

# Loanwords in Vietnamese

Mark J. Alves

## 1. Vietnamese language situation and background

The current population of Vietnam is over 80 million, of which approximately 90% consists of ethnic Vietnamese. There are also some three million overseas Vietnamese living primarily in North America and Europe. As a national language, Vietnamese is the sole language used in the news and media outlets throughout Vietnam, and it is also the primary language of education.

Vietnamese is considered to have three general regional varieties – Northern, Central, and Southern Vietnamese – which are represented by the speech of the cities of Hanoi (northern), Hue (central), and Ho Chi Minh City (southern) respectively.<sup>1</sup> While Hanoi speech is considered representative of the official standard pronunciation accent in Vietnam, the variety spoken in Ho Chi Minh City is also a highly influential variety in Vietnam and one likely to last in the face of dialect leveling due to the economic prosperity of the south. As the Vietnamese romanized alphabet, called *Quốc c Ngữ* or “National Language” (from Chinese 國語 (*guó yǔ*)), does not fully represent the phonology of any single variety of Vietnamese, but rather an idealized system, the data in the online database connected to this article do not refer to regional phonological differences.

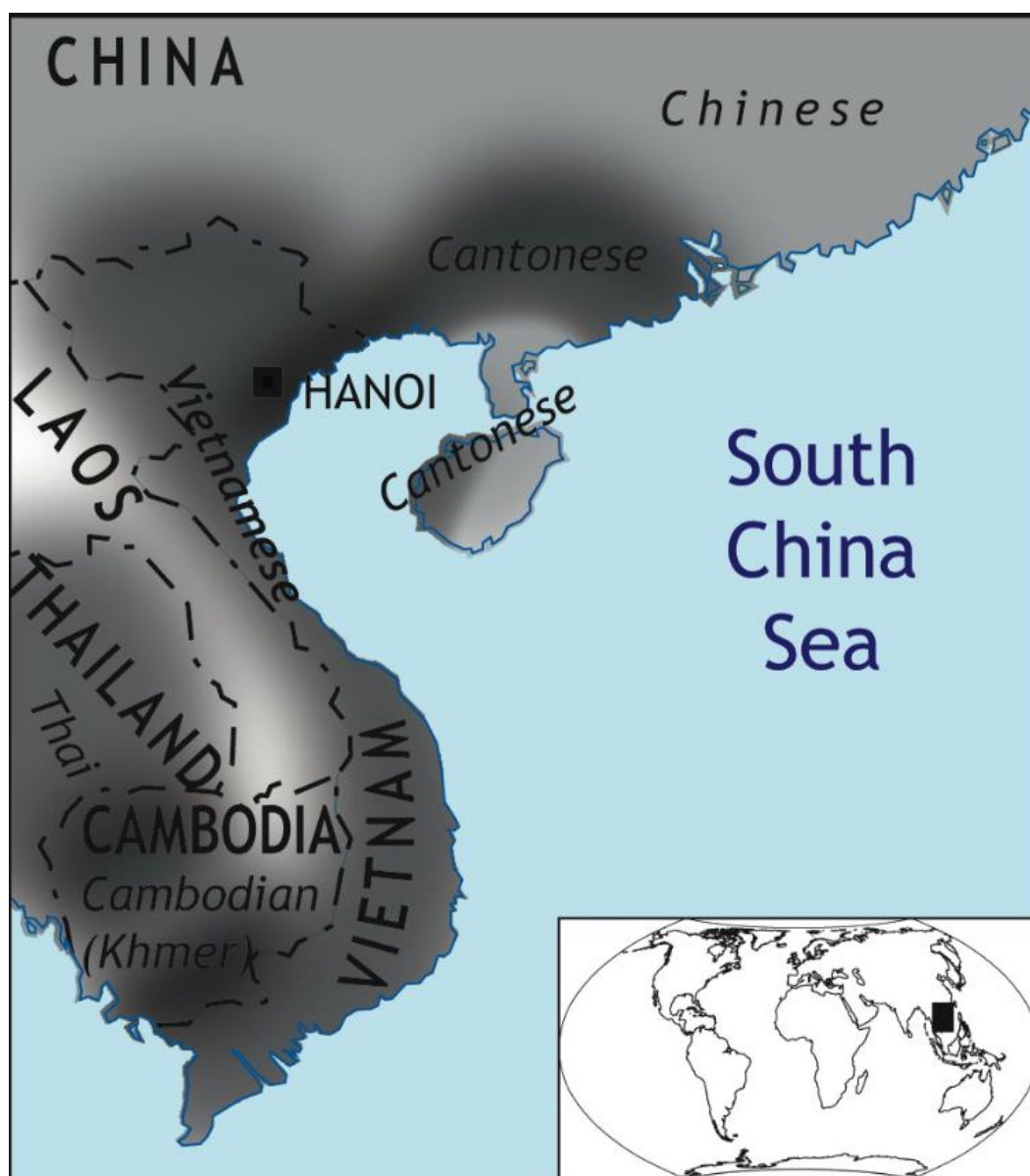
Studying the loanword situation in Vietnamese and sorting out what is borrowed and what is retained is a complex matter due to the extensive time frame involved (over two millennia) and the number of potential language contact situations in a region with hundreds of languages. The linguistic affiliation of Vietnamese has been contested over the past century, but currently, available linguistic evidence continues to confirm that Vietnamese has a solid core of basic vocabulary cognates in common with Mon-Khmer languages (Alves 2006), and not Tai-Kadai, Austronesian, Sino-Tibetan, nor is it a mixture of languages, as some have claimed.<sup>2</sup> Overall, the Vietnamese lexicon is highly sinified, though other language contact situations with Tai languages,<sup>3</sup> French, and English have left their imprints as well.

### Map 1: Geographical setting of Vietnamese

<sup>1</sup> This three-way division represents more of somewhat arbitrary points on a dialect chain continuum.

<sup>2</sup> Still, the notion that the Vietnamese share their origins with the many hill-tribes of Southeast Asia does not have popular acceptance in Vietnam. This can be seen in the fact that government-accepted maps of language groups in Vietnam show Vietnamese and Muong as distinct from any other language families.

<sup>3</sup> The terms *Tai* and *Thai* are distinct. *Thai* refers to the national language of modern Thailand, while *Tai* refers to the main language group within the Tai-Kadai language family, to which Thai belongs. In this article, *Proto-Tai* and *Tai* will both refer to the language source from which Vietnamese is posited to have borrowed.



## 2. Sources of data

The core data sources include dictionaries and databases of historical reconstructions. The primary Vietnamese dictionaries consulted for this study include that of Nguyễn Đình Hoa (1966) and the online Free Vietnamese Dictionary Project. The former was useful in identifying core meanings of Sino-Vietnamese characters instead of only the bisyllabic compounds and in listing standard Sino-Vietnamese counterparts to native Vietnamese words, some of which turn out to be early Chinese loans. The latter dictionary provided more modern entries. As for Sino-Vietnamese readings, the substantive “Từ điển Việt Hán” (Vietnamese-Chinese Dictionary) (1991) was the main source. Lastly, various reconstructions were consulted, including Starostin’s online Old Chinese and Middle Chinese reconstructions, Li Fang-Kuei’s Proto-Tai (Li 1977), and Shorto’s Proto-Austroasiatic (Shorto 2006). The entries in these sources were compared with posited loans for consideration of the likelihood of those words’ being borrowed.

As there are several sources of loanwords in Vietnamese, the sources of information vary. Tai loanwords are much more speculative than the other sources of loanwords as they were most likely borrowed in the Pre-Han Dynasty era and thus have undergone tremendous phonological change. Moreover, there are no written records to document the assumed contact. The list of a few dozen possible Tai loanwords considered for this database is in the work of Nguyễn Tài Cẩn (1995: 322). Regarding Chinese loanwords, which were borrowed over a period of two thousand years, only those words listed in Sino-Vietnamese dictionaries are considered to be Chinese with absolute certainty. Other forms, including those borrowed in the earliest period of Sino-Vietnamese contact, dialectal borrowing, and some nativized words are judged to be likely loanwords only when enough supporting phonological, semantic, and historical evidence could be provided. The words of this Old Sino-Vietnamese layer come from a range of studies, crucially, the works of Wang Li (1958), Pulleyblank (1981, 1984), Haudricourt (1954), and Nguyễn Tài Cẩn (1995). Finally, resources on western loanwords, primarily French but also some English loanwords, include Nguyễn Đình Hoa’s dictionary (1966), Barker’s (1969) study of the phonology of French loanwords in Vietnamese based on over 130 words, and a 2006 dictionary of new words in Vietnamese published by the Vietnamese Viện Ngôn Ngữ Học (Institute of Linguistics).

3. Language contact and borrowing in Vietnamese

There is both comparative linguistic data and written documentation showing that lexical borrowing in Vietnamese extends back more than two thousand years, though the initial period of language contact is less certain. 28% of the Vietnamese vocabulary in the sample, or about 415 words of the 1,477 words in the database, is either clearly or probably borrowed, with another 2% which are less likely to be loanwords. Of the loanwords in Vietnamese in the database, 90% are from Chinese (less than one-third from the Han dynasty era and more than two-thirds from the Tang Dynasty era onward), 4% from French, 3% from Tai languages, and a small handful from English and Chamic. There are only 16 clear or probable calques, most of which are likely related to Chinese words. Of the remaining vocabulary in the study, about 370 (about 25%) of the vocabulary in the lexical data, are retentions from within the Austroasiatic language family at various levels on the sub-branches, and nearly 670 (about 44%) cannot be connected to any particular source and must be considered strictly Vietnamese, though undoubtedly, other data will further identify sources of this dataset.

In the database, 31% of the loanwords were labeled as being added to the Vietnamese lexicon, meaning such words did not exist in Vietnamese before being borrowed, and 21% were labeled as coexisting with native Vietnamese vocabulary. The loanwords which were added to Vietnamese were either new technologies or culturally specific concepts. This was clearly the case for the more recent Western loanwords, but loanwords from Han and Pre-Han Dynasty era are, naturally, harder to assert with certainty, for example, the Chinese loan “spoon” and the Tai “canal.” New concepts and technology were introduced primarily by the Chinese. More recently, within the past hundred years, French and most recently English borrowings reflect borrowing under conditions of the introduction of new technologies.

This study divides the eras of borrowing into five general categories, though in some periods, more than one language was the source of borrowing. The general categories are summarized in Table 1, after which additional discussion is given.

Table 1: Timeline of Vietnamese language contact

Era	Details
<i>Pre-Han Dynasty</i> (pre-200 BCE)	<ul style="list-style-type: none"><li>• No historical documents about Vietnam</li><li>• Probably contact among ethnic groups in the region, including at least Austroasiatic, Tai-Kadai, and Non-Chinese Sino-Tibetan</li><li>• Likely borrowing from Proto-Tai</li></ul>
<i>Han Dynasty</i> (200 BCE to 200 CE)	<ul style="list-style-type: none"><li>• Near the end of the era of Old Chinese</li><li>• 100 BCE: beginning of Chinese political domination of modern-day Southern China</li><li>• Initial Sino-Vietnamese contact and forced adoption of Chinese political and cultural systems in the region</li><li>• Establishment of the Sino-Vietnamese groups through intermarriage in influential families</li></ul>
<i>Tang</i> (600s to 900s CE)	<ul style="list-style-type: none"><li>• Beginning of the era of Middle Chinese</li><li>• Spread of the Chinese rhyming dictionaries<sup>4</sup> throughout East Asia and Vietnam, thereby providing a substantive source of lexical borrowing</li><li>• The end of the Tang Dynasty is followed by Vietnamese political independence from China.</li></ul>
<i>Independence</i> (900s to early 1800s)	<ul style="list-style-type: none"><li>• Chinese cultural and linguistic foundation already established, particularly among elite Vietnamese</li><li>• Continuing, deepening Chinese cultural influence despite political independence</li><li>• Gradually increasing numbers of Chinese words in written and spoken Vietnamese</li><li>• Some dialect borrowings over the centuries, extending into the modern era</li></ul>
<i>Modern Era</i> (late 1800s to 1900s)	<ul style="list-style-type: none"><li>• French colonialism in the region (late 1800s to mid-1900s), resulting in some lexical borrowing, primarily of new terms for clothing, food, and technology</li><li>• From the late 1800s, rapid spread of Sino-neologisms (terms created with Chinese morphs initially by both Chinese and Japanese to translate Western concepts) into Vietnamese</li><li>• Spread of Communism in early to mid-1900s partially under Chinese influence</li><li>• From the 1960s, borrowing of some English words during American-supported war in Vietnam</li><li>• End of 20<sup>th</sup> century English borrowings of technological terms</li></ul>

Based on somewhat tentative data, essentially a few dozen of loanwords at most (Nguyễn Tài Cẩn 1995: 322), Proto-Tai peoples may have come into contact with the ancestors of the Vietnamese before the Han Dynasty (200 BCE to 200 CE),<sup>5</sup> at which time they presumably shared with the Vietnamese some technology related to agriculture and animal husbandry and the associated vocabulary (e.g., *vịt* “duck” and *mương* “canal”). Some complications in dealing with the origins of the Vietnamese include etyma appearing in multiple language families (e.g., *mắt* “eye” and *này* “this” with possible cognates in both Austronesian and Sino-Tibetan). Such etyma are weak if not useless support for claims of genetic linguistic affiliation other than Austroasiatic until other sources of data can be used to clarify their widespread appearance in the region.<sup>6</sup> Overall,

<sup>4</sup> The “rhyming dictionaries” in China included various categories, but overall, these were dictionaries that indicated common categories of initial consonants, rhymes, and tones, though without indication of explicit phonetic features of the sounds.

<sup>5</sup> In general, two assumptions are the guidelines for this study: (a) a form that is recognized in Proto-Mon-Khmer is not a loanword even when it appears in other language families, such as Tai-Kadai or Sino-Tibetan, and (b) non-Mon-Khmer etyma that appear in Chinese are loanwords from Chinese, not other languages. These, of course, are subject to corroborating evidence, such as archeological discoveries or written records from the era.

<sup>6</sup> Indeed, it was precisely the fact that Chinese words are seen in both Vietnamese and Tai languages that Haudricourt (1954) was able to argue against the Tai-Vietnamese affiliation posited by Maspero (1912).

the number of Tai loanwords is small (less than 1% of the database) and suggests the occurrence of meaningful if ultimately light Tai-Vietnamese contact in pre-recorded history.

The socio-historical conditions under which the Chinese became the dominant sociolinguistic influence are clearer. A more detailed description of Sino-Vietnamese language contact is provided in §4, so only a brief summary is given here. In general, historical records clearly show a southern expansion of Chinese which resulted in both a mixture of Chinese and Proto-Tai peoples and a southward migration of the Proto-Tai peoples into modern-day Southeast Asia. However, the point at which the Chinese took northern Vietnam changed the sociolinguistic dynamic dramatically. From the Han Dynasty and probably some centuries afterwards, Vietnamese borrowed hundreds of words that form a notable portion of Vietnamese vocabulary today (113 words, about 8% of the database), including a semantically wide range of words which have become entirely nativized in Vietnamese. The actual number of Han era loanwords beyond the database is most likely substantially higher, at least 300, many of which are highly integrated into Vietnamese. The imposition of Chinese customs and introduced technologies, trade and large numbers of Chinese who settled in Vietnam were the impetus for these borrowings. Words borrowed in this era are not recognized by Vietnamese as anything other than native Vietnamese and are difficult to identify as loanwords even by historical linguistic specialists in Vietnamese. The number of loanwords in the next millennium is substantially higher, as discussed in the next section.

While Europeans missionaries were present in Vietnam from the 1500s, it was not until the around the end of the 19<sup>th</sup> century that French colonialization began to have noticeable effects on the Vietnamese lexicon. The loanwords from the French colonial period include terms related to clothing, food, and household goods (e.g., *ca vát* “tie” (French *cravate*), *bơ* “butter” (French *beurre*), etc.) and then emerging technology (e.g., *ga* “train station” (French *gare*) and *pin* “battery” (French *pile*)), all of which are reflective of the social presence the French had in Vietnam. There was a solid French community in parts of Vietnam, and a French-style education system was influential in some parts of Vietnam. In Barker’s 1969 study, there were over 130 French loanwords. Finally, there have been loanwords from English since the period of substantive American presence in Vietnam since the 1960s (e.g., *mít tỉnh* “meeting”, *ti vi* “TV,” and *top* “top (in reference to the highest in a list)”). Over 70 English loanwords appear in the Institute of Linguistics’ 2006 dictionary of new words in Vietnamese. Interest in learning as a second language has been extremely high in recent years, and current developments in computer technology have led to borrowings without a large community of English speakers in Vietnam (*in-tơ -nét* “internet”, *CPU* “central processing unit,” etc.).

4. Numbers and kinds of loanwords

Tables 2 and 3 summarize the quantitative distribution of loanwords over donor languages, semantic word classes and semantic fields.

**Table 2:** *Loanwords in Vietnamese by donor language and semantic word class (percentages)*

	Chinese	French	Proto-Tai	English	Indo-European	Cantonese	Chamic	Total loanwords	Non-loanwords
Nouns	27.5	2	0.7	0.5	0.4	0.1	0.1	31.3	68.7
Verbs	23.8	-	0.3	-	-	0.6	0.3	25	75

Function words	17.6	-	-	-	-	-	-	17.6	82.4
Adjectives	19.3	-	0.8	-	-	-	0.8	20.9	79.1
Adverbs	25	-	-	-	-	-	-	25	75
all words	<b>25.3</b>	<b>1.2</b>	<b>0.5</b>	<b>0.3</b>	<b>0.3</b>	<b>0.2</b>	<b>0.2</b>	<b>28.1</b>	<b>71.9</b>

**Table 3:** *Loanwords in Vietnamese by donor language and semantic field (percentages)*

	Chinese	French	Proto-Tai	English	Indo-European	Cantonese	Chamic	Total loanwords	Non-loanwords
The physical world	23.1	-	-	-	-	-	1.3	24.4	75.6
Kinship	27.1	-	-	-	-	-	-	27.1	72.9
Animals	7.9	-	3.5	1.7	1.7	-	-	14.8	85.2
The body	19.4	-	-	-	-	-	-	19.4	80.6
Food and drink	17.1	4.7	-	-	1.2	2.4	-	25.3	74.7
Clothing and grooming	16.5	5	-	-	1.7	-	-	23.1	76.9
The house	21.9	-	-	-	-	-	-	21.9	78.1
Agriculture and vegetation	15.4	1.4	1.4	-	-	-	1.4	19.6	80.4
Basic actions and technology	26.9	-	1.3	1.3	-	1.3	-	30.6	69.4
Motion	18.3	-	2.4	-	-	-	-	20.7	79.3
Possession	32.6	-	-	-	-	-	-	32.6	67.4
Spatial relations	26.8	-	-	-	-	-	-	26.8	73.2
Quantity	15	-	-	-	-	-	-	15	85
Time	26.7	-	-	-	-	-	-	26.7	73.3
Sense perception	12	-	-	-	-	-	-	12	88
Emotions and values	29.2	-	-	-	-	-	2.1	31.2	68.8
Cognition	44.6	-	-	-	-	-	-	44.6	55.4
Speech and language	39.3	-	-	-	-	-	-	39.3	60.7
Social and political relations	60	-	-	-	-	-	-	60	40
Warfare and hunting	47.5	-	-	-	-	-	-	47.5	52.5
Law	45.1	-	-	-	-	-	-	45.1	54.9
Religion and belief	49.1	-	-	-	-	-	-	49.1	50.9
Modern world	41.7	16.7	-	3.3	-	-	-	61.7	38.3
Miscellaneous									
function words	18.5	-	-	-	-	-	-	18.5	81.5
	<b>25.3</b>	<b>1.2</b>	<b>0.5</b>	<b>0.3</b>	<b>0.3</b>	<b>0.2</b>	<b>0.2</b>	<b>28.1</b>	<b>71.9</b>

5. Sino-Vietnamese

As Chinese is the chief source of loanwords in Vietnamese, it is necessary to explain in more detail the historical sociolinguistic relationship between Chinese and Vietnamese. Over a quarter of the database in this study consists of words of

Chinese origin. However, an even larger percentage of words have entered Vietnamese via written transmission, as estimates based on Vietnamese language dictionaries suggest that 70% or higher of the Vietnamese lexicon is Chinese in origin. A majority of Chinese vocabulary entered Chinese without the presence of a large bilingual community, and so, it can be assumed that Sino-Vietnamese borrowing occurred largely via written language. The two primary eras of borrowing via spoken contact include the Han Dynasty and the modern era, both of which are periods in which large Chinese communities resided in major cities in Vietnam, though neither era was a time when most of the Chinese loanwords entered Vietnamese. About 30% of Chinese loanwords in this database were borrowed via this spoken contact, while about 70% in this study were most likely spread into Vietnamese since the end of the Tang dynasty largely through the educated elite in Vietnam. These include many borrowed since the early 20<sup>th</sup> century spread of Sino-neologisms, combinations of Chinese morphs to create new words, which spread throughout China, Japan, Korea, and Vietnam. There are five primary layers of Sino-Vietnamese, as listed and described briefly in Table 4.

**Table 4:** Eras of Transmission of Chinese into Vietnamese

Layer	Period	Characteristics
<i>Old Sino-Vietnamese</i>	Han Dynasty (mainly the first two centuries CE, extending into the Six Dynasties era)	Colloquial/daily vocabulary borrowed; Borrowed through spoken transmission; Considered by Vietnamese to be native Vietnamese vocabulary
<i>Sino-Vietnamese Proper</i>	Tang Dynasty, 7 <sup>th</sup> century onward	Pronunciation based on rhyme dictionary categories; Borrowed largely through texts shared by Vietnamese literary elite; Recognized by Vietnamese as Chinese in origin; Gradually incorporated into Vietnamese over following centuries, thereby allowing additional layers of doublet readings
<i>Post-Tang Sino-Vietnamese (pre-modern)</i>	10 <sup>th</sup> to 19 <sup>th</sup> centuries	Nativized readings of highly integrated Tang Dynasty Sino-Vietnamese; Considered to be Nôm readings (Chinese-based orthography to represent native Vietnamese vocabulary) rather than Chinese
<i>Post-Tang Sino-Vietnamese Colloquial (modern)</i>	19 <sup>th</sup> to 20 <sup>th</sup> centuries	Smaller numbers of Chinese words borrowed in this era; Some likely Cantonese culinary terms (e.g., <i>lạp p xư ố ng</i> “sausage” and <i>mì chính</i> “monosodium glutamate”) <sup>7</sup>
<i>Sino-neologisms</i>	Beginning of the 20 <sup>th</sup> century	Words created by Chinese and Japanese utilizing Chinese morphs to translate newly introduced Western concepts and terms, and then brought into Vietnamese; Largely abstract and/or technological in nature; Thousands of words borrowed in a short period

The earliest period of Sino-Vietnamese contact began more than two thousand years ago when the Han Dynasty, which marks the end of the period of Old Chinese, vastly expanded Chinese political control as far south as modern-day Northern Vietnam. This era begins officially in 111 BCE, though the more intense Sino-Vietnamese language contact occurred somewhat later, with the era of significant cultural influence taking place in the later Eastern Han era, when

<sup>7</sup> The forms are considered Cantonese in origin because they are not standard Sino-Vietnamese literary pronunciations, they are phonetic close to Cantonese, and Cantonese has been dominant variety of Chinese spoken among Chinese living in Vietnam since the 20<sup>th</sup> century (and likely before).

there were both large-scale immigrations of Chinese into Vietnam and administrative orders aimed at the sinicization of the newly settled territories. Near the very beginning of the Christian era, 20,000 Chinese soldiers (8,000 from northern China and 12,000 from Southern China) settled in modern-day Northern Vietnam (Taylor 1983: 49). These Chinese soldier-settlers, along with periodic influential Chinese escaping turmoil in China, intermarried with Vietnamese and established an influential community in Vietnam. At the same time, Chinese customs were imposed on the Vietnamese, including Chinese style marriages, clothing, culinary traditions and accoutrements, and other customs affecting daily life (*ibid.* 33), as reflected in the lexical evidence in Table 3. Many of the words borrowed in this period which belong to the Old Sino-Vietnamese layer of were borrowed again through standardized dictionaries during and after the Tang Dynasty, the period of Middle Chinese, creating numerous Sino-Vietnamese doublets, as listed in Table 5.

**Table 5:** *Old Sino-Vietnamese Data (Starostin’s Reconstructions)*

OSV	Meaning	SV	Meaning	OC	MC	CH
<i>mũ</i>	hat	<i>mạ o</i>	hat	<i>māwh</i>	<i>ma<sup>^</sup> w</i>	帽
<i>giày</i>	shoe	<i>hài</i>	shoe	<i>γ riē</i>	<i>γ ä</i>	鞋
<i>gả</i>	marry	<i>giá</i>	marry	<i>krǎ h</i>	<i>kạ<sup>`</sup></i>	嫁
<i>vợ</i>	wife	<i>phụ</i>	woman	<i>bwə</i>	<i>bə w</i>	婦
<i>cúi</i>	bow, prostrate oneself	<i>quị</i>	kneel	<i>ghwáj</i>	<i>wé</i>	跪
<i>lạ y</i>	kowtow	<i>lễ</i>	ceremony	<i>riǎ́ j</i>	<i>líej</i>	禮
<i>phép</i>	rule, law	<i>pháp</i>	rule, law	<i>prap</i>	<i>pwi p</i>	法

The combination of forced adoption of cultural customs and the creation of a bilingual community founded a layer of Chinese linguistic and cultural influence that was to serve as the foundation upon which a more substantive quantity of borrowing was to follow. After the fall of the Han Dynasty, Chinese political domination of Vietnam continued, and the Sino-Vietnamese families maintained their social status in Vietnam, though it appears that they were firstly Vietnamese in ethnic identity (Taylor *ibid.*). In addition to the historical evidence, phonological evidence has been provided showing patterns in initials and tones, such as the work of Wang (1958), Mei (1970), and Tryon (1979), which further increases the likelihood of these being borrowed from Chinese in this era.

From this point onward, Chinese became the primary donor of loanwords into Vietnamese. In the 7<sup>th</sup> century CE, the powerful Tang Dynasty began, and Chinese rhyming dictionaries were spread throughout modern East Asia, including Vietnam. These dictionaries, which indicated the approximate readings of Chinese characters by comparison with the initial consonants, tone categories, and syllable shapes, effectively shared the entirety of the Chinese lexicon. The spread of these dictionaries, and other Chinese literary traditions, and the influence on the languages of the recipient cultures took place over a period of centuries. During the Tang Dynasty, and the subsequent Song Dynasty, in addition to the collection of classic Chinese writings, Chinese rhyme dictionaries containing the approximate pronunciations of dozens of thousands of Chinese characters spread throughout China, Korea, Japan, and Vietnam. In the following centuries, thousands of Chinese words gradually entered both the written and spoken languages in the region. This layer of Chinese loanwords is recognized by



the Vietnamese as Chinese loanwords since they are documented as such. This large intake of Chinese vocabulary has had some effects on Vietnamese phonology, but in the area of grammar, the influence is largely limited to borrowing of grammatical words with perhaps some effect on Vietnamese noun phrase structure due to the borrowing of dozens of measure words from Chinese (Alves 2001).

Borrowings from Chinese dialects, with which Vietnamese would have had the most immediate spoken contact, are relatively few in number. Instead, the larger source again came via texts. Throughout East Asia, there was a tremendous increase of the use of Chinese morphs to translate western texts, a movement which started initially in Japan and China starting in the late 1800s (Masini 1993). These newly coined Sinitic words made their way in massive quantities, certainly numbering in the thousands, into Korea and Vietnam (Sinh 1993). Such “Sino-neologisms” consisted largely of bisyllabic forms which in some cases utilized ancient words but which also include the combining of these monosyllabic morphs in ways similar to the use of Greek and Latin lexical material among European languages.

In the accompanying database, the influence of this borrowing covers all semantic fields, from the physical world, to body parts, to spatial relations and social and political relations. All told, this already highlights the tremendous influence Chinese has had in Vietnamese. However, as some researchers have suggested that as much as 70% of the Vietnamese vocabulary is Chinese in origin, the number 26% in the current must be compared with other semantic areas not included. For example, the database does not include a number of semantic fields which can reveal and accurately highlight the influence that Chinese has had on the Vietnamese lexicon. While family relations are included in the study, a more systematic study of terms for family relations reveals that Chinese is a dominant portion of those terms and indeed the overall system of familial terms and the hierarchical structure (e.g. Benedict 1947). In addition, the hierarchical relationships obtaining among those terms contributed substantially to the complex system of pronominal terms of address in Vietnam.

Another lexical category not included in the accompanying database is names. This borrowing may extend back to the Han Dynasty when the administration imposed Chinese marriage customs and a system of managing households for tax purposes. The borrowing of Chinese into Vietnamese includes not only the names but also the entire naming system, including the order (family, middle, and given names) and indication of gender. Some of the most common Chinese surnames are also common throughout Vietnam (e.g. Trần 陳, Hoàng 黃, Lý 李, etc.). The use of gender-marking in names (e.g. the Vietnamese intercalary/middle name Vietnamese Mỹ, Chinese měi 美 “beautiful” for females or Vietnamese Văn, Chinese wén 文 “literature” for males) is also part of the Vietnamese system of personal names.

**Table 6:** Sample of Chinese Name System in Vietnamese<sup>8</sup>

	Family	Middle	Given
Vietnamese	Trần	Mỹ	Lan
Pinyin pronunciation	Chén	Měi	Lán
Chinese Characters	陳	美	蘭

Of particular note is the fact that the percentage of Chinese loanwords in this study is significantly smaller than the purported 70% or higher percentages of Chinese loanwords based on word counts in dictionaries. Certainly, considering all of the words in dictionaries rather than a selected portion of a language, such as high frequency words in a language, gives a skewed picture of the loanword situation of any language. At the same time, however, this study provides a

<sup>8</sup> Both the family name and given name in the table have alternate pronunciations as indicated.

number that may not fully clarify the depth of the Sino-Vietnamese language contact situation, which will be more fully addressed in the next section.

## 6. Grammatical borrowing in Vietnamese

Chinese has left a solid grammatical imprint on Vietnamese in the areas of grammatical vocabulary and word-formation patterns. While some of this influence may have been the result of spoken contact, most appears to have occurred since or after the period of Vietnamese political independence, suggesting that in fact, most grammatical vocabulary of Chinese origins was borrowed via written language. This detail corresponds to the idea that Vietnamese syntax has been only minimally influenced by contact with Chinese since more intense spoken contact would likely result in more significant grammatical transfer. The one instance of syntactic borrowing is the ordering of numerals and classifiers and measure words in general before nouns, following the Chinese pattern but which is the opposite of other major languages in Southeast Asia (e.g., Thai and Cambodian, in which numerals and measures follow head nouns) (Alves 2001). The borrowing of dozens of measure words and some classifiers from Chinese implies the effect of trade, a kind of spoken contact between Chinese and Vietnamese even during political independence from China.

Rather than structural transfer, Sino-Vietnamese grammatical influence is primarily in the area of grammatical vocabulary. Alves (2005) lists over two dozen core grammatical words in Vietnamese of Chinese origin, not including the dozens of measure words noted above. Some of these were borrowed grammaticalized forms (e.g, *tạ* 在 ‘at,’ *như* 如 ‘as’, etc.), but there is a sizeable number of grammatical words in Vietnamese of Chinese origin words which were not originally grammatical in Chinese (e.g., *quá* ‘very’ (Chinese 過 *guò* ‘exceed’) and *nhất* ‘most’ (Chinese — *yī* ‘one’)) (Ibid.). Among the Old Sino-Vietnamese loanword layer, of which there are over 400 solid etyma, very few grammatical loanwords appear. Most of the grammatical borrowings took place after independence from China took place, in the Post-Tang Dynasty era, during which ultimately thousands upon thousands of Chinese loanwords entered the speech of the Vietnamese population at large. Many of the grammatical words borrowed from Chinese occurred in the period after the Tang Dynasty, most likely through the literary elite of Vietnam. Grammatical words such as *từ* “from” (standard Sino-Vietnamese *tự*, Chinese 自 *zì*) and *bị* “(passive marker)” (Chinese 被 *bèi*), which are not spoken Cantonese Yue dialect loans, were likely transmitted by the educated upper class in Vietnamese who were often required to be literate in Chinese. Of the Chinese dialect loanwords, many are related to food, suggesting a spoken means transmission through trade situations.

In addition to grammatical vocabulary, Vietnamese word formation patterns have also been at least partially influenced by Chinese. Of the over 550 of the words in the database which are analyzable compounds, about 150 are Chinese borrowings. This figure does not include the many loanblends that also contain borrowed Chinese lexical material though which are not strictly loanwords. This is in part due to the strong compounding word formation pattern in Chinese, a natural pattern which has been enhanced in Vietnamese. In addition, the monomorphemic words in Chinese have come in many instances to serve as affixes. In some cases, new words have been created in Vietnamese which do not exist in Chinese, such as *trực tuyến* “on-line” from the Chinese syllables 直 *zhí* “direct” and 線 *xiàn* “line,” (Bùi 2001: 153) but which in Chinese means only “a straight line,” while “on-line” in Chinese is 在線 *zài xiàn*, with a different initial morpheme. Word-formation and Chinese morphological material is discussed further in §7.

## 7. Integration of loanwords in Vietnamese

The general perception by Vietnamese speakers as to whether or not words are non-Chinese foreign in origin depends on time depth and the source language and culture. In general, time-depth first determines the perception by native speakers of the “foreignness” of words, though whether or not the word appears to be culturally indigenous is also a factor. Finally, the association with a written tradition is another way that loanwords are readily identified, with Chinese as dominant example. The more education speakers of Vietnamese have, the more awareness they have of the origins of standard, literary Sino-Vietnamese words as well as more recent Western loanwords.

The probability that some of the posited Tai loanwords from the earliest period of language contact, possibly even predating the earliest Sino-Vietnamese contact, is high, though this cannot be asserted with absolute certainty. That they are loanwords is supported by the number of words within a semantic clustering (i.e., they are primarily related to agriculture and animal husbandry) in addition to phonological patterns. However, due to the vast time depth and the fact that such words are solidly part of native Vietnamese culture, words of Tai origin are generally viewed as native Vietnamese words by Vietnamese speakers. In addition, these words must have been spread via spoken transmission, and thus there is no written documentation.

Chinese words borrowed during initial Sino-Vietnamese contact over fifteen hundred years ago are so well phonologically integrated in Vietnamese that such forms are perceived as native vocabulary (e.g., *đũa* “chopsticks” from the archaic Chinese word *zhù* 箸 “chopsticks”). In contrast, Chinese words borrowed during from the Tang dynasty on, the time of the spread of the Chinese rhyming dictionaries which left a permanent literary tradition, are formally recognized by educated Vietnamese as Chinese in origin. Indeed, such words are listed in specialized Sino-Vietnamese dictionaries. There is also a set of Chinese words which were likely nativized after the Tang dynasty that are considered native Vietnamese words (e.g., *dùng* “to use” is considered native Vietnamese and was most likely nativized phonetically at some point after the borrowing of the synonymous *dụ ng*, considered to be Sino-Vietnamese, from Chinese 用 *yòng*).<sup>9</sup> Finally, there are Chinese dialect borrowings, presumably borrowed through spoken contact, which are not generally recognized as foreign in origin (e.g., *mì* “noodles” possibly from a variety of a Chinese Southern Min dialect pronunciation of 麵 *miàn*).

Another aspect of the massive borrowing of Chinese vocabulary is the large number of synonyms, though the Chinese morphs tend to have more restricted usage. For example, the native word for “water” *nước* in Vietnamese can only be replaced with the Sino-Vietnamese *thủy*, Chinese 水 *shǔi*, in names or within highly artistic or literary language or in specific lexical compounds in which *thủy* is functioning as a bound morpheme rather than a word. When there are native and Sino-Vietnamese synonyms, the Chinese vocabulary tends to have a higher status in Vietnamese. For example, there are two words for “kidney” in Vietnamese, *cật* and *thận*. *Cật* is the native and tends to be used in referring to animal kidneys, while *thận*, from Chinese 腎 *shèn*, tends to refer to human kidneys.

Within the last century of borrowings from Western languages, primarily French but also some from English, there do appear a solid number of permanent loanwords. Most of these permanent elements are French in origin (e.g., *áo sơ*

---

<sup>9</sup> This category of Sino-Vietnamese words is particularly uncertain since they differ in only single phonological aspect, though in some instances, such is also the case for Han Dynasty borrowings, though generally, Han Dynasty loans, due to the substantial time depth, have more than one phonological variant. More work needs to be done in this area to clarify these categories.

*mi* “shirt” (from French *chemise*),<sup>10</sup> *bơ* “butter” (French *beurre*) etc.). However, some French words that were common in the early 20<sup>th</sup> century have gone out of use since the end of the French colonial era (e.g., *bông* “coupon” from French *bon* or *mãng-đã* “money order” from French *mandat*). Some more recent borrowings are from English, as noted in §3. It is too soon to tell whether or not these words will become permanent items in the Vietnamese lexicon. Some of these forms are more likely to have required additional phonological adaption (e.g., the French loanword *pin* “battery” as unaspirated /p/ is not part of any native Vietnamese or Sino-Vietnamese syllables), thereby marking them as foreign.

Another aspect to consider in the integration of Chinese loanwords is ways in which lexical material is seen in Vietnamese grammar, including both morphological material and grammatical vocabulary. A noticeable number of grammatical words in Vietnamese are of Chinese origins, as discussed in §6. As for lexical material contributing to source of morphological material, compounding is an extremely common means of word-formation in Vietnamese, of which there are three general categories: numerous instances with native material only (e.g., *đá ngà m* (stone-underwater) “reef”), the many thousands with borrowed material only, most of which are Chinese in origin (e.g., *triết học* “philosophy” from Chinese 哲學 *zhé xué*), and loanblends containing both native and non-native material, which again are most often Chinese in origin (see Table 5). In terms of statistical distribution, loanblends are the least common of the three categories. It may be the case that in numbers based on dictionary counts that the bisyllabic words with two Chinese etyma are the most common, though in daily spoken Vietnamese, native only and Chinese only may be similar.

**Table 5:** *Samples of loanblends*

Vietnamese	Meaning	Source of 1 <sup>st</sup>	Source of 2 <sup>nd</sup>
linh thiêng	to have supernatural power	Chinese 靈 <i>ling</i>	native
ma quỷ	ghost	native	Chinese 鬼 <i>guǐ</i>
ăn cắp	to steal	native	Chinese 劫 <i>jié</i>

8.Vietnamese attitudes toward borrowing

Vietnamese has been receptive to loanwords in periods of substantive sociolinguistic contact. However, loanwords borrowed mainly by spoken means of transmission are in the minority, while large-scale lexical borrowing occurred through written sources. The borrowing of Chinese written words in particular, in contrast with the early Han Dynasty borrowings or more recent dialectal borrowings, has been connected to a highly influential tradition of high culture that has allowed many thousands of Chinese words to become part of the Vietnamese lexicon. While there was a late twentieth century movement towards the reduction of Chinese vocabulary in particular and the increased use of “native” Vietnamese vocabulary (a conscious attempt to remove what is specifically Sino-Vietnamese neologisms after 1975),<sup>11</sup> generally, both the Vietnamese lexicon and consequently the Vietnamese morphological system has continued to be to influenced by Chinese, though in substantially smaller

<sup>10</sup> In this compound, *áo*, a native Vietnamese word referring to clothing on the upper part of the body, is the head noun.  
<sup>11</sup> In some cases ironically, words of Chinese origin were still replaced with words of Chinese origin. For example, in replacing the Sino-Vietnamese *hỏa xa* (Chinese 火車, *huǒ chē*, (fire-vehicle)) with *xe lửa* “train” (vehicle-fire), the standard dictionary Sino-Vietnamese reading *xa* ‘vehicle’ was replaced with nativized Sino-Vietnamese term *xe* ‘vehicle’.

numbers than those which entered Vietnamese at the beginning of the twentieth century.

Over the past 150 years, many words and names have been borrowed as a result of the spread of technology. There appears to be readiness to borrow terms when no native equivalent exists, particularly in the case of modern technology. This may, in part, be due to the use of a romanized alphabet, which allows even words that violate Vietnamese phonotactic rules to be printed (e.g., *fax* or *făc* “fax machine” as Vietnamese does not have any final fricatives) and which even appear to have some currency in spoken Vietnamese as well. This is in contrast with Chinese, which relies on sometimes somewhat challenging methods of using Chinese characters to transcribe foreign terms. While in some instances such terms have been spread into Vietnamese (e.g., *hũ u cơ* from Chinese 有機 “organic”), Vietnamese is able to utilize its alphabet to readily borrow foreign words and sounds (e.g., *in tơ nết* “internet”, unlike Chinese). The major period of borrowing from Chinese has largely passed, though the vast quantity of Chinese lexical and morphological material has been so thoroughly incorporated into Vietnamese that it continues to be reused in the creation of new terms, similar to the use of Latin and Greek words and morphology in the creation of new words in English.

While there is some concern expressed by some Vietnamese about foreign loanwords, the use of foreign terms is seen by some Vietnamese as fashionable. Knowledge of foreign languages is a status symbol and a sign of education, as can be the use of foreign terms in spoken Vietnamese. Foreign language schools are increasingly popular, particularly in major cities, and computers and televisions are increasingly seen in Vietnamese homes.<sup>12</sup> In addition to access to foreign languages through the media and the internet, the overseas Vietnamese also bring back foreign terms. Recently, more English words have entered Vietnamese, mostly in the area of technology, though it is too soon to see how many of these will become a permanent part of Vietnamese.

While non-Chinese loanwords are generally readily recognized by Vietnamese, the degree to which Vietnamese recognize Sino-Vietnamese vocabulary depends on the level of education. Chinese is to Vietnamese as English is to Latin, and just as speakers of English who have more education and, particularly those who have studied Latin or any of the various Romance languages, educated Vietnamese and especially those who are familiar with Chinese and have even studied Sino-Vietnamese as a subject are keenly aware of words of Chinese origin in Vietnamese and are able to take advantage of that knowledge to understand Sino-Vietnamese compounds and even create new terms using those monosyllabic morphemes.

However, while modern western terms are being increasingly borrowed, it is still the case that Chinese is often the first source for the adoption of new terms. The substantive borrowing of Chinese over a vast period came with a social status and both literary and cultural tradition that will continue to be an integral part of the Vietnamese lexicon.

## 9. Concluding thoughts

Based on this specific research project, Chinese loanwords constitute over a quarter of Vietnamese, though in fact, it does not take into account a number of

---

<sup>12</sup> The percentage of Vietnamese using the internet in 2000 was 0.2%, while by 2006, it had risen to 14.68% (Nguyễn Mạnh Tươi ng).

other factors that highlight the overall influence of this language contact situation. If the study involved a more representative daily, spoken sample of Vietnamese with Vietnamese cultural attributes as the center, the number would certainly be higher. Nevertheless, the number is substantially lower than if entire dictionaries are considered, so this is a useful approach to give some perspective. On the matter of means of transmission, a solid majority of Chinese loanwords in Vietnamese were not borrowed by groups in spoken contact but rather via texts. However, at an early stage, Chinese communities popularized Chinese cultural customs, along with the vocabulary, thereby setting the stage for these later eras of continued Sino-Vietnamese borrowing.

## References

- Alves, Mark J. 2001. What's so Chinese about Vietnamese? in *Papers from the Ninth Annual Meeting of the Southeast Asian Linguistics Society*. ed. Graham W. Thurgood. Tempe, Arizona: Arizona State University. 221-242.
- Alves, Mark J. 2005. Sino-Vietnamese grammatical vocabulary and triggers for grammaticalization. *Hội Thảo Quốc Tế Ngôn Ngữ Học Liên Á Lần Thứ VI (The 6th Pan-Asiatic International Symposium on Linguistics)*. Hanoi: Nhà Xuất bản Khoa Học Xã Hội (Social Sciences Publishing House). 315-332.
- Alves, Mark J. 2006. Linguistic Research on the Origins of the Vietnamese Language: An Overview. *The Journal of Vietnamese Studies*, Vol. 1, No.1-2: 104-130. University of California Press.
- Barker, Milton E. 1969. The phonological adaptation of French loanwords in Vietnamese. *Mon-Khmer Studies* 3:138-147.
- Benedict, Paul K. 1947. An analysis of Annamese kinship terms. *Southwestern Journal of Anthropology* 3: 371-390.
- Bùi Khánh Thế, et. al. 2001. *Mâu thuẫn về Tiếng Việt Hiện Đại* (Several Questions regarding Modern Vietnamese). Hồ Chí Minh City: Nhà Xuất bản, Đại Học Quốc Gia, Thành Phố Hồ Chí Minh.
- Đào, Duy Anh. 1979. *Chữ Nôm: Nguồn Gốc, Cấu Tạo, Diễn Biến* (Chu Nom: origins, formation, and transformations). Hanoi: Nhà Xuất bản Khoa Học Xã Hội.
- Diffloth, Gérard. 1989. Proto-Austroasiatic Creaky Voice. *Mon-Khmer Studies* 15:139-154.
- The Free Vietnamese Dictionary Project*. <http://www.informatik.uni-leipzig.de/~duc/Dict/>.
- Phạm Văn An and Lưu Văn Đương. 1990. *Từ Điển Việt Hán* (Vietnamese-Chinese Dictionary). Hanoi: Nhà Xuất bản Đại Học và Giáo Dục Chuyên Nghiệp.
- Ferlus, Michel. *Proto-Vietic Reconstruction*. Unpublished manuscript.
- Gordon, Raymond G., Jr. (ed.), 2005. *Ethnologue: Languages of the World, Fifteenth edition*. Dallas, Tex.: SIL International. Online version: <http://www.ethnologue.com/>.
- Haudricourt, André G. 1954. Comment reconstruire le Chinois Archaïque. *Word* 10.2-3:351-364.
- \_\_\_\_\_. 1954b. Sur l'origine de la ton de Vietnamien. *Journal Asiatique* 242.69-82.
- Lê, Ngọc Trụ. 1967. *Việt-Ngữ Chánh-Tả Tự-Vị* (Vietnamese Language Spelling Dictionary) (2<sup>nd</sup> ed.). Saigon: Than-Tan.
- Li, Fang Kuei. 1977. *A Handbook of Comparative Tai*. Honolulu: University Press of Hawaii.

- Masini, Federico. 1993. The formation of Modern Chinese lexicon and its evolution toward a national language: The period from 1840 to 1898. *Journal of Chinese Linguistics*, monograph series no. 6.
- Maspero, Henri. 1912. Études sur la phonétique historique de la langue Annamite: les initiales. *Bulletin de l'École Française d'Extrême-Orient* 12:1-127.
- Mei, Tsu-Lin. 1970. Tones and prosody in Middle Chinese and the origin of the rising tone. *Harvard Journal of Asiatic Studies*, Vol. 30: 86-110.
- Nguyễn Mạnh Tư`ng. 2007. Website: Mỏ Vàng Mới (Website: A New Goldmine). *Thành Đạt*, July: 22.
- Nguyễn n, Đình Hòa. 1966. Vietnamese-English Dictionary. Rutland, Vermont: Charles E. Tuttle Co.
- Nguyễn n, Tài Cầ n. 1995. *Giáo trình lịch sử ngữ âm tiếng Việt* (Textbook of Vietnamese historical phonology). Hanoi: Nhà Xuất bản Giáo Dục.
- Pulleyblank, Edwin G. 1981. Some notes on Chinese historical phonology. *Bulletin de l'ecole Française d'Extreme-Orient* 277-288.
- Pulleyblank, Edwin G. 1984. *Middle Chinese: A Study in Historical Phonology*. University of British Columbia Press.
- Shorto, Harry. 2006. *A Mon-Khmer Comparative Dictionary*. Edited by Paul Sidwell with Doug Cooper and Christian Bauer. Pacific Linguistics 579. Canberra: Research School of Pacific and Asian Studies, The Australian National University.
- Sinh, Vinh. 1993. Chinese characters as the medium for transmitting the vocabulary of modernization from Japan to Vietnam in Early 20<sup>th</sup> century. *Asian Pacific Quarterly*, 25.1:1-16.
- Starostin, Sergei. *StarLing Database Server: Old Chinese Database*. Accessed in August 2006.
- Taylor, Keith W. 1983. *The Birth of Vietnam*. Berkeley: University of California Press.
- Tryon, Ray. 1979. *Sources of middle Chinese phonology: a prolegomenon to the study of Vietnamized Chinese*. Southern Illinois University MA Thesis.
- Việ n Ngôn Ngữ Họ c (Institute of Linguistics). 2006. *Từ Điển Từ Mới i Tiếng Việt* (Dictionary of New Words in Vietnamese). Ho Chi Minh: Nhà Xuất bản Phư ơ ng Đông.
- Wang, Li (王力). 1958. Hanyueyu Yanjiu 汉越语研究 (A study on Sino-Vietnamese). *Hanyu Shi Lunwen Ji* (汉语史论文集). Beijing: Kexue Chubanshe. 290-406.