 3.2). Examples of such forms are *gi*³³, *gu*³³, *dzi*²¹, *tei*²², and *dzi*³³ *mo*²¹.

[6] The J- type stems have the initial approximant *j-*. The etymology is unclear. Jacques (2008: 45-46) mentions that the original meaning of the root *jám* in Situ rGyalrong is “to be sunny.” In this paper, I separated the J- type, which is classified together with the TS- type in Shirai et al. (2016), because it is highly doubtful whether the two types can be traced back to the same etymology. Examples of such forms are *kəjam*, *kə’jam*, *kə²²jaʔm⁵⁵*, and *kəjam*.

[7] TS- type stems have an initial fricative or affricate (s/ts/tsh/dz/ʈ) followed by a [-front] vowel. The most possible etymology is PTB **tsyar* ‘sunshine’ (Matisoff 2013). It is well-known that Burling (1983: 2, 20) proposes the Sal group that shares the common root **sal* for ‘sun,’ and points out that **sal* is derived from PTB **tsyar* ‘sunshine.’ Examples of such forms are *san*, *jan*, *a³³ tshu⁵⁵*, and *A⁵⁵ tshu²¹*.

[8] The **riŋ* type is traced back to PTB **r(i/u)ŋ* ‘sun / sunshine’ (STEDT #6225). The distribution is so limited that our list has only one example of the **riŋ* stem type from Darang: *ɹun*⁵³.

Besides the stems above, Matisoff (2003: 281) provides another PTB root for ‘sun’: **(m-)ka(-n)* ‘heavens / sky / sun’ (STEDT #697). However, according to STEDT, the words derived from **(m-)ka(-n)* rarely mean ‘sun,’ but do mean ‘heaven’ or ‘sky’ in most modern varieties. STEDT provides a compound in Magar *nam-khan* (**gnam* + **(m-)ka(-n)*) that means ‘sun.’ This suggests that its primary meaning is ‘heavens / sky,’ rather than ‘sun.’

3.1.2 Compound types

There are a variety of compounds that consist of the stems discussed in Section 3.1.1. Moreover, there are several stems found only in compounds: [9] L-, [10] HW-, and others (‘god/deity,’ ‘eye,’ etc.)⁷.

[9] The L- type stem is found in several types of compound, but its etymology is unclear. Possible etymologies are PTB **s-la(m/p)* ‘dry (by fire / sun)’ (STEDT #3515) and PTB **s-li* > PLB **ʔ-li^l* ‘old (of things)’ (STEDT #7183). Examples of such forms are *nam³¹ loŋ⁵⁵*, *lo²¹ gu³³*, *lumū*, *li⁵⁵ tei¹¹*, etc.

[10] The HW- type stem includes *ho* / *xo* in some Loloish varieties and stems with initial *w* / *v*. The etymology is unclear, but a possible etymon is PTB **hwaŋ* ‘shine / bright / yellow’ (STEDT #5518). It is considered a related form (allofam) of PTB **(p/b/h)(w/y)a(:)(r/l/n/t)* ‘fire / burn / shine / bright / light / white’ (STEDT #6713)—that is, the same root as a possible etymology of the B- type. Examples of such forms are *ho³³ bu³³*, *xo²¹ dzi²¹ mo²¹*, *nji⁵⁵ vɔ³³*, *nju⁵⁵ we³³*, etc.

Morphemes that mean ‘god/deity’ are found in the second syllable of some Tibetic varieties (e.g., *n̥n̥ʂla*, which literally means ‘sky-deity’) and the third syllable of some Newar varieties (e.g., *nibhaa: dya:*, which literally means ‘sunlight-deity’).

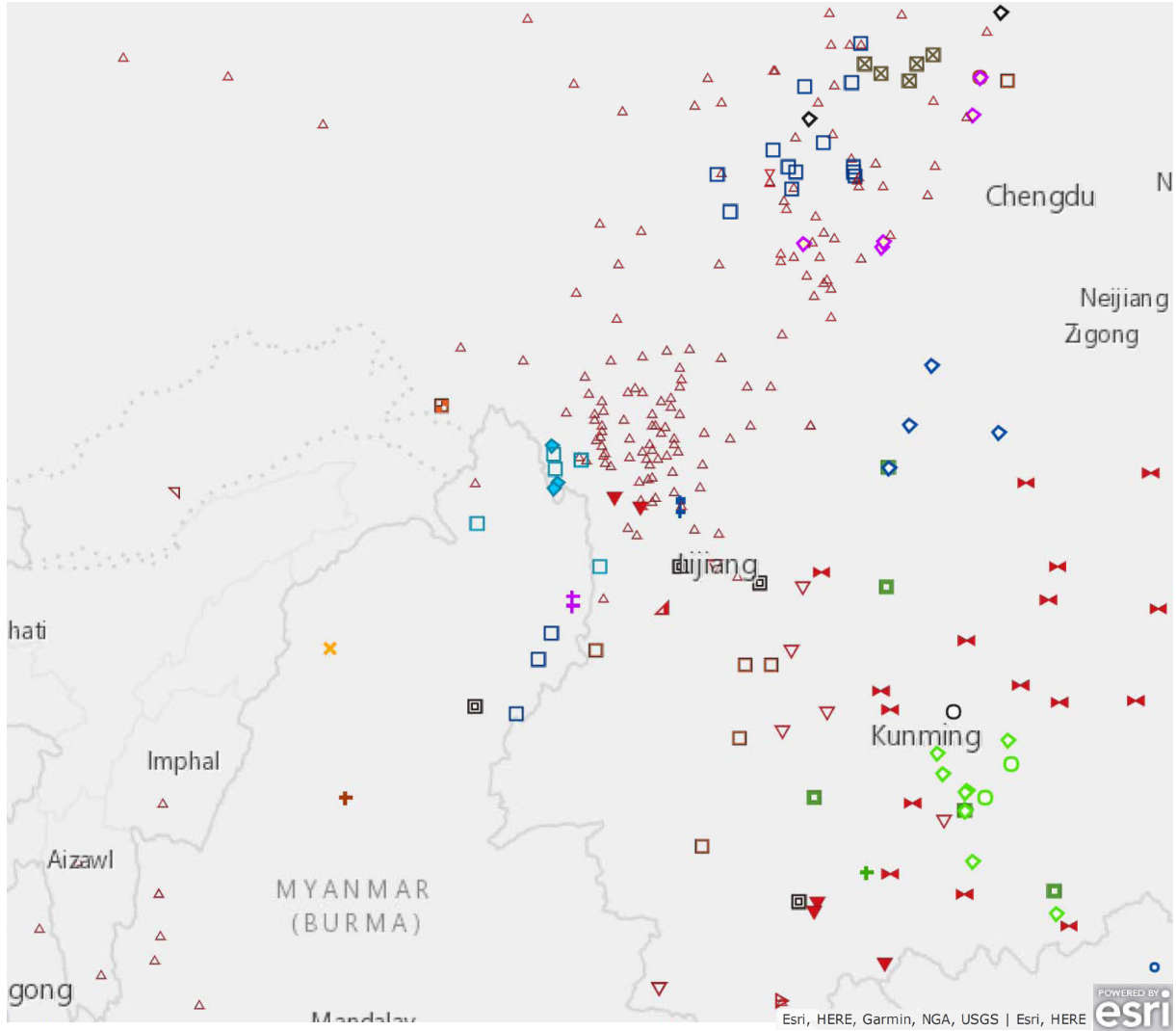
Morphemes that mean ‘eye’ are found in the second syllable of some Luish (Asakian) varieties (e.g., *cəmiʔ*, which literally means ‘sun(light)-eye’).

3.2 Geographical distribution and interpretation

In this section, I re-examine the geographical distribution of the types of the word ‘sun’ in TB and revise some of the conclusions that we made in Shirai et al. (2016), such as the time-depth of the **TS-* type, J- type (both of which are classified into the D type in Shirai et al. 2016) and the **gnam* type.

Map 1 shows the geographical distribution of the types of the word ‘sun’ throughout the TB area based on the data from 373 varieties. All data sources apart from the field data provided by the members of the TB team are listed at the end of this paper. Since the spots are heavily concentrated in the central area, an enlarged map of the area is provided as Map 2.

⁷ In this paper, I made the following changes from Shirai et al. (2016): (i) excluded the second syllable with initial *m-* in Luish (Asakian) varieties from [3], which means ‘eye’ (Matisoff 2013), and (ii) excluded the second syllable with initial voiceless lateral in Tibetic from [9], which means ‘god/deity’ (Suzuki 2016).



Map 2: ‘Sun’ in Tibeto-Burman: The central area enlarged

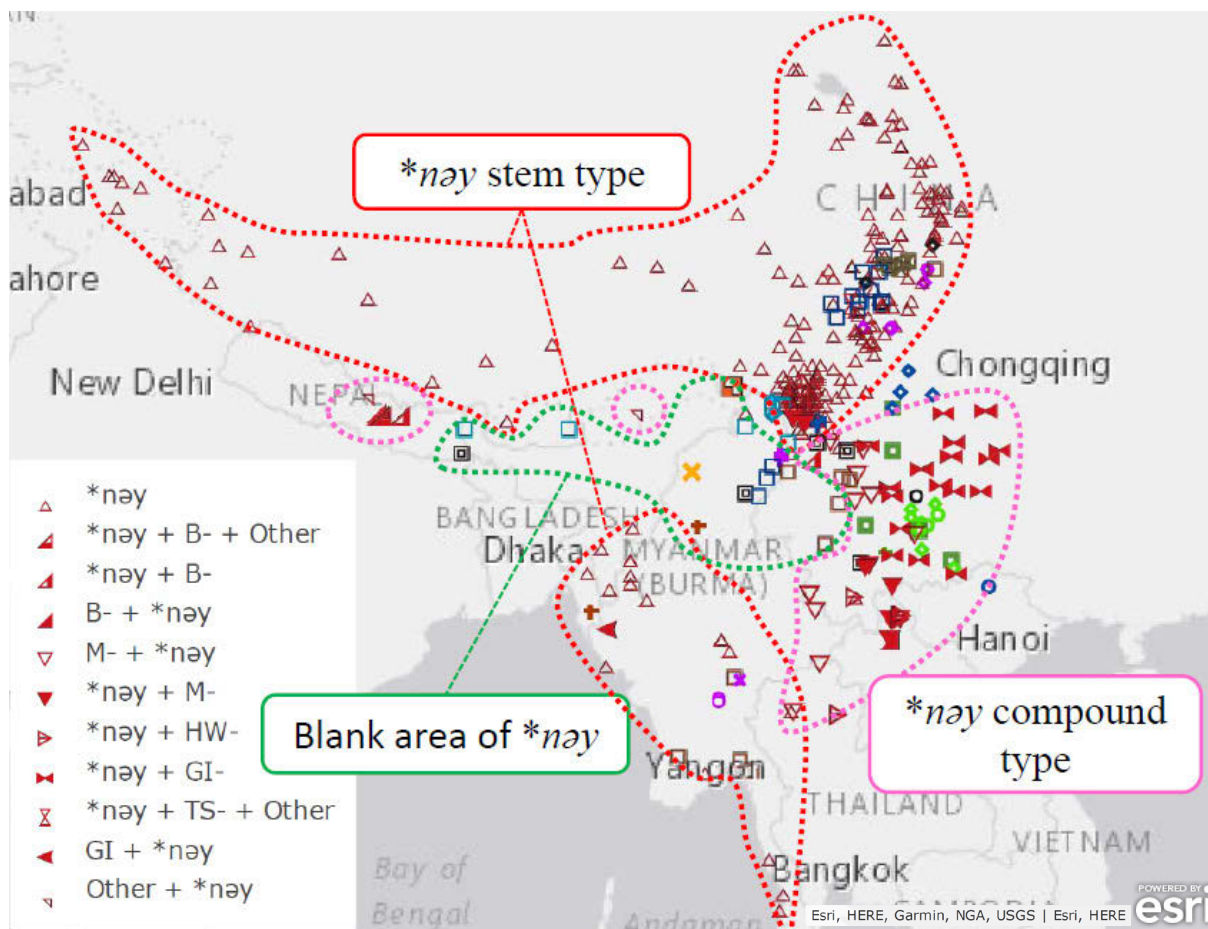
3.2.1 *nəy

See Map 3. The **nəy* type is the most widespread beyond the branches of TB. The **nəy*-stem type is found in the Tibetic, rGyalrongic, and Qiangic groups that are located in the northern and western parts of the TB area, and in Burmic and Loloish in the southern part. Moreover, compound types with **nəy*-type morpheme are also widespread. The main types of **nəy*-compounds are shown in Table 1.

Table 1: Compound types with a **nəy*-type morpheme

Compound Type	Language Group	Area in TB	Example
* <i>nəy</i> + GI-	Yi varieties	East	<i>ne</i> ²¹ <i>ge</i> ²¹ (Naso Yi)
* <i>nəy</i> + M-	Loloish varieties	Southeast	<i>nu</i> ⁵⁵ <i>ma</i> ³³ (Hani)
M- + * <i>nəy</i>	Loloish varieties	Center and South	<i>mùŋ hnunŋ</i> (Bisu)
* <i>nəy</i> + HW-	Loloish varieties	Southeast	<i>ni</i> ⁶ <i>wo</i> ⁴ (Mpi)
* <i>nəy</i> + B-	Newar varieties	West	<i>nibaar</i> (Dolakha Newar)

The **nəy*-stem type and **nəy*-compound type cover the whole TB area. However, interestingly, they are not found in the south-central area. Consequently, we can conclude that the **nəy* type is the oldest in terms of time-depth (Shirai et al. 2016).



Map 3: Distribution of **nəy*

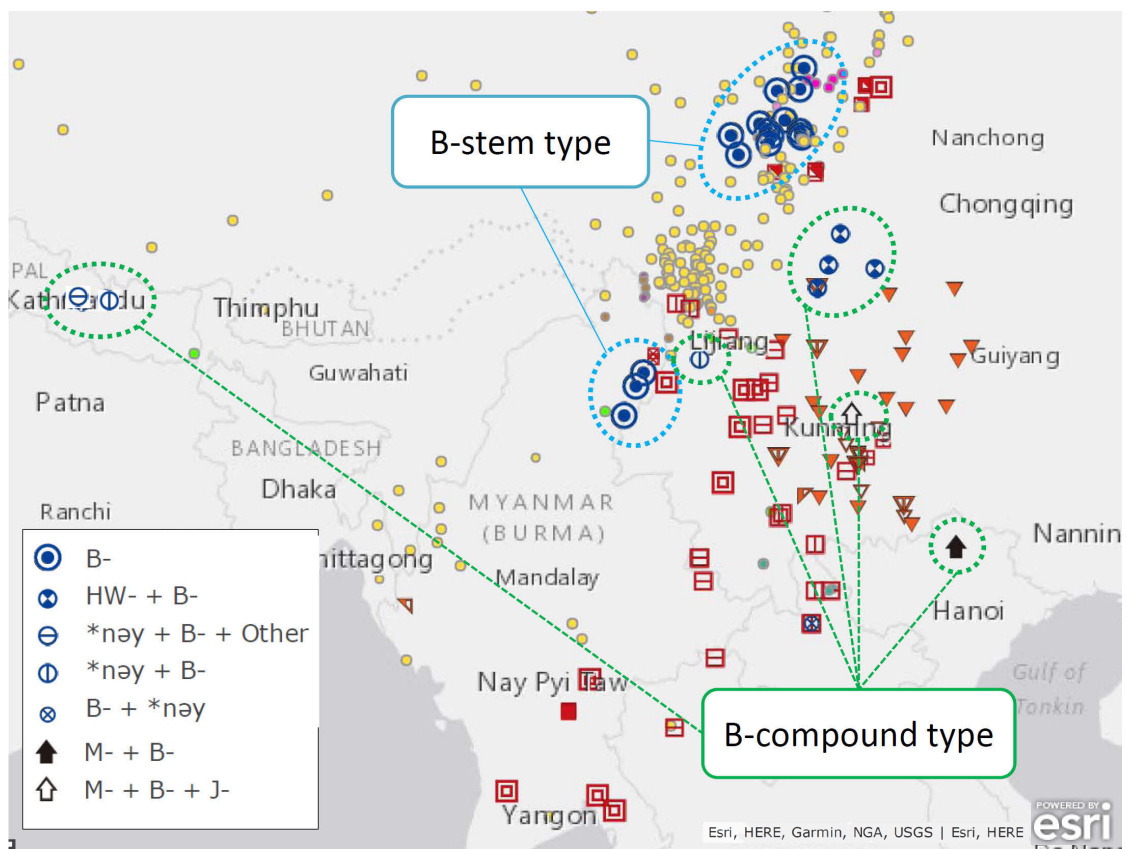
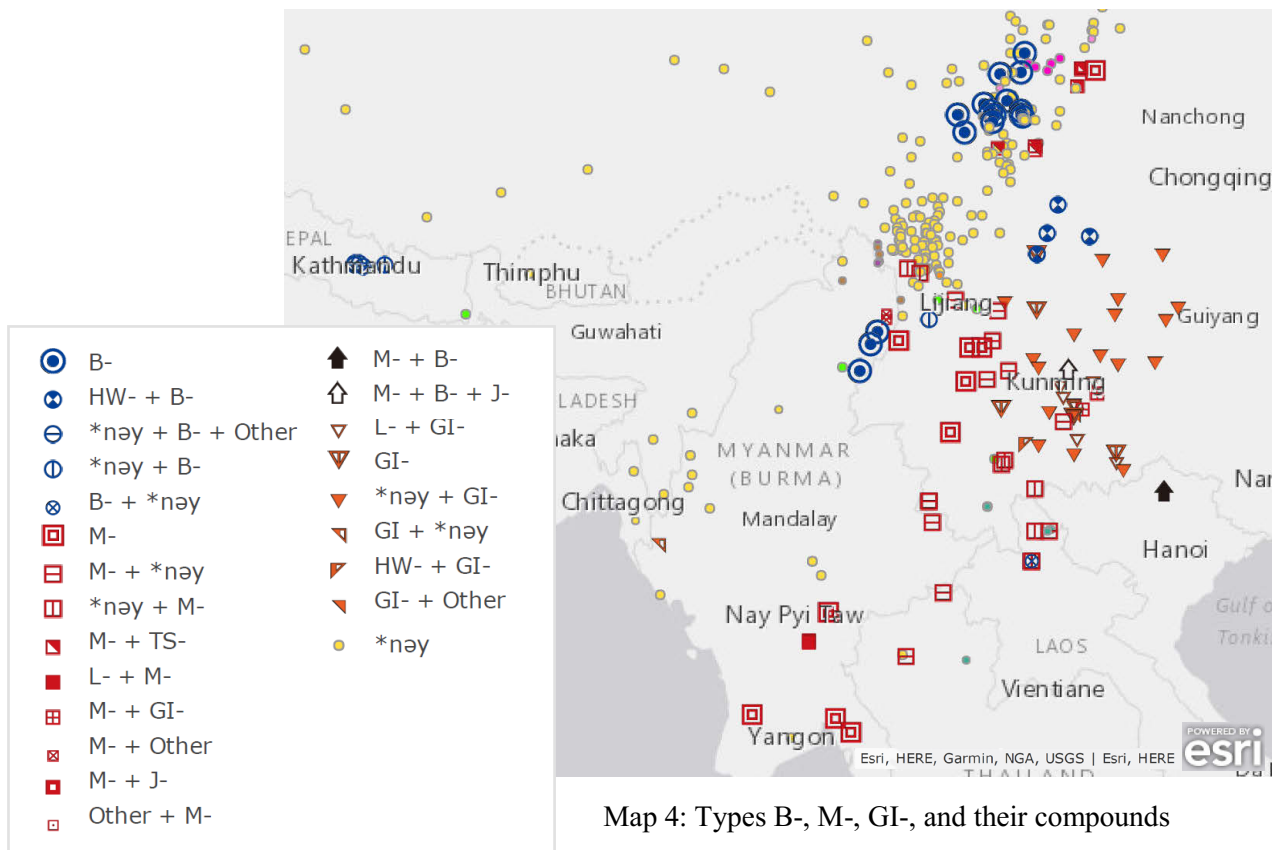
3.2.2 B-, M-, and GI-: Relative time-depth of stem types and compounds

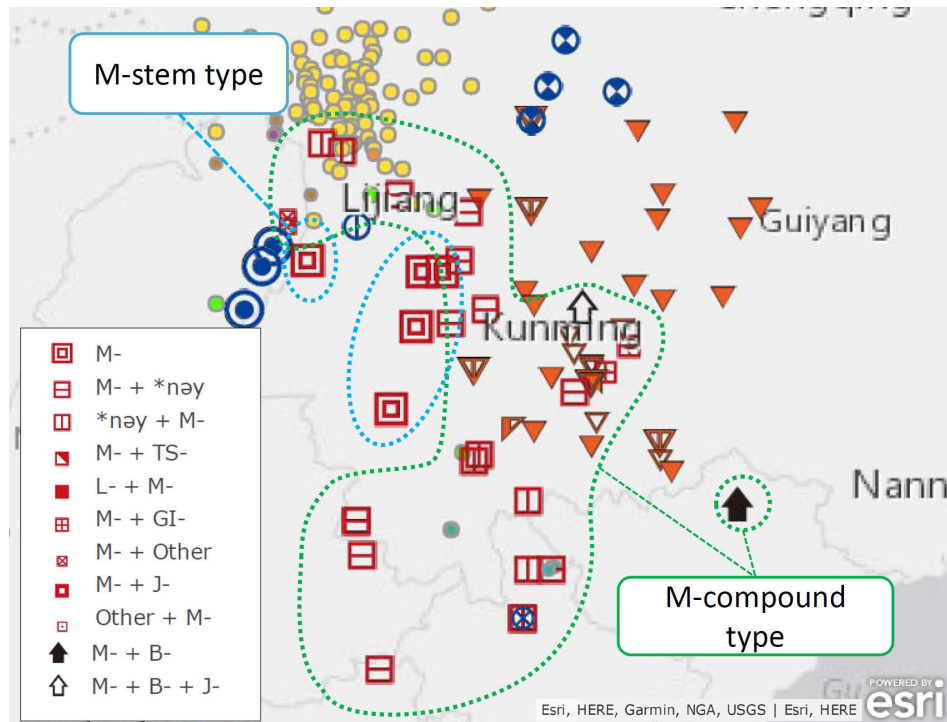
Map 4 shows the distribution of the types B-, M-, GI-, and their compounds. The other types including the **nəy* stem type are indicated with small dots.

Newar dialects (of eastern Nepal) that are newly added to our present list have compounds with a B-type stem: **nəy* + B- (+ ‘deity’). For example, Kathmandu Newar *nibhaa*:-*dya*: and Dolakha Newar *nibaar*. Geographically, they are quite remote from other B- types, as indicated in Map 5. Moreover, compound types involving B- are relatively rare (Shirai et al. 2016). I now examine the distribution of **nəy* + X (stems other than **nəy*) and the X-stem type.

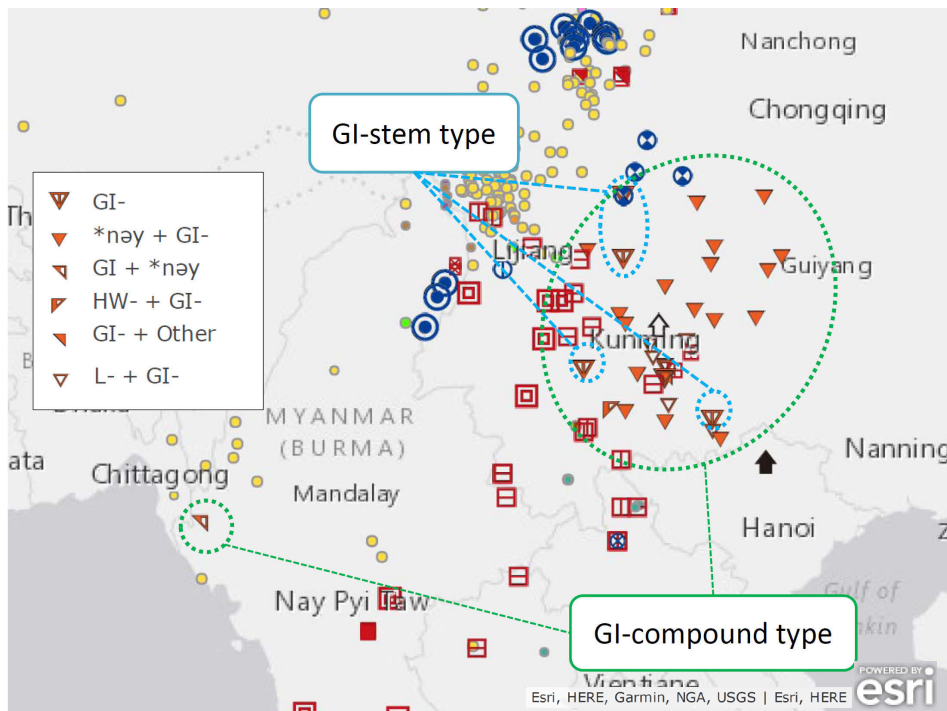
See Map 6. When we focus on the distribution of the M- + **nəy* compounds and the M-stem type within the Loloish varieties (from southwestern China to the China-Myanmar-Laos-Vietnam border area), they show a significant distribution: the compound types are found around the spots of the M-stem type⁸. Thus, we can propose a hypothesis that the M-stem type is newer than M- + **nəy* type compounds, at least in the Loloish group.

⁸ In geolinguistics, such distribution is called ‘ABA distribution.’ This means that spots with character A are located on both sides (or around) of spots with character B. It may mean that character B rose after character A spread through the area, if character A is not likely to rise independently in different areas. Thus, ABA distribution goes to show that A is older than B.





Map 6: Type M- and its compounds in southwestern China



Map 7: Type GI- and its compounds

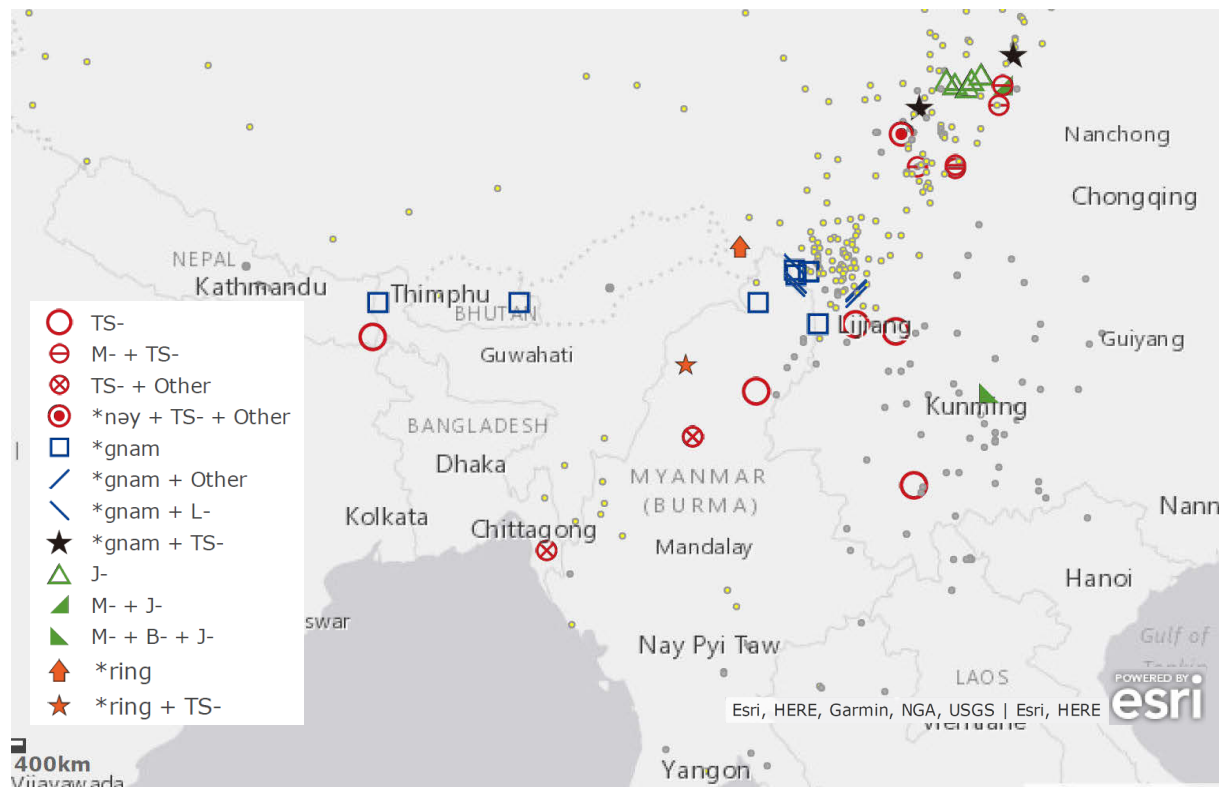
See map 7. As we have pointed out in Shirai et al. (2016), the GI-stem type and *nəy + GI-compound type show a similar distribution: the compound type is more widespread and the stem type is found sporadically among them. Consequently, we hypothesized that GI- was first used as part of a compound (especially the *nəy + GI- type compound), and later became an independent word meaning ‘sun.’

Compounding may trigger a semantic shift and may generate a new stem that means the same as the compound. We have now found a parallel case in the M- type and GI- type. Thus, the **nəy*-compound types may be older than other stems. Moreover, **nəy* itself means ‘sun.’ It may also have derived compounds that mean ‘sun’ individually in several TB varieties. Thus, we tentatively conclude that the **nəy* + B- (+ ‘deity’) type compound in Newar is formed individually from the B-stem type in the central area.⁹

3.2.3 TS- and J-

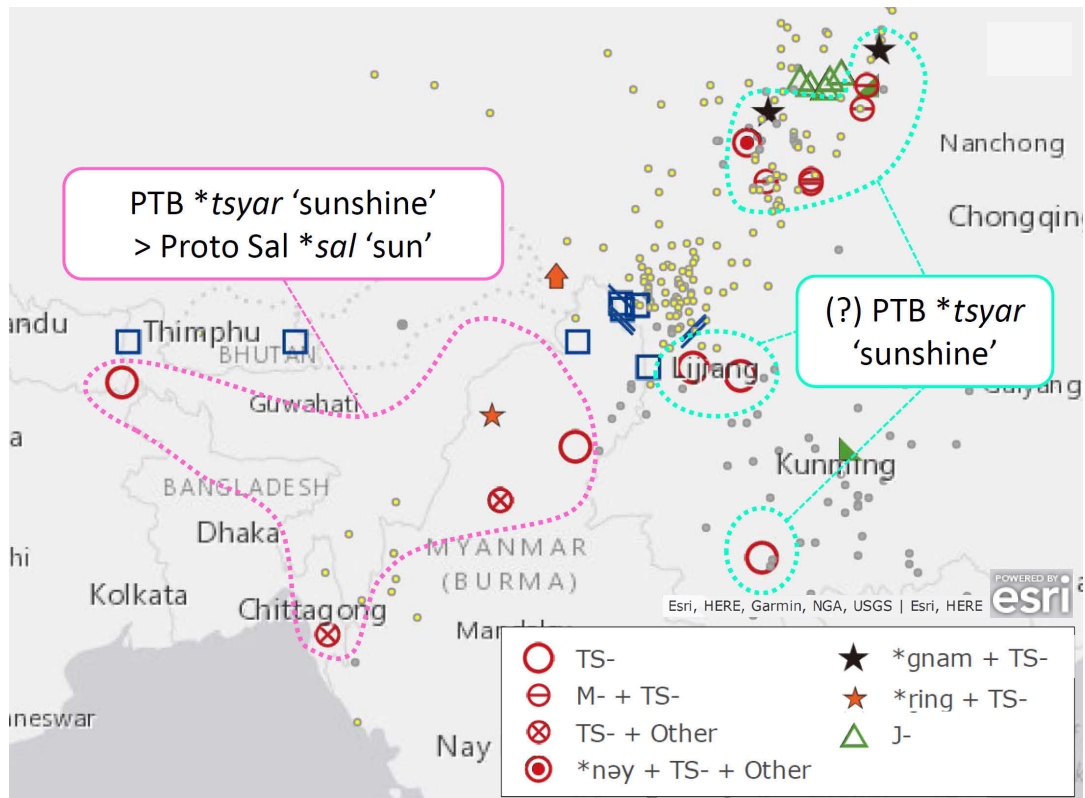
See maps 8 and 9. Although the number of varieties that have the TS- type for ‘sun’ is rather small, their location is considerably widespread: Varieties of Bodo-Garo, Jingpho-Luish, and Northern Naga have the cognate form of **tsyar* (STEDT). Moreover, some varieties of the Loloish and Qiangic group have a similar form, although the etymology is unclear. Besides, we have found scarcely any examples of compounds of **nəy* and TS-.¹⁰ The TS- compounds typically consist of a stem with a different meaning from ‘sun,’ e.g., gSerpo Tibetan *na^xtsa* (**gnam* ‘sky’ + TS-) and Cak *cəmi?* (TS- + ‘eye’). This suggests that the TS- type has appeared as a word for ‘sun’ independently of the **nəy* type. Consequently, we can conclude that the TS-stem type is not older than its compounds. Therefore, it might be possible to make a tentative conclusion that the TS- type is not newer than the types B-, M-, and GI-, although the issue of their relative time-depth awaits further investigation.

The J- type is found only in rGyalrongic varieties of the Tangut-Qiang group (northwestern Sichuan). I can conclude that the J- type is a relatively new form for ‘sun.’

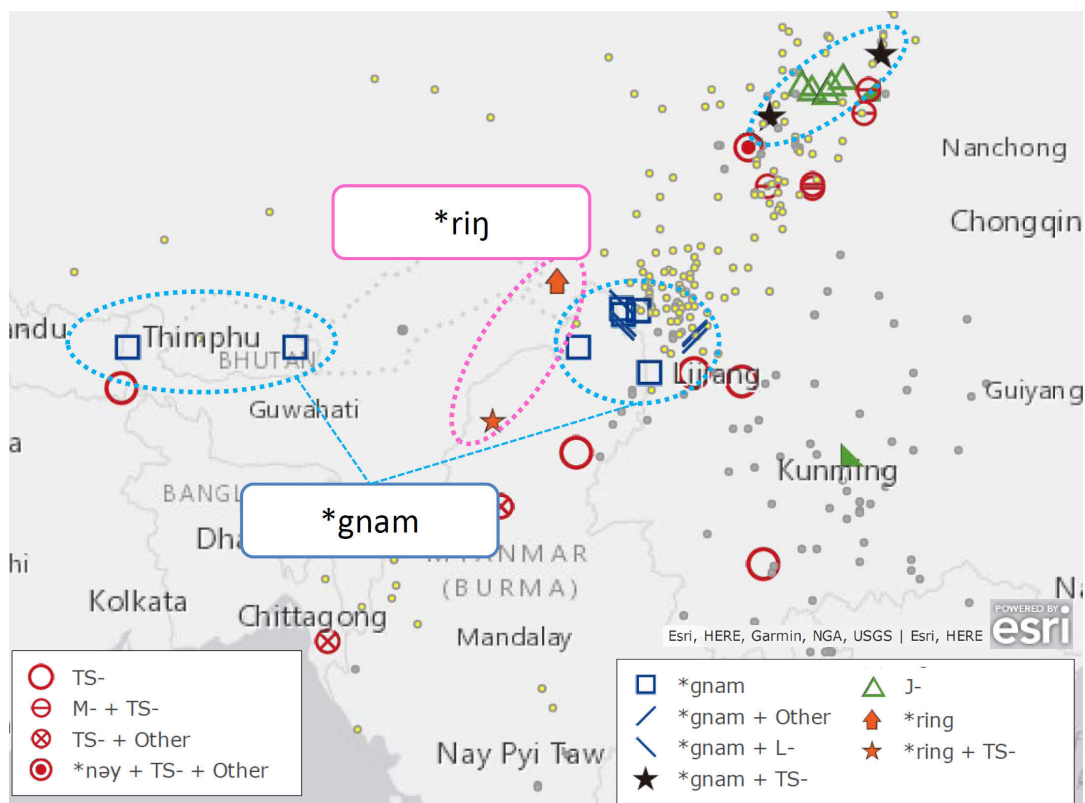


⁹ If the B-type morphemes in Newar (west of the TB area), Loloish (center) and Qiangic (north) can be traced back to one etymology, the B- type is considerably old. However, this is not yet certain. The time-depth of type B- is a subject for future investigation.

¹⁰ The only exception is sTau *nətsə'kətsosi*.



Map 9: The TS- type



Map 10: The **gnam* type and **ring* type

3.2.4 *gnam and *riŋ

See Maps 8 and 10. The **gnam* type is found along the China-Myanmar border area (spots of Nungic and Tibetic varieties), eastern Nepal (Kiranti varieties) and eastern Bhutan (Tsangla varieties).¹¹ The **gnam* compounds include Yangthang Tibetan ⁶*n̄̃ ʃa* (**gnam* ‘sky’¹² + ‘deity’), gSerpo Tibetan *na^x tsa* (**gnam* ‘sky’ + TS-), and Trung *nam³¹ loŋ⁵⁵* (**gnam* + L-). It is difficult to determine whether the stem type or compound type was the first to mean ‘sun.’

The **gnam* type of ‘sun’ in the China-Myanmar border area is a good example of an areal feature: the same type is shared by both the Nungic and Tibetic varieties. Moreover, the other Tibetic varieties do not have this type for ‘sun.’

The spots of the **gnam* type in the China-Myanmar border area are apparently separated from the other spots in Bhutan and Nepal. However, we can find the **riŋ* type (as well as some **nəy*-compound types) between them. The **riŋ* type shows the most limited distribution among the eight stem-types. The **riŋ*-stem type is found in the North Assam group, and a **riŋ*-compound type that consists of **riŋ* + TS- (Tangsa *rong shar*) is found in the Northern Naga group. Both are located in the China-Myanmar border area. The fact that the **gnam* type is located separately on both sides of the **riŋ* type suggests that the **riŋ* type is newer than the **gnam* type.

4 A cross-study of the items: review of our past studies

In this section, I provide a brief review of our project on TB vocabulary conducted so far and re-interpret the data from a cross-cutting perspective.

4.1 Types with wide distribution

Table 2 shows the most widespread types both geographically and genetically; that is, the types that are expected to be the ‘oldest’ from the geolinguistic point of view. Such forms are represented as the form of the shared etyma in the ‘PTB’ column.

Item	PTB	Area
‘sun’	* <i>nəy</i> ‘sun / day / dwell’	Everywhere except for south-central
‘rice plant’	(* <i>b-ras</i> ‘rice / fruit / bear fruit / round object’)	North/West
‘milk’	* <i>s-nəw</i> ‘breast / milk / suck’	North/South/East
‘wind’	* <i>g-ləy</i> ‘wind’	North/South/East
‘iron’	* <i>syam</i> ‘iron’	Center/North/South/East
	or * <i>l-tsyak</i> ‘iron’	North/South/West

They are not chosen simply according to their geographic distribution because the Tibetic varieties are extremely widespread in the TB area and retain considerable coherence in language features. Consequently, though a feature that is found only in Tibetic (and around it as a Tibetan loan) is considerably widespread in terms of geography, it may be newer than another feature that is found in a number of other genetic groups. For example, ‘milk’ in TB involves the type derived from the Written Tibetan *’oma*, which is most widely distributed in terms of geography. However, Ebihara et al. (2016) concluded that the type derived from PTB **s-nəw* in Table 2 is older than the *’oma* type, because the **s-nəw* type appears in at least six genetic groups and is geographically the second-most widespread. In addition, the same etymon is shared by varieties of Tamangish, Kiranti, and Tsangla (STEDT), which our map in Ebihara et al. (2016) lacks. This supports Ebihara et al.’s (2016) conclusion that the **s-nəw* type is the oldest among the types of word for ‘milk’ in TB.

¹¹ Shirai et al. (2016) lacks data on Kiranti and Tsangla and concludes that the **gnam* type is geographically limited and historically new. However, we now know that this is incorrect.

¹² In fact, in many of the modern Tibetan dialects, the morpheme derived from **gnam* means ‘sky.’ Note that **gnam* is attested as Written Tibetan *gnam* ‘sky’ and many of the modern varieties reflect this.

The PTB root **nəy* for ‘sun’ shows the widest distribution both geographically and genetically (3.2.1).

In contrast, the chronological order of the words for ‘rice plant’ is difficult to ascertain; thus, one of the etyma is shown within parentheses in Table 2. **b-ras* is geographically the most widespread, but it is shared only by Tibetic varieties and several other languages that have probably borrowed it from Tibetan (Suzuki et al. 2016). Other roots are also distributed among limited genetic groups; thus, it seems that almost every cultural group has a unique root for ‘rice plant.’¹³

Our first attempts at studying the distribution of the words for ‘wind’ (Iwasa et al., forthcoming) and ‘iron’ (Kurabe et al., forthcoming) do not provide conclusive information on the chronological order. Here, I choose PTB **g-ləy* for ‘wind,’ and propose two candidates for ‘iron’: PTB **syam* and PTB **l-tsyak*.

The words derived from PTB **g-ləy* ‘wind’ show striking diversity in their forms. We found all patterns of initial lateral, velar, and both (disyllabic forms) (Iwasa et al. forthcoming).

The chronological order of the two roots for ‘iron’ is difficult to ascertain. In the eastern area, the **syam*-type is predominant, but in the northern and western areas, the **l-tsyak*-type is predominant with Tibetan residents. In the southern area, both are found. Interestingly, sTau dialects (rGyalrongic), which are located between the areas of the two roots, have mixed forms of the two types (Kurabe et al. forthcoming).

4.2 Areas with unique types

Contrary to the ‘old’ types, we have also found some types with limited distribution for each item. When a number of such types are piled up, they show an interesting distribution. I exclude the obvious loanwords from other language families (e.g., Chinese) from the discussion here.

Examining the five geolinguistic maps of TB vocabulary, some areas appear to be significant with unique types that are not found in other areas. See Map 11.

Northwestern Sichuan of China is one such area. For example, rGyalrongic varieties in this area share a unique root for ‘sun’ that is characterized by the initial palatal approximant (Section 3.2.3). Moreover, many of the Tangut-Qiang branch varieties, which involve the rGyalrongic and Qiangic groups, have a word for ‘milk’ that is derived from Proto-Tangut-Qiangic **s-lu* ‘milk’ (STEDT #1786) (Ebihara et al. 2016). Moreover, the **s-lu* type of words for ‘milk’ is attested only in this branch. STEDT provides a PTB root **s-lwayŋ* for ‘breast’ (STEDT #3458) but clearly mentions that **s-lwayŋ* ‘always means BREAST, not MILK.’¹⁴

Another significant area is found at the China-Myanmar-India border area. The Northern Naga group is known as an ‘areal’ group of this area. The words for ‘sun’ that are derived from PTB **r(i/u)ŋ* are found concentrated only in this area, and not only the Northern Naga group, but also the North Assam group in this area have the **riŋ* type of words for ‘sun’ (Section 3.2.4).

Each root for ‘iron’ that can be traced back to a PTB etymon shows considerably wide distribution. Among them, the words derived from PTB **sya:l* \bowtie **syi:r* ‘iron’ are distributed in a relatively narrow area. They are concentrated in western Myanmar, northeastern India, and eastern Nepal (Kurabe forthcoming). Thus, the eastern area of the distribution of PTB **sya:l* \bowtie **syi:r* ‘iron’ overlaps with the distribution of PTB **r(i/u)ŋ* ‘sun.’

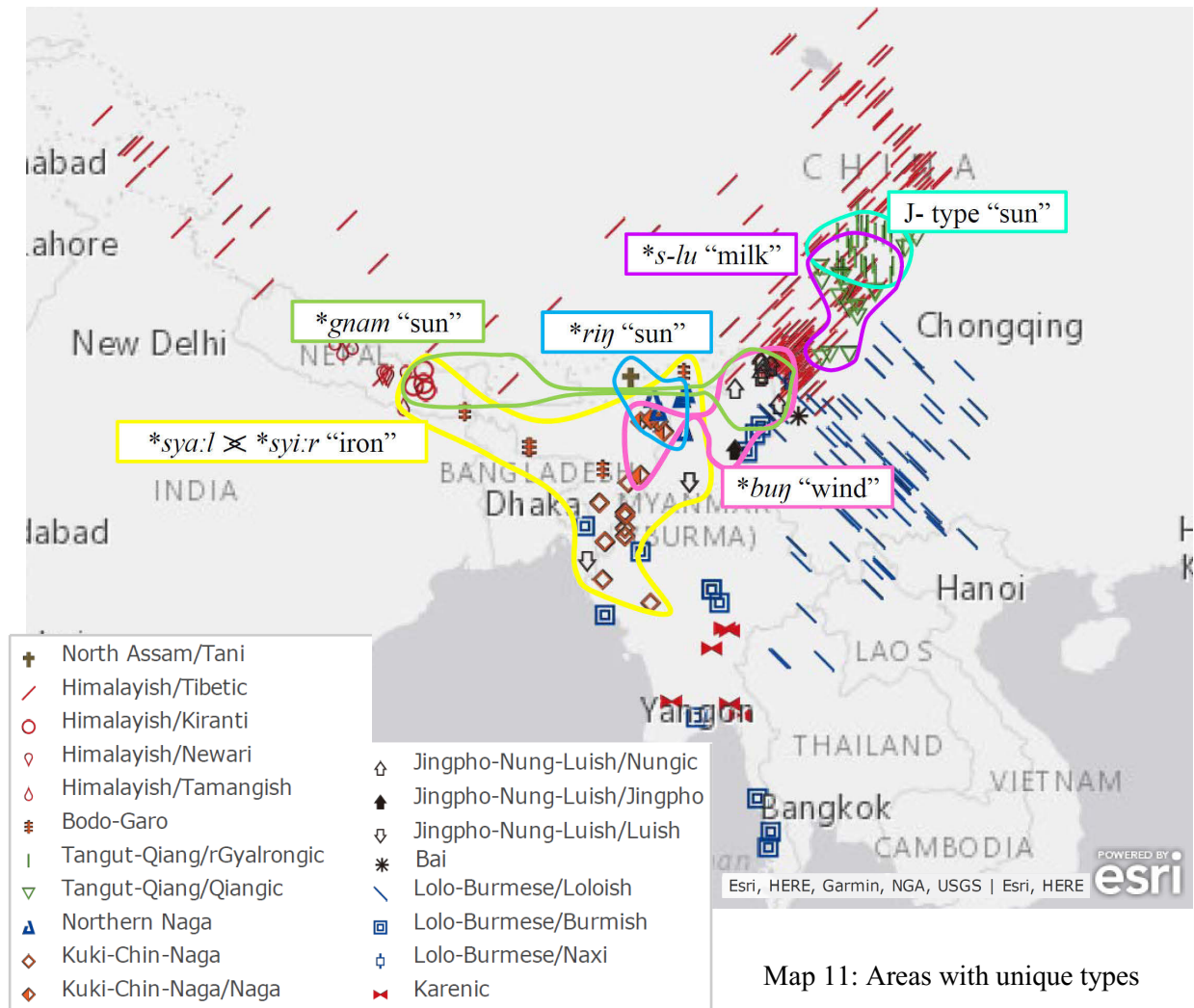
The distribution of words for ‘wind’ that are derived from PTB **buŋ* ‘wind’ is noteworthy. It is concentrated around this area, with interruption in the middle of the area. As is mentioned in Iwasa et al. (forthcoming), Nungic and Jingpho languages (of the China-Myanmar border area) have the **buŋ* type.

¹³ Our first attempt at studying the distribution of words for ‘rice plant,’ which appears as Suzuki et al. (2016), lacks information about the languages of Tamangish, Kiranti, and Tsangla. Here, I briefly survey these varieties using the information from the STEDT database, although the distribution is scattered and it is difficult to find the oldest type. Words for ‘sun’ in Tsangla can be traced back to PTB **kuk* and this etymon is shared with Bai and several Lolo-Burmese languages. Words for ‘sun’ in Kiranti can be mainly traced back to PTB **m-dz(y)a-k/n/t/s*, which is shared only with Nusu. Words for ‘sun’ in Tamangish languages mainly have voiceless fricative initials. They are similar to the ‘S-’ type in Suzuki et al. (2016), which is found in the southeastern area.

¹⁴ <http://stedt.berkeley.edu/~stedt/cgi/rootcanal.pl/etymon/1786>, accessed on April 23, 2017.

In addition, in the western side of the Myanmar-India border area, the Central Naga group has the **bung* type of words for ‘wind,’ for example, Chungli Ao *mu²puŋ²* (STEDT). However, the Northern Naga group, who resides in the middle of the area, do not have the **bung* type words for sun, but do have the **g-lay* type, that is, the most widespread type of ‘wind’ (Table 2).¹⁵ This suggests that the **g-lay* type is the newer in this area. Consequently, the **g-lay* type has only lately been brought into this area either through borrowing or migration.

The isoglosses in Map 11 suggest that the language groups in these areas diverged from the other TB groups at an early stage and underwent an individual process of language change. We should concentrate on the historical process of their formation in future research.



5 Conclusion

First, this paper has introduced the geolinguistic approach to TB vocabulary conducted as part of the Asian Geolinguistic Project. I have pointed out that some of our past research lacked information on several groups of the TB language family, which should be covered in later revisions.

Next, I re-examined the distribution of the words for ‘sun’ in TB with additional data. The data supported our conclusion on the **nəy* type: it is expected to be the oldest type from the geolinguistic perspective. I also revised our earlier conclusion, especially on the time-depth of the M- + **nəy* type,

¹⁵ This is a clear ABA distribution (See Note 8).

TS- type, and **gnam* type. The M- + **nəy* types of compound are expected to be older than the M- stem type at least in the Loloish group (Southwestern China), while the **nəy* + GI- types are expected to be older than the GI- stem type. The TS- type and **gnam* type are geographically widespread and expected to be considerably old. Moreover, I analyzed three new types of morphemes: the J- type, **riŋ* type, and HW type. The J- type and **riŋ* type are found as stem types, but their limited distribution suggests that they are relatively new forms for ‘sun’.

Finally, I reviewed our past attempt at researching TB vocabulary. I listed the widespread types, and simultaneously pointed out the difficulty in deciding such types. Moreover, I picked up two significant areas with unique vocabulary: western Sichuan and the China-Myanmar-India border.

We can expect that the continuous investigation into more items of vocabulary could clarify the time depth of the divergence of language groups in TB: this is one of our goals. Another expected goal is to clarify the process of formation of the linguistic areas. To achieve this, we need to compile more data (especially with regard to varieties of all groups) and learn more about the history of each word, each area, and each people.

Acknowledgments

This work is supported by the Japan Society for the Promotion of Science Kakenhi #16H03414. I would like to thank Prof. Mitsuaki Endo and the members of the TB team of the Asian Geolinguistic Project for their advice and support in this work. I also would like to thank Editage (www.editage.jp) for English language editing, and Esri (www.esri.com) for the ArcGIS Online service for drawing maps.

References

- Bradley, David. 1979. *Proto-Loloish*. London: Curzon Press.
- Burling, Robbins. 1967. *Proto-Lolo-Burmese*. The Hague: Mouton.
- Burling, Robbins. 1983. The sal languages. *Linguistics of the Tibeto-Burman Area* 7(2):1-32.
- Coblin, Weldon South. 1986. *A sinologist's handlist of Sino-Tibetan lexical comparisons*. (Monumenta Serica Monograph Series, 18). Nettetal: Steyler Verlag.
- Dryer, Mathew S. 2008. Word order in Tibeto-Burman languages. *Linguistics of the Tibeto-Burman Area*. 31(1): 1-83.
- Dryer, Mathew S. 2017. Word order in Sino-Tibetan languages from a typological and geographical perspective. Graham Thurgood and Randy J. LaPolla (eds.) *The Sino-Tibetan Languages*. 2nd ed. pp. 70-82. London/New York: Routledge.
- Ebihara, Shiho, K. Iwasa, K. Kurabe, S. Shirai, H. Suzuki and I. Matsuse. 2016. Milk: Tibeto-Burman. M. Endo (ed.) *Studies in Asian Geolinguistics III*. Fuchu: ILCAA, Tokyo University of Foreign Studies.
- Iwasa, Kazue, H. Suzuki, K. Kurabe, S. Ebihara, S. Shirai and I. Matsuse. forthcoming. Wind: Tibeto-Burman. M. Endo (ed.) *Studies in Asian Geolinguistics IV*. Fuchu: ILCAA, Tokyo University of Foreign Studies.
- Jacques, Guillaume [向柏霖]. 2008. *A Study of rGyalrong* [嘉絨語研究]. 中国新發現語言研究叢書. Beijing: Minzu Chubanshe [民族出版社].
- Kurabe, Keita, H. Suzuki, K. Iwasa, S. Ebihara, S. Shirai and I. Matsuse. forthcoming. Iron in Tibeto-Burman. M. Endo (ed.) *Studies in Asian Geolinguistics V*. Fuchu: ILCAA, Tokyo University of Foreign Studies.
- Matisoff, James A. 1990. On Megalocomparison. *Language* 66(1): 106-120.
- Matisoff, James A. 1991. Sino-Tibetan Linguistics: Present State and Future Prospects. *Annual Review of Anthropology* 20: 469-504.
- Matisoff, James A. 1978. *Mpi and Lolo-Burmese microlinguistics*. (Monumenta Serindica 4). Fuchu: ILCAA, Tokyo University of Foreign Studies.
- Matisoff, James A. 2003. *Handbook of Proto-Tibeto-Burman: System and philosophy of Sino-Tibetan reconstruction*. Berkeley / Los Angeles / London: University of California Press.

- Matisoff, James A. 2013. Re-examining the genetic position of Jingpho: Putting flesh on the bones of the Jingpho/Luish relationship. *Linguistics of the Tibeto-Burman Area* 36.2: 15-95.
- Opgenort, Jean Robert. 2011. A note on Tilung and its position within Kiranti. *Himalayan Linguistics* 10.1: 253-271. [<http://escholarship.org/uc/himalayanlinguistics>]
- Shirai, Satoko, K. Kurabe, K. Iwasa, H. Suzuki and S. Ebihara. 2016. Sun: Tibeto-Burman. In: Endo (ed.) *Studies in Asian Geolinguistics I*, pp. 14-17. Fuchu: ILCAA, Tokyo University of Foreign Studies.
- STEDT. *The Sino-Tibetan Etymological Dictionary and Thesaurus*. (<http://stedt.berkeley.edu/>, accessed on 2017-04-24).
- Suzuki, Hiroyuki. 2016. A geolinguistic description of terms for 'sun' in Tibetic languages of the eastern Tibetosphere. In: Endo (ed.) *Studies in Asian Geolinguistics I*, pp. 79-85. Fuchu: ILCAA, Tokyo University of Foreign Studies.
- Suzuki, Hiroyuki, S. Shirai, K. Kurabe, K. Iwasa, S. Ebihara and I. Matsuse. 2016. Rice plant: Tibeto-Burman. In: M. Endo (ed.) *Studies in Asian Geolinguistics II*. Fuchu: ILCAA, pp. 11-14. Tokyo University of Foreign Studies.

Data Sources (except for those in the references listed above and the field data provided by the members of the TB team: H. Suzuki, S. Ebihara, K. Kurabe, K. Iwasa, I. Matsuse, and S. Shirai.)

- Bielmeier, Roland (ed.) unpublished. Comparative Dictionary of Tibetan Dialects (CDTD) Nouns. Bern.
- Chen Kang [陳康]. 2010. *A Study on the Yi Dialects*. [彝語方言研究.] Beijing: Zhongyang Minzu Daxue Chubanshe [中央民族大学出版社].
- Duo'erji [多爾吉]. 1998. *A Study of Geshizha sTau* [道孚語格什扎話研究]. Beijing: Zhongguo Zangxue Chubanshe [中國藏學出版社].
- Ebihara, Shiho. 2010. *Amdo-Tibetan Pronunciation and Conversation* [アムド・チベット語の発音と会話]. Tokyo: ILCAA.
- Ebihara, Shiho. 2008. A Descriptive Study on the Amdo Dialect of Tibetan Spoken in Gonghe County, Qinghai Province [青海省共和県のチベット語アムド方言]. PhD dissertation. Tokyo: The University of Tokyo.
- Genetti, Carol. 2007. *A Grammar of Dolakha Newar*. Berlin/New York: Mouton de Gruyter.
- Haller, Felix. 2000. *Dialekt und Erzählungen von Shigatse* Wissenschaftsverlag. Bonn: VGH.
- Haller, Felix. 2004. *Dialekt und Erzählungen von Themchen: Sprachwissenschaftliche Beschreibung eines Nomadendialektes aus Nord-Amdo*. Bonn: VGH Wissenschaftsverlag.
- Hanson, Ola. 1906. *A Dictionary of the Kachin Language*. Rangoon: Baptist Mission Press.
- Hansson, Inga-Lill. 1989. A comparison of Akha, Hani, Khatu, and Pijo. *Linguistics of the Tibeto-Burman Area* 12.1: 1-91.
- Henderson, Eugénie J. A. 1997. *Bwe Karen Dictionary, with Texts and English-Karen word list*. London: School of Oriental and African Studies, University of London.
- Herrmann, Silke. 1989. *Dialekt und Erzählungen von Dinri*. Bonn: VGH Wissenschaftsverlag.
- Hoshi, Michiyo. 1978. *Zangskar Vocabulary*. Beijing: 中国社会科学出版社.
- Hoshi, Michiyo. 1987. *A Sharchok Vocabulary*. Tokyo: ILCAA.
- Hoshi, Michiyo & Hoshi, Izumi. 1995. *Tibetan Vocabulary* [チベット語語彙集]. Tokyo: ILCAA.
- Hua, Kan [華侃]. 2002. *Vocabularies of Amdo Dialect of Tibetan* [藏語安多方言詞匯]. Gansu: Gansu Minzu Chubanshe [甘肅民族出版社].
- Huang, Bufan [黃布凡]. 2007. *A Study of Lavrung* [拉烏絨語研究]. 中国新發現語言研究叢書. Beijing: Minzu Chubanshe [民族出版社].

- Huang, Bufan ed. [黄布凡主編]. 1992. *A Tibeto-Burman Lexicon* [藏緬語族語言詞匯]. Beijing: Zhongyang Minzuxueyuan Chubanshe [中央民族学院出版社].
- Huang, Chenglong. 2007. *A Study of Puxi Qiang* [浦溪羌語研究]. 中国少数民族語言方言研究叢書. Beijing: Minzu Chubanshe [民族出版社].
- Huziware, Keisuke [藤原敬介]. 2008. A descriptive linguistic study of the Cak language [チャック語の記述言語学的研究] PhD Dissertation. Kyoto Univesity.
- Huziware, Keisuke [藤原敬介]. 2012. Towards a reconstruction of Proto-Luish [レイ祖語の再構にむけて]. *Kyoto University Linguistic Research* [京都大学言語学研究] 31: 25-131.
- Iwasa, Kazue. 2003. A Wordlist of the Ma Ndzi Language. In: *Descriptive and Theoretical Studies in Minority Languages of East and Southeast Asia* [論集：東・東南アジアの少数言語の現地調査] 3.A3-16: 69-76. Kyoto: Faculty of Informatics, Osaka Gakuin University [大阪学院大学情報学部].
- Iwasa, Kazue. 2004. Axi and Azha - Descriptive, Comparative and Sociolinguistic Analyses of Two Lolo Dialects of China. PhD Dissertation. Kobe: Kobe City University of Foreign Studies.
- Jiang, Li. 2015. *A Grammar of Guiqiong: A Language of Sichuan*. Leiden/Boston: Brill.
- Kato, Atsuhiko. 2008. *Basic data on the Geba language* [ゲーバー語基礎資料]. *Asian and African Languages and Linguistics* 3: 169-219.
- Kato, Atsuhiko. 2009. A basic vocabulary of Htoklibang Pwo Karen with Hpa-an, Kyonbyaw, and Proto-Pwo Karen forms. *Asian and African Languages and Linguistics* 4: 169-218.
- Kato, Atsuhiko and Khin Pale. 2012. The Myeik (Beik) dialect of Burmese. *Journal of Asian and African Studies* 83: 117-160.
- Kato, Takashi. 2008. *Linguistic Survey of Tibeto-Burman Languages in Lao P.D.R.* Fuchu: ILCAA, Tokyo University of Foreign Studies.
- Kiryu, Kazuyuki. Meche. (ms).
- LaPolla, Randy J. 2003. Rawang glossary. (ms.)
- LaPolla, Randy with Huang Chenglong. 2003. *A Gramar of Qiang*. Berlin/New York: Mouton de Gruyter.
- Liu, Guangkun [劉光坤]. 1998. *A Study of Mawo Qiang* [麻窩羌語研究]. 中国少数民族語言方言研究叢書. Chengdu: Sichuan Minzu Chubanshe [四川民族出版社].
- Matisoff, James A. 2006. *English-Lahu Lexicon*. California: University of California Press.
- Nagano, Yasuhiko and Marielle Prins. 2013. rGyalrongic Languages Database. (<http://htq.minpaku.ac.jp/databases/rGyalrong/>).
- Nishida, Tatsuo [西田龍雄]. 1966. A Preliminary Study on the Bisu Language: A Language of Northern Thailand, Recently Discovered by Us [ビス語の研究：タイ国北部におけるビス族の言語の予備的研究]. *Southeast Asian Studies* [東南アジア研究] 4.1: 65-87.
- Norman, Rebecca. 2001. *Getting Started in Ladakhi: A Phrasebook for Learning Ladakhi* (Revised and Expanded Edition) Leh: Melong Publications.
- Ohno, Toru [大野徹]. 2000. *A Burmese (Myanmar) - Japanese Dictionary* [ビルマ語（ミャンマー語）辞典]. 東京: 大学書林.
- Otsuka, Kosei [大塚行誠]. 2014. A basic vocabulary of the Palaw dialect of Burmese [ビルマ語パロー方言基礎語彙]. *Asian and African Languages and Linguistics* [アジア・アフリカの言語と言語学] 8: 163-200. Fuchu: ILCAA, Tokyo University of Foreign Studies.
- Qin Liying. unpublished. Trung Vocabulary. (ms.)
- Qu, Aitang and KerangTan [瞿霽堂、譚克讓] 1983. *Ngari Tibetan* [阿里藏語]. Beijing: Chinese Academy of Social Sciences Press [中国社会科学出版社].

- Sawada, Hideo [澤田英夫]. 2004. A tentative etymological word-list of Lhaovo (Maru) language. *Approaches to Eurasian Linguistic Areas*, pp. 61-122. Kobe: Kobe City College of Nursing.
- Shintani, Tadahiko L.A. 2009. *The Pyen (or Phen) language: Its classified lexicon*. Fuchu: ILCAA, Tokyo University of Foreign Studies.
- Shintani, Tadahiko L.A. 2014. *The Zayein Language*. Fuchu: ILCAA, Tokyo University of Foreign Studies.
- Solnit, David. 1997. *Eastern Kayah Li: Grammar, Texts, Glossary*. Honolulu: University of Hawaii Press.
- Sprigg, Richard Keith. 2002. *Balti-English English-Balti dictionary*. Richmond: Routledge Curzon.
- Srinuan Duanghom. 1976. *An Mpi Dictionary*. Bangkok: Indigenous Languages of Thailand Research Project, Central Institute of English Language.
- Sun, Hongkai, Katia Chirkova and Guangkun Liu [孫宏開、齊卡佳、劉光坤]. 2007. *A Study of Baima* [白馬語研究]. 中國新發現語言研究叢書. Beijing: Minzu Chubanshe [民族出版社].
- Sun, Hongkai & Guangkun Liu [孫宏開、劉光坤]. 2009. *A Grammar of Anong: Language Death under Intense Contact*. Leiden: Brill.
- Sun, Hongkai, Dan Xu, Guangkun Liu and Lurongduoding [孫宏開、徐丹、劉光坤、魯絨多丁]. 2014. *A Study of Shihing* [史興語研究]. 中國新發現語言研究叢書. Beijing: Minzu Chubanshe [民族出版社].
- Taylor, L. F. 1921. The dialects of Burmese (I). *The Journal of the Burma Research Society* 11: 89-97.
- van Driem, George. 1992. *The Grammar of Dzongkha*. Bhutan: RGoB, Dzongkha Development Commission.
- VanBik, Kenneth. 2009. *Proto-Kuki-Chin: a Reconstructed Ancestor of the Kuki-Chin Languages*. Berkeley: University of California.
- Wannemacher, Mark and Zau Mo. 2000. *A Preliminary Zaiwa - English Lexicon with English - Zaiwa Glossary*. Chiang Mai: Payap University.
- Xu Xianming, Bai Bibo et al. [許鮮明, 白碧波 等]. 2013. *Studies on the Hlersu Language* [山蘇彝語研究]. Beijing: Minzu Chubanshe [民族出版社].
- Yabu, Shiro [藪司郎]. 1980. Linguistic data of the Yaw dialect of the Burmese language [ビルマ語ヨー方言の資料]. *Journal of Asian and African Studies* [アジア・アフリカ言語文化研究] 19: 164-182. Fuchu: ILCAA, Tokyo University of Foreign Studies.
- Yabu, Shiro [藪司郎]. 1981. Linguistic data of the Taung'yo dialect of the Burmese language [ビルマ語タウンヨウ方言の資料]. *Journal of Asian and African Studies* [アジア・アフリカ言語文化研究] 21: 154-187. Fuchu: ILCAA, Tokyo University of Foreign Studies.
- Yabu, Shiro [藪司郎]. 1988. A preliminary report on the study of the Maru, Lashi and Atsi languages of Burma. *Historical and cultural studies in Burma*, pp. 65-132. Tokyo: Sophia University.
- Yuan, Jiahua [袁家驊]. 1953. *Folksongs and Language of Axi* [阿細民歌及其語言]. Beijing: Zhongguo Kexue Chubanshe [中国科学院出版]
- Yunnansheng Lunan Yizu zizhixian Wenshi Yanjiushi (ed.) [云南省路南彝族自治州縣文史研究室編]. 1982. *A Concise Yi-Chinese Dictionary* [彝漢簡明詞典]. Kunming: Yunnan Minzu Chubanshe [云南民族出版社].
- Zangmianyu Yuyin he Cihui Bianxiezhu [《藏緬語語音和詞匯》編寫組]. 1991. *Tibeto-Burman Phonology and Lexicon* [藏緬語語音和詞匯]. Beijing: Zhongguo Shehuikexue Chubanshe [中国社会科学出版社].
- Zhongyangminyuan Yizu Lishi Wenxian Bianyishi, Zhongyangminyuan Yizu Lishiwenxianban [中央民院彝族歷史文獻編譯室, 中央民院彝族歷史文獻班]. 1984. *Yunnan-Sichuan-Guizhou-Guangxi: A Comparative Yi-Chinese Dictionary of Basic Vocabulary* [滇川黔桂 彝漢基本詞匯對照詞典]. (Mimeographed version [油印本]).
- Zhou mao cao [周毛草]. 2003. *A Study on rMachu Tibetan* [瑪曲藏語研究]. Beijing: Minzu Chubanshe [民族出版社].

The distribution of diphthongs in Vietnamese dialects

Masaaki Shimizu^a and Mika Kondo^b

^aGraduate School of Language and Culture, Osaka University,

^bJoint Researcher, ILCAA, Tokyo University of Foreign Studies

Abstract

This study aims to show the geographical distribution of diphthongs in Vietnamese dialects for the purpose of rethinking their status in the syllable structure of each dialect and pointing out the changes that occur in the Southern dialect. The focus of the analysis are the certain diphthongs in the Northern and Southern dialects and the monophthongs that tend to be diphthongized in the Southern dialect.

1 Introduction

3 diphthongs in the vowel system of the Northern dialect (NV) and Southern dialect (SV) of Vietnamese (see (1)), also called *contour vowels* in the recent literature (Emerich 2012), are quite problematic in that, especially in the Southern dialect, they tend to be monophthongized when they occur in the closed syllables (i.e., they merge with some monophthongs in the phoneme inventory) while they stay diphthongs when they occur in the open syllables. The vowel inventory of these two dialects is given in (1).

(1)	ie	uɤ	uo	iə	uə	uə
	i	u	u	i	i/u	u
	e	ɤ	o	e	ɤ	o
	ɜ	ʌ	ɔ	ɜ	ʌ	ɔ
	a	ɐ		a	ɐ	
	NV (HN16)			SV (SG05)		

Some examples of the NV diphthongs and their correspondence in SV are shown in (2), where the cases that are monophthongized are labeled as (a) and those that are not are labeled as (b).

(2)	Orthography	NV	SV	
(a)	<i>tiêm</i>	tiem1	tim1	‘to inject’
	<i>tuôm</i>	tuɤm1	tum1	‘to emit’
	<i>buôm</i>	buom2	buum2	‘sail’
	<i>tiền</i>	tien2	tiŋ2	‘money’
	<i>lươn</i>	luɤn1	luŋ1	‘eel’
	<i>cuồn</i>	kuon5	kuŋ5	‘to roll’
	<i>tiếng</i>	tien5	tiŋ5	‘voice’
	<i>thương</i>	tʰuɤŋ1	tʰuŋ1	‘injured’
	<i>uống</i>	ʔuon5	ʔuŋ5	‘to drink’
(b)	<i>tia</i>	tie1	tiə1	‘ray’
	<i>tư</i>	tuɤ1	tuə1	‘fur’
	<i>tua</i>	tuo1	tuə1	‘fringe’



Fig.1 *tiem* (tiem1: ①, tim1: ◆)



Fig.2 *tia* (tiə1: ①, ti1: ◆)

The distribution of (a) and (b) in the dialects is shown in Fig.1—the case with final *m*— and Fig.2. It is quite obvious that an isogloss can be recognized at the location of Hải Vân Pass which lies between Huế and Đà Nẵng City.

The question here is why the difference between (a) and (b) occurs in SV. We will think of other phenomena observed in SV below to explain this difference.

2 Diphthongization

Another problem concerning the vowels in SV is the intrusive sound clearly recognized between the initials and the high vowels /i u/ in the open syllables. Their actual value is as follows:

(3)	Orthography	NV	SV	
	<i>tí</i>	[ti:ʔ]	[tɕi:ʔ]	‘service’
	<i>tư</i>	[tu:ʔ]	[tɕu:ʔ]	‘private’
	<i>tu</i>	[tu:ʔ]	[tɕu:ʔ]	‘to practice’

Furthermore, it is noteworthy that the distribution of the SV forms in (3) is exactly same as that of SV forms in (2) (see Fig.3~5.).

Another kind of vowel change can be found in the case of lower vowels /e ɤ o/ in the open syllables, but their geographical distribution is different from that of (3). Their actual value is follows:

(4)	Orthography	NV	SV	
	<i>tê</i>	[te:ʔ]	[te:ĩʔ]	‘numbed’
	<i>tơ</i>	[tɤ:ʔ]	[tɤ:ốʔ]	‘silk’
	<i>tô</i>	[to:ʔ]	[to:ũʔ]	‘bowl’



Fig.3. *ti*
([ti:1]:◆, [tɿ:1]:⊙)



Fig.4. *tur*
([tur:1]:◆, [tʰur:1]:⊙)



Fig.5. *tu*
([tu:1]:◆, [tʰu:1]:⊙)



Fig.6. *tê*
([te:1]:◆, [te:ĩ]:⊙)



Fig.7. *tor*
([tr:1]:◆, [tr:õ]:⊙)



Fig.8. *tô*
([to:1]:◆, [to:ũ]:⊙)

Both (3) and (4) can be regarded as kinds of diphthongization, in which the sprouting short vowels are not phonemic.

Considering the phenomena in (2), (3), and (4), we can think of two isoglosses as follows:

(5)

Isoglosses	phonological evidences	(Shinizu 2016)
① Hải Vân Pass	(2), (3)	
② Ninh Thuận-Bình Thuận	(4)	v [w/j~]

The isogloss between Ninh Thuận and Bình Thuận is also the border between [v] and [v/w/j] for the initial *v* (Shimizu 2016).

3 Vowel shift

The next phenomena that are concerned with the vowels in SV can be taken as a vowel shift to avoid the phonemic merger. The cases are those forms containing the rhymes *ây/âu*, *ay/au*, and *ai/ao* in orthography. Their realization in each dialect is as follows:

(6)	Orthography	NV	SV	Quảng Nam	
	<i>cây</i>	kaj1	kɛj1	kɛj1	‘tree’
	<i>cay</i>	kɛj1	kaj1	ka1	‘hot’
	<i>cái</i>	kaj5	kaj5	kəa5	‘female’
	<i>cầu</i>	kaw2	kɛw2	kɛw2	‘bridge’
	<i>sáu</i>	səw5	ɣaw5	ɣa5	‘six’
	<i>sao</i>	saw1	ɣaw1	ɣo1	‘star’



Fig.9. cây
(kaj1: ♀, kɛj1: ○)



Fig.10. cay
(kɛj1: ○, kaj1: ●, ka1: ★)



Fig.11. cái
(kaj5: ●, kəa5: ★)



Fig.12. cầu
(kaw2: ♀, kɛw2: ○)



Fig.13. sáu
(səw5: ○, ɣaw5: ●, ɣa5: ★)



Fig.14. sao
(saw1, ɣaw1: ●, ɣo1: ★)

Their geographical distribution is shown in Fig.9~14.

Fig.9 and Fig.10 show that the distribution of $k\Lambda j1/k\epsilon j1$ and $k\epsilon j1/kaj1$ is quite similar, and this is also the case for $k\Lambda w2/k\epsilon w2$ and $s\epsilon w5/saw5$ (Fig.12, 13). This means, when $\Lambda j > \epsilon j$ and $\Lambda w > \epsilon w$ occurred, $\epsilon j > aj$ and $\epsilon w > aw$ also occurred at the same time so as to avoid the phonemic merger. As a result, it was inevitable to cause the merger between **ay/ai** and **au/ao** in the same regions (see Fig.10/11, 13/14).

4 Interpretation

So far those processes that occurred in SV vowels have been introduced. Here we try to find the correlation among those processes which can be summarized as follows:

- | | | |
|-----|---|------|
| (7) | Monophthongization | (2a) |
| | Diphthongization I | (3) |
| | Diphthongization II | (4) |
| | Vowel shift: $\Lambda > \epsilon > a / _ \{j, w\}$ | (6) |

We suppose that (2a) and (3) are highly correlated in that they are concerned with the characteristics of the syllable structure of SV. Based on the moraic theory, each syllable may have two moras at most, and SV has the tendency of ‘one mora = one simple element’ (simple: [i], [u], [u], etc.), while that of NV may contain the internal contour ([ie], [ur], [uo]) within one mora (Emerich 2012).

- (8)
- | | |
|---|---|
| <pre> graph TD sigma[σ] --- t[t] sigma --- mu1[μ] sigma --- mu2[μ] mu1 --- ie[ie] mu2 --- m[m] </pre> <p>NV ‘to inject’</p> | <pre> graph TD sigma[σ] --- t[t] sigma --- mu1[μ] sigma --- mu2[μ] mu1 --- i[i] mu2 --- m[m] </pre> <p>SV ‘to inject’</p> |
|---|---|

In SV, according to the principle ‘one mora = one simple element,’ two elements of the NV diphthongs in the open syllables may have been reanalyzed as in (9), and the values of diphthongs remained as they are.

- (9)
- | | |
|---|---|
| <pre> graph TD sigma[σ] --- t[t] sigma --- mu1[μ] sigma --- mu2[μ] mu1 --- ie[ie] mu2 --- ie[ie] </pre> <p>NV ‘ray’</p> | <pre> graph TD sigma[σ] --- t[t] sigma --- mu1[μ] sigma --- mu2[μ] mu1 --- i[i] mu2 --- e[e] </pre> <p>SV ‘ray’</p> |
|---|---|

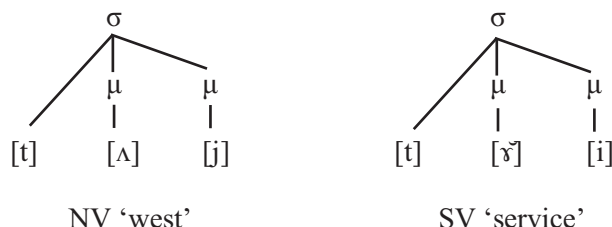
Concerning Diphthongization I, SV still has the tendency of ‘one mora = one simple element’ and fills one of the mora with an intrusive vowel.

- (10)
- | | |
|---|---|
| <pre> graph TD sigma[σ] --- t[t] sigma --- mu1[μ] sigma --- mu2[μ] mu1 --- i[i] mu2 --- i[i] </pre> <p>NV ‘service’</p> | <pre> graph TD sigma[σ] --- t[t] sigma --- mu1[μ] sigma --- mu2[μ] mu1 --- schwa[ə] mu2 --- i[i] </pre> <p>SV ‘service’</p> |
|---|---|

The geographical range adapting Diphthongization II is more limited than Diphthongization I, which might be concerned with the sonority of vowels.

Once Diphthongization I has occurred, it is highly possible that the syllables with rhymes *-i* and *-u* may have auditory similarity with those of *-ây* and *-âu*, just as in the case of (11), which might have been the trigger for the changes in (6): $\Lambda j > \text{ej}$, $\Lambda w > \text{ew}$.

(11)



5 Conclusion

This study has attempted to determine why not all of the NV diphthongs had undertaken the process of monophthongization in SV. It also supposes a characteristic of the syllable structure in SV, which can be summarized as: ‘one mora = one simple element.’

References

- Emerich, G.H. 2012. *The Vietnamese Vowel System*. A Dissertation in Linguistics Presented to the Faculties of the University of Pennsylvania, University of Pennsylvania, USA.
- Hoàng, T.C. 2004. *Phương Ngữ Học Tiếng Việt* (Vietnamese dialectology). Nxb. Đại học Quốc gia Hà Nội: Hà Nội.
- Martinet, A. 1952. Function, structure, and sound change. *Word*, vol.8, no.1, pp.1-32.
- Shimizu, M. 2014. *The Distribution of Final Palatals in Vietnamese Dialects*. Papers from the 2nd International Conference on Asian Geolinguistics, Bangkok, Thailand, 24-25 May, 2014.
- Shimizu, M. 2016. *A Phonological Basis for Rethinking Vietnamese Isoglosses*. Papers from the 3rd International Conference on Asian Geolinguistics, Royal University of Phnom Penh (RUPP), 23-24 May, 2016.
- Thompson, L.C. 1959. Saigon Phonemics. *Language*, vol.35, no.3, pp.454-476.

No.	Province	Birth Place	<i>tia</i>	<i>ti</i>	<i>tê</i>	<i>te</i>	<i>tua</i>	<i>tur</i>	<i>tơ</i>	<i>ta</i>	<i>tua</i>	<i>tu</i>	<i>tô</i>	<i>to</i>	<i>tiêm</i>
HN-01	Thái Nguyên	ucky	ti:ə1	ti:1	te:1	te: ^a 1	ti:ə1	ti:1	tə:1	ta:1	tux1	tu:1	to:1	tə:1	tiəm1
HN-09	Vĩnh Phúc	ducbac	ti:ə1	ti:1	tê:1	te:1	ti:ə1	ti:1	tə:1	ta:1	tux1	tu:1	to:1	tə:1	tiəm1
HN-15	Bắc Giang	quenham	tia1	ti:1	te:1	te:1	tua:1	tua:1	tr:1	ta:1	tua1	tu:1	to:1	tə:1	tiəm1
HN-06	Bắc Ninh	bacninh	ti:ə1	ti:1	tê:1	te: ^a 1	ti:ə1	ti:1	tr:1	ta:1	tua1	tu:1	to:1	tə:1	tiəm1
HN-08	Hải Dương	namsach	ti:ə1	ti:1	tê:1	tea1	ti:ə1	ti:1	tr:1	ta:1	tux1	tu:1	to:1	tə: ^a 1	tiəm1
HN-19	Hải Dương	ngocchau	tiə1	ti:1	te:1	te: ^a 1	tua1	tua:1	tr:1	ta:1	tua1	tu:1	to:1	tə:1	tiəm1
HN-05	Hung Yên	vinhkhuc	tiə1	ti:1	te:1	te:1	tiə1	ti:1	tr:1	ta:1	tua1	tu:1	to:1	tə:1	tiəm1
HN-20	Hải Dương	trungkhanh	tiə1	ti:1	te:1	te: ^a 1	tua: ^a 1	tua:1	tr:1	ta:1	tua1	tu:1	to:1	tə:1	tiəm1
HN-13	Hung Yên	thuanhung	ti:ə1	ti:1	te:1	te:1	tua:1	tua:1	tr:1	ta:1	tua1	tu:1	to:1	tə:1	tiəm1
HN-03	Hà Tây	hatay	ti:ə1	ti:1	te:1	te: ^a 1	ti:ə1	ti:1	tr:1	ta:1	tux1	tu:1	to:1	tə:1	tiəm1
HN-11	Hải Phòng	dongphuong	ti:ə1	ti:1	te:1	te: ^a 1	tua:1	tua:1	tr:1	ta:1	tua1	tu:1	to:1	tə:1	tiəm1
HN-18	Hải Dương	tucuong	tiə1	ti:1	te:1	te: ^a 1	tua1	tua:1	tr:1	ta:1	tua1	tu:1	to:1	tə:1	tiəm1
HN-17	Hải Dương	huthai	tiə1	ti:1	te:1	te: ^a 1	tua1	tua:1	tr:1	ta:1	tua1	tu:1	to:1	tə: ^a 1	tiəm1
HN-10	Hà Tây	unghoa	ti:ə1	ti:1	tê:1	te: ^a 1	ti:ə1	ti:1	tə:1	ta:1	tua:1	tu:1	to:1	tə:1	tiəm1
HN-02	Hung Yên	antao	ti:ə1	ti:1	te:1	te:1	ti:ə1	ti:1	tr:1	ta:1	tux1	tu:1	to:1	tə:1	tiəm1
HN-07	Hà Nam	lynhan	ti:ə1	ti:1	tê:1	tea1	ti:ə1	ti:1	tr:1	ta:1	tua1	tu:1	to:1	tə:1	tiəm1
HN-04	Hà Nam	thanhliem	ti:ə1	ti:1	tê:1	te: ^a 1	ti:ə1	ti:1	tr:1	ta:1	tua1	tu:1	to:1	tə:1	tiəm1
HN-16	Nam Định	daithang	ti:a1	ti:1	te:1	te: ^a 1	tua:1	tua:1	tr:1	ta:1	tua:1	tu:1	to:1	tə:1	ti: ^e m1
HN-30	Thanh Hoá	thanhtho	tiə1	ti:1	te:1	te: ^a 1	tua1	tua:1	tr:1	ta:1	tua1	tu:1	to:1	tə:1	tiəm1
HN-14	Nam Định	nghiahunh	ti:a1	ti:1	te:1	te:1	tua:1	tua:1	tr:1	ta:1	tua:1	tu:1	to:1	tə:1	tiəm1
HN-28	Thanh Hoá	hunglocth	tia1	ti:1	te:1	te:1	tua1	tua:1	tr:1	ta:1	tua1	tu:1	to:1	tə:1	tiəm1
HN-27	Thanh Hoá	hoangkhanh	ḍi:ə1	ḍi:1	tê:1	te: ^a 1	ti:ə1	ti:1	tə:1	ta:1	tua:1	tu:1	to:1	tə:1	ti: ^e m1
HN-24	Thanh Hoá	hauloc	ti:ə1	ti:1	tê:1	te: ^a 1	ti:ə1	ti:1	tə:1	ta:1	tua:1	tou1	too1	tə:1	tiəm1
HN-33	Nghệ An	quyhop	tiə1	ti:1	te:1	te:1	tua1	tua:1	tr:1	ta:1	tua1	tu:1	to:1	tə:1	tiəm1
HN-29	Nghệ An	dienthap	tiə1	ti:1	te:1	te:1	tua1	tua:1	tr: ^a 1	ta:1	tua1	tu:1	to:1	tə:1	tiəm1
HN-32	Nghệ An	phuthanh	tiə1	ti:1	te:1	te:1	tua1	tua:1	tr:1	ta:1	tua1	tu:1	to:1	tə:1	tiəm1
HN-22	Nghệ An	vandien	ti:ə1	ti:1	tê:1	te:1	ti:ə1	ti:1	tə:1	ta:1	tux1	tu:1	to:1	tə:1	tiəm1
HN-25	Nghệ An	hungloc	ti:ə1	ti:1	tê:1	tea1	ti:ə1	ti:1	tə:1	ta:1	tux1	tu:1	to:1	tə:1	tiəm1
HN-31	Nghệ An	vinh	tia1	ti:1	te:1	te:1	tua1	tua:1	tr:1	ta:1	tua1	tu:1	to:1	tə:1	tiəm1
HN-12	Hà Tĩnh	buixa	ti:ə1	ti:1	te:1	te:1	ti:ə1	ti:1	tr:1	ta:1	tua1	tu:1	to:1	tə:1	tiəm1
HU-03	Quảng Bình	lytrach	tix1	ti:1	te:1	te:1	tux1	tua:1	tr:1	ta:1	tux1	tu:1	to:1	tə:1	ti: ^y m1
HN-21	Quảng Bình	haidinh	tiə1	ti:1	tê:1	tea1	ti:ə1	tua:1	tr:1	ta:1	tux1	tu:1	to:1	tə:1	tiəm1
HU-01	Quảng Trị	vinhkim	tix1	ti:1	te:1	te:1	tua1	tua:1	tr:1	ta:1	tux1	tu:1	to:1	tə:1	tiəm1
HU-02	Huế	quangdien	ti:x1	ti:1	te:1	te:1	tua:x1	tua:1	tr:1	ta:1	tux1	tu:1	to:1	tə:1	ti: ^y m1
HU-07	Đà Nẵng	thoquang	ti: ^a 1	ti:1	te:1	te:1	ti: ^a 1	ti:1	tr:1	ta:1	tua1	ṭu: ^a 1	to:1	tə:1	ti: ^y m1
HU-17	Đà Nẵng	danang1	ti: ^a 1	ti:1	te:1	te:1	ti: ^a 1	ti:1	tr:1	ta:1	tua1	ṭu: ^a 1	to:1	tə:1	ti: ^y m1
HU-13	Đà Nẵng	camle1	tix1	ṭi:1	te:1	te:1	tix1	ti:1	tr:1	ta:1	tux1	ṭu: ^a 1	to:1	tə:1	ti: ^y m1
HU-12	Hội An	cudai	ti:ə1	ṭi:1	te:1	te:1	tua:1	ṭu: ^a 1	tr:1	ta:1	tua1	ṭu: ^a 1	to:1	tə:1	ti: ^y m1
HU-16	Hội An	hoian	ti:ə1	ṭi:1	tê:1	ṭi:1	ti:ə1	ṭu: ^a 1	tə:1	ta:1	tua1	ṭu: ^a 1	to:1	tə:1	ti: ^y m1
HU-09	Quảng Nam	daihoa	ti:x1	ṭi:1	te:1	te:1	ti:ə1	ṭu: ^a 1	tr:1	ṭa:1	tua1	ṭu: ^a 1	to:1	tə:1	ti: ^y m1
HU-06	Quảng Nam	duyhoa	ti:ə1	ṭi:1	te:1	te:1	tua:1	ṭu: ^a 1	tr:1	ṭa:1	tua1	ṭu: ^a 1	to:1	tə:1	ti: ^y m1
HU-15	Quảng Nam	tamgiang	ti:ə1	ṭi:1	tê:1	te:1	ti:ə1	ti:1	tə:1	ta:1	tua:1	to:1	to:1	tə:1	ti: ^y m1
HU-11	Quảng Nam	tammydong	ti:ə1	ṭi:1	te:1	te:1	tua:1	ṭu: ^a 1	tr:1	ṭa:1	tua1	ṭu: ^a 1	to:1	tə:1	ti: ^y m1
HU-05	Quảng Ngãi	tungghia	tiə1	ṭi:1	te:1	te:1	tua1	ṭu: ^a 1	tr:1	ta:1	tua1	ṭu: ^a 1	to:1	tə:1	ti: ^y m1
HU-19	Bình Định	quynhon	tix1	ṭi:1	te:1	te:1	tux1	ṭu: ^a 1	tr:1	ta:1	tux1	ṭu: ^a 1	to:1	tə:1	ti: ^y m1
HU-20	Phú Yên	tuyhoa	tix1	ṭi:1	te:1	ṭe:1	tux1	ṭu: ^a 1	tr:1	ta:1	tux1	ṭu: ^a 1	to:1	tə:1	ti: ^y m1
HU-21	Ninh Thuận	phuocson	ti:1	ṭi:1	te:1	te:1	tua:1	ṭu: ^a 1	tr:1	ta:1	tu:1	ṭu: ^a 1	to:1	tə:1	ti: ^y m1
SG-11	Long An	myhanhac	ti:ə1	ṭi:1	te: ^a 1	tea1	ti:ə1	ti: ^a 1	tr: ^a 1	ta:1	to:x1	ṭu: ^a 1	to: ^a 1	tə:1	ti: ^y m1
SG-04	Bình Thuận	thuanquy	ti: ^a 1	ṭi:1	te:1	te:1	tua: ^a 1	ṭu: ^a 1	tr:1	ta:1	tua: ^a 1	ṭu: ^a 1	to:1	tə:1	ti: ^y m1
SG-06	tp HCM	binhthanh1	ti:ə1	ṭi:1	te:1	te:1	tua:1	ṭu: ^a 1	tr:1	ta:1	tua1	ṭu: ^a 1	to:1	tə:1	ti: ^y m1
SG-07	Long An	tpHCM2	ti:ə1	ti:1	te: ^a 1	te:1	tua: ^a 1	ṭu: ^a 1	tr:1	ta:1	tua: ^a 1	ṭu: ^a 1	to: ^a 1	tə:1	ti: ^y m1
SG-05	Tiền Giang	tpHCM1	ti:x1	ṭi:1	te:1	te:1	tua: ^a 1	ṭu: ^a 1	tr:1	ta:1	tua: ^a 1	tu:1	to:1	tə:1	ti: ^y m1
SG-03	tp HCM	binhchanh	ti:ə1	ti:1	te: ^a 1	te: ^a 1	tua:1	ṭu: ^a 1	tr: ^a 1	ta:1	tua:1	tu:1	to: ^a 1	tə:1	ti: ^y m1
SG-10	BR-VT	xuyenmoc	ti:ə1	ti:1	te: ^a 1	te:1	tua:1	ṭu: ^a 1	tr:1	ta:1	tua1	tu:1	to:1	tə:1	ti: ^y m1
SG-12	An Giang	kienan	ti:ə1	ṭi:1	te: ^a 1	te:1	ti: ^a 1	tua: ^a 1	tr:1	ta:1	tua: ^a 1	ṭu: ^a 1	ṭu: ^a 1	tə:1	ti: ^y m1
SG-01	Bến Tre	p3bentre	ti:ə1	ṭi:1	te: ^a 1	te: ^a 1	ti:ə1	ṭu: ^a 1	tr: ^a 1	ta:1	tua:1	ṭu: ^a 1	to: ^a 1	tə:1	ti: ^y m1
SG-15	Bến Tre	luongquoi	ti:ə1	ṭi:1	te: ^a 1	te: ^a 1	ti:ə1	ṭu: ^a 1	tr: ^a 1	ta:1	tux1	ṭu: ^a 1	to: ^a 1	tə:1	ti: ^y m1
SG-02	Bến Tre	phuochiep	ti:ə1	ṭi:1	te: ^a 1	te: ^a 1	tua:1	ṭu: ^a 1	tr: ^a 1	ta:1	tua1	ṭu: ^a 1	to: ^a 1	tə: ^a 1	ti: ^y m1
SG-16	Vĩnh Phúc	tambinh	ti:ə1	ṭi:1	tê:1	tea: ^a 1	ti:ə1	ṭu: ^a 1	tr: ^a 1	ta:1	tua:1	ṭu: ^a 1	to: ^a 1	tə:1	ti: ^y m1
SG-08	Sóc Trăng	tanlong	ti:ə1	ti:1	te: ^a 1	te: ^a 1	tua:1	tua:1	tr:1	ta:1	tua:1	tu:1	to: ^a 1	tə:1	ti: ^y m1

No.	tuom	buom	tien	luon	cuon	tieng	thuong	uong	cau	sao	sau	cay	cai	cay
HN-01	tiam1	duy2	tien2	luon1	kuyn5	tieng5	thieng1	zuyn5	kau2	sa:ol	sao:5	kail	ka:is	kail
HN-09	tiam1	duam2	tien2	luon1	kuan5	tieng5	thieng1	zuang5	kao2	sa:ol	sao:5	kail	ka:is	kail
HN-15	tuym1	duam2	tien2	luyn1	kuon5	tieng5	thiuyng1	zuang5	kau2	sa:ul	sau5	kail	ka:is	kail
HN-06	tiam1	duom2	tien2	luon1	kuon5	tieng5	thieng1	zuang5	kau2	sa:ol	sao:5	kail	ka:is	kail
HN-08	tiam1	duam2	tien2	luon1	kuan5	tieng5	thieng1	zuang5	kau2	sa:ol	sao:5	kail	ka:is	kail
HN-19	tuym1	duam2	tien2	luyn1	kuon5	tieng5	thiuyng1	zuang5	kau2	ga:ul	gau5	kail	ka:is	kail
HN-05	tiam1	duam2	tien2	luon1	kuon5	tieng5	thieng1	zuang5	kau2	sa:ol	sao:5	kail	ka:is	kail
HN-20	tu'm1	duam2	tien2	nuyn1	kuon5	tieng5	thi'u'ng1	zuang5	kau2	ga:ul	gau5	kail	ka:is	kail
HN-13	tuym1	duam2	tien2	nuyn1	kuon5	tieng5	thiuyng1	zuang5	kau2	sa:ul	sau5	kaeil	ka:is	kail
HN-03	tiam1	duam2	tien2	luon1	kuan5	tieng5	thieng1	zuang5	kau2	sa:ol	sau:5	kail	ka:is	kail
HN-11	tuym1	du'm2	tien2	luon1	kuon5	tieng5	thiuyng1	zuang5	kau2	sa:ul	sau5	kail	ka:is	kail
HN-18	tuym1	duam2	tien2	nuyn1	kuon5	tieng5	thiuyng1	zuang5	kau2	sa:ul	sau5	kail	ka:is	kail
HN-17	tuym1	duam2	tien2	luyn1	kuon5	tieng5	thiuyng1	zuang5	kau'2	ga:ul	ga'u'1	kail	ka:is	kail
HN-10	tiam1	duom2	tien2	luon1	kuan5	tieng5	thieng1	zuang5	kau2	sa:ol	sao:5	kail	ka:is	kail
HN-02	tiam1	duam2	tien2	luon1	kuan5	tieng5	thieng1	zuang5	kau2	sa:ol	sao:5	kail	ka:is	kail
HN-07	tiam1	duom2	tien2	luon1	kuan5	tieng5	thieng1	zuang5	kau2	sa:ol	sao:5	kail	ka:is	kail
HN-04	tiam1	duam2	tien2	luon1	kuan5	tieng5	thieng1	zuang5	kau2	sa:ol	sao:5	kail	ka:is	kail
HN-16	tu:'m1	du'm2	tien2	lu'u'n1	kuon5	tieng5	thiuyng1	zuang5	kau2	fa:ul	fau5	kail	ka:is	kail
HN-30	tuym1	duam2	tien2	luyn1	kuon5	tieng5	thiuyng1	zuang5	kau2	sa:ul	sau5	kail	ka:is	kail
HN-14	tuym1	duam2	tien2	luyn1	kuon5	tieng5	thiuyng1	zuang5	kau2	sa:ul	sau5	kaeil	ka:is	kail
HN-28	tuym1	duam2	tien2	luyn1	kuon5	tieng5	thiuyng1	zuang5	kau2	sa:ul	sau5	kail	ka:is	kail
HN-27	ti'm1	duam2	tien2	li'n1	kuyn5	ti'h'ng5	thieng1	zu'ng5	kao2	sa:ol	sao5	kail	ka:is	kail
HN-24	tiam1	duam2	tien2	luon1	kuan5	tieng5	thieng1	zuang5	kao2	sa:ol	sao:5	kail	ka:is	kail
HN-33	tuym1	duam2	tien2	luyn1	kuon5	tieng5	thiuyng1	zuang5	kau2	sa:ul	sau5	kail	ka:is	kail
HN-29	tuym1	duam2	tien2	luyn1	kuon5	tieng5	thiuyng1	zuang5	kau2	ga:ul	gau5	kail	ka:is	kail
HN-32	tuym1	duam2	tien2	luyn1	kuon5	tieng5	thiuyng1	zuang5	kau2	ga:ul	ga:u5	kail	ka:is	kail
HN-22	tiam1	duam2	tien2	luon1	kuan5	tieng5	thieng1	zuang5	kao2	fa:ol	fao:5	kail	ka:is	kail
HN-25	tiam1	duym2	tien2	luon1	kuon5	tieng5	thieng1	zuang5	kao2	sa:ol	sao:5	kail	ka:is	kail
HN-31	tuym1	duam2	tien2	luyn1	kuon5	tieng5	thiuyng1	zuang5	kau2	sa:ul	sau5	kaeil	ka:is	kail
HN-12	tiym1	duom2	tien2	liyn1	kuon5	tieng5	thieng1	zuang5	kau2	ga:ul	gau5	kail	ka:is	kail
HU-03	tuym1	duom2	tien2	luyn1	kuon5	tieng5	thiuyng1	zuang5	kau2	ga:ul	gau5	kail	ka:is	kail
HN-21	tiam1	duom2	tien2	luon1	kuon5	tieng5	thieng1	zuang5	kao2	ga:ol	ga:ol	kail	ka:is	kail
HU-01	tuym1	duom2	tien2	luyn1	kuon5	tieng5	thiuyng1	zuang5	kau2	sa:ul	gau5	kaeil	ka:is	kail

Mapping the Yi characters – an attempt to explore a route of their historical change and origin –

Kazue Iwasa

Research Fellow of the Japan Society for the Promotion of Science,
Institute for Research in Humanities, Kyoto University

Abstract

This paper aims firstly at showing some intriguing outcomes as a result of a geolinguistic approach by mapping exemplary Yi characters used in nine vernacular areas, then secondly at considering a possible route of its origin and historical changes which would have happened upon the Yi script hitherto. Finally, current problems and future perspectives of this study will be mentioned.

1 Introduction

The Yi script has been used among the Yi ethnic group living in China, at least since the 15th century¹. Yiyu (彝语), the language this nationality speaks, belongs to the Lolo-Burmese language group of the Tibeto-Burman language family. According to the official classification in China, ‘Yiyu’ is divided into six dialects: Northern, Eastern, Southern, Southeastern, Western and Central. The four dialects of them are said to possess the Yi script and abundant manuscripts written by it.

The Yi script has been used among ‘Bimos’- shamans in the Yi communities -, and inherited from generation to generation within every Bimo’s clan. This sacred script seems to play an important role as a kind of mnemonic device when Bimos recite a scripture written in it during a religious ritual or ceremony². Since Yi characters are used exclusively among those Bimos and their usage tends to be extremely dependent upon an every single Bimo, consequently, Yi characters show huge diversity from a village to a village, or more precisely from a Bimo to a Bimo.

2 The Yi language and its script

Yi people live in the Southwest part of China, i.e., Sichuan, Guizhou, Yunnan and Guangxi Zhuang Autonomous Region, then, northern parts of Viet Nam and Laos. Except four dialects in China, neither script nor manuscript has been found in Viet Nam and Laos so far.

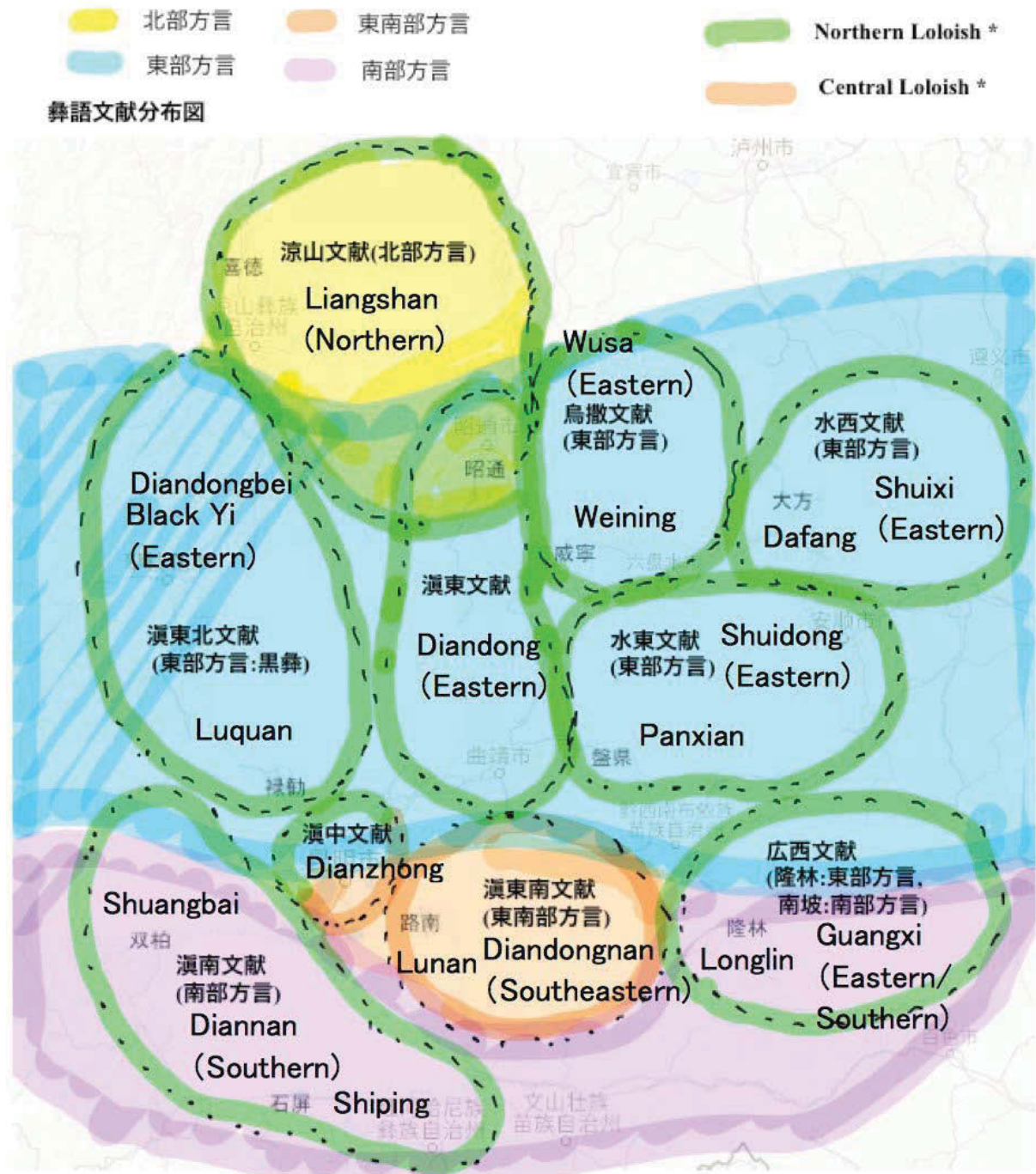
The dialects with and without the script in China are shown below:

With or without the script	Official classification in China	Classification by Bradley (2002)
with	Northern (Nosu)	Northern Loloish
	Eastern (Nasu)	
	Southern (Nisu)	
	Southeasterne (Sani, Azha)	
without	Westrn	Central Loloish
	Central	

¹ Shimizu 2014:3.

² Wasilewska 2012 and 2015.

Here is a map showing the areas with the Yi script and dialectal classification.



Yi characters are basically all syllabic. Their forms are relatively simpler than Chinese characters. It is said that Yi characters would have been created around Yuan or Ming dynasty, and they can also be classified by *Liushu* (六书), six methods of forming Chinese characters. In fact, one of them, *Jiajie* (假借), or phonetic loan is very common and happens so often that there is a high tendency since Yi characters to become syllabic from ideographic throughout all the areas³, for it seems that Bimos use a

³ The standardised Yi script, Guifan Yiwen (规范彝文), is completely syllabic.

homophonous⁴ character in copying manuscripts quite freely. There are some special circumstances for this tendency as stated below.

Bimos are religious leaders and at the same time intellectuals among Yi people. Hence, such a sacred script has been dealt with exclusively by their religious specialists. Consequently, the usage of Yi characters is extremely dependent upon Bimos.

Concerning some practical use, it is very plausible that these holy characters function as a kind of mnemonic sign, when Bimos recite sutras in a ceremony or a ritual. In other words, this writing system seems not to be a communication tool. Or, the Yi script is not created for being easily understood by everyone, but for handing such divine scriptures down to posterity⁵. More specifically, it is highly probable that these sacred characters should serve unequivocally as an effective and practical means for preventing leakage of a precious heritage to every Bimo's clan.

Due to the situation aforesaid, it is considerably difficult to decipher any Yi manuscript, even in cooperation with highly proficient and enthusiastic Bimos. Yi characters and manuscripts are then very personal to some extent. For example, a Bimo said that they would omit a certain character intentionally or even write a wrong word when copying texts, lest other Bimos should be able to sneak a look at their manuscripts. The reason why they can play such a trick is undoubtedly because they recite all the scriptures taught by their religious master.

3 The details and goals of mapping Yi characters

For mapping Yi characters, the data concerning the Northern, Eastern, Southern and Southeastern dialects are cited from *DianChuanQianGui Yihan Jiben Cihui Duizhao cidian* (《滇川黔桂彝汉基本词汇对照词典》), and those of the Southern dialect are also from three dictionaries: *Diannan Yiwen Zidian* (《滇南彝文字典》), *Guyiwen Changyong Zidian* (《古彝文常用字典》), *Yunnan Guifan Yiwen Yihan Cidian* (《云南规范彝文彝汉词典》)⁶.

Nine vernacular areas recorded on the maps are as follows:

Northern: Liangshan

Eastern: Diandongbei (Luquan), Shuixi (Dafang), Wusa (Weining), Shuidong (Panxian), Guangxi (Longlin)

Southeastern: Diandongnan (Lunan)

Southern: Diannan (Shuangbai and around Shiping)

On the maps, every Yi character is written with its phonetic value in IPA⁷. When there are allographs, they are also written aside.

This is the very first geolinguistic analysis of Yi characters. The goal of this new approach is to explore a possible route by which this script would have spread, and to clarify what kind of historical changes would have happened through its diffusion.

This time, all the data are based upon nouns. Totally, 516 words have been mapped and analysed.

⁴ Or, a character with very similar phonetic value is used.

⁵ It is said that Bimo apprentices copy by hand their master's text and recite it.

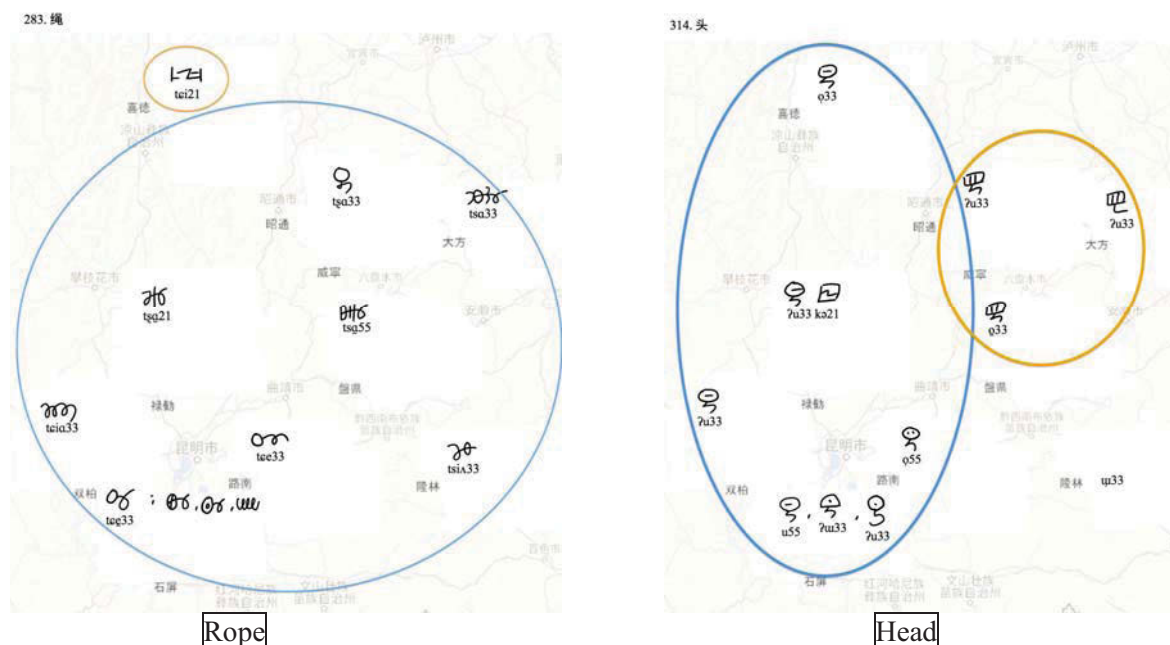
⁶ See References for other resources.

⁷ Certain phonetic descriptions in the original data have been changed, because several fonts are either unclearly or wrongly written, or again, due to the limit of my device, several particular fonts cannot be used in inputting on the maps. For example, original [ɿ] and [ʮ] are written as [i] throughout all the maps. Needless to say, any errors or inadequate expressions on this paper and maps are my sole responsibility.

4 Findings from the maps

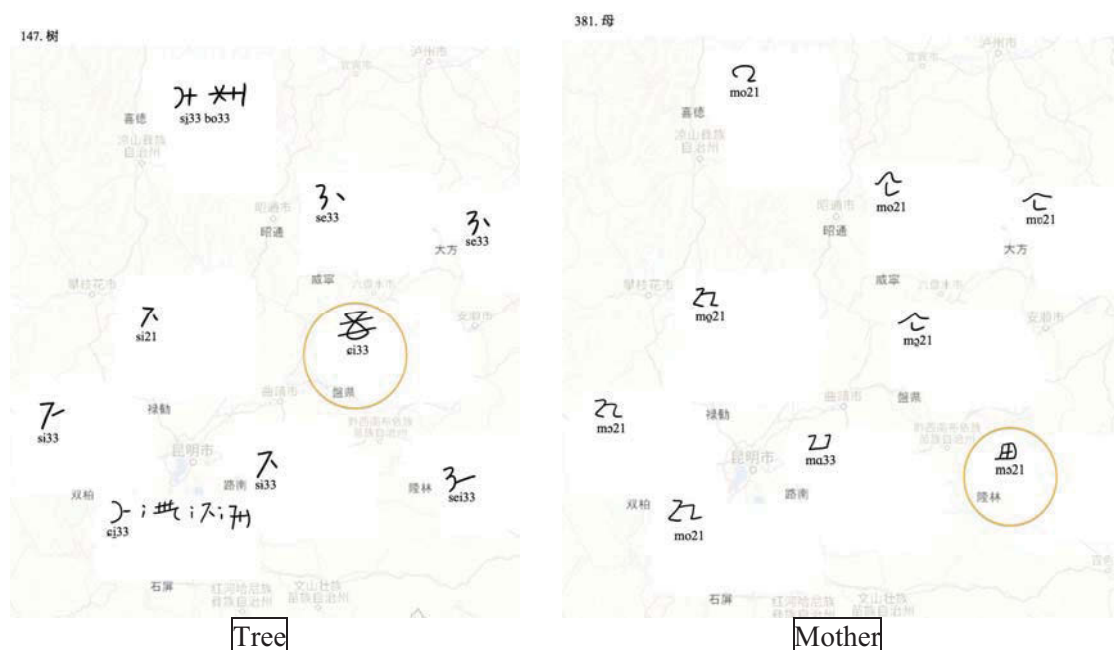
Examining all the maps gives us interesting findings such as AB distribution, phonetic loan (假借), and historical changes in phonetic/lexical aspects, in addition to intriguing cases where almost identical characters are found in north and south poles within the Yi-script-diffusion area.

4.1 AB distribution



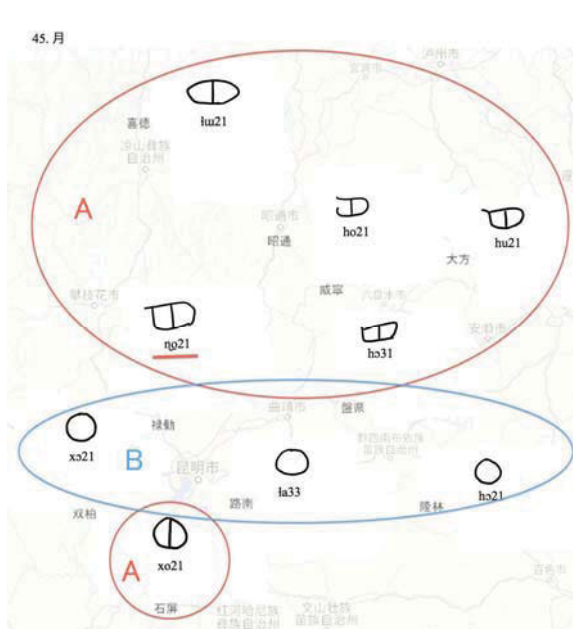
The maps illustrating AB distribution are given above.

4.2 Phonetic loan: Usage of another character with a homophone or similar phonetic property

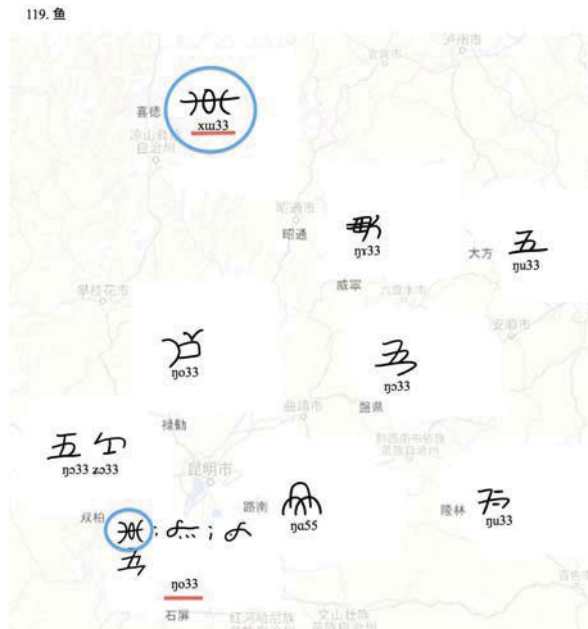


In the map left above, only the character in Panxian is different from others. It is highly probable that it should be a case of a phonetic loan. The map right above shows only in Longlin a different character is used for expressing ‘mother’, actually, it is utilised for negation in many other areas.

4.3 Lexical replacement: An identical character with the same meaning but different phonetic value



Month



Fish



Pear



Hemp, flax

Comparison with other areas in the map of ‘Month’, an initial consonant of Luquan (Diandongbei, Black Yi) is nasal /ŋ-/ , very different from the other eight areas where all words have initial fricatives. According to STEDT⁸, the etymon of the words with an initial fricative is PTB *s/g-la (MOON / MONTH), while that of Luquan is unclear.

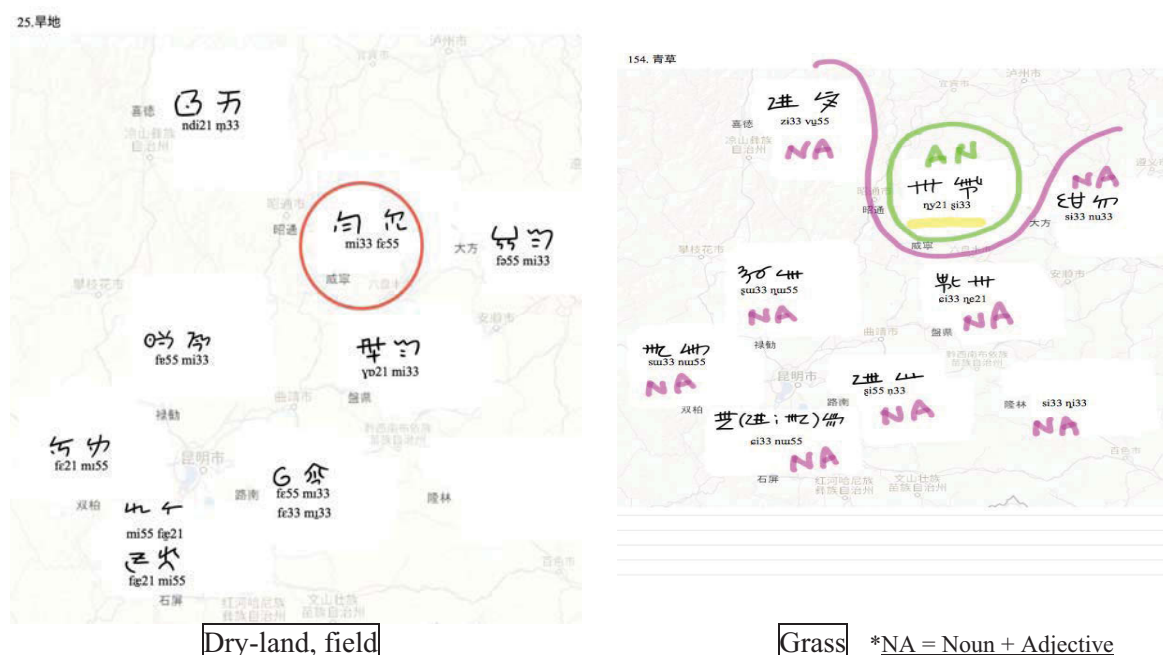
However, in spite of its phonetic difference, the character shape of Luquan still remains the same feature as others. It suggests that this character, or grapheme, is originally ideographic, and that the pronunciation or quite probably the word itself must have changed only in Luquan.

Besides, this map shows ABA distribution, where there is a belt between Shuangbai and Longlin from East to West.

In the map of ‘Fish’, an identical character is found both in Liangshan and Shiping, although its pronunciation is not the same. The etymon of [xu³³] in Liangshan is uncertain, on the other hand, that of the words in the other areas is PLB *s-ŋya (FISH)⁹. If their phonetic representations are taken into an account, a lexical change would have happened in Liangshan region, whereas the meaning that this character holds would have been consistently respected. That would be the reason why this character is still used for expressing ‘fish’ in Liangshan.

Two more interesting examples are also demonstrated for further reference.

4.4 Different word order



In Weining, Wusa area of the Eastern dialect, the word order of ‘dry-land’ is Noun+Adjective. In the Yi language, this is a basic word order of a noun phrase:

mi ³³	fe ⁵⁵
earth, land	dry

On the other hand, other areas such as Dafang, Luquan, Shuangbai, Shiping and Lunan, shows Adjective+Noun order. It is possible to say that this lexicon has already become a compound in these areas. It is noteworthy that both expressions seem to be possible in Shiping.

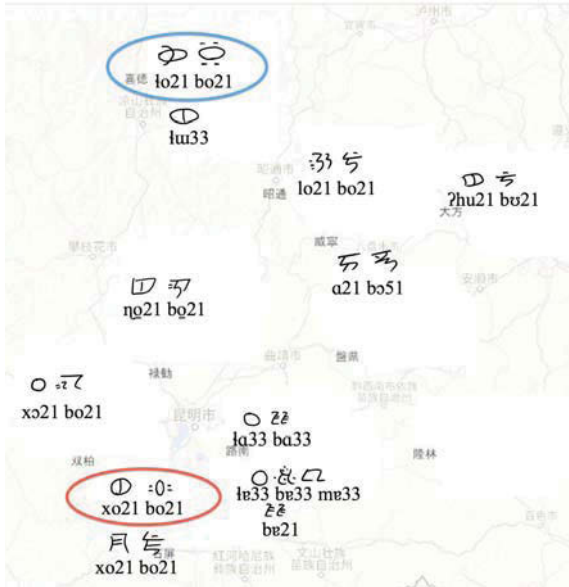
⁸ <http://stedt.berkeley.edu/~stedt/cgi/rootcanal.pl/gnis?t=month> (accessed on 14th July 2017.)

⁹ <http://stedt.berkeley.edu/~stedt/cgi/rootcanal.pl/gnis?t=fish> (accessed on 14th July 2017.)

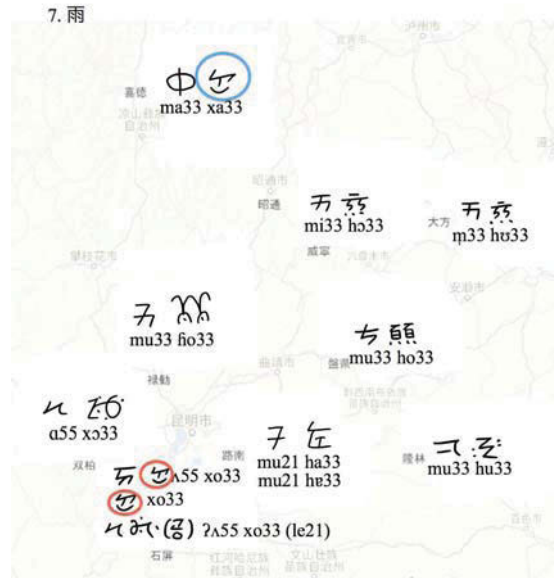
However, on the map of ‘Grass’, the situation is totally opposite. Further investigation is needed about this matter.

4.5 Identical characters found in north and south poles

3. 月亮



Moon



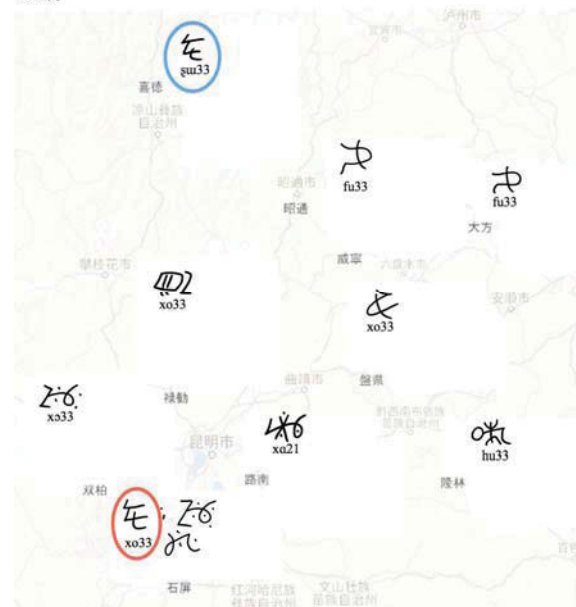
Rain

40. 年



Year

207. 肉



Meat

In all the maps above, the characters respectively of Liangshan in the northernmost part and Shiping in the southernmost point are almost identical, although there exists slight phonetic difference. The other characters found between these two regions are more or less different in their forms, in spite of their phonetic similarity¹⁰.

¹⁰ As for ‘Moon’, except the word of Luquan whose etymon is unclear, the etymon of the rest is PTB **s/g-la* (MOON / MONTH), PLB **s/ɿ-la³* (MOON / MONTH) and also PL **bə la³* (MOON). Therefore, the first syllables,

5 Conclusion

The Liangshan Yi tribe is said to be so conservative that Bimos there would not show their own sutras to others¹¹. It is unassumable that Yi characters of this area would be brought to the far enough, southernmost region, Shiping. Yet, there exist not a few examples where we can find identical Yi characters in the northernmost and southernmost poles of Yi-character-diffusion area, Liangshan and Shiping. This might suggest that these two areas should remain relatively older script type than the others to some extent, although due to the limit of the data at present work, very few maps demonstrate clear ABA distribution. However, at least, it seems to indicate the possibility that they should share the much nearer origin of the script than the rest.

According to Zhilujing (指路经), Scripture of showing the way¹², it is said that six ancestors of the Yi ethnic group started to migrate from Zhaotong (昭通) to Sichuan, Yunnan, Guizhou and Guangxi¹³. If this legend shows such a historical fact to some extent, each group settling down in Liangshan and Shiping might have protected its own script well as it is for a long time, after a journey from their homeland, Zhaotong.

At present, all the data have been cited from published materials. In the very near future, I hope to pick up data from original manuscripts and to cover as many places as possible on each vocabulary.

Finally, and perhaps most importantly, it is urgently needed to establish grapheme to each Yi character, without depending upon too much existing classification such as radicals or strokes. This is an inevitable step for analysing the various characters spread among the Yi people more precisely and accurately.

Acknowledgements

Here, I express my deep gratitude to Professor Mitsuaki Endo, who has introduced me the field of Geolinguistics and generously offered me many invaluable materials. If it had not been for his suggestion, I could not have dreamt to start this incredibly intriguing study.

References

- Bradley, David. 2002. The subgrouping of Tibeto-Burman. In Christopher I. Beckwith (ed.), *Medieval Tibeto-Burman languages: proceedings of a symposium held in Leiden, June 26, 2000, at the 9th Seminar of the International Association of Tibetan Studies*. Brill's Tibetan studies library. 1. Brill:73–112.
- Chen, Shilin et al. 陈士林他著. 1985. 『彝语简志』民族出版社.
- Ding, Chunshou. 丁椿寿. 1993. 『彝文论』. 四川民族出版社.
- Fan, Xiuli. 樊秀丽. 2001. 「中国彝族指路经における送霊・招魂/祖先移住経路言説に関する研究—民族アイデンティティ形成との関連について」. 国際協力研究誌.8-1. 広島大学大学院国際協力研究科編:67-90.

[xo²¹] in Shiping and [ɬo²¹] in Liangshan, are cognate.

In the case of ‘Rain’, the etymon is unclear, however, phonetic similarity is observed throughout all the areas.

As for the case of ‘Year’, the etymon of all the examples on the map is PTB *(k/?)uk (BEND / CROOKED / RETURN / BACK / YEAR).

About ‘Meat’, the etymon of all the data on the map is PTB *sya-n (FLESH / MEAT / GAME ANIMAL).

¹¹ Wasilewska 2015.

¹² Wasilewska 2015:237.

¹³ Fan 2001, also Guoji and Ling 1993.

- Guoji · Ningha, Ling Fuxiang. 果吉 · 宁哈 岭福祥. 1993. 『彝文《指路经》译集』. 中央民族学院出版社.
- Huang, Jianming. 黄建明. 1993. 『彝族古籍文献概要』. 云南民族出版社.
- Iwasa, Kazue. 2005. Mamuteyi –Lolo manuscript no.6(2) in the library of the Ecole Française d’Extrême-Orient, Paris. In Christopher I. Beckwith (ed.), *Medieval Tibeto-Burman Languages II: proceedings of the 10th Seminar of the IATS*. Brill:153-169.
- Iwasa, Kazue. 2008. Allegory of Mrs Xian-Zheng - A Preliminary Analysis and Decipherment of the Lolo Manuscript, Ms.CHI.YI.26 [Ms.CHI.80] in the Bibliothèque Interuniversitaire des Langues Orientales in Paris -. In Christopher I. Beckwith (ed.), *Medieval Tibeto-Burman Languages III: proceedings of the 11th Seminar of the IATS*:111-148.
- Kong, Xiangqing. 孔祥卿. 2005. 『彝文的源流』. 民族出版社.
- Shimizu, Toru. 清水享. 2014. 「近代から現代における彝族社会の変化と文化変容における総合的研究-涼山彝族を中心として-」 Nihon University, Ph.D. Dissertation.
- Wasilewska Halina. 2012. Orthography of Traditional Yi Writing. Conventions of Sign Use in the Ritual Scriptures of the Yi. In Nathan W. Hill (ed.), *Medieval Tibeto-Burman Languages IV*. Brill:447-480.
- Wasilewska Halina. 2015. Ritual manuscripts of the Yi in China – the case study of the Nuosu Yi. *Rocznik Orientalistyczny*, 2015:225-241.
- Zhu, Chongxian. 朱崇先. 1994. 『彝族典籍文化』. 中央民族大学出版社.
- Zhu, Chongxian. 朱崇先. 1996. 『彝族典籍文化研究』. 中央民族大学出版社.
- Dictionaries:
- 李增华主编. 2014. 『古彝文常用字典』. 云南民族出版社.
- 普梅丽 张辉 普梅笑 编著. 2014. 『云南规范彝文彝汉词典』. 云南民族出版社.
- 普璋开等编著. 2005. 『滇南彝文字典』. 云南民族出版社.
- 武自立 纪嘉发 编著. 2011. 『汉彝简明词典』. 四川民族出版社.
- 云南省路南彝族自治县文史研究室编. 1984. 『彝汉简明词典』. 云南民族出版社.
- 中央民院彝族历史文献班编. 1984. 『滇川黔桂彝汉基本词汇对照词典』. 中央民院彝族历史文献编译室.

How a Geolinguistic Approach can Contribute to a Study of Tibetic Languages: Examining the Cases of ‘Wind’ and ‘Rice’

Hiroyuki Suzuki^a

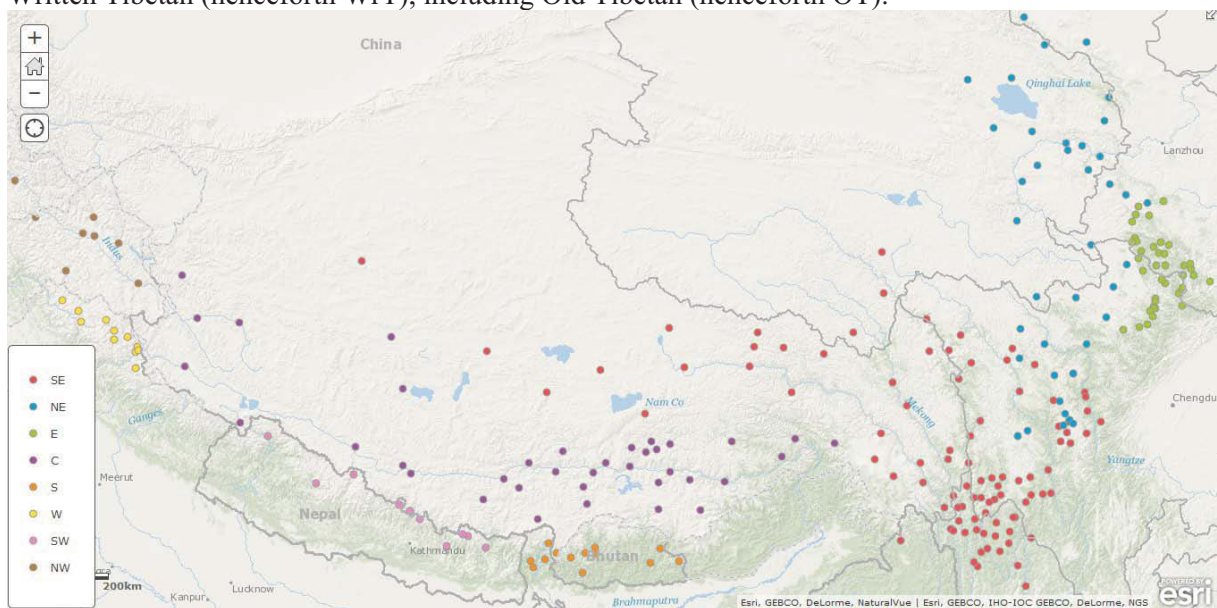
^aIKOS, University of Oslo / National Museum of Ethnology

Abstract

This article presents a brief introduction to a geolinguistic study regarding Tibetic languages, focusing on the eastern Tibetosphere (Yunnan, Sichuan, and Gansu). With around 230 dialects of the eastern Tibetic languages, it discusses basic issues on simple word forms shared over most varieties, and then two case studies on ‘wind’ and ‘rice’ are described.

1 Introduction

Tibetic languages are spoken across six countries including China, Pakistan, India, Nepal, Bhutan, and Myanmar (Fig. 1).¹ They are spoken in the largest area among the Tibeto-Burman languages. Despite this width of distribution and linguistic diversity (8 sections as a dialect continuum, see Tournadre 2014 and Tournadre & Suzuki forthcoming; Fig. 1), their lexical traits are fairly stable and homogeneous. According to Jin ed. (1983:144), more than 70 % of words have their lexical correspondence with Written Tibetan (henceforth WrT), including Old Tibetan (henceforth OT).²



Legend: SE=Southeastern; NE=Northeastern; E=Eastern; C=Central; S=Southern; W=Western; SW=Southwestern; NW=Northwestern.

Fig. 1: Distribution and diversity of Tibetic languages (based on Tournadre & Suzuki forthcoming)

¹ All the maps included in the article have been designed with ArcGIS online.

² One should note that the Written Tibetan forms, i.e., all the forms written in the Tibetan script, might not form a whole language. Some of them are of various local origins. It thus means that an existence of a given lexical form does not directly indicate a word used in a wider range of the Tibetosphere. Therefore, the high proportion of word forms corresponding to WrT does not mean that most Tibetic languages have *one* common origin.

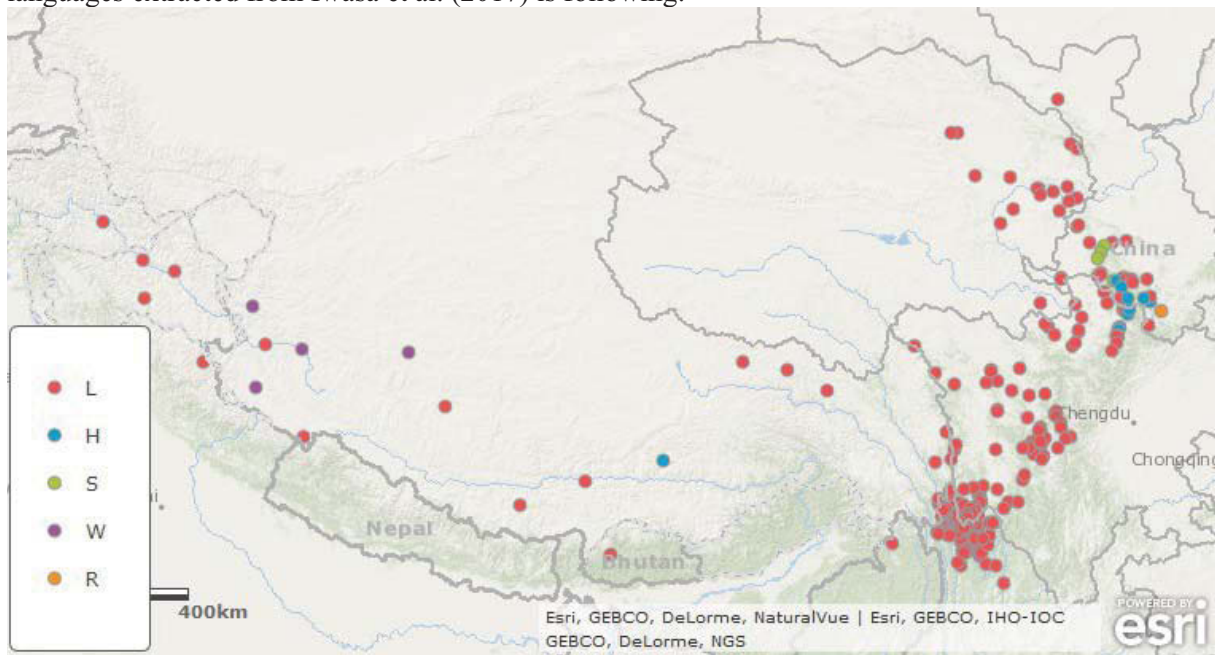
Contrary to the linguistic diversity, the lexical variation is less evident; principally, phonetic variation is outstanding all over the Tibetic languages. However, geolinguistic studies have just started for Tibetic languages even though the methodology itself has already been mentioned in sKal-bzang 'Gyur-med & sKal-bzang dByangs-can (2002) and Jiang (2002); hence it is an interesting issue how the linguistic variation within them appears based on a geolinguistic aspect.

This article examines two cases regarding the lexical feature: 'wind' and 'rice,' with a focus on the eastern Tibetosphere. These two words present different features from each other; therefore we can see how a geolinguistic approach can contribute to a study of Tibetic languages. They have been discussed in Suzuki (2017a) and Suzuki (2016a) respectively, and the description of the article includes a repetition.

The data used to create the linguistics maps of the whole Tibetosphere is the same as those used in the project *Studies in Asian Geolinguistics*, which are presented in Iwasa et al. (2017) and Suzuki et al. (2016). The counterpart of the maps of the eastern Tibetosphere only includes first-hand materials collected by the author from 2003 to 2017. Because of this, as well as because of time constraints on the part of the present author, the data points are not equally distributed within this area, and the points on the map only reflect the current research situation. The linguistic maps reflect so-called 'regiolects,' i.e., dialects with regional differences. Sociolects, which certainly exist in the given area, are not dealt with in this article.

2 'Wind'

A linguistic map representing a difference of word forms (WrT etyma if applicable) of the Tibetic languages extracted from Iwasa et al. (2017) is following:³



Legend: L= WrT *rlung*; H= WrT *lhags pa*; S= WrT *bser bu*; W= WrT *'ur*; R=others.

Fig. 2: Distribution of word forms for 'wind' of Tibetic languages.

The most widespread form is Type L, which is derived from WrT *rlung*. However, Figure 2 ignores the formation of different word forms, e.g., a form of the only root and a dissyllabic counterpart including a suffix, as well as phonetic discrepancy. Hence, there are to greater extent differences within Type L.

The semantic difference between *rlung* and *lhags pa* in the literary language is concerned with two aspects: semantic field and degree of strength of wind. The WrT word *rlung* also denotes 'air,' 'breath'

³ The map of Figure 2 was re-drawn and legend was also modified.

as well as ‘air element (one of the four elements in the world)’, and it means stronger wind than *lhags pa* ‘breeze’. However, it seems that there are only a few oral varieties which still maintain the distinction of meaning by using different lexical forms.

Looking at the eastern Tibetosphere, we can find that there are four roots among the five, and the variation clearly appears in the northeastern area. Of them, Type R is just attested in the Liping Jiuyuanzhai dialect recorded in Yang (1995) and does not appear in my own dataset. Additionally, I have not found its WrT etymon or any convincing etymology yet. Therefore, I will omit this in the following discussion.

Within my dataset from the eastern Tibetosphere, three principal types are attested: (A) WrT *rlung*-type; (B) WrT *lhags pa*-type; and (C) WrT *bser bu*-type. The last one mainly appears as /s^ha rə/, thus the phonetic form does not show a direct sound correspondence with WrT. In order to analyse the distribution of each type more in detail, I divide them into several subcategories. Examples are as follows:⁴

A-type (Type L in Figure 2)

A1a: monosyllabic form corresponding WrT *rlung*

[^hl̥ɿ̯], [^hl̥oŋ], [ɬoŋ], [^hl̥ä], [^wloŋ], [l̥ö], [^hl̥v̥], [^hj̥ɿ̯], [^hj̥ö], etc.

A1b: monosyllabic form including a voiceless lateral initial /l̥/

[l̥ɿ̯] etc.

A2a: disyllabic form (compound) corresponding to WrT *rlung dmar*

[^hl̥ɿ̯ mɛ:], [^wl̥ɿ̯ mɛ:], [j̥ö mɛ̃:], etc.

A2b: disyllabic form (compound) corresponding to WrT *rlung ma*

[^hl̥o mɛ], [^wl̥o ma], etc.

A2c: disyllabic form (compound) related to WrT *rlung dmar*

[^hl̥ɿ̯ peʔ], [^hl̥ɿ̯ m̥beʔ], [^hl̥ɔ: beʔ], [^hj̥ö pje], [^hj̥ö mjeʔ], etc.

A3: disyllabic form corresponding to WrT *rlung kha*

[^hl̥ö k^ha], [^hl̥oŋ k^ha], etc.

A4: other types

[^hl̥o wo], [^hl̥oŋ ^hdzə]

B-type: a form corresponding to WrT *lhags pa* (Type H in Figure 2)

[hɜ kə], [hə hə], [həʔ pa], [həʔ pa], [hə pa], etc.

C-type: probably related to WrT *bser bu* (Type S in Figure 2)

[s^ha rə], [s^hɛ lə:], etc.

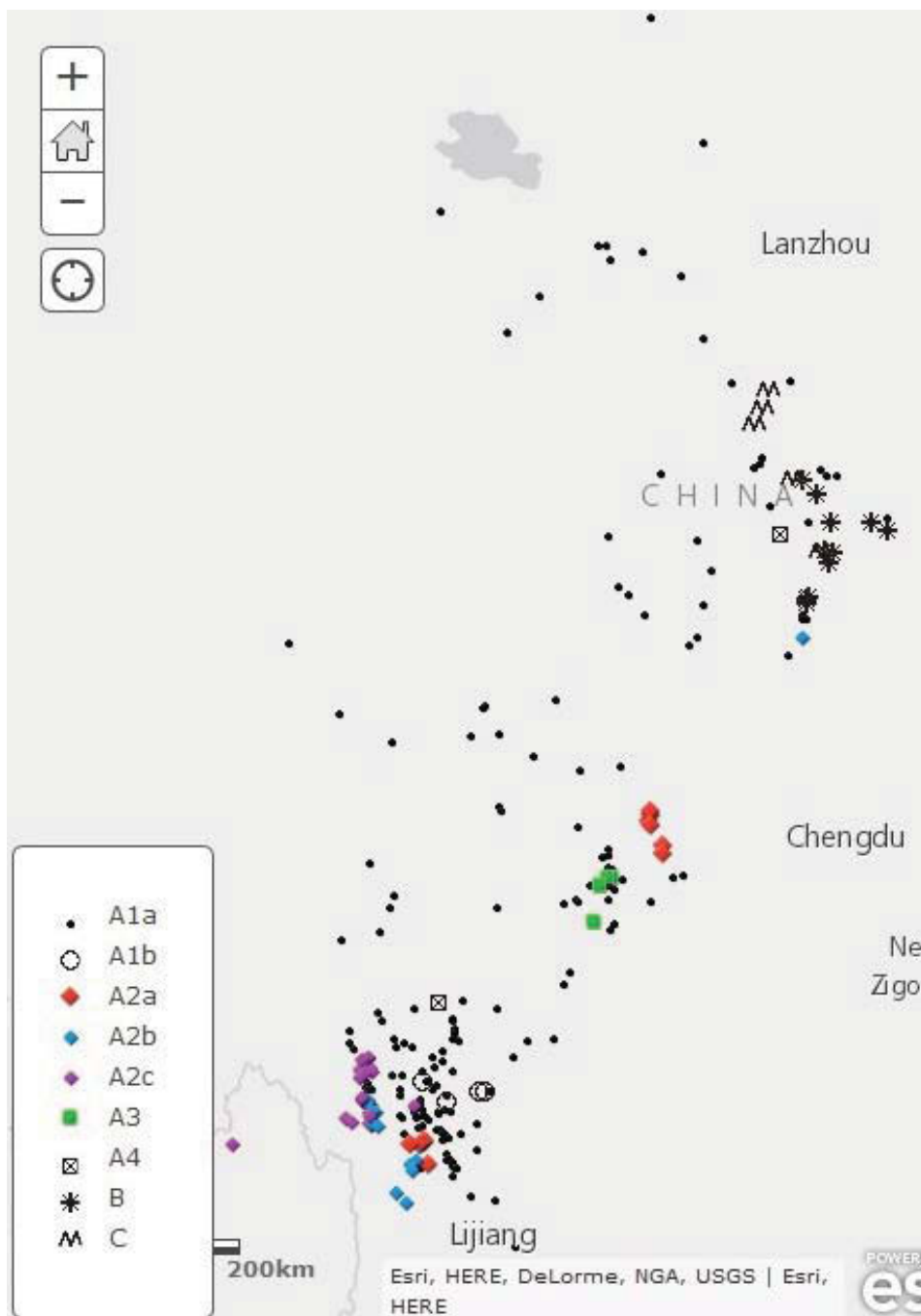
Note that the difference of initials (/l/ or /j/), depending on the whole system of the sound correspondence between spoken varieties and WrT. The chronological order should be: /l/ > /j/, however, this is not reflected in the classification above. The voiceless counterpart of the initial /l̥/ (A1b, a part of A4) might have appeared through another rule of sound change. WrT *rlung dmar* generally denotes ‘stormy wind’ (Zhang 1985:2738). It is a little complicated to distinguish a form corresponding to WrT *rlung dmar* from one corresponding to WrT *rlung ma*. For example, the dGudzong dialect (Rongbrag Khams) uses /^hl̥ä maʔ/, which is close to WrT *rlung dmar* because WrT *a* in an open syllable in this dialect generally corresponds to /o/.

Type B always appears as a form ‘root+suffix *pa*.’ This word formation is a difference feature from Type A, which can form a word by using the root itself.⁵

Figure 3 is an overall distribution of word forms for ‘wind’ in the eastern Tibetosphere based on the classification presented above.

⁴ A suprasegmental description is uniformly omitted. For a methodology of how to describe segmental sounds, see Zhu (2010) and Suzuki (2016c).

⁵ There are several dialects from Rongbrag which employ a form corresponding to WrT *lhags pa* for ‘frost’, not ‘wind.’ This use has already been attested in the 18th century, because it is recorded in *Muping Yiyu*, one of the nine texts known as Ding-series *Xifan Yiyu* (Suzuki 2007).



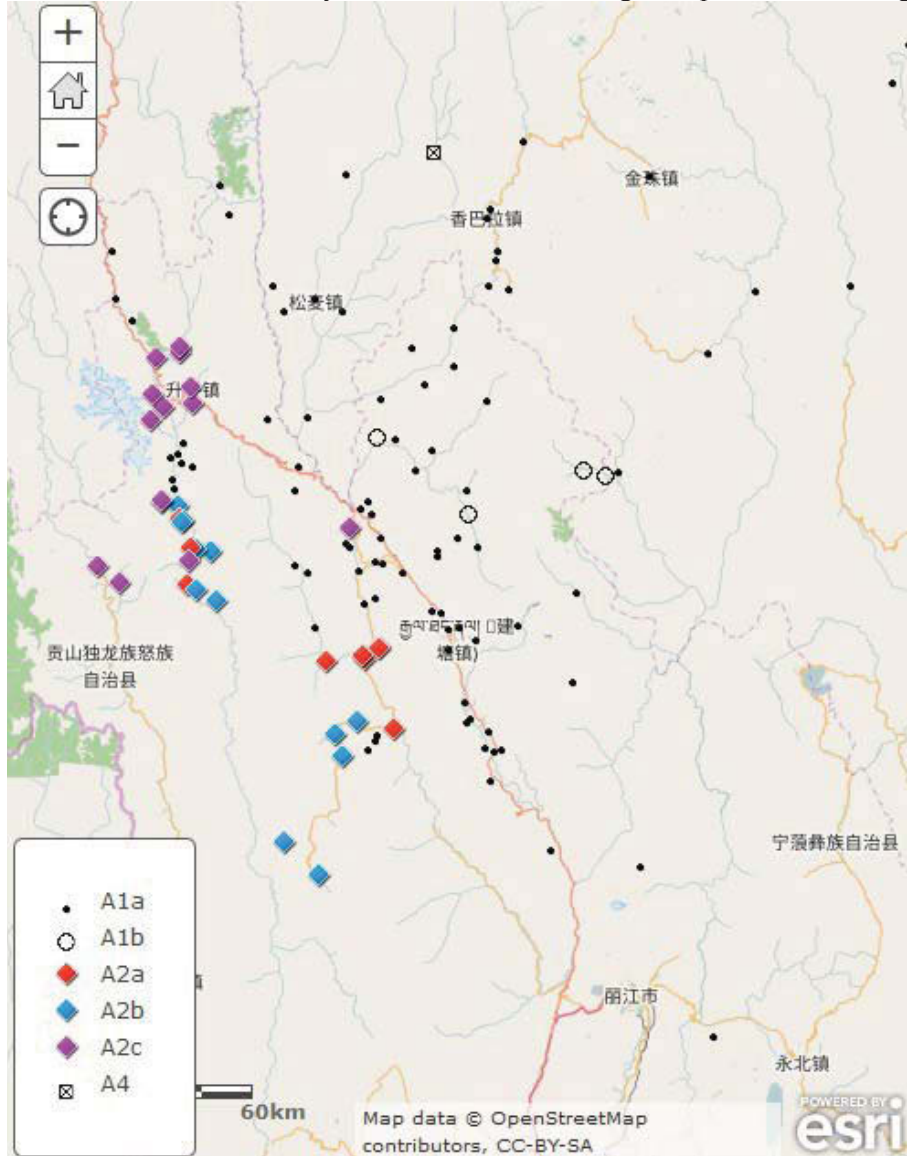
Legend: See the classification above.

Fig. 3: Overall distribution of word forms for ‘wind’ in the eastern Tibetsphere.

As Figure 3 displays, Type A (using a word including WrT *rlung* form) is widespread in the eastern Tibetsphere. The B and C types are both distributed in the north-eastern area of this region. They are used in the varieties which are linguistically classified as ‘Eastern Section’ (Tournadre 2014, Tournadre & Suzuki forthcoming). Speakers of these varieties are said to be descendants of immigrants from somewhere in Central Tibet in the period of Tibet Empire (Yang 2009:94-95; Sum-bha Don-grub Tshe-ring 2011:37-38).

According to the data of modern varieties spoken in Central Tibet (Figure 1), the use of Type B is attested even in Lhasa. However, Type B is registered in WrT and not regarded as a dialectal word, hence sharing the word form is not strong evidence to connect the varieties spoken in Central Tibet with those in Eastern Section. Additionally, several varieties in Eastern Section also use Type A. Their distribution is scattered; hence, the A type might not have been acquired by the influence of surrounding languages (mainly Amdo). In Literary Tibetan, in fact, both Type A (*rlung*) and Type B (*lhags pa*) are used, denoting ‘wind’ and ‘breeze’ respectively. Even at present, the coexistence of either the ‘A and B’ type or the ‘B and C’ type is attested in a few varieties. However, this difference is not reflected on the map.

Another interesting phenomenon is found in the southernmost area of this map: dialects spoken in the southern Khams, mainly in Yunnan Province. Figure 4 presents an enlarged map of this area.



Legend: See the classification above.

Fig. 4: Distribution of word forms for ‘wind’ in the southeastern Khams region.

Type A is divided into seven subgroups in total, based on word formation patterns (A1, A2, A3, A4) and phonetic realisations (A1a, A1b; A2a, A2b, A2c). Of the seven subcategories, A1a and A1b only consist of a word stem. The formation of A1b is irregular, just attested in the southern Khams area. The varieties using A1b are spoken in the border area of dialect groups such as Sems-kyi-Nyila, Chaphreng,

Muli-nDappa, and sDerong-nJol. A2a, A2b, and A2c are similar to each other regarding the second morpheme of the word. However, it is not certain whether A2c is genetically closer to A2a or A2b. The A2 form is originally related to two WrT forms, i.e., *rlung dmar* ‘strong wind, hurricane’ and *rlung ma* ‘wind.’ This observation means that A2 has two origins. However, because of the existence A2c, they are dealt with together. In Section 2, I mention that A2c is closer to A2a, but it is just an assumption. Since A2a and A2b have a WrT correspondence for each, it is probably correct to claim that the common form distributed in two or more places is not because of a shared innovation but because of coincidence. Particularly, it is the case that several places are geographically far from each other, e.g., several varieties of Rongbrag Khams and various varieties spoken in Yunnan for A2a.

More interestingly, one should note the distribution of A2a, A2b, and A2c in Yunnan. That of A2a is surrounded by A2b and A2c (see Figure 4). If this is considered as an example of the ABA-distribution, A2a is more recent form than A2b and A2c. However, paying attention to the distribution of A2c, we find that it is concentrated in the area to the north-west of A2a and A2b, it can be treated separately. Only one A2c form attested along the Jinshajiang River (mBukha dialect; Sems-kyi-nyila Khams) implies that this form originates from sDerong-nJol Khams spoken in the northwest to the region through where a main traffic road passes. If we accept this observation, the relationship between A2a and A2b will be a key question. Looking at the distribution along the Lancangjiang River, we can notice that A2a and A2b look like an ABA-distribution. Since A2a is situated in the centre, it might be a more recent form than surrounding A2b forms. If these word forms are not originally different and related to each other, the geographical distribution will mean that the form *rlung ma* has changed into *rlung dmar* because of confusion of the sound structure (see Suzuki 2011). People in younger generation might have forgotten the original form and have made an analogy regarding the second syllable, and begin to confuse one form with the other.

To the contrary, the dialects spoken along the Jinshajiang River are a quite different case. The dialects using A2a belong to the East Yunling Mountain subgroup, whereas those using A2b, to the Melung subgroup. The latter group is likely to have A2b originally based on its phonetic realisation, which suggests a lack of the final *r* in WrT. The former group is more sensitive to the pronunciation corresponding to the WrT final *r*, which is maintained as a consonantal feature, or omitted with influence to the preceding vowel. Therefore, this case can be analysed as an existence of two different word forms. A3 is mainly found in the Minyang Rabgang area, regardless of the languages. Some varieties of Minyang Rabgang Khams use A3, and some surrounding varieties of Amdo also use it. The expansion of A3 could have started from Minyang Rabgang Khams, which is regarded as a sedentary, more archaic variety in the local historical context (Sonam Wangmo 2013, Suzuki & Sonam Wangmo 2015). A4, including two exceptional forms [ʰlo wo] and [ʰloŋ ʰdzə], are attested in the Babzo dialect (dPal skyid Tibetan) and the Rwata dialect (Chaphreng Khams), respectively. The origin of these word forms is still unclear.

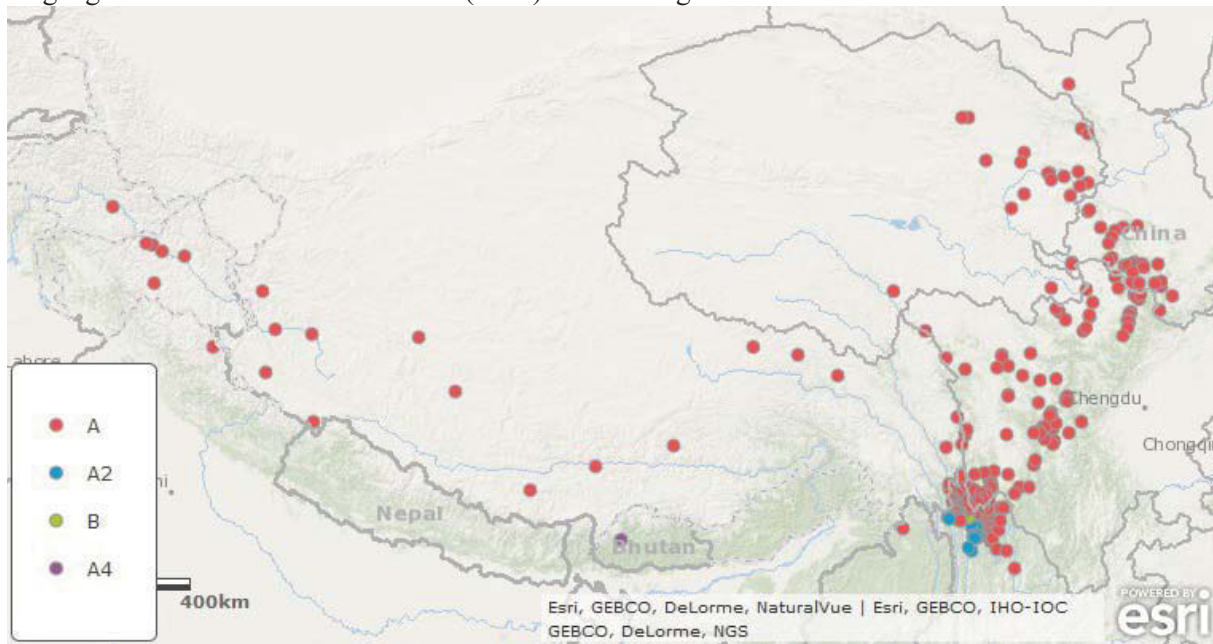
To conclude the discussion regarding the word for ‘wind,’ we can point out that a certain geographical range exists for lexical minorities as well as various word formations. This suggests that the development of word forms is related to a geographical continuum, and even though some forms are minor within the Tibetic languages, there can be reasons why they exist in a certain area. A geolinguistic approach can reveal how a lexical variation exists in these languages.

3 ‘Rice’

The English term ‘rice’ is a polysemic word for many Asian languages; hence we should specify which kind of ‘rice’ is discussed. Suzuki et al. (2016) examine how terms of ‘rice’ are diverse within the Tibeto-Burman language. However, since the majority of the Tibetosphere is out of the rice-cultivating area, many varieties of the Tibetic languages do not possess rich terms regarding the rice.

The majority of Tibetic languages use the form derived from WrT *’bras* for any forms of ‘rice’ (plant, grain, or cooked rice). Nevertheless, there are some minorities which display different patterns from the

majority. For this point, I present a linguistic map representing a difference of WrT etyma of the Tibetic languages extracted from Suzuki et al. (2016) is following:⁶



Legend: A=WrT *'bras*; A2=two roots (WrT *'bras* and *drus ma*); B=WrT *drus ma*; A4=four roots.
Fig. 5: Distribution of word forms for 'rice' of Tibetic languages.

As Figure 5 shows, most Tibetic languages certainly use the same WrT etymon *'bras*. The majority of the Tibetosphere does not belong to the rice-cultivating area and WrT *'bras* is not a part of basic words in the Tibetic languages. Suzuki & Sonam Wangmo (2016) assume that the word *'bras* was expanded everywhere in the Tibetosphere as a religious term.

However, several dialects spoken in the south-eastern area of the Tibetosphere, which corresponds to the north-western part of Yunnan Province as well as Bhutan, show a different situation. The latter has four terms for 'rice.' This classification of the terminology is common to Burmese and Newar (Suzuki et al. 2016). The former can be divided into two types as follows:

- (A) one semantic category for 'rice'
- (B) two semantic categories for 'rice'

Type A possesses only one single stem as in English, whereas Type B distinguishes 'rice grain (hulled, polished, and cooked)' from 'rice plant' or 'general species' name for rice' by differing stems.

Variation of phonetic realisations is, indeed, not crucial for a classification; however, we can note the regularity of sound correspondence with WrT from an irregular one. A partial discussion of the irregular phonetic form of WrT *'bras* 'rice' was provided in Suzuki (2012). The classification proposed in the article is as follows:

A-type

A-1: showing a regular sound correspondence of WrT *'bras*

[^hdʑe:], [^hdʑi:], [^hʃi:], [^hdʑe:], [^hbɛː:], [^hbɛːː:], [^hbɛːː:], etc.

A-2: showing an irregular sound correspondence of *'bras*

[^hdʑu ɦu], [^hguː:], [^hgiː:], etc.

A-3: correspondence of WrT *drus ma* 'polished grain'

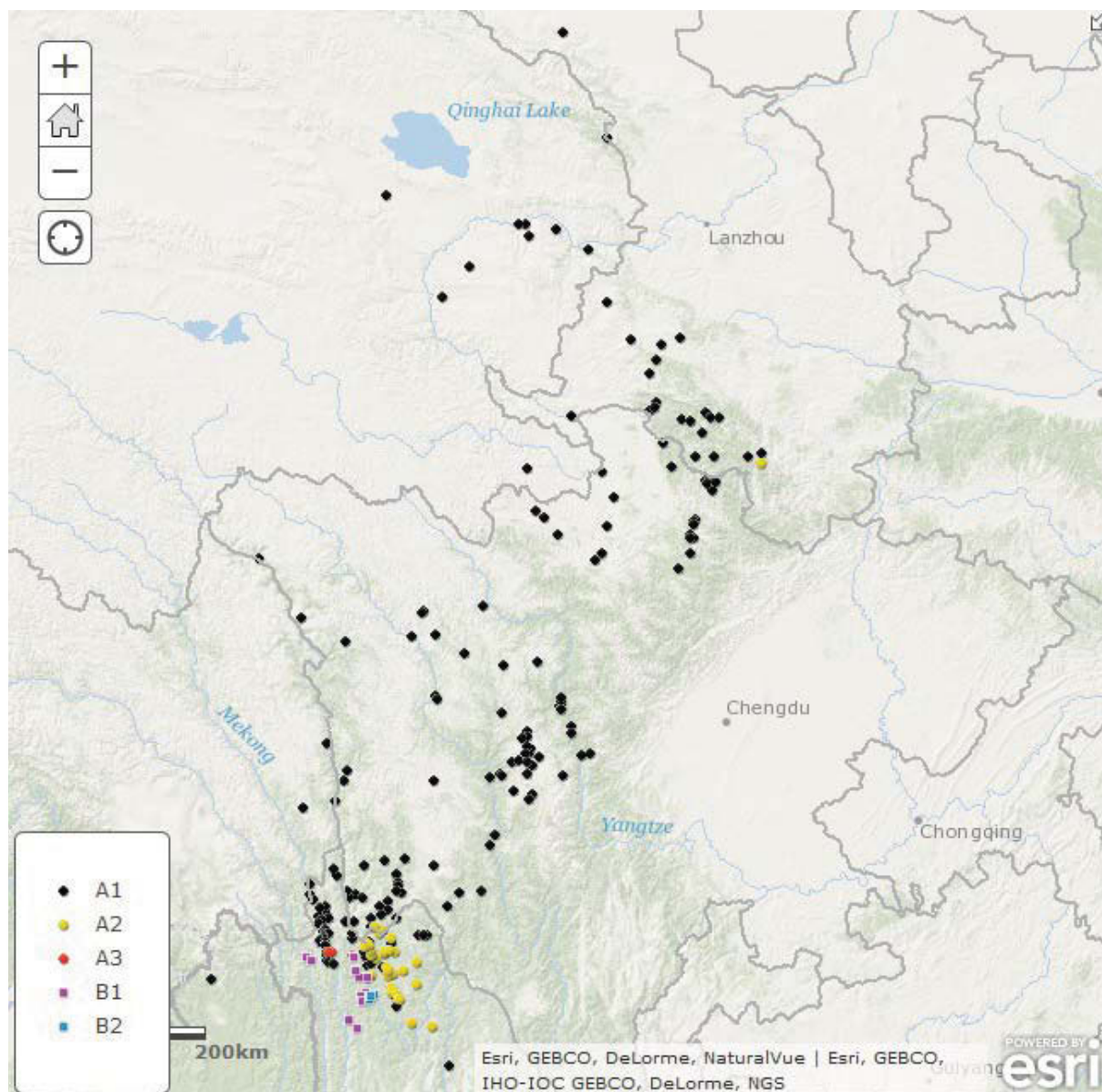
[tɛ ma], [tɛː ma]

B-type

⁶ Only the difference of etyma is distinguished. The map reflects neither an existence of affixes nor various phonetic realisations.

B-1: '*bras*' 'general name for rice' and *drus ma* 'rice grain' with a regular sound correspondence
 [ʰdze:] + [ʰtu: ma], [ʰdze:] + [ti: ma], [ʰdze:] + [təŋa], [ᵐbɛ:] + [tə mɛ], etc.

B-2: '*bras*' 'general name for rice' and *drus ma* 'rice grain' with an irregular sound correspondence
 [ʰge:] + [ʰtə: ma],

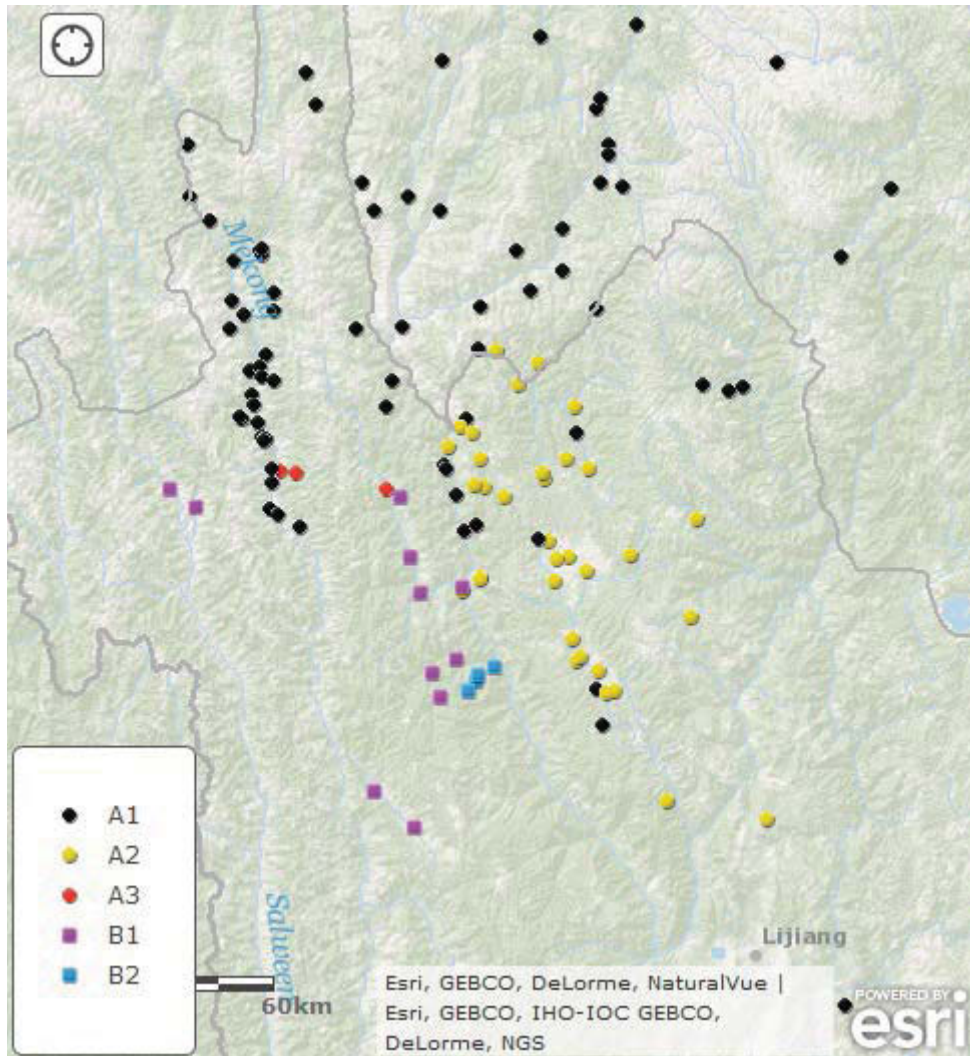


Legend: See the classification above.

Fig. 6: Overall distribution of word forms for 'rice' in the eastern Tibetsphere.

Figure 6 displays an overall distribution of the word forms for 'rice,' reflecting semantic differences as well as phonetic realisations, that is, the map distinguishes the classifications presented above from each other. It displays that the varieties using Type A are distributed in the majority of the eastern Tibetsphere with an evident exception from Yunnan, where those using Type B concentrate. The area of Type B belongs to a rice cultivation culture, and Tibetans there also plant rice. Therefore, the distribution of Type B is highly related to this cultural background, where a classification of 'rice plant' and 'rice grain' must have been needed. However, as mentioned in Suzuki et al. (2016), the rice does not grow in many parts of the Tibetsphere because the climate condition is inappropriate for rice-growing, but the varieties share the same root of this word. This implies that the rice is not a basic word but a

cultural one which can be related to the religious purpose. We can also note that the WrT form *'bras* corresponds to Proto-Tibeto-Burman (PTB) **b-ras* 'RICE / FRUIT / BEAR FRUIT / ROUND OBJECT' (STEDT⁷), and it is principally Tibetic languages that employ this PTB etymon for 'rice' among the Tibeto-Burman languages.



Legend: See the classification above.

Fig. 7: Distribution of word forms for 'rice' in the southeastern Kham region.⁸

Figure 7 is an enlarged version of the southeastern Kham area, where the word form for 'rice' is complicated in the eastern Tibetosphere. The minor groups of the classification above, which are A-2, A-3, B-1, and B-2, appear mainly in the rGyalthag dialect group spoken on the rGyalthag-Yangthang plain and the adjacent area of the Jinshajiang River. Some varieties spoken along the Lancangjiang River and the Nujiang River also have either the A3 or B1 type.

Firstly, it is certain that Type B appears in varieties spoken in a rice cultivation culture, including Wujing, Tuoding, and Xiaruo townships as well as Tacheng Town (belonging to the Jinshajiang drainage system), Yongchun and Pantianghe townships (belonging to the Lancangjiang drainage system),

⁷ <http://stedt.berkeley.edu/~stedt-cgi/rootcanal.pl/etymon/2071>, accessed on 14th April 2017.

⁸ Unfortunately, the map automatically generated by ArcGIS does not reflect the factual borderline dividing Yunnan Province from Sichuan Province. The actual line should be further to the north; on the map, Dongwang Township belongs to Sichuan, which should be within Yunnan.

and Bingzhongluo and Bangdang townships (alongside Nujinag). Note that the dialectal relationship among the varieties is not so close to every other because these varieties include the Sems-kyi-nyila and sDerong-nJol groups.

Secondly, we should also pay attention to the distribution of the A-3 type, with a single stem corresponding to WrT *drus ma*, which only appears in three varieties in a mountainous area which does not belong to a rice cultivation culture. Considering the geographical condition and genetic position of dialects, we can see that these varieties probably once had Type B system and lost the form corresponding to WrT *'bras* with a replacement of WrT *drus ma*. Following this, it is also noted that Type B is distributed in two different dialect groups as mentioned above. However, the lexical varieties for 'rice' imply that they might have had a mutual relationship. Suzuki (2014) mentions that the Bodgrong dialect (spoken along the Nujiang) is spoken by immigrants from some villages along the Lancangjiang, among which two villages, gYanggril and Tshodrug, are nominated as candidates based on the local tradition. The case of 'rice' suggests that speakers of the Bodgrong dialect might be related to those of Tshodrug, for the dialects with the A-3 type are spoken in the close area to it. Now the Tshodrug dialect does not maintain Type B and employs the A-type; however, it is possible that the elder generation of the speakers of the Tshodrug dialect used Type B. A more detailed discussion is provided in Suzuki (2017b).

Finally, we look at the A-2 and B-2 types, both of which are characterised by an irregular sound correspondence of WrT *'bras*. These types have a /g/ as the main initial, which is considered as an irregular form. However, the velar sound /g/ attested in the form for 'rice' has a close relation to /j/ and /dz/ as discussed in Suzuki (2015a, 2016b). Based on each phonetic form, /g/ must be related with /j/, not with /dz/. Taking the process of sound development discussed in Suzuki (2016b) into consideration, /j/ is the most conservative sound and /dz/ is the most innovative. The rGyalthag dialect, an example of the A-2 type, normally has a /dz/ initial for a WrT *'br* initial as seen in /ⁿdzɔʔ/ for WrT *'brug* 'dragon', while the form of 'rice' is /ⁿgu:/, which can be considered as an exception. Then, how did the rGyalthag dialect obtain this velar initial attested in 'rice'? Figure 7 suggests that the form for 'rice' with a /g/ initial might have spread from south to north in the rGyalthag-Yangthang plain. This route of expansion may be related to that of Naxi from the 15th to 18th centuries. According to Suzuki (2015b), the sound change regarding the WrT *r*-glide should have been influenced by Naxi after its intense contact began in the 15th century; thus the expansion of the word form for 'rice' might be related to Naxi's rule for the rGyalthag area at that period.⁹ In this case, 'rice' is not likely to be used for a kind of staple food but for a religious purpose, as rice cultivation is not practised on the rGyalthag-Yangthang plain. This explanation can also be applied in the case of the B-2 types attested along the Jinshajiang. The region once functioned as an 'entrance' from the Naxi cultural area to the Tibetosphere and has a religious site. Naxis and Tibetans still live together in this region.¹⁰

The word form of 'rice' in the Tibetic languages in the eastern Tibetosphere mainly corresponds to WrT *'bras*, and its geographical distribution is nearly pervasive. Most regions do not belong to the rice cultivation area; however, varieties have the same stem for rice. It is probably because the rice is used for religious rituals, whether they are of Bon or Buddhism. The 'rice' seems to be a kind of staple food, but in the case of Tibet, it can be for a religious purpose (Suzuki & Sonam Wangmo 2016).

In the Tibetosphere in Yunnan, however, a complicated system is attested. Several dialects spoken under the rice cultivation culture distinguish 'rice grain' from 'rice plant' by using different stems. The irregular sound correspondence of WrT *'bras* is also seen in Yunnan, which might be spread from the Naxi area to the north. The case of the Bodgrong dialect, spoken along the Nujiang, can be related to the varieties with Type B spoken along the Jinshajiang. Because Type B is attested in the limited range among the Tibetic languages, it is difficult to suppose that varieties with Type B developed independently in several places. The migration history of the Bodgrong Tibetans also indicates the origin where the varieties using Type B are spoken.

⁹ See Wang (1995) for a detail.

¹⁰ See Wu (2009) for a detail. However, the varieties that were influenced by Naxi the most belong to the Melung subgroup of the Sems-kyi-nyila group, and this fact appears in the Melung's systematic phonetic development. See also Suzuki (2013).

To conclude the discussion regarding the word for ‘rice,’ we can find some exceptions which can be considered from an areal viewpoint even though the word form is quite pervasive in the whole Tibetosphere. A development of lexical items is related to geographical, areal, and cultural features, thus necessity can produce new forms. The word form for ‘rice’ tells us an existence of the minor but important areal variation within the Tibetic languages.

4 Conclusion

This article presented two case studies on a geolinguistic analysis of ‘wind’ and ‘rice’ in the Tibetic languages. From a viewpoint of lexical forms, ‘wind’ has various roots, whereas ‘rice’ has a majority with minorities. The maps of the whole Tibetosphere (Figures 2 and 5) show the difference of lexical diversity, however, the situation in the eastern Tibetosphere is complicated. This observation is partially because of the difference of density of research points; we might find a lexical variation in other parts of the Tibetosphere following the progress of field research. However, it is also the fact that the formation of the eastern Tibetosphere is by far complicated because of intermittent contacts of languages as well as ethnic groups.

Acknowledgements

I should like to express my gratitude to my Tibetan friends who helped me and taught me their mother tongue. Field research was funded by four Grants-in-Aid for Scientific Research from the Japan Society for the Promotion of Science [JSPS]: “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. The article is also a part of the results of a Grant-in-Aid for Scientific Research from JSPS “Investigation of Undocumented Languages in the Eastern Tibetosphere and their Geolinguistic Research” (headed by the present author, No. 17H04774).

References

- Iwasa, K., H. Suzuki, K. Kurabe, S. Ebihara, S. Shirai, and I. Matsuse. 2017. Wind: Tibeto-Burman. *Studies in Asian Geolinguistics IV —Wind—*. (in press)
- Jiang, D. 2002. *Zangyu Yuyinshi Yanjiu*. Minzu Chubanshe.
- Jin, P. ed. 1983. *Zangyu Jianzhi*. Minzu Chubanshe.
- sKal-bzang 'Gyur-med & sKal-bzang dByangs-can. 2002. *Zangyu Fangyan Gailun*. Minzu Chubanshe.
- Sonam Wangmo. 2013. *Lhagang Monastery in Myth, History and Contemporary Society*. MPhil. Thesis, Universitetet i Oslo.
- Sum-bha Don-grub Tshe-ring. 2011. *Bod-kyi yul-skad rnam-bshad*. Zhongguo Zangxue Chubanshe.
- Suzuki, H. 2007. Qingdai Muping Tusi suoguan diqu de Zangyu fangyan. *Kangding Minzu Shifan Gaodeng Zhuanke Xuexiao Xuebao* 3:1-5 [Also found in: H. Suzuki, *Dongfang Zangqu Zhuyuyan Yanjiu*, 233-242, Sichuan Minzu Chubanshe, 2015].
- Suzuki, H. 2011. Zai yinbian guocheng zhong chansheng you xiaoshi de ruanehua yuanyin---Yunnan Deqin Yanmen xiang Guzha Zangyu zhi li. *Kyoto University Linguistic Research* 30:35-49.
- Suzuki, H. 2012. À propos du terme ‘riz’ et de l’hypothèse du groupe dialectal Sems-kyi-nyila en tibétain du Kham. *Revue d'études tibétaines* 23:107-115.
- Suzuki, H. 2013. Yunnan Weixi Zangyu de r jieyin yuyin yanbian --- Jiantan “erhua” yu “jinhou” zhi jiaocha guanxi. *Dongfang Yuyanxue* 13:20-35.

- Suzuki, H. 2014. Kamu-Tibettogo Bingzhongluo [Bodgrong] hoogen no onsei bunseki. *Asian and African Languages and Linguistics (AALL)* 9:153-193.
- Suzuki, H. 2015a. Jiantang Zangyu tuhua yanjiu de jige yiyi. *Yunnan Zangxue Yanjiu* 2:184-197.
- Suzuki, H. 2015b. *Lijiang Yongshengxian Daan Zangyu de laili chutan: tongguo Naxizu de jiechu ruhe yanbian*. Lecture given at Southwest University for Nationalities (Chengdu). [A revised version published in *Zangxue Xuekan* 17:250-263, 2016]
- Suzuki, H. 2016a. A geolinguistic description of terms for ‘rice’ in Tibetic languages of the eastern Tibetosphere. *Studies in Asian Geolinguistics II — Rice—*, 52-59. Online: https://publication.aa-ken.jp/sag2_rice_2016.pdf, accessed 14 March 2017.
- Suzuki, H. 2016b. /j/ ga kataru onhenkasi --- Khams-Tibettogo Xianggelila hoogengun ni okeru kookoogai keiretu onso ni tuite no oboegaki. *Journal of Kijutsuken, Descriptive Linguistics Study Group* 8:91-103. Online: <http://id.nii.ac.jp/1422/00000898/>, accessed 6 July 2017.
- Suzuki, H. 2016c. In defense of prepalatal non-fricative sounds and symbols: towards the Tibetan dialectology. *Researches in Asian Languages* 10:99-125.
- Suzuki, H. 2017a. Geolinguistic analysis of ‘wind’ in Tibetic languages in the eastern Tibetosphere. *Studies in Asian Geolinguistics IV — Wind—*. (in press)
- Suzuki, H. 2017b. Historical development of Bodgrong Tibetan (Gongshan, Yunnan): From a geolinguistic perspective. *Studies in Asian Geolinguistics VI — Means to count nouns—*. (in press)
- Suzuki, H. and Sonam Wangmo. 2015. Quelques remarques linguistiques sur le tibétain de Lhagang, «l’endroit préféré par le Bodhisattva». *Revue d’études tibétaines* 32:153-175. Online: http://himalaya.socanth.cam.ac.uk/collections/journals/ret/pdf/ret_32_05.pdf, accessed 14 April 2017.
- Suzuki, H. and Sonam Wangmo. 2016. Cultural contexts of the expansion of a Tibetan word ‘bras ‘rice’ in the easternmost Tibetosphere. In Mitsuaki Endo (ed.) *Papers from the Third International Conference on Asian Geolinguistics*, 72-79, Research Institute for Languages and Cultures of Asian and Africa. Online: https://publication.aa-ken.jp/papers_3IC_Aasian_geolinguistics_2016.pdf, accessed 6 July 2017.
- Suzuki, H., S. Shirai, K. Kurabe, K. Iwasa, S. Ebihara, and I. Matsuse. 2016. Rice plant: Tibeto-Burman. *Studies in Asian Geolinguistics II — Rice—*, 11-14. Online: https://publication.aa-ken.jp/sag2_rice_2016.pdf, accessed 14 March 2017.
- Tournadre, N. 2014. The Tibetic languages and their classification. In Th. Owen-Smith & N. W. Hill (eds.) *Trans-Himalayan Linguistics: Historical and Descriptive Linguistics of the Himalayan Area*, 105-129. Walter de Gruyter.
- Tournadre, N. and H. Suzuki. forthcoming. *The Tibetic Languages: An Introduction to the Family of Languages Derived from Old Tibetan* (with collaboration of Konchok Gyatsho and X. Becker)
- Wang, H. 1995. *Diqing Zangzu Shehuishi*. Zhongguo Zangxue Chubanshe.
- Wu, G. 2009. *Diqing Xianggelila Lüyou Fengwuzhi — Yanzhe Diming de Xiansuo*. Yunnan Renmin Chubanshe.
- Yang, Sh. 1995. *Yi he liang jiang liuyu Zangyu fangyan huiyao*. Gansu Minzu Chubanshe.
- Yang, Sh. 2009. *Anmuduo Dongbu Zangzu Lishi Wenhua Yanjiu*. Minzu Chubanshe.
- Zhang, Y. ed. 1985. *Bod rgya tshig mdzod chen mo*. Minzu Chubanshe.
- Zhu, X. 2010. *Yuyinxue*. Shangwu Yinshuguan.

Contributors

Mitsuaki ENDO, Aoyama Gakuin University

Chitsuko FUKUSHIMA, University of Niigata Prefecture

Mika FUKAZAWA, The Foundation for Research and Promotion of Ainu Culture

Satoko SHIRAI, JSPS / University of Tsukuba

Masaaki SHIMIZU, Osaka University

Mika KONDO, Joint Researcher, ILCAA, Tokyo University of Foreign Studies

Kazue IWASA, JSPS / Kyoto University

Hiroyuki SUZUKI, University of Oslo / National Museum of Ethnology

Acknowledgements

This volume is a part of the results of the project on “Studies in Asian Geolinguistics” 2015 - 2017 at the ILCAA, TUFS. The articles were partially supported by JSPS Grants-in-Aid for Scientific Research (KAKENHI) 15K02525, 16J01055, 17J40087, respectively.