

The Vietnamese Language

THANH VIET CAO (高成越)

INSTITUTE OF LINGUISTICS
NATIONAL TSING HUA UNIVERSITY

MAY 1ST, 2025



OUTLINE

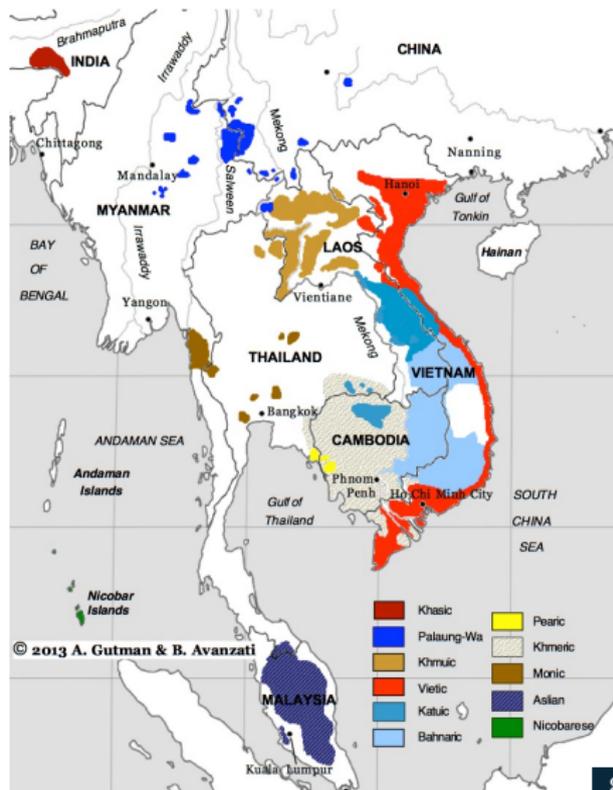
- 1 History of Vietnamese
- 2 Linguistic profile of Vietnamese
- 3 Sound System
- 4 Writing Systems
- 5 Lexicon
- 6 Morphology & Grammar
- 7 Cultural Notes



History of Vietnamese

VIETNAMESE AS AN AUSTROASIATIC (AA) LANGUAGE (1)

- Vietnamese belongs to the *Viet-Muong* group, the *Vietic* branch, the *Austroasiatic* language family.
- However, unlike most other Austroasiatic languages, Vietnamese has **tones** and doesn't have a proper **inflectional morphology**.



VIETNAMESE AS AN AA LANGUAGE (2)

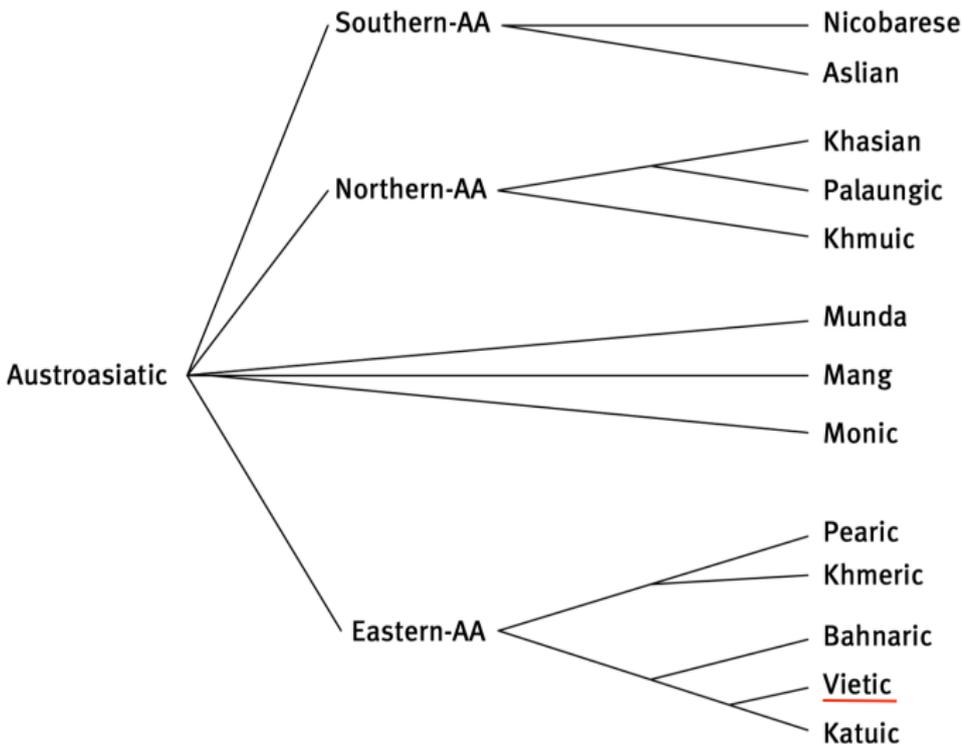


Figure 1: A classification of AA branches (Sidwell 2021)

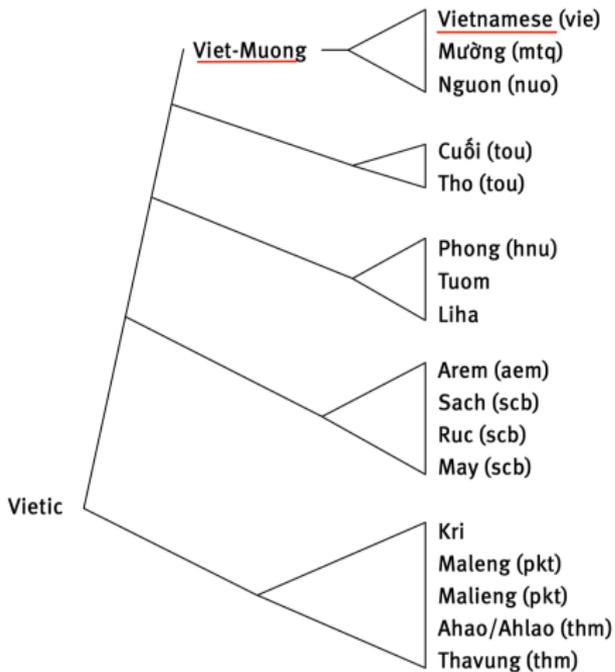


Figure 2: Vietic classification following Sidwell & Alves (2021)

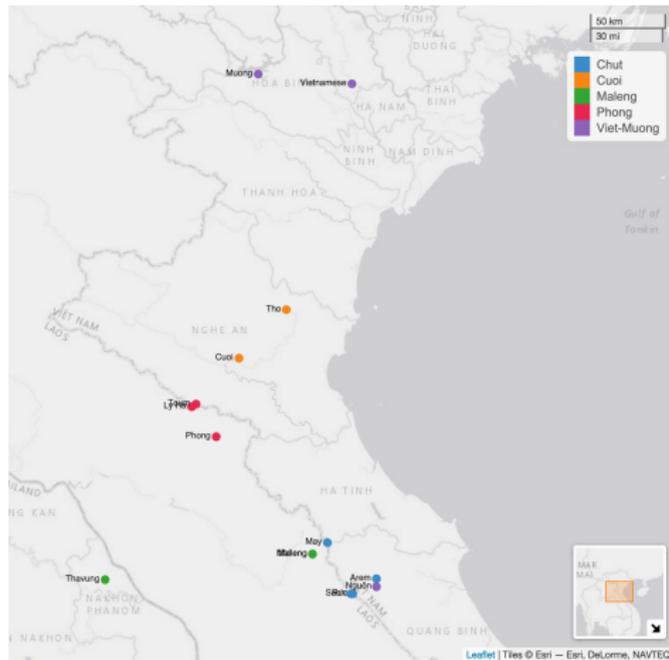


Figure 3: Vietic languages distribution (drawn via R's lingtypology)

VIETNAMESE AS AN AUSTROASIATIC LANGUAGE: BASIC VOCABULARY

Table 1: Basic vocabulary across Austroasiatic languages
(cf. SEALang Mon-Khmer Etymological Dictionary)

	'one'	'two'	'three'	'boat'	'husked rice'	'taro'
Proto-AA	*muəjʔ/ *mo:jʔ	*baar	*piʔ/ *ʔuj	NA	*rəŋ.ʰko:ʔ	*sə.ʰroʔ
Vietnamese [Vietic]	mot ^{d2}	haj ^{a1} / vaj ^{a2}	ba: ^{a1}	nok ^{d1} < nốc >	ɣa:w ^{c2}	sə: ^{c2}
Bahnar [Bahnaric]	mu:ŋ	ba:r	pe:ŋ	du:k	hɔw/həw	traw
Mon [Monic]	mɔwaj	ba:	pɔeʔ	kləŋ	həoʔ	krao
Khmer [Khmeric]	muəj	pɿ:	baj	tu:ʔ	ŋkɔ:	tra:v
Katu [Katuic]	mij	bəər	pɛ:	du:k	hŋɔʔ	ʔarəw [Bru]
Khmu [Khmuic]	mɔ:j	ba:r	peʔ	clɔ:ŋ	rəŋ.kɔʔ	sroʔ
Chong-H [Pearic]	mù:ʔj	p ^h ə:ʔj	p ^h e:ʔw	toʔ	kək ^h ɔ /kək ^h o:	tra:w

(also see Diffloth (2011) for more details)

VIETNAMESE AS AN AA LANGUAGE: TONOGENESIS (1)

- Vietnamese, however, was widely & **mistakenly** recognized as a Sino-Tibetan or Kra-Dai language since it has lexical tones, whereas most other Austroasiatic languages don't.
- To prove Vietnamese is an AA, we have to account for the origin of its tones (among other things). This was exactly what Haudricourt (1954) did in his breakthrough paper, i.e., to show that a **tonogenesis** did occur in the history of Vietnamese.



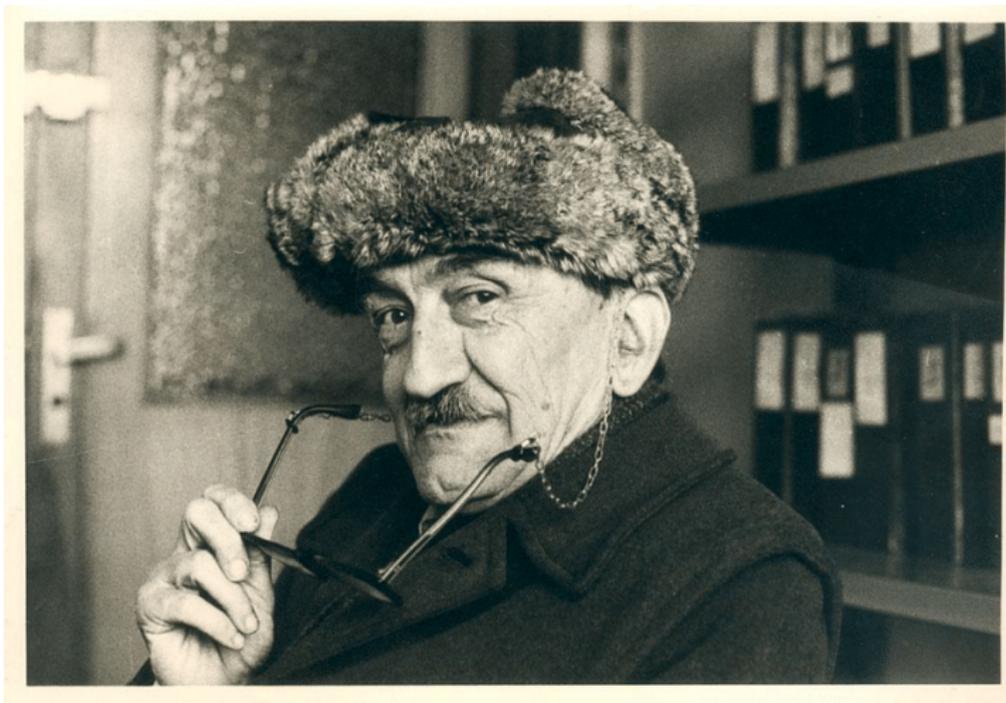


Figure 4: André-Georges Haudricourt (1911 – 1996)

VIETNAMESE AS AN AA LANGUAGE: TONOGENESIS (2)

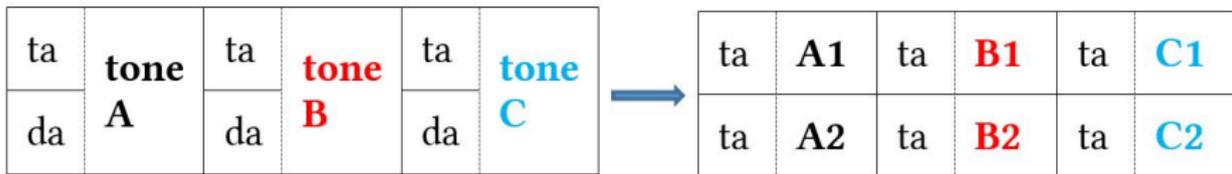
- Under Haudricourt's (1954) account, the tonal system developed in Vietnamese was a compensation for the loss of two codas in Proto-Vietic following an areal tendency (also happened in Sinitic, Kra-Dai, and Hmong-Mien languages by the same mechanism).

ta	ta [?]	ta ^h	→	ta	tone	ta	tone	ta	tone
da	da [?]	da ^h		da	A	da	B	da	C

Development from atonal period to 3 tones in unchecked syllables
(i.e. without stop codas /-p, -t, -k/)

👉 Question: Given that modern Vietnamese has 6 tones, what happened with three tones A, B and C in the next period?

VIETNAMESE AS AN AA LANGUAGE: TONOGENESIS (3)



Answer: Development from 3 to 6 tones in un-checked syllables

PERIODIZATION OF VIETNAMESE

1. **Pre-Vietnamese:** before the 8th century (Proto-VietMuong era)
2. **Proto-Vietnamese:** 8th –10th century
(the import of Literary Sino-Vietnamese readings started)
3. **Archaic Vietnamese:** 10th –13th century
(marked by the individualization of Literary Sino-Vietnamese)
4. **Ancient Vietnamese:** 13th –16th century
(exemplified by *Hua Yi Yi Yu* 華夷譯語 Chinese - Vietnamese Glossary in the XVI century – see next slide)¹
5. **Middle Vietnamese:** 16th –19th century
(exemplified by *A.de Rhodes's (1651) Viet-Bo-La Dictionary*)
6. **Modern Vietnamese:** the language of the Nguyễn (1802-1945) period up until now²

(cf. Henri Maspero 1912)

¹Also see *Nguyễn Trãi Quốc Âm Từ Điển* (15th century) for another look.

²cf. *Taberd's Vietnamese - Latin Dictionary* (published in 1838).

Linguistic profile of Vietnamese

VIETNAMESE: FACTS & FIGURES

- **Tiếng Việt** (*Vietnamese, Annamese, 越南語*): the national and official language of Vietnam & mother tongue of Viet (aka Kinh) people.
- **Speakers**: 96 million in Vietnam (2017) and approximately 4.5 million overseas Vietnamese (about 2 million in the US).
- **Genetic Relationship**:
Austro-Asiatic > Eastern-AA > Vietic > Viet-Muong.
- **Language Typology**: isolating, tonal (6 tones in the standard language), SVO, analytic, classifier language.

THE CURRENT WRITING SYSTEM OF VIETNAMESE (I.E. THE *Quốc Ngữ* ORTHOGRAPHY)



Figure 6: The alphabet of modern Vietnamese

PRACTICE: TONE DIACRITICS OF VIETNAMESE

System of Vietnamese tones: 2 (tonal registers) x 3 (tonal categories)

	A	B	C
High (1)	ngang (A1)	sắc (B1)	hỏi (C1)
Low (2)	huyền (A2)	nặng (B2)	ngã (C2)

 Practice: specify the tones of the following words.³

- **lạ, lã, mả, là**
- **mo, mỏ, mớ, mở**
- **ngu, ngủ, ngư, ngữ**
- **mo, mó, mò, mô**
- **che, ché, chề, chê**
- **ma, mà, má, măm, mâm**

³Attention: do not confuse tone diacritics with vowel-quality diacritics.

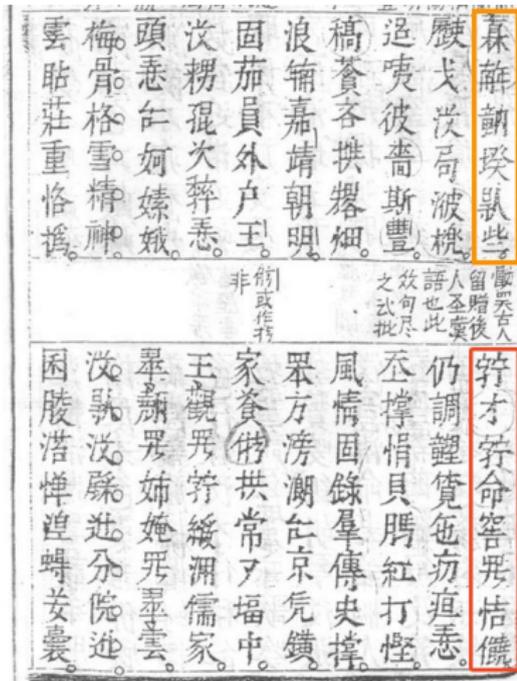
THE TALE OF KIEU (19TH CENTURY)

BEST KNOWN LITERARY CLASSIC IN VIETNAMESE

Truyện Kiều (The Tale of Kieu)

稟 辭 勳 燹 馱 些

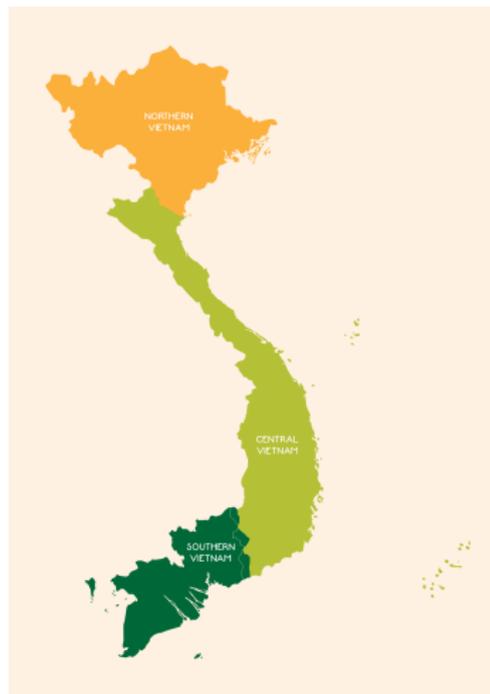
- (1) *Trăm năm trong cõi người-ta,*
hundred year in world other people
'In the hundred year span of human life'
- (2) *Chữ tài chữ mệnh khéo-là ghét*
word talent word fate such hate
nhau.
each-other
'Talent and fate always seem to conflict
with each other.'



Source: The Vietnamese Nôm Preservation

VIETNAMESE: THREE MAIN DIALECTS

- Vietnamese has **three** main varieties with many local subdialects derived from them:
 - ▶ **Northern:** e.g. Ha Noi variety
 - ▶ **Central:** e.g. Nghe An variety
 - ▶ **Southern:** e.g. Sai Gon variety
- Despite their dialectal differences (mainly lexical), Vietnamese generally manage to understand each other in daily conversations.



EXAMPLES: LEXICAL VARIATIONS

- **Northern:** một **bát** cơm
- **Central:** một **đội** cơm
- **Southern:** một **chén** cơm



'A/one **bowl** of rice'

👉 Question: What is the dialectal difference in this case?

EXAMPLES: PHONOLOGICAL VARIATIONS (1)

Table 2: dialectal variations of Vietnamese

Orthography	Northern	Central	Southern	Gloss
<i>lính</i>	[liɲB1]	[liɲB1]	[linB1]	'soldier'
<i>anh</i>	[ɛɲA1]	[ɛ:ɲA1]	[anA1]	'brother'
<i>vai</i>	[va:jA1]	[va:jA1]	[ja:jA1]	'shoulder'
<i>sạch</i>	[sɛcD2]	[sɛ:cD2]	[satD2]	'clean'
<i>tốt</i>	[totD1]	[totD1]	[tokD1]	'good'
<i>tay</i>	[tajA1]	[tajA1]	[ta:jA1]	'hand'
<i>mũi</i>	[muɲC2]	[muɲB2]	[muɲC1]	'nose'

EXAMPLES: DIALECTAL PRONUNCIATIONS (2)



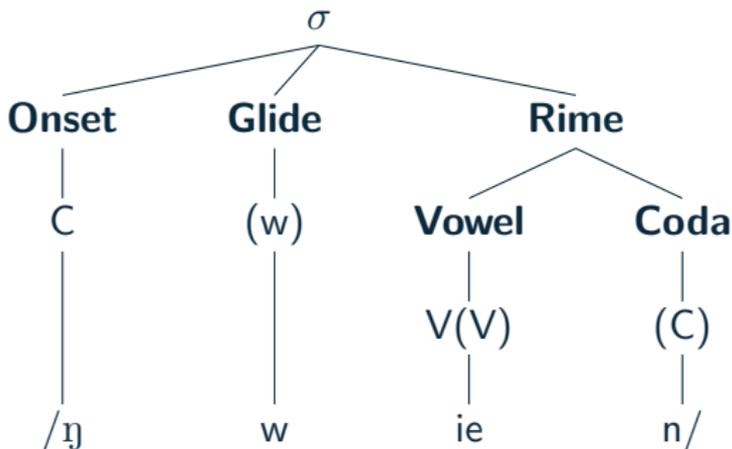
Cộng hòa Xã hội Chủ nghĩa Việt Nam
Độc lập - Tự do - Hạnh phúc

☞ The official title (*Socialist Republic of Vietnam*) and the motto (*Independence - Freedom - Happiness*) of Vietnam.

Sound System

SYLLABLE STRUCTURE OF VIETNAMESE

- Every syllable in Vietnamese bears a **Tone**.
- Apart from the **Tone**, the (segmental) syllabic structure of Vietnamese follows the template: **C(w)V(V)(C)** (elements in parentheses are not compulsory).



(cf. D.T. Thuat 2007)

ONSET INVENTORY: 22 CONSONANTS (INCLUDING 3 RETROFLEXES)

Table 3: Onset Inventory of (Northern) Vietnamese⁴

<i>Bilabial</i>	<i>Labiodental</i>	<i>Dental</i>	<i>Alveolar</i>	<i>Palatal</i>	<i>Velar</i>	<i>Glottal</i>
		t ^h <th>				
		t <t>		tʃ <ch>	k <k/c/q>	ʔ <∅>
β 			d̪ <d̪>			
m <m>			n <n>	ɲ <nh>	ŋ <ng/ngh>	
	f <ph>	s <x>			x <kh>	h <h>
	v <v>	z <gi/d/r>			ɣ <g/gh>	
		l <l>				

⁴Conventions: Yellow = plosives; Red = affricates; Green = implosives; Orange = nasals; Blue = fricatives; Olive = approximants. Note that the symbols in angled brackets <> are the orthography.

👉 Question: given the **onset** inventory in **Table 3**, what **phonetic feature** below should we pay special attention to (in order to speak Northern Vietnamese properly)?

- (A) [+/- **implosive**] (implosives vs. plosives)
- (B) [+/- **nasal**] (nasal vs. non-nasal)
- (C) [+/- **voice**] (voiced vs. voiceless)
- (D) [+/- **dorsal**] (velar vs. non-velar)

- Hint: let's count the pairs of phonemes involving the same contrast.

VOWEL INVENTORY: 11 MONOPHTHONGS, 3 DIPHTHONGS

Table 4: Vowel Inventory of (Northern) Vietnamese

ie < iê/ia >	ɨə < ươ/ưà >	uo < uô/ua >
i < i >	ɨ (ɯ) < ư >	u < u >
e < ê >	ə/ɛ: < â/ơ >	o < ô >
ɛ/(ɛ:) < e >		ɔ/(ɔ:) < o >
	a/a: < ă/a >	

CODA INVENTORY: 6 CONSONANTS, 2 SEMI-VOWELS

Table 5: Coda Inventory of (Northern) Vietnamese⁵

p < p >	t < t >	c/k < ch/c >
m < m >	n < n >	ɲ/ŋ < nh/ng >
w < u/o >	j < y >	

⁵Unlike English, the codas in Vietnamese are spoken in an *unreleased* manner (in IPA, e.g. *tóc* [tɔk̚˧˧˨ D1] 'hair'.)

PHONOLOGICAL FEATURES OF CENTRAL VIET-NAMESE

- Features of **Central dialects** (e.g. Nghe An dialect):
 - ▶ having three initial retroflexes <tr> /ṭ-/, <s> /ʂ-/, <r> /ʐ-/
 - ▶ having 5 tones only: Ngã (C2) is merged into Nặng (B2) or Hỏi (C1)

Tonal system of Central Vietnamese: 5 tones

	A	B	C
High (1)	ngang (A1)	sắc (B1)	hỏi (C1)
Low (2)	huyền (A2)	nặng (B2)	

PHONOLOGICAL FEATURES OF SOUTHERN VIET-NAMESE

- Features of **Southern dialects** (e.g. Sai Gon dialect):⁶
 - ▶ having three retroflexes but /v-/ and <gi/d> /z-/ are merged into the glide /j-/ while <qu> /kw-/ becomes the glide /w-/.
 - ▶ the three front vowels /i, e/ and /ɛ/ become centralized as [i, ə] and lowered as [a] before finals /-n, -t/, respectively.
 - ▶ dental finals /-t/ and /-n/ become velars as [-k], [-ŋ] after central and back vowels, respectively.
 - ▶ having 5 tones only: Ngã (C2) is merged into Hỏi (C1).

⁶Pham, A. H. (2008). The non-issue of dialect in teaching Vietnamese. *Journal of Southeast Asian Language Teaching*, 14, pp. 22-39.

SIDE-NOTES: ALLOPHONES (1)

- In Northern and Central Vietnamese, codas <ch> [-c] and <nh> [-ŋ] are considered as allophones of /-k/ and /-ŋ/, i.e. [-k^j, -ŋ^j], after front vowels /i, e, ε/.

- ▶ **Palatalization:**

/-k,-ŋ/ → [+front;-back] / [+syllabic;+front;-back]__#

- ▶ Examples: *bích* 'jade', *sách* 'book', *kếch* 'big', *kính* 'glass', *kênh* 'canal', etc.

SIDE-NOTES: ALLOPHONES (2)

- In Northern and Southern Vietnamese, codas /-k/ and /-ŋ/ become **labialized** as [-k^w, -ŋ^w] after back rounded vowels /u, o, ɔ/.
 - ▶ **Labialization:**
/-k,-ŋ/ → [+labial] / [+syllabic;+labial] ___#
 - ▶ Examples: *Úc* 'Australia', *ông* 'grandpa', *ong* 'bee', etc.

INVENTORY OF RIMES (VOWELS + CODAS)

Vowels	∅	/-j/ -i/y	/-w/ -o/u	/-m/ -m	/-p/ -p	/-n/ -n	/-t/ -t	[-ŋ] -nh	[-c] -ch	/-ŋ/ -ng	/-k/ -c
/i/	i	-	iu	im	ip	in	it	inh	ich	-	-
/iə/	ia	-	iêu	iêm	iêp	iên	iêt	-	-	iêng	iêc
/e/	ê	-	êu	êm	êp	ên	êt	ênh	êch	-	-
/ɛ/(ĕ)	-	-	-	-	-	-	-	anh	ach	-	-
/ɛ:/	e	-	eo	em	ep	en	et	-	-	eng	ec
/i/(ɨ)	ư	ưi	ưu	-	-	-	ưt	-	-	ưng	ưc
/iə/	ưa	ươi	ươu	ươm	ươp	ươn	ươt	-	-	ương	ươc
/ə/	-	ây	âu	âm	âp	ân	ât	-	-	âng	âc
/ə:/	ơ	ơi	-	ơm	ơp	ơn	ơt	-	-	-	-
/a/(ă)	-	ay	au	ăm	ăp	ăn	ăt	-	-	ăng	ăc
/a:/	a	ai	ao	am	ap	an	at	-	-	ang	ac
/u/	u	ui	-	um	up	un	ut	-	-	ung	uc
/uə/	ua	uôi	-	uôm	-	uôn	uôt	-	-	uông	uôc
/o/	ô	ôi	-	ôm	ôp	ôn	ôt	-	-	ông	ôc
/ɔ/	o	oi	-	om	op	on	ot	-	-	ong	oc
/o:/	-	-	-	-	-	-	-	-	-	ông	-
/ɔ:/	-	-	-	-	-	-	-	-	-	oong	ooc

PRACTICE: THE SPELLING RULE OF /k-/

- Problem: Specify the rule for writing the onset /k-/ in Quốc Ngữ script, viz. <k, c, q>, given the following dataset.
- Dataset:
kéo 'scissors', *kính* 'glass', *cao* 'high', *kể* 'tell', *cũ* 'old', *có* 'have', *quả* 'fruit', *kiến* 'ant', *cua* 'crab', *quý* 'valuable', *quê* 'countryside'.
- Hint: specify the natural classes/phonological categories of the segments following each orthographic realization of /k-/. Notice the glide /-w-/ is written as <-u-> or <-o-> in Vietnamese.

TONES OF VIETNAMESE (1)

MA (A1) <i>ghost</i>	MÁ (B1) <i>mother</i>	MẢ (C1) <i>tomb</i>
MÃ (A2) <i>but</i>	MẠ (B2) <i>rice seedlings</i>	MÃ (C2) <i>horse</i>

- Tones in Vietnamese are **lexical tones**, i.e. they help us distinguish different lexical items (i.e. words).
- Northern Vietnamese tones make use of both **pitch** and **phonation contrast** (clear voice vs. creaky voice).
- Some authors suggest that an 8-tone, not 6-tone, system might be **phonetically finer-grained** for the distinction between *checked syllables* (having stop codas) vs. *un-checked syllables* (not having stop codas) in Vietnamese (e.g. Michaud 2004).
- **Checked syllables** only can bear **Sắc** (B1/D1) or **Nặng** (B2/D2) tones.

TONES OF VIETNAMESE (2)

Table 6: The 8-tone system of Northern Vietnamese⁷

<i>Category</i>	A	B	C	D
<i>Register</i>				
High (1)	a	á	ǎ	áp
Low (2)	à	ạ	ã	ạp

⁷The tone C2 - **Ngã** (in shade) is phonemic only in Northern Vietnamese.

TONES OF VIETNAMESE (3)

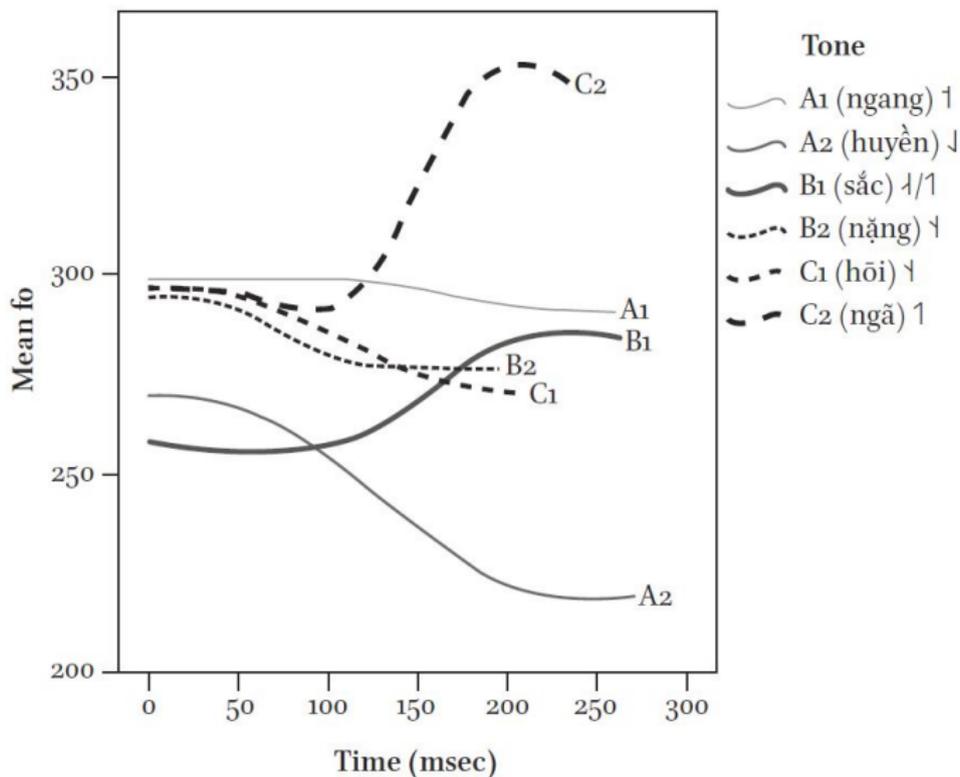


Figure 7: Northern Vietnamese tone contours in unchecked syllables

PRACTICE: TONES OF VIETNAMESE

- (1) Ông Lữ đi một đôi giày cũ.
grandpa Lữ wear one pair shoes old
'Mr. Lữ wears a pair of old shoes.'
呂先生穿了一雙舊的鞋子.
- (2) *Thỉnh thoảng tâm tư tôi thẫn thờ.*
Sometimes mind 1.SG distracted
'Sometimes my mind got distracted.'
有時候我心中很悵惘.

PRACTICE: TONGUE TWISTER

Try to speak clearly and properly the following sentences :

- (3) *Nồi đồng nấu ốc nồi đất nấu ếch.*
pot bronze cook snail pot earthen cook frog
'The bronze pot is used for cooking snails, the earthen pot is used for cooking frogs.'
- (4) *Anh Thanh ăn thạch anh Thành ăn hành.*
brother Thanh eat jelly brother Thành eat onion
'Brother Thanh eats jelly and brother Thành eats onion.'

SIDE-NOTES ON VIETNAMESE PHONOLOGY (1)

- Vietnamese also can have a medial glide /-w-/. However, it is phonotactically banned after labial onsets, e.g.: *toán* [twan B1] ‘math’, *qua* [kwa A1] ‘pass’, etc. but not *[vwa].
- Spelling rules for the medial glide /-w-/:
 - ▶ /-w-/ → <-u-> / k ___ or ___ [+syllabic; -low; +tense].
 - ▶ /-w-/ → <-o-> / elsewhere.
- Examples: *hoa* ‘flower’, *tuần* ‘week’, etc.

SIDE-NOTES ON VIETNAMESE PHONOLOGY (2)

- Unlike Chinese, Vietnamese doesn't have the *tone sandhi*, but have segmental *assimilation*, *dissimilation* as well as *cliticization/reduction*.
 - ▶ *Assimilation*: **sắc** 'sharp' + **xảo** '巧 ingenious' = **sắc xảo** 'sharp, ingenious' (not **sắc xảo*)
 - ▶ *Dissimilation*: **mười** (10) + **năm** (5) = **mười lăm** 'fifteen' (not **mười năm*)
 - ▶ *Cliticization*: **phải** 'right/yes' + **không** 'no/not' = **phỏng** 'right?' (A-not-A question)⁸

⁸Pham, A. H. (2007). Cliticization in casual speech in Vietnamese. *Cahiers de linguistique Asie orientale*, 36(2), 219–244.

Writing Systems

NÔM CHARACTERS (10TH - 19TH CENTURY)

- **Nôm characters** (*chữ Nôm* 字喃) had been stabilized around the 11th century and was already widely used to transcribe vernacular Vietnamese since the Tran dynasty (1225 - 1400 CE) (cf. Nguyen Dinh Hoa et al. 2018).
- Make use of 漢字 (Chinese characters) and devised in accordance with principles of Chinese characters (i.e. 六書 'Six Scripts').

A large, bold Nôm character, which is a stylized Chinese character '天' (tiān), representing the word 'trời' (sky).

trời 'sky'

A large, bold Nôm character, which is a stylized Chinese character '龍' (lóng), representing the word 'sông' (river).

sông 'river'

Table 9: Examples of compound ideograms

Nôm graph	Vietnamese morpheme	Source of components
在 (於)	ở ('in')	在 (in) + 於 (in)
請 (請)	xin ('to beg')	口 (mouth) + 請 (beg)
世 (世)	đời ('era', 'age')	世 (era) + 代 (age)
事 (事)	việc ('work')	事 (work) + 役 (work)
年 (年)	tuổi ('year' old)	年 (year) + 歲 (old)

Data source: Lục Vân Tiên (1870s); Nguyễn TC (2006).

NÔM CHARACTERS: SEMANTIC-PHONETIC COMPOUNDS

Table 10: Examples of semantic-phonetic compounds

Nôm graph	Vietnamese morpheme	Source of components
罛	Bốn ('four')	四 (four) + 本 <i>bản</i>
糲	Bún ('noodle')	米 (rice) + 罛 <i>bốn</i>
𠂔	Trời ('sky')	天 (sky) + 上 (above)
𠂔	Lời ('speech')	口 (mouth) + 𠂔 <i>trời</i>
𠂔	Mười ('ten')	什 (ten) + 邁 <i>mại</i>
𠂔	Mời ('to invite')	口 (mouth) + 𠂔 <i>mười</i>
𠂔	Trong ('inside')	中 (inside) + 竜 <i>long</i>
𠂔	Trong ('pure')	氵 (water) + 𠂔 <i>trong</i>

Source: Trần, TD (2019). The history of Nôm: A periodization. *Journal of Chinese Writing Systems*, 3(3), 175-188.

PRACTICE: NÔM CHARACTERS

👉 Can you guess the (rough) meaning and pronunciation for the following *semantic-phonetic* Nôm characters?

𠂇 𠂈

Hint: the semantic component of the left character is on the right, whereas that of the right one is on the left.

PRACTICE: NÔM CHARACTERS

☞ Can you guess the (rough) meaning and pronunciation for the following *semantic-phonetic* Nôm characters?

[ʃa:] 'three'

[an] 'eat'

Answer: Nôm characters for “three” [ʃa:], “eat” [an]

THE VALUE OF NÔM CHARACTERS

*Khen ai khéo dựng nên dừa,
'Those who put up the coconut tree
such ingenious.'
Đẩy trèo đây hứng cho vừa một đôi.
'You climb, I catch, just like a couple.'*



Figure 8: *Hứng dừa* 'coconuts catching' (Đông-Hồ painting).

CHỮ QUỐC NGỮ (17TH - PRESENT)

- invented by the Portuguese and French missionaries in the 17th century (Francisco de Pina, Alexandre de Rhodes, etc.)
- granted as the official writing system of Vietnamese since 1910s (under the French colonial era).

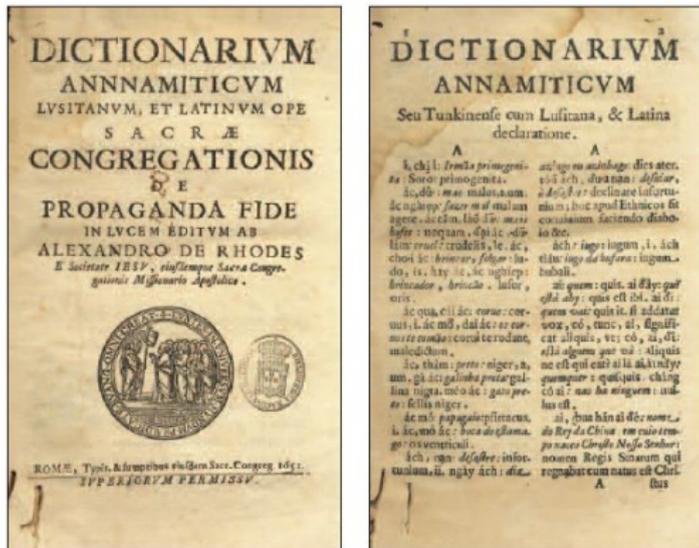


Figure 9: A.de Rhodes's (1651) Viet-Bo-Li Dictionary

Lexicon

LAYERS OF LEXICON

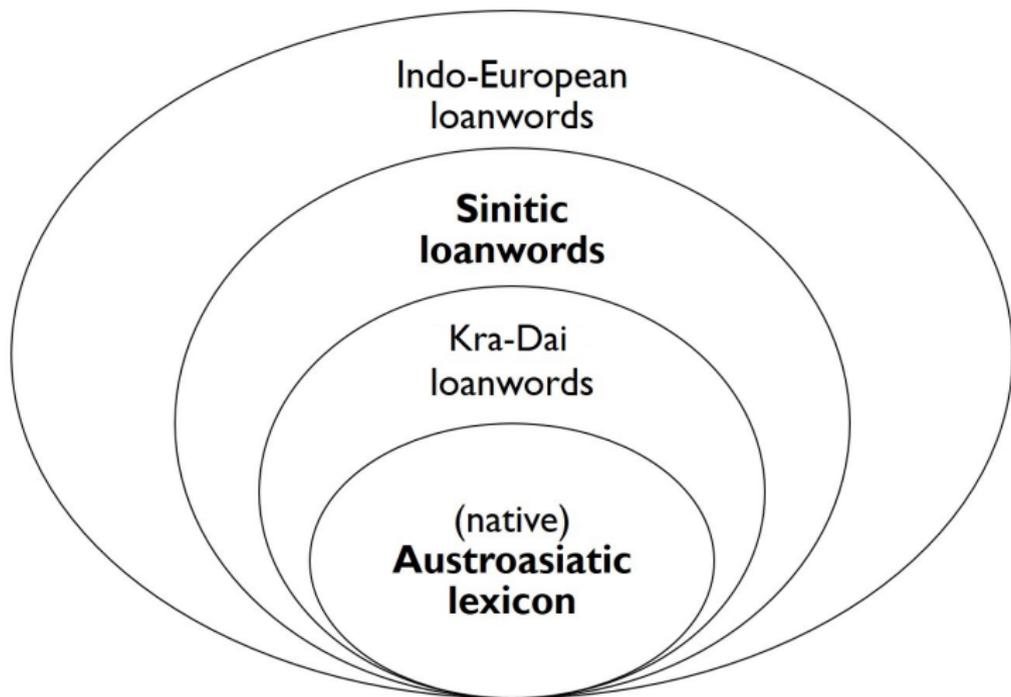


Figure 10: Four essential lexical layers in Vietnamese (from most ancient to most recent outwardly)

AUSTROASIATIC LEXICON: THE MOST ANCIENT LAYER IN VIETNAMESE

Table 11: Basic vocabulary across Austroasiatic languages

Vietnamese	Muong	Khmer	Mon	Gloss
[kən]	[kən]	[koun]	[kon]	'child'
[cim]	[cim]	-	[hacem]	'bird'
[ba]	[pa]	[bəj]	[pae ^ʔ]	'three'
[tǎj]	[t ^h ǎj]	[daj]	[toa]	'hand'
[sɔŋ]	[sɔŋ]	-	[krəŋ]	'river'
[tu]	[tlu]	[mluu]	[həplu ^h]	'betel'

- Most **basic vocabulary** in Vietnamese still retains their *native* Austroasiatic origins. (cf. F.E. Huffman 1977)

- Vietnamese lexicon has a small handful of Tai loanwords presumably borrowed even before Han Dynasty (200 BCE - 200 CE) when Proto-Tai people lent the ancestors of Vietnamese some technology related to *agriculture* and *animal husbandry* with the relevant vocabulary (cf. Alves 2009).

Table 12: Potential cognates between Vietnamese and Tai Yo

Vietnamese	Tai Yo ⁹	Gloss
(chó) má	mả	'dog'
rẫy	hày	'mountain field'
chuối	kuôi	'banana'
mương	mường	'trench'
muống	bung	'water spinach'

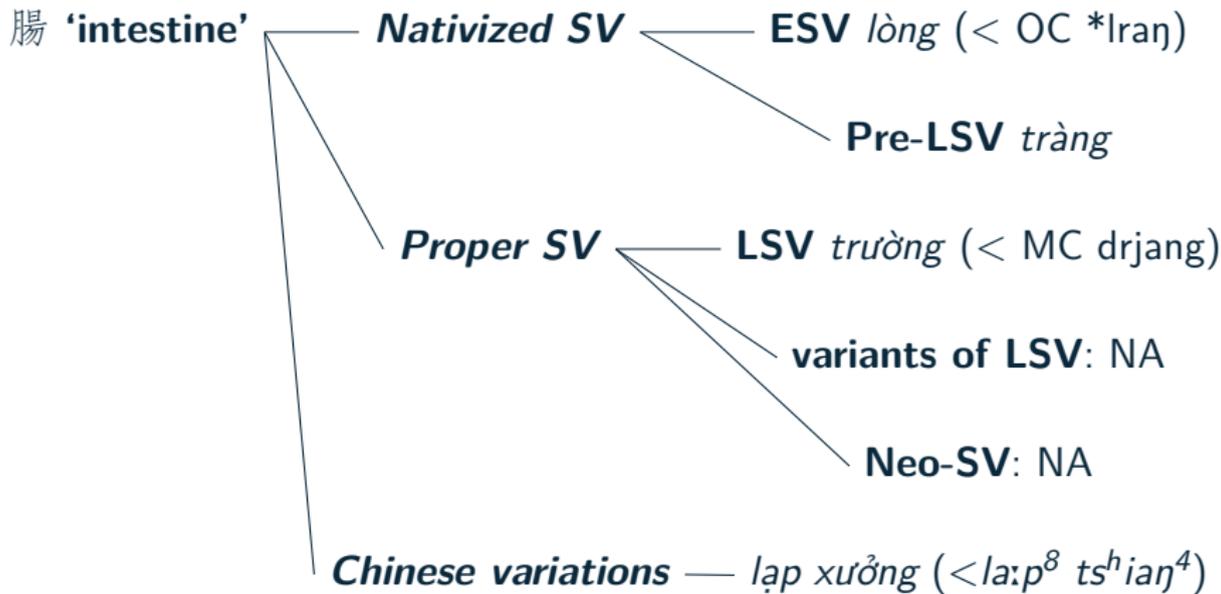
⁹A “big” Tai language spoken in Central Vietnam.

SINITIC LOANWORDS: THE RICHEST LAYER IN VIETNAMESE

- Approximately 70% of modern Vietnamese lexicon are borrowed from Chinese (cf. Alves 2009).
- There are 3 macro-layers of Sinitic loans in Vietnamese:
 1. **Early Sino-Vietnamese**
 - *Early Sino-Vietnamese I* (ESV-I)
 - *Early Sino-Vietnamese* (ESV-II)
 2. **Sino-Vietnamese Proper**
 - *Literary Sino-Vietnamese* (i.e. Sino-Vietnamese Proper, LSV)
 - *Variants of LSV* (i.e. Post-Tang SV)
 - *Neo-Sino-Vietnamese graphic loans* (i.e. Sino-neologisms in 20th century)
 3. **Chinese variations loans:** loans from Cantonese, Min languages, etc. in recent time.

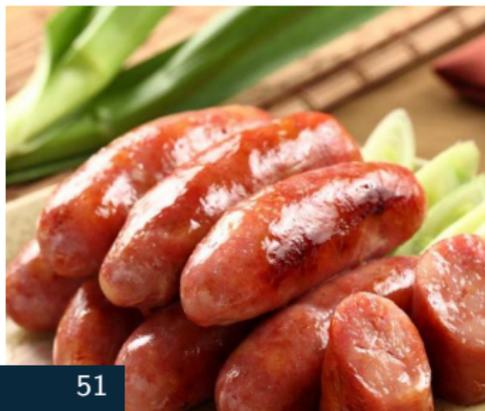
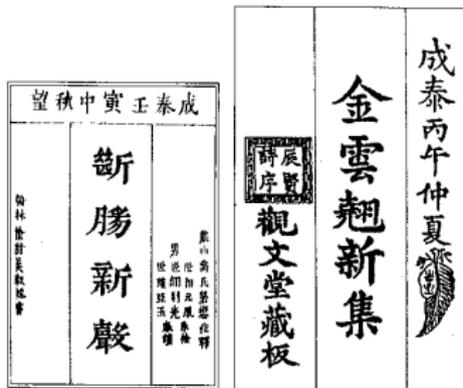
SINITIC LOANWORDS (2): EARLY SINO-VIETNAMESE

- Among micro-layers of Sinitic loans, **Early Sino-Vietnamese** are the earliest layer with some even borrowed from Old Chinese.



SINITIC LOANWORDS (3): THE “QUADRUPLETS”

腸 ‘intestines’: lòng - trường - tràng - xướng



Tĩnh dạ tứ

Sàng tiền minh nguyệt quang

Nghi thị địa thượng sương

Cử đầu vọng minh nguyệt

Đê đầu tư cố hương

Lý Bạch

靜夜思

床前明月光

疑是地上霜

舉頭望明月

低頭思故鄉

李白

¹⁰For comparison, see how to read this poem in Taiwanese here.

SINITIC LOANWORDS (5): LITERARY SINO-VIETNAMESE

- **Literary Sino-Vietnamese (LSV) lexicon** (漢越詞) is a part of the vocabulary of **Late Middle Chinese** (7th – mid-13th century) accompanied with Chinese characters taught in school and used in administration in Vietnam (交州) that was incorporated into the Vietnamese lexicon since the Tang–Song era.
- LSV items should not be confused with Sino-neologist **graphic loans**, i.e. loans without pronunciations, from Chinese (most coined in Japanese) in the early 20th century. E.g.: *kinh-tế* 經濟 ‘economy’, *văn-hóa* 文化 ‘culture’, *chính-trị* 政治 ‘politics’, etc.¹¹

¹¹cf. Tranter, N. (2009). Graphic loans: East Asia and beyond, *Word*, 60(1),1-37.

SINITIC LOANWORDS (6): LITERARY SINO-VIETNAMESE FROM “SINOSPHERE” PERSPECTIVE

Table 13: 德 dé, 多 duō, and 得 dé in Middle Chinese and selected varieties of Chinese¹²

	德	多	得
Middle Chinese	<i>tok</i>	<i>ta</i>	<i>tok</i>
Mandarin	[tʰ35]	[tuo55]	[tʰ35]
Cantonese	[tɛk5]	[to55]	[tɛk5]
Sūzhōu	[tʰʔ4]	[to44]	[tʰʔ4]
Sino-Korean	<i>tōk</i> [tʌk]	<i>ta</i> [ta]	<i>tūk</i> [tuɤk]
Sino-Japanese	<i>toku</i>	<i>ta</i>	<i>toku</i>
Sino-Vietnamese	<i>đưc</i> [dʰuɤk D1]	<i>đa</i> [dʰa A1]	<i>đác</i> [dʰaɤk D1]

(Cf. Baxter & Sagart 2014:11)

¹²Some authors suggest a close relationship between LSV and some southern dialects of Chinese, e.g. *Seiyap Cantonese* (cf. Ding 2015)

SINITIC LOANWORDS (7): SEMANTIC CHANGES IN LITERARY SINO-VIETNAMESE

- Note that a few Literary Sino-Vietnamese items have undergone noticeable semantic changes after being borrowed into Vietnamese for a long time.

Table 14: Examples for semantic changes in LSV

Character	Original meaning	LSV	Vietnamese usage
困難	'difficult'	khốn-nạn	'mean'
所以	'therefore'	sở-dĩ	'if, the reason that'
曲折	'winding'	khúc-chiết	'clear'
意思	'meaning'	ý-tứ	'considerate'

INDO-EUROPEAN LOANWORDS

- Most Indo-European loanwords in Vietnamese come from French¹³ (during the French colonial era 1884 – 1945) and English (due to recent globalization).

Table 15: French loans

French	Viet	Gloss
beurre	bơ	'butter'
cravate	cà-vạt	'tie'
gare	ga	'train station'
pile	pin	'battery'
chaland	sà-lan	'barge'
pâté	ba-tê	'liver pate'

Table 16: English loans

English	Viet
meeting	mít-tinh
television	ti-vi
radio	ra-đi-ô
cowboy	cao-bồi
Internet	internet
(e-)mail	meo/mêu

¹³cf. Dictionnaire vietnamien - français: Les mots vietnamiens d'origine française.

AUSTRONESIAN LOANWORDS

- Beyond four essential layers above, Vietnamese also imports a bunch of Austronesian etyma. Most of them belong to basic vocabulary and are assumed to be borrowed very early.

Table 17: Potential lexical correspondences among Bahasa Indonesia, PR Cham and Vietnamese

Bahasa Indonesia	PR Cham ¹⁴	Vietnamese	Gloss
hari	huray	ngày	'day'
bulan	bulan	trăng	'month/moon'
mata	mita	mắt	'eye'
pulau	palau	cù-lao	'island'

☞ Some of these cognates also imply the contact between Austroasiatic and Austronesian languages.

¹⁴Phan Rang Cham is a Chamic language spoken in South Central Coast of Vietnam.

FUNCTIONAL LAYERS OF LEXICON

Besides the major lexical categories (i.e. N, V, Adj), we also should pay attention to the following functional layers in Vietnamese:

- *Personal Pronouns* : **tôi, chúng-tôi, chúng-ta**, etc.
- *Numerals* : **một, hai, ba**, etc.
- *TAM (Tense/Aspect/Modality) particles*: **đã, đang, sẽ, rồi, nhỉ, mà**, etc.
- *Connectives*: **và, với, hoặc, hay**, etc.

👉 In the next slides, we will talk about the first two basic groups: i.e. Personal Pronouns & Numerals.

*'On the sociolinguistic level, great languages (in MSEA) tend to reflect hierarchically stratified cultures, often with clearly distinguished formal vs. colloquial styles of speaking, including such features as **elaborate pronominal systems** that encode differences in age, gender, and social status...'*
(cf. Matisoff 2019)

THE TWO PRONOMINAL SYSTEMS IN VIETNAMESE

Vietnamese employs two sub-systems of personal pronouns:

- Proper personal pronouns: **tôi** 'I 我', **chúng-tôi** 'we-exclusive 我們', **chúng-ta** 'we-inclusive 咱們' etc.
- Derivational personal pronouns: **anh** 'elder brother 哥', **chị** 'elder sister 姐', **thầy** 'teacher 師', **sếp** 'boss', etc.

Table 18: Proper Personal Pronouns in Vietnamese

	Singular	Plural
First person	<i>tôi</i> (could be formal)	<i>chúng tôi</i> (excluding the addressed subject)
	<i>ta</i> (neutral, non-formal)	<i>chúng ta</i> (including the addressed subject)
	<i>tao</i> (superior to subordinate, familiar)	<i>chúng tao</i> (vulgar, excluding the addressed subject)
	<i>minh</i> (intimate)	<i>minh</i> or <i>chúng minh</i> (intimate, including the addressed subject)
Second Person	<i>mày</i> or <i>mi</i> (superior to inferior, familiar)	<i>bay</i> , <i>chúng mày</i> , <i>tụi mày</i> , <i>chúng bay</i> (superior to subordinate, familiar)
Third Person	<i>nó</i> (superior to subordinate, familiar)	<i>chúng nó</i>

NUMERALS

Number	IPA	Vietnamese	Gloss
1	[mot D2]	một	'one'
2	[ha:j A1]	hai	'two'
3	[ʃa: A1]	ba	'three'
4	[ʃon B1]	bốn	'four'
5	[nam A1]	năm	'five'
6	[saw B1]	sáu	'six'
7	[ʃaj C1]	bảy	'seven'
8	[ta:m B1]	tám	'eight'
9	[tɕin B1]	chín	'nine'
10	[miəj A2]	mười	'ten'
100	[mot D2 tɕam A1]	một trăm	'one hundred'
1,000	[mot D2 ɲin A2]	một nghìn	'one thousand'
1,000,000	[mot D2 tɕiew B2]	một triệu	'one million'

PRACTICE WITH THE BANKNOTES (VND)



Morphology & Grammar

- Vietnamese is a typical *isolating language* in the sense that most of its **words** roughly correspond to **morphemes** (i.e. smallest meaningful units) and its morphology (i.e. study of morphemes) only has the **derivational** component, and lacks the **inflectional** one.
- Regarding the **word formation**, there are three types of words in Vietnamese:¹⁵
 - ▶ **simple (monosyllabic) words**: *nước* 'water', *đất* 'land', etc.
 - ▶ **compounds**: *đất-nước* 'country', *nước-mắt* 'tear', etc.
 - ▶ **reduplicatives**: *đỏ-đỏ* 'reddish', *vui-vẻ* 'cheerful', etc.

¹⁵Henceforth, compounds and reduplicatives will be indicated by the hyphen(-).

- All simple words are monosyllables (i.e. having one syllable only) but not all monosyllables are simple words, i.e. free morphemes. Consider the following contrast:

(5) *Tôi uống **nước**/***thủy**.*

1.SG drink water

'I drink water'.

(*nước*: native; *thủy* = 水)

- A sentence that is wholly composed of simple words:

(6) *Tôi rất thích ăn cơm.*

1.SG very like eat rice

'I like to eat rice very much.'

COMPOUNDS (1)

1. **Co-ordinate compounds** are words whose each component has equal contribution to the meaning of the whole: *mua-bán* 'buy and sell', *núi-sông* 'mountain and river', *quần-áo* 'pants and shirts', *phải-trái* 'right and wrong', etc.¹⁶
2. **Sub-ordinate compounds** are words whose one constituent modifying (modifier) the meaning of the other (modified): *xe-đạp* 'bicycle', *vui-tính* 'humorous', *nhà-văn* 'writer', etc.

¹⁶Unlike Chinese, the word-order of co-ordinate compounds in Vietnamese is not strictly fixed, e.g. *núi-sông* = *sông-núi* 'mountain and river', BUT *bố-mẹ* 'parents' ≠ **mẹ-bố* (unavailable compound).

3. “**Random**” **compounds** are formed by combining two or more meaningless, bound components without any phonetic correlations, e.g.: *bù-nhìn* ‘scarecrow’, *ếch-ương* ‘frog’, *mồ-hôi* ‘sweat’, *mặc-cả* ‘bargain’, *cà-phê* ‘coffee’, etc.¹⁷

¹⁷Some authors treat these words as **simple** **polysyllabic** words.

REDUPLICATIVES (1)

- Vietnamese is among the languages which make use of reduplicatives most frequently.
- Types of Vietnamese reduplicatives:
 - ▶ **Total reduplicatives** : **đỏ** 'red' → **đỏ-đỏ** 'reddish'
 - ▶ **Partial reduplicatives** :
 - Onset reduplicatives: **lạnh** 'cold' → **lạnh-lùng** 'cold, glacial'
 - Rime reduplicatives: **khóc** 'cry' → **khóc-lóc** 'cry, lament'¹⁸

¹⁸Both 'lùng' and 'lóc' do not have specific lexical meaning.

REDUPLICATIVES (2)

- Interestingly, reduplicatives in Vietnamese strictly follow tonal harmony. Thus, the tones of both syllables in a reduplicative must belong to the same register, i.e. the First-High (*ngang, sắc, hỏi*) or the Second-Low (*huyền, nặng, ngã*) category. For instance: **nã-o-nề** [ngã + huyền] ‘depressing’, **lạnh-lùng** [nặng + huyền] ‘glacial’, etc.

Table 7: Tonal Harmony in Reduplicatives

	EVEN	SHORT	LONG
REGISTER			
FIRST	ngang	sắc	hỏi
SECOND	huyền	nặng	ngã

(cf. Thompson 1965:141)

👉 Question: given the tonal harmony in Table 7, which one among the following forms is a proper reduplicative in Vietnamese?

- (A) *lá-lạ*
- (B) *lả-lạ*
- (C) *la-lạ*
- (D) *là-lạ*

SYNTAX OF VIETNAMESE: AN OVERVIEW

- Vietnamese has the same basic SVO word order like English.
- Vietnamese is a typical **analytic language** in the sense that its grammatical relationships are mainly indicated by **word order** and **functional particles/words** (not by inflectional means as in synthetic languages).
- For instance, Vietnamese is found to employ many pre-verbal and post-verbal particles to convey Tense-, Aspect-, and Modality-related uses.
- Tense-Aspect particles in Vietnamese:
 - ▶ ***đã*** (pre-verbal): past tense
 - ▶ ***đang*** (pre-verbal): progressive aspect
 - ▶ ***sẽ*** (pre-verbal): future tense
 - ▶ ***rồi*** (clause-final): perfective aspect

SYNTAX (1): EXAMPLES

- (7) a. *Tôi đã ở Đài Loan hai năm.*
1SG PST live Taiwan two year
'I **lived** in Taiwan for two years.'
- b. *Tôi ở Đài Loan hai năm rồi.*
1SG live Taiwan two year already
'I **have lived** in Taiwan for two years.'
- (8) **Notice:** Different word orders involving temporal adverbs can give rise to different TAM-related interpretations.
- a. *Cô ấy mới về tối qua.*
3SG just come.back last night
'She **just came back** last night.'
- b. *Tối qua cô ấy mới về.*
last night 3SG just come.back
'She **did not come back until** last night.'

SYNTAX (2): VIETNAMESE AS A PRO-DROP LANGUAGE

- Like Chinese and Thai, Vietnamese is a **pro-drop** language.

(9) A: *Bác đi đâu đấy ạ?* 'Where are you going?' (greeting)
B: *Đi ăn cơm.* 'I go to eat (lunch/dinner).'

- Vietnamese is an **SVO** and **topic-oriented** language (like Chinese). Consider the contrast below:

(10) Subject vs. Topic

a. Tên **tôi** là Nam

name 1SG COP Nam

'My name is Nam.' (Sub = *tên tôi*)

b. **Tôi** tên là Nam

1SG.TOP name COP Nam

'I have name which is Nam.' (Top = *tôi*; Sub = *tên*)

SYNTAX (3): HEAD-INITIAL OR HEAD-FINAL?

- The basic word order regarding Modifiers and Head in Vietnamese is **head-initial**: **Head + Modifier**.
- However, due to lexical borrowing, Vietnamese also makes use of the Chinese **head-final** pattern, i.e. **Modifier + Head**.
- Compare: Chinese pattern vs. native pattern.
 - **mãnh hổ** vs. **hổ dữ** 'frightening tiger 猛虎'
 - **đại ca** vs. **anh cả** 'big brother 大哥'

THE NOMINAL STRUCTURE OF VIETNAMESE

Numeral	Classifier	Noun	Modifier	Poss.P	Demonstrative
<i>ba</i>	<i>con</i>	<i>mèo</i>	<i>đen</i>	<i>của John</i>	<i>ấy</i>
three	CLF	cat	black	John's	that

'those three black cat-s of John'

- Like Chinese, Vietnamese nouns (not pronouns) also lack plural morphology, i.e. the bare nouns like *mèo* 'cat' denote both singularities and pluralities.
- Vietnamese employs the so-called classifiers like *con* (and *measure words*) to make (bare) nouns countable.
- Unlike IE languages, Vietnamese nominals do not have a counterpart of English *the*.

Cultural Notes

1. **Xin chào! Tôi tên là _____** ('Hello! My name is ...')
2. **Tôi là người _____** ('I come from ...')
(Việt Nam / Đài Loan / Hồng Kông/ In-đô/ Thái Lan/ Ma-lai)
3. **Xin lỗi, bạn tên là gì?** ('Excuse me, what is your name?')
4. **Cám ơn!** ('Thank you!')
5. **Anh ăn cơm chưa ạ?** ('Have you had lunch/dinner?')

 What is the function of the sentence-final particle “**ạ**”?

TRADITIONAL FOODS IN VIETNAM(1)



phở 'pho'



bún đậu 'tofu vermicelli'

TRADITIONAL FOODS IN VIETNAM(2)



nem cuốn 'summer roll'



nem rán/chả giò 'spring roll'

TRADITIONAL FOODS IN VIETNAM(3)



bánh da lợn 'pig-skin cake'



bánh cốm 'green rice flake cake'

DOS AND DONT'S IN VIETNAM (1)



Be prepared for the shocking weather!
(especially in the spring/summer
transition in the North/Central)



Short clothes at sacred places mean
'short' civilization

DOS AND DONT'S IN VIETNAM (2)

☞ Most importantly, don't you forget to **bargain** whenever you buy something from the traditional markets, vendors and transport services!



Xe ôm 'motorbike taxi'

ACKNOWLEDGEMENT

Deep gratitude toward Prof.Dr. Hsiu-chuan Liao, Dr. Andre Goderich and Noppakao Sirintranon for their illuminating comments and suggestions.

THANK YOU!

Cám ơn các bạn đã chú ý lắng nghe!

SELECTED REFERENCES (1)

- Alves, M.J. (2009). Loanwords in Vietnamese. *Loanwords in the World's Languages: A Comparative Handbook*, 617-637.
- Brunelle, M. (2014). Vietnamese (Tiếng Việt). Chapter 13 in *The Handbook of Austroasiatic Languages* (2 vols). Brill.
- Diffloth, G. (2011). Austroasiatic word histories: boat, husked rice and taro, In N. J. Enfield (ed.), *Dynamics of Human Diversity*, Pacific Linguistics, 295-313.
- Ding, P. (2015). Chinese influence on Vietnamese: a Sinospheric tale. In *Language Empires in Comparative Perspective*, 55-75.
- Doan, Thien-Thuat (2007). *A Course in Vietnamese Phonetics*. Hà Nội: Education Publishing House.
- Haudricourt, A.G. (1954). De l'origine des tons en Vietnamien. *Imprimerie Nationale*.
- Michaud, A. (2004). Final consonants and glottalization: new perspectives from Hanoi Vietnamese. *Phonetica*, 61(2-3), 119-146.

SELECTED REFERENCES (2)

- Nguyen, Dinh-Hoa, Alves, M.J., Nguyen, Hong-Con (2018). Vietnamese. In *The World's Major Languages*, 696-712.
- Nguyen, Tai-Can (1995). *A Course in Vietnamese Historical Phonology*. Hà Nội: Education Publishing House.
- Sidwell, P., & Alves, M. (2021). The Vietic Languages: A Phylogenetic Analysis. *Journal of Language Relationship*, 166-194.
- Sidwell, P. & Jenny, M. (Eds.) (2021). *The Languages and Linguistics of Mainland Southeast Asia: A Comprehensive Guide* (Vol. 8). Walter de Gruyter.
- Thompson, L.C. (1965). *A Vietnamese Grammar*. University of Washington Press.
- Vittrant, A. & Watkins, J. (Eds.) (2019). *The Mainland Southeast Asia Linguistic Area* (Vol. 314). Walter de Gruyter GmbH & Co KG.