NOTES ON THANH-CHƯƠNG VIETNAMESE IN NGHỆ-AN PROVINCE[1]

Mark J. Alves
Montgomery College, Maryland
<Mark.Alves@montgomerycollege.edu>

Nguyễn Duy Hương
Institute of Linguistics, Hanoi

1 Introduction

Vietnamese dialectal variation in terms of pronunciation is generally divided into three main regions, namely Northern, Central, and Southern Vietnamese (NV, CV, and SV hereafter).[2] Typically, each region is characterized by the varieties[3] of Vietnamese spoken in major cities of those regions, Hà-Nội (considered in Vietnam to be 'Official Vietnamese,' OV hereafter), Huế, and Hồ-Chí-Minh respectively, though Vietnamese in Vinh is also sometimes recognized as a Central variety distinct from that of Huế. However, few dialect studies on Vietnamese have concentrated on the numerous tiny and isolated villages scattered throughout Vietnam. Reports of their conservative characteristics (cf. Châu 1989, Nguyễn T. C. 1995) suggest that some of these geographically remote varieties

---

[1] We wish to thank all the consultants who participated in this study. The data was collected in September of 1997. One of the researchers in this study, Mr. Nguyen, is a native of that region and was crucial in eliciting the local vocabulary. Speech samples were recorded using the Summer Institute of Linguistic's 'Cecil' acoustic phonetic software. To elicit the data, three word-lists were used, including one-syllable words (46 words), two-syllable words (57 words), and complete sentences (15 sentences). The word-lists were chosen primarily for tonal contrasts, though substantial segmental variety was also included. The sentence list--which contained simple, daily sentences--was created based on known regional vocabulary. Most of the recorded data came from two native speakers of TC Vietnamese. Recordings made with the help of speakers from other regions other than Nghệ-An were representative of Official Northern and Southern Vietnamese, which allowed a point of reference for comparison.

We recognized the hazard of using wordlists to elicit natural speech, so we employed a few devices to reduce this interference as much as possible. First, we persuaded our consultants to use their most natural 'family' pronunciation, which most readily admitted was more comfortable than reading in OV. Next, Mr. Nguyễn, as a native speaker of TCV with linguistic training and familiarity with dialect differences, helped to identify the use of OV. Finally, the sentence list was actually 'translated' by the consultants into their native variety, and was not read by the consultants, which created a point of reference to help screen out reading pronunciations.

[2] See Thompson 1984-85 for a general phonetic description of those regions.

[3] The term 'variety' is used throughout most of the paper rather than 'accent' or 'dialect' in order to avoid the controversy in deciding which category is accurate. Based on the significant phonological and lexical differences between TCV and OV, TCV is more like a 'dialect' than a minor regional 'accent,' though this is still relative to the region in which it is spoken where the differences may be said to be simply accent. Perhaps the TCV region in Nghệ-An province, including different accents spoken in neighboring areas can, altogether, be considered a 'dialect' with varying local 'accents,' though further research is needed to clarify the situation.

of Vietnamese are largely untapped storehouses of historical linguistic information about the Vietnamese language.

This paper describes both the modern and historical characteristics of one variety, Thanh Chương Vietnamese (TCV hereafter), a variety of Vietnamese from the dialectally diverse region of North-Central Vietnam.[4] Thanh Chương is a district in the western hills of Nghệ-An province, North-Central Vietnam, a short distance from the border of Laos. Field data has shown TCV to be a highly conservative variety of Vietnamese. In the following sections, the archaic nature of TCV is illustrated in terms of Vietnamese phonology and etymologies. After a summary of the phonological characteristics of TCV, those features are compared with Mường, Rục[5] (a Minor Vietic[6] language), and Pacoh (a language of the Katuic branch of Mon-Khmer). Then, the etymological layers of TCV (regional and non-regional Vietnamese, Vietic, and Mon-Khmer) are discussed, with highlights of Mon-Khmer vocabulary not present in modern OV.[7]

2 TCV Phoneme System

TCV has maintained all of the segmental distinctions as represented by the national Quốc-Ngữ orthography. This is generally true among varieties of North-Central Vietnamese, possibly making this the regions with the highest degree of conservatism, according to currently available data. The sound systems of Vietnamese in Hồ-Chí-Minh, Ha-Nội, and Huế have all undergone various phonological mergers resulting in smaller phonemic inventories. While TCV has 24 consonants, SV and mid-CV have 22, and NV has only 20. NV has no retroflex consonants, which merged with palatals, and /j/ and /r/ merged with /z/. From Huế to Hồ-Chí-Minh city, [v] and [z] have merged with [j], and dentals have merged with velars in syllable-final position. TCV has retained all of those phonemic distinctions, as shown in Table 1.

Most varieties of Vietnamese have preserved all the vowel phoneme categories seen in the Quốc-Ngữ orthography, though phonetic differences exist.[8] TCV vowels are, as are NV vowels, phonetically close to the indicated orthographic representation, having few allophonic variants. This differs from the phonetic characteristics of SV and mid-CV

---

[4] Both North and Central Vietnam show a higher degree of dialectal diversity than the much younger South, which had only been inhabited by the Vietnamese after the fall of the Champa Empire in the end of the fifteenth century. Whether the North or Central region contains more diversity is not something that has been investigated to our knowledge, though our studies of the Vietnamese of several villages in Nghệ-An suggest that this area is indeed highly diverse and may contain at least a few distinct dialects.

[5] The primary source is Nguyễn V. L. 1993.

[6] The term 'Vietic' (Hayes 1992) refers to the group of languages including Vietnamese and Mường, which together form a subgroup of Vietic, and the two dozen or so groups of archaic languages (see Ferlus 1974 and 1975 for a list and subgrouping), which have been called 'Minor Vietic' (Alves, 2003).

[7] Tai-Kadai cognates are not discussed as they are neither genetically nor statistically relevant. This paper was written with the assumption that sufficient data has been collected to date in order to firmly place Vietnamese in the Mon-Khmer language group. Gage (1985) claimed to have found only 3% of Vietnamese vocabulary as possibly being cognate with Tai-Kadai forms, while Thomas and Headley (1970) found about 2X% of Mon-Khmer forms.

[8] Again, see Thompson 1965 for a discussion of the differences.

(e.g. vowel centralization in closed syllables, such as /i/ to /ɨ/ and /e/ to /ə/ and monophthongization of diphthongs).

Table 1: *Thanh-Chương Consonants*[9]

	lab	den	pal	ret	vel	glot
voiceless stop	-p	-t-	-c-	ṭ-	k-	
voiced stops	b-	d-			g-	
Voiceless continuants/ aspirates	f-	tʰ-	s-	ṣ-	x-	h-
voiced continuants	v-	z-				
nasal	-m-	-n-	-ɲ-		-ŋ-	
glides/liquids	-w-	l-	-j-	ṛ-		

Besides consonants and vowels, TCV tones also differ from other mainstream Vietnamese varieties. Table 2 lists the phonetic realizations of tones in Hà-Nội, Nha-Trang, and Thanh-Chương. Among many varieties of Vietnamese, only five tonal phonemes remain of the six represented orthographically, most often due to a collapsing together of the hỏi and ngã tones. Such is the case for TCV as well as SV and CV. The sắc tone does have more than one phonetic realization in TCV and Hà-Nội Vietnamese depending on the syllable final, whether a voiceless stop or a sonorant (vowel or nasal). The exact value of the open-sắc tone in TCV is left undetermined in Table 2 because that tone showed two different phonetic forms, a low-level tone (11) and a low-rising glottalized tone (13g).

Table 2: *Tone Systems of Vietnamese as Spoken in Hà-Nội, Nha-Trang, and Thanh-Chương*[10]

TONE	HN	NT	TC
ngang	33	33	35
huyền	32	32	33
sắc, open	24	45	11/13g
sắc, closed	45		45
nặng	22g	23	22
Hỏi	31	24	31
Ngã	35g		

---

[9] IPA is used for much of the paper, though the Vietnamese Quốc-Ngữ orthography has been used for ease of reading and reference. Hyphens are used in the table to indicate the range of distribution of the phonemes in syllables, whether strictly in word-initial position, such as the continuant series (e.g. /x-/), word-final position (only /-p/), or both positions, such as the nasal series (e.g. /-ŋ-/).

[10] The Y.R. Chao system is being used to represent tone throughout this paper. The number 5 represents the highest pitch level, and 1 is the lowest. The first number represents the starting point and the second, the end point. In this paper, the letter 'g' indicates that the tone is clearly glottalized.

Table 3: *Correspondences of Standard and TCV*

Gloss	Comparison QN[11]	IPA	OV	IPA	TCV	IPA
1. gourd	âu : u	əw : uː	bầu	bəw²¹	bù	buː¹¹
2. bear			gấu	gəw²⁴	gụ	guː¹¹
3. deep			sâu	səw¹¹	su	ʂuː¹¹
4. buffalo			trâu	cəw¹¹	tru	ʈuː¹¹
5. road	ươ : a	iə : aː	đường	diəŋ²¹	đàng	daːŋ³³
6. person			người	ŋɨəj²¹	ngài	ŋaːj³³
7. grill			nướng	nɨəŋ²⁴	náng	naːŋ¹¹
8. fire	ưa : a		lửa	lɨə³¹	lả	laː³¹
9. itchy			ngứa	ŋɨə²⁴	ngá	ŋaː¹¹
10. female	ay : ây	aj : əj	cái	kaːj²⁴	cấy	kəj¹¹
11. urinate			đái	daːj²⁴	đấy	dəj¹¹
12. fruit			trái	caːj²⁴	trấy	ʈəj¹¹
13. you	ay : i	aj : iː	mày	maj²¹	mi	miː³⁵
14. this			này	naj²¹	ni	niː³⁵
15. same	ô : u	oː : uː	giống	zəuŋ^m24	giũm	zuːm³¹
16. knee			gối	goːj²⁴	cúi	kuːj¹¹
17. I			tôi	toːj³³	tui	tuːj³⁵
18. goat	ê : ia	eː : iə	dê	zeː³³	dia	jiə³⁵
19. son-in-law			rể	zeː³¹	rỉa	riə³³
20. pastry	anh : eng	aɲ : ɛːŋ	bánh	baɲ²⁴	béng	bɛːŋ¹¹
21. soup			canh	kaɲ¹¹	keng	kɛːŋ³⁵
22. fishy			tanh	taɲ¹¹	teng	tɛːŋ³⁵
23. thatch			tranh	caɲ¹¹	treng	ʈɛːŋ³⁵
24. blue /green			xanh	saɲ¹¹	xeng	xɛːŋ³⁵
25. pile	-ông/-ôông	əuŋ^m : oːŋ	chồng	cəwŋ^m21	nhôông	ɲoːŋ³⁵
26. trunk			gốc	gəuk^p45	côộc	koːk¹¹
27. to plant			trồng	ʈəuŋ^m21	lôông	loːŋ³⁵

A quick examination of the features of earlier stages of Vietnamese and Vietic[12] shows how conservative TCV phonology is. There are a few general patterns of

---

[11] Some of the Quốc-Ngữ symbols used in this table are part of a non-standard, vernacular convention used to transcribe differences in regional pronunciations.
[12] For a description of earlier stages of Vietnamese and Vietic, see Nguyễn T.C. 1995.

phonological correspondences in certain lexical items between OV and TCV,[13] which are illustrated in table 3.[14] First, TCV has preserved forms that have not undergone diphthongization (examples 1 to 9, 13, and 14). Next, as mentioned, TCV has preserved initial retroflex sounds, a characteristic of SV and CV (examples 4 and 12). Then, there are two cases of final vowel-consonant assimilation (examples 21 to 25). TCV has preserved a few cases of the earlier Vietic sequence /-ɛŋ/ for which OV has only the palatalized /-æɲ/. Interestingly enough, two of the forms are actually old Chinese loanwords (examples 20 and 24). Also, TCV has not completed the process of rounding final consonants after backround vowels (examples 25 to 27), a process completed in OV but not in Mường dialects (cf. Nguyễn M.D. 1972).

3 TCV Etymological Layers

TCV vocabulary has at least five identifiable etymological layers. These include (1) non-regional vocabulary (Vietnamese common to all regions of Vietnam), (2) regional vocabulary generally used in Central Vietnam, (3) local vocabulary used in and around Thanh-Chương district, (4) Vietic vocabulary not seen in OV, and (5) some Mon-Khmer vocabulary not seen in OV. The significance of the latter four categories is that they are in contrast with OV, which shows that vocabulary in TCV's region has historically gone down a somewhat different etymological road. In some cases, the development is very localized with no identifiable external source. In other cases, TCV appears to have preserved genetically higher-level forms (i.e. from Vietic and Mon-Khmer) that were not maintained in mainstream Vietnamese. Finally, some TCV words are archaic forms that may predate Chinese loanwords and help identify an early layer of Chinese loanwords not belonging to the literary Tang-Song era Sino-Vietnamese.

3.1 *Non-Regional Vietnamese Vocabulary*

For the most part, the TCV lexicon consists of basic and culturally specific words used throughout Vietnam. This pan-Vietnam vocabulary is also multi-layered, including that which is Vietic, Mon-Khmer,[15] Chinese, Tai-Kadai, and purely Vietnamese in origin. As a result, though Vietnamese speakers from outside the North-Central region may have difficulty comprehending the TCV accent, there is overall mutual intelligibility. Furthermore, speakers of TCV are aware of the lexical differences between their local vocabulary that of OV, which is often used on Vietnamese TV, radio, and in schools. Speakers of TCV are even aware of the lexical and phonological differences between their speech and that of nearby neighbors. Regardless of the differences between TCV and OV, all TCV speakers recognize their linguistic and ethnic connections, a fact supported by lexical and phonological evidence.

---

[13] Maspero 1912 included data on various dialects throughout Vietnam. Some of the characteristics that occurred in our data were still present in Maspero's time.

[14] The acronyms used in the table are as follows. QN (Quốc-Ngữ), IPA (International Phonetic Alphabet), OV (Official Vietnamese), and TCV (Thanh Chương Vietnamese).

[15] For a good list of the core Mon-Khmer vocabulary in Vietnamese, see Huffman 1977.

3.2 *Regional Vocabulary: Central Vietnamese*

TCV contains numerous vocabulary items that are associated with Central Vietnam. Speakers in both Vinh and Huế either use such forms or are aware that these are lexical items restricted to Central Vietnam. Table 4 contains a few examples.

Table 4: *Central Vocabulary in Thanh-Chương Vietnamese*

Gloss	OV	IPA	TCV	IPA
1. (reciprocal)	nhau	ɲaw^{33}	chắc	cak^{11}
2. no	không	xauŋ^{m33}	nỏ	nɔː^{31}
3. this	này	naj^{21}	ni	niː^{35}
4. where	đâu	dəw^{33}	mô	moː^{35}

These forms are not seen with complete consistency throughout the region. The negation word was claimed by some speakers in urban areas to be restricted to rural areas.

3.3 *Local Vocabulary: Thanh-Chương*

Some TCV vocabulary is restricted further to the region in and around Thanh-Chương. A few examples are shown in Table 5.

Table 5: *Local Thanh-Chương Vocabulary*

Gloss	OV	IPA	TCV	IPA
1. we (excl.)	chúng ta	$\text{cuːŋ}^{m24}\,\text{taː}^{33}$	choa	cwaː^{35}
2. no need	không bảo	$\text{xəuŋ}^{m33}\,\text{baːw}^{31}$	hung nhủ	$\text{huŋ}^{m35}\,\text{ɲuː}^{31}$
3. which	nào	naːw^{21}	mồ	moː^{33}
4. remain	còn	kɔːn^{21}	nưng	nɨːŋ^{35}

More definitive statements about the precise geographic range of the usage of these forms can only be made after more dialect studies are conducted in this region. Tables 4 and 5 are just a taste of the wide-ranging lexical diversity that has been witnessed in Vietnamese dialect studies.

3.4 *Vietic Vocabulary: Muong and Minor Vietic*

TCV contains some lexical forms that appear to be cognate with Mường and a Minor Vietic language, Rục. Some of those cognates illustrate phonological changes that occurred between Proto-Vietic and modern OV. In Table 6, the change from the cluster *[tl][16] to a single retroflex [ʈ] is seen in example 2. Examples 3 and 9 illustrate the change from a final *[l] to [n] in TCV, which is lost in OV. Example 8 illustrates the rounding of finals after back round vowels.

Examples 2 and 7, and possibly 4 and 5, are examples of lexical preservations in which the original Vietnamese forms have been replaced by Sino-Vietnamese terms. The words 'head' and 'tiger' are lexical preservations in TCV, while the modern OV form is

---

[16] These clusters are attested in seventeenth century Vietnamese as seen in Alexandre de Rhodes' dictionary (1651).

indisputably Chinese in origin.[17] Rục, an extremely conservative language even among Minor Vietic languages, lends support to the claim that these TCV lexical forms are genuine Vietnamese words and not Chinese loans.

Table 6: *Thanh-Chương Cognates with Mường and Rục*

Gloss	OV	IPA	TCV	IPA	Others
1. day	hôm	hoːm^{33}	bựa	bɨːə^{22}	bưa (Rục)
2. head	đầu	dəw^{21}	trôôc	ʈoːk^{13}	tlok (Mường) kulôok^{4} (Rục)
3. light	nhẹ	ɲɛː^{22}	nhẹn	ɲɛn^{22}	nhel^{3} (Rục)
4. place	chỗ	coː^{35}	lỗ	lo^{31}	lô^{4} (Rục)
5. still	còn	kɔn^{21}	lưa	lɨə^{35}	lơ^{1} (Rục)
6. there	kia	kiːə^{33}	nớ	nəː^{13}	na^{3} (Rục)
7. tiger	hổ	hoː^{31}	khái	xaːj^{13}	khal (Mường)
8. tired	mệt	met^{22}	nhọc	ɲaukᵖ^{2}₂	nhôok^{4} (Rục)
9. tree	cây	kəj^{33}	cơn	kəːn^{35}	kăl (Mường)

Table 7: *Comparison of TCV and Non-OV Vietic Vocabulary*

Gloss	HN	IPA	TC	IPA	Katuic/Vietic
Far	xa	saː^{33}	ngái	ŋaːj^{11}	*səŋaːj (Proto-Katuic) cŋaay (Khmer) chơngaj^{35} (Rục)
to slice	thái	tʰaːj^{24}	sát	ʂat^{11}	*ciət (Proto-Katuic) klat (Pacoh) sat (Bru)
which	nào	naːw^{21}	mô	moː^{35}	*ʔdəməː (Proto-Katuic) chamơ/mô (Rục)
defecate	iả	ʔiə^{35}	e	ʔɛː^{35}	*ʔɛh (Proto-Katuic)
to like	thích	tʰic^{45}	ưng	ʔɨːŋ^{35}	ính (Pacoh)
small pot	niêu	niəw^{33}	tréc	ʈɛk^{11}	trek (Pacoh)
to prune	thiến	tʰiən^{24}	lặt	lat^{22}	lăt (Pacoh)

3.5 *Mon-Khmer Vocabulary not in OV*

The final etymologies discussed here include some possible Mon-Khmer cognates that are not present in OV. In table 7, the TCV forms are compared with apparent Katuic

---

[17] Such Sino-Vietnamese forms are readily checked in Sino-Vietnamese dictionaries, though the possibility still remains that these are coincidental look-alikes. However, contrasting evidence in other Vietic languages suggest otherwise.

cognates,[18] and when notable, other non-Katuic forms. This evidence is suggestive of a close historical relationship between Vietic languages and the Katuic branch, though whether it is a genetic relationship, contact relationship, or both certainly cannot be determined by this meager amount of data.

4   Concluding Comments

Clearly, the amount of lexical and phonological difference between OV and TCV suggests that further research on Vietnamese dialects would turn up more diversity that could broaden our understanding of the relationship between Vietnamese and other neighboring languages and language groups in the region. Unfortunately, recent years have seen the neutralization of dialects along the increasingly modernized roads along the Vietnamese coastline. It is only a matter of time before that neutralization reaches places like Thanh-Chương and thereby eradicates what is a wealth of historical information.

References

Alves, Mark J. 2003. Ruc and other Minor Vietic languages: Linguistic strands between Vietnamese and the rest of the Mon-Khmer language family. *Papers from the Seventh Annual Meeting of the Southeast Asian Linguistics Society*, ed. by F. K. Lehman (Chit Hlaing).

Hoàng, Thị Châu. 1989. *Tiếng Việt trên các miền đất nước (phương ngữ học)* (Vietnamese in regions of our land (dialect studies)). Hà Nội: Nhà Xuất Bản Khoa Học Xã Hội.

De Rhodes, Alexandre. 1991 (Originally published in 1651). *Từ-điển Annam-Lusitan-Latinh (Thường gọi Từ điển Việt-Bồ-La)*. Hồ Chí Minh: Nhà Xuất Bản Khoa Học Xã Hội. (Both original Latin and Vietnamese translation).

Ferlus, Michel. 1974. Le groupe Viet-Muong. *Asie de Sud-Est Monde Insulindien* 5.1:69-77.

_____. 1975. Vietnamien et Proto-Viet-Muong. *Asia Sud Est un Monde Insulindien* 6.4:21-55.

Gage, William W. 1985. Vietnamese in a Mon-Khmer perspective. *Southeast Asian Linguistics Presented to Andre-G. Haudricourt*, ed. by S. Ratankul, D. Thomas, and S. Premisirat, 493-524.

Hayes, La Vaughn H. 1992. Vietic and Việt-Mường: a new subgrouping in Mon-Khmer. *Mon-Khmer Studies* 21:211-228.

Hoàng, Thị Châu. 1989. *Tiếng Việt Trên Các Miền Đất Nước (Phương Ngữ Học)* (Vietnamese in regions of the nation (dialectology)). Hà Nội: Nhà Xuất Bản Khoa Học Xã Hội.

Huffman, Franklin. 1977. An examination of lexical correspondences between Vietnamese and some other Austroasiatic languages. *Lingua* 43:171-198.

Nguyễn, Minh Đức. 1972. *Một vài nét về các thổ ngữ của tiếng Mường Hòa Bình* (Some notes of the indigenous languages of the Muong of Hoa Binh), Tìm hiểu ngôn ngữ các dân tộc thiểu số ở Việt Nam (Understanding the languages of the ethnic minorities of Vietnam): 6-24. Hà Nội: Viện Ngôn Ngữ Học.

---

[18] All reconstructed forms are taken from Pejros 1996.

Nguyễn, Tài Cẩn. 1995. *Giáo trình lịch sử ngữ âm tiếng Việt* (Textbook of Vietnamese historical phonology). Hà Nội: Nhà Xuất Bản Gíao Dục.

Nguyễn, Văn Lợi. 1993. *Tiếng Rục* (The Ruc language). Hà Nội: Nhà Xuất Bản Khoa Học Xã Hội.

Pejros, Ilia. 1996. *Katuic Comparative Dictionary*. The Australian National University, Pacific Linguistics Series C-132.

Thomas, David D and Robert K. Headley Jr. 1970. More on Mon-Khmer subgroupings. *Lingua* 25:398-418.

Thompson, Laurence C. 1984-85. A Vietnamese reference grammar. *Mon-Khmer Studies* 13-14. University of Hawai'i Press.