

# THE CURRENT STATUS OF VIETNAMESE GENETIC LINGUISTIC STUDIES

*Mark J. Alves -Nan Jeon<sup>(\*)</sup>*

## INTRODUCTION

Throughout the twentieth century, controversy surrounded the origins of the Vietnamese language. Currently, there are several primary schools of thought, in which Vietnamese is considered to be an Austroasiatic, Tai-Kadai, or Austronesian language. In another approach, Vietnamese is considered to be a 'mixed' language which either cannot be shown to belong to any particular group or can be shown to have a less obvious linguistic substratum. The dominant and most well supported hypothesis so far--based on solid lexical and phonological evidence--has been that Vietnamese is a language of the Mon-Khmer branch of Austroasiatic. However, even at the end of the twentieth century, there are those who still consider Mon-Khmer to be a kind of superstratum and that deeper exploration reveals a Tai or Austronesian base.

This paper explores these issues, summarizing available evidence for each viewpoint, using primarily phonological and lexical evidence, though other factors (typology and historical feasibility) are considered as well. The conclusion reached in this paper is that, though some provocative evidence exists showing possible ancient connections with Austronesian and Tai-Kadai, Vietnamese is nonetheless a Mon-Khmer language, particularly in light of recently collected data from Minor Vietic languages as well as other language contact issues. This paper first discusses the methods used to determine genetic linguistic affiliation. Then, for each of the three main hypotheses (that Vietnamese is Mon-Khmer, Tai-Kadai, or Austronesian), previous work is discussed, stages of linguistic development are hypothesized, and a general summary is given.

For this paper, wordlists by Diguët (1910) and Maspero (1912) for Vietnamese and Tai-Kadai, Matsumoto (1928) for Vietnamese and Austronesian/Japanese, and Luce (1965) and Huffman (1977) for Vietnamese and Mon-Khmer have served as the primary sources of lexical material. The works of Nguyễn T. G. (1978, 1984) on Vietnamese lexical studies, Nguyễn V. L. (1995) on Vietnamese and the Minor Vietic language Rục, and Nguyễn T. C. (1995) on Vietnamese historical phonology were also influential. Data from each source is reviewed, and the said criteria have been applied to restrict what should or should not be used.

---

<sup>(\*)</sup>College of Technology and Commerce



# **METHODOLOGY FOR DETERMINING GENETIC LINGUISTIC AFFILIATION**

A widely held view in modern historical linguistics is that claiming certain languages share common genetic linguistic origins requires the identification of systematic correspondences in the lexical, phonological, morphological, and syntactic systems of those languages (as stated in Thomason and Kaufman 1988:8). In Vietnamese, the most useful evidence is lexical and phonological, while morphological and syntactic evidence is less useable. Vietnamese lacks agglutination and inflection, using primarily lexical compounding and reduplication. Syntactic evidence is complicated by the overall similarity of syntax in Southeast Asia. The basic sentence order 'subject-verb-object', topic-comment structure, and head-modifier word order is seen in a majority of languages. Thus, the primary focus of this work (as has been the case with previous works) is lexical and phonological, though a few assertions can still be made from morphological and syntactic data.

In the application of lexical and phonological data, two aspects have been standard in historical linguistics: (1) the use of basic vocabulary and (2) the identification of regular sound correspondences. Basic vocabulary is the most stable part of languages--it is least likely to be lost and least likely to be borrowed--and thus is the primary viable source of genetic linguistic research. Basic vocabulary must refer to semantically general phenomena common to a group of languages in a selected geographic region. Having identified possible cognates within this portion of the vocabulary, patterns of sound correspondences must be found to validate these words as true cognates. The task is also complicated by the fact that additional sound changes may have occurred over time, causing the sound patterns to be less and less obvious. In such cases, additional evidence is needed, such as different historical reconstructions of languages.

Another criterion is based on observations of the use of lexical data in searching for Vietnamese origins. In various works, words are argued to be cognates with data chosen inconsistently from varying sources or from varying historical stages of languages. To provide convincing evidence of shared genetic linguistic origins, words must be selected consistently from one language or one period of a language and not multiple sources. Words taken randomly from differing layers or subgroups increases the possibility of coincidental look-alikes.

While the theorems of basic vocabulary and regular sound correspondences are still the basis for any judgment of genetic linguistic relationships, in the specific case of Vietnamese, other criteria are needed to further reduce possible uncertainties. In particular, categorizing the type of vocabulary considered results in the inclusion or exclusion of certain words classes. Consider the following criteria.

## **Criteria for inclusion or exclusion**

1. Quantity (requirement)
2. Number of languages having shared cognates (requirement)
3. Multiple possibilities (exclusion)
4. Sound-meaning association (exclusion)

First of all, without a minimum requirement of a number of sound correspondences or a number of words that actually support the claims for sound correspondences, any chance similarities could be used. A sound correspondence is only valid if at least a few words

support it, and conversely, a word only supports claims of sound correspondences if at least a few other words show the same pattern. Next, when dealing with a group of related or potentially related languages, we should expect to find cognates not merely in one language of a language family, but in several languages from a few subbranches. As for exclusions, first, there are words that are associated with more than one language family (cf. /ma/ or some such for 'mother' and /ba/ or some such for 'father', or the words for 'eye' and 'this' which show similar phonological forms in many language families in and around Southeast Asia). Second, words that are unmarked in terms of their sound-meaning association (e.g., words indicating contact through hitting, such as 'tap', 'slap', 'whap', etc.). These criteria, though still subjective, should help to reduce ambiguity and uncertainty.

In addition to general methodology, human migration and contact must also be considered. If linguistic evidence points towards shared historical background between groups, what kind of relationship does it suggest? How viable are hypotheses in the face of available historical resources and social and geographic circumstances? Though certainly speculative, the viability of each of the linguistic-genetic hypotheses in human terms is given some consideration.

**AUSTRONESIAN EVIDENCE**

The primary works supporting the view that Vietnamese shares identifiable roots with Austronesian are those of Matsumoto (1928), Nguyễn B. L. (1971), and Nguyễn N. B. (1994). Matsumoto (Ibid.), suggest a relationship between Vietnamese and both Japanese and Malayo-Polynesian, a view Nguyễn N. B. supports. Matsumoto provided over seventy possible cognates--consisting mainly of basic vocabulary (though closer inspection shows that these are not always semantically basic)--to link Vietnamese and Japanese, but only nine posited forms to support the Vietnamese-Austronesian link. The most obvious problem confronting this list of words is the lack of phonological correspondences. However, for these words to be useable, they must provide consistent phonological correspondences--either language to language or stage to stage of languages--neither of which is the case.

Consider the data (Nguyễn B. L. 1971:477) in Table 1. 'Chin' is not a very basic body part term, and other body parts are lacking. 'Friend' is a Sino-Vietnamese form (cf. Mandarin 伴 'companion'). 'Honey' is seen in various languages of Asia and cannot be used for supporting claims of shared historical origins or even direct contact. 'Island' is most likely a Cham loan. 'Leaf' is most likely Mon-Khmer, though the Cham form is interesting. Excluding most of these examples makes 'tree' and 'mountain' seem like a chance similarity. Notice the random selection from either Malay or Cham. There should be consistent cognate evidence in both Malay and Chamic. If there is a significant time depth, then the sound correspondences should not be so direct.

**Table 1: Supposed Vietnamese-Japanese-Austronesian forms**

Gloss	Vietnamese	Malayo-Polynesian	Japanese	Source
Chin	cầm	dagu, anka', anko'	ago, agi	?
friend	bọn, bạn	ṭman	tomona-fu	Chinese
honey	mật	kemet, kemut	kimo	SEA



island	cù lao	(Cham)kalau, kulau, pulau	ikuri	Cham
Leaf	lá	(Cham)hala	*pa > ha	Mon-Khmer
mountain	non	m̃num, ɓnum, ɓnom	mine	?
tree	cây	(Malay)kayu	ko, ke	?

In another approach, Vietnamese could be considered linked to Austronesian in view of the Austric hypothesis (Schmidt 1906), which links the Austroasiatic and Austronesian language families. The best evidence linking Austronesian and Austroasiatic is primarily morphological (Reid 1994), while lexical evidence is extremely scarce (Diffloth 1994). Only indirectly does Vietnamese show any possible morphology, as seen in the closely related Minor Vietic languages, which show remnants of the causative /pa-/ prefix. The lexical evidence, too, is quite scarce, with the only possibilities being 'dog', 'eye', 'fish', and a few others. Still, even in this view, the most immediate relationship is still that Vietnamese is a Mon-Khmer language, not Austronesian.

Considering the issue of geography, Vietnamese would indeed be in the right position for Austronesian origins. The Northern Vietnamese coast would have been accessible to the Austronesian travelers, who, in current theories (e.g., Blust 1996), passed from southern China to Taiwan and then expanded southward. The problem then is to account for the huge amount of basic Mon-Khmer vocabulary and the possible development of a registral and tonal system,<sup>1</sup> even at early stages of contact with Chinese.

## TAI-KADAI EVIDENCE

Though most of the previously considered evidence linking Vietnamese to Tai-Kadai has been primarily lexical, a more recent proposal using word families (Hồng 1998) has provided some food for thought regarding the Vietnamese-Tai relationship. Maspero (1912)<sup>2</sup> listed numerous potential cognates that Vietnamese shared with Tai-Kadai or Mon-Khmer. Haudricourt (1954) readily noted that some of the supposed Tai-Kadai forms are actually Chinese in origin. However, this does not account for the more than 100 instances of Tai cognates Maspero posited to be in Vietnamese. In Maspero's data, there are many instances of weak lexical or phonological evidence, cases of onomatopoeia, or forms that are seen throughout languages in Southeast Asia. Of the Tai languages Maspero used for comparison, some are within the sphere of Vietnamese influence and may have borrowed from Vietnamese.

Consider Table 2, which contains words Maspero posited as Vietnamese-Tai cognates. 'Chicken' in Tai languages is likely the Southern Chinese pronunciation of the Chinese word. Moreover, the supposed phonological link in Vietnamese appears weak, lacking the final off-glide. Supposed cognates for 'mouse' and 'partridge' occur in Tai languages not close to Vietnamese territory, while those close to Vietnamese territory do not have them, which makes them less likely Vietnamese-Tai candidates. Moreover,

<sup>1</sup> The likelihood that Vietic had phonologically significant phonation and possibly a precursor to a four-way registral/tonal system is suggested by the data of Minor Vietic languages (see Nguyễn V. L. 1988 for data on Rục and Arem).

<sup>2</sup> A less commonly cited work is that of Diguët 1910, which included what he considered cognates between Vietnamese and a few Tai languages.



'partridge' does not fall under the category of basic vocabulary (as does 'bird', a Mon-Khmer cognate). Finally, 'cat', in addition to being somewhat onomatopoeic, is seen in Tai and Sinitic languages, as well as Vietnamese, and is more of a Pan-Southeast Asian word than one belonging to any specific language group.

**Table 2: Proposed Tai-Vietnamese cognates**

Gloss	Viet	Exclusionary reason
chicken	gà	Chinese (Maspero 1912:22)
mouse	chuột	weak lexical evidence (Ibid. 28)
partridge	đà	weak lexical evidence (Ibid. 34)
cat	mèo	Southeast Asian (Ibid. 64)

Of the approximately 100 words used for evidence of a Tai-Vietnamese link by Maspero, at least 75 can be excluded based on criteria discussed in section 2. About a dozen are Chinese loans. Several are words that could be considered onomatopoeia, especially ones related to hitting. Weak lexical evidence excludes about 30 words, weak phonological evidence excludes another 15. Eight of the words are seen throughout languages of Southeast Asia (such as 'cat', 'ivory', and 'crossbow'). 25 words form the remainder. These 25, however, are troublesome since not all are basic vocabulary and they do not form any particular class of vocabulary.

The phonological aspect is not unproblematic. Firstly, there are no significant patterns of phonological correspondence. This could potentially be due to the time depth involved, but then this also increases the likelihood of coincidental resemblances. Next, the differences in the tone system arrangements and their historical development are significant. Whereas Tai languages are generally seen as having four syllable-initial types that condition tones (Chamberlain 1971), Vietnamese has two (Haudricourt 1954). Moreover, the Minor Vietic languages are primarily four-tone systems, something not seen among Tai languages.

Some provocative evidence was recently presented in a paper Hoàng T. C. (1998) on the issue of word families, with a focus on Tai languages and Vietnamese. Her claim is that Vietnamese has a Mon-Khmer substratum, but that it belongs instead to what she called the Han-Tai family based on 'language structure', which is a result of extended contact with Tai and Chinese. She showed several groups of words in Vietnamese related both phonologically and semantically and then showed similar sets in Tai languages (based on Nghiêu 1996 and Luo 1996). Some sets of words that Hồng posited are not useable due to the constraints discussed in Section 2. There are forms that could belong to language families other than Tai, such as Chinese or Mon-Khmer. There are onomatopoeic forms. Some show weak phonological correspondences. However, there are still some forms that cannot be excluded under these criteria. Hồng's data is thought provoking since such word families should be found in related languages separated by significant time depth but within the same linguistic area. The issue becomes what the word family cognates in Vietnamese are also seen as in Tai and Mon-Khmer.<sup>1</sup> Although Hoàng's theory is tempting, her claim that Mon-Khmer is a substratum needs some

<sup>1</sup> Indeed, if one wishes to support the claim that Vietnamese is a Mon-Khmer language, similar evidence should be able to be identified, a study this author has not yet seen or tried to undertake.



additional lexical evidence to support language shift, evidence which is lacking, and solid structural evidence, evidence which has many problems. A tone system, though a significantly distinguishing phonological characteristic, is not enough to posit a linguistic relationship. Word order, too, has its problems, again, since we are dealing with the Southeast Asian linguistic area. There is also the matter of time depth and chronology. The timing of tonogenesis in Vietnamese and the Tai languages is significantly different, and many of the phonological changes in Vietnamese that make it typologically similar to Sinitic probably mainly took place over the past several centuries with no records of extensive Vietnamese-Tai bilingualism (at least for the Vietnamese). Overall, her proposal is worthy of following, but her claims based on that partial and meager evidence are certainly premature.

Based on the proximity of Vietnamese with Tai groups, the ancestors of the Vietnamese have definitely been in contact with Tai peoples in the region of modern day Southern China and Southeast Asia. Phạm 1982 discusses what appears to be an agricultural basis for showing a relationship between Vietnamese and Tai. Biogenetic evidence also suggests possible close links between the Vietnamese and Tai groups; however, other studies show links between Mon-Khmer and Tai-Kadai groups (Cavalli-Sforza 1994:225, 234). Based on the kind of contact between many groups, it would be hard to definitively sort out who provided what, when, where, and how. This means that, based on current methods, the Vietnamese-Tai relationship can only be considered a contact one at best, and likely a relatively recent one (roughly speaking, the past several centuries), unless some solid evidence can be provided to prove otherwise.

## MON-KHMER EVIDENCE

For the relationship between Vietnamese and Mon-Khmer, ample lexical and phonological evidence exists. Some of the more comparative lexical studies include those of Maspero (1912), Luce (1965), Headley and Thomas (1970), and Huffman (1977). From these sources, we can find over 100 Mon-Khmer etyma in Vietnamese. There are phonologically oriented works by Haudricourt (1954), Gage (1985), Diffloth (1989), Nguyễn T. C. (1995), and Alves (forthcoming). These works show the path of phonological development Vietnamese has undergone over the past two millennia.

### Lexical Evidence

The lexical evidence linking Vietnamese to Mon-Khmer is solid not only since it is basic vocabulary, but also because it is much the same vocabulary that relates Mon-Khmer languages to each other. That is, much of the Mon-Khmer vocabulary in Vietnamese is part of proto-Mon-Khmer vocabulary. Huffman's 1977 wordlist showed several dozen potential Vietnamese/Mon-Khmer cognates that have numerous sound correspondences with more than a dozen other Mon-Khmer languages. What this evidence shows is (1) core vocabulary, (2) sound correspondences, and (3) a spread of this vocabulary in languages both within and between sub-branches of Mon-Khmer. That fact makes this data consistent and thus methodologically meaningful.

The basic nature of the vocabulary is also notable. Instead of, say, 'chicken' or 'peacock' (as was suggested by Maspero (1912)) being Tai cognates, we find the semantically least marked word *chim* 'bird'. We also find the truly basic animal terms chó 'dog', cá 'fish', and rắn 'snake'. Rather than highly specific body parts, we find chân



'leg/foot', tay 'arm/hand', tai 'ear', and mũi 'nose'. Though showing fewer examples, there are basic MK verbs, such as n 'to sit', lắng 'to listen',<sup>1</sup> chết 'to die'. Thomas (1976) shows the numerals from one through ten as all belonging to Mon-Khmer or a sub-branch thereof. Though the pronoun system in Vietnamese has largely been replaced by a system of social terms of reference, the intimate/vulgar terms mày 'you' and nó 'he/she/it' are Austroasiatic (Pinnow 1965). In addition to these, đó '3s', the demonstratives ấy and đấy 'that/there', the Central Vietnamese interrogative mô 'what', and possibly the interrogatives nao/bao/sao may also be MK in origin, though monosyllabification has occurred.

**Phonological Evidence**

The phonological correspondences of this basic vocabulary are plentiful. Gage (1985) listed dozens of individual sound correspondences that can be grouped into several broad categories. These correspondences not only provide evidence to link Vietnamese and Mon-Khmer, they also provide a means of deconstructing Vietnamese to stages several centuries ago and earlier. At least several words support each of these sound correspondences, though there are still exceptions. Table 3 lists a few of the general categories of sound correspondences, which also indicate the kinds of sound changes that have occurred in Vietnamese.

Minor Vietic data, such as Rục (Nguyễn V. L. 1995), provide evidence that significantly changes previously supposed hypotheses about the stages and timing of certain phonological developments in Vietnamese. Table 4 shows the hypothesized range of syllable shapes and the maximum number of tones in the Vietnamese tone systems of different historical periods.

**Table 3: Vietnamese-Mon-Khmer sound correspondences**

Vietnamese	Mon-Khmer
retroflex (tr, s, r)	cluster (khl)
voiced stop (b, d)	voiceless stop (p, t)
sắc/nặng	final stop (p/t/c/k/?)
hỏi/ngã	final fricative (h/s)

**Table 4: Earlier stages of Vietnamese syllable types and tone systems**

Century	Syllable Shape	Tone System
up to the 16 <sup>th</sup>	CVCVC, CCVC, CVC	four to five tones
from the 16 <sup>th</sup> to 19 <sup>th</sup>	CCVC, CVC	five to six tones
the 20 <sup>th</sup>	CVC	five to six tones

That Vietnamese had consonant clusters is non-controversial since some have had them into this century as well as written records from the 17<sup>th</sup> century (as noted by Maspero

<sup>1</sup> This is an older, less used form.

1912). That Vietnamese had presyllables is seen indirectly from evidence of the Minor Vietic languages (Alves 1997) and the Ming dynasty Vietnamese-Chinese glossary. The smaller number of tones in previous centuries is suggested again by evidence from the Minor Vietic languages, but also the wide distribution of five-tone systems in both Vietnamese and varieties of Mường.

In showing these earlier stages of Vietnamese phonology, the apparent modern similarity between Vietnamese and either Tai or Chinese is trivialized. The modern stage is simply the recent endpoint of a long change in progress that was as much due to internal typological reduction tendencies (if not more, see Alves forthcoming) as it was the result of contact with Chinese.

## **Morphological Evidence**

There are in this category two additional pieces of information that show Mon-Khmer origins. First, Vietnamese has evidence not only of word-initial clusters, but also some evidence of prefixes. Second, though the patterns of alternating reduplication (Vietnamese *từ láy*) are common both to Mon-Khmer and Tai-Kadai, the amount of productivity is apparently much higher among Mon-Khmer languages. In addition, there is the phenomenon of phonologically separable reduplicants, a characteristic seen in Vietnamese and the Mon-Khmer language Pacoh (as discussed in Vu 1998).

## **Migration and Intergroup Contact**

The above discussed evidence allows the hypothesis that a group of Mon-Khmer peoples, proto-Vietic speakers, lived in the region of modern day northern or north central Vietnam and settled in lowland regions. The unknown factor is what kind of contact these ancestors of the Vietnamese had with other non-Mon-Khmer groups in that period, though it can be said that the contact was not significant enough to leave a solid imprint on the Vietnamese lexicon, as can be seen by the scant solid evidence for contact with non-Chinese languages in the region. Contact with Chinese did take its toll, especially over the past several centuries in which the daily Vietnamese lexicon has been saturated with Sino-Vietnamese and did end up with phonological and morphological systems that resemble modern day Chinese.<sup>1</sup>

## **CONCLUDING THOUGHTS**

For the basic Mon-Khmer vocabulary seen in Vietnamese to be dismissed would require the exclusion of basic comparative methodologies. Furthermore, in most cases, the evidence linking Vietnamese origins with language groups other than Mon-Khmer is not well supported in light of the criteria proposed in this paper. Indeed, it cannot be proved that shift did NOT occur, but then that is always the case with ANY language.

### **Table 5: Feasibility of Vietic origins**

---

<sup>1</sup> Although, Alves (forthcoming) makes claims that (a) much of the modern-day linguistic structure of Vietnamese can be attributed to natural linguistic processes rather than contact with Chinese, (b) Chinese played a relatively minor role in the non-lexical aspects of Vietnamese, and (c) many of the changes that make Vietnamese resemble Chinese structurally were only realized in the recent past few centuries.



	Mon-Khmer	Tai-Kadai	Austronesian
<b>Lexical</b>	Solid	interesting, but with problems	speculative
<b>Phonological</b>	solid	no clear patterns	no clear patterns
<b>Typological</b>	morphological evidence affixes and split eduplication)	similar syllable shape, tone systems	general SEA only
<b>Historical</b>	Clear and logical direction of change	unclear and less accountable	nothing clear

Nguyễn T. G. (1978:122-124), in a study of etyma in Vietnamese, cited Trần's 1977 lists of possible Tai-Viet cognates, of which there were only 31 (many of which could be excluded), whereas there were 65 Mon-Khmer forms and nearly 300 for Mon-Khmer languages spoken in the Central Highlands. If Vietnamese is a 'mixed' language (if such a term can be applied to any language), why then is the core vocabulary still solidly Mon-Khmer? If credible explanations can be given to account for the current phonological shape of Vietnamese, the need for a 'mixed language' hypothesis is significantly reduced.

One aspect that must be considered in positing genetic linguistic origins is the feasibility of those claims based on existing data. Existing evidence shows that Vietnamese definitely had initial consonant clusters even just a few hundred years ago, and several hundred years ago, Vietnamese probably had more clusters and possibly presyllables. Data from Minor Vietic languages show the likelihood that Vietnamese went through a four-tone stage in its tonal development. This accounts for perhaps the past thousand years. Before that, say from one to two thousand years ago, Vietnamese must have had its core Mon-Khmer vocabulary as during that time, historical records suggest contact primarily with Chinese, Cham, and Tai peoples. Thus, we enter prerecorded history and three possible genetic linguistic sources: Mon-Khmer, Tai-Kadai, and Austronesian. For the latter two hypotheses to work, there would have to have been a period in which the ancestors of modern Vietnamese speakers were in an extremely socially subordinate position. Such terminology is potentially inflammatory, however, that would indeed have to be the case.

If Vietnamese is Tai-Kadai or Austronesian, then perhaps a small group underwent some kind of relexification. There are some historical gaps into which we could fall, such as the question of the 100 Yue and whether or not Austroasiatic peoples lived in modern day Southern China (Mei and Norman 1973), though neither would clearly determine what the origin of the Vietnamese language is.

The Mon-Khmer hypothesis could be described as follows. A group of Mon-Khmer speakers settled in the north central region of Vietnam. Mon-Khmer type register and/or Austroasiatic phonation have had a chance to influence the sound system of Vietnamese. Some 'Vietnamese' speakers in the region of the Red River Valley came into contact with the Chinese. There was probably some contact between them and Tai peoples to the northwest, but this contact would have been restricted by the mountain ranges. The contact with the Chamic peoples, who arrived in Southern Vietnam after the Han dynasty, and with whom the Vietnamese typically had an unfriendly political relationship at best, would have been fairly minimal.

## REFERENCES

1. Alves, Mark J. forthcoming. What's so Chinese about Vietnamese? Papers from the Seventh Annual Meeting of the Southeast Asian Linguistics Society. Berkeley: University of California Press.
2. Blust, Robert. 1996. Beyond the Austronesian homeland: the Austric hypothesis and its implications for archaeology. in Goodenough, W (ed.) Prehistoric Settlement of the Pacific. Philadelphia: Transactions of the American Philosophical Society 86.5:117-140.
3. Chamberlain, James R. 1971. A workbook in comparative and historical Tai linguistics. Bangkok, Thailand: English Language Center of the University Development Commission.
4. Cavalli-Sforza, Luigi L. 1994. The history and geography of human genes. Princeton: Princeton University Press.
5. Diffloth, Gerard. 1989. Proto-Austroasiatic Creaky Voice. Mon-Khmer Studies 15: 139-154.
6. Diffloth, Gerard. 1994. The lexical evidence for Austric, so far. Oceanic Linguistics 33.2:309-321.
7. Diguët, Edouard Jacques Joseph. 1910. Tude de la Langue Thơ. Paris: Librairie Maritime et Coloniale.
8. Gage, William W. 1985. Vietnamese in Mon-Khmer perspective. in S. Ratankul, D. Thomas, and S. Premisirat, eds. Southeast Asian Linguistics Presented to Andre-G. Haudricourt. 493-524.
9. Haudricourt, Andre G. 1954. L'origine des tones en Vietnamien. Journal Asiatique 242:69-82.
10. Hogan, Lee C. 1993. A comparison of reconstructed Austronesian, Old Chinese, and Austro-Thai. Linguistics of the Tibeto-Burman Area 16.2: 1-55.
11. Huffman, Franklin E. 1976. The relevance of lexicostatistics to Mon-Khmer languages. Lingua 43:171-198.
12. Luce, G. H. 1965. Danaw, a dying Austroasiatic language. Lingua 14:98-129.
13. Luo, Yongxian. 1996. Word families in Tai: a preliminary account. Pan-Asiatic Linguistics. Proceedings of the 4<sup>th</sup> International Symposium on Languages and Linguistics, vol. 3:850-883.
14. Matsumoto, Nobuhiro. 1928. Le Japonais et les langues Austroasiatiques: tude de vocabulaire compař. Paris: P. Geuthner.
15. Nguyễn, Bình Lộc. 1987. Nguồn gốc Mã Lai của dân tộc Việt Nam (The Malay origins of the Vietnamese). Los Alamitos, CA:Xuân Thu. Originally published: Saigon: Bách Bộc, 1971.
16. Nguyễn Ngọc Bích. 1994. "Tiếng Việt, Tiếng Nhật và hệ Mã Lai đa đảo," Tuyển tập Ngôn ngữ và văn học Việt Nam, vol. 2:437-480. San Jose: Dòng Viện.



17. Nguyễn Tài Cẩn. 1995. Giáo trình lịch sử ngữ âm tiếng Việt (Textbook of Vietnamese historical phonology). Hà Nội: Nhà Xuất Bản Giáo Dục.
18. Nguyễn Thiện Giáp. 1978. Từ Vựng Tiếng Việt (The Vietnamese lexicon). Hà Nội: Trường Đại Học Tổng Hợp Hà Nội Xuất Bản (University of Hanoi Press).
19. Nguyễn, Văn Lợi. 1988. Sự hình thành đối lập đường nét thanh điệu bằng/không bằng trong các ngôn ngữ Việt-Mường (Trên tư liệu tiếng Arem và Rục) (The features of the comparative formation of the level/non-level tones in some Viet-Muong languages (in materials of the Arem and Ruc languages)). Ngôn Ngữ 2:3-9.
20. Norman, Jerry and Tsu-Lin Mei. 1976. The Austroasiatics in ancient South China: some lexical evidence. *Monumentica Serica* 32:274-301.
21. Pinnow, H.J. 1965. Personal pronouns in the Austroasiatic languages: a historical study. *Linga* 14:3-42.
22. Reid, Lawrence A. 1994. Morphological evidence for Austric. *Oceanic Linguistics* 33.2:323-344.
23. Schmidt, Pater Wilhelm. 1906. *Die Mon-Khmer-Völker*. Braunschweig: Druck und Verlag von Friedrich Vieweg und Sohn.
24. Thomas, David. 1976. South Bahnaric and other Mon-Khmer numeral systems. *Linguistics* 174:65-80.
25. Thomas, David D. and Robert K. Headley Jr. 1970. More on Mon-Khmer subgroupings. *Lingua* 25:398-418.
26. Thomason, Sarah Grey and Terrence Kaufman. 1988. *Language contact, creolization, and genetic linguistics*. Berkeley: University of California Press.
27. Vũ Đức Nghiệu. 1996. Những từ có liên hệ với nhau với nghĩa lịch sử trong tiếng Việt (Some words that are related in meaning historically with initials in Vietnamese). The National University of Hanoi, Dissertation.

# CÁCH ĐÁNH DẤU TRUNG TÂM VÀ PHỤ NGỮ TRONG CÁC NGÔN NGỮ ĐƠN LẬP: TRƯỜNG HỢP CẤP SO SÁNH TRONG NHÓM HÁN NGỮ (SINITIC) (TÓM TẮT)

*Umberto Ansaldo và Stephen Matthews*

Sự phân biệt về mặt loại hình học giữa cách đánh dấu trung tâm và phụ ngữ (Nichols 1986) thường được sử dụng để đánh dấu các mối quan hệ ngữ pháp về mặt hình thái học, chẳng hạn như quan hệ sở hữu cách trong ví dụ (1) và (2):

- (1) sensei -no hon (yếu tố đánh dấu phụ : tiếng Nhật)  
teacher-POSS book  
'the teacher's book'
- (2) István kocsi -ja (yếu tố đánh dấu chính: tiếng Hungari)  
Steven car -3sgPoSS  
'Steven's car'

Sự phân biệt này cần được ứng dụng một cách có hệ thống đối với các chỉ tố đánh dấu quan hệ ngữ pháp trên bình diện cú pháp trong các ngôn ngữ đơn lập, vốn không có những phụ tố hình vị ngữ pháp. Vì mục đích này, chúng tôi tập trung vào việc trình bày các cấu trúc so sánh hơn/kém trong các ngôn ngữ Hán/Sinitic. Trong các ngôn ngữ này, chỉ tố đánh dấu quan hệ so sánh có thể xuất hiện trong cương vị thành tố của danh ngữ chính (ví dụ 3) hoặc của vị ngữ (ví dụ 4).

- (3) [bi wo] lao (yếu tố đánh dấu phụ: tiếng Quan thoại)  
compare I old Bắc Kinh)  
'older than me'
- (4) [khah kuā] I (yếu tố đánh dấu chính: tiếng Minnan,  
more tall s/he tiếng Fujian)  
'taller than him'

Khái niệm “cương vị thành tố” – được hiểu như quan hệ kề cận về cú pháp giữa chỉ tố đánh dấu với từ trung tâm hoặc từ làm phụ ngữ trong một kết cấu ngữ pháp nhất định – ở đây được xem như một sự mở rộng của khái niệm hình thái học về ‘chỉ tố’. Trong các ngôn ngữ đơn lập, cương vị thành tố có thể được xác lập thông qua phép trắc nghiệm cương vị thành tố như các phép trắc nghiệm mà chúng tôi trình bày trong trường hợp về quan hệ so sánh. Chẳng hạn, ví dụ (3) cho phép chen thêm các phó từ vào như trong ví dụ (5), trong khi đó, ví dụ (4) cho phép lược bỏ danh ngữ chính như trong ví dụ (6)

- (5) [bi wo] [hai lao] (6) [khah kua]  
compare me still old more tall  
'even older than me' 'taller'

Trong bài viết này, chúng tôi chứng minh rằng có 4 khả năng logic được miêu tả theo quan niệm loại hình học của Nichols – đó là: cách dùng chỉ tố đánh dấu từ trung tâm, dùng chỉ tố đánh dấu từ phụ thuộc, dùng chỉ tố đánh dấu cả hai vế, dùng chỉ tố zero - được thể hiện trong các biến thể của tiếng Hán, đã minh họa tính đa dạng về mặt loại hình học trong ngữ pháp tiếng Sinitic. Hàm ý về mặt loại hình học của cách tiếp cận này, đặc biệt trong các ngôn ngữ đơn lập như tiếng Sinitic, cho thấy có thể tạo ra được những mô hình ngữ pháp về các chỉ tố đánh dấu từ trung tâm và chỉ tố đánh dấu từ phụ thuộc.