

# What's so Chinese about Vietnamese?

Mark J. Alves

University of Hawaii at Manoa

## 1 Introduction

This study explores the ways in which Chinese<sup>1</sup> has and has not affected the language spoken by the Vietnamese and their ancestors over two thousand years of language contact in what is an example of borrowing rather than shift. Based on comparative lexical, phonological, morphological, and syntactic evidence, the influence of Chinese, though lexically significant, is best viewed as structurally superficial. This paper demonstrates that, at each linguistic level, Chinese influence is primarily restricted to non-structural aspects of Vietnamese, and the various linguistic elements of Chinese have been fit onto a primarily Southeast Asian and Mon-Khmer<sup>2</sup> linguistic template.

Early language contact between the ancestors of speakers of Vietnamese and speakers of Chinese came through administrative, commercial, and individual contact, though the amount of documentation of each type of contact varies (Taylor 1983). Administrative control over the region of modern northern and central Vietnam, which lasted from the Han dynasty through the Tang dynasty, is the best-documented type of contact. The other two means of language contact have significantly less documentation. Contact through commerce was certainly common throughout the history of Sino-Vietnamese relations; though, it was probably not continuous enough to result in widespread language shift. Finally, language contact came through the periodic influx of Chinese immigrants into Vietnam who were assimilated into the old Viets or Lac Lords culture.

This latter aspect may be the most important one in exploring the effect of Chinese on Vietnamese since it provided a means of linguistic borrowing through direct social contact and pressure. In Keith Taylor's *The Birth of Vietnam* (1983), it is shown how, within a few generations in modern day northern Vietnam, the early immigrant Chinese largely abandoned their loyalty to the rulers to the north and created what Taylor calls the 'Sino-Vietnamese families', a wealthy and influential social class. These Sino-Vietnamese families may have helped introduce and maintain some parts of Chinese culture and accompanying vocabulary, but it can be assumed that they shifted to the language spoken in that region (an early ancestor of modern Vietnamese) with little overall effect on that language as they adopted the local lifestyle. In terms of language contact, then, Chinese influence on the language of old Viets was through linguistic borrowing. It has been hypothesized that the results of borrowing consist primarily of

lexical influence with some accompanying phonological influence (Thomason and Kaufmann 1988:39). Indeed, that simple statement appears to be the case for Chinese-Vietnamese language contact even two thousand years later, that is, lexical borrowing with some phonological influence. Unless widespread bilingualism was present over following centuries, a claim for which there is no evidence, the mostly illiterate population in that region would have had little direct contact with Chinese, and the structural linguistic changes probably occurred very slowly mostly in the post-Tang era, after the era of Chinese political domination.

What was the language of the Vietnamese ancestors like two thousand years ago at the beginning contact with Chinese? Based on core vocabulary, the Vietnamese language is clearly, as are other Vietic languages (a term coined by Hayes in 1984),<sup>3</sup> a Mon-Khmer language (Huffman 1977; Gage 1985).<sup>4</sup> Thus, the language in question was most likely bisyllabic and did not have a fully-developed tonal system, the prototypical Mon-Khmer structure. Over the past two millennia, Vietnamese<sup>5</sup> has borrowed extensively from the Chinese lexicon, much more so than neighboring languages in the modern mainland Southeast Asian region, such as Thai and Cambodian. However, despite assumptions that significant changes in Vietnamese linguistic structure—particularly its phonology—are due to contact with Chinese, most linguistic characteristics of 20<sup>th</sup> century Vietnamese can be viewed as the result of ordinary language internal changes seen among languages in Southeast Asia, changes in which Chinese played a smaller part than has been previously thought. Some structural changes that have given Vietnamese its Chinese-like appearance are actually quite recent, developing only in the past several centuries without direct Chinese political domination.

In this paper, Vietnamese phonology, morphology, and syntax are shown to have primarily Southeast Asian (and in some cases specifically Mon-Khmer) typological traits, despite heavy lexical borrowing from Chinese. Chinese influence on Vietnamese linguistic structure is shown to be minimal, and much of it, relatively recent (within the last several centuries).

## 2 The Vietnamese lexicon

As noted in the previous section, the core vocabulary of Vietnamese is Mon-Khmer in origin. Numerals, body parts, unmarked terms for natural phenomena (e.g., dog, bird, root, leaf), a few personal and demonstrative pronouns are, with few exceptions, Mon-Khmer etyma. In addition to Mon-Khmer vocabulary, Vietnamese has been said to have a significant amount of Tai loans (Maspero 1912; Vương; Hoàng T. C. 1998) and possibly some vocabulary connected with Austronesian languages (Gregerson 1992). Finally, Vietnamese has two layers of Chinese vocabulary: the commonly studied Sino-Vietnamese layer, borrowed as

literary readings during the Tang dynasty, and Old Sino-Vietnamese vocabulary<sup>6</sup> that was borrowed before the Tang dynasty (Wang 1958; Lê 1959; Tryon 1979), probably during the Han dynasty and the rise of the Sino-Vietnamese families. It is in part due to this mixture of etymological sources that Vietnamese has been claimed by some scholars to be a language for which no single parent source can be posited. In this paper, Vietnamese is not considered to be a 'mixed' language since its core vocabulary is solidly Mon-Khmer (Huffman 1977) and comparison with more conservative Vietic languages has already clearly demonstrated that Vietnamese is Mon-Khmer (Nguyễn V. L. 1995; Alves to appear).

How much of the Vietnamese lexicon is Chinese? Some estimates put the number at 60 percent (Nguyễn D. H. 1961). Other Vietnamese linguists have put that number even higher (in personal communication). However, considering the nature of dictionaries, on which these figures are based, literary entries necessarily constitute a disproportional amount, thereby misrepresenting the amount of words actually used in the Vietnamese vernacular. In daily speech, which is a clearer indicator of genetic origin and language contact, the percentage of Chinese loanwords is significantly lower. One part of the Vietnamese lexicon includes reduplicant forms, which constitute a notable part of the vernacular and which are non-Chinese in origin. There are 5,000 entries in the recently published dictionary of Vietnamese reduplicants (Viện Ngôn Ngữ Học 1995),<sup>7</sup> which, compared to a regular dictionary of 50,000 entries, is a significant amount. Basic vocabulary may constitute a smaller portion of a dictionary than literary vocabulary, but it will be a larger portion of the spoken language. Hence, for purposes of understanding language contact at the colloquial level, the Chinese element, though still prominent, is not nearly as high when the literary language—a portion not directly available to non-literate Vietnamese in previous centuries—is excluded from statistics.

What lexical areas of spoken Vietnamese has Chinese influenced? The two areas, content words (non-grammaticalized) and function words (grammaticalized vocabulary), are discussed in the following two subsections. The focus is on vernacular Vietnamese.

## 2.1 Content words

Many Vietnamese content words that are similar to words in modern Thai and Khmer are actually Chinese in origin, as shown in extensive lists in a 1973 article by Pou and Jenner. Words related to trade, such as 'pastry' and 'bag/packet', are seen in many languages throughout the region of modern mainland Southeast Asia and are better seen as regional loans that were not necessarily borrowed directly from Chinese.<sup>8</sup>

Many other Chinese words that were not spread through Southeast Asia present in Vietnamese are those forms related to administrative

matters and cultural traits. Many of those words come from the Han era, the Old-Sino-Vietnamese (OSV hereafter) layer, all of which have lexical counterparts in the later Tang dynasty Sino-Vietnamese loans (SV hereafter). Taylor (Ibid.) described a situation in which the Han Chinese imposed marriage, education, and certain farming practices on the Vietnamese. Loanwords from that period still remain. Examples of OSV administrative vocabulary are ghi ‘record’, tuổi ‘years old’, and họ ‘family name’. Beyond a few basic familial terms that are Mon-Khmer in origin, Benedict (1947) noted that most Vietnamese kinship terms are Chinese and noted the two layers of Sino-Vietnamese, the OSV and SV loans. However, those borrowed familial terms have very different grammatical functions than do the forms in Chinese, functioning like pronouns, as discussed in section 5.1.

Examples of OSV words related to marriage include gả ‘give a woman in marriage’, góa ‘widow’, chồng ‘husband’, and vợ ‘wife’. Vietnamese vocabulary related to studying are generally Chinese loans (e.g., giấy ‘paper’, bút ‘writing utensil’, and tủ ‘bookshelf’ are OSV, while sách ‘book’ and học ‘study’ are SV). Most of the OSV words were borrowed again around the Tang dynasty era, but the OSV forms have remained the winners in the competition for daily use, while the later Sino-Vietnamese vocabulary remained part of the literary language only. Only in the 20<sup>th</sup> century, with the rise of widespread literacy and the spread of East Asian modern vocabulary (see Sinh 1993), has Chinese taken a more prominent role in the Vietnamese lexicon.

## 2.2 Function words

Vietnamese function words of Chinese origin consist of two types, (a) those that were grammatical in Chinese and have maintained their Chinese semantic functions and patterns of syntactic distribution, and (b) those that were non-grammatical in Chinese but have been grammaticalized after entering Vietnamese (see sections 5.1 and 5.2 for related discussion). Regardless, most of the Sino-Vietnamese loans have changed both syntactically and semantically, often appearing in bisyllabic words with indigenous Vietnamese forms.

Sino-Vietnamese words that have maintained roughly the same semantic functions and syntactic distribution as in Chinese include some measure words and clause-linking words. In Table 1, underlined portions of the clause-linking words are of Chinese origin while the others are non-Chinese. Mandarin readings of the words of Chinese origin are given for reference. Their status as Chinese words is verifiable through their written form (for each a Chinese character exists) and their phonological correspondences.<sup>9</sup>