The Languages of Vietnam: Mosaics and Expansions

Jerold A. Edmondson and Kenneth J. Gregerson

The University of Texas at Arlington

Abstract

The languages of Vietnam reflect the historical migrations of prehistoric people groups as well as successive waves of cultural influence by world civilizations over many thousands of years. In addition to cultural spread, certain of Vietnam’s ethnicities also constitutes social mosaics in the sense of Renfrew (2000), growing and flourishing in place. The linguistic fabric of the nation is woven of widely different strands, including Austroasiatic and Austronesian speech communities, Sino-Tibetan and Kra-Dai (Tai-Kadai), as well as Hmong-Mien. Despite the chaos of the last half century, national and international scholars, building on French studies of the nineteenth century, have been able to document language names, places, populations, and linguistic relationships to a degree never achieved in the past. The story of the peoples of Vietnam is a rich one that invites renewed understanding.

1. Introduction

The multi-ethnic nation-state of Vietnam with 53 official minority nationalities (plus Vietnamese) has been transfigured by a turbulent history of invasions, migrations, demic expansions, and the rise and fall of cultures. The result of these forces is a populace divided into: (i) a majority ethnicity, the Kinh or ethnic Vietnamese, who expanded from a traditional geographic base in the Red River Basin, and (ii) a large number of minority ethnicities, with smaller populations, homelands in highland enclaves, less developed economies, and slowly changing cultures.

The rise of the Mon Khmer (MK) Vietnamese majority can be tracked back with cladistic study of their DNA and the names of plant and animal species to the remote Austro-Asiatics (AA), the ancestor of the MK and ultimately the Vietic branch populations (Diffloth 2005). For instance, by studying the genetic material from contemporary Vietnamese, we can determine from what place and at what time they arrived in Vietnam (their origin and development, i.e. their biogenesis). Archeology informs us about the development and evolution of Vietnamese culture (culturo-genesis) and linguistic reconstruction (linguo-genesis) can uncover language
structures from past times. All these fields use contemporary data to tell us about things long past (Bellwood 2006) (see Note 1 for more details about cladistic, phylogeographic, and archeological studies). In the case of many minority peoples of Vietnam genetic and artifactual research is still quite incomplete.

The AA peoples of Southeast Asia, as determined from studies of India’s tribal populations, are descended from a common male line originating in the Mundari peoples, who are thought to be the oldest population group of India and first AA’s (Kumar et al. 2007). Some of the Mundari men, perhaps on hunting or war expeditions, intermarried with women of contiguous areas of Southeast Asia in contemporary Myanmar. Links from these pioneers to the contemporary Vietnamese using phylogeographic analysis were established by Ballinger et al. (1992), who has shown that the Vietnamese populace possesses – of all Southeast Asian people groupings – the greatest genetic diversity in its maternally transmitted mitochondrial DNA (mtDNA). This diversity is the product of regular but random mutations over the last 25,000 years obligatorily passed on from mothers to sons and daughters up to contemporary generations. Thus, these mtDNA patterns determine a set of distinctive genetic ‘fingerprints’ that fixes the source and arrival times of the proto-populations of Vietnam from MK forbearers. Ballinger et al. (1992: 5) observed that the great diversity among the Vietnamese is mirrored in great diversity of the aboriginal populations of Malaysia, including certain modern MK-speaking aboriginals, for example, the Temiar and Semai. Thus, the pre-Vietic and the Orang Asli people of Malaysia appear to be among the oldest within Southeast Asian MK peoples and may have constituted a physical population unity in the deep past of the Malay Peninsula (cf. also www3.nationalgeographic.com/genographic, specifically the mtDNA haplogroup F).

So following this proposed model, perhaps 17 millennia BP people with these shared genetic features moved up the Malay Peninsula as far as northern Vietnam near Hòa Bình to dwell in rock shelters and caves where the French archeologist Madeleine Colani first discovered this Hoabinhian Culture of 15,000–5000 BP and noted that its unique features extended throughout the region (Colani 1927). Later these populations mixed with the inhabitants of Bạc Sơn near the Red River in Bắc Thái and Lạng Sơn Provinces. Bacsonians possessed tools that were made of ground and polished stone, a significant improvement over Hòa Bình tools and pottery. During this same period, there was another culture called Quỳnh Vân, whose inhabitants occupied the coast of north-central Vietnam, largely living from fishing (cf. http://www.bvom.com/resource/vn_history.asp?pContent=Pre-History). They were followed in the Late Neolithic and Early Bronze Age by people of the Phùng Nguyên Culture of Vietnam, who were wet rice farmers and who lived in sedentary villages on the Hòng (Red), Ðà (Black), and Lô Rivers 2000–
1500 BC. In further triangulation of the peopling of Vietnam, we know from linguistic comparison that pre-Vietics also have a connection to groups of north-central Vietnam found today in Nghệ An, Hà Tĩnh, and Quảng Bình Provinces near the Lao border where some Vietic people still live. Perhaps some 4500 years ago a portion of these Vietic people groups appear to have arrived in the Red River Valley, and prospered. Later still they established a joint kingdom – known in Chinese records as Âu Lạc – at Cổ Loa near Hanoi together with the indigenous Kra-Dai (Tai-Kadai) peoples (cf. Section 2). Li Hui et al. (2006) have analyzed comparatively palm prints, fingerprints, and sole prints between Kra-Dai and MK and decide for an MK genetic heritage.

But the reign of the kingdom of Âu Lạc was destined to be short-lived. In 221 BC an invasion force from the north under General Zhao Tuó with an army of 100,000 took northern Vietnam in the name of the emperor of China, and it remained a province of the Middle Kingdom until AD 939 with much mixing of genetic, linguistic, and cultural features. Then subsequently, free from the Chinese yoke, Vietnam prospered under the institutions and a written tradition it had inherited and was able to fend off periodic attacks from Chinese authority. In AD 1471 they succeeded in a final conquest of the kingdom of Chăm-speaking people of Vietnam. As a result they enlarged their control to the present boundaries of the country.

The large-scale dispersal of the Vietnamese language southward resulted in the gradual retreat of smaller languages into highland areas or into ones that had been traditionally occupied. And, there their speakers remain and make up about 13 per cent of the total population of contemporary Vietnam.

The territorial topography is thus divided into linguistic expansions and linguistic contractions back into heartland areas. Colin Renfrew (2000) called the two ‘linguistic spread zones’ and ‘linguistic mosaic zones’. To explain the catalysts for such vast changes of demography and social structure, Bellwood and Renfrew have proposed the **Dispersal/Farming Hypothesis**, according to which demic expansion ensues from changes such as the transition to agrarianism (Bellwood and Renfrew 2002; Bellwood 2005). Others suggest that beyond better nutrition additional powerful developments may be in play, for example, disease, wars, climatic changes, changes of hunting conditions, new technology (spear points, bows, and weapons of war), or political consolidations.

The linguistic spread zones of Vietnam are today occupied by the Kinh, while the mosaic areas of Vietnam are the territories along the China and Lao borders, the Central Highlands and, as well, territory in Nghệ An and Thanh Hóa Provinces and surrounding provinces, where one finds groups who speak languages of the Kra-Dai (Tai-Kadai), MK, Hmong-Mien (Miao-Yao), Austronesian, and Sino-Tibetan stocks. There is good
reason to believe that the fertile Red, Black, and Lô river valleys produced higher population densities, greater infant survival, and longer life expectancy. The result was a need for new lands. The dispersal of the Vietnamese to the south that followed involved the significant physical movement of people with some cultural assimilation of other ethnicities (see Edmund Leach 1954). The most powerful argument favoring population dispersal as opposed to language and cultural shift is the distance involved. From Hanoi to Saigon it is more than 1800 km and it would have been difficult to maintain momentum for the hardships of such a military campaign without social cohesion, without the driving force of kinship and cultural proximity behind it (Bellwood 2006: 100). Thus, from the fifteenth century onward the Vietnamese populations occupied the territory from the Red River Valley to the mouths of the Mekong across a variety of landscapes, while speaking a common, though not completely homogeneous, but varying language that served as the vehicle of government and education, whereas the non-Kinh populations were pressed into small enclaves, with few major cities, more sluggish economic development, cultural and linguistic norms changing much more slowly, and bilingualism needed to participate in the larger Vietnamese world.

Bilingualism was also fostered in part by the typological similarity of grammatical features among the languages of Vietnam. Many of the languages possess isolating monosyllabic typology with SVO and head–modifier word order. Nearly all are classifier languages. The Kra stock is unique in having a sentence final-negative word. Tibeto-Burman languages demonstrate SOV word order. The idea of a Sprachbund or zone of syntactic convergence is strongly manifest over many landscapes, while phonetic/phonological as well as lexical differences have been less affected (Figure 1).

2. Kra-Dai (Also Called Tai-Kadai)

This grouping has a total population of several million speakers. They are the descendants of what the Chinese records called the Bác Việt, a group whom genetic evidence suggests came into Myanmar, and then Yunnan Province, China, from India about 20,000 BP (cf. Figure 2 and Su et al., 1999 and, especially, Li 2002a,b, 2004). They then veered south into Thailand and Laos, looping back north into Vietnam and then advanced into Chinese territory settling along the border and continuing up to the mouth of the Yangtze River near Shanghai 8–10,000 BP.2

They are precursor people of such familiar groups as the Thai, Lao, Shan of Myanmar, and all the Kra-Dai people of Vietnam. Many of them live in northern Vietnam near the China border, belong to the Tai stock of languages and, are today agriculturalist practicing wet paddy rice cultivation (Figure 2).
2.1. THE TÀY

One of the most populous of Tai minority groups of Vietnam are the Tày (population 1,477,514), an autochthonous group dwelling very near the China border and extending south and west from their heartland in Cao Bằng Province (Figure 1). This group was formerly called Thọ, now considered pejorative, meaning ‘landed/local people’. For its size this group is relatively understudied.

As a sample of language diversity, the Tày language can be divided into four geographic areas by linguistic features as well as by their customs and practices: Area 1 includes northern Cao Bằng Province (our data were gathered from Trùng Khánh and Nguyên Bình Districts) and preserves original features such as voiced initial consonants from *b-, *d-, *g-, as well as, *r-, and *hr- (where * indicates historical or ancestral sounds); Area 2 comprises central northern Vietnam, southern Cao Bằng, northern Bác Thái, as well as, northern Lạng Sơn; Area 3 encompasses southern Lạng Sơn and southern Bác Thái (data from Bác Giang); and finally, Area 4 represents a Tày expansion into Hà Giang and Tuyên Quang Provinces to the west (our data are from Bát Xát District of Cao Bằng Province). Table 1 provides some distinguishing linguistic features of the four areas (tones indicated with raised numerals, Li 1977, cf. also Hoàng 1997).

As Table 1 shows, Tày at all locations preserves the Tai proto diphthongs *ie, *uo, and *ua, as in küiə1 ‘salt’. Also Area 1 preserves the original voiced consonant initials when these occurred in low tones, for example, ‘thatch grass’ yə1 < *y and ‘river, wharf’ də2 < *d.
Path of Migration and Division of Some of the Major Ethnicities with their genetically distinctive markers

Fig. 2. Migration pathways of Chinese (Han), Tibeto-Burman, Hmong-Mien (Miao-Yao), Kam-Tai based upon phylogeographic study of the Y chromosome polymorphisms of contemporary speakers (see Su et al. 1999, Li 2002a,b, 2004; and https://www3.nationalgeographic.com/genographic/atlas.html (cf. especially the genetic markers M119 and M122.)

Table 1. Comparative table of Tày vocabulary showing the original and more archaic forms of Area 1 and the mutated forms, especially of Areas 3 and 4, which are presumably the result of population movements to the South (3) and West (4).

<table>
<thead>
<tr>
<th>River</th>
<th>Scold</th>
<th>Knife</th>
<th>Nest</th>
<th>Eye</th>
<th>House</th>
<th>Sky</th>
<th>Garden</th>
<th>Stone</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ᵁdaᵇ₂</td>
<td>ᵇjaᶜ₂</td>
<td>ᵇaⁿᵃ₂</td>
<td>ᵇaⁿʰᵃ¹</td>
<td>ᵇaⁿᵃ₁</td>
<td>ᵇaⁿʰᵃ₁</td>
<td>ᵇaⁿʰᵃ₁</td>
<td>ᵇaⁿʰᵃ₁</td>
</tr>
<tr>
<td>2</td>
<td>ᵁdaᵇ₁</td>
<td>ᵇjaᶜ₂</td>
<td>ᵇaⁿᵃ₂</td>
<td>ᵇaⁿʰᵃ¹</td>
<td>ᵇaⁿᵃ₁</td>
<td>ᵇaⁿʰᵃ₁</td>
<td>ᵇaⁿʰᵃ₁</td>
<td>ᵇaⁿʰᵃ₁</td>
</tr>
<tr>
<td>3</td>
<td>ᵁdaᵇ₁</td>
<td>ᵇjaᶜ₂</td>
<td>ᵇaⁿᵃ₂</td>
<td>ᵇaⁿʰᵃ¹</td>
<td>ᵇaⁿᵃ₁</td>
<td>ᵇaⁿʰᵃ₁</td>
<td>ᵇaⁿʰᵃ₁</td>
<td>ᵇaⁿʰᵃ₁</td>
</tr>
<tr>
<td>4</td>
<td>ᵁdaᵇ₁</td>
<td>ᵇjaᶜ₂</td>
<td>ᵇaⁿᵃ₂</td>
<td>ᵇaⁿʰᵃ¹</td>
<td>ᵇaⁿᵃ₁</td>
<td>ᵇaⁿʰᵃ₁</td>
<td>ᵇaⁿʰᵃ₁</td>
<td>ᵇaⁿʰᵃ₁</td>
</tr>
</tbody>
</table>
2.2. NÜNG

The Nùng (population 856,412) are linguistically close and live geographically adjacent to the Tày but are historically and culturally distinct (Figure 1). The Nùng (except for the Nùng Châu) are thought to be immigrants from territory across the Sino-Vietnamese border who arrived after the defeat of the Tày rebel Nùng Trí Cao (AD 1025–1055). Unlike Tày, Nùng does not preserve the proto-Tai diphthongs (cf. the systematic contrast illustrated by Tày rɯan\textsuperscript{42} ‘house’ in opposition to Nùng Châu Ɂən\textsuperscript{42}, which reflects simple monophthongs for many cognate form).

The Nùng refer to themselves generally as [nuŋ], followed by an attribute signifying the original homeland (in China) or features of their women’s dress. According to Các dân tộc ít người ở Việt Nam (1978: 201) (Minority peoples of Vietnam, the northern provinces), the Nùng Châu (Nùng of Lóngzhōu Guāngxī originally extended into China territory just east of their current Vietnamese areas in Lăng Sơn Province), Nùng Lôi (Nùng of Xiàléizhōu, stemmed from the territory abutting on Vietnam just about half way between Jingxī and Dàxìn Counties), Nùng Inh (Nùng Longying, were once located north of Lóngzhōu and to the west of Tàiponge Fù); other types Nùng Dín, Nùng An, Nùng Phán Sình, Nùng Quý Rǐn, Nùng Xuòng (Nùng Zhuang), and Nùng Giang (Nùng Yang) also are so-named from original location or dress. In fact, Các dân tộc ít người ở Việt Nam (1978: 201) reports that there are fifteen different subgroups among the Nùng, many corresponding to the Zhuang ethnicity of South Guāngxī, China.

When Vietnam became independent of China, there were also changes in the balance of power among the Tai groups along the frontier. The Nùng, in particular, had assumed control of Yŏng Prefecture located southwest of contemporary Nánning bordering Vietnam. In AD 1053 Nùng Trí Cao attacked Chinese outposts and repelled the Vietnamese sent to arrest him. Finally, after several failed plots by the Chinese to assassinate him, Nùng and his friends burned their own villages one night, saying to the masses that Heaven had destroyed everything they possessed and thus they were all sworn to the death to rebel. So it was that they began the greatest Tai insurrection of all time. His Tai army marched on Nánning, Guāngxī’s seat of government, and moved east pillaging and putting officials to the sword. After reaching Guāngzhōu (Canton) the Chinese and Vietnamese jointly allied against him. He was defeated and harsh terms of settlement were imposed.

In the sixteenth century the descendants of Nùng Trí Cao’s original followers sought relief through immigration into Vietnam (Lajonquière 1906: 186; Abadie 1924: 77). The exodus continued to a lesser extent up to the twentieth century. Once in Vietnamese territory, the Nùng encountered the Tày, already entrenched on the best lands. Nevertheless, the two groups were said to have made contracts that allowed the Nùng title to the hills around the Tày and to paddies of lesser value near the pre-existing Tày.
settlements. The industrious Nùng, however, had soon turned this formerly less valuable land into bountiful and productive paddies. This history explains why the Nùng and the Tày are still unlike; as Tày communities were aboriginal, they were established at lower altitude, and their practices and clothing were and are much more uniform place-to-place.  

2.3. THE MILLION-STRONG THÁI OF THE SOUTHWEST TAI STOCK

The million-strong Thái of the Southwest Tai stock (near relatives of the Siamese and Lao) are also immigrants into northwestern Vietnam beginning prehistorically and lasting until AD 1600. They are found today in Lai Châu, Sơn La, and in older settlement areas such as Thanh Hóa and Nghệ An Provinces. Black and White Thái of Lai Châu have six tones, whereas the kinds in Thanh Hóa and Nghệ An have five.

There are several types of Thái in Vietnam and these groups have been called by a variety of names: the Thái Dam (Black Thái/Tai) (population 699,000), Thái Don (White Thái/Tai) (population 280,000), and Thái Daeng/Deng (Red Thái/Tai). There is dispute about the last group, the Red Thái. Some Vietnamese ethnologists, notably Vi Văn An and Cấn Trong, have argued that the name Thái Daeng for groups along Highway 7 of Thanh Hóa and Nghệ An Provinces is only a local name and in reality the people of this area are ancient White Thái immigrants from Bắc Hà District of Lào Cai Province (cf. http://www.columbia.edu/cu/china/Vi.html).

2.4. THE SMALLER KRA-DAI GROUPS

Smaller Kra-Dai groups include Sán Chay, also called Cao Lan-Sán Chí (population 114,000), probably a Central Tai language showing contact with Han Chinese and possibly Yao, Bố Y, Già, Tú Di (nineteenth century Bouyei immigrants from Guizhōu, China, affiliated with the North Tai stock (total population 40,000), and the Southwest Tai languages Lừ (population 1400), Lào (population 9600), in Lào Cai Province Tây Sa Pa, and Pa Di (population 200). Little is known about these groups in Vietnam.

2.5. THE KAM–SUI GROUPS

The Kam and Shui of Vietnam (called ‘Đồng’ and ‘Thủy’) are highly endangered with one or two speakers and 120 speakers, respectively; they live in Tuyên Quang Province where they migrated about 150 years ago as a consequence of the Taiping Rebellion in China.

2.6. THE KRA GROUPS

There are also several people groups living in Vietnam and southwestern China with small populations who speak Kra languages: La Chí (population
2.7. LINGUISTIC FEATURES OF THE KRA-DAI GROUPS

Linguistic features of Kra-Dai groups are generally six tonal categories in open syllables for Tai and Kam-Sui speech forms, Kra languages have important archaic lexical and phonological features. La Ha has two series of voice stops, \( /\breve{\text{q}}-\text{\breve{d}}-/ \) contrasting with \( /\text{b}-\text{d}/ \) and \( /-\text{l}/ \) (Figure 4).

3. Hmông-Mien (Miao-Yao)

In Vietnam languages from this stock were formerly generically called Mèo, but are known today as Mông or Hmông (Hmong) (population 787,604), Dao [zau] (Yao), and Pa-hng (Pà Thén) (population 3700). The Hmông-Mien groups in their pre-history came with the precursors of the Han and Tibeto-Burman peoples to Yùnnán Province about 20,000 years ago (Figure 2). On the journey northward the Hmong-Mien branched off at the headwaters of the Yangtze and traveled upriver. They are thought to be linked to neolithic societies including the Dàxì Culture 大溪文化 (5300–6400 BP) and the Qūjiālíng Culture 屈家嶺文化 (4600–5000 BP), who settled in China along the middle reaches of the
Yangtze River (for which genetic evidence has just been provided in Li Hui et al. 2007, also Edmondson 2007).

3.1. THE HMÔNG OR MÔNG

The Hmông are located in Hà Giang, Lai Châu, Lào Cai, Sơn La, Yên Bái, Tuyên Quang, Bắc Thái, Cao Bằng, Thanh Hóa, and Nghệ An Provinces, whereas the Dao are also found along the China and Lao borders. Non-specialists may be tempted to confuse Mông with Mùô, a large Vietic MK group settled near Hòa Bình.

All Hmông groups of Vietnam belong to the ‘Western’ or ‘Guìzhōu–Yúnnán–Sìchuan’ types save for the Ná Miao [na³miao³⁵] (population 1200), who live in Cao Bằng and Lạng Sơn Provinces (near Thạch An). Based on linguistic comparison their language seems to be affiliated with the Central Hmongic grouping with origin near 梧州 Liúzhōu of Guǎngxī Province.

Different investigators have reported different subgroups of Hmông in Vietnam. The information here is from Professor Nguyễn Văn Lợi of the Linguistic Institute of Vietnam, who spent several years working in Hmông villages (cf. also Jacques Lemoine 1995).

First, there are the Hmông Lênh (Striped Hmông); they are also very similar to the Hmông Ntsjua (Flowery or Green Hmông) and Hmông Pua. The women of these kinds of Hmông wear skirts with green background overlaid with white embroidery work. The Hmông Lênh inhabit the territory far from the Chinese border and are also located in Lào Cai.

The second group of Hmông in Vietnam is the Hmông Daw or (in Vietnamese) Mèo Tráng, meaning ‘White Hmông’. Note that this group is not to be confused with the Dao [zau] or Yao, who represents a sister group to the Hmông as a whole. The Hmông Daw live mostly in Hà Giang Province in the districts of Mèo Văc, Động Văn, Yên Minh, and Quàn Bà, as well as in some districts of Lào Cai, Sơn La, Thanh Hóa, and Nghệ An. This group has perhaps the largest population among the Hmông of Vietnam. These were the first Hmông group to come to Vietnam from China starting about 300 years ago. From China they first moved to Mèo Văc; a part of them then followed the border going to Lào Cai and there they then moved south into Thanh Hóa and Nghệ An. Hmông custom says that spirits of the dead must first travel to Mèo Văc before they can pass on to their ultimate resting places in China. In China we can find this variety of Hmông in eastern Wénshān near to the border with Vietnam. They are traditionally called White Hmông because the women wear white skirts without embroidered patterns.

The third kind of Hmông in Vietnam is the Hmông Du or Black Hmông. The population is relatively small. They live mostly in Hòa Bình, Ngân Bái, and some in Sơn La, and Thanh Hóa all further to the South. This kind is also found in China south of Sìchuān on the border with Guìzhōu as well as in small groups to the southwest of Wénshān of Yúnnán Province near to the border with Vietnam. The Hmông Du often wear black trousers.

A fourth variety is the Hmông Si, also known as Red Hmông – it is possible that the Hmông Pe have also started wearing this costume although they belong to the first variety (cf. above). The Women often wear headgear with red balls bound with wool twine into the hair. But Si means ‘light’ as they find other kinds of Hmông language to be ‘heavy’ in pronunciation. They have only a small population and are mixed with Hmông Lềnh and live far from the China border.

A fifth subtype of Hmông in Vietnam is the Hmông Súa, also called the Chinese Hmông. They are found in only three villages of Hà Giang Province, at Thượng Lan in the district Bắc Mê and also in Lào Cai and Lai Châu. We have heard that there are also some families in Yên Bai and
Quàn Bạ of Hà Giang. There are Hmông Súa west of Guiyáng near Huìshuǐ near Guiyáng. The Hmông Súa wear a skirt with an apron with black piping. The men kept the custom of wearing a pigtail in the fashion of the Han during the Qing Dynasty, as recently as 1971.


3.2. THE DAO

The Dao or Yao groups in Vietnam fall into two subtypes: Mienic (autonym Yu Mien, population 350,000) and Munic (autonym Kim Mun, population 170,000). The Yao are also found in large numbers in China having a population there of more than 2.1 million. There is a large variety of cultural and linguistic subgroups among the Yu Mien and Kim Mun, who also arrived in Vietnam from the thirteenth to the twentieth century by different routes and who then settled in different areas including lowland plains to mountainous areas. They also differ in many aspects of dress, custom, and practice. Notably, many of the Mienic groups make use of Chinese characters to record important things. The main linguistic difference between Mien and Mun has to do with the presences of clusters. Oversimplifying considerably, Mun preserves the complex initials kl-/kj-, whereas in Mien these have become palatalized to the affricates ts/-/tc (cf. ‘road’ in Yu Mien tsau but in Kim Mun of Lao Cai kjau). The Yu Mien spoken at Ogang in Bác Thái Province is very close to many kinds spoken in Guǎngxī Province, China, so it is thought that this group must have come more recently. The pattern of variation within Yao is very simple in comparison to the very complex internal structure of variation within the Miao (Hmông) language, especially the western types. We were able to study Kim Mun at two locations, Lào Cai at Bát Xát District, where we were told that there are six Kim Mun villages and also in Hà Giang Province at Na Khê Village of Yên Minh District. The Kim Mun of Lào Cai accords with the rest of the Yao branch by preserving most of the consonant stop codas at the end of syllables, except for final -k, which has been lost. Kim Mun is also known for its rich system of tonal developments and these developments yield a secondary series of tones often accompanied by voice quality differences.

3.3. PÀ THÊN

Pà Thênh, Pa-hng, or 巴咩 Bá-hêng [pa31 ŋ35] (population 5569) are immigrants from northeastern Guǎngxī Province where they are found in Sānjiāng County and adjoining areas and in Guǐzhōu and Húnán
Provinces. They were often called Red Yao in past literature. Despite this exonym, their language split off early from general Hmongic and innovated considerably showing features such as syllabic nasals, breathy vocal register in many of the low tones, and uvular initials.

4. Mon-Khmer

The MK stock along with the Munda(ri) of India makes up what Pater Wilhelm Schmidt (1906) called the AA language family. MK includes not only such widely known languages as Vietnamese (Việt) and Khmer (Cambodian), but also a large number of other lesser known groups in Vietnam.

4.1. Vietic Subgroup

The most populous MK group in Vietnam are the ethnic Vietnamese/Việt (65,795,718), living everywhere in the country and the Mường (1,137,515), who call themselves Mol or Mô and are found in Hòa Bình and Thanh Hóa Provinces in northern Vietnam. The traditional (over-simplified) characterization of Vietnamese dialectology distinguishes northern (especially Hanoi), central and southern varieties. Recently, Nguyễn (1995) and Alves (2007) have noted the further existence of a special north-central regional area differing from all three classically defined areas. The Central variety spoken in Huế has long been known as a very special type and there are a number of further locations that still need careful documentation to round out the Vietnamese dialect picture.

An early view on the classification of Vietnamese grouped it as a Tai-related language, sharing certain words and phonological structure, especially its tone system (Li et al. 2007). A further proposal even posited an affinity with Tibeto-Burman. Yet, other suggestions that Vietnamese could be related to Khmer were dismissed on the grounds that Khmer’s lack of tone would constitute a ‘... chasm to broad ...’ to allow any genetic relationship. In 1954, however, the French scholar, André Haudricourt, broke new ground by demonstrating how a tone language like Vietnamese arose within MK, most of whose languages are non-tonal in nature. This theory of ‘tonogenesis’ (as Matisoff 1973 has called it) was shown to predict each of the six Vietnamese tones based on whether MK cognates of Vietnamese forms possessed an original voiced or voiceless initial consonant and also depending on the nature of the original final of the word. From an orthographic perspective Vietnamese represents the oldest example in mainland Southeast Asia of the use of a romanized writing system. While a traditional Chinese-like demotic script called chữ Nôm ‘southern script’ was used by an elite minority, the roman Quốc Ngữ ‘national script’ has over more than three centuries achieved the status of an accepted orthography, which has helped contribute to an almost universal literacy among the ethnic Vietnamese (Kinh). This almost
unique alphabet in Mainland Asia results from the seventeenth-century work of Alexander De Rhodes who codified romanizations that had been in use by Roman Catholic padres for a number of decades earlier (Gregerson 1969). Spellings like gi- come from Italian conventions, nh- from Portuguese, and tone diacritics from European classical languages.

Smaller Vietic groups include the: Thô (68,394) and Hung (700) in Nghệ An Province as well as the Chúrt (3829) in Quảng Bình parallel with Mu Gia Pass, Quồn (2000?) and Arem (20) also in Quảng Bình Province, and the Maleng (200) in Quảng Bình and Hà Tĩnh Provinces. Some classifications use the form Chúrt (which means ‘mountain’) to designate the entire grouping of the smaller Vietic languages. The smaller Vietic groups above are historically interesting because they reflect simplified prosodic systems that stand as precursors to the more tonally elaborated five and six systems of Vietnamese and Mường. We note that there is recent unpublished fieldwork on Karìì, a newly discovered language of the Vietic group in Khammouan Province, Laos. Karìì has a large number of vowels distinguished by vocal registers and length, has weak–strong stress in the word structure, and its final consonants system is richer than many less archaic languages (Enfield and Diffloth forthcoming).

4.2. KATUIC

E.Bru (population 55,559), Pacoh (population 16,000), Phuong (population 15,112) and E.Katu (population 50,458), and Taoih (population 19,000) are all found in the highland provinces of Central Vietnam. Diffloth (2005) groups Katuic as the closest other MK branch to Vietic.

4.3. KHMERIC

A large population (1,055,174) of Central Khmer speakers is to be found primarily in the southern delta of Vietnam. Their speech is basically the same as that spoken across the border in Cambodia, where it is the national language.

4.4. BAHNARIC

North Bahnaric (NB) groups living in the central highlands of southern Vietnam include the Sedang (population 100,648), Todrah (population 9142), Monam (population 5000?), Katua (population 3000), Kayong (population 2000?), Hre (population 113,111), Jeh (population 15,243), Halang (population 13,500), Halang Doan (population 2000), Rengao (population 16,000), as well as tentatively ?Ramam (population 250) and ?Trieng (population 15,000). A hallmark of NB languages generally is their phonological register systems in which all words in these languages are pronounced with either a ‘breathy’ vs. ‘clear’ or ‘plain’ vs ‘tense’ phonation articulation (Smith 1972).
Cua (population 27,766) is a Northeastern Bahnaric language and Bahnar (Bana) (population 158,456), whose name has been recruited to label this entire MK subgroup, has itself been classified as Central Bahnaric. Living farther to the south of the other Bahnaric groups in southern Vietnam are speech communities classified as south Bahnaric, which include the Central Mnong (population 32,451), Eastern Mnong (population 30,000), southern Mnong (population 30,000), Stieng (Budeh, Bulo; population 50,000+?), Chrau (population 22,567), Koho (population 128,729), and Maa (population 33,338). This South Bahnaric grouping stands geographically separated from other Bahnaric groups to the north by a major east–west incursion of two large Chamic languages, namely, Rade and Jarai, into the highlands.

4.5. KHMUIC

Khmu (population 56,542), O’du/Iduh (population 301), Puoc (population 18,018), Khao (population 10,000?), and Kháng (population 4000). Interestingly, Kháng, found in Sơn La and Lai Châu Provinces, is a tonal language (five tones), while other Khmu dialects in Southeast Asia have breathy vs. tense register (like North Bahnaric languages above) and still others reflect no systemic prosody contrasts at all in their words.

4.6. OTHER [NORTHERN (?) MK]

Mảng (2663) is an MK language of uncertain classification located in Lai Châu Province in northern Vietnam (and in China).

5. Austronesian

The Austronesian language family is represented in Vietnam by the Chamic subgroup, which for a millennium as the kingdom of Champa controlled all or portions of the (pre-Vietnamese) south of Vietnam. In the twelfth century Champa attacked the Khmer from Tonle Sap lake, and sacked royal Angkor. There were earlier defeats at the hands of the Vietnamese in AD 982 in central Vietnam (Indrapura) and finally in AD 1471 in southern Vietnam (Vijaya) ultimately reduced Chăm populations to enclaves in the south such as those in Ninh Thuận around Phan Rang and Bình Định at Phan Thiết along the coast as well as An Giang and Tây Ninh to the west. It is thought from their ceramics and iron utensils that the Chăm parent groups arrived from Borneo in the first or second century BC. That hypothesis sees them as an end-point of a larger southward Austronesian migration beginning a number of millennia ago in southern China via Taiwan, the Philippines, and Borneo – and thence to the coasts of mainland Southeast Asia (Bellwood 1978: 85–8) (Figure 5).
Fig. 5. Southern Vietnam showing locations of many Austronesian and MK languages (from Gordon 2005, editor of Ethnologue.com).
5.1. THE HIGHLAND CHAMIC SPEECH COMMUNITIES

The Highland Chamic vs. Lowland Chamic speech communities include the Rade (population 270,348), Jarai (population 317,557), and Haroi (population 35,000). These languages, especially Rade and Jarai, as noted earlier, constitute a geographical phalanx that has expanded in a westerly direction into the highlands centering around Banmethuot on the south and Pleiku/Kontum on the north. This intrusion has separated the South Bahnaric group of MK languages from Bahnar itself further north.

5.2. COASTAL CHAMIC LANGUAGES

Coastal Chamic languages include Eastern Chăm (72,873), Western Chăm (25,000), Chru (14,978), Northern Roglai (52,931), Southern Roglai (41,000), and Cac Gia Roglai (3000). It has been suggested (Thurgood 1999) that defeats in battles with the Vietnamese motivated Chăm migration as far as Aceh (Sumatra). This view, or at least its timeline, has been contested (see, for example, Sidwell 2006). It is known that the seafaring Châms also colonized certain areas on Hainan island off the coast of China, where, in contact with neighboring tone languages, Chăm, called there Utsat, also developed a five tone system (Benedict 1941, 1984). Indeed, Chamic constitutes a kind of ‘language lab’ for studying prosodic evolution, for while the Eastern Chăm (Phan Rang) may well have the makings of a pitch-oriented (tonal?) system; in contrast, their Western Chăm brethren in Vietnam and in Cambodia (and possibly Haroi too) have developed more of a breathy vs. tense register system reminiscent of the phonology of Khmer. See Pittman (1959), Lee (1966), Blood (1962), Thurgood (1999) for further linguistic history of the Chamic group.

6. Sino-Tibetan

The Sino-Tibetan groups of Vietnam all belong to one of two major subgroupings: Sinitic or Tibeto-Burman.

6.1. THE SINITIC GROUPS

The Sinitic group according to Vietnamese classification are divided into the Chinese Vietnamese (Hoa in Vietnamese and 華 húa in Chinese), who, as a whole, have swelled to become the largest minority group in all of Vietnam with a reported population of 2.6 million. Most of these people live in Chợ Lớn 堤岸 di’àn, in Standard Chinese, which is the Chinese area of greater Hồ Chí Minh City/formerly Sài Gòn. The Hoa people groups include the descendents of the eighteenth to twentieth century persecuted immigrants from Guǎngdōng Province in China and

The Ngái of Quảng Ninh Province are perhaps the same as the Danka or sea islanders of Fujian, Quảngdong, and Guǎngxī Provinces, or塞家 dânjia in Standard Chinese and dan6ga1 in Cantonese; also called 水上人家 seis2 seung5 yan4 or water dwelling people. Today their population in Vietnam is thought to number about 1430 with three subgroups. They live on small watercraft and subsist by fishing or pearl diving. According to their legends they are descendents of an 东晋朝 Eastern Jin rebel army in AD 399–411. Their language is thought to be a form of Cantonese sometimes with unusual vocabulary, although other scholars have suggested that the Ngái originated from minorities of the Southeast China coast, perhaps the Hlai or 黎 lí of Hainan island.

Sán Diu are also immigrants from Guǎngdong Province in the sixteenth century and speak a more archaic kind of Cantonese than the Hoa. They were said to number 114,000 in 2000. Today then live in Quảng Ninh, Hải Hưng, Hà Bắc, Bắc Thái, Vĩnh Phúc, Hà Giang, Tuyên Quang, and Thanh Hóa Provinces.

Details on the Sán Diu or Ngái are largely lacking, but we suspect that there might be influence from an archaic form of Chinese resembling Cantonese but found in Guǎngxī Province called 平话 Pinghua, which is thought to be derived from the vernacular forms of speech spoken by the garrison troops first sent to Guǎngxī and Vietnam in the second century BC.

### 6.2. TIBETO-BURMAN GROUPS

In Vietnam the Tibeto-Burman groups are all speakers of the Lolo-Burmese subgroup and include the Hà Nhi, Sila, Công, Phù Lá, Lô Lô, and La Hủ. These languages all have monosyllabic isolating word structure with SOV word order.

#### 6.2.1. The Lô Lô Group

The Lô Lô (population about 1500) are found in the northeastern sections of Hà Giang and western Cao Bằng. There are, in fact, three different forms of this language: (i) the Black Lô Lô, (ii) the Red Lô Lô, and (iii) the Flowery Lô Lô, with these names taken again from the clothing of their women. There has been until recently little information about this language group. The Lô Lô of Vietnam have uvular stops and are allied with groups just across the border in China whose languages are very similar (cf. Edmondson 2005).

#### 6.2.2. The Phù Lá

The Phù Lá (population 9046), though they count as a single nationality, comprises two languages and cultures, Phù Lá (autonym Phu Kh(ŋ)a)
and Xá Pho. Both dwell in Lào Cai Province near Sa Pa District. More information about them is found in Edmondson (2005).

6.2.3. Other Tibeto-Burman Groups
The Sìla (population 840), Côông (population 2000), Hà Nhi (population 17,535), and La Hà (population 6874) all live in the extreme northwest of Vietnam in Lai Châu Province, many of them in the Mùông Tế District on the Black River. All four belong to the central or southern subgroupings of the Loloish language. They live very close to the China border. For more discussion, see Edmondson (2003, 2005).

The Côông language had been regarded as a variety of Phunoi mostly spoken in Laos (Bradley 1979: 68). However, from our data Côông seems more different from Phunoi than expected. It has independent tone on all syllables and no minor syllable in front of a major syllable; it has no voiceless nasal consonants and no final oral stops, only /-m -n -ŋ/. As with a number of the lesser-known languages of Vietnam, we look for further cooperation between Vietnamese and international scholars to refine our understanding of Côông.

7. Conclusion
The rich fabric of languages and cultures of Vietnam has been woven by the many processes of history that have been sketched above. From the deep past and the paths of its genetic trajectories through later prehistoric beach heads laid bare at renowned archeological sites, the texture of this country and its people have received their underlying form. Further successive waves of civilizations sweeping as they have over the region have left tell-tale traces of ancient India and China, as well as of Islam and the West thus lending color and character to the nation.

Appreciating the effects of this variegated past will occupy Vietnamese and international scholarship for many generations hence. And what of the future? How will all these strands come together in the years ahead? After more than a century of conflict and turmoil, can majority culture members reach out to the easily forgotten and marginalized minority peoples in a way that bespeaks mutual respect and genuine good will towards their interests? In an era of rapidly developing nation states with their new-found mandates to deliver education to all, provide communication infrastructure to the farthest provinces, and support a means of livelihood for each citizen – in this context, can Vietnam distinguish itself as a model for the region and even the world? The sociolinguistic complexity of this challenge is daunting, but, like rays of light, success stories here and there in the region do, indeed, give hope in the days ahead for a developing and dynamic multilingual, multicultural Vietnam in our times.
Acknowledgements

The authors wish to acknowledge the support of the National Endowment for the Humanities with grant NEH RT-21754-05 for 1995 and the National Science Foundation with grants SBR 9511285 for 1995 and SBR9729043 for 1998 for the research reported on here under the name ‘Languages of the Vietnam–China Borderlands’. For the results of the fieldwork readers can also consult our website http://ling.uta.edu:16080/~jerry/research/. Thanks also go to many speakers of those languages we studied as well as to Professor Lý Tấn Thắng and Professor Nguyễn Văn Lợi of the Linguistic Institute of Vietnam who helped us greatly in arranging our fieldwork in Vietnam.

Short Biography

Jerold A. Edmondson researches the minority people of East Asia, Southeast Asia, East Africa, and Mexico, specializing in tone and voice quality phonetics. He co-edited two books on Kra Dai (Kadai) languages with David P. Solnit. He is also interested in using the laryngoscope for studying the very special sounds of the lower throat, see his 2006 paper with John H. Esling ‘Sounds of the lower throat . . .’ Phonology 23.157–91. He holds a PhD from University of California, Los Angeles in Germanic Languages but has done research in Asia since 1980 including sponsoring grants from the National Endowment for the Humanities and the National Science Foundation for work on the ‘Languages of the Vietnam–China Borderlands’ (1995–1997) and (1998–2000). He is currently Professor of Linguistics at The University of Texas at Arlington, Arlington, TX, USA.

Kenneth J. Gregerson holds a PhD in Linguistics from the University of Washington (Seattle, 1971). He has served with SIL International as a researcher, consultant, and linguistics professor on languages of Southeast Asia since 1965. He is also an Adjunct Professor of Linguistics at the University of Texas at Arlington. His special focus has been on Mon-Khmer languages of the region with special emphasis on Vietnamese, Khmer, Rengao, and others. Along with his wife, Marilyn, he has over more than 40 years lived and worked in the countries of Vietnam, Cambodia, Thailand, and Indonesia. From 1979–1990, Dr. Gregerson served as President of SIL International. His most recent work has been in Cambodia as a consultant on language survey, analysis, and orthography development. He has lectured at many universities around the world highlighting the special features of the languages of Southeast Asia and advocating for the development of ethnic minority speech communities.

Notes

* Correspondence address: Jerold A. Edmondson, Department of Linguistics and TESOL, Box 19559, University of Texas at Arlington, Arlington, TX 76019-0559, USA. Email: j.edmondson@sbcglobal.net.
The authors, listed here in alphabetical order. Moreover, throughout we will use Vietnamese spelling of the names of groups unless they are better known by another name. We will also not attempt here to treat all minority groups in the same depth but instead single out some lesser known, lesser studied language groups for special attention. It should also be noted that there are perhaps as many as 100 distinct languages in Vietnam despite there being only 54 officially recognized groups. Often an ethnicity/nationality has been decided on political, cultural, historical, and social as well as linguistic criteria.

We offer here a few words in introduction to some ideas that figure into this essay. Demic expansions are increases in populations that result from some innovation. They are to be distinguished from cultural diffusion, as, unlike cultural diffusion, demic expansion usually leads to cultural transmission from parents to children as a part of absorbing new territory from rapid population growth.

Cladistics is a system of classification that arranges organisms by their order of branching in an evolutionary tree. Phylogeography is the analysis of the genetic markers for the geographic distributions of lineages of people as manifest in the pattern of their genetic variations. For example, Humans can have polymorphisms such as black to blond colored hair, black to blue eyes, and dense or light body hair growth. Most phylogeographic studies examine polymorphic forms of the mtDNA or Y chromosome DNA structures.

Long ago, genes, cultures, and languages were strongly linked. But in the last 10,000 years hunter-gather social forms transitioned to farming and gene mixing overtook the integrity of social organization along blood lines (Oppenheimer 2006).

2 For a discussion of the term Kra-Dai, see O斯塔pirat (2000). The US Department of State webpage (http://www.state.gov/r/pa/ei/bgn/4130.htm) states that ‘Vietnam’s approximately 2.3 million ethnic Chinese, concentrated mostly in southern Vietnam, constitute Vietnam’s largest minority group.’ It may be difficult to decide as there is considerable intermarriage between these groups and the ethnic Kinh or Vietnamese population especially in urban centers and they are not as frequently found in compact ethnic villages or towns.

3 Designations of Nùng and other groups by location and/or dress are of long-standing tradition and accord fairly faithfully with independent socio-linguistic facts.

4 In February 2007, a 200-page book written in an ancient form the Thái script was discovered in Mường Đan Village of Nghệ An Province. There have been reports before of the existence of such an ancient Thái script in use near the city of Quý Châu (Huỳnh-Láltanavong 1998). The Thái are thought to have migrated from further north to this area more than 2000 years ago.

The term Thai refers to Siamese as well as other related minority languages located in Vietnam and Laos. The form Tai is used by linguists to refer to the entire stock and is also used as a variant of Thái by groups in Vietnam. Tày designates a specific Tai group closely related to Nùng in North Vietnam.

Works Cited


Li Hui, Ying Huang, Laura F, Mustavich, Fan Zhang, Jing-Ze Tan, Ling-E Wang, Ji Qian, and Li Jin. 2007. Y chromosomes of prehistoric people along the Yangtze River. Human Genetics 10.1007/s00439-007-0407-2 online publication.


Thurgood, Graham. 1999. From ancient Cham to modern dialects: two thousand years of language contact and change. Honolulu, HI: University of Hawai Press.


WEB SOURCES


