Sino-Vietnamese Grammatical Vocabulary and Triggers for Grammaticalization[1]

Mark J. Alves
Montgomery College

A significant amount of the grammatical vocabulary in Vietnamese is of Chinese origins, despite the fact that much of it was either not originally grammatical in Chinese or went in semantico-syntactic directions not seen in Chinese. This fact can be accounted for by expected issues of grammaticalization, such as typologically common tendencies of grammaticalization (both regional and more universal) and semantico-syntactic specialization following the development of social hierarchical system. However, in addition, Sino-Vietnamese vocabulary has a special status of social and academic prestige, which may be another significant trigger for the grammaticalization of these terms. This paper looks at grammatical words in several categories, including expressions of quantity, intensification/comparison, location/direction, preverbal modal functions, and personal reference. Support for claims made come from texts from the 15th, 17th and 19th centuries.

1.   MAIN ISSUES

Of the grammatical vocabulary of Vietnamese, a sizeable number (perhaps four dozen measure words and a few dozen more cumulatively in other grammatical categories) are of Chinese[2] origins. Such vocabulary constitutes three types: (a) those words borrowed from Chinese as fully-developed grammatical vocabulary, (b) words which had some grammatical functions in Chinese but traveled down their own path of grammaticalization after entering Vietnamese, and (c) words which had not been grammaticalized in Chinese but which became grammatical at some point after entering Vietnamese. Such a scenario begs the question of why Chinese vocabulary in particular (as opposed to other etymological strata, such as Mon-Khmer or Tai-Kadai) has provided Vietnamese with such a seeming disproportionate amount of its grammatical vocabulary, some of which was not originally grammatical in Chinese.

Three key aspects provide some possible answers. First, cross-linguistically, certain properties of words lend themselves toward grammaticalization (e.g., words meaning "true" grammaticalizing into intensifying words (e.g., English "very/truly"). Second, some words may come to fit within discourse systems (e.g., grammatically specialized terms of address with familial term origins, a phenomenon commonly seen among languages of Southeast Asia). Finally, the prestige of Chinese vocabulary in various facets of Vietnamese society may have also been a factor in this overall process. The socially marked nature of that vocabulary was most likely a trigger in at least some of the situations discussed in this paper.

The content of this paper is explanation of some of the basic concepts of this study (Sino-Vietnamese vocabulary, language contact, and grammaticalization) and then to discuss examples of Sino-Vietnamese grammatical words of a few grammatical categories (quantity and measure words, intensifying and comparative words, locational and directional words, preverbal modal words, socially-conditioned terms of address and reference, and clause-linking words. It must be noted that the exact syntactic and semantic details of many of these

---

[1] I must thank David Branner and Jerold Edmondson for their comments on this paper. Marc Miyake generously tested the historical phonological validity of all of the claims regarding words with less well-established origins and provided useful thoughts on various points in this paper. Any errors are naturally mine alone. Comments can be sent via e-mail to malves98@yahoo.com.

[2] It is best to note at the outset that the term "Chinese" is entirely ambiguous since it can represent variously (1) all varieties of Chinese, (2) written Chinese, or (3) some specific variety of Chinese or historical period of the Chinese language, depending on the perspective of the speaker. This author attempts to be clear about varieties of Chinese by mentioning specific varieties or uses the term generally when it does not interfere with overall clarity or the direction of a specific argument. Certainly, points of contention regarding the term's usage will likely remain.

words are certainly deserving of further research and debate, and so the very brief discussion in this article can only touch minimally on such issues but hopefully not in a way that interferes with the presentation of the material.

1.1. SINO-VIETNAMESE VOCABULARY

The Vietnamese term *tự Hán-Việt* "Sino-Vietnamese words" can be applied generally to all Chinese loanwords in Vietnamese. However, more precisely, Sino-Vietnamese vocabulary is considered to be vocabulary borrowed by the Vietnamese during the Tang dynasty (roughly the 7th to 10th centuries), primarily through the spread of literacy in Chinese by language schools in (and possibly outside but near) Vietnam. It was quite possibly a now extinct southern Chinese literary reading standard, a kind of Southern Chinese koine (Hashimoto 1978:9). Hence, this definition of Sino-Vietnamese refers to the readings of Chinese characters by Vietnamese speakers learned in an academic setting, as opposed to non-literary transmission. Such words are easily identified and confirmed as Chinese in origin through Sino-Vietnamese dictionaries.

Two other subcategories of Sino-Vietnamese vocabulary exist, though their origins in Chinese are somewhat controversial and uncertain as these words tend to have more phonological irregularities, and chance similarity cannot always be ruled out. The first of these, Old Sino-Vietnamese vocabulary, are the words that presumably entered Vietnamese early in contact with China, possibly as early as the Han dynasty but certainly before the Tang dynasty and the Middle Chinese era.[3] This is also the period in which the first substantial number of Chinese immigrants came to Vietnam, leaving ancestors whose social prestige remained significant in following centuries, the so-called Sino-Vietnamese families (Taylor 1983). Thus, Old Sino-Vietnamese words are essentially loanwords that came through spoken contact with Chinese. Another category of colloquial forms are the so-called 'nativized' Sino-Vietnamese, generally, words that have phonetically similar counterparts in Tang-Song dynasty Sino-Vietnamese but with somewhat unpredictable phonetic variation.[4] Whether their primary means of transmission was spoken or written is less clear, but regardless, they have lengthy status as spoken Vietnamese, as much as several centuries.

The grammatical vocabulary discussed in this paper comes from all three categories (Sino-Vietnamese, Old Sino-Vietnamese, and nativized Sino-Vietnamese), suggesting that this vocabulary in some way, regardless of time of entry, developed special status whether or not native Vietnamese speakers were aware of the origins of the words. The primary sources of written materials used in this paper are (1) 15th century poems of Nguyễn Trãi (taken from Nguyễn Đình Hòa 1985), (2) the 17th century Vietnamese-Latin dictionary by Alexandre de Rhodes, and (3) the 19th century story *Truyện Kiều* by Nguyễn Du. Throughout the discussion of specific Sino-Vietnamese grammatical words, suggestions for which the above-mentioned categories they belong to are made.[5]

---

[3] Some important works on the lexical and phonological characteristics of Old-Sino-Vietnamese are those of Maspero 1912, Wang 1948, Mei 1970, Đào 1979, Tryon 1979, and Pulleyblank 1981.

[4] Similarly, Sino-Vietnamese dictionaries list doublets, phonetic variations of the same character, which have no particular phonological generalizations. I know of no simple way to determine whether these variants are the result of mistakes or regional varieties of Chinese as the source of transmission.

[5] Sino-Vietnamese vocabulary not dealt with in this article include words that have undergone little or no substantial change from the original Chinese form and those that are simply too speculative in terms of phonological shape and/or meaning. Words in the former category include adverbs such as *chỉ* (只) "only", *hiện tại* (現在) "presently", and *tuy nhiên* (雖然) "although". The word *sở dĩ* (所以) "The reason why…" is different from the Chinese meaning "therefore". While *nhưng* (仍) "but" is daily in Vietnamese but more classical in Chinese, it has still not changed significantly. While those words have solid standing as Sino-Vietnamese forms that can be confirmed in dictionaries, there are words that should be mentioned but only with hesitation as their phonological forms are in question in some cases. They may represent earlier or nativized borrowings, but they

1.2. GRAMMATICALIZATION

The literature on grammaticalization is substantial and need not be discussed in great depth here. Grammaticalization is, at minimum, a subset of semantic changes that result in more specialized functions of words and morphs within the grammar of a language. Both basic and functional vocabulary can be the source of grammatical vocabulary. Indeed, once grammaticalized, vocabulary often continues these semantico-syntactic changes and increased specialization (see Hopper and Traugott 1993). The most important and simultaneously difficult question to answer is why such changes take place. Analogical extension and reanalysis may be the routes of change, but the impetus for changes tends to focus on or at least touch on pragmatic needs. Thus, for example, in Vietnamese, the development of a system of terms of address and reference, and the associated grammaticalization of words within that system, is entirely related to the hierarchical system of a Southeast Asian culture with Confucian values. The use of family terms with inherent semantic clarity of hierarchical relations and associated pragmatic communicative needs makes them prime targets for grammaticalization.

A useful concept in dealing with grammaticalization is the cline, essentially a path of semantico-syntactic change. Typological tendencies and universality among languages of grammaticalization clines are used in this paper as support to suggest clear links between the source words/meanings and the grammatical words discussed. Clines noted in this paper follow the notation of Heine and Kuteva 2002 (abbreviated as H&K elsewhere in this paper), using all capital letters for the categories, for example, ALL > PLURAL, the cline for *các* in §2. A good portion of grammaticalized vocabulary in Vietnamese follows not only clines seen in other Southeast Asian languages (Matisoff 1991) but also more universal categories of such change (as listed in H&K Ibid.).

1.3. LANGUAGE CONTACT AND MEANS OF TRANSMISSION

Language can only be spread in a substantial way as Chinese was into Vietnamese by some kind of long-term and intense contact. Three clear possible situations emerge which differ from loanword to loanword. First, some Sino-Vietnamese words entered Vietnamese by direct spoken contact. Next, such words were spread by native Vietnamese speakers who were literate in Chinese and spread that vocabulary without Chinese speakers' participation. Finally, a combination of these two situations explains the borrowing of these words. Determining which factor is most appropriate has to be dealt with on a word-to-word basis but in many cases simply cannot be ascertained.

When the means of transmission is speech contact, namely, when speakers of Vietnamese interacted with Chinese speakers, the crucial detail is which varieties of Chinese were spoken. In fact, this makes the whole situation much more complex than this paper can account for, but as a general rule, the assumed types of Chinese with which the Vietnamese were in primary contact with are those spoken in bordering parts of Southern China in

---

could just as well come from other sources and simply be look-alike forms. Words in this category include *bởi* "because" (cf. Chinese *wèi* (為) with phonological reconstruction conflicts), *để* (cf. Chinese *dǐ* (底) "bottom"), *đã* (cf. Chinese *yǐ* (已) but most likely derived from the basic Vietnamese mean "to satisfy"), *đừng* "do not" (probably not a nativized form of Sino-Vietnamese *đình* (停) "cease" but rather possibly a grammaticalized form of native Vietnamese *đứng* "to stand"), *lại* "to come" and "again" (cf. Chinese *lái* (來), as there are both semantic and phonological problems), *ở* "to be at" "to live at" (cf. classical Chinese *yú* (於)), *thấy* "to see" (cf. colloquial Cantonese *tái*), and *trong* (cf. Chinese *zhōng* (中)), Sino-Vietnamese *trung*, but also likely a Mon-Khmer cognate, cf. Pacoh *kallúng* inside). A few of the forms listed in this article may have some questionable details, but in general, the author has tried to be conservative in positing what can be reasonably supported as Chinese in origin.

Guangdong and Guangxi provinces. Thus, it is best to consider Chinese dialect groups such as Yue (of which Cantonese is most well known), Pinghua, or possibly Hakka/Kejiahua, all of which likely have elements of the southern Chinese koine.[6]

Finally, it should be kept in mind that, regardless of era of contact, Chinese culture and language has maintained a status of education, wealth, and power in Vietnam for two thousand years. Chinese was borrowed through learning of the Chinese literary tradition; Chinese vocabulary entered Vietnam through trade and the wealth of Sino-Vietnamese families; Chinese words came through force of military and political power. Hence, a majority of Sino-Vietnamese vocabulary entered carrying a certain prestige, marking it off from the indigenous Mon-Khmer stratum or other strata of Vietnamese vocabulary. The next sections identify and discuss that vocabulary.

2.   QUANTITY AND MEASURE WORDS

The highly developed system of Vietnamese classifiers and measure words has accepted a wide range of words of Chinese origin. The word for "10,000" was borrowed twice (i.e., Old-Sino-Vietnamese *muôn* and Sino-Vietnamese *vạn* (萬)), and the Sino-Vietnamese word *tư* (四) "four"[7] has taken on special functions in the Vietnamese calendar and the naming of children in families. Of a list of 150 classifiers and measure words listed by Nguyễn Đình Hòa (1957), about three dozen can be identified as Chinese in origin (Alves 2001). These most likely represent a combination of both oral and written transmission from Chinese to Vietnamese. In addition, the prenominal position of these words in noun phrases, paralleling that of Chinese but contrasting with Tai languages and Mon-Khmer languages outside of Vietnam, also suggests that trade contact with Chinese has led both to lexical and accompanying syntactic patterns of borrowing. Two grammatical words used in quantity expressions, the pluralizing *các* and the classifier *cái*, are discussed below.

2.1. CÁC

The pluralizing *các* currently has taken on the meaning "the various", in contrast with the original Chinese word 各 (Mandarin *gè*), meaning, "each/every". In Nguyễn Trãi's writings, the pluralizer used was *phô*, no longer in modern Vietnamese. In de Rhodes' dictionary and *Truyện Kiều*, *các* still bore the meaning "all". Taken together, this suggests that the word remained a prenominal pluralizer from the time of its borrowing, but since the 17th century, it has shifted from a distributive or inclusive meaning to a more generalized plurality with the addition of definiteness. H&K (2002:36) note the grammaticalization cline ALL > PLURAL in several languages, to which Vietnamese may be added.

In a similar kind of grammaticalization, the Sino-Vietnamese word *số* (數) "number" has come to be used in the phrase *một số* "a number of", again, a grammaticalized step away from the original non-grammatical Chinese usage.

2.2. CÁI

A word of somewhat less certain historical status is the highly grammaticalized classifier *cái*, which itself appears to be in a class on its own outside of other classifiers, being able to precede even other classifiers (Nguyễn Tài Cẩn 1975:239-250). A folk etymology for this

---

[6] Mandarin, having developed significant status relatively recently, could not have been a solid source of influence before a couple of centuries ago; indeed, Mandarin itself is considered something of a "baby" in its part in Chinese linguistic history. Such a statement is still simplifying the matter since no language springs from nowhere but rather is simply the ancestor of an earlier language. Any apparent similarities between functions in Vietnamese and Mandarin are not likely due to any kind of language contact between Vietnamese and Mandarin, and any attempt to support hypotheses of borrowings solely from Mandarin data is weak at best.

[7] The actual Sino-Vietnamese form is *tứ*, but the level-tone variant may be a nativized form.

form is the homophonous adjective *cái* "female", but such a semantic shift is difficult support. One reasonable possibility is that it comes from the Sino-Vietnamese *cá* (個), which is similarly a "super-classifier" in many varieties of Chinese, including the Yue dialects spoken in Chinese provinces bordering Vietnam. The phonetic variation from *cá* could be explained by a parallel development in some parts of Yue territory (cf., Toisanese /kwaj/) and indeed supports the possibility of spoken language as the means of transmission into Vietnamese. This word has taken on special grammatical functions in Cantonese not present in Vietnamese,[8] but in fact, both Cantonese and Vietnamese have noun phrases with the said etymology in phrase initial position (without a preceding numeral quantifier) and similar definiteness indicated, a fact generally true of classifiers in both languages.

The process may have begun with the borrowing of the unit classifier through contact with Chinese speakers. Eventually, however, the noun-like functions of *cái* emerged in constructions such as *cái này* "this thing". *Cái* began to precede other classifiers, taking on a virtual article-like function, as it had by the 19th century as seen in *Truyện Kiều*.

cái điều bạc mệnh có chừa ai đâu
the item misfortune have avoid whoever question
1. "Never does misfortune miss anyone". (*Truyện Kiều*, line 108)

Clearly, *cái* presents a solid example of grammatical change and increasing grammatical expansion and specialization.

3. INTENSIFYING AND EQUATIVE WORDS

Basic comparison in Vietnamese is made with a native word *hơn* in X+*descriptor*+*hơn*+Y constructions, and intensification of descriptive stative verbs is made with preverbal *rất* and postverbal *lắm*. However, there is a solid handful of Sino-Vietnamese words in this grammatical category, including the intensifying *quá* and *thật*, the superlative *nhất*, and the equative *bằng*, *giống*, and *như*.

3.1. QUÁ

The Sino-Vietnamese word *quá* (過) "to cross" has a phonetic doublet form *qua*, which has remained the basic verb of movement in Vietnamese (and indeed itself is a kind of grammaticalized directional marker (e.g., *bay qua* (fly-cross) "to fly over") and past time (e.g., *hôm qua* (day-past) "yesterday"). In Chinese, this etymon does have some special functions apparently through the overlapping senses of "passing" and "exceeding" as seen in both Mandarin, *tài gùo fèn* "going too far (in one's behavior)" and more relevantly, Cantonese, which has the post-descriptor adverb *gwo taùh* "exceedingly".[9] Determining whether this indicates borrowing from Cantonese or another Southern variety of Chinese would require evidence not yet available, but it is also quite possible that this is a language-internal innovation in both Vietnamese and Cantonese.

Regardless, the direction of change from verb to adverb is typologically feasible, and the phenomenon of words having the meaning "to exceed" as the basis for comparison is common (see H&K Ibid.:126-127, who list the cline EXCEED > ELATIVE with examples from several languages). Thus, the original verb *quá*, different phonetically from *qua*, came to be reanalyzed for as an adverb within a close analogical range ("pass over" to "overly"). In *Truyện Kiều*, *quá* was still used mainly as a verb, so this change is relatively recent.

---

[8] In Cantonese, classifiers and possessive marking are connected in general.

[9] It is significant to note, though, that the more universal Chinese intensifier, seen in many varieties of Chinese, 好 (Mandarin *hǎo*, Cantonese *hóu*) GOOD > VERY, does not appear in Vietnamese.

3.2. THẬT

The original meaning of the nativized Sino-Vietnamese *thật* "true" is a common source of intensification among the world's languages (see H&K 2002:302, which lists the cline TRUE > INTENSIFIER), making *thật* predisposed towards such a development. Interestingly, the Sino-Vietnamese proper counterparts, *thiệt* and *thực* (實), did not grammaticalize in this way, nor did *phải* "true", a native Vietnamese word. It is possible that *thật*, a word with more perceived formality or abstraction (i.e., "truth" as a philosophical entity), was more likely to become grammaticalized.

While the word in Chinese has been used in expressions to identify actual situations ("in truth" or "actually", as it does as well in Vietnamese, *thật ra* "actually"), in Vietnamese, the word has taken on an intensifying function, expressing "truly/very" (see §2, sentence 1). De Rhode's dictionary only lists the basic, non-grammatical meaning "true", and in *Truyện Kiều*, its adverbial function is seen strictly before main verbs, not adjectives, suggesting that this is a late-19th or perhaps post-19th century development.

In Chinese, the etymon 實 (Mandarin *shí*) does appear in lexical compounds with intensifying adverb functions (*shí zài* 實在 "really" (Sino-Vietnamese *thực tại*) and *quē shí* 缺實 "really"). In contrast, though, *thật* functions alone as a free morpheme and has gone through some additional semantic specialization, appearing before or after descriptors or sentence finally with slight semantic differences (as described in the entry for *thật* in the Từ Điển Tiếng Việt (1996)).

3.3. NHẤT

The Sino-Vietnamese numeral *nhất* (一) "one" has come to express the superlative in Vietnamese, following adjectival descriptors (e.g., *to nhất* (big-most) "biggest") and certain mood verbs (e.g., *thích nhất* (like-most) "to like the most"). Though H&K (Ibid.) do not list a grammaticalization cline ONE > SUPERLATIVE, the semantic overlap is clear since "one" expresses extremity.[10] De Rhode's dictionary lists *nhất* as *nhít*, an understandable alternate considering phonologically parallel Sino-Vietnamese pairs such as *bệnh*:*bịnh* "ill" and *chính*:*chánh* "main". Its first gloss in de Rhodes' dictionary is "one", but it is also listed as with the meaning "primary" and "best". It is not used in *Truyện Kiều* as a superlative. Apparently, it was after the start of the 19th century that *nhất* began to express the superlative with a range of adjectives, good or bad, a good example of reduced semantic specifications and widened grammatical function.

3.4. BẰNG

With some trepidation regarding the phonetics, it can be posited that *bằng* is a nativized Sino-Vietnamese grammatical word.[11] The general meaning of the proposed Chinese character 平 (Sino-Vietnamese proper *bình*, Mandarin *píng*) is "fair" or "equal". Originally a kind of stative verb in Chinese, *bằng* has attained special status in Vietnamese as a prepositional marker of equality, entirely unlike Chinese usage. Shift from verb to preposition is a common direction of grammaticalization, and the shared semantics are clear. In the 15th century, Nguyễn Trãi used *bằng* with equative meanings, suggesting an early development of that function in Vietnamese. By the 19th century, it had developed a wider grammatical range of functions, as seen in *Truyện Kiều*.

---

[10] Similarly, the meaning "most" in Japanese is expressed by Sino-Japanese *ichiban* in which *ichi* is Sino-Japanese (一) "one".

[11] An actual Sino-Vietnamese word *bằng* is the Chinese word *píng* "rely on" (憑), which serves as only incidental support at best for considering *bình* and *bằng* as related.

Bằng has come to take descriptive stative verbs before and nouns after to indicate equality between the items compared. A less certain but still possible development is the expansion from "equal" to "composed of" (e.g., *Cái này làm bằng gì* (thing-this-make-of-what) "What is this made of?") and from there to instrumental "by means of" (e.g., *Bàn này làm bằng tay* (table-this-make-by-hand) "This table was made by hand".).

In sum, it is proposed here that *bằng* has moved from a stative verb meaning "equal", to a preposition indicating equality in comparative constructions, and then to prepositions expressing means, although this last shift is a hypothesis that requires more evidence to confirm.

3.5. GIỐNG

The form *giống* is likely an early colloquial borrowing of Sino-Vietnamese proper *chủng* (種) "species/type". Similarly, in varieties of Chinese, the character 樣 (Mandarin *yàng*, Sino-Vietnamese *dường* "semblance") is used in the word *yī yàng* "the same". Interestingly, the nativized colloquial form is the grammaticalized element, not the more literary form. De Rhodes' dictionary lists the stative verb meaning, and so perhaps this usage started before the 17th century. In this situation, a noun has grammaticalized to a verb, not an uncommon sort of derivation. The cline TYPE > SAME is not listed in H&K, though somewhat parallel is the cline MANNER > SIMILE (Ibid. 210).

3.6. NHƯ

The word *như* (如) "as/like" is uncontroversially a Sino-Vietnamese word proper. In fact, in general, the word has maintained the same meaning and overall similar function as in Chinese. In *Truyện Kiều*, it was used several dozen times, demonstrating its solid position in Vietnamese by the 19th century. Vietnamese, however, has given it one development not seen in Chinese; it can be followed with items as examples, effectively expressing "such as", similar to the use of "like" in English. A cline can be posited for this semantico-syntactic expansion, namely, SIMILAR > EXEMPLIFY.

4. LOCATIONAL AND DIRECTIONAL WORDS

Locational and directional words discussed here include *tại* "at", *gần* "near", *bên* "side", and *ngoài* "outside".

4.1. TẠI

The somewhat literary *tại* (在) "to be at" is uncontroversial in its status as a Sino-Vietnamese word. In Nguyễn Trãi's poetry, the synonymous *ở* "to be at" was the primary locative verb used. [12] It primarily functions as a locational preposition and was listed (with *ở* co-referenced) in de Rhodes' 17th century dictionary. In *Truyện Kiều*, it was used fourteen times, mostly with the meaning "because of". In that story, *tại* was generally followed by nouns indicating humans (e.g., *người* "people", *ta* "we", *ai* "someone", and *mình* "oneself"). Interestingly, *tại* is not used in spoken Cantonese but is used in Mandarin as well as written Chinese texts. This suggests that this is a purely literary borrowing, one that did not move into Vietnamese through spoken interchanges with Chinese speakers.

It is a word that continues to grammaticalize in Chinese (cf., the same word *zài* in Mandarin indicates progressive action) and does so as well in Vietnamese but in a different way. It has developed a meaning related to cause, as in *tại sao* "why" and can be used alone

---

[12] The word *ở* itself is cognate with the verb "to reside at", which illustrates a common Southeast Asian grammaticalization pattern (see Matisoff 1991:414-418).

to mean "because" (e.g., *tôi nghĩ tại tôi mệt* (I-rest-because-I-tired) "I'm resting because I'm tired"). This is an example of the cline that H&K (Ibid. 200) posited, LOCATIVE > CAUSE.

4.2. GẦN

The form *gần* "near" is likely a nativized form of the written Sino-Vietnamese *cận* (近) "near". The presence of the *huyền* tone suggests the possibility of borrowing through contact with Chinese speakers (a case of phonetic uncertainty), though the absence of a colloquial Sino-Vietnamese word meaning "far" is a problem in maintaining such a position, and *gần* still could have been spread from writings as a nativized form of literary readings. Its basic meaning and general syntactic distribution is no different from that of Chinese, though it does, in Chinese, tend to be preposed with locational prepositions when it is used as a stative verb.

In addition to the general locational meaning, *gần* has analogical meanings in the semantic sense of extent, meaning "nearly" or "almost", in contrast with Chinese usage. This development may have occurred after *Truyện Kiều*, which used the word in three dozen instances, all with either the stative verb or locational function, not the adverbial usage. In Vietnamese, *gần* can be used with numbers (e.g., *gần 400 trang* (nearly-400-page) "almost 400 pages") or verbs or descriptors of completion (e.g., *gần xong* "almost done"). Matching this category, H&K (Ibid. 214-215) list the common cline NEAR > PROXIMATIVE. The dictionary *Từ Điển Tiếng Việt* (Trung Tâm Từ Điển Học 1996) lists several senses of this word, demonstrating its substantial semantic range and inherent semantico-syntactic flexibility.

4.3. BÊN

The locational noun *bên* "side" is a nativized version of the Sino-Vietnamese proper form *biên* (遍) "side" and also serves as a locational relator noun with a more abstract interpretation (i.e., grammatical locus). In Vietnamese, *bên* has generally maintained both the same core and grammaticalized meanings as in written Chinese, but the word is used within a strictly Vietnamese syntactic framework, even when both terms are originally Chinese (e.g., *bên ngoài* "outside" in which *ngoài* is also nativized Sino-Vietnamese). In *Truyện Kiều*, the approximately fifty instances of *bên* show it in a wide range of grammatical uses.

In Vietnamese, *bên* is followed by modifying and compounding elements (e.g., *bên đó* (location-that) "there") rather than being preceded by them, as in Chinese (e.g., *zài nà bīan* (at-that-location) "there")). It occurs in a variety of locational compounds such as *bên trong* "inside" and *bên trái* "left side". As for the means of transmission, *bên* could come from spoken language contact since spoken Cantonese does use the related *bīn*, suggesting a possible southern usage.[13]

4.4. NGOÀI

The word *ngoài* "outside" is a nativized form of Sino-Vietnamese proper *ngoại* (外). This pair is phonologically parallel to other nativized-original doublets. Consider the doublet pairs *loại* versus *loài* (類) "type", *tự* and *từ* (自) "from", and *dụng* versus *dùng* (用) "utilize", which in all three cases, the forms with the *nặng* tone are considered standard Sino-Vietnamese words while with those with the *huyền* tone are considered *Nôm* words, essentially nativized Sino-Vietnamese (also see *vì* in §7). Considering the phonological

---

[13] Cantonese also has a commonly used colloquial counterpart, *douh*, which phonetically resembles Vietnamese locational *đâu* "where", but the tone type (low versus mid) and the grammatical functions of these words differ too much to posit a connection.

patterning, and lack of such a pattern outside of Chinese, these most likely come from nativized literary readings rather than through direct speech contact.

As with *bên* "side" mentioned above, the syntax of this word strictly follows Vietnamese parameters. *Ngoài* is also used in the additive phrase *ngoài ra* "besides" in a way different from the formal Chinese construction (i.e, Mandarin *zhī wài* (之外)). Interestingly, while *ngoại* has it nativized form, its Sino-Vietnamese antonym *nội* "inside" does not (i.e., there is no form *nồi*).

5. PREVERBAL WORDS

Preverbal elements discussed here include passive markers (*được* and *bị*) and negation (*không* and *đừng*). The group of words that represent functions similar to European passive voice (including Sino-Vietnamese *do* (由)) deserves special mention. The development of these particles' passive-like function began largely at the end of the 19th century or early 20th century.[14] The speed at which the means to mark passive-like constructions[15] in the past century developed is surprising but nevertheless a phenomenon seen throughout China and Southeast Asia. Indeed, contact with the West can be seen as a partial catalyst for grammaticalization in these instances.

5.1. ĐƯỢC

The word *được* is most likely a nativized borrowing of standard Vietnamese *đắc* (得) and may have originally (before the development of a passive meaning) entered Vietnamese through speech contact as opposed to being a literary borrowing.[16] The phonological form of *được* shows that it was likely borrowed before a similar change led from /-a-/ to /-ượ-/ in words such as *đường* and *nước* (see Nguyễn Tài Cẩn 1978 for other relevant discussion). Indeed, the forms *đàng* and *nác* have been kept in some North-Central Vietnamese dialects.

The use of *được* both as a verb "obtain" and with an abilitative meaning was firmly established by the 15th century, as seen in Nguyễn Trãi's writings. In de Rhodes' dictionary, *được* expresses capability, end result of an action (e.g., *tìm được* translated into modern Vietnamese as *tìm thấy* (seek-see) "to find"), and possession,[17] thereby confirming the recentness of the development of a passive meaning. In *Truyện Kiều*, *được* is seen commonly (about forty instances), expressing variously ability, result, and, additionally, acceptability of a situation, but not the passive voice.

Được probably began with the abilitative meaning and possibly the resultative meaning. When it began to be used as a main verb meaning "okay" must have been by the 19th century. The grammaticalization of *được* away from the Chinese usage begins with the use of *được* to indicate passive voice in desirable situations. It further moved away by diverging semantically from *bị*, which implies an undesirable situation, as discussed next.

---

[14] For discussion about the spread of passive voice in Thai, see Prasithrasint 1985. Also, Sinh 1993 discusses the spread of Chinese throughout East Asia from Japan in the early 20th century, quite possibly the time for the spread of usage of the passive usage in Vietnamese.

[15] Passivization as seen in European languages does not always match the types of constructions seen in Vietnamese, hence the use of the term "passive-like".

[16] The highly conservative Vietic language Rục has, for Vietnamese *được*, the form *tược* with an unvoiced initial, as opposed to a voiced initial for the passive-marking *bị*, which suggests an earlier entry of *được* than *bị*, probably by some centuries, but the passive meaning probably came into use at near the same time.

[17] The Thai passive marker *thùuk* has been suggested as somehow being related to Vietnamese *được*, but connecting these two words in terms of language contact would require unreasonable phonological changes (e.g., /th/ versus /d/ and /u/ versus /ɨə/). Moreover, Chinese 得 (Mandarin *dé*) has additional semantic-syntactic functions in common with *được* but in contrast with Thai *thùuk*.

5.2. BỊ

    The preverbal word *bị*, indisputably of Chinese origins (被, Mandarin *bèi*),[18] overall has a negative connotation, that the result is undesirable. The following two sentences show the semantic distinction with *được*; whereas the sentence with *bị* is an unlucky situation, the sentence with *được* indicates that it is a positive situation for some reason. This usage of *bị* as passive is clearly absent in early references (*Truyện Kiều* shows zero instances).[19]

anh ấy	bị	chọn	anh ấy	được	chọn
he	(passive)	select	he	(passive)	select

2a. "He was (unluckily) selected".   2b. "He was (luckily) selected".

    However, in addition to its basic passive-like function, it has taken on other meanings, distributional patterns, such as being followed by stative verbs (e.g., *nó bị bệnh* (he-passive-sick) "He is sick"). Finally, *bị* has even recently entered popular colloquial speech in the emphatic *hơi bị* "rather" preverbal adverbial phrase. Thus, the word in just a century has taken on a grammatical life and direction of its own.

5.3. KHÔNG

    The negation word *không* is most likely cognate with the homophonous Sino-Vietnamese word *không* (空) "void". The typological evidence is solid as H&K (Ibid.:188) list the grammaticalization cline LACK > NEGATION. Nguyễn Trãi used the word as a verb meaning "empty of/lacking", the basic meaning of the verb even in de Rhodes' dictionary.

túi	đã	không	tiền	khôn	chác	rượu
pocket	already	empty (of)	money	difficult	buy	wine

3. "Lacking money, it's difficult to buy wine". (from *Nguyễn Trãi Toàn Tập*, in Nguyễn Đình Hòa 1985:464, my own translation.)

    Nguyễn Phú Phong (1996), in an article discussing the historical development and semantic properties of negation in Vietnamese, claims that *không* became grammaticalized at some point after European merchants came to Vietnam in the 17th century. If so, the original meaning of "empty" came analogically to mean "no/not", though as Nguyễn Phú Phong (1996) noted, it specifically negates existence rather than truth values, as the native Vietnamese negation word *chẳng* does (a word mostly supplanted by *không* in modern Vietnamese). This semantic specification follows logically from the semantics of a verb meaning essentially "lacking". It can be followed by verbs (e.g., *tôi không biết* (I-no-know) "I don't know") and sometimes by nouns (e.g., *anh ấy không tiền* (fellow-that-no-money) "He has no money"). Its complete grammaticalization came with its combining with *phải* "true" regularly in collocations such as *đó không phải là sách của tôi* (that-not-true-be-book-of-I) "That's not my book". It is currently seen in various grammatical phrases and gambits.

    Its usage in *Truyện Kiều* is substantial, appearing thirty-nine times and functioning variously as (1) a morpheme in a compound, (2) a main verb meaning "empty", (3) a noun-negating verb (i.e., "lacking X"), and (3) a final question particle. However, it is also

---

[18] Somewhat ironically, equivalent passive sentences in spoken Cantonese use the similar sounding syllable *béi* with a rising tone, but that word actually is derived from the homophonous verb meaning "to give", a common grammaticalization cline in Chinese, GIVE > PASSIVE. The standard, written word 被 (not spoken in Cantonese) is pronounced *beih* in Cantonese with a low tone.

[19] Nguyễn, Đình Hòa (1990:101) notes the listing in the dictionary of the phrase *bị phũ ba* and translates it as "run into a storm" while the 1991 Vietnamese translation lists *bão táp ở biển* "a storm in the sea" with no verb. Regardless, this is not a passive expression, though it does suggest misfortune.

significant to note that the negation word *chẳng* appeared one hundred times, and its related *chăng*, twenty-five times, demonstrating the earlier state of competition between these forms, but a competition in which *không* has largely won since.

6.   SOCIALLY-CONDITIONED TERMS OF ADDRESS

While personal pronouns of the world's languages typically fit into systems of person, number, and, to a lesser extent, gender, the daily Vietnamese system of Vietnamese terms of address and reference is bound to an explicitly organized social system. These terms are most typically derived from familial terms (e.g., *anh* "brother"[20] or *cô* (姑) "aunt"), though in fact, the overall system has extended to professional situations and other non-family relations (e.g., *thầy* "teacher" or *ông chủ* "boss".). A substantial number of the familial terms that have both pragmatic and grammatical specialization are Chinese in origin. Benedict (1947) posited that over a dozen of the family terms, especially those of older generations (e.g., *ông* (翁) "grandfather" and *bà* (婆) "grandmother"),[21] are Chinese in origins, though another portion is likely Mon-Khmer cognates (e.g., *con* "child" and *cháu* "grandchild").[22] In fact, a handful of the terms are nativized Sino-Vietnamese vocabulary, such as *cậu* "younger uncle" (standard Sino-Vietnamese *cữu* (舅)) and *chị* "older sister" (standard Sino-Vietnamese *tỉ* (姐)).[23]

How exactly these words came historically to be the primary means of address in Vietnamese is unclear. Some of the words may have come through personal interaction with Chinese immigrants or their ancestors in the post-Han era of the Sino-Vietnamese families (see Taylor 1983, chapter 3). Still, some may simply be high-level vocabulary brought from written texts alone. Regardless, these terms have taken a place of dominant usage over native Vietnamese pronouns, which have a relatively limited usage, used in intimate, very informal, or even confrontational situations.

The pattern of grammaticalization here is a shift from common nouns to nouns with pronominal reference. Each term keeps some of the semantic features of the original term, such as gender and age relative to the speaker or listener. These terms do not, however, inherently indicate person and have what are sometimes called "floating" reference, referring variously to first, second, or third person, depending upon the discourse context. Unlike typical personal pronouns, these words can combine with numbers and demonstrative pronouns to provide the otherwise missing plurality and deictic properties (e.g., *các cô đó* (the various-miss-that) "those ladies").

It is important to recognize the difference between a clause-external and clause-internal function (TOA indicates a pronominal term of address).

---

[20] The word *anh* "elder brother", unlike other Vietnamese familial terms, is of uncertain etymological origins. Despite its homophony with the Sino-Vietnamese *anh* (英) "a person of outstanding virtue", the semantic shift is still odd. Pinnow (1965) reconstructed a proto-Mon-Khmer 1st person pronoun **iŋ*, which again has substantial semantic distance from the Vietnamese meaning.

[21] In fact, the terms more generally mean "old man" and "old woman", and in Mandarin, there are not the terms for grandparents. Cantonese uses 婆 for "grandmother" but not 翁 for "grandfather".

[22] Benedict described the origins of Vietnamese familial terms as follows: "1. extensive nuclear material of undetermined affiliations; 2. core of nuclear Mon-Khmer elements; 3. scattering of Thai elements, clearly less nuclear than the above; 4. peripheral mass of Chinese lexical material" (1947:371). Published several years before Haudricourt's tonogenesis hypothesis, such a statement is significant. It describes well the situation of Vietnamese lexical origins overall, though with more Chinese in the core.

[23] These particular examples could in fact belong to the Old-Sino-Vietnamese substratum at time before the third category of tonal development (i.e., tones *hỏi* and *ngã*) had developed in Vietnamese, at a time when Vietnamese had the four-tone systems seen in Chứt languages, such as Rục and Arem (Nguyễn Văn Lợi 1991).

đi	đâu	đó	cô		cô	đi	đâu	đó
go	where	there	ma'am		TOA	go	where	there

4a. "Where are (you) going, ma'am?" 4b. "Where are you (ma'am) going?"

The first sentence contains a usage (i.e., a familial word used outside the main clause to clarify reference) common to many languages, and the use of familial terms in this function is seen in Chinese. However, of more interest is the second sentence in which the term has a fully developed pronominal usage as a subject within the sentence matrix, something not generally the case in Chinese speech.[24] It is this development of these terms that puts them out of a more pragmatic function into solid grammatical territory.

The nativized Sino-Vietnamese form *bác* (Sino-Vietnamese proper *bá* (伯)), the familial meaning of which is "uncle (father's older brother)", is an exceptional example of grammaticalization in that it has taken one more step beyond the other grammaticalized familial terms in becoming a gender-neutral term of respectful address. Table 1 shows the derivation in steps, changes of meaning and semantico-syntactic features (TOA=term of address). Arrows indicate changed lexical properties between derivations.

BÁC₁	BÁC₂	BÁC₃	BÁC₄
"uncle" →	TOA	TOA	TOA
older	older	older	older
noun	noun →	pronoun	pronoun
male	male	male →	gender-neutral

Table 1: Grammaticalization of "bác"

Bác began as a basic noun with the Chinese meaning "uncle". It came to have to a meaning of respect with older male reference, but possibly kept only a clause-external function. Next, it developed a pronominal function but still kept the features "older" and "male". Finally, it developed a gender-neutral function, thereby functioning as a general term of respectful address regardless of gender. The constant, other than phonetic form, was the use of the term with reference to one older than another, generally coinciding with required respect. It is that fourth step that is beyond the scope of other such terms in Vietnamese and is also a further step of grammaticalization, that is, removal of a semantic feature but an increased range of usage in the language in general.

7.   CLAUSE-LINKING WORDS

The last words discussed are words that connect clauses, including here the multi-functional *thì* and the cause-marking *vì*.

7.1. THÌ

The topic-comment linking *thì* is quite possibly derived from Sino-Vietnamese *thì* (時) "time", the phonetic doublet of which is *thời*. In Chinese, the cognate 时 (Mandarin *shí*) is a noun phrase head, appearing at the ends of clauses and creating a time clause (i.e, it means "When…"). The Vietnamese usage of *thì* is somewhat similar, and the position is similarly interclausal, but in fact, it developed a more generalized meaning and did not always having a purely temporal function. Nguyễn Trãi's poems contain sentences with *thì* with this topic-comment function, linking clause to clause. Maintaining its noun status would have been

---

[24] The use of terms such as "auntie" (cf. Mandarin *a-yí*) with floating reference is typically in informal speech between adults and children, whereas in Vietnamese, this system is used throughout Vietnamese society regardless of age or social situation.

problematic as Vietnamese noun phrase structure is strictly left-branching. True time words in Vietnamese must be clause initial (e.g., *khi* "when" as in *khi cô đói thì ăn đi* (when-miss-hungry-then-eat-(imperative)) "When you (miss) are hungry, go ahead and eat").

In *Truyện Kiều*, *thì* showed common usage, appearing 87 times with a variety of preceding elements, including clauses, single verbs, and nouns. In addition, it started sentences, showing its complete reanalysis from its presumed original noun status. In modern Vietnamese, *thì* is preceded variously by clauses, phrases, or just words (including agents, patients, locations, time, and instruments, among other semantico-syntactic categories), thereby decreasing its semantic features and increasing its usefulness in Vietnamese syntax (see Cao (1992) and Clark (1992)). These three general stages are shown in Table 2, which highlights the changed lexical features with arrows.

THÌ₁	THÌ₂	THÌ₃
noun →	particle	particle
"time/when" →	(linking word)	(linking word)
clause-clause	clause-clause →	ANY-clause

Table 2: Grammaticalization of "thì"

7.2. VÌ

The nativized *vì* "because/due to" is most likely cognate with Sino-Vietnamese proper *vị* (為) (see comments on parallel phonological relationships for ngoài in §4). It appears in de Rhodes' dictionary (spelled *uì*), and is commonly used in *Truyện Kiều*, in which there are thirty-five instances of *vì* followed most commonly by nouns but also, sometimes, complete clauses. Its general semantico-syntactic function has changed little from its Chinese origins, but its usage in lexical compounds of cause (e.g., *bởi vì*, *tại vì*, *vì sao*, etc.), compounds expressing effect (e.g., *vì vậy* and *vì thế*, meaning essentially "Due to the mentioned situation"), and even linked grammar words (e.g., *sở dĩ…là vì…* "the reason for…is that…"), is noteworthy. It clearly plays an active role in grammatical expressions, which leaves it in a position for other developments.

8.   RESTATEMENT AND FINAL CONSIDERATION

Again, the question was why so many words specifically of Chinese origin, many of which are nativized and not literary, have grammaticalized in Vietnamese. Three interesting facts stand out. First, many of the grammaticalized forms are the nativized forms while the literary counterparts did not grammaticalize (emphasizing the importance of actual speech usage in this process). Second, some of the Sino-Vietnamese grammaticalized words have virtually supplanted other native Vietnamese grammatical vocabulary (e.g., *không* over *chẳng* and *các* over *phô*). Finally, many of the forms discussed here either were not grammaticalized in Chinese or traveled down their own route of grammaticalization. The hypothesis here is that social prestige carried with the borrowing of that vocabulary, regardless of speakers' awareness of origins of the words, may have contributed to perceived special status of those words and that such attention to these words may have served as a kind of trigger interacting with the other two factors of pragmatic function and typological tendencies.

I leave the readers with a caveat. While the words presented in all cases have fairly solid or sometimes complete status as Chinese in origin, strong statements about direction of change, timing, and development were kept to a minimum since, in fact, more data would have to be collected from ancient Nôm writings to clarify and verify these issues. Indeed,

interested readers are encouraged to test and improve upon my statements of the timing of these instances of grammaticalization.

9.   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. by Graham Thurgood, 221-241.

Benedict, Paul K. 1947. "An Analysis of Annamese Kinship Terms". *Southwestern Journal of Anthropology* 3:371-390.

Cao, Xuan Hao. 1992. "Some Preliminaries to the Syntactic Analysis of the Vietnamese Sentence". *Mon-Khmer Studies* 20:137-151.

Đào, Duy Anh. 1979. *Chữ Nôm: Nguồn Gốc, Cấu Tạo, Diễn Biến* (Chu Nom: origins, formation, and transformations). Hà Nội: Nhà Xuất Bản Khoa Học Xã Hội.

De Rhodes, Alexandre. 1991 (originally 1651). *Từ Điển Annam-Lusitan-Latinh (Thường Gọi là Từ Điển Việt-Bồ-La)*. Ho Chi Minh City: Nhà Xuất Bản Khoa Học Xã Hội.

Hashimoto, Mantaro J. 1978. "Current Developments in Sino-Vietnamese Studies". *Journal of Chinese Linguistics* 6:1-26.

Heine, Bernd and Tania Kuteva. 2002. *World Lexicon of Grammaticalization*. New York: Cambridge University Press.

Hopper, Paul J. and Elizabeth Closs Traugott. 1993. *Grammaticalization. Cambridge Textbooks in Linguistics*. Cambridge University Press.

Lê, Văn Lý. 1948. *Le Parler Vietnamien*. Paris.

Clark, Marybeth. 1992. "Conjunction as Topicalizer in Vietnamese". *Mon-Khmer Studies* 20:91-109.

Maspero, Henri. 1912. "Études sur la Phonétique Historique de la Langue Annamite: Les Initiales". *Bulletin de l'École Françoise d'Extrême-Orient* 12:1-127.

Matisoff, James A. 1991. "Areal and Universal Dimensions of Grammatization in Lahu". ed. by Traugott and Heine, *Approaches to Grammaticalization*, Vol. 1: 383-453. Philadelphia: John Benjamins.

Matthews, Stephen and Virginia Yip. 1994. *Cantonese: A Comprehensive Grammar*. New York: Routledge Grammars.

Mei, Tsu-Lin. 1970. "Tones and Prosody in Middle Chinese and the Origin of the Rising Tone". *Harvard Journal of Asiatic Studies* 30: 86-110.

Nguyễn, Đình Hòa. 1966. *Vietnamese-English Dictionary*. Rutland, Vermont: Charles E. Tuttle Co.

Nguyễn, Đình Hòa. 1985. "Some Archaic Vietnamese in Nguyễn Trãi's Poems". ed. by Graham Thurgood et. al., *Linguistics of the Sino-Tibetan Area: The State of the Art*, 463-473. Canberra, Australia: Australian National University.

Nguyễn, Đình Hòa. 1990. "Seventeenth-Century Vietnamese Lexicon: Preliminary Gleanings from Alexandre de Rhodes' Writings". ed, by J.H.C.S. Davidson, *Austroasiatic Languages: Essays in Honour of H. L. Shorto*, 95-104.

Nguyễn, Phú Phong. 1996. "Negation in Vietnamese and in Some of the Viet-Muong Languages". *Pan-Asiatic Linguistics: Proceedings of the Fourth International Symposium on Languages and Linguistics, January 8-10, 1996*, vol. II:563-568. Thailand: Mahidol University at Salaya, Institute of Language and Culture for Rural Development.

Nguyễn, Tài Cẩn. 1975. *Từ Loại Danh Từ trong Tiếng Việt Hiện Đại* (Nouns in Modern Vietnamese). Hà Nội: Nhà Xuất Bản Khoa Học Xã Hội.

Nguyễn, Tài Cẩn. 1978. "Quá Trình Hình Thành Thế Đối Lập giữa Ba Từ 'Được, Bị, Phải' (Processes of Formation among the Three Words 'Được, Bị, Phải'). *Tạp Chí Ngôn Ngữ* (Journal of Linguistics) 1978.2:19-22.

Nguyễn, Tài Cẩn. 1979. *Nguồn Gốc và Quá Trình Hình Thành Cách Đọc Tiếng Hán Việt* (The Origins and Process of Development of Sino-Vietnamese Readings). Hà Nội: Nhà Xuất Bản Khoa Học Xã Hội.

Nguyễn, Văn Lợi. 1991. "Về Quá Trình Hình Sự Đối Lập Âm Vực Thanh Điệu trong các Ngôn Ngữ Việt Mường" (On the Process of the Distinctions of Tone Contour in the Viet-Muong Languages). *Tạp Chí Ngôn Ngữ* (Journal of Linguistics) 1991.1:49-59.

Nhà Xuất Bản Đại Học và Giáo Dục Chuyên Nghiệp. 1990. *Từ Điển Hán-Việt* (A Chinese-Vietnamese Dictionary). Hà Nội.

Pinnow, H.J. 1965. "Personal Pronouns in the Austroasiatic Languages: a Historical Study". *Linga* 14:3-42.

Prasithrathsint, Amara. 1985. *Change in the Passive Constructions in Written Thai during the Bangkok Period*. Unpublished dissertation. Honolulu, HI: University of Hawai'i at Mānoa.

Pulleyblank, Edwin G. 1981. "Some Notes on Chinese Historical Phonology". *Bulletin de l'ecole Françoise d'Extreme-Orient* 277-288.

Sinh, Vinh. 1993. "Chinese Characters as the Medium for Transmitting the Vocabulary of Modernization from Japan to Vietnam in Early 20th Century". *Asian Pacific Quarterly* 25.1:1-16.

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.

Trung Tâm Từ Điển Học. 1996. *Từ Điển Tiếng Việt*. Hà Nội: Nhà Xuất Bản Đà Nẵng.

Wang, Li. 1948. "Hanyueyu yanjiu" (Research on Sino-Vietnamese). *Lingnan Xuebao* 9.1.1-96. Reprinted 1958 in the *Hanyushi Lunwenji*, 290-406. Beijing Kexue Chubanshe.