Re-Imagining "Annam": A New Analysis of Sino–Viet–Muong Linguistic Contact

John D. Phan©2010[·]

Abstract:
This article examines the linguistic boundaries that separated (or united) Medieval China's southern territories and the river plains of northern Vietnam at the end of the first millennium C.E. New evidence from Sino–Vietnamese vocabulary demonstrates the existence of a regional dialect of Middle Chinese, spoken in the Ma, Ca, and Red River plains. Preliminary analysis suggests that a "language shift" away from this "Annamese Middle Chinese" in favor of the local, non-Chinese language, was largely responsible for the highly sinicized lexicon of modern Vietnamese. This theory, which challenges the tradition of an essentially literary source for Sino–Vietnamese, may help to explain some of the sinicized features of Vietnamese phonology and syntax as well. The last section of the article presents a tentative hypothesis for the formal emergence of Vietnamese contra its closest relative, Muong. These hypotheses require further testing, and are presented here as a first look at the history of the languages of "Annam".

Key Words:
Ancient Vietnam; Sino–Vietnamese; Muong; historical phonology; language contact

Introduction

This article revisits the notions of "Chinese" and "Vietnamese" in a linguistic context, and as they pertain to the transitional period linking the first and second millennia C.E. New evidence from Sino–Vietnamese (Chinese words borrowed into Vietnamese), and the Vietnamese language's closest living relative, Muong, demonstrate that traditional notions of the "survival" of the Vietnamese language under centuries of Chinese domination create a false imagining of its history and evolution—one that has been tailored to a political agenda of national identity. Preliminary analysis of these two bodies of data suggests, rather, that Chinese speakers native to Annam (by which I mean the Red, Ca, and Ma River plains) shifted to the local "Proto-Viet–Muong" language around the turn of the first millennium CE, transforming it in the process. Among the dialects of this new, hybridized Proto-Viet–Muong, one group eventually mutated, through generations of regional strife, into a unique new language: Vietnamese. This revised narrative challenges traditional concepts of the multi-millennial "survival" of the Vietnamese language, suggesting instead a "birth" of Vietnamese, significantly *not* in the depths of pre-Chinese history but during and immediately following Annam's long centuries of membership within the Chinese imperial order.

Before presenting my hypothesis and engaging with the technical arguments in the article, it may be useful to visit briefly the notion of Vietnamese identity. Research into Vietnam's ancient past has suffered from a multitude of historiographic fantasies, guided and nurtured by a range of social and political agendas. At the heart of these fantasies lies the notion of Vietnamese identity, a nebulously defined spirit that, like the stone in stone soup or the Wizard of Oz, derives its extraordinary power from the force of the idea of its existence.

[·] John D. Phan is a doctoral candidate at Cornell University. His email address is jdp49@cornell.edu.

But what exactly is the secret in the soup; just who is the man behind the curtain?

Anxious to defend the reality of a Vietnamese quintessence, a number of scholars have committed serious logical errors in their direct—or indirect—cultivation of the idea. Such classic works as Neil Jamieson's *Understanding Vietnam*, or Nguyễn Khắc Viện's enormously popular *Vietnam: A Long History* have served to inform a powerful vision of how and why the Vietnamese came to be (Jamieson, 1993; Nguyễn, 1993). Yet, as these and other examples indicate, the strongest claims appear motivated by a desire to explain the Indochinese conflicts of the twentieth century. Reduced to its base form, the equation is simple: Hanoi's success in the twentieth century is best explained by a distinct and indomitable "Vietnamese" spirit which has survived centuries (and in some cases, millennia) of invasion, occupation, and war. When thus stripped of its historical camouflage, the argument is revealed as both internally contradictory—how do we understand South Vietnam?—as well as reliant on flawed assumptions. Among its hidden premises lies the major assumption that the "Vietnamese" of the first-century CE Trưng Sisters rebellion, the "Vietnamese" of the third-century Triệu Thị Trinh rebellion, and the "Vietnamese" of the tenth-century Ngô Quyền "liberation" are essentially the same people, linked by a psycho-cultural connective tissue that makes each of these battles essentially just different rounds of the same epic struggle—a struggle that would dramatically repeat itself in a twentieth century theatre. Until very recently, this kind of political mythologising has crippled our attempts to understand the origins of the Vietnamese people and their culture.[1]

How, then, do we pierce the veil? How can we begin to understand the roots of a culture so recently named *Vietnamese*? Surprisingly, one of the most favoured weapons in the arsenal of historiographic fantasy may provide an answer: language. If a continuous Vietnamese identity is a 'Wizard of Oz' illusion, then language has been the smoke and mirrors that give the illusion life. Despite massive lexical borrowings—currently, over 70 percent of Vietnamese words are of Chinese origin (Trần, 1997, p. 555)—the language itself is genealogically unrelated to Chinese. Its perceived "survival" has therefore been used over and again as "empirical" proof for a distinct—and enduring—Vietnamese essence. To be successful, enlisting language in this way requires an ignorance of the promiscuity of language change, the subtlety of language contact, and the ultimately arbitrary nature of language genealogy. And yet, the study of language can indeed tell us volumes about the nature, history, and evolution of a group of human beings. The complex ways in which language evolves—both independently, and through contact with other languages—comprise a very special record of the people who speak it, and unlocking that record can sometimes provide extraordinary insights into realms beyond the study of pure linguistics. The history of the Vietnamese language is such a case, and the implications of its unique evolution tell us a great deal about how the modern Vietnamese emerged.

Based on a preliminary analysis of Sino–Vietnamese, and novel data on the Muong language, I will make two basic arguments concerning the evolution of the Vietnamese language. First, I argue that Annam witnessed a "language shift" among sectors of its elite population *away from* a local variety of Middle Chinese (spoken regionally), and *towards* the local, non-Chinese language of "Proto-Viet–Muong" (i.e. the immediate ancestor of the modern Vietnamese and Muong languages). This process of "language shift" radically transformed the grammar and vocabulary of Proto-Viet–Muong. Second, I argue that a subgroup of dialects among the (by then already) diversified array of this hybridized Proto-Viet–Muong, mutated in isolation, and so ultimately evolved into the modern Vietnamese language. My argument thus dates the "birth" of Vietnamese to the early centuries of the second millennium—over a thousand years after the first Han colonization of the Red River plain.

[1] Michael Churchman's contribution to this volume is a notable exception.

The article consists of three main sections. Part One reviews the basic arguments for language shift by examining the origins of Late Sino–Vietnamese words. Part Two summarizes the controversy over the Muong language, and presents a new model for Viet–Muong speciation. These analyses are then brought together and contextualized in the conclusion, wherein I revisit the notions of "Chinese" and "Vietnamese" as applied to the linguistic climate of tenth-century Annam.

Part One: Multilingualism and Language Shift in the Red River Plain

It is a serious but common error to imagine ancient polities as monolingual. Indeed, when examining the Vietnamese language, one is immediately confronted by vast numbers of Chinese loanwords. Because all language contact is facilitated by the exchange of words, this huge body of vocabulary represents a kind of fossil record which, under examination, can reveal the nature of the contact through which it was transfused.

It is rather telling that Sino–Vietnamese was traditionally considered the product of Tang Dynasty 唐朝 [618–907 CE] literary loans (Maspero, 1912; Wang, 1948). This theory matches explanations for other Sino–Xenic phenomena (like Sino–Korean *Hanja-eo*한자어/漢字語, and Japanese *Kan'on* 漢音). It is almost certain that instances of literary transfusion occurred several times in the history of Vietnamese, and continue to occur today. Yet despite being thought of as literary transfusions, these loans were also generally treated as representing contemporaneous Chinese *speech*. Mantaro Hashimoto attempted to explain this apparent paradox as "wholesale" loans, i.e. characters borrowed together with their "reading" (Hashimoto, 1968). Hashimoto makes a point of distinguishing this type of "Sino–Xenic" loan from other examples of Chinese loanwords outside of Japan, Korea, and Vietnam.

The language spoken in China roughly from the Northern and Southern Dynasties南北朝era [420–-589 CE] through to the end of the Tang Dynasty is generally termed Middle Chinese, which E. G. Pulleyblank reconstructed into two phases: Early Middle Chinese (pre-Tang), and Late Middle Chinese (Tang). All modern Chinese languages aside from the *Min* 閩 group (Fujianese) are traditionally thought to have evolved from *Late* Middle Chinese, including the donor of Sino–Vietnamese, which was used by Pulleyblank extensively in his reconstruction (Pulleyblank, 1991).[2] Maspero theorized that Sino–Vietnamese originated from the Chinese taught in the classrooms of the Annamese, and was based on the major dialect of Chang'an 長安, the Tang capital (Maspero, 1920). Others later echoed the notion of a northern origin for Sino–Vietnamese. However, the claim was seriously challenged by Mantaro Hashimoto who, using data from Wang, demonstrated six provocative similarities between the phonology of Sino–Vietnamese and a number of southern Chinese dialects (Hashimoto, 1968). This theory was partially supported and corroborated by Miyake, who rejected some details of Hashimoto's argument but accepted a southern origin for Sino–Vietnamese (Miyake, 2003, p. 127).

[2] Pulleyblank's use of Sino–Vietnamese, as well as Hashimoto's "wholesale" definition of Sino–Xenic loans, blurs the distinction between a "literary" and a "spoken" borrowing. As Hashimoto implies, part of this ambiguity lies in exactly how a word might be transfused through literature. Another part of this ambiguity seems to stem from the fact that Rhyme Books actually reflect a number of major sound changes that transformed Early Middle Chinese into Late Middle Chinese. What is really at issue, therefore, is that what we understand as "Late Middle Chinese" is largely based on literary standards, and so we do not have a sufficiently diverse reconstruction of Middle Chinese dialects to be able to distinguish between claims of a "literary" or a spoken modicum of borrowing. Norman (1988) objected to the Karlgrenian notion of a homogenous Late Middle Chinese ancestor for most modern Chinese languages. While Norman's theories have been well-received, evidence for and articulation of Middle Chinese diversity has been slow to follow until this article. Its new data allows us to pinpoint with some accuracy the nature of the donor of Sino–Vietnamese, and thus to articulate at least one dialectal variety of Middle Chinese.

In agreement with Miyake, and in support of Hashimoto's larger claim, I will demonstrate that Sino–Vietnamese must rather have originated from a *regional dialect* of Middle Chinese, native (although not necessarily restricted) to the Red, Ma, and Ca River plains (i.e. "Annam"). This claim rests on two sound-changes known to have affected Late Middle Chinese, which I will show did not affect the donor of Sino–Vietnamese. The absence of these sound-changes in Sino–Vietnamese phonology requires its ancestor to have diverged before the formation of Late Middle Chinese.

Before turning to these sound-changes, let us first establish a more precise definition of "Sino–Vietnamese".

Introduction to Sino–Vietnamese

Broadly speaking, Sino–Vietnamese (*Hán–Việt*) refers to any words of Chinese origin that have become fixed in the Vietnamese lexicon. This is more complicated than it first appears because of the large number of calques (literal translations of a word or phrase, e.g. "flea market" from French *marché au puces*) and neologisms that emerged in the modern era. Most of these are recognizably "Chinese" to the average native speaker. They are generally elevated vocabulary with either literary or intellectual flavour, and may not have a "native" alternative in the language:

Table 1

#	字	Sino-Viet.	Gloss
1.	兄弟	huynh đệ	"brothers"
2.	說明	thuyết minh	"to explain"
3.	有益	hữu ích	"useful"
4.	文化	văn hóa	"culture"
5.	政府	trính phụ	"government"
6.	社會	xã hội	"society"
7.	工產	Cộng Sản	"Communism"
8.	生學	sinh hộc	"biology"
9.	移動	di động	"mobile phone"
10.	風俗	phong tục	"custom"

In fact, a great number of the cases in Table 1 were borrowed through Japanese (e. g., Japanese *bunka* for "văn hóa", *seihu* for "trính phụ", *seikai* for "xã hội"), while others were either translated from Mandarin neologisms (e. g., Mandarin *gongchan* for "Cộng Sản"), or were locally created (e. g., "di động").[3] The vast majority of modern terminology has either been borrowed or fashioned in this manner—a habit shared by many Sinospheric languages, and analogous to the role of Latin and Greek in Europe (e. g., words like "ophthalmology" or "neutrino"). As a phenomenon, these calques and neologisms only matter to us because they must derive from a phonological system that predated their creation—and it is precisely the origin of this system that we are trying to determine.

[3] Words borrowed through Japanese in this fashion were "calqued" via literary routes. Thus, the written form of a word like *bunka* ("culture") was transliterated into *văn hóa*, with the (Chinese) grapheme acting as the conduit between the Japanese word and its Vietnamese incarnation.

For this reason, we can dismiss the type of data in Table 1, and move on to words directly borrowed from Chinese. Even among these direct loans, we find a very great body of elevated vocabulary:

Table 2

#	字	Sino-Viet.	Gloss
1.	文	văn	"literature; pattern"
2.	仁	nhân	"humaneness"
3.	義	nghĩa	"righteousness"
4.	福	phúc	"fortune"
5.	祿	lộc	"prosperity"
6.	壽	thọ	"longevity"
7.	財	tài	"wealth"
8.	法	pháp	"law"
9.	禮	lễ	"ritual"
10.	樂	nhạc	"music"

Virtually all words for intellectual concepts are of Chinese origin, and it is largely for this reason that Sino–Vietnamese has long been considered the product of literary transfusions.

For a language to possess a vocabulary "bejeweled" in this manner is itself neither surprising nor unique. Among other Asian languages, Japanese and Korean both share an elevated Chinese register of vocabulary, while Korean also contains a certain degree of Japanese prestige loans. Similarly, Thai reveals an elevated stratum of Khmer words while Khmer in turn demonstrates a similar stratum of Pali words. What is striking about the Sino–Vietnamese case is that an extraordinary number of basic vocabulary words are also of Chinese origin, as Table 3 shows.

Table 3

#	字	Sino-Viet.	Gloss
1.	婆	bà	"grandmother"
2.	裙	quần	"trousers"
3.	感	cảm	"to feel"
4.	冷	lệnh	"cold"
5.	頭	đầu	"head"
6.	古	cũ	"old"
7.	心	tim	"heart"
8.	外	ngoài	"outside"
9.	面	miền	"face/side"
10.	旬	tuần	"week"

These household words are seldom recognized as Chinese in origin. To these, we may also add an extraordinary range of functional and grammatical words, as typified in Table 4 (over page).

Table 4

#	字	Sino-Viet.	Gloss
1.	在	tại	prep. "at"
2.	如	như	"like, as"
3.	為	vì	comp. "because"
4.	被	bị	advers. marker
5.	得	được	"to get"
6.	只	chỉ	"only"
7.	仍	nhưng	"but"
8.	雖	tùy	"although"
9.	個	cái	gen. classifier
10.	卷	cuốn	classifier
11.	類	loại	classifier/"type"
12.	條	điều	classifier
13,	封	phong	classifier
14.	分	phần	classifier/"part"
15.	當	đương/đang	present/progressive marker
16.	來	lại	aspect marker
17.	過	quá	adv. "excessively"
18.	實	thực	adv, "truly, really"
19.	各	các	quantifier, "every, all"
20.	每	mỗi	quantifier, "each"

Alves (2008) noted that most Sino–Vietnamese grammatical vocabulary is comprised of connectives and measure words. The entries in Table 3 are mostly classifiers (see numbers 9–14), with some complementizers (e.g., 3 and 7–8). Also apparent are prepositions (1), adverbs (17), some morphemes used in derivations (18), as well as quantifiers (19 and 20). All of these words are thoroughly perceived as Vietnamese.

It is this astonishing range and pervasiveness that sets Sino–Vietnamese apart, not only from other Sino–Xenic phenomena but also from many other cases of language contact and borrowing known to the field.

Fortunately, this very diverse body of vocabulary is relatively datable, and Sino–Vietnamese has long been divided into two categories depending on the time-depth of the borrowing (Maspero, 1912; Wang, 1948; Miyake, 2003). These are: "Early Borrowings", which include any borrowings from Chinese ranging from Old Chinese (spoken in the Later Han 後漢 Dynasty of the second century BCE) through the Early Middle Chinese (EMC) of the early Tang 唐 Dynasty (sixth and seventh centuries CE); and "Late Borrowings", which are ostensibly borrowings from Late Middle Chinese (LMC), spoken during the latter half of the Tang Dynasty, that were thought to descend from literary loans. "Early Borrowings" were taken before the completion of sound-changes that transformed Early Middle Chinese into Late Middle Chinese (to use Pulleyblank's terminology), while "Late Borrowings" were taken after. For this reason, I propose to use "Early Sino–Vietnamese" for early borrowings and "Late Sino–Vietnamese" for later borrowings, for their ease of comparison with the Early and Late Middle Chinese of Pulleyblank's reconstruction.

Most Sino-Vietnamese words were borrowed during the late period—a very unsurprising fact since Tang China exported huge numbers of words around the continent at this time (cf. *Hanja-eo*, *Kan'on*). We are thus concerned here only with Late Sino–Vietnamese, leaving aside for the moment the origins of Early Sino–Vietnamese.

The existence of such a vast number of basic Chinese loanwords indicates a form of contact that was far more intimate, and far more pervasive, than the types of literary exchanges responsible for Japanese *Kan'on*. The pertinent question, therefore, is precisely what kind of contact would have produced words of the number and nature seen in Late Sino-Vietnamese?

An Annamese Middle Chinese

The origins of Late Sino–Vietnamese are retrievable through an application of the comparative method, a form of analysis that compares modern related dialects or languages in order to reconstruct their evolutionary history. Although Sino–Vietnamese itself represents a body of loanwords that derive from a genetically unrelated language—an "adopted" versus a genetically inherited vocabulary—the comparative method is still a powerful tool for articulating its nature. If we screen out innovations known to belong to the evolution of Vietnamese, then we should be able to test Late Sino–Vietnamese for innovations known to be shared with the rest of the descendents of Late Middle Chinese (i.e., all the modern languages of China aside from Min languages). Furthermore, if Late Middle Chinese resulted primarily from literary loans it should adhere to the literary standard of Late Middle Chinese as represented by the Rhyme Tables. What we seek is an innovation (or innovations) shared both by comparative evidence (in the form of modern Chinese languages) and by the phonology described in the Rhyme Tables.

There are two such innovations: the merger of palatal and retroflex affricates; and a "Grassman's-like" dissimilation. If these innovations are shown to be reflected in Late Sino–Vietnamese as well, then the claim that Late Sino–Vietnamese derives from Late Middle Chinese (or its literary standard) stands on solid linguistic ground. However, if (as I will show) these innovations are not shared with Late Sino–Vietnamese, then Late Sino–Vietnamese must derive from another source. I will treat each of these innovations in turn, first describing their mechanics before turning to the Sino–Vietnamese data in search of their presence.

a) The Merger of Palatal and Retroflex Affricates. The first innovation involves two types of consonants called *palatal* and *retroflex* affricates. Affricates are a type of compound consonant (think of the initial in Japanese *tsunami* as ***t-*** + ***s-***), while the terms *palatal* and *retroflex* refer to the point in the oral cavity at which the consonants are articulated. These two types of sounds had merged by the end of the Tang Dynasty, and were recorded as a single class in the Rhyme Tables, labelled *zhengchiyin* 正齒音 (Baxter, 1992).

Table 5. Merger of Palatal and Retroflex Affricates[4]

字	EMC	LMC	E. Man.	Man.	Gloss
懺	tʂʰəɨmʰ	tʂʰaːmˋ	tʂʰaːmˋ	tʂʰan51	"repent"
臭	tɕʰuwʰ	tʂʰiẇ	tʂʰiw	tʂʰow51	"smelly"

[4] All reconstructions, unless otherwise noted, are taken from Pulleyblank's 1991 *Lexicon*.

As shown in Table 5, the forms listed in the Early Middle Chinese column are distinct. The form in red (top row) corresponds to a retroflex affricate, while the form in blue (bottom row) corresponds to a palatal affricate. Note however, that by Late Middle Chinese (LMC), the palatal form has merged with the retroflex series. This merger is still reflected in modern Mandarin: these two characters are represented in Pinyin as **ch**an4 and **ch**ou4 respectively.

Let us now ask: is this innovation shared with Late Sino–Vietnamese? The answer is unequivocally no.

Table 6. Merger of Palatal & Retroflex Affricates

字	EMC	LMC	E. Man.	Man.	LSV		Gloss
懺	$tʂ^{h}$əɨm^{h}	$tʂ^{h}$a:m̀	$tʂ^{h}$a:m̀	$tʂ^{h}$an51	**ʂam31ʔ**	**sạm**	"repent"
臭	**tɕ^{h}**uw^{h}	$tʂ^{h}$iẁ	$tʂ^{h}$iw	$tʂ^{h}$ow51	**su35**	**xú**	"smelly"

As the data in yellow (far right) demonstrates, the retroflex series (in red) remain distinct from the palatal series (in blue).[5] In all of the cases above, the Late Sino–Vietnamese tokens demonstrate a clear distinction between initials derived from Early Middle Chinese palatal affricates, and those derived from Early Middle Chinese retroflex affricates.

The fact that these two sound classes did not merge in Sino–Vietnamese is the first piece of evidence that challenges the standard assumptions of its origins. The second involves the deletion of "voiced aspiration" in syllables bearing a final -^{h}.

b) A "Grassman's-like" Dissimilation.

The second shared innovation involves a dissimilatory process similar to the principle called Grassman's Law (Pulleyblank, 1984). Pulleyblank argued that Late Middle Chinese voiced aspiration (transcribed here as a period of "breathy phonation" -̤-) is deleted in syllables with oblique (*ce* 側) tones derived from final –̤.[6] This breathiness was the result of voiced initials that had lost their voiced quality by Late Middle Chinese.

Table 7

字		Tone-type	EMC	LMC	Man.	Gloss
1.	卞	oblique-departing	bian^{h}	p̤iaǹ	pian51	impetuous
2.	便	level	**b**jan	p̤ian	p**^{h}**ian35	easy, cheap

[5] Note that subsequent sound-changes have transformed the originally palatal onsets into alveolar fricative –s-. The retroflex series did in fact merge with the alveolar in northern dialects, but only very recently (in perhaps the last twenty years). The long held distinction of these two classes can still be seen in the conventional orthography, which maintains <s> for the original retroflexes, and <x> for the original palatals.

[6] There were three oblique tones: the oblique tone deriving from -̤ was known as the "departing tone" 去聲; the tone arising from a final glottal (-') was known as the "rising tone" 上聲; and the neutralization (or "clipping") of full tones in syllables ending with (-p, -t, or –k) was called the "entering tone" 入聲. These three oblique tones, along with "level tone" 平聲, comprised the four classical tone groups in Middle Chinese. Note also that Pulleyblank originally transcribed voiced aspiration as a separate segment, -ɦ-. For reasons explained in Phan (2008, manuscript), I prefer both to analyze this quality and to transcribe it as breathy phonation type. In this article, all Late Middle Chinese data transcribed with breathy phonation type -̤- correspond to (-ɦ-) in Pulleyblank's *Lexicon* (Pulleyblank, 1991).

As shown above, the voiced *b*- in Early Middle Chinese becomes a voiceless *p*- in Late Middle Chinese, accompanied by breathiness (-ˌˌ-). This breathiness is then deleted in syllables with oblique-departing tone (i.e., EMC final -ˌˌ-), yielding a voiceless, unaspirated stop in modern Mandarin (cf. Mandarin plain *p*- in row 1, spelled as *b*- in modern Pinyin). Alternatively, syllables with "level" (*ping* 平) tones deriving from open syllables or syllables with final sonorants (in this case, final –n) do not delete the breathiness resulting from an Early Middle Chinese voiced onset, but transform it into voiceless aspiration (also transcribed as -^{h}) in the modern language (see Mandarin *p^{h}*- in row 2, spelled as *p*- in modern Pinyin).

In essence, the (-^{h}) found at the end of the syllable is perceived of as "too similar" to the breathiness (-ˌˌ-) found at the beginning of the syllable, leading to the deletion of (-ˌˌ-), and the subsequent prevention of its transformation into voiceless aspiration (-^{h}) in the modern language.[7]

Pulleyblank's observations are supported by good evidence in the Mandarin and Cantonese comparative data (as well as the philological record). The backbone of Pulleyblank's arguments stems from the fact that Early Middle Chinese voiced onsets bear aspirated reflexes in syllables with level tone (as in row 2 of Table 7), a very robust pattern, which is further shown in the next table.

Table 8

字		Tone-type	EMC	LMC	Man.	Cant.	Gloss
1.	便	level	**b**jan	pˌˌian	p**ʰ**ian35	p**ʰ**in21	easy, cheap
2.	嘲	level	**b**jin	pˌˌin	p**ʰ**in35	p**ʰ**an21	knit brows
3.	田	level	**d**ɛn	tˌˌian	t**ʰ**ian35	t**ʰ**in21	field
4.	停	level	**d**ɛŋ	tˌˌiajŋ	t**ʰ**iŋ35	t**ʰ**iŋ21	stop
5.	狂	level	**g**uaŋ	kˌˌyaŋ	k**ʰ**waŋ35	k**ʰ**uɔ ŋ21	mad, crazy

In Table 8, all tokens are level tone (i.e. no final -^{h}), and thus would be expected to bear aspirated onsets in the modern languages—an expectation clearly borne out by the Mandarin and Cantonese data.

Again, these level tone cases contrast with oblique (departing) tone cases, in which the final (-^{h}) triggers the deletion of breathiness, thus preventing the development of aspirated onsets in the modern languages.

Table 9

字		Tone-type	EMC	LMC	Man.	Gloss
1.	卞	oblique-departing	bian^{h}	pˌˌiaǹ	pian51	impetuous
2.	辨	oblique-rising	bian’	pˌˌiaǹ	pian51	distinguish
3.	電	oblique-departing	dɛm^{h}	tˌˌiam̀	tian51	lightning

[7] Please note that both initial voiceless aspiration, and the departing tone's "final –h", are transcribed as a superscripted (-^{h}). This is because in all likelihood they were essentially quite similar, except that one occurred at the beginning of the syllable and the other at the end.

4.	簟	oblique-rising	dɛm'	t̪.iam̀	tian51	woven grass matt
5.	墐	oblique-departing	gɨn^{h}	k̪.iǹ	tɕin51	plaster with mud
6.	近	oblique-rising	gɨn'	k̪.iǹ	tɕin51	near

Thus, none of the Mandarin tokens in Table 9 demonstrate aspirated onsets (as opposed to the data shown in Table 8).

At this point, the condition of "final –^{h}" needs to be refined to include both oblique departing (*qu* 去) tone, which, as already stated, derives from final -^{h}, and oblique rising (*shang* 上) tone, which derives from syllables bearing a final glottalization (-'). At least in Mandarin, departing and rising tones have been shown to merge. (Pulleyblank, 1984, Jacques, 2005) Data 1, 3, 5 above all demonstrate departing tone (transcribed in EMC as a final –^{h}); data 2, 4, 6 above all demonstrate rising tone (transcribed as final glottal –'). Because the final glottal (-') is assimilated to the final (-^{h}), all syllables originally bearing it were then subject to the form of Grassman's Law described above.

The Sino–Vietnamese Data

The question that pertains here, of course, is whether we see the same kinds of reflexes in Late Sino–Vietnamese as we do in Mandarin or Cantonese. To refresh memories of the crucial reflexes, Table 7 is reproduced as Table 10 below.

Table 10

字		Tone-type	EMC	LMC	Man.	Gloss
1.	卞	oblique-departing	bian^{h}	p̪.iaǹ	pian51	impetuous
2.	便	level	bjian	p̪.ian	pʰian35	easy, cheap

In summary, Early Middle Chinese level tone syllables (crucially lacking final –^{h}) which bear voiced onsets, yield aspirated onsets in modern Chinese languages (see Mandarin in row 2 above). In contrast, Early Middle Chinese departing tone syllables (and the rising tone syllables that assimilated to them) which bear voiced onsets, delete the "voiced aspiration"/breathiness resulting from onset voicing, and yield unaspirated reflexes in modern Chinese languages (see Mandarin in row 1 above). In other words, for Late Sino–Vietnamese to have been borrowed from the metropolitan standard of Middle Chinese, we ought to see *unaspirated* onsets in syllables historically bearing voiced onsets and departing tone (final –^{h})—just as in the modern Chinese comparative data; and *aspirated* onsets in syllables historically bearing voiced onsets and level tone—just as in the modern Chinese comparative data.

Most Late Sino-Vietnamese syllables historically bearing voiced onsets and departing tone (final –^{h}), do appear to demonstrate unaspirated onsets.

:

Table 11

#	字	Tone-type	EMC	LMC	Man.	LSV	Gloss
1.	罷	oblique-rising	baɨj'	p̪.a:j̀	pa51	ɓaj3ʔ5	cease
2.	稗	oblique-departing	baɨj^{h}	p̪.a:j̀	paj51	ɓaj35	barnyard millet

3.	辦	oblique-departing	baɨn^{h}	p̣.aːǹ	pan51	ɓiən31^{ʔ}	handle, deal with
4.	傍	oblique-departing	baŋ^{h}	p̣.aŋ̀	ban4	ɓaŋ21	near, approaching
5.	暴	oblique-departing	baw^{h}	p̣.uaẁ	bao4	baw31^{ʔ}	sudden, violent

However, this correspondence fails in syllables bearing palatal affricate onsets:[8]

Table 12. Grassman's-like Dissimilation and SV Palatal Affricates

字	Tone-type	EMC	LMC	Man.	LSV	Gloss
上	oblique-departing	**dʝ**ianŋ^{h}	**ʂ**.ianŋ̀	**ʂ**aŋ C2	**t^{h}**iəŋ31^{ʔ}	on

As the Late Sino–Vietnamese datum demonstrates, the Sino–Vietnamese correspondence for Early Middle Chinese voiced palatal affricates appears to be an aspirated onset. This is exactly what we would not expect, since the Grassman's-like process should have deleted the breathiness from which the modern voiceless aspiration ostensibly derives. Nonetheless, this correspondence is borne out robustly in the data.

Table 13. EMC Voiced Palatal Affricates & Grassman's Law

字		Tone-type	EMC	LMC	E. Yuan	Mand.	LSV	Gloss
1.	善	oblique-departing	dʝian^{h}	ʂ.ianˋ	ʂɛǹ	ʂan51	t^{h}iən35	good (virtuous)
2.	樹	oblique-departing	dʝuə^{h}	ʂ.yə̀	ʂỳ	ʂu51	t^{h}u31^{ʔ}	tree
3.	上	oblique-departing	dʝianŋ^{h}	ʂ.ianŋ̀	ʂaŋ̀	ʂaŋ51	t^{h}iəŋ31^{ʔ}	on
4.	召	oblique-departing	dʝiaw^{h}	ʂ.iaẁ	ʂhɛẁ	ʂaw51	t^{h}iu31^{ʔ}	summon
5.	盛	oblique-departing	dʝiajŋ^{h}	ʂ.iajŋ̀	ʂiŋ̀	ʂəŋ51	t^{h}iŋ31^{ʔ}	flourish

It is important to note that these data all bear affricate onsets. Because affricates are compound consonants, they often behave differently than regular consonants like *p*-, *t*-, *k*-. Note, for example, that Early Middle Chinese dʝ- becomes a fricative ʂ- (a throaty version of English *sh*-) in Mandarin, which makes it quite difficult to determine whether any form of aspiration survived. In fact, the aspiration present in the Sino–Vietnamese reflexes is something of a mystery; on its face, it seems clearly to contradict what we would expect out of Pulleyblank's Grassman's-like dissimilation, and yet affrication introduces a number of possible alternative causes for aspiration that may have developed later.[9] It is important, therefore, to look also to the "control" cases—those cases involving level tone, in which we would expect unequivocally to see aspirated onsets. But as Table 14 indicates, none of these demonstrate aspirated onsets:

Table 14

字		Tone-type	EMC	LMC	Man.	Cant.	LSV	Gloss
1.	便	level	**b**jan	p̣.ian	p**ʰ**ian35	p**ʰ**in21	tiən31^{ʔ}	convenient
2.	嘲	level	**b**jin	p̣.in	p**ʰ**in35	p**ʰ**an21	tən21	knit brows

[8] This is the same class of consonants that was discussed above (see page 10).
[9] It remains unclear where this aspiration comes from. It is possible that there was a contrastive, aspirated fricative in Proto-Viet–Muong, as is the case in some Burmese languages.

3.	田	level	**d**ɛn	t̤.ian	t^{h}ian35	t^{h}in21	ɗi̤ən21	field
4.	停	level	**d**ɛŋ	t̤.iajŋ	t^{h}iŋ35	t^{h}iŋ21	ɗi̤ŋ21	stop
5.	狂	level	**g**uaŋ	k̤.yaŋ	k^{h}waŋ35	k^{h}uɔŋ21	kuəŋ21	mad, crazy

What we see is that level-tone syllables bearing voiced onsets surprisingly bear *unaspirated* onsets in Sino–Vietnamese. These data suggest that aspiration (whether voiced or otherwise) was never a reflex of voicing in the donor of Late Sino–Vietnamese, meaning that the breathiness which resulted from voicing took an entirely different path in the donor of Sino–Vietnamese. For our purposes, it is enough to note that this form of Grassman's Law clearly did not occur in the donor of Late Sino–Vietnamese.

If neither innovation can be shown to have affected Late Sino–Vietnamese, then the donor of these words must have been distinct both from the literary standard of Middle Chinese and from the spoken dialects from which most modern Chinese languages evolved. Therefore, Late Sino–Vietnamese must have been borrowed from a regional dialect of Middle Chinese spoken in Annam. It remains possible that this regional dialect of Middle Chinese was part of a more broadly-spoken southern *koine*, as was once posited by Mantaro Hashimoto (Hashimoto, 1968), and the intriguing possibility remains that it may be a sort of "cousin" of the modern Xiang 湘 language of Hunan 湖南 Province. At the moment, however, all that can be demonstrated with confidence is that it was spoken in the Red, Ma, and Ca River plains, and for this reason I call it "Annamese Middle Chinese", after the Tang designation for the region (*Annan* 安南/ Viet. *An Nam*).

Annam in Flux

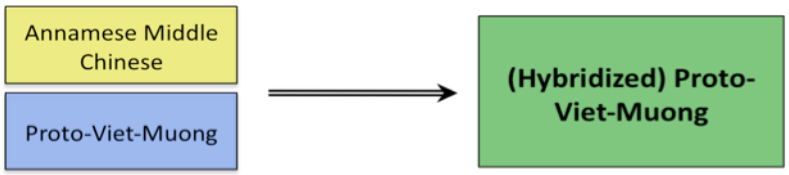
We are now beginning to accumulate a richer picture of the languages of Annam: not only were sectors of society literate in Chinese, some appear to have been entirely comprised of *native* Chinese speakers. These people spoke no imported northern dialect, however, but apparently a home-grown variety of Middle Chinese that had developed in ways peculiar to the area. Yet an important question remains unanswered: how exactly did Annamese Middle Chinese donate such a profundity of words into the ancestor of modern Vietnamese? Here I propose that Late Sino–Vietnamese resulted from a language shift, in which Annamese Middle Chinese speakers switched linguistic allegiance from their own language to Proto-Viet–Muong. This process effectively "dragged" large numbers of words (as well as some linguistic features) from their native language into the language they adopted, thereby transforming it.

The phenomenon of language shift does not appear very uncommon in the history of languages. Some of the best-documented examples include the Alsatian shift from German to French, the Finnish switch from Swedish to Finnish, and the Norman shift from French to English.[10] Put concisely, language shift refers to the "partial or total abandonment of a group's native language in favour of another" (Winford, 2003, p. 15). The native language is referred to as Language 1 (or L1), from which the speaker population shifts allegiance in favour of a Target Language (TL or L2). Under this scheme, the term "interference" is used to measure the effect that L1 exerts over TL, i. e., the amount and character of change undergone by TL as a result of language shift.

The conditions under which a speaker population might shift are quite varied, as the cases above suggest: in Alsace, native German speakers suffered French bans on the German language, while in Norman-conquered England, the French-speaking aristocracy

[10] For basic treatments of these examples, see Thompson and Kaufman, *Language Contact, Creolization and Genetic Linguistics* (1988), and Winford, *An Introduction to Contact Linguistics*.

shifted to English as political and cultural ties to the European mainland faded.[11] Winford describes two principle forms of language shift, both characterized by L1 exerting a linguistic interference on TL but occurring under different sociolinguistic conditions. The second (and for our purposes, less interesting) category involves a native population adopting an invasive language, and thereby "indigenizing" the adopted language (for instance, Hibernian, Indian, or Singaporean English) (Winford, 2003, p. 15). The more interesting category for us involves a shifting group that is completely absorbed into the TL community, a case in which "the innovations that [the shifting group] introduced are imitated by the TL community as a whole, thus becoming permanently established in the language" (Winford, 2003, p. 15). Applied to turn-of-the-millennium Annam, this case suggests that a population of Annamese Middle Chinese speakers shifted to Proto-Viet–Muong, thereby inducing a number of linguistic innovations that were subsequently imitated by the Proto-Viet–Muong community as a whole and thus "permanently established in the language". This post-shift Proto-Viet–Muong would have consequently demonstrated a number of superficial "Chinese" qualities, as well as containing a large number of Chinese loanwords as relics of the abandoned L1. The entire process can be schematized quite simply as follows:



Under this scheme, Annamese Middle Chinese induced an *adstratum* effect on Proto-Viet–Muong. Essentially, Annamese Middle Chinese donated both words and structures to Proto-Viet–Muong throughout a prolongued period of intimate bilingualism. As generations of bilingual speakers gradually began to favour Proto-Viet–Muong, the variety of Proto-Viet–Muong spoken by these people became heavily influenced by their native Annamese Middle Chinese—a form of mixing similar to the English spoken by immigrant populations in the United States. However, unlike the analogy of, say, Cuban immigrants in modern America, the hybridized form of Proto-Viet–Muong spoken in Annam was perceived as a prestige dialect since it was undoubtedly spoken by the social and cultural elite. As this elite class of hybridized Viet–Muong/Annamese Middle Chinese speakers began slowly to abandon their native language, their hybridized Proto-Viet–Muong spread throughout the community until, as Winford describes, the entire TL population had adopted the unique characteristics of the prestige group. The final stage of language shift involved the total abandonment of Annamese Middle Chinese, leaving only (a highly sinicized) Proto-Viet–Muong spoken in the Red River plains. It was this hybridized Proto-Viet–Muong that would eventually split into modern Vietnamese and Muong languages.

[11] The case of Norman England provides a fascinating comparison with the tenth-century Red River plains. In many ways, the political and social history of tenth-century Vietnam greatly resembles the situation of Norman-held England a century later. In each case, an aristocratic minority ruled a linguistically foreign majority in geographically isolated lands.

This hypothesis best fits the available facts at this point, but it requires a strong and thorough comparison with other documented cases of language shift, as well as conventional borrowing, to solidify the claim.[12]

Part Two. Vietnamese, Muong, and the Proto-Viet–Muong Language

Part One determined the presence of a regional dialect of Middle Chinese, and proposed its extinction via language shift. Yet what of the non-Chinese contemporary of that language, the language adopted (and thereby transformed) by the elite of the Red River plains? Claims surrounding the millennial "survival" of the Vietnamese language have assumed that what was spoken throughout the centuries of "Chinese domination" should be called (and considered) "Vietnamese", although even a cursory examination reveals the grave error in this assumption. In order to better understand the substantive origins of the Vietnamese language, we must look to its living relatives.

The Muong are the closest living relatives of Vietnam's majority Kinh (or ethnically Vietnamese) population, and their language may be considered a sister to Vietnamese. As alluded to above, the shared ancestor of Vietnamese and Muong is a language called Proto-Viet–Muong, a term which, in the context of this paper, describes the language both before and after the transformative shift of Annamese Middle Chinese speakers to it. In other words, Proto-Viet–Muong was the language that long coexisted with Annamese Middle Chinese, and it was the language that was subsequently hybridized in its adoption by the shifting Annamese Middle Chinese elite.

The Muong have long been victims of a special kind of intellectual neglect which unfortunately results from their close kinship with the Vietnamese. While some important work has been done on the Muong,[13] a great deal of ambiguity still surrounds the nature of their language and its relationship to Vietnamese.

Figure 1. Distribution of the Muong (and other Vietic) Languages



[12] For a fuller defense of this hypothesis, see my forthcoming dissertation.
[13] Map reproduced with permission from Paul Sidwell's *mon-khmer.com* website (Sidwell, 2009).

Introduction to the Muong and their Language

The Muong are currently the third-largest ethnicity in Vietnam (after the majority Kinh, and the highland Tay), with a population of roughly 1,140,000 (1999 census) spread out over an area west, southwest, and south of the Red River. They are most heavily concentrated in the provinces of Hoa Bình, Thanh Hóa, and Phú Thọ[14] with communities as far west as Yên Bái and Sơn La, and as far south as Nghệ An (Lewis, 2009). Muong communities are generally situated in low mountain valleys surrounded by peaks, which places them in geographical zones contiguous with the Kinh majority (as opposed to the higher elevations inhabited by the Hmong or Dao). The Muong are subsistence farmers who cultivate rice and corn as staples, alongside a number of small cash-crops including tea (Phú Thọ), sugarcane (Thanh Hóa, Phú Thọ, Hoa Bình), and, recently, acacia lumber (Phú Thọ, Hoa Bình).

The term "Muong," as both a language and ethnicity, has suffered from a long history of ambiguity. Cuisinier described the Muong as culturally and linguistically *les Annamites attardés* ("backward Annamese"), establishing a strong bias toward understanding the Muong as primitive, fossilized cousins of the lowland Vietnamese.[15] (Cuisinier, 1946, pp. 562–63) Keith Taylor ultimately challenged this notion, arguing that French ethnological categories forced the Muong into an artificially distinct subgroup that was necessarily subordinate to the lowland Vietnamese (Taylor, 2001). Although not linguistic in character, Taylor's criticism is largely applicable to early scholarship on Muong and other ethnic minority languages related to Vietnamese.

Vương Hoàng Tuyên (1963) counted at least three Vietic languages, identified by place name, as "Muong", which he listed alongside yet another (simply) "Muong" heading. Thompson later likened the value of this approach to the "presentation [of] three or four Swiss German dialects, identified by village names, besides a list headed simply 'German'". (Thompson, 1976, p. 1115). Thompson was right to point out the ambiguity in Vương's practice, but his German analogy only illustrates a confusion of nomenclature; the real error here lies in what both Vương and Thompson appear to accept as "Muong". As Thompson notes, the plain "Muong" listed was probably a Hoa Bình dialect of Muong (proper). However, the other three "forms" (May/Rục, Arem, Tay Pong) are in fact, distinct languages more closely related to each other than to either Vietnamese or—significantly—to Muong (Ferlus, 1975). Evidently, the habit of labelling any "primitive" relatives of the Vietnamese as "Muong" had led to the creation of something Taylor rightly called an artificial category which lumped together groups of very different lineages in opposition to the lowland, urbanized civilization of the Kinh. In fact, the "Muong" spoken in Hoa Bình (as well as Thanh Hóa and Phú Thọ) is better grouped with Vietnamese than with these other languages.

Ferlus eventually clarified this issue by identifying a subgroup of "conservative" (*conservatrice*) Vietic languages, including Rục, Arem, Pong, Maliêng, Thavưng, and others, all opposed to Vietnamese and Muong, both of which notably share full-blown tone systems, pervasive monosyllabicism, and an impoverished cluster inventory (Ferlus, 1975). To this we may add (as Maspero noted nearly a century ago), large bodies of Chinese loanwords (Maspero, 1912), Nevertheless, the impression that Muong represents a cohesive subgroup opposed to Vietnamese has survived into modern scholarship (Rischel, 1995).

[14] Vinh Phú province was split into Phú Thọ and Vinh Phúc provinces in 1996; the *ethnologue* description of Muong still lists "Vinh Phú" (rather than Phú Thọ) as a Muong area.
[15] For an excellent review of early French ethnographic scholarship on the Muong, see Keith Taylor's 2001 article, "On Being Muonged".

The genealogical status of Vietnamese (and by association, Muong) was once a subject of some controversy. Initially long held to be a simple offshoot of Chinese (see Taberd, 1838), Henri Maspero later recognized an affinity between Viet–Muong and other Southeast Asian language groups like Tai and Mon–Khmer, arguing for a principle connection with Tai because of the presence of tone (Maspero, 1912). The issue was finally laid to rest by André Haudricourt in his seminal 1954 account of tonogenesis, which demonstrated not only that Vietnamese (and Muong) tones evolved from final consonants (and not a unique genetic inheritance), but that Vietnamese (and Muong) was a member of the Mon–Khmer family, and completely unrelated either to Tai or Chinese languages (Haudricourt, 1954).

This genetic classification was further refined in the 1990s, when works by Michel Ferlus and Gerard Diffloth argued for the placement of all Vietic languages in the Katuic subgroup of the Mon–Khmer family, the main branch of the Austro–Asiatic super-family (Diffloth, 1991; Ferlus, 1989-90; Ferlus, 1994). This classification is still somewhat contested, while the diversification of Vietnamese and Muong themselves remains completely obscure.

For our purposes, the question of whether or not Muong is a subgroup is essential to an understanding of how modern Vietnamese emerged out of a hybridized Proto-Viet–Muong, and resolving it will allow us to complete a picture of how and when Vietnamese came to be. Before confronting this issue, however, it is important to consider something Maspero claimed in 1912, that Muong also inherited the brunt of Chinese loans borrowed into Vietnamese. Throughout Part One, I described the language shift event as one from Annamese Middle Chinese to Proto-Viet–Muong, and have yet to demonstrate that Muong was also a recipient of large-scale Chinese borrowings. Indeed, an unspoken assumption that Muong is the "unsinicized" cousin of Vietnamese is not uncommon. Nevertheless, Muong also seems to attest large numbers of Chinese loanwords:

Table 15. Examples of Sino–Vietnamese Grammatical Words[16]

#	字	Sino-Viet.	Sino-Muong (Khen)	Gloss
1.	在	tại	taj1	prep. "at"
2.	如	như	ɳə1	"like, as"
3.	為	vì	bi2	comp. "because"
4.	只	chỉ	ci5	"only"
5.	個	cái	kaj3	gen. classifier
6.	分	phần	fən2	classifier/"part"
7.	當	đương/đang	taŋ1	present/progressive marker
8.	來	lại	laj4	aspect marker
9.	實	thật	t^hət4	adv, "truly, really"
10.	每	mỗi	moj4	quantifier, "each"

As Table 15 shows, a large number of basic Sino–Vietnamese words also bear Sino–Muong counterparts. There do seem to be important holes (Vietnamese "to get" *được* from Chinese 得, and Vietnamese adversative particle *bị* from Chinese 被, as well as a number of family terms), which may indicate a stratified Chinese effect. However, a large proportion

[16] All Muong Khen data is taken from L. C. Thompson, *Proto-Viet-Muong Phonology* (1976). All tonal transcriptions are non-IPA, and follow a system developed in the field specifically for Muong. A key to this system and a description of its values is included in my forthcoming dissertation.

of these words also occur in Muong, where they do not appear to have been borrowed via Vietnamese (note the non-spirantized b- initial in 3). Direct Muong loans of this kind contradict Maspero's belief that these words were borrowed through Vietnamese. This means that the processes of Viet–Muong diversification and language shift must have overlapped to a great degree.

But how did Vietnamese emerge from this language? As mentioned above, Muong is habitually treated as a cohesive subgroup of the Viet–Muong language family. This requires that the modern Muong dialects mutated together, sharing evolutionary innovations that distinguished them from what eventually became Vietnamese. Contrary to this, and based on newly collected evidence, I propose that what we call "Muong" today does not represent a unified subgroup which split off from Proto-Viet–Muong, but is rather a constructed label artificially uniting a number of linguistic varieties that were already divergent at the "Proto-Viet–Muong" stage.

Linguistic subgroups are determined on the basis of shared innovations, which are linguistic changes shared by a given group of dialects or languages. The following section discusses three salient innovations found in modern Muong, none of which are shared among the Muong dialects under study. The distribution of these innovations suggests a number of distinct lineages instead of a single Muong subgroup.

The Problem of Modern Muong Diversity

This section briefly summarizes my analysis of three sets of data: two sets that correspond to Proto-Viet–Muong implosives *ɓ- and *ɗ- respectively, and one set that corresponds to Proto-Viet–Muong *r-.[17]

Our main set of data is comprised of modern reflexes for Proto-Viet–Muong implosives. An *implosive* is a consonant that is formed through an intake of breath, rather than the conventional explosive. For example, modern Vietnamese <b-> and <d-> are actually implosives when compared with standard English. Ferlus argued that modern Vietnamese nasals (m-, n-) actually derived from an ancient set of implosives, while modern Vietnamese implosives evolved from ancient Proto-Viet–Muong explosives (Ferlus, 1986) These two transformations are part of a complex chain of events that only occurred in Vietnamese, without affecting the other ancient dialects of Proto-Viet–Muong, making them one of the defining innovations of the Vietnamese subgroup[18]. Let us examine the Muong reflexes for these ancient implosives. Table 16 presents data from four dialects of Muong: one was collected by the Barkers and subsequently used by Thompson in his 1975 analysis (Khen), while three were collected by the author in 2009–2010 (Muot, Nabai, and Choi).

Table 16. Muong Correspondences for Ancient PVM Implosive *ɗ-

	Gloss	Viet orth.	Muot	Nabai	Choi	Khen
1.	ought to, must	**n**ên	**r**eiɳ1	**ɗ**eiɳ	**r**ein1	**d**eiɳ1
2.	deer	**n**ai	**r**aj1	**ɗ**aj1	**r**aj1	**d**aj1
3.	sunny/bright	**n**ắng	**r**ăŋ2	**ɗ**ăŋ3	**r**ăŋ3	**d**ăŋ4
4.	water	**n**ước	**r**ak3	**ɗ**ak3*	**r**ak3	**d**ak3

[17] I give fuller treatment to these issues in "Muong is Not a Subgroup" (forthcoming).

[18] Note that the absence of this innovation among all of the Muong dialects does not count as evidence for a Muong "subgroup", because this can be understood as the retention of conservative features. Subgrouping depends on innovations shared by a group of dialects or languages; since all descendents of a particular proto-language could potentially retain any or all features of that language, retentions are not reliable for the articulation of subgroups.

5.	crushed	**n**át	**r**aʲc3	**ɗ**aʲc3	**(p**et6)	**d**ac3

As Table 16 shows, the regular Vietnamese reflex for Proto-Viet–Muong *ɗ- is /n-/. This is true across all dialects of Vietnamese. However, the Muong data clearly demonstrates at least two major correspondences for Proto-Viet–Muong *ɗ- : /r-/ in Muot and Choi, and either /ɗ-/ or /d-/ in Nabai and Khen[19]. Already we see that the innovation of * ɗ- to r- is not shared universally in Muong.

Reflexes for Proto-Viet–Muong labial implosive ɓ- also support a Muot-Choi/Nabai-Khen divide:

Table 17. Muong Correspondences for Ancient PVM *ɓ-

Gloss		Viet orth.	Muot	Nabai	Choi	Khen
1.	lose	**m**ất	**v**ăt3	**ɓ**ət3*	**v**ət3/6	**b**ət4
2.	carry	**m**ang1	**v**aŋ1	**ɓ**aŋ1	**v**aŋ1	**b**aŋ1
3.	salt	**m**uối	**v**oj3	**ɓ**oəj3	**v**waj3	**b**ɔj3
4.	misshapen	**m**éo	**v**ɛw2	**ɓ**ɛw3	**v**ɛw3	**b**ɛw3
5.	salted fish	**m**ắm	**v**ăm3	**ɓ**ăm3	**v**ăm3	**b**ăm3

In these cases, Muot and Choi demonstrate a spirantized *v-*, whereas Nabai and Khen again demonstrate a conservative ɓ/*b*-. Thus, the only thing that groups the four Muong dialects together so far is their lack of the nasalization seen in Vietnamese. But as already noted, the lack of an innovation does not constitute a basis for sub-grouping, while unshared *ɗ- → r and *ɓ- → v- innovations suggest distinct, rather than shared lineages.

It is possible to argue for a Muot–Choi dialectal subgroup since they seem to share two innovations: the rhoticization of dental implosive ɗ-, and the spirantization of labial implosive ɓ-. Interestingly, each of these dialects is spoken on either side of Hoa Bình–Phú Thọ to the north, and Thanh Hóa to the south. Yet if these do form a subgroup, it will be a challenge to explain why members of the same innovating group are split by what appears to be a large body of conservative, non-innovating dialects spoken in Hoa Bình.

Let us now briefly look at the reflexes for Proto-Viet–Muong *r-.

Table 18. Muong Correspondences for Ancient PVM *r-

Gloss		Viet orth.	Muot	Nabai	Choi	Khen
1.	fence	rào	raw2	haw2	raw2	raw2
2.	forest	rừng	rɯŋ2	həŋ2	rɯŋ2	rəŋ2
3.	stubble (rice)	rạ	ra4	ha4	ra4	ra-
4.	intestine	ruột	roʲc4	hwoaʲc4*	roʲc4	rɔc-
5.	spirits	rượi	raw4	haw4	raw4	raw-

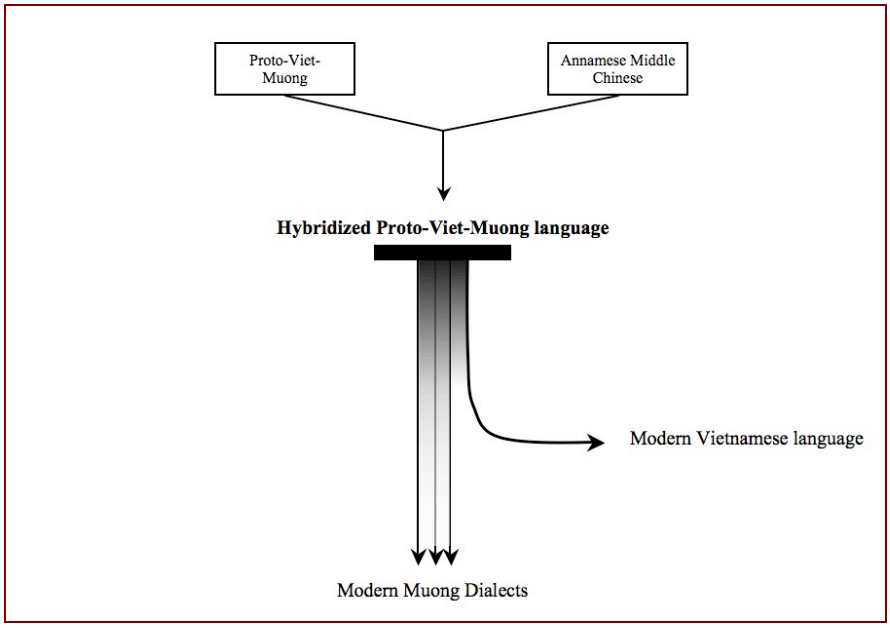
[19]All Khen data in this section is taken from Thompson (1975), who relied on the data of Milton and Muriel Barker. For simplicity, I have rendered the Thompson/Barker data in my own transcriptional system, numbering tones according to their correspondences. It is unclear whether or not the dentals in the Muong Khen data are implosive or not; for this reason, I have faithfully rendered them as "d-".

We see fairly consistent r- realizations across Muong, except for the surprising h- reflexes in Nabai. This is the first positive innovation we have seen in the Hoa Bình dialects, and not a simple retention of Proto-Viet–Muong. At some point, the Nabai dialect transformed r- into h-. Interestingly, this innovation was not shared by Khen, and indeed there is no reason to expect that it would have been, since (once again) the retentions that Tables 16 and 17 indicate that they share, do not demonstrate a sub-grouping relationship.

As already noted, the basis of sub-grouping is shared innovations. In the case of Vietnamese, we have a complex set of changes, a part of which transformed Proto-Viet–Muong implosives into nasals. This set of changes divided Vietnamese from Proto-Viet–Muong, and provides the basis for a Vietnamese subgroup. These facts suggest that it was *Vietnamese* that split off from Proto-Viet–Muong, while an already diverse array of Proto-Viet–Muong varieties eventually evolved into the modern Muong dialects.

In this model, Vietnamese is properly understood as a subgroup while Muong is not. Because the Muong dialects are mutually intelligible, they might be understood as a single language, although this is misleading in terms of their evolutionary history. All Muong dialects, as well as Vietnamese, descend from an array of Proto-Viet–Muong dialects that had already been sinicized. Then, a single lowland dialect or possibly a subset of them split off together, forming their own sub-group and eventually evolving into modern Vietnamese. The remainder continued to evolve, but never mutually innovated into a new group (until possibly very much later).

Figure 2. Model for Viet–Muong Diversification[20]



[20] While I use the term "hybridized" in this scheme, I do not here make the claim that Proto-Viet–Muong was a true hybrid language, i. e., a creole that developed from a pidgin. The sense is weaker here, and is meant only to denote a strong adstratum effect from Annamese Middle Chinese.

What is required to prove the validity of a Muong subgroup is an innovation shared by all Muong dialects, but not shared by Vietnamese. And yet, all we seem to have that unifies Muong are retentions, not innovations. To name a few, these include the preservation of medial liquid clusters tl- and kl- (cf. Muot **tl**eiɳ for Viet. "trên", "on"; **tl**ɯj for "trời", "heaven"), final liquid –l (Khen po**l** for Viet. "vôi"; "lime"), and low-front vowel –a- (Choi r**a**k for Viet. "nước"; "water"). These characteristics are shared by all the Muong dialects shown above; yet as retentions all they really demonstrate is the validity of a *Vietnamese* subgroup.

In Part Two we identified three true innovations, but found none of them was shared universally. They were:

1) Rhoticization of PVM *ɗ-: *ɗ- → r-

2) Spirantization of PVM *ɓ-: *ɓ- → v-

3) Laryngealization of PVM *r-: *r- → h-

Rather, Muong only appears to assume unity in converse with Vietnamese, and then only by lacking the innovations that substantively define Vietnamese. In summary, we found no innovations to be shared across Muong, but we did find three innovations distributed unevenly across the dialects.

In light of these observations, we can only conclude that Muong is not a subgroup, but that it represents the ongoing evolution of an already-diversified group of Proto-Viet–Muong dialects, from which Vietnamese split sometime *after* the language shift event described in the last section.

Of course, today the Muong dialects are mutually-intelligible whereas Vietnamese is not mutually-intelligible with any of them, and this seems to imply that there probably are real innovations shared potentially among all dialects. One potential candidate is a flip-flop of high and low tone features, in which historically "high" tones are realized as low, and visa versa. It should be noted, however, that this process is also currently affecting northern Vietnamese, and it could simply be a regionally shared feature spontaneously (rather than promiscuously) occurring around the area. Nevertheless, other innovations might be discovered. Given the divergent innovations examined above, however, it seems likely that these would be late innovations, meaning that they demonstrate a kind of convergence, rather than a shared divergent history. For now, it is safe to conclude that Muong is not a single subgroup.

Summary and Conclusion

In Part One, we populated Red River Plain society with an elite, Chinese-speaking sector, who spoke a local dialect of Middle Chinese, and who eventually shifted allegiance away from this language in favour of the contemporary, non-Chinese language that coexisted with it. Part Two determined that this non-Chinese language was Proto-Viet–Muong, and that an already sinicized Proto-Viet–Muong had diversified into a number of dialects before a subset of those dialects innovated into a new language: Vietnamese. The remaining dialects eventually became what we now call Muong.

These facts allow us to re-imagine the landscape of Annam in a way that frees it from the binding, essentialist narrative of Vietnamese defiance and survival, and allows us to glimpse the processes through which Vietnamese emerged as a unique language. Let us summarize the principle features of this hypothesis. The socio-political elite of the Red River plains shifted from Annamese Middle Chinese to Proto-Viet–Muong. Throughout this

multi-generational process, Chinese words were diffused throughout Proto-Viet–Muong, and across a presumed dialectal diversity of the language. At this point, we know that a subset of dialects innovated together to form Vietnamese, and the question of what catalysed this mutation—and why certain dialects were left out of it—that is tantalizing. Perhaps the absence of a distinct (and mutually unintelligible) elite language spurred a socio-linguistic transformation. More likely, the intense regional turmoil of the tenth century fractured the stability—and thus, linguistic interconnectivity—of the Red River plains, leading to the speciation of Vietnamese within a narrower region of Annam, while other areas evolved along different paths. Further research promises to shed light on this fascinating question.

Nevertheless, this new model for tenth-century Red River plains' Annamese society allows us to challenge the applicability to it of notions like "Chinese" and "Vietnamese" on a very basic level. There is a very good reason why the idea of the survival of the Vietnamese language is so powerful: like genetics, language is sometimes perceived as a kind of trans-temporal vessel for essential identity—both a repository and a tether, linking a particular group back through the centuries to a purer, more fundamental form. Perceived origin is privileged over the intervening evolution, and the whole history of a group or language becomes reduced to a kind of "fall from grace" storyline. This is only made slightly more complex by a transparent elegy for the "adaptability" or "borrowing ingenuity" of a given culture or language, which essentially serves to mask a survival narrative in more palatable terms.[21]

In fact, the privileging of origin over evolution is fundamentally problematic. It can reify an imagination of the distant past over a traceable, quantifiable, describable history, in the process promoting a kind of teleological levelling that erases the richness and complexity of the past by collapsing multi-dimensionality into a featureless plane. Ironically, language can indeed act as a kind of repository—not of a national quintessence, but of the marks of change, cross-pollination, and evolution that comprise the special history of a given group of speakers.

The linguistic record of the Red River plains challenges a set of paradoxical assumptions that have infiltrated the atmosphere of Vietnamese studies. On the one hand, it is popular to believe—even just a little—in the "Vietnameseness" of the first millennium polities, even (and often) as far back as the Dongsonian rebellions of the Trưng sisters. On the other hand, it is also popular to understand lowland Vietnamese culture as sinicized and therefore, civilized, as opposed to the primitive cultures of the highlands. These are clearly simple or ugly stereotypes that are logically and historically indefensible, yet they find their way into the most sophisticated works, a sort of nebulous cloud that informs the premises of arguments bent on other issues.

The arguments I have made here—the existence of a local dialect of Middle Chinese, its obsolescence in favour of Proto-Viet–Muong, and the eventual emergence of a new language from among the hybridized dialects that followed—plainly challenge the notion of a multi-millennial "Vietnamese" identity. They strongly imply that a recognizable culture for the Vietnamese, like their language, formed during the first few centuries of independent kingship, rather than in a vague and distant, pre-Chinese era. They also challenge our understanding of the depth and breadth of Chinese influence on the peoples of the Red, Ma, and Ca River plains, not only by proposing a workable model of language shift, but also by demonstrating how the consequences of that shift form the common heritage of both Vietnamese and Muong languages. Most critically, they comprise a new, if as yet still preliminary, narrative for the birth of the Vietnamese language, not in the depths of an

[21] See Nguyễn Đình Hoa's very important "Vietnamese Creativity in Borrowing Foreign Elements", in *Borrowing and Adaptation in Vietnamese Culture*, ed. T. B. Lam (Honolulu: Center for Asian and Pacific Studies, University of Hawaii, 1987), pp. 22–40.

imagined past but via a rich and complex history that comes down to us through the speech of the people still living in the plains of northern Vietnam.

Sources:

Alves, M.J. (1999). "What's so Chinese about Vietnamese?" in *Papers from the ninth annual meeting of the Southeast Asian Linguistics Society*, ed. *G.* Thurgood, pp. 221–242.

Baxter, W. H. (1992). *A Handbook of Old Chinese Phonology.* Berlin & New York: Mouton de Gruyter.

Cuisinier, J. (1946). *Les Muong: geographie humaine et sociologie.* Paris: Institut d'Ethnologie.

Diffloth, G. (1991). "Vietnamese as a Mon-Khmer Language." *Papers from the First Annual Meeting of the Southeast Asian Linguistics Society* , pp. 125–39.

Ferlus, M. (1994). "Contact ancien entre Viet-Muong et Austroasiatic nord." *Kristina Lindell Symposium on Southeast Asian Studies*, pp. 1–4.

Ferlus, M. (1986). "Histoire abrégée de l'évolution des consonnes initiales du Vietnamien et du Sino-Vietnamien." *Mon–Khmer Studies, 20*: 111–25.

Ferlus, M. (1989-90). "Sur l'origine des langues Viet-Muong." *Mon–Khmer Studies* 18–19: 52–59.

Ferlus, M. (1975). "Vietnamien et proto-viet-muong." *Asie du Sud-Est et monde insulindien 6,* 4: 21 55.

Guillaume, J. (2006, March). "An Introduction to Chinese Historical Phonology." Handout for a workshop on historical Chinese phonology. Leiden: Leiden University.

Hashimoto, Mantaro J. (1968). "Current Developments in Sino-Vietnmaese Studies." *Journal of Chinese Linguistics,* 6: 1-26.

Haudricourt, A. (1954). "De l'origine de tons en vietnamien." *Journal Asiatique,* 242: 69–82.

Hock, H. H. (1991). *Principles of Historical Linguistics* (second revised & updated edition). Berlin & New York: Mouton de Gruyter.

Jamieson, N. (1993). *Understanding Vietnam.* Berkeley: University of California Press.

Ladefoged, P. (1996). *The Sounds of the World's Languages.* Cambridge, Massachusettes: Blackwell Publishing.

Lewis, M. P. (2009). *Ethnologue, SIL International*. Retrieved July 16, 2010, from *Ethnologue*: Languages of the World, 16th ed: http://www.ethnologue.com/show_language.asp?code=mtq

Maspero, H. (1912). "Etudes sur la phonétique historique de la langue annamite, Les initiales." *BEFEO 12*: 121–26.

Miyake, M. (2003). *Old Japanese: A Phonetic Reconstruction.* London & New York: RoutledgeCurzon.

Nguyen, D. H. (1987) "Vietnamese Creativity in Borrowing Foreign Elements", in *Borrowing and Adaptation in Vietnamese Culture*, ed. T. B. Lam (Ed.), Honolulu, Hawaii: Southeast Asian Studies, Center for Asian & Pacific Studies, University of Hawaii. pp. 22-40.

Nguyễn, K. V. (1993). *Vietnam: A Long History.* Hanoi: The Gioi Publishing.

Nguyễn, V. T. (1982). *Ngữ âm tiếng Mường qua các phương ngôn.* [English translation], Hanoi.

Pulleyblank, E. (1984). *Middle Chinese, a Study in Historical Phonology.* Vancouver: University of British Columbia Press.

Pulleyblank, E. (1991). *Lexicon of Reconstructed Pronunciation in Early Middle Chinese, Late Middle Chinese, and Early Mandarin.* Vancouver: University of british Columbia Press.

Rischel, Jørgen. (1995). *Minor Mlabri: a Hunter-Gatherer Language of Northern Indochina*. Copenhagen : Museum Tusculanum Press, University of Copenhagen.

Sidwell, P. (2009). *Mon-Khmer.com: Vietic*, retrieved July 19, 2010, from Mon-Khmer.com: http://people.anu.edu.au/~u9907217/languages/Vietic.html.

Taberd, J. (1838). *Dictionarium Anamitico-Latinum.* Serampore: Marshman.

Taylor, K. (1985). *The Birth of Vietnam.* Berkeley, California: Berkeley University Press.

Taylor, K. W. (2001). "On Being Muonged." *Asian Ethnicity,* 2: 25–34.

Thompson, L. C. (1976). "Proto-Viet-Muong Phonology." *Oceanic LInguistics Special Publications, No. 13, Austroasiatic Studies Part II* , 1113–1203.

Trần, Đ. S. (1997). "Từ Hán Việt Gốc Nhật trong Tiếng Việt." ["Sino-Vietnamese words of Japanese origin in Vietnamese"] *Thông Báo Hán Nôm Học* , 552–59.

Wang, L. (1948). 漢語越語研究 Hanyueyu yanjiu [Research on Sino-Vietnamese]. *Lingnan Xuebao* 91: 1–96.

Winford, D. (2003). *An Introduction to Contact Linguistics.* Malden, Massachusettes: Blackwell Publishing.

Young Lives: Maps & Regions. (n.d.), retrieved July 19, 2010, from Young Lives: An International Study of Childhood Poverty: http://www.younglives.org.uk/countries/vietnam/map-and-regions