

The place of Vietnamese in Austroasiatic

André-Georges Haudricourt

▶ **To cite this version:**

André-Georges Haudricourt. The place of Vietnamese in Austroasiatic. 2017. halshs-01631477

HAL Id: halshs-01631477
https://shs.hal.science/halshs-01631477

Preprint submitted on 9 Nov 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Non-final version (November 9th, 2017). In preparation for: Haudricourt, André-Georges. Evolution of languages and techniques*. (Ed.) Martine Mazaudon, Boyd Michailovsky & Alexis Michaud. (Trends in Linguistics. Studies and Monographs [TiLSM] 270). Berlin: De Gruyter Mouton.*

The place of Vietnamese in Austroasiatic (1953)

Originally published as: La place du vietnamien dans les langues austroasiatiques, *Bulletin de la Société de Linguistique de Paris* 49(1): 122–128.

translated by Alexis Michaud

Abstract

Vietnamese must be classified in the Austroasiatic language family, between the Palaung-Wa group to the north-west and the Mon-Khmer group to the south-west.

The Vietnamese language has been considered as related to the Mon language for more than a century (Logan 1852: 658; Forbes 1881; Müller 1888, vol. 4, p. 222; Kuhn 1889; Schmidt 1905: 5; Przyluski 1924: 395; Benedict 1942). However, over the past forty years, the opinion (based on work by Henri Maspero) that Vietnamese is related to the Tai[1] languages has been expressed with some frequency. Looking at H. Maspero's article, we actually find a more nuanced opinion:

> "Pre-Annamese was born out of the fusion of a Mon-Khmer dialect with a Tai dialect; the fusion may even have involved a third language, which remains unidentified. At a later period, the Annamese [=Vietnamese] language borrowed a huge number of Chinese words. But the language whose influence dominated and gave Annamese its modern form was

1. ["Tai" was substituted for Haudricourt's "Thai" to refer to the language family. Haudricourt emphasized a distinction between Dioi (also known as "Zhuang", after the Chinese ethnonym *Zhuàng* 壮) and closely related languages on the one hand, and on the other hand the rest of the family, referred to as "*le thai proprement dit*" (1956: 313). Haudricourt's views on Tai language classification can be found in "On the reconstruction of initial consonants in monosyllabic languages: the case of Proto-Tai" (1956; this volume) and his 1967 article about Lakkia.]

> definitely a Tai language, in my opinion; I think the Annamese language must be related to the Tai family." (Maspero 1912: 117)

If we accept that there is no such thing as "fusion" between languages, and that genealogical relatedness must be assessed on the basis of core vocabulary and grammatical structure, we are led to consider that the modern form of a language is not determined by its genealogical origin, but by the influences to which it is subjected in the course of its history. For a thousand years, from the first or second centuries B.C. until the 10th century, the Tai and Vietnamese peoples lived under Chinese domination; to me, this is a sufficient explanation for the converging evolution of the Tai and Vietnamese languages.

The Tai languages located to the west of the Red River (Siamese, Shan, Lao, White Tai, Black Tai) are extremely close to one another, whereas to the east of this river the languages diverge to a greater or lesser extent: Dioi [a.k.a. Yoy, Buyi 布依, Zhuang 壮], Cao Lan, Mak, Sui, and the less-closely related Gelao, Tulao [土老], Lati (La Chi) and Laqua. It therefore seems that the Tai languages originate in southern China, and only crossed the Red River around the 10th century.

The Muong languages, archaic languages belonging to the same branch as Vietnamese, are spoken from the Red River down to the latitude of the city of Vinh to the south.

Likewise, the Austroasiatic languages proper do not extend beyond the Red River to the east. The Palaung-Wa group (see Shafer 1952) extends into Laos, with the Lamet and Khmu languages.[2] The Phong dialects[3] are

2. The sources used here are Shafer (1952) and an unpublished Riang vocabulary collected by G.H. Luce. As for Lamet, I have used unpublished vocabulary collected by K.G. Izikowitz, who has published some words and texts (1951); there is also a short vocabulary list in Doudard de Lagrée & Garnier (1873, vol. 2, p. 510). The Khmu data come from an unpublished vocabulary list collected by W.A. Smalley in Luang Prabang (Laos); from Roux & Trần 1927 (data on "Tsa Khmu", spoken in the neighbourhood of Dien Bien Phu, Vietnam); from the entries "Mi of Xieng cang" and "Khmous of Louangprabang" in Doudard de Lagrée & Garnier (1873); and from the "Kha khao des Houa phan" vocabulary in Macey (1907b: 871-874). The following two varieties must also be considered part of the Khmu language group: (i) the dialect spoken by the Phu Theng in Cammon (Laos), for which we have a word list collected by the cao müöng [district chief] of Kham Keut; and (ii) the dialect of Nghe-An (Vietnam) documented by H. Maspero (1912), under the name "Kha".

spoken in an area located between the Khmu and Muong areas; to the south, the Kuy dialects [a.k.a. Kui, Kuai, Suai, Souei][4] extend along the Mekong river down to Cambodia. The other languages that will be used in the present study are Bahnar (the best-preserved of the South-Eastern Austroasiatic languages), Mnong (Biêt, Gar, Köho), [5] Khmer, Samre (spoken by the Chong, Por and Pear),[6] Mon, and Khasi.

---

3. On Phong, see Macey (1907a: 1407); the variety described is the language of Muong ċam (spoken in Cammon, Laos). This language is very close to Hung and Khong-kheng, two dialects of Nghe-An (Vietnam) described by Guignard (1907). Maspero (1912) used these two dialects, considering them as Muong dialects. See also the Kha muong ben vocabulary in Fraisse (1950a), and the manuscript vocabulary list of the Tum-on-phong and the Kha tong-luong collected by the cao müöng [district chief] of Kham Keut. [Haudricourt sometimes cites the language as *Phong* (the more commonly used name), and sometimes as *Pong*, following Macey (1907a). In the translation, this has been standardized to *Phong.*]
   These articles also contain vocabularies of more divergent dialects, such as P. Macey's Kha Phong and Kha Pou-Hok vocabularies; the language of the Harèmes [Arem] has been documented by A. Rivière (Pavie 1902: 285-290).
   I have not taken into account the Sek language: it is insufficiently well documented. One may wonder whether the Sek dialect collected by L. Cadière in Quang-Binh (Cadière 1905) and classified by H. Maspero among the Muong languages belongs to the same language as that spoken by the Sek of Laos, who live along the banks of the Mekong river and whose vocabulary is documented by Rivière (Pavie 1902), Macey (1907a), Fraisse (1950a), and in the manuscript of the cao müöng of Kham Keut cited above). More than half of the vocabulary collected by Cadière appears to be of Thai-Tay stock, and some of it is even more archaic than Proto-Thai-Tay in some respects, calling to mind the Tai dialects spoken to the east of the Red River: for instance, 'house' is ran; compare Proto-Tai *rɯən, Dioi ran, Cao Lan lan, Mak zan, Sui ɣan.
4. On Kuy, see (i) the Kuy O and Kuy Ntra vocabularies in Lévy (1943); (ii) the Kha Tairi, Kha Mong-Khong and Sô vocabularies in Macey (1907b); (iii) A. Rivière (Pavie 1902); (iv) the So and Nanhang vocabulary in Doudard de Lagrée & Garnier (1873); (v) Fraisse (1950b), and the Kha Tiali and Kha Salang vocabulary in Fraisse (1950a). Lastly, I have also used a manuscript vocabulary list of Kha Bourou.
5. The Mnong data are based on unpublished documents by Georges Condominas about the Gar dialect [see also Condominas (1952)]. On other dialects (Biet and Köho), see Hoeffel (1936), Cassaigne (1936), and Dournes (1950).
6. On Samre, see Brengues (1905:41-46) [Haudricourt's attribution to Fr. Petit-Huguenin is an error]; Isarangura (1935); and Baradat (1941).

Looking at grammatical similarities, there are so many parallels between all of these languages that one cannot tell which of them is closest to Vietnamese. However, if one examines the remnants of earlier morphology, one finds, among other examples, a causative relationship between 'to die' and 'to kill', which is absent from Proto-Thai (Table 1).

[Table 1. 'To die' and 'to kill' in Vietnamese and selected Austroasiatic languages vs. Proto-Thai.]

language	'to die'	'to kill'
Vietnamese (orthography)	*chết*	*giết*
Vietnamese (pronunciation)[7]	čet	ǯet
Phong	cet	pøcet
Khmu	han	phan[8]
Lamet	yam	pyam
Kuy Ntra	kcet	komcet
Khmer	*slap*	*samlap*
Mon	*khyüt*	*pacüt*
as against		
Proto-Thai	*tay	*qa

Of course, one could object that 'to die' and 'to kill' might have been borrowed into Vietnamese together from a language such as Phong. The decisive evidence comes from basic vocabulary.

H. Maspero discussed the origin of Vietnamese vocabulary in the following terms:

> "Words of Mon-Khmer origin are undeniably very numerous in Annamese [Vietnamese], and they are among the most common words: the numeral system is entirely Mon-Khmer (…) However, there is not a single complete set of Mon-Khmer words in any semantic field: words of Mon-Khmer origin are blended with words of Tai origin, to an extent that varies across word sets. For instance, the words *trăng* 'moon', *mưa* 'rain', *gió* 'wind' and *nước* 'water' are of Mon-Khmer origin, but the words *mọc* 'fog' and *mùa* 'season' are Tai; *rú* 'mountain, forest' and *sông* 'river' are Mon-Khmer, but *đồng* 'field', *rẫy* 'dry ricefield' and *mó* 'spring' are Tai; *mắt* 'eye' and *chân* 'foot' are Mon-Khmer, but *lưng* 'back', *bụng* 'belly', *ức* 'chest', *cằm* 'chin',

7. [International Phonetic Alphabet transcriptions: cet^{B1} and ʒet^{B1}, respectively.]
8. [Under the influence of Lao, the old clusters C + h in Khmu have become aspirated consonants.]

> *bi* 'calf (body part)',[9] *cổ* 'neck' are Tai. The word for 'clothes', *áo*, is Mon-Khmer, but 'to sew', *nhíp*, is Tai. Of the two words for 'rice', *lúa* and *gạo*, the first is Mon-Khmer, the second is Tai. The word for 'bird', *chim*, is Mon-Khmer, but the names of the two domestic birds, 'chicken' *gà* and 'duck' *vịt*, are Tai, etc."[10] (Maspero 1912: 115)

The list of body part names could be Maspero's most convincing argument: two Austroasiatic words as opposed to seven Tai words. But an examination of the most common words, such as those cited in the questionnaire of the École Française d'Extrême-Orient[11] (questions 178 through 189), reveals the following facts:

'Head': the word *đầu* [dɤ̃w^{A2}] comes from Chinese; the older Vietnamese word is *trốc* [ʈok^{D1}], which is also found in Muong and in Phong: **klok**. The other Mon-Khmer languages have different words for 'head': Kuy **plew**, Bahnar **kol**, Mnong **bok**, Samre **tos**, Riang **kiŋ**, Khmu **kampoŋ**, Mon *k'düp*, as against Proto-Thai-Tay ***hruo**.

'Hair': Vietnamese *tóc* [tɔk^{D1}], Muong **t^{h}ak**, Phong and Samre **suk**, Kuy, Mon, Bahnar, Mong **sok**, Riang **huk**, Khmu **glə?**, Khasi **sniuh**, Khmer *sa'k*. Proto-Thai-Tay: ***p^{h}rom**.

'Eye': Vietnamese *mắt* [măt^{D1}], Muong, Phong, Kuy, Khmu, Bahnar, Mnong, Samre, Mon **mat**, Khasi **k^{h}mat**, Khmer *bnêk*… as opposed to Proto-Thai-Tay ***ta~t^{h}a**. (The form **t^{h}a** is observed to the north of the Red River: in Dioi: **ta**, in Mak and Sui: **da**, in Laqua: **te**; Benedict suggests that these words are related to Western Austronesian *mata*.)

'Ear': Vietnamese *tai* [tay^{A1}], Muong **t^{h}ai**, Phong **sai**, Kuy **ktor**, Bahnar **don**, Mnong **tor**, Samre **prlaŋ**, Khmer *traciek*, Riang **ktiek**, Khmu **rmwəy**, Khasi **skor**. Proto-Thai-Tay: ***ɽ̥u~k^{h}ru**, Dioi **rwə**, Mak **eha**, Sui **q^{h}a**, Laqua **rə**, Lati (La Chi) **lu**, Kelao **rao**.

---

9. [Haudricourt points out further below that the form *bi* is not found in Vietnamese dictionaries. The usual word for 'calf' is *bê* [ɓeA1].]
10. [Throughout the present article] Vietnamese words are given in the common orthography, Mon and Khmer are provided in transliteration, and the other languages are transcribed phonologically. The Proto-Thai-Tay words all have the level tone, except for 'tooth', 'belly' and 'rice' which have a rising tone.
11. [This questionnaire was created and printed as a leaflet by the École Française d'Extrême-Orient (1938) for a broad enquiry, which was interrupted by the war in 1940. It contained 484 entries arranged by semantic fields. These leaflets, filled in by civil servants of the colonial administration, constituted one of the main bases of Haudricourt's comparative linguistic research during his stay at the Ecole Française d'Extrême-Orient in Hanoi at the end of the 1940s.]

'Nose': Vietnamese *múi* [muy^{B1}], Muong **mui**, Phong **muic**, Kuy, Mon, Khmu, Bahnar **muh**, Mnong **moh**, Khmer *cramuh*, Samre **mtut**, Riang **kɗo'**, Khasi **k^{h}mut**. Proto-Thai-Tay and Dioi: *ɗaŋ, Mak: **naŋ**.

'Mouth': Vietnamese *miệng* [miəŋ^{B2}], Muong and Phong **mɛŋ**, Riang **mɔŋ**, Khmu **tnoh**, Bahnar **ɓər**, Mnong **bər**, Samre **trnɛŋ**, Khmer *mā't*, Khasi **ʃntur**. Proto-Thai-Tay, Dioi, Sui: **pa:k**, Mak: **ɓup**.

'Tooth': Vietnamese *răng* [răŋ^{A1}], Riang **raŋ**, Khmu **r̥aŋ**, Phong and Kuy **kneŋ**, Samre **k^{h}o:y**, Khmer *dmeñ* [dmeŋ], Mnong **sek**, Mon **ŋek**, Khasi **ɓniat**. Proto-Thai-Tay: **k^{h}ɛw**, Dioi **hew**.

'Tongue': Vietnamese *lưỡi* [lɯəy^{C2}], Muong **la:i**, Phong **la:ic**, Khmu **n̥ta:k**, Riang **ta:k**, Kuy **loai**, Khmer *antāt*, Samre **krta:t**, **slta:t**, Bahnar **lpiet**, Khasi **tl̥liet**. Proto-Thai-Tay: ***lin**, Sui and Mak **ma**.

'Neck': Vietnamese *cổ* [ko^{C1}], Muong **kel**, **kok**, Phong **kiko**, Riang **kok**, Kuy **tkoŋ**, Mon and Mnong **kɔ**, Bahnar **akɔ**, Khmer **koah** [in transliteration: *gah*; the form provided by Haudricourt is that in Modern Khmer], Samre **kɔk**, Khasi **rndaŋ**. Proto-Thai-Tay: *ɢɔ, Dioi and Mak **ho**, Laqua **ku**, Lati (La Chi) **k^{h}i**.

'Lip': Vietnamese *môi* [moy^{A1}], Phong **kbui**, Kuy **tmor**, Samre **sənya:t**,[12] Khmer *papur, babir*, Riang **tor**. Proto-Thai-Tay does not have any specific word for 'lip'; 'rim, edge' is ***rim**, 'skin' is *n̥ăŋ.

'Chin': Vietnamese *cằm* [kăm^{A2}], Muong and Phong **kăŋ**, Kuy **skăŋ**, Bahnar and Mnong **ka:ŋ**, Samre **krmaŋ**, Khmer ('jaw') *dgām*, Riang and Khmu **ka:p**. Proto-Thai-Tay: ***ga:ŋ**, Dioi **haŋ**, Mak ('jaw') **ga:ŋ**.

'Arm' (or 'hand'): Vietnamese *tay* [tăy^{A1}], Muong **t^{h}ai**, Phong **si**, Kuy, Bahnar, Mnong and Samre **ti**, Mon and Khmer **tai**, Riang and Khmu **ti?**, Khasi **kti**. Proto-Thai-Tay: ***muu**, Dioi **fuŋ**, Sui **mya**, Mak **mi**, Kelao **mao** (Western Austronesian: **rima**[13]).

In the above list, there are two words, 'neck' and 'chin', to which H. Maspero, in the passage cited earlier, assigned a Tai origin. In fact, in view of the data adduced above, there is no reason to doubt that these two words

12. [Haudricourt's notation for Samre is *sn'yat*. This is a typographical error: 'lip' in the source used by Haudricourt (Baradat 1941: 3) is written as *Săn yat*, for all the dialects in which the word had not yet been replaced by a Khmer loan.]
13. [The form lima, with initial l, having the meanings 'hand' and 'five', is widespread among Austronesian languages. Proto-Austronesian 'hand' is reconstructed as *qa-lima by Blust and Trussel (2010), who suggest the likelihood that the proto-form is a derivation (by an otherwise unattested prefix) from *lima 'five'.]

have a Mon-Khmer origin. Indeed, earlier on in his article, H. Maspero (1912: 22) had provided a Mon-Khmer origin for Vietnamese ‘chin’. To sum up, the twelve Vietnamese words in the above list fall into three categories:
(i) Words that are common to almost all Austroasiatic languages: ‘hair’, ‘eye’, ‘nose’ and ‘hand’;
(ii) Words that only have cognates in Phong; however, for these words, the other languages differ from one another: ‘hair’, ‘eye’, ‘nose’, and ‘hand’;
(iii) Words that have cognates neither in Muong nor in Phong, but which are observed in Mon-Khmer: ‘neck’ and ‘chin’, or in Palaung-Wa (Khmu, Riang): ‘tooth’.

Lexically, Vietnamese thus behaves like any other Austroasiatic language.

Several of the Vietnamese-Tai cognates proposed by H. Maspero need to be revised. For instance, the Vietnamese word provided for ‘calf (body part)’, *bi*, is not found in standard Vietnamese dictionaries such as Hue (1937),[14] and there is no word for ‘calf’ in Proto-Thai-Tay; *bụng* [buŋ] ‘belly’ has cognates in Kuy and Samre: puŋ, and in Khmu: buŋ, whereas the Proto-Thai-Tay word for ‘belly’ is *dɔŋ. Siamese and Lao are the only two Thai-Tay languages that have borrowed the form buŋ. Muong has a different word for ‘belly’: *trộng* [ʈoŋ^{B2}], which has cognates in Phong and Riang: kluŋ.

Another case in point is ‘rice’: there is a word in Proto-Thai-Tay, *k^{h}ăw, which shows some phonetic similarity with Vietnamese *gạo* [ɣaw^{B2}]; but it is misleading to say that there are “two words” for rice in Vietnamese [Maspero 1912: 115]: these two words refer to two different concepts, which are found in all Austroasiatic languages. These are shown in Table 2.

[Table 2. ‘Paddy’ and ‘husked rice’ in Austroasiatic languages.]

language	‘rice plant; paddy’	‘husked rice’
Vietnamese	*lúa* [luə^{B1}]	*gạo* [ɣaw^{B2}]
Muong, Phong	lo	kao
Kuy	alo	asə
Bahnar	ɓa	p^{h}e
Mnong (Gar, Biet)	ba	p^{h}e
Mnong (Köho)	koe	p^{h}e
Samre	hal	rk^{h}oː

14. [The usual word is *bê* [ɓeA1].]

Khmer	sruː	ŋkɑː
Mon	sroʔ	sŋŭ
Khmu	ŋoʔ	rŋkoʔ
Riang	ŋoʔ	koʔ
Khasi	kɓa	k^{h}aw

The agreement both in content and in form between Khasi, Palaung-Wa (Khmu and Riang), Samre, and Khmer clearly shows that the origin of Vietnamese *gạo* 'husked rice' is to be sought in Austroasiatic. Progress in our knowledge of the northern Austroasiatic languages reveals that some words considered as Tai by H. Maspero are in fact borrowed by Siamese and Lao from an Austroasiatic substratum; these words constitute as many additions to the list of Vietnamese words of Austroasiatic origin. Furthermore, we may question whether the words that are common to Vietnamese and Tai really are borrowings from Tai into Vietnamese, or whether there may not be some borrowings from Vietnamese into Tai; at present, the languages related to Thai are too poorly documented for this issue to be resolved.

Lastly, the argument whereby Vietnamese, having six tones, should be separated from the toneless Austroasiatic languages is not valid. As early as 1913, Blagden pointed out that, under such a hypothesis, one would have to place the Scandinavian languages, which have tones, in a different language family from the rest of the Germanic languages. And indeed, on the basis of the unpublished documents of Luce on Riang and Izikowitz on Lamet, we now know that these undoubtedly Austroasiatic languages have two tones.[15]

Comments

Over a decade after the publication of his 1953 article, Haudricourt noted that the conclusion that Vietnamese must belong in the Austroasiatic language family was still far from being widely accepted. "Since writing

15. [In his later research, Haudricourt sensed that the tones of Riang and Lamet may in fact be register oppositions rather than tone oppositions (Haudricourt 1965: 169-170, translated in this volume, p. xyz). This was confirmed since: on Riang, see Shorto and Luce (2013) and Shintani (2014); the synchronic facts for Lamet are more complex (Lindell et al. 1978) because some dialects have become predominantly tonal (Ferlus, p.c. 2016).]

my article, of the writers on this subject that I have seen, only N.D. Andreev (1958) agrees with me" (Haudricourt 1966: 51). For instance, Pinnow (1959) takes up Maspero's point of view that Vietnamese must be related to Tai. Specialists had a hard time recognizing Vietnamese as an Austroasiatic language because of specific evolutions that made it an outlier within this language family. "Vietnamese is the outcome of a millenium of Chinese influence over a language of the Viet-Muong group of the Austroasiatic family" (Ferlus 1992: 111). The variety of Chinese spoken in Annam (present-day northern Vietnam) during the first millenium was a southern dialect (Hashimoto 1978), rather than the standard variety of the time (the language of the Chinese capital). This Chinese superstratum was much richer phonetically than the Proto-Viet-Muong inherited layer. At the beginning of the second millenium, Proto-Vietnamese also underwent a spirantization of medial obstruents, as shown by Ferlus (1992), building on an observation by Haudricourt (see "Consonant shifts in Mon-Khmer initial stops", this volume). When the language became phonologically monosyllabic, medial spirants were phonologized as an additional series of initial consonants; the resulting system contained many more phonemes than are commonly found in Austroasiatic.

A further factor that complicates the establishment of sound correspondences between Vietnamese and other Austroasiatic languages, and could have delayed acceptance of Haudricourt's conclusion, is the existence of irregular correspondences among initial consonants. 'Foot' is in tone category A1 in Vietnamese: **cɤ̆n**^{A1} (orthographic *chân*), reconstructing to an unvoiced initial consonant ***c** at the stage preceding the tone split (about which see Haudricourt 1961). But Austroasiatic cognates of 'foot' either retain a voiced initial, as in Khmu **ɟiaŋ**, or belong in the low/breathy phonation-type register (*e.g.* Khmer **cəəŋ**, Mon **càŋ**, Lamet **cèŋ**), reconstructed as originating in voiced initials. Both Maspero and Haudricourt list 'foot' among the Vietnamese words of Austroasiatic origin, but in detail the correspondence seems unconvincing. The solution proposed by Ferlus (1999) is that, when Proto-Viet-Muong spread into the Red River delta from a homeland located further south, it replaced earlier Austroasiatic languages which lacked voiced initial stops, hence the sporadic loss of initial voicing (nineteen examples, including 'foot', are listed by Ferlus 1999).

In light of these and other advances in understanding Vietnamese historical phonology, its Austroasiatic affiliation is no longer a subject of

dispute among specialists. Haudricourt's article, together with his 1954 article "The origin of tones in Vietnamese" (this volume), is now recognized as a landmark offering "compelling empirical and theoretical bases" for relating Vietnamese to Austroasiatic (Sidwell 2009: 37; see also Gage 1985: 493-494).

References

Baradat, Raymond. 1941. Les dialectes des tribus sâmrê [The dialects of Sâmrê tribes]. Unpublished manuscript of the Ecole Française d'Extrême-Orient, Paris. 267 p. http://sealang.net/monkhmer/database (accessed in 2010).

Benedict, Paul K. 1942. Thai, Kadai and Indonesian: a new alignment in southeastern Asia. *American Anthropologist* 44. 576–601.

Blagden, Charles Otto. 1913. The classification of the Annamese language. *Journal of the Royal Asiatic Society* 1913. 427–432.

Brengues, Jean. 1905. Notes sur les populations de la région des Montagnes des Cardamomes [Notes about the populations of the region of the Cardamom mountains]. *Journal of the Siam Society* 2. 19–47.

Cadière, Léopold. 1905. Les hautes vallées du Sông-Gianh [The upper reaches of the Gianh river]. *Bulletin de l'École Française d'Extrême-Orient* 1. 349–367.

Cao Müöng of Kham Keut [i.e. the Head of the county of Kham Keut, *Bolikhamxay Province, Laos*; *Cao* means 'prince, dignitary', and *müöng* is the administrative level below the province]. n.d. Manuscript vocabulary list of the Tum-on-phong and the Kha tong-luong languages, collected as part of the 1938–1940 linguistic data collection project of the École Française d'Extrême-Orient.

Cassaigne, Jean. 1936. *Dictionnaire Koho* [A Koho dictionary]. Saigon: Imprimerie de l'Union.

Doudard de Lagrée, Ernest & Francis Garnier. 1873. *Voyage d'exploration en Indochine* [An exploratory voyage through Southeast Asia]. Paris: Hachette.

Dournes, Jacques. 1950. *Dictionnaire sré (köho) - français* [A Sre (Koho) - French dictionary]. Saigon: Imprimerie d'Extrême-Orient.

École Française d'Extrême-Orient. 1938. *Questionnaire linguistique* [Linguistic questionnaire]. Hanoi: Imprimerie d'Extrême-Orient.

Forbes, Charles J. F. S. 1881. *Comparative Grammar of Further India*. London: W. H. Allen.

Fraisse, André. 1950a. Les tribus Sek et Kha de la province de Cammon (Laos) [The Sek and Kha tribes of Cammon province, Laos]. *Bulletin de la Société d'Etudes Indochinoises* 25(3). 410–413.

Fraisse, André. 1950b. Les tribus Sô de la province de Cammon, Laos [The Sô tribes of Cammon province, Laos]. *Bulletin de la Société d'Etudes Indochinoises*, nouvelle série, 5(25). 171–185.

Guignard, Théodore. 1907. Vocabulaire Hung et Khong-kheng [Hung and Khong-kheng word lists]. École Française d'Extrême-Orient manuscript (reference in the Hanoi library of the École Française d'Extrême-Orient: E.F.E.O. 8°395).

Hoeffel, Ernest. 1936. *Lexique franco-biat* [A French-Biat lexicon]. Saigon: Imprimerie de l'Union Nguyên Văn Cua. xix, 5, 189 p.

Hue, Gustave. 1937. *Dictionnaire annamite-chinois-français* [Annamese-Chinese-French dictionary]. Hanoi: Trung-Hoà.

Isarangura, Nai Noe. 1935. Vocabulary of Chawng words collected in Krat province. *Journal of the Siam Society* 28(2). 173–180.

Izikowitz, Karl Gustav. 1951. *Lamet Hill Peasants in French Indochina.* Göteborg: Etnografiska Museet.

Kuhn, Ernst. 1889. Beiträge zur Sprachenkunde Hinterindiens [Contributions to Southeast Asian linguistics]. *Sitzungsberichten der Münchener Akademie der Wissenschaften* 2. 190–236.

Lévy, Paul. 1943. *Recherches préhistoriques dans la région de Mlu Prei [Cambodge], accompagnées de comparaisons archéologiques et suivies d'un vocabulaire français-kuy* [Research on the prehistory of the area of Mlu Prei (Cambodia), with archaeological comparisons, and followed by a French-Kuy glossary]. Hanoi: Imprimerie d'Extrême-Orient.

Logan, John R. 1852. Ethnology of the Indo-Pacific islands. *Journal of the Indian Archipelago* 6.

Macey, Paul. 1907a. Etude ethnographique et linguistique sur les K'Katiam-pong-houk, dits 'Thai-Pong' (province de Cammon, Laos) [An ethnographic and linguistic study of the K'Katiam-pong-houk, a.k.a. 'Thai-Pong' (Cammon province, Laos)]. *Revue Indochinoise* 5. 1417–1424.

Macey, Paul. 1907b. Etude ethnographique sur les khas [An ethnographic study of the Kha people]. *Revue Indochinoise.* 869–874.

Maspero, Henri. 1912. Etude sur la phonétique historique de la langue annamite. Les initiales [Studies in Annamese historical phonetics. Initial consonants]. *Bulletin de l'École Française d'Extrême-Orient* 12. 1–127.

Müller, Friedrich. 1888. *Grundriß der Sprachwissenschaft* [Outline of linguistics]. Wien: Hölder.

Pavie, Auguste. 1902. *Mission Pavie, 1879-1895, Géographie et voyages, volume 4 : Voyages au centre de l'Annam et du Laos, et dans les régions sauvages de l'Est de l'Indo-Chine* [Report of the 1879-1895 Pavie mission – Geography and travel routes, volume 4: Journeys to the centre of Annam and Laos, and into the wild regions of eastern Indo-China]. Paris: E. Leroux.

Przyluski, Jean. 1924. Les langues austroasiatiques [The Austroasiatic languages]. In Antoine Meillet & Marcel Cohen (eds.), *Les langues du monde* [Languages of the World], 385–403. Paris: Librairie Édouard Champion.

Roux, Henri & Trần Văn Chu. 1927. Les Tsa Khmu [The Tsa Khmu people]. *Bulletin de l'Ecole Française d'Extrême-Orient* 27. 169–222.

Schmidt, Wilhelm. 1905. *Grundzüge einer Lautlehre der Mon-Khmer Sprachen* [Principles of a phonology of the Mon-Khmer languages]. Wien: Hölder.

Shafer, Robert. 1952. Etudes sur l'austroasien [Studies in Austroasian]. *Bulletin de la Société de Linguistique de Paris* 48. 111–158.

Additional references

Andreev, Nikolaj Dmitrievič (Андреев, Н. Д.). 1958. K voprosu o proishozhdenii v'etnamskogo jazyka (К вопросу о происхождении вьетнамского языка) [On the origins of the Vietnamese language]. *Sovetskoe vostokovenie* 2. 101–111.

Blust, Robert & Stephen Trussel. 2010 (updated 2016). Austronesian Comparative Dictionary. http://www.trussel2.com/acd/ (accessed in 2016).

Brengues, Jean. Notes sur les populations de la région des Montagnes des Cardamomes [Notes about the populations of the Cardamom Mountains area]. *Journal of the Siam Society* 2. 19–47

Condominas, Georges. Rapport d'une mission ethnologique en pays Mnong Gar [Report on an ethnological field trip in Mnong Gar territory]. *Bulletin de l'Ecole française d'Extrême-Orient* 46(1). 303–313.

Ferlus, Michel. 1992. Histoire abrégée de l'évolution des consonnes initiales du vietnamien et du sino-vietnamien [A short history of the evolution of Vietnamese and Sino-Vietnamese initial consonants]. *Mon-Khmer Studies* 20. 111–125.

Ferlus, Michel. 1996. Langues et peuples viet-muong [Viet-Muong languages and peoples]. *Mon-Khmer Studies* 26. 7–28.

Ferlus, Michel. 1999. Les disharmonies tonales en viet-muong et leurs implications historiques [Irregular tonal correspondences within Viet-Muong (Vietic) and their historical implications]. *Cahiers de Linguistique Asie Orientale* 28(1). 83–99.

Gage, William W. 1985. Vietnamese in Mon-Khmer perspective. In R. Suriya, D. Thomas & P. Suwilai (eds.), *Southeast Asian Linguistic Studies presented to André-G. Haudricourt,* 493–524. Bangkok: Mahidol University.

Hashimoto, Mantaro. 1978. Current developments in Sino-Vietnamese studies. *Journal of Chinese Linguistics* 6(1). 1–26.

Haudricourt, André-Georges. 1956. De la restitution des initiales dans les langues monosyllabiques: le problème du thai commun [On the reconstruction of initial consonants in monosyllabic languages: the case of Proto-Tai]. *Bulletin de la Société de Linguistique de Paris* 52(1). 307–322.

Haudricourt, André-Georges. 1966. The limits and connections of Austroasiatic in the northeast. In N. H. Zide (ed.), *Studies in comparative Austroasiatic linguistics*, 44–56. The Hague: Mouton.

Haudricourt, André-Georges. 1967. La langue lakkia [The Lakkia language]. *Bulletin de la Société de Linguistique de Paris* 62(1). 165–182.

Lindell, Kristina, Jan-Olof Svantesson & Damrong Tayanin. 1978. Two dialects of the Rəmeet (Lamet) language. *Cahiers de Linguistique Asie orientale* 4(1). 5–22.

Nguyễn Như Ý. 1999. *Đại từ điển tiếng Việt* [Comprehensive Vietnamese dictionary]. Hanoi: Nhà Xuất bản văn hoá thông tin.

Phan, John D. 2013. Lacquered Words: the evolution of Vietnamese under Sinitic influences from the 1st century BCE to the 17th century CE. Ph.D., Cornell University.

Pinnow, Heinz-Jürgen. 1959. *Versuch einer historischen Lautlehre der Kharia-Sprache* [Towards a historical phonology of the Kharia language]. Wiesbaden: Otto Harrassowitz.

Sidwell, Paul. 2009. Classifying the Austroasiatic languages: History and state of the art. Munich: Lincom.

Thomas, David & Robert Headley. 1970. More on Mon-Khmer subgroupings. *Lingua* 25. 398–418.

Watkins, Calvert. 1990. Etymologies, equations, and comparanda: Types and values, and criteria for judgment. In Philip Baldi (ed.), *Linguistic Change and Reconstruction Methodology*, 289–303. Berlin: Mouton de Gruyter.

Wulff, Kurt. 1934. *Chinesisch und Tai: sprachvergleichende Untersuchungen* [Chinese and Tai: comparative linguistic studies]. Det Kongelige Danske Videnskabernes Selskab. Historik-filologiske Meddelelser 20. 3. Copenhagen: Levin & Munksgaard.