

## On the relationship between the Austroasiatic and Austronesian languages in Southeast Asia

Tran Tri Doi\*

*College of Social Sciences and Humanities, VNU  
336 Nguyen Trai, Thanh Xuan, Hanoi, Vietnam*

Received 20 June 2009

**Abstract.** As have been known, the Austroasiatic and Austronesian languages phonetically and lexically have correspondence. However, whether this is *borrowed or inherited* relationship has long been an issue of controversy, because plausible evidence to these points of view is still unavailable.

In this paper, equivalent data from basic vocabulary between the Viet Muong disyllabic / sesquisyllabic languages (e.g. Arem, Ma Lieng, Sach, Ruc, Aheu) and the Chamic languages are carefully investigated. Despite the shared basic vocabulary, this kind of equivalence lends further weight to the view that preference is given to the borrowed relationship. For that reason, these lexical resemblances are of a restricted range which supports the relation of *special borrowings* between the two languages.

I. P.K. Benedict, in his 1973 paper, argued for a 'substratum' relationship between Austronesian (abbreviated as AN) and Austroasiatic (AA) languages when he offered a further explanation of the so-called 'Austro-Thai' and its relatedness to Austroasiatic languages [1]. In his terminology, 'Austro-Thai' refers to the languages of Miao-Yao, Tai-Kadai, and Austronesian.

In the same year of 1973, the link between these two language families was also put forward for consideration by S.E. Jakhontov. Having suggested that Vietnamese is part of the Austroasiatic (Mon-Khmer, in his words) language family and Tai is inheritance-related to the Austronesian (or Indonesian, in his

terminology) rather than to Sino languages, this link is regarded as a borrowing relationship [2].

A year later, A.G. Haudicourt shared the same line of analysis with P.K. Benedict and S.E. Jakhontov [3,4]. He further emphasized that this is a special borrowing relationship since the shared words of the two language families belong to the basic vocabulary and regularly appear in different sub-groups. As he pointed out, there are some Mon-Khmer words in the Malaysian language, which are unavailable in the Cham languages, for instance: 'crab' *ketam*, Bahnar *kotam*, Khmer *ktam*, Mon *gatam*, Samre *tham*, Khasi *tham*, Wa *tam*" [3, tr.33]. Also, the vocabulary of some Mon-Khmer languages including Maa, Mnong, Bahnar consists essentially of loans made from the Cham languages as a consequence of the long-term dominance of the

\*DI: 84-4- 35588603.  
E-mail: doihanh@yahoo.com

Cham people. Noteworthy, 'the borrowed lexicon originated either from Sanskrit ...; or Indonesian languages' [3, tr.33].

The issue seems to be sorted out until lately some other proposals of the classification of Southeast Asian languages have been released. Most recently, from sound correspondences, shared vocabulary and shared morphology, L. Sagart (2004) has argued that the Tai-Kadai languages are a subgroup of Austronesian and additionally, Austronesian and Chinese are genetically related within a macrophyllum which he called 'Proto-Sino-Tibetan-Austronesian' (PSTAN) [5]. Accordingly, in the area of pre-literate Southeast Asia and South China, Sino-Tibetan, Austronesian, and Tai-Kadai languages all belonged to a macro-family, which excluding Austroasiatic.

Therefore, despite of different approaches, previous scholars have shared the same idea that the correspondence between Austroasiatic (particularly Mon-Khmer) and Austronesian languages is not as proof of an inheritance relationship, but as a result of borrowings.

2. On the other hand, Phạm Đức Dương, in his effort to build up a model of a mixed language which arises through language contact, has been argued for a hypothesis of 'ngữ hệ Đông Nam Á' (literally means: Southeast Asian language-family) or Proto-Austroasiatic as he further explained. In his system, this language family consists of the languages of Austroasiatic, Austronesian and Dong Tai (equivalent to Tai-Kadai in other systems) [6, tr.30]. It means that Austroasiatic and Austronesian initially originated from the same family, which subsequently divided into three present-day sub-groups, namely Austroasiatic, Austronesian and Dong Tai.

Phạm Đức Dương, however, provided only ethnolinguistic arguments to support his position. Strictly speaking, linguistic evidence for a kin relationship between Austroasiatic and

Austronesian has yet to be presented. Although many aspects of this hypothesis remain unclear, there is solid indication that the special relatedness between the two language families, which was first noticed by A.G. Haudricourt, is worthy of deeper consideration. This paper, therefore, is concerned with empirical data from different languages in order to shed some light on this unusual relationship.

3. Having investigated some languages that still keep in existence characteristics of the ancient Viet-Muong languages, which belong to Mon-Khmer, a branch of the Austroasiatic family [7], we realized that these languages have much of core vocabulary cognate shared with some other Austronesian languages. Taken into account, the lexical equivalence is devoted to illuminate the nature of the relationship between Viet-Muong and Austronesian languages, or more specifically, between northeastern Mon-Khmer and mainland Austronesian languages of the Southeast Asia region.

The languages investigated including Arem (Ar), Malieng (ML), Sach (S), Ruc (R) or Aheu (Ah), Kha Phong (Kh), etc. are well known for having a 'sesqui-syllabic pattern', which is supportive of the fact that they well sustain the linguistic peculiarities of the Proto-Viet Muong (PVM) period. Since Proto-Viet Muong is the most conservative member of the Mon-Khmer branch, the lexical correspondence between Viet-Muong and Austronesian languages is of great significance.

Speakers of sesqui-syllabic Viet-Muong languages are primarily located in the provinces of Quang Binh, Ha Tinh, and south Nghe An, where they live along mountainous areas of the Vietnam-Laos border (see the map below). Due to rugged terrain, tormentous weather and limited transportation in this geographical region, these languages consist of some well-preserved phonetic patterns of the Proto-Viet Muong language [8].

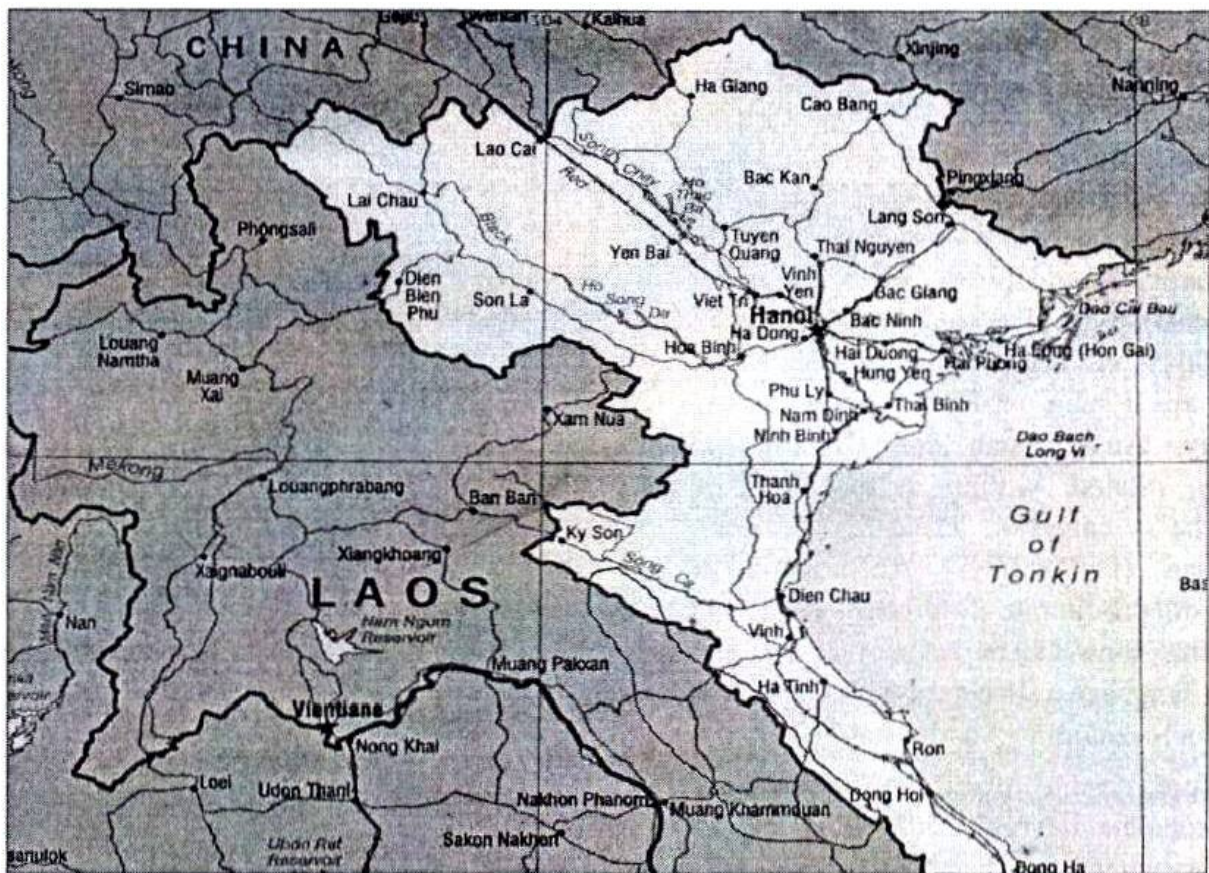


Figure of the residential territory of the sesqui-syllabic Viet-Muong languages (red-lined area).

### 3.1. Presentation of data

Data observation reveals that the *sesqui-syllabic* Viet-Muong and the Cham languages lexically have correspondence. More

interestingly, most of the shared words are essential components of the vocabulary of the two languages under comparison. Two sub-categories of lexicon are given with respect to land and temporal elements.

#### 3.1.1. Land-related words

Cham	Vietnamese, Muong	Sesqui-syllabic Viet-Muong
chok (C) "rocky mountain"	núi đá (V), núi tá (M) "rocky mountain"	ci t (R), lakú:ɲ ʔatɛʔ (Ar) "rocky mountain"
paɪəu (C) "stone"	đá (V), tá (M) "stone"	latá (R, S), ʔatɛʔ (Ar) "stone"
haluk (C) "earth"	đất (V), tát (M) "earth"	ban (R, S), ʔaták (Ar) "earth"
haluk ləm (C) "clay"	đất sét (V) "clay"	ban tɬet (R), ʔaták kuɲec (ML) "clay"
chuah (C) "sand"	cát (V), kách (M) "sand"	takác (R), ʔatɛʔ kát, taka:c (Ar) "sand"

### 3.1.2. Time-related words

Cham	Vietnamese, Muong	Sesqui-syllabic VietMuong
haray (C) "day"	ngày (V), ngày (M) "day"	pakəh (S, R), bəh (Ar) pakət (Kh) "day"
gok page (C) "early morning"	sáng sớm (V), lạng khóm (M) "early morning"	ʃam (R, S), ʔaram' (Ar) "early morning"
jalá (C) "noon"	trưa (V), trưa (M) "noon"	pakəh (S, R), cilia (Ar) kalía (ML) "noon"
mudom (C) "night"	đêm (V), tēm (M) "night"	li m (S, R), lám (Ar) "night"
mudom mudom (C) "late night"	đêm hôm (V), tēm (M) "late night"	hom (S, R), təjɔp (ML) "late night"
bilan (C) "month"	tháng (V), khàng (M) "month"	tʰɛŋ' (S, R), tʰɛŋ' (Ar) "month"
bilan (C) "moon"	trăng (V), tlàng (M) "moon"	palian (S, R), ʔmɾɛah (Ar) "moon"
thun (C) "year"	năm (V), năm (M) "year"	năm (S, R), tʰun (Ar), sanăm (Kh) "year"

Note: Cham data follow Bùi Khánh Thế [9]; Muong data follow [10]; Ruc data follow Nguyễn Phú Phong [11]; Ruc, Arem, Sach, Malieng, Kha Phong data come from our fieldwork. Cham and Muong transcriptions follow the original texts, others are transcribed in IPA.

### 3.2. Discussion of data

Apparently, the two wordlists are parts of the core vocabulary. Following the treatments of P.K. Benedict, S.E. Jakhontov, A. G. Haudricourt, and L. Sagart, these lexical correspondences, at first glance, are more likely indicative of inherited relationship. Thoroughly considered, this is probably not the case.

3.2.1. In the first place, regarding land-related terms, the Ruc words for "rocky mountain" and "earth" are quite similar to the equivalent Cham words. Regarding words referring to time, the correspondence between Cham and Viet-Muong is more varied. In particular, Cham and Viet are equivalent in terms of the word for 'day'; Cham and Viet, Muong, Arem, Malieng have the term for 'noon' in common; while Cham and Viet, Muong, Sach, Ruc correspondingly share the word for 'moon', which is synonymous to 'month' in Cham. On the other hand, the term for 'year' is shared only by Cham and Arem;

while the chosen word for 'early morning', to some extent, is jointly employed by Cham, Sach, Ruc, Kha Phong.

It is fairly easy to see the phonetic regularity among these basic words, such as Cham *bilan*, Viet *trăng*, Muong *blăng/tlăng*, Sach và Ruc *palian*. Based on this kind of equivalence between Cham and VietMuong only, it seems to lend further weight to the view that preference is given to the genetic relationship.

3.2.2. A careful examination of the data, however, reveals different results. The Cham and Viet-Muong forms do not truly correspond regularly in the sense that some Viet-Muong words are similar to Cham's, but not to those of other languages in the same Viet-Muong group. For example, with regard to the words for "year", *tʰun* in Arem is related to *thun* in Cham, yet has no connection with *năm* or *sanăm* in other Viet-Muong. Similarly, the Cham word for "moon" forms identifiable patterns of

equivalence with those of Viet, Muong, Sach, Ruc, but not with *Amreah* in Arem. The terms for “noon” are mostly shared by the languages under comparison, except Sach, Ruc. In parallel, the Cham words for “rocky mountain”, and probably for “earth” exhibit pretty direct correspondence with those of Sach, Ruc, but not with those of other Viet-Muong.

The irregular similarities within Viet-Muong languages show that Viet-Muong languages only correspond *individually* with Cham. Despite the fact that Cham and Viet-Muong have the core vocabulary in common (as shown in 3.2.1), this kind of equivalence is hardly indicative of inherited phonetic relationship.

This irregularity, however, can be explained in the following way: some Cham words are sustainable in some Viet-Muong languages, but are lost in other languages of the same Viet-Muong group. Therefore, the individual correspondence, by itself, would be insufficient to cast grave doubt on the kin link between Cham and Viet-Muong.

3.2.3. A more thorough investigation of the data, however, gives out an interesting point, namely *the consistency of each wordlist* of the Cham and sesqui-syllabic Viet-Muong languages under comparison. This is probably the most apparent evidence of the borrowing relationship between the two groups.

In the first list of land-related words, the terminologies for “stone”, “rocky mountain”, “earth”, “clay”, and “sand” in the sesqui-syllabic Viet-Muong languages uniformly have the matching forms in Vietnamese and Muong. Although the words for “rocky mountain” and “earth” in Ruc are different in detail, they cannot in principle be considered as a violation of the consistency relation of the whole group.

The same holds true for the second list of time-related words. Terminologies for “early

morning”, “noon”, “late night”, “month”, and “year” exhibit a consistent pattern of behavior throughout the Viet-Muong group. The exceptional cases of the words for “day” and “moon” do not fundamentally fail to agree with this pattern.

In summary, the consistent correspondence within the Viet-Muong group provides strong evidence to confirm that the Viet-Muong basic words, which are quite similar to those of the Cham languages, are of borrowed origin. Due to the non-indigenous origin, they are preserved in some, not all languages of the entire Viet-Muong group. In other words, the Cham group of Austronesian does share with the Viet-Muong group of Austroasiatic some core vocabularies, and this similarity is best taken as a solid indicator of special borrowings.

Consequently, Southeast Asia, in our point of view, is home to five language families, namely Austroasiatic, Austronesian, Tai – Kadai, Sino – Tibetan and Miao-Yao; which is also strongly held by other scholars. It is not incompatible with the view that the Southeast Asia region exhibits a consistent cultural-linguistic resemblance. This cultural-linguistic correspondence, however, should be strictly differentiated from the similarity of linguistic origins.

## References

- [1] P.K. Benedict, *Austro - Thai and Austroasiatic*, Austroasiatic Studies, part 1, 1976, pp 4-36.
- [2] S.E. Jakhontov, Về sự phân loại các ngôn ngữ ở Đông Nam châu Á, *Ngôn ngữ* n<sup>o</sup>1 (1991) 73-77.
- [3] A.G. Haudricourt, Limites et connexions de l'austroasiatique au Nord - Est, *Asie du Sud-Est et monde insulinduen* 5, N<sup>o</sup>1 (1974) 1-14.
- [4] A.G. Haudricourt, Giới hạn và nối kết của ngôn ngữ Nam Á ở Đông Bắc, *Ngôn ngữ* n<sup>o</sup>1 (1991) 33-40.

- [5] L. Sargat, *The higher phylogeny of Austronesian and the position of Tai - Kadai*, Workshop on "Premieres austronésien: langues, gènes, systèmes de parenté", Paris, May 5, 2004.
- [6] Phạm Đức Dương, *Học tranh ngôn ngữ-văn hoá tộc người ở Việt Nam và Đông Nam Á*, NXB Đại học Quốc gia Hà Nội, 2007.
- [7] Trần Tri Dõi, *Giáo trình lịch sử tiếng Việt (sơ thảo)*, NXB Đại học Quốc gia Hà Nội, 2005.
- [8] M. Ferlus, *The Origin of Tones in Viet-Muong*, SALS XI<sup>th</sup> Conference, Mahidol University, Bangkok, Thailand, May 16-18-2001.
- [9] Bùi Khánh Thế (Chủ biên) (1996), *Từ điển Việt - Chăm*, NXB Khoa học Xã hội, 1996.
- [10] Nguyễn Văn Khang (Chủ biên), *Từ điển Mường - Việt*, NXB Văn hoá dân tộc, Hà Nội, 2002.
- [11] Nguyễn Phú Phong et al, *Lexique Vietnamien - Ruc - Français*, Université de Paris VII, Paris 1988.