

THE LIMITS AND CONNECTIONS OF AUSTROASIATIC IN THE NORTHEAST

by

ANDRÉ G. HAUDRICOURT

We shall examine the problems presented by the connections of the Austroasiatic languages with Cham, Vietnamese, the Thai languages, and the Miao languages.

Distinguishing between families of languages, and ascertaining to what family a given language belongs are both done more easily in Europe than in eastern Asia. In Europe the morphology and the grammatical structure make it easy to distinguish between an Indo-European language and a Finno-Ugric or a Semitic language. Because the languages of eastern Asia have neither regular nor irregular inflections, and because the syntax of these languages is often the same, comparative linguists must work almost entirely with vocabulary. Until now there have been ambiguous cases which each linguist has resolved according to his own theories and attitudes. When a language 'A' has a part of its vocabulary in common with a language 'B' and another part in common with a language 'C', one can say that language 'A' is a 'mixed language' (*Mischsprache*). Both W. Schmidt and H. Maspero speak of 'mixed languages' – Schmidt in referring to Cham, and Maspero to Vietnamese; but, still wanting to give a genetic classification for these languages, they arbitrarily classified 'A' with 'C' (Cham with Mon-Khmer, and Vietnamese with Thai).

Today we have other theories. From the genetic point of view there is no such thing as a 'mixed language'. Either the two parts of the vocabulary are equivalent, in which case 'A', 'B', and 'C' are members of the same language family, or the two parts of the vocabulary are not equivalent. One part is original and the other is borrowed. 'A' which is related to 'B' has borrowed a part of its vocabulary from 'C' (Cham is Austronesian, and Vietnamese is Austroasiatic).

Since 1875 when vocabularies and texts on Cham and its dialects were

published,¹ speculations on the nature of the relationship between Cham and Malay have begun to appear. We can classify these as belonging to either of two points of view, one concerned with linguistic affiliations, and the other with geographic origins.

Genetic affiliations with the Indonesian languages were shown by H. Kern,² Ernst Kuhn,³ and G. K. Niemann,⁴ but K. Himly⁵ and W. Schmidt⁶ preferred to consider Cham a mixed language related to Austroasiatic. W. Schmidt was recognized as the originator of the term 'Austroasiatic', and the creator of the 'Austrie' group which brought together Austroasiatic and Austronesian, and his opinion on the place of Cham was often accepted by non-specialists. One even finds it in Salzner's Atlas.⁷ How-

¹ Dr. A. Morice, "Étude sur deux dialectes de l'Indo-Chine. Les Tiams et les Stiengs", *Revue de Linguistique* (Paris, Maisonneuve, 1875). R. Humann, *Vocabulaire français-tiam* (Saigon, autogr., n.d.). A. Landes, *Contes Tjames*, textes et caractères tjames, accompagnés de la transcription du premier conte en caractères romains et d'un lexique (Saigon, autogr., 1886). E. Aymonier, *Grammaire de la langue Chame ou Tchame* (Paris, Leroux, 1889) (This has also appeared in *Excursions et Reconnaissances*, 31, Saigon, 1889). – Note that the genealogical connections of the Cham people have already been discussed in Crawford, *Grammar and Dictionary of the Malay Language* (London, 1852) and in A. Bastian, *Sprachvergleichende Studien mit besonderer Berücksichtigung der Indochinesischen Sprachen* (Leipzig, 1870) and particularly by a French missionary, M. C. Fontaine (1815-71) who stated: "The greater part of these dialects, especially Charay, Redai, Candio and Penong have so many common points of contact that they can only be considered as branches of a single trunk. After having lived among these tribes for several years, and having been obliged to return to Singapore because of precarious health, I was surprised, after a short study of Malay, to find that it contained a great number of Charay words, and an even greater number of words such as the numerals, which showed a marked similarity in the two languages. I have no doubt that these similarities would be found to be still more striking by those who undertook a 'study in depth' of these two languages; their grammatical construction is absolutely identical." *Annales de l'Extrême Orient*, Paris, Challamel, 1882-83, 5, 264. The first Jarai vocabulary published was that of J. Moura, *Le royaume du Cambodge* (Paris, Leroux, 1883), tome I, pp. 440-47 and 501-5.

² H. Kern, "Taalkundige gegevens ter bepaling van het stamland der Malaisch-Polynesische volken", *Versl. en Med. der Kon. Akad. van Wetenschappen*, III r., 6, p. 270.

³ Ernst Kuhn, "Beiträge zur Sprachenkunde Hinterindiens", *Sitzungber. d. k. bayer. Ak. d. Wiss. Philos.-philol. Cl.*, München, 1889, 2 teil, pp. 190-236.

⁴ G. K. Niemann, "Bijdrage tot de kennis der verhouding van het Tjam tot de talen van Indonesië", *Bijdr. tot de Taal-, Land- en Volkenkunde van Ned. Indië*, 1891, pp. 27-44.

⁵ K. Himly, "Sprachvergleichende Untersuchung des Wortschatzes der Tscham-Sprache", *Sitzungber. d. k. bayer. Akad. d. Wis. Philos.-philol. Cl.*, München, 1890, 3. Heft.

⁶ W. Schmidt, "Die Mon-Khmer-Völker, ein Bindeglied zwischen Völkern Zentralasiens und Austronesiens", *Archiv für Anthropologie*, 5 (Braunschweig, 1906), and a French translation in *BEFEO*, 7, p. 223 (Hanoi 1907), and also *Die Sprachfamilien und Sprachenkreise der Erde* (Heidelberg, 1926), pp. 135-47.

⁷ R. Salzner, *Sprachenatlas des indopazifischen Raumes* (Wiesbaden, Harrassowitz, 1960).

ever, Kern's opinion: 'Cham is as much Mon-Khmer as – say – Tahitian',⁸ has been accepted by the specialists.⁹

As early as 1877 Dr. Hamy drew the following conclusions from this view:

If the Chams are not the only tribes on the peninsula who speak a Malay language, and if all the tribes of the principal range of the western mountains have this language in common with the Chams, we must regard the ethnic group of which the Piaks, the Charays, and the Chams are the branches as a true continental Malay group; the immigration of the Menang Kabeu was no more than a simple episode in the history of the race, and it is in the Indo-Chinese mountains and not in Sumatra that one must seek the origin of a people who have played the most important role in the history of Western Oceania.¹⁰

This opinion was also shared by both Kern and Schmidt, and for half a century the mountain peoples of Vietnam, Laos, and Cambodia were called Indonesian by French authors, even though they spoke Austro-asiatic languages.

Only twenty years ago, P. K. Benedict¹¹ showed that the continental origins of the Austronesian languages were to be sought further north toward the coast of Kuangtung. Later, the study of I. Dyen on *q in Austronesian¹² showed that the treatment of *q > h was peculiar to western Indonesia (Java and Sumatra). We may conclude from this that Cham and its dialects belong to this group. The change of *q > h must have been made in Java and Sumatra even though in the extreme north-western extension of the group one finds the Selung language of the Mergui islands of Burma where *q > k.¹³

⁸ E. Aymonier and A. Cabaton, *Dictionnaire Cam-français* (Paris, Leroux, 1906), page vi.

⁹ For example by Henri Maspero in the article which he wrote in his brother's book, G. Maspero, *Un empire colonial français, l'Indochine* (Paris, Van Oest, 1929), t. 1, p. 66. One must always remember that in 1912, H. Maspero seemed to follow implicitly W. Schmidt in putting Cham with Mon-Khmer in his tables, *BEFEO*, 12.

¹⁰ In *La Nature*, 8, IX (1877), p. 292, according to *Annales de l'Extrême Orient*, 5, 264 (Paris, 1882-83).

¹¹ P. K. Benedict, "Thai, Kadai and Indonesian: a new alignment in Southeastern Asia", *Am. Anthr.*, 44, pp. 576-601.

¹² I. Dyen, *The Proto-Malayo-Polynesian Laryngeals* (= *W. D. Whitney Ling. Ser.*) (Baltimore, 1953), with the correction that q is not a laryngeal but a uvular.

¹³ Walter G. White, *An Introduction to the Mawken Language* (Toungoo, SPG. Mission Press, 1911), and some words in J. R. Logan, "The Silong tribe of the Mergui Archipelago", *J. of the Ind. Arch.*, 4 (1850), pp. 411-412; John Anderson, *The Selungs of the Margui Archipelago* (1889); W. J. S. Carrapiett, *The Salons, Ethnographical Survey of India, Burma 2* (Rangoon, Gov. Press, 1909).

	<i>Selung</i>	<i>Malay</i>	<i>Cham</i>
liver	katai	hati	hatai
shrimp	kodang	hudang	hudang
rain	kujang	hujang	hujang
black	ketam	hitam	hitam
thigh	paka	paha	pha
trunk	pokon	pohon	phun
year	takon	tahun	thun
branch	dakan	dahan	dhan
earth	tanak	tanah	tanäh
blood	dalak	darah	darah
to beat	bonok	bunoh	banuh

Finally, there exists, at least in Sumatra, an Austroasiatic substratum which has contributed to the similarity between Cham and the laungage of Acheh.¹⁴ (There even exist in Malay Mon-Khmer words unknown in Cham, for example, 'crab', *ketam*,¹⁵ Bahnar *kötam*, Khmer *ktam*, Mon *gatam*, Samre *tham*, Khasi *tham*, Wa *tam*.)

The delicate problem with which the Austroasiatic comparative linguist is faced is that of finding the words in the languages which surround Cham: Maa, Mnong, Bahnar, which have been borrowed in the course of centuries of Cham domination. The borrowed words may come from Sanskrit, but one also finds them in Mon and in Khmer and in the dialects influenced by Mon and Khmer; they may come from Indonesian but in order to be certain of their Indonesian origin, one must do more than find them in Java and Sumatra, since the Indonesian languages of these regions have borrowed Austroasiatic words.

The affiliation of Vietnamese has been the occasion for even more protracted misunderstandings. On the basis of syntax and vocabulary this language has been related to Mon by J. R. Logan,¹⁶ C. J. F. S. Forbes,¹⁷ Fr. Müller,¹⁸ E. Kuhn,¹⁹ and W. Schmidt,²⁰ this in spite of the

¹⁴ H. J. K. Cowan, "Aantekeningen betreffende de verhouding van het Atjehsch tot de Mon-Khmer talen", *Bijdragen tot de Taal-, Land- en Volkenkunde van Nederl. Indië*, 1948, t. 104, pp. 429-514. Summarised in: H. J. K. Cowan, "Prospects of a 'Papuan' Comparative linguistics", *Bijdr.* ..., t. 113, pp. 73-75.

¹⁵ H. J. Pinnow, *Versuch einer Historischen Lautlehre der Khariasprache*, p. 77 (59).

¹⁶ J. R. Logan, "Ethnology of the Indo-Pacific Islands", *Journ. of the Indian Archipelago*, 6 (1852), p. 658.

¹⁷ C. J. F. S. Forbes, *Comparative Grammar of Further India* (London, Allen, 1881).

¹⁸ Fr. Müller, *Grundriss der Sprachwissenschaft*, Bd. 4 (Wien, 1888), p. 222.

¹⁹ Ernst Kuhn, *Beiträge* ... id. (3).

²⁰ Wilhelm Schmidt, *Grundzüge einer Lautlehre der Mon-Khmer Sprachen* (Wien, Holder, 1905).

opinion of A. H. Keane,²¹ who classifies Vietnamese as one of the tone languages spoken by the yellow race, which differ radically from the languages without tones spoken by the brown race.

Henri Maspero, Sinologist and jurist by training, and self-taught linguist, while he was a member of l'Ecole française d'Extrême Orient, wrote an article on Vietnamese²² in which he concluded that this language was the result of mixture, but that the existence of a tone system must connect it with the Thai family.

For him as for Keane, tones are a permanent characteristic of a language; he did not ask whether, in the course of history, a language without tones could acquire them, or whether a language with tones could lose them. The linguists who preceded him did not take into account the Vietnamese tones when they compared the words of this language with those of Mon or Khmer. However, in a work on Thai²³ Maspero had noted a regular correspondence between tones in comparing languages of the same family. It was therefore legitimate to consider tones to be as important as vowels or consonants, besides which he looked widely for a relationship between tones and (non-tonal) phonemes. For him the tone languages of the Far East were characterized by a relationship between the tone and the initial phoneme of the word,²⁴ but he did not distinguish between synchronic and diachronic differences nor between tonemes and tones (which was excusable in view of the date of his work).

From his comparisons he deduced:

When these words entered Annamite they all received a tone. The tone, however, differed according to whether the Mon-Khmer initial phoneme was voiceless, voiceless and aspirated, or voiced. In general, words with initial voiceless or half-voiceless sounds or with asyllabic voiceless prefixes have taken in Annamite the 'băng' tone or the 'sắc' tone, and the words with a voiced nasal or liquid initial take the 'huyền' tone; the words with initial *s* or *h*, or with liquid or nasal initial preceded by a prefix *h* or *s* are classed as aspirates and take the 'sắc' tone; *before other initials of any kind, prefixes are dropped without affecting the tone.* Words with final stop take the 'sắc' and 'nặng' tones respectively. Lastly, final *h* (or perhaps *s*) frequently resulted in Annamite 'nặng' or 'ngã' tones –

²¹ A. H. Keane, "On the relations of the Indo-Oceanic races and languages", *Journ. of the Anthr. Inst. of Great Britain and Ireland* (London, Trubner, 1880), French translation in *Annales de l'Extrême orient revue asiatique et océanienne*, Paris, Challamel, t. 5 (1882-83), pp. 238-50 and 264-78 under the title "Rapport ethnologiques et linguistiques des races indochinoises et indo-pacifiques".

²² Henri Maspero, "Etudes sur la phonétique historique de la langue annamite, les initiales", *BEFEO*, 12 (1912).

²³ Henri Maspero, "Contribution à l'étude phonétique des langues thai", *BEFEO*, 10 (1910).

²⁴ *BEFEO*, 12, p. 89, note (1).

sometimes these, and sometimes the 'hỏi' tone, depending upon whether the Mon-Khmer initial was voiced or voiceless.²⁵

H. Maspero did not consider that it was necessary to distinguish toneme languages where the tone was sufficient to distinguish between words, from tonemeless languages where the tone was automatically determined by the nature of the consonant. One understands how far away he was from modern linguistics when he states:

...A very important fact seems to me to be established: that the Annamite tone system is essentially the same as those of the Thai languages, of Chinese, and of the Tibeto-Burman languages. The tone depends upon the distinction of the pitch given to the initial, and upon the inflection given to the final. Historical evolution has, however, led to the dissolution of the old system: the six tones have survived individually but they are only survivals: for ten centuries tones have existed in Annamite, but properly speaking a tonal system no longer exists.²⁶

It seems clear to us today that since the pitch of the tone is determined by the nature of the initial consonant and since its inflection is determined by the nature of the final consonant, there is no toneme, tones are simply the phonetic consequence of the phonemes of the word; where Maspero saw a system of tones, we see only a language without tones (tonemes). On the contrary, it is when the system has dissolved, when the automatic connection between the consonant and the tone is broken following the consonantal change at the beginning or end of the word, that – for us – tonemes appear and we have a true tone (tonemic) language. It follows from the very terms used by Maspero, that the tone languages of the Far East did not originally have tonemes. Thus the argument from the tonal system is invalid for separating Vietnamese from the Austroasiatic languages on basic grounds.

It only remains for us to find out if the correspondences discovered by Maspero enable us to explain the six tonemes of Vietnamese. According to his terminology the tones are produced by the combination of two pitches and three inflections; we have:

Height			
High	bằng	sắc	hỏi
Low	huyền	nặng	ngã
	Inflection 1	Inflection 2	Inflection 3

²⁵ BEFEO, 12, pp. 99-100.

²⁶ BEFEO, 12, pp. 102-103.

The origin of the pitch distinctions is clear, they are due to the initial consonant. The old voiceless initials gave rise to a high toneme. This may be due to an earlier voiceless stop: *con* 'child', *chim* 'bird', *dan* 'to weave', *ba* 'three', *cá* 'fish', *chây* 'louse', *chiết* 'to die' . . . , a former voiceless spirant; *tám* 'eight', *tóc* 'hair' . . . an earlier liquid or nasal made voiceless by a preceding spirant: *năm* 'year', *lá* 'leaf', *lúa* 'rice', *răng* 'tooth', or an earlier preglottalized consonant: *năm* 'five', *nước* 'water', *muôi* 'salt', *măng* 'bamboo shoot', . . .

The old voiced initials became low tonemes where there were old voiced stops: *cầm* 'chin', *đọng* 'stagnant', *bụng* 'belly', liquids: *ruôi* 'fly', *lưỡi* 'tongue', *rễ* 'root', . . . or a nasal: *mũi* 'nose', *muoi* 'mosquito', *một* 'one', *mặt* 'face', . . .

The origin of the differentiation of inflection is less easy to show. H. Maspero made the origin of Inflection 3 obvious when he stated that it corresponded in Mon and Bahnar to some words ending in *h*. By way of contrast, he wanted to look for the origin of Inflection 2, or at least for that of the *sắc* tone, in the aspiration of the initials; such an investigation would seem to be in contradiction with his assertion that inflections depend on finals. However, at that time, Maspero did not have the evidence for a solution. It was only in 1954, after having taken into consideration the unpublished materials of G. H. Luce on Riang, G. K. Izikowitz on Lamet, and W. A. Smalley on Khmu, that I was able to present a solution.²⁷ Inflection 2 corresponds to the words ending in Glottal stop (-ʔ) in many languages of the Palaung-Wa group; however this result might have been predicted from the fact that this inflection was the only one noted for words which preserved the final stops *-c*, *-t*, *-p*.

The origin of the inflections can, then, be shown thus:

<i>Vietnamese Inflections</i>	1	2	3
<i>Austroasiatic endings</i>	voiced sounds	voiceless stops	voiceless spirants

The proposed connection between Vietnamese and Thai was questioned by C. O. Blagden,²⁸ and J. Przyluski,²⁹ but accepted by W. Schmidt,³⁰ and adopted by R. Salzner.³¹ Even though Pinnow has frequently cited

²⁷ "L'origine des tons en vietnamien", *Journal Asiatique*, 242, 1, pp. 69-82.

²⁸ C. O. Blagden, "The Classification of the Annamese Language", *Journ. of the Royal As. Soc.* (1913), pp. 427-432.

²⁹ A. Meillet and M. Cohen, *Les langues du monde*, 1st ed. (1924), p. 395.

³⁰ W. Schmidt, *Die Sprachfamilien und Sprachenkreise der Erde* (1926), p. 126.

³¹ See fn. 7.

Vietnamese in his work,³² he has not indexed it, and since he has indexed all the other Austroasiatic languages, he seems not to consider it as such.

Since writing my article, of the writers on this subject that I have seen, only N. D. Andreev agrees with me.³³

The problem of the relation of the Thai languages and the Austroasiatic languages is, then, very closely related to the above; if Vietnamese is Austroasiatic, what then is the significance of the Thai vocabulary that Maspero claims is to be found there?

In examining this vocabulary, as I have begun to do in order to classify Vietnamese,³⁴ one is struck by its heterogeneity.

In the first place, I do not consider it certain that Thai is genetically connected with Chinese, so that the many words which resemble Chinese could be ancient loans in Thai, and if they are found in Vietnamese, this language could also have borrowed them from Chinese.

The way in which most Vietnamese look at the relation between Chinese and Vietnamese should be noted here. For them, Chinese is the language of learning; classical Chinese is pronounced in a Vietnamese way according to a tradition dating back to the tenth century, the date of the independence of Vietnam. For the Vietnamese the words of the vernacular language, Vietnamese proper, are borrowed from Chinese, pronounced in this fashion. Apparently, they do not seem to suspect that the loans, made for the most part during the Chinese domination, before the tenth century, were from an older form of Chinese, pronounced differently.³⁵ I have tried in two articles^{36, 37} to show the importance of these loans. We may cite from Maspero's list the words common to Thai and Vietnamese which must be of Chinese origin: *chè* 'to row', *chì* 'lead (metal)', *đôi* 'yoke', *bè* 'raft', *bẩn* 'to divide', *bánh* 'bread', *phân* 'dust', *tiếng* 'noise', *khem* 'to refrain from', *kim* 'needle', *đuă* 'chop stick', *dức* 'male', *ngà* 'ivory', *nàng* 'Mrs.', *mèo* 'cat' ...

³² See fn. 15.

³³ N. D. Andreev, "K voprosu o proisxoždenii vjetnamskogo jazyka", *Sovetskoe vostokovenie*, 2 (1958), pp. 101-111.

³⁴ "La place du vietnamien dans les langues austroasiatiques", *BSL*, 49, 1 (1953), pp. 122-128.

³⁵ According to the writings of L. Cadière, *BEFEO*, 2 (1902), pp. 1-113, *BEFEO*, 4 (1904), pp. 1065-81, *BEFEO*, 8 (1908), pp. 93-148, 381-481, pp. *BEFEO*, 10 (1910), pp. 61-92, 287-337, *BEFEO*, 11 (1911), pp. 67-110, and the more recent work of Nguyễn Bạt-Tụy, *Chữ và văn Việt khoa-học* (Saigon, 1950), *Ngôn-ngữ-học Việt-nam* (Saigon, 1959).

³⁶ See fn. 27.

³⁷ "Comment reconstruire le chinois archaïque", *Word*, 10, pp. 358-64.

Otherwise most of the Thai languages used by Maspero are spoken by conquering populations established in territories of Austroasiatic peoples. The Shan are found among the Palaung-Wa, the Siamese separate the Mon from the Khmer, the Laotians are in the minority in their own country, etc. We must expect to find many Austroasiatic loans in these Thai languages, and even these are often more difficult to bring to light than Thai loans in Austroasiatic. In Maspero's list we shall cite: *bung* 'belly', which should be compared with Khmu, *bung*, Samrê *pung*, (in Samre: $p < *b$, $t < *d$, $k < *g$, $ph < *p$, $th < *t$, $kh < *k$), *nghe* 'to hear', with Riang *ngar*, *nganh* 'hook', with Khmer *ngieng*, ...

Lastly, Maspero's bias – that Vietnamese is related to Thai – has led him to compare Vietnamese words with Thai forms rather than with Austroasiatic forms quite as satisfactory in both form and meaning. Thai and Austroasiatic, such as *sa* for *cổ* 'neck', *cầm* 'chin' and some plant and animal names of southern origin which Thai must have borrowed from Austroasiatic such as *gạo* 'paddy', *gà* 'chicken', *cà* 'eggplant', ...

The precision with which Thai and Austroasiatic can be compared depends also upon the extent to which the old forms of current words of the Thai languages can be reconstructed. We must admit that the efforts of K. Wulff³⁸ and of P. K. Benedict³⁹ do not provide enough regular correspondences with Indonesian, and do not permit an explanation of the origin of tones (which the Vietnamese-Austroasiatic correspondences do). We have scarcely begun to have enough material on the Thai (or Daic) family and we still know nothing of practically all the Kadai languages. It is in large part due to the investigations of Li Fang Kuei and more recently to the remarkable work of the linguists of the Democratic Peoples' Republic of China, that we can temporarily classify the Daic languages into four groups.

The Southern or Thai proper group can be divided into subgroups: one to the west of the Red River, in Vietnam (White and Black Tay), in Yunnan (Tay Lü and Tay Nüa), in Laos, in Thailand (Siamese and Yuon), in Burma (Khün, Shan, and Khamti), and in India (Ahom), the languages of the conquering and literate peoples, and the others to the east of the same river, in Vietnam (Tay and Thô) and in China (Nung and Thô of Lung-Chou), dialects of illiterate peasants. The regularity

³⁸ K. Wulff, "Über das Verhältnis des Malayo-polynesischen zum Indochinesischen", *Det Kgl. Danske Videnskabernes Selskab. Hist.-Philol. Meddelelser*, 27, 2 (København, 1942), p. 157.

³⁹ See fn. 31.

of correspondences allows the reconstruction of a monosyllabic Proto-Thai.⁴⁰

The Eastern group is still not well known, but it includes in Laos the Sêk language⁴¹ and in Vietnam, the languages of the Man Cao-lan and the Ts'ün-lao.⁴² The pronunciation of some words is different, which suggests that the mother tongue of these two groups might be disyllabic. This group certainly exists in China, and perhaps one must connect with it the dialects of Tien-pao and of Yung-shun, collected by Fang Kuei Li.

The Western group, long known by the name of Dioi, is Fang Kuei Li's Northern group. It is spoken in Vietnam by certain immigrant peoples (Nhang, Giây, Qui-châu) and particularly in China, by large minorities at Kuang-hsi (Chuang) and at Kueichou (Pu-Yi). This group often coincides in pronunciation with the preceding, but differs in initial consonants in some words. It seems to me that there also one must infer the existence of disyllabic words, the present initial consonants possibly having been intervocalic.

The Northern group includes the languages in central China, called Kam-Sui by Fang Kuei Li and Dong-Shui by Serdyuchenko,⁴³ to which probably should be added Be spoken on the Northern coast of Hainan.⁴⁴ This group, which has been unknown until recently, is particularly important for comparative linguists because of its archaism. Here are two examples: the common Thai word for dog **hma*, which isolates Thai from the other languages; the Kam form: *hng^{wa}* shows that here an onomatopoeic word must have superseded some other word. Thai proper

⁴⁰ The reconstruction was started by Maspero (23), continued by K. Wulff, "Chinesisch und Tai, Sprachvergleichende Untersuchungen", *Det Kgl. Danske Videnskabernes Selskab. Hist.-Philol. Meddelelser*, 20, 3, pp. 34-90; and in my articles: "Les phonèmes et le vocabulaire du Thai commun", *Journal Asiatique*, 236 (1948), pp. 197-238; "Les consonnes uvulaires in Thai", *BSL*, 48 (1952), 1, pp. 84-5; "Les consonnes préglottalisées en Indochine", *BSL*, 46 (1950) 1, pp. 172-82; "De la restitution des initiales dans les langues monosyllabiques, le problème du Thai commun", *BSL*, 49 (1956), 1, pp. 307-22; this last article was a reply to Fang Kuei Li, "Consonant Clusters in Tai", *Language*, 30 (1954), 3, pp. 368-79.

⁴¹ My communication to the Congress of Orientalists in Moscow (1960), "La langue Sêk", will appear in the Proceedings of the Congress.

⁴² See my "Note sur les dialectes de la région de Moncay", *BEFEO*, 50 (1960), 1, pp. 168-73.

⁴³ G. P. Serdyuchenko, *A survey of the Zhuang-Tai languages in the Chinese people's republic* (Papers presented by the USSR delegation, 25th International Congress of Orientalists, Moscow, 1960), 18 pp. and Fang Kuei Li, *Language*, 24 (1948), 2, note 2, p. 165. Note also the appearance of *Gaeml-Gax jianming cidian* (Kam-Chinese Pocket Dictionary) (Guizhou, 1959), 254 pp.

⁴⁴ Dictionnaire Ong-Bê, manuscript by F. Savina, which I am in the process of editing.

has no specific word for 'bamboo shoot', so that we have: Sêk: *nāng*, Mak: *nāng*, Sui: *nāng*, Kam: *nāng*, Bê: *nan*, Dioi: *rāng*, Chuang: *rāng*, which must be compared not only to the Đày (Kadai) *nüōng* but also to Vietnamese: *măng*, Müông: *Phăng*, Bahnar: *tôPhăng*, Mon: *Phăng*, and Khmer: *lampăng*.

At the present time, the relation between the Daic languages and Austroasiatic is hardly demonstrable, but the comparison with Austroasiatic is just as useful as that with Indonesian in reconstructing a polysyllabic Proto-Daic.

Although the Miao-Yao languages have been very imperfectly studied, writers have used them in comparative work even at a very early period. R. M. Davies⁴⁵ did not hesitate to classify them as Austroasiatic, and his view was adopted more recently by R. D. A. Forrest⁴⁶ even though Maspero and many others have preferred to put them in the Tibeto-Burman family.⁴⁷ W. Schmidt and J. Przyluski⁴⁸ continued, as did the ancient Chinese, to confuse them with Thai.

We are now beginning, as I have shown,⁴⁹ to have a clear idea of the phonological history of these languages, and some dictionaries have recently appeared.⁵⁰ The tone system is comparable to that of Chinese, Daic, and Vietnamese, but unlike these languages, the Miao-Yao finals are kept only in the two Yao languages; the others, the Miao languages proper, are composed of open syllables, nasalized or not. In contrast to the other languages, the initials in certain Miao dialects are more conservative having half-nasals: *mp*, *nt* . . . probably indicating a nasal prefix, uvular *q*, *g*, . . . kept even in certain loan words or related to Chinese, groups of initial consonants not only with *l*: *pl*, *npl*, *ml*, *kl*, . . . but with *r*:

⁴⁵ R. M. Davies, *Yünnan, the link between India and the Yangtze* (Cambridge, 1909), p. 341.

⁴⁶ R. A. D. Forrest, *The Chinese Language* (London, Faber), pp. 89-96.

⁴⁷ A. Meillet and M. Cohen, *Les langues du monde*, 2 ed. (1952), pp. 563-6.

⁴⁸ See fn. 29.

⁴⁹ "Introduction à la phonologie historique des langues miao-yao", *BEFEO*, 44 (1954), 2, pp. 555-76.

⁵⁰ *Qiangdong fangyan Hmub-Diel jianming cidian* (Eastern Kuei-chou Miao-Chinese Pocket Dictionary) (Guizhou, 1958), which permits us to make more precise the pronunciations given in J. Esquirol, *Dictionnaire Kanao-français et français-kanao* (Hongkong, 1931), and *Chuanqiandian fangyan Hmongb-Shuad jianming cidian* (Szechuan-Kueichou-Yunnan Miao-Chinese Pocket Dictionary) (Guizhou, 1958), which renders obsolete Savina's "Dictionnaire Miao-tseu-français", *BEFEO*, 16 (1916) since the only Europeans who have recorded this language properly are David Crockett Graham, *Vocabulary of the Ch'uan miao* (Chungking, 1939), and G. Linwood Barney and W. A. Smalley, *Third Report on Meo (Miao)* (Xianghuang, 1953). On the other hand, on Yao we still have Savina, "Dictionnaire français-man", *BEFEO*, 26 (1926), pp. 12-255.

pr, *mpr*, *tr*, *ntr* ... The Proto-Miao word therefore has many more relevant features than the Thai word, and is that much more useful for comparison.

If by analogy with Vietnamese *-p* reconstructed for Inflection 2 and *-h* for Inflection 3, a reconstruction of Proto-Miao is obtained which may be compared with Austroasiatic. Several of Davies' comparisons may thus be verified:

1. 'dog' – Hmong *tle*, Hmu *tla*, Kanao *hla*, Yao-mien *kyo*, Yao-mun *klo*, give in Proto-Miao **kloP* but in Pinnow, page 112: Khmu *soP*, Riang *shoP*, Vietnamese *chô* produce **čōP*, now since the *ch* of that language sometimes corresponds to *kl*, we have 'banana' – Vietnamese *chuôi*, Thai *kluoi*, the comparison is better justified with 'dog' in Kharia, *soloP* (Old Mon *clüw*, Mon *klüw*).

2. 'eye' – Hmong *mua*, Hmu *ma*, Kanao *mai*, Yao-mien *muoi*, Yao-mun *mei*, > Proto-Miao **muaih*. There is no more reason to compare words with Austroasiatic forms having final *-t* than with Tibeto-Burman forms having final *-k*. (R) VN *măt*

3. 'nose' – Hmong *ncü*, Hmu *nai*, Kanao *nyi*, Yao-mien 'blui', > (R) *mäh* Proto-Miao **mbriuh* (this reconstruction of an *-r-* is justified by the following words which in certain languages do have an *-r-*: 'five' Hmong *či*, Hmu *tsa*, Kanao *kya*, Yao-mien *pla*, Yao-mun *pya*, > Proto-Miao **pra*. 'House' Hmong *če*, Hmu *tsai*, Kanao *kyi*, Yao-mien *plau*, Yao-mun *pyau*, > Proto-Miao **prauP*. 'Fish' – Hmong *nče*, Hmu *nai*, Kanao *nyi*, Yao-mien *plau*, Yao-mun *pyau*, Proto-Miao **mbrauP*. The final sound of the word agrees with that of the Austroasiatic word for 'nose', Vietnamese *mũi* < **muih*.

4. 'blood' – Hmong *nšā*, Hmu *šhā*, Kanao *šhie*, Yao-mien *žham*, Yao-mun *šam*, > Proto-Miao **nčhamP* which is to be compared with: Santali: *maŋam*, Bahnar: *maham*, Khmer: *jham*. (R) *moham*

5. 'bone' – Hmong *tshā*, Hmu *shō*, Kanao *shō*, Yao-mien *pbung*, Yao-mun *sung*, > Proto-Miao **tshung* but since Hmong needs an *-h* tone and Yao-mien needs another initial, we have in Austroasiatic, Santal *jang*, Khasi *šPəng*, Khmer *cPəng*, Khmu *cPang*, Vietnamese *xỉống*, and Riang *khruang*. (R) *Kesəng*

6. 'water' – Hmong *tle*, Hmu *ö*, Kanao *au*, Yao-mien *uom*, Yao-mun *wam*, > Proto-Miao **Puom* which is to be compared with Palaung-Wa *om*; on the other hand since *tle* represents an ancient **gloP*, it therefore does not prove that Miao is closer to the Burmese *re* than to the Tibetan *ču* as Maspero thought. In some dialects of Hmong *tle* becomes *pde*, but this does not prove a connection with the Austroasiatic *pḍāk*. (R) *hām*

7. 'kite (eagle)' – Hmong *tlǎ*, Hmu *tlǎ*, Kanao *hle*, Yao-mun *klang* > Proto-Miao **klangʔ*, which is to be compared with Bahnar *klāng*, (R) *klā*, Khmer *klēng*, and Khasi *khlieng*.

8. 'name' – Hmong *mpe*, Hmu *pi*, Kanao *pie*, Yao-mien *puo*, Yao-mun *pu*, > Proto-Miao **mpuoh* which is to be compared with Khmer *jmoh*, Riang *mus*, but there also is in Old Mon the form *simoʔ*.

The Miao-Yao languages are not in contact with Austroasiatic languages, they are separated from them by the Tibeto-Burman languages, Thai and Kadai. The arrival of the Yao (Man) peoples in Vietnam dates only from the seventeenth century, and that of the Miao (Meo) from the nineteenth.

Some other words like 'moon' – Hmong *hli*, Hmu *hla*, Kanao *hla*, Yao-mien *hla*, Yao-mun *la* > Proto-Miao word **hlah* and 'pig' Hmong *mpua*, Hmu *pa*, Kanao *pa* > Proto-Miao **mpaih* are rather to be compared with Burmese *la'* and *wak*, Tibetan *zlaba* and *phag*, while 'to kill' – Hmong *tua*, Yao *tǎy* > the Proto-Miao word **taih* and 'to die' Hmong *tua*, Hmu *ta*, Kanao *ta*, Yao *tǎy*, > Proto-Miao the word **daih* can be compared with Thai *tāi*, Indonesian *matai*, just as the name of birds – Hmong *nong*, Hmu *nō*, Kanao *nau*, Yao-mien *nōk*, Yao-mun *noʔ*, > Proto-Miao **nongh/nok* can be compared with Thai *nōk*, Indonesian *manuk*.

The Miao-yao languages seen to form a link between the Austroasiatic and the Tibeto-Burman families in the same way that the Karen languages do, and their phonological richness is useful in reconstruction.

(1961; translated from French)