

CLASSIFICATION OF THE SINO-TIBETAN LANGUAGES

ROBERT SHAFER

The present world-wide misconception that the Sino-Tibetan family of languages is divided into an eastern "Chinese-Siamese" sub-family and a western "Tibeto-Burman" sub-family¹ has resulted from two distinguished scholars working on opposite sides of the area and their almost total ignorance of the languages in their colleague's field.

Since about 300 Sino-Tibetan languages and dialects have been recorded, some division of labor has naturally occurred in their investigation. And this has followed geographical lines. In southeast Asia, great streams—the Mekong, Salween, Irrawaddy, and Chindwin—flow in a generally southerly direction, and the intermediate ridges, the monsoon forests, the Malay peninsula extending far southward, and the political division with Indo-China under France and Burma and India under Great Britain have hindered communication and have tended to compartmentize knowledge into one division facing toward the Pacific Ocean and another facing toward the Indian Ocean.

Henri Maspero, the last great scholar to hold to the "Siamese-Chinese" division, was a product of the *École Française d'Extrême-Orient* at Hanoi. He published brilliant work on both the Chinese

¹ Konow carelessly stated that Daic and Chinese "form one distinct family as compared with the Tibeto-Burman forms of speech" (*Linguistic Survey of India*, 3 [1909], p. 1). Of course he meant sub-family, but the absurdity of having families within a family of languages has been repeated in the *Encyclopedia Britannica* down to the last edition, which refers to the "Tibeto-Burman family" and the "Siamese-Chinese family" (1953, v. 22, p. 187, and v. 20, p. 596 a).

Ambrogio Ballini and Carlo Tagliavini correctly refer to sub-families but these are Konow's Sino-Siamese and Tibeto-Burman (*Enciclopedia Italiana*, v. 19, pp. 46 and 129ff. The *Dizionario Enciclopedico U. T. E. H. A.* mistakenly applies "tibeto-birmano" and "siamochino" to race. All have obviously been following Konow and the *Linguistic Survey of India*.

and Daic (Thai) languages. If one knows two languages one is bound to note some resemblances between them whether they are related or not.² Maspero noted a considerable number of parallels between Daic words and the corresponding Chinese forms—enough, he thought, to consider these languages to be closely related genetically. But Maspero knew practically nothing of the “Tibeto-Burman” languages which had been studied for some time primarily under the patronage of the British in India, and so Maspero naturally accepted “Tibeto-Burmic” as a sub-family on the authority of those who had been studying those languages.

And Sten Konow, a Norwegian scholar primarily interested in Iranian languages, was engaged by the British government in India to handle the non-Aryan languages for the *Linguistic Survey of India*. Konow was so fully occupied in making grammatical analyses from the specimens of text of the many non-Aryan languages and dialects that he had little time for anything else.³ Konow had very little to do with Chinese and Daic for the *Linguistic Survey*, and since the Sino-Daic numerals correspond so closely that even an amateur can see the resemblance, it was only natural for him to accept the “Chinese-Siamese” division for these languages that he knew so little about.

But Konow, the Iranist, brought to his work on the Sino-Tibetan languages the Indo-Europeanist's point of view of comparative grammar: that morphology is the primary criterion of genetic relationship. Now some Sino-Tibetan languages, like Chinese, have almost no morphology. So Konow substituted what seemed to him the closest thing to morphology—the word order—as a criterion.

Thus the “division,” actually made by scholars working at opposite sides of the field and ignorant of the languages at the opposite end, was rationalized by Konow as resting upon the position of words in the sentence; he stated⁴ that “the Tibeto-Burman family arranges the words of a sentence in the order of subject, object, verb, while the order in Chinese and Tai is subject, verb, object.”

This statement that the fundamental division in the Sino-Tibetan family rested almost altogether upon the shift in position of one

² Thus after 20 years the close resemblance of Old Japanese *wata* “sea” to English *water* still sticks in my memory, although I do not believe any genetic relationship exists between the languages.

³ Personal communication from Konow.

⁴ *Encyclopedia Britannica*, 11th ed. (1911), v. 26, p. 929.

part of speech—let us say that the object precedes the verb in one division and follows it in another—is rather surprising, coming from an Indo-Europeanist. For any English-speaking beginning student of German can describe how annoying the German “displacement” of the verb is. Yet would anyone contend that German and English should be placed in different divisions of the Indo-European family because of differences in the position of the verb in the sentence?

If Sino-Tibetan languages do not have anything resembling Indo-European morphology, naturally we cannot use morphological correspondences as a primary criterion of linguistic relationship. But an alternative solution is possible. Students of Chinese will be familiar with the concept of “empty words”, those which theoretically, at least, have lost their specific meaning and have come to be used as particles. Since the precise use of many of these “empty words” is still disputed even in languages studied for so long as Chinese, no special study of them in all the Sino-Tibetan languages has yet been made. But they are separate words and generally follow the same phonetic development as “full words,” and since some of them are found over widely scattered parts of the Sino-Tibetan area, this writer has included a considerable number of them in his work on the phonetic development of the Sino-Tibetan languages, where they at the same time offer evidence of a common Sino-Tibetan “morphology”—as nearly as the family can be said to have one. If any one group showed no, or very few, morphological elements (pronouns, interrogatives, etc.) in common with other Sino-Tibetan languages, we should be very skeptical of genetic relationship.

But Indo-European comparative grammar is not based on morphology alone. A great part of it consists in working out the phonetic equations for the various languages. Neither Konow nor Maspero had attempted to work out such equations and their comparisons consisted only of “look-alikes,” a good many of which were erroneous. The greater part of this writer’s work on Sino-Tibetan has consisted in working out such equations.⁵ Only common words, numerals, parts of the body, verbs, adjectives, morphological elements, etc., have been considered and comparisons showing pronounced semantic divergence have been rejected.

⁵ For bibliography, see Shafer, “East Himalayish,” *Bull. Sch. Or. Afr. St.* 15 (1953), 357 n., or “Newari and Sino-Tibetan,” *Studia Linguistica* (Lund, 1952), 92, n. 1, and 93, n. 3, par. 2 ff.

A rough tally of such comparisons so far published⁶ shows 216 between Bodish⁷ and Chinese, 191 between Bodish and Burmese, 122 between Chinese and Daic, 101 between Burmese and Chinese, 63 between Bodish and Daic, and 38 between Burmese and Daic. Of course this will not be the final figure, but there is no reason to believe that the relative figures will vary greatly from those given above.

If we take the first two figures—216 comparisons between Bodish and Chinese but 191 between Bodish and Burmese—it would indicate that Bodish is genetically closer to Chinese than it is to Burmese. To anyone not led by the exotic appearance of Chinese characters to regard the language as a thing apart, this conclusion should not come as a surprise in view of geography and history. For while Tibet is somewhat closer to Burma than to China, to go from Lhasa to Mandalay by the shortest route one would have to cross the Himalayas and the mountain range on the Indo-Burmese frontier and cross two rather formidable rivers, the Brahmaputra and Irrawaddy. Historically we know there has been contact between China and Tibet for many centuries, but I know of no such contact between Tibet and Burma.

We may also note that in contrast to the 216 comparisons between Chinese and Bodish, we find only 122 between Chinese and Daic. That is, Chinese is considerably closer to Bodish than it is to Daic. In fact, almost since I began work on the Sino-Tibetan languages some 20 years ago, I have held that if Daic is related to the (other?) Sino-Tibetan languages, the relationship is very distant.⁸ When Benedict cast doubt on the genetic relation-

⁶ "The Vocalism of Sino-Tibetan," *Journ. Amer. Or. Soc.* 60 (1940), 302-337; 61 (1941), 18-31; "Problems in Sino-Tibetan Phonetics," *JAOS* 64 (1944), 137-143; and "The Initials of Sino-Tibetan," *JAOS* 70 (1950), 96-103.

⁷ Bodish refers to Old Bodish (classical Tibetan) and languages closely related to it; for definition see below.

⁸ In 1938 I prepared a list of words showing the lack of precise phonetic and semantic correspondence between very common words in Daic and other (?) Sino-Tibetan languages, words such as those for parts of the body, celestial luminaries, pronouns, etc., and I tried to convince Maspero that Daic was not Sino-Tibetan. It was Maspero's insistence on Sino-Daic genetic relationship which caused me to work over Maspero's and K. Wulff's comparative data and to add my own, which nearly doubled the number of comparisons and filled in to some extent the gaps Wulff's work left in many series of finals. These Sino-Daic comparisons were presented in "The Vocalism of Sino-Tibetan" and subsequent articles with no statement regarding interrelationships.

I had discussed my skepticism regarding the relationship of Daic to Sino-Tibetan both before leaving for Europe, after talking with Maspero, and after publication of

ship of Daic to Sino-Tibetan, he was merely repeating and expanding my own unpublished view and, ironically enough, in part with my own unpublished materials.

Benedict dismissed the some 200 comparisons between Daic and "other" Sino-Tibetan languages which had been adduced by Maspero, Wulff, and myself as non-basic, while he held that his own 30 Daic-Indonesian comparisons (including such cultural words as those for "rice-field" and "door") were basic.

More recently André G. Haudricourt has noted, regarding the Daic languages, that "les mots de la langue commune incontestablement proches de mots chinois sont les noms de nombres, des techniques militaires (cheval, selle, éléphant, jouget) et des techniques artisanales (métier à tisser, ouvrier, papier), bref un vocabulaire de civilisation susceptible d'emprunt. Au contraire le nom des parties du corps et le vocabulaire agricole ont peu d'affinité avec le vocabulaire chinois correspondant."⁹ Haudricourt thought that the relationship of the Daic languages was to be sought in the north in the Man-Yao languages, and in the south in the Annamese-Muong. Since Haudricourt is working in these latter fields it is to be hoped that he will bring forward Daic comparisons with these groups of languages. For it is only by searching for lexical and morphological parallels on all sides and by establishing the phonetic equations for such parallels that we can finally decide the genetic relationship of a doubtful group such as Daic.

While we cannot dismiss Haudricourt's suggestion of Chinese loan words in Daic without investigation, such precise phonetic correspondences as Burmese *k'ye'*, Siamese and Lao *k'i'* "dung", or Lao *hnū'*, Luṣei *hnu-* 'breast',¹⁰ for example, cannot by any semantic juggling be considered words of civilization likely to be borrowed. And too many of the comparisons are of that type to dismiss them off-hand. So in consideration of the very substantial number of comparisons between Daic and "other"

my article, so that there was no basis for Benedict's statement that I followed Maspero and Wulff in setting up an Eastern Division composed of Chinese and Daic in opposition to Tibeto-Burmic (Benedict, "Thai, Kadai, and Indonesian," *American Anthropologist* 44 (1942), 588).

⁹ "Les phonèmes et le vocabulaire du thai commun," *Journal Asiatique* (1948), 235 ff.

¹⁰ Cited in "Vocalism" (see n. 4 above), Table 4, no. 18, and Table 6, no. 26 respectively. Many comparisons of a similar nature will be found in the essays referred to above.

Sino-Tibetan languages I shall continue to present Daic as Sino-Tibetan—if only as a challenge to accept such evidence as conclusive or to produce contrary evidence.

And while we have seen above that on a statistical basis Daic is closer to Chinese than to any other Sino-Tibetan group, we have also seen Chinese is not closest to Daic but to Bodish. So we should abandon the Sino-Daic division in contrast to the Tibeto-Burmic division. Rather we should set up the following main divisions of Sino-Tibetan: Sinitic (Chinese), Daic, Bodic, Burmic, Baric, and Karenic.¹¹ The languages composing each division, as far as known, will be given below. Altogether we have some published material on about 300 languages and dialects of this family. Most of them are unknown even to scholars and it seems imperative to give the reader some clue to the relative position of groups in the descending scale of importance by adopting a uniform system of nomenclature; thus the descending order of the groups within a family were designated as follows:

Family (ending *-an*, as Sino-Tibetan); division (ending *-ic*, as Sinitic); section (ending *-ish*, as Bodish); branch (no specific ending); unit (no specific ending). Moreover, some of the languages are known under various names—the name they give themselves and the names other peoples give them, or under various spellings, as the Siamese spelling Dai, which is today pronounced Thai in Siam, for example, but Tai in some other languages of the division. It has been the custom for many years on the railroad maps of Europe to place the names in the language of the country, and scholars could do well to accord the same courtesy to each people, which will at the same time give the scholar a standard. This means abandoning the names given a people by its neighbors, changing Aka to Hruso, Miri to Mišing, Digaro to Taying, etc. I have also taken as standard the oldest spelling of the name of a people, usually the written rather than the spoken word, when it is known; thus the Bodish dialect Lhoskad rather than the spoken Hloke, and Dbus rather than Ü.

¹¹ I use Sino-Daic, Tibeto-Baric, Tibeto-Burmic always to refer to two or more divisions.

SINO-TIBETAN FAMILY

Sinitic Division [China]

MANDARIN SECTION¹² [N and W China]

WU SECTION [E coast of China] : Shanghai, Ningpo, Wenchow.
S. ANHWEI.

MIN SECTION [SE coast of China] : Foochow, Amoy, Swatow.
HUNAN.

YUE SECTION [S coastal] : Canton, Kiangsi, Hakka.

Daic Division¹³ [W China, Tonkin, Kgd. of Laos, Siam, Shan St. of Burma].

SOUTH : Siamese, Lao.

NORTH : Tai noir, Tai blanc, Tho, Dioi, Ahom.

NORTHWEST : Shan, Khamti.

Bodic Division.

BODISH SECTION.

*Bodish Branch*¹⁴.

West Bodish Unit [Kashmir] : Sbalti (Balti) ; Burig (Purik) ; Ladwags (Ladakhi) : Sam, Leh, Rong ;¹⁵ Lahul.

Central Bodish Unit [S Tibet, N. India, Nepal, Bhutan] : Lhoskad, (Lhoke), Šarpa, Kagate, Garhwal, Spiti, Mnyamskad (Nyamkat), Džad, Gtsang, Dbus (Ü) : Lhasa, Sikkim, Åba (Batang), Chonj, Tseku, Dartsemdo (Tatsienlu), Nganshuenkuan, Sotati-po, Paurong, Dru, Panakha, Panags, Nyarong, Ngamdo (Amdo), Khams¹⁶.

¹² For some of the recorded Mandarin dialects see Bernhard Karlgren, "Études sur la phonologie chinoise," *Archives d'Études Orientales* 15 (1915), pp. 230-1. The classification of the Chinese dialects given here is Karlgren's but with certain modifications by Yuen Renn Chao. For a dialect map of China see *Shun Pao*, 60th anniversary edition.

¹³ Daic languages show little divergence except in phonetic development, which Maspero used in his classification, "Contribution à l'étude du système phonétique des langues thai," *Bull. Éc. Fr. Ext.-Or.* 11 (1911), 158, n. 1. Tai Noir, according to Maspero, and Ahom, in my opinion, are intermediate. A good part of the vocabularies of many of the Hainan languages and dialects is also probably Daic.

¹⁴ Progressive phonetic degeneration of Bodish dialects from west to east through the dialect of Nganshuenkuan, after which archaic aspects increase through Khams.

¹⁵ Not to be confused with the Rong which is geographically in the Himalayas and linguistically in the Northern Naga Branch of the Kukish Section (see below).

¹⁶ The Hanniu of von Rosthorn also belongs somewhere in the Central Bodish Unit ; it is not a Rgyarong dialect.

South Bodish Unit¹⁷ [Sikkim].

Groma (Tromowa) : Upper, Lower.

Sikkimese¹⁸

Dandzongka.

East Bodish Unit : Dwags (Takpa).

Tsangla Branch.

Rgyarong Branch.

Rgyarong : Pati, Wassu.

Gurung Branch [C Nepal] : Gurung, Murmi, Thaksya.

WEST HIMALAYISH SECTION [scattered, mostly W Himalaya].

NNW Branch.

Bunan.

Thebor : Sumtsú, Zangram, Sungham, Kanam, Lippa.

NW Branch.

Kanauri : Upper, Lower, Tsítkhuli, Tukpa, Kanaši.

Mantšati, Tšamba Lahuli, Rangloi.

Almora Branch.

Rangkas, Darmiya.

Tšaudangsi, Byangsi.

Džanggali Branch.

*Eastern Branch*¹⁹ : Thami, Bhamu.

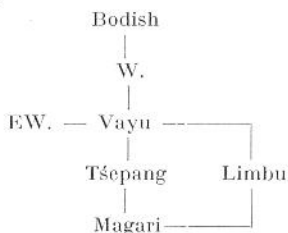
WEST CENTRAL HIMALAYISH SECTION²⁰ [C Nepal] : Vayu, Tšepang, Magari.

¹⁷ Characterized by the shift of *-r-* to *-y-*, also a characteristic of most of Burmish.

¹⁸ Not to be confused with Sikkim in Central Bodish.

¹⁹ Thami and Bhamu, which had been separated by Konow, belong together. From the limited vocabularies of them one can only say that they are placed in West Himalayish because they appear to be closer to that group than any other.

²⁰ A poorly defined section of which it can only be said that the languages are related to each other more than to languages of other groups; see "Classification of Some Languages of the Himalayas," *J. Bihar Res. Soc.* 36 (1950), 192ff. Interrelationships are approximately :



W. is West Himalayish, and EW. is the Western Branch of East Himalayish.

EAST HIMALAYISH SECTION²¹ [E. Nepal].*Western Branch.*Bahing, Sunwari²², Thulung, Tšaurasya.Dumi Unit : ²³ Dumi, Khaling, Rai.*Eastern Branch*²⁴.Khambu Unit : Khambu,²⁵ Natšhereng.Bontawa Unit : Rodong; Waling : Rungtšhenbung, Kiranti, Dungmali; Lambitšhong;²⁶ Lohorong,²⁷ Limbu,²⁸ Yakha.Not definitely classified in a division; probably sections of Bodic, possibly of Burmic, certainly not of Baric : ²⁹

NEWARISH [C Nepal] : Newari, Pahri.

DIGARISH [NE Assam and into Tibet] : Taying (Digaro), Midu.

MIDŽUIŠH [in Tibet beyond NE Assam].

HRUSHISH [N Assam].

Hruso (Aka) : Dialect A,³⁰ Dialect B³¹.

DHIMALISH [Darjeeling and Jalpaiguri] : Dhimal, Toŋo.

MIŠINGISH [NE Assam and into Tibet].

Mišing (Miri), Abor.

Yano.

Nyising (Dafla) : E,³² C,³³ Tagen.DZORGAISH (DZORGAIC?)³⁴ [NE Tibetan plateau and Szechuan, Kansu] : Dzorgai, Kortse, "Outer Mantse," Pingfang.

²¹ Konow mixed in some languages that do not belong here and omitted some that do. See my "East Himalayish," *Bull. Sch. Or. Afr. St.* 15 (1953), 356 f.

²² Phonetically degenerate compared to Bahing.

²³ Diverges considerably.

²⁴ Approaches Burmish phonetically.

²⁵ Kulung and Sangpang are the same language recorded in different localities.

²⁶ Tšhingtang is the same language in another locality.

²⁷ Balali is the same language in a different locality.

²⁸ Limbu and Yakha diverge somewhat from other East Himalayish languages.

²⁹ None of these groups seems to possess features distinctive from Bodic or Burmic and it is a question of classifying them under one or the other. This may have to be done on a weighted statistical basis, similar in principle, perhaps, to that proposed by A. L. Kroeber and C. D. Chrétien in "Quantitative Classification of Indo-European Languages," *Language* 13 (1937), 83-103, but weighted to allow for the criticisms made by A. Meillet.

³⁰ By Campbell. Comparatively archaic. See my "Hruso," *Bull. Sch. Or. Afr. St.* 12 (1947), 184-196.

³¹ By the other three recorders of Hruso. Phonetically degenerate.

³² By Needham, Hamilton.

³³ By Robinson.

³⁴ This group occupies at least a portion of the Sifan plateau. It probably forms the most northeastern outpost of the Tibeto-Burmic peoples today. The vocabularies

Burmie Division.

BURMISH SECTION.

Burma Branch [Burma].

Southern Unit : Burmese, Arakanese, Tavoy, Taungyo, Intha, Danu, Yaw.

Northern Unit : Phun (Megyaw, Samong dialects), Atsang (Ngatsang, Maingtha), Lawng (Maru), Letsi (Lashi), Tsaiwa (Atsi, Szi).

Lolo Branch [N Burma, Tonkin, SW China].

Southern Unit :³⁵

Phunoi : Phunoi, Pyen, Khaskhong, Hwethom. Akha : Akha, Ako, Asong, Phana, Menghwa, Woni, Lahu, Lahuna, Lahushi, Kui.

Central Unit.

Lisu : Lisu, Lishaw, Lipha, Liphoh, Kesopho, Kosopho.

Nyi, Tsökö, Weining, Ahi, Lolopho.

Northern Unit : Thongho, Pakishan, Kangsiangying, Kiaokio, Nee, Ulu, Laichau, Tudza, Nuoku.

Tonkin Unit : Mung ;³⁶ White, Black, Khoany.

Unclassified :³⁷ Manyak (Menia) ; Mossö, Dion ; Duampu ; Phupha ; Nameji.

Hor Branch [E Tibet?] : Hor (Horpa)³⁸.

*Hsihsia (Sihia) Branch*³⁹.

I have seen have been short and poorly recorded. Lexically these languages seem to be predominantly Tibeto-Burmie but with some peculiar features. In recent years, Chinese scholars—notably Wên Yu in *Studia Serica*—have given some attention to these languages under the name of Ch'iang, assuming that they are spoken by a people mentioned in old documents. But only one bit of the recent literature on the subject has come to my attention. Perhaps some parts of the vocabulary of Sotati-po are Dzorgaish ; see Central Bodish Unit above.

³⁵ The Northern Unit of the Burmese Branch approaches phonetically and sometimes lexically the Southern Unit of the Lolo Branch ; see Shafer, "Phonétique historique des langues lolo," *T'oung Pao* 41 (1952), 191-229. In fact the transition may be said to be gradual in the loss of final consonants from the Burmese Branch to the Northern Unit of the Lolo Branch, while inversely the loss of initial sonancy is progressive from the Northern Unit of the Lolo Branch to the Burmese Branch. But the languages of the Burmish Section are remarkably uniform lexically considering the vast stretch of territory they cover.

³⁶ Perhaps partly non-Lolo.

³⁷ Too little data or too irregularly recorded.

³⁸ Preserves some prefixes lost in the Burmese and Lolo Branches ; somewhat degenerate regarding finals.

³⁹ Uncertainty regarding transcription and limited vocabulary make definite classification impossible at present, but Hsihsia appears to be the most northern extension of the Burmish section.

MRUISH SECTION⁴⁰ [Arakan Hills].

NUNGISH SECTION⁴¹ [N Burma] : Rawang, Metu, Melam, Tamalu, Tukiumu.

KATŠINISH SECTION [N Burma] : Katšin, Khauri ; Džili.⁴²

TŠAIRELISH SECTION⁴³ [SE Assam].

LUISH SECTION [SE Assam and adjoining Burma districts] : Andro, Sengmai, Sak, Kadu.

TAMAN.⁴⁴

KUKISH SECTION⁴⁵ [Indo-Burmese frontier regions].

⁴⁰ See my article on "The Linguistic Position of Mru", *Journ. Burma Res. Soc.* 31 (1941), pt. 2, no. 2.

⁴¹ Closer genetically to the Burmish than to the Kukish Section.

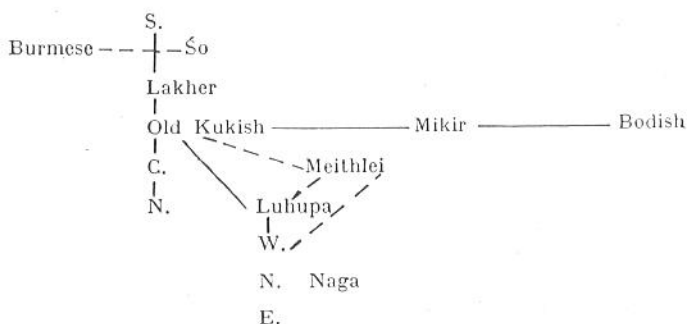
⁴² More archaic in its prefixes. Only a extremely short vocabulary has been published.

⁴³ Records of the Luish languages vary so in extent and the ability of the recorder that it is difficult to establish interrelationships within the section. The finals of Sak are considerably altered compared to those of other Luish languages, and this is perhaps not altogether due to poor recording. It preserves medial *l as r, this phoneme being lost in the rest of the section. Certain Sak forms common to Kukish, such as those for "mother" and "bird," and not found in the other Luish languages, may be borrowings from Kukish.

⁴⁴ The Taman recorded by R. Grant Brown, *JRAI* 41 (1911), 305 f., is archaic in prefixes but the phonetics of Taman is not generally clear from the brief vocabulary and probably will not be until we have sufficient data to be able to eliminate loan words. We may not be able to classify it until we have larger vocabularies of it and its dialects and of the surrounding languages. It appears to have been under Burmish influence but has a number of rather rare stems found in Luish, but not exclusively there, as words for "buffalo," "elephant," "horse," "salt," and "father."

⁴⁵ One of the major points on which Konow's classification errs is in setting up a Naga group. The Naga languages are all Kukish except the northeastern-most, which is Baric. The proof of the first part of this statement was contained in my article on "The Naga Branches of Kukish," *Rocznik Orientalistyczny*¹⁶ (Krakow, 1950), 467-530, and of the last part of the statement in "Classification of the Northernmost Naga Languages," *J. Bihar Res. Soc.* 39 (1953), 225-264.

The interrelationships within Kukish are approximately as follows :



Southern Branch.

Śo : Sandoway, Thayetmyo, Minbu, Tśinbon, Chittagong,⁴⁶
Lemyo.⁴⁶

Yawdwin ; Tśinbok.

Khami : S., Khimi,⁴⁷ N.

Lakher Branch.

Mara, Tlongsai,⁴⁸ Hawthai.

Sabeu.⁴⁹

Zeuhngang.

Śandu.⁵⁰

Old Kuki Branch.

Central Unit : Tśiru, Aimol, Purum, Langrong.⁵¹

Kyau.⁵²

Western Unit : Hrangkhoh, Biate, Hallam ; S Luhupa.⁵³

Kolhreng (C. Peripheral) Unit : Kolhreng, Kom, Tarao.⁵⁴

Solid lines represent genetic relationship, broken lines borrowing. Burmese loan words in Śo, and Meithlei loan words in Old Kuki, Luhupa, and West Kukish, are largely due to administrative dominance of Burmese and Meithlei.

Old Kukish, taken collectively, appears to very nearly represent proto-Kukish, while phonetic degeneration, particularly in finals, has occurred to the south in Lakher and S. Kukish, and particularly in prefixes in the north in Central and Northern Kukish. But all the groups along the vertical axis are essentially the same except for varying degrees of decay and the extent of borrowing. The number of stems peculiar to individual branches along the main vertical axis are almost negligible. So I term these languages Central Core languages or Kukish proper. Kukish languages to the right of the vertical axis have decayed and diverged semantically and morphologically.

⁴⁶ Slightly better preserved than most recordings of Śo.

⁴⁷ A slight link with N. Khami. Data on Khimi will be found in "Khimi Grammar and Vocabulary," *Bull. Sch. Or. Afr. St.* 11 (1944), 386-434.

⁴⁸ Although Parry thought he was recording the same dialect as the Mara of Savidge, there are a few phonetic differences.

⁴⁹ The Lakher vocabulary of Lewin approaches the Sabeu of Parry but does not correspond exactly. It is probably somewhat intermediate between Sabeu and some of the other dialects recorded by Parry.

⁵⁰ Slightly more archaic than the other Lakher recordings, as it was taken down earlier.

⁵¹ Precise classification here is doubtful.

⁵² Belongs here ? Eliminating borrowings from Southern Kukish, Kyau is probably close to the Western Unit of the Old Kuki Branch.

⁵³ Provisional classification. Certainly not a Luhupa language, although it has borrowed a few words from that group, as those for "horse" and "cow," and has undergone some vocalic changes—particularly to *u and *ui—which are not characteristic of Old Kuki.

⁵⁴ Precise classification doubtful. Probably transitional to Lamgang Unit, but in its main characteristics belongs with the Kolhreng Unit.

Lamgang (Southern) Unit : Lamgang, Anal.⁵⁵

*Langel.*⁵⁶

Central Branch.

Luśei Unit : Luśei (Dulien dialect, Ngente dialect), Zahao Hmar, Pankhu,⁵⁷ Bom.

Haka Unit : Haka (Lai), Śonśe, Taungtha,⁵⁸ Bandžogi.

Kapwi Unit.⁵⁹

Northern Branch.

Thado, Ralte;⁶⁰ Śiyang (Siyin), Vuite (Paite).

*Luhupa Branch.*⁶¹

Maring Unit :⁶² Maring, Khoibu.

Tangkhu Unit : Ukhrul, Phadang; Tśamphung.⁶³

Kupome Unit : Kupome, Khunggoi, Central Luhupa, Northern Luhupa.

*Western Branch.*⁶⁴

Maram, Kwoireng; Kabui, Khoirao; Empeo.⁶⁵

Northern Naga Branch.

Hlota;⁶⁶ Ao Unit : Tśungli; Longla; Mongsen, Tśangki, Khari; Tengsa, Rong (Leptśa)⁶⁷ (Sikkim), Yatśam; Yatśumi,⁶⁸ Thukumi.

⁵⁵ A slightly degenerate form of Lamgang; see the author's "Phonetik der Alt-Kuki-Mundarten," *ZDMG* 102 (1952), 262-279.

⁵⁶ A "Central Core" Kukish language whose precise classification is doubtful. It is spoken in southern territory but probably by migrants from the Central Branch of Kukish.*

⁵⁷ Has some special links with Bandžogi, but probably mainly in the vocabularies of Lewin, because both languages are more archaic there than in later recordings.

⁵⁸ Spoken in Southern Kukish territory but belongs essentially to the Haka Unit of Central Kukish. This is not clear from a cursory examination, as Taungtha has borrowed some words from Southern Kukish languages and has developed to some extent along individual lines.

⁵⁹ Precise classification here questionable. But not to be confused with Kabui, as Konow did.

⁶⁰ Has a few minor links with Vuite.

⁶¹ Probably genetically closest to Old Kuki, but not its Western Unit.

⁶² Diverges toward Kukish proper more than the two following Luhupa units, as already noted by Konow, p. 451.

⁶³ Sharply divergent.

⁶⁴ The subordination of Kwoireng and Khoirao noted below is not marked.

⁶⁵ Butler's Empeo is a different dialect from that of Soppitt and Stewart.

⁶⁶ Kyō or Kyontsū would perhaps be a better designation.

⁶⁷ Rong is slightly closer to Tengsa than Yatśam, but the latter are closer to each other.

⁶⁸ Perhaps better designated Yimtsurr.

*Eastern Branch.*Rengma.⁶⁹

Simi (Sema) : Kežama, Sopvoma ; Zumomi, Dayang.

Angami : Tengima ; Tšakrima : Dzuna, Kehena, Mima.

*Meihlei Branch.*⁷⁰*Mikir Branch.*⁷¹**Baric Division** (Assam).

BARISH SECTION.

North Central Branch.

Garo : Ačsik (Standard), Kamrup, Abeng, Dacca.

*Jalpaiguri Branch.**South Central Branch* : Atong, Rabha, Ruga, Kontš, Tintikeya, Cooch Behar, Kotš.*Western Branch* : Bodo, Metš ; Dimasa, Hodžai ; Tipura ; Lalung ; Moran.*Eastern Branch* : Tšutiya.

NAGISH SECTION.

Mošang, Šangge.

Namsangia.

Banpara, Mutonia.

Tšingmegnu (Tamlu).

Angwanku (Tableng), Mulung.

Tšang.

Karenic Division⁷² [C and Southerly Burma].

Yeinbaw ; Karenni : Yintale, Manö, Sinhmaw Mapauk ;

⁶⁹ Perhaps better Nzong or Nzonyu. Grierson's Rengma is slightly more altered than Butler's. Ntenyi, of the northern group of Western "Rengma," actually forms a connecting link between Hlota and Simi, while the Iseni-Kotsenu of Hutton is intermediate between Ntenyi and Anyo (Eastern or Naked "Rengma"). These meagerly recorded languages cannot be more definitely placed at present.

⁷⁰ In accord with the standard outlined above of using the oldest form of the native name known, I use *Meihlei* instead of the usual *Meithei*, the present pronunciation. For the Assamese call these people *Mēklē* (Konow, p. 20), which is about as close to 'Meihlei as the Assamese could make (the Kukish lateral affricates are frequently written *kl*, *k'l*). Singh (p. 71) recorded the Thado name of the Manipuris as *Mailhai*, i. e., *Mei-hlei*; and 'l' is usually recorded as *hl* in Thado. And since 'l' became *l'* in the language of Manipur, we must conclude that the earlier name was 'Meihlei, which became the present *Meithei* by regular phonetic change.

⁷¹ Greatly aberrant, but definitely Kukish.

⁷² Lexically predominantly Burmic, but also strongly Bodic, although it is one of the most southern of the Sino-Tibetan groups. It might almost be called intermediate

Pwo : Bassein, Maulmein ; Mopwa : Dermuha, Bilitsi ; Taungthu ;⁷³

Zayein : Sawntung, Padang, Banyang ; Kawnsawng ; Gheko, Bwe, Sgaw, Wewaw, Padaung, Karenbyu.

CONCLUSION

Sten Konow was the first scholar to attempt to cover almost the entire Sino-Tibetan field. But his routine duties on the *Linguistic Survey of India* and his grammatical analyses made from texts permitted him to make lexical comparisons almost exclusively on modern Bodish dialects, and he did not attempt to work out phonetic equations in a serious way. Thus he did not have a broad basis of comparative work for his classification, but he appears instead to have taken into consideration geographical location, grammatical similarities, and a few random lexical comparisons made by other authors or observed by himself but not published.

Konow's classification was sometimes remarkably good when correct alignment did not involve a knowledge of phonetic shifts, a knowledge Konow did not possess. Thus his classification of the West Himalayish languages is excellent, and all his Kuki-Chin languages are indeed Kukish. It was when severe phonetic shifts occurred so that parallels were disguised, as in the Naga languages, or when there were no closely related languages apparent, that Konow was most likely to go astray.

Of the Naga languages Konow confused the West Kukish and Luhupa branches and misclassified Mikir. He included Empeo, Kabui, Khoirao, Kapwi (which he confused with Kabui),⁷⁴ and Mikir in what he called the "Nāgā-Bodo group, bridging over the difference between the characteristic features of the two forms of speech"⁷⁵ while "Mikir clearly belongs to the same group as Kachchā Nāgā,⁷⁶ Kabui, and Khoirao." But the position of

between Bodic and Burmic, but certain phonetic and lexical peculiarities make it preferable to regard it for the present as a separate division. Due to lack of any sharp division of the Karenic languages and the inadequate materials on some of them, the following classification is only tentative.

⁷³ Not to be confused with Taungtha in Central Kukish.

⁷⁴ P. 418.

⁷⁵ P. 379.

⁷⁶ I. e., Empeo.

Mikir could not have been as clear as he thought, for later he transferred Mikir from his Nāgā-Bodo to his Nāgā-Kuki group,⁷⁷ though still recognizing⁷⁸ Mikir "affinities with Bodo," he considered it "much more closely connected with Kuki," yet that "it occupies a somewhat independent position."

Actually Mikir occupies the most independent position of any of the Kukish languages, both lexically and phonetically. And instead of being a connecting link with Barish, as one might expect from its geographical position and as Konow thought, it is a slight connecting link with the more distant Bodish. And whether Konow included or omitted Mikir, his Nāgā-Bodo group had no more connection with Barish (his Bodo) than have any of the other Kukish languages.

Konow's transfer of Mikir to his Nāgā-Kuki group only aggravated the latter's heterogeneity, for the group already included members of three different branches : Sopvoma (really Eastern Kukish),⁷⁹ Maram and Kwoireng (really Western Kukish), and Maring, Tangkhul (Ukhrul), Phadang, and Khunggoi (really Luhupa), while the addition of Mikir brought in a fourth branch. With such diversity of languages in one supposed group, Konow's statement⁸⁰ that the Nāgā-Kuki bridges over the gulf between Angami and the Kuki languages is meaningless, for languages from three or four groups would naturally link almost any distantly related languages.

As Rong has borrowed perhaps the greater part of its vocabulary from Bodish, with some elements from surrounding Himalayish languages, and as it is separated by some 400 miles from the Ao languages to which it belongs, its separation from the latter in the *Linguistic Survey* is understandable. Yet one may classify Rong rather precisely as belonging not only to the Ao Unit of the Northern Naga Branch, but as belonging to the Tengsa subunit, and within this subunit as being slightly closer to Tengsa than to

⁷⁷ Vol. 1, pt. 1, p. 66, n. 2.

⁷⁸ P. 69.

⁷⁹ But Konow observed (p. 451) that "Sopvomā is so closely connected with all the languages of the Western sub-group, that it might with equal propriety be classed as belonging to it as to the Nāgā-Kuki one" and in following pages all Konow's comparisons are between Sopvoma and Angami. One may suspect that Konow discovered Sopvoma's correct classification too late to change it without inconsistency in the text, so he resorted to compromise.

⁸⁰ P. 451.

Yatšam.⁸¹ However, Tengsa and Yatšam are often closer in form to each other than either is to Rong—for during its period of separation Rong has developed some independent phonetic peculiarities.

The correct classification of Rong presents an interesting problem for Indian anthropologists and historians: Were the Rong left behind when the Northern Naga branch (and perhaps all the Kukish peoples) migrated from the Himalayas to their present location on the Indo-Burmese border, or are the Rong a remnant left behind from a time when the Northern Naga extended clear across the Valley of Assam?

I shall not take up here all the other languages which Konow mis-classified as the reader may note them by comparing Konow's classification with the one presented here, and I have rectified some of his errors in papers already published.

Relationships Outside Sino-Tibetan

I have presented a considerable number of comparisons, with some phonetic equations, between Sino-Tibetan and Vietnamese,⁸² Athapaskan,⁸³ Khasic and Palaungic (as representing Austroasian).⁸⁴ I am inclined to agree now with André G. Haudricourt, however, that Vietnamese is probably basically Austroasian, and I think that the comparisons I made between Vietnamese and Sino-Tibetan should probably be added to those I made between Sino-Tibetan and two of the Austroasian languages.⁸⁵

Yet one might infer that I was attempting to set up a macro-Sino-Tibetan family consisting of Sino-Tibetan, Austroasian, and Athapaskan. I have not had a chance to examine the evidence critically, but I believe if it were done it would show that the

⁸¹ One may get some idea of the correct classification of Rong in „Classification of Some Languages of the Himalayas,” *Journ. Bihar Res. Soc.* 36 (1950), insert between pp. 173 and 174.

⁸² “L’annamite et le tibéto-birman,” *Bull. Éc. Fr. Ext.-Or.* 40 (1940), 439-442, or “Annamese and Tibeto-Burmic,” *Harv. Journ. As. St.* 6 (1942), 399-402; “Le vietnamien et le tibéto-birman,” *Dân Việt Nam*, No. 1 (Hanoi, 1948), 1-10.

⁸³ “Athapaskan and Sino-Tibetan,” *Internat. Journ. Am. Ling.* 18 (1952), 12-19.

⁸⁴ “Études sur l’austroasien,” *Bull. Soc. Ling.* 48 (1952), fasc. 1, pp. 133-158.

⁸⁵ I make this statement on the basis of some manuscript pages sent me by M. Haudricourt. It is probable that the Manic languages are also Austroasian, judging from his “Introduction à la phonologie historique des langues miao-yao,” *BEFEO* 44 (1954), 554-576.

Sino-Tibetan and Athapaskan roots were practically all different from those in the Sino-Tibetan and Austroasian comparisons. Does this indicate that Sino-Tibetan is really some kind of mixture from two more primitive families of languages, let us say of Pacifican and Austroasian?

San Francisco.
