

Some Remarks on Reconstructing the Prehistoric Linguistic Relationships of the Tai-Kadai Language Family and Its Putative Linguistic Affiliations with Other Language Families: A Review Article

Hugo Yu-Hsiu Lee

Graduate School of Language and Communication,
National Institute of Development Administration (NIDA)
118 Seri Thai Road, Bangkok, Bangkok 10240 Thailand
e-mail: YL15@umail.iu.edu

(Received: 9-9-11/ Accepted: 20-10-11)

Abstract

From the outset the researcher has approached the topic of this review article with great trepidation. Obviously, the paper does not purport to provide final answers or solutions to problems associated with the reconstruction of the linguistic prehistory of the Tai-Kadai language family or even less definitive conclusions concerning the ethnolinguistic affiliations or affinities with other language families. Even so, the researcher has the temerity to review research investigations impinging on this topic in the work of archaeologists, geneticists, historians and linguists. Special attention is directed to recent advances in genetic investigations of the peopling of Asia. It is in the context that the researcher discusses what can be inferred from these investigations concerning the prehistoric linguistic relationships between the Tai-Kadai the Austronesian, and the Austro-Tai language families, on the one hand, and the languages of the Baiyue (Chinese: 百越; Bǎiyuè) peoples on the other. As such, this review article offers suggestions for reclassifying the Tai-Kadai language family and related language families with which it was associated or had affinities so as to include the Bǎiyuè (or Hundred Yue) languages and some of its subdivisions, including proto-Yaht and proto-Vietnamese (Việt) in the prehistory of these language groups.

Keywords: Peopling of Asia, the Austronesian language family, the Austro-Tai hypothesis, Baiyue, Tai-Kadai language family

1. Introduction

Both concrete and hypothetical evidence suggest that the Tai-Kadai languages were widespread in southeast China prior to 500 BCE when the dominance of paelo-Sinic languages commences (Forrest, 1965, 1973). Even nowadays Tai languages are found in Guangdong, Guanzou and Guangxi. In modern times, the Tai-Kadai language family—also known as the Daic, Kadai, Kradai or Krai-Dai language family—is comprised of approximately 92 languages, including 55 languages spoken in mainland Southeast Asia with 100 million speakers. Among these languages are the Thai and Lao languages as the official and national languages of the Kingdom of Thailand (Thailand) and the Lao People's Democratic Republic (Laos), respectively. The geographic distribution of Tai-Kadai or Kadai language speakers ranges from southern China (Guangxi [Guǎngxī] and Yunnan

[Yúnnán], and the island of Hainan (Hǎinán) to parts of the former French Indochina (Northern Vietnam and Laos), Thailand, and Burma (Myanmar) on the mainland of Southeast Asia, not to speak of speakers in northeastern India (Assam) (Fan, 2000).

Little doubt remains that branches of the Tai peoples migrated southward into those parts of mainland Southeast Asia which are now Laos, Myanmar (Burma) and Thailand approximately a thousand years ago or during the first millennium CE (Fan, 2000). However, details of the prehistoric spread of the Tai-Kadai language families and contacts with related language families still remain obscure. Definite knowledge of this prehistory is unavailable, not least because the further into the past we venture the less confidence we have in being able to reconstruct the linguistic history of the languages in question. It seems, therefore, that at best only speculative and tentative hypotheses for relationships between these language families can be projected with varying degrees of plausibility in view of the paucity of concrete evidence.

Those who have been involved in research—archaeologists, anthropologists, geneticists, historians and linguists—connected to a greater or lesser degree with the ethnolinguistic affiliations of speakers of Tai-Kadai tongues have never even approached a consensus on possible affiliations with the tongues spoken by the speakers of what purport to be related languages. Still, some scholars have insisted that there are well-established affinities or even affiliations between the Tai-Kadai family of languages and other language families, the speakers of which would have contact with Tai-Kadai speakers. However, it must be roundly said that there is no scholarly consensus on the nature of putative ethnolinguistic relationships between the Tai-Kadai language family and other language families. Nonetheless, some scholars still insist that affinities—if not affiliations—exist between the Tai-Kadai language family and other language families.

The question of the origins of the Tai-Kadai or Kam-Tai peoples is much more than a complex academic question. For more than a century there has been an active concern with the connections or relationships between the peoples who are speakers of different Tai-Kadai tongues and possibly related tongues. Noteworthy is it that answers to this question directly bear on the origins of the Thai people of the Kingdom of Thailand, a question having significant cultural, social and political implications. This question has only at best been partially answered in view of the current state of linguistic and scientific inquiry.

Accordingly, some scholars (e.g., Oppenheimer, 1998) find that the linguistic evidence concerning alleged linguistic affinities or affiliation is hardly compelling because the linguistic data with which scholars have to work ordinarily derives from reconstructions of the presumed prehistoric characteristics of these languages. However, other linguists are convinced of these affinities or affiliations on the basis of close similarity in the lexis of these language families as particularly evidenced by parallels, at least in the roots, of basic, non-culturally specific words used in order to mention bodily parts (e.g., “eye” or “feet”), natural phenomena (e.g., “water”), numbers (1-10) and verbs used in everyday life (e.g., “to die”). In addition, considerable similarities have been found in kinship terms (e.g., “older sister”) and in modes of address and titles (especially those which are gender-coded or age-coded). As is obvious, less fundamental words cannot be compared, especially since such words often refer to objects or phenomena for which there are no ancient counterparts (e.g., “automobile”). Accordingly, it is of importance to note that such very basic words in a language’s lexis normally remain in constant use for hundreds or thousands of years with very few changes made in the words themselves.

In the last decade or so, much more interest has been shown by linguists in this area of historical linguistics. In large measure this is not only because new viewpoints have been developed, but also because nowadays linguists are making use of the findings of archaeologists, anthropologists, geneticists and historians in developing an interdisciplinary approach to the prehistory and evolution of the Tai-Kadai language family. Current efforts to reconstruct the linguistic prehistory of the Tai-Kadai language family and its ethnolinguistic affiliations principally rely on findings that can be subsumed under the rubrics of archaeological investigations, anthropological studies, research in historical linguistics, and biological inquires using the findings of genetics.

Archaeologists have been useful in being able to show how technology and culture diffused throughout the geographical area in question by the comparative minute examination of ancient sites, especially any artifacts that remain in situ. Anthropologists have made a contribution by virtue of conducting comparative studies of handicraft production and use. Historical linguists have compared the modern languages descended from the ancient Tai-Kadai language family with those of other language families to which they may bear affinities and affiliations. They have comparatively examined morphology and syntax, lexes, and phonologies in an effort to establish putative cognates and especially common roots. These investigations are used in efforts to establish the contours of the evolution over time of these languages from common ancestors. However, in following Oppenheimer (1998), the researcher remains convinced of the inadequacy of this reconstructed linguistic evidence. Although linguistic evidence is considered in this article, this is offered in the spirit of *caveat emptor*. For the researcher is convinced of the limitations of trying to establish a common origin for these language families on the basis of a purely linguistic reconstruction of their linguistic prehistory.

Geneticists have investigated the diffusion of minor or secondary mutations and Y-chromosome DNA haplogroups in populations selected as sample populations. In applying findings based on these genetic investigations to the groups in question, it has been frequently found that explanations and interpretations based on the genetic findings cannot be reconciled or are not in consonance with the corresponding evidence from linguistics or cultural anthropology. This may well be because members of one language family have been assimilated by the members of another language family through conquest, domination, or through a combination of various means of establishing hegemony. Thus, when the evidence from genetics shows that ethnic groups have a mixed genetic inheritance and yet speak the same language, one must conclude in such cases that minority groups have been assimilated and have adopted the language of the dominant groups with the mother tongue of the minority suffering eventual extinction.

Accordingly, while taking cognizance of the four approaches just outlined, the researcher assumes in this review article that it is indisputable that the Tai-Kadai language family bears close relationships with four other major language families. These four other major language families are the hypothetical Austric language superfamily, the Austro-Tai language family (suggested as a possible subdivision of the proposed Austric super language family), the Austronesian language family (also suggested as a possible subdivision of the Austric super language family) and the Baiyue: (百越, Bǎiyuè) languages. Accordingly, the researcher holds the view that all four language families are closely related to the Tai-Kadai language family in spite of the fact that those belonging to other schools of thought would peremptorily dismiss such claims. In this light, then, this review article examines the putative prehistoric connections between the Tai-Kadai language and the four aforementioned language families.

2. The Evidence for Peopling of Asia

In an attempt to substantiate the overall views of the researcher, this review article commences with research that charts recent work examining early Asian migrations of humans on the basis of DNA and mitochondrial genome analyses. In the past decade, it has been significantly claimed that East and Southeast Asians share a common origin and thus share DNA. The overall claim is that a single major wave of migration into Asia from India via the southern route peopled Southeast and East Asia instead of the received view that this peopling occurred in two separate waves of migration involving other areas. In this connection, highly germane is the large scale research conducted by the HUGO Pan-Asian SNP Consortium (2009) in the course of which mapped were genetic divergences in Asia. After adopting the Affymetrix Gene Chip Human Mapping 50K Xba Array technique to genotype 1,928 individuals from 73 Asian (Southeast Asian and East Asian) populations scattered across 10 Asian countries at 54,794 autosomal single nucleotide polymorphisms (SNPs) sites, some 2,000 samples with 50,000 SNP markers were collected by 90 scientists.

On this basis of this evidence, it was concluded that an influx of Southeast Asian individuals contributed genetically in the highest proportion to the peopling of East Asia. Further evidence for this conclusion is that haplotype diversity increases as one moves from the north to the south and all the East Asian haplotypes are present in extant Southeast Asian populations. It is accordingly logical to argue that the older populations with more anomalies have been in one place longer than the younger populations, because the older populations have had more opportunities for divergences in the basic genetic stock to occur in contrast to younger genetic lines which have had less time for the development of genetic anomalies. Thus, these DNA findings indicate that East Asians have genetically originated and spread from Southeast Asia and Southeast Asia is the major contributor to the gene pools for East Asia with minor contributions coming from the southern regions of Central Asia.

In view of these findings based on HUGO Pan-Asian SNP Consortium's extensive investigation, it is unlikely that many linguists would deny that many proto-modern Asian languages can be reconstructed and shown to be genetically related in the sense of having common ancestors. If so, a hypothesized common ancestral language derived from a higher super-stock language family at a much earlier time may well have involved numerous and widespread populations with genetic and linguistic diversifications occurring within the members of a single language family. It is most likely that the major migratory wave via the southern route not only contributed to the peopling of Asia, but brought in its wake the great super-family of languages ancestral to many, if not all, proto-modern Asian languages including members of the proto-Tai-Kadai language family. At this stage it is not clear what languages these very early Asian settlers spoke as they gradually spread across what is now Asia. However, those linguists who accept the Greater Austric Hypothesis (e.g., Bengtson, et al.) find little difficulty in agreeing on what proto-languages these early settlers from the south spoke in spite of trenchant opposition by other linguists.

3. The Evidence for the Descent of Tai-Kadai Languages from the Austric Languages

Those who hold the view that the Tai-Kadai languages are descendents of the hypothesized Austric languages believe that proto-Tai-Kadai speakers came to what is now modern Asia in the first great prehistorical migratory wave. Those who hold this view propose affinities between the Austro-Tai language family and the postulated Austric languages. As a super language family, the Austric languages have been proposed as a huge group of related languages involving a combination of Austroasiatic and Austronesian languages. This so-called "greater Austric hypothesis" was first proposed by the German Lutheran missionary and linguistics scholar Wilhelm Schmidt in 1906.

In any event, the highly controversial greater Austric hypothesis has not been validated by sufficient empirical evidence and is, therefore, not widely accepted by linguists. The empirical lacunae stem from the fact that it has so far proven infeasible to compile a detailed proto-Austric glossary clearly applicable to related languages spoken by large numbers of prehistorical groups dispersed over a vast area stretching from the eastern part of the Indian subcontinent to mainland Southeast Asia (including what became the erstwhile French Indo-China) and insular Southeast Asia (Indonesia and the Philippines) and thence to southern China and even northeast Asia in the area now encompassed by Manchuria, Korea and Japan. Linguists who accept the Austric super language family hypothesis believe they have found compelling evidence in being able to compile what they take as cogent compilations of comparative lexeis. They have reconstructed purported etymologies suggesting that there are intimate connections and specific overlaps between the lexeis of languages held to fall under the rubric of the Austric super language family. They claim to have shown such correspondences between Austronesian and Austro-Asiatic lexeis and have accordingly commenced compiling an Austric glossary on this basis. For example, Hayes (1992, 1997, 1999, 2000, and 2001) has posted a 92-page Austric glossary online. Moreover, in the Tower

of Babel Project/the Evolution of Human Language, linguists have developed an online database of Austric comparisons containing 903 records with 46 pages.

In view of the classification scheme used by the Mansfield Library at the University of Montana (n.d.), the Austric language families would be divided into two branches, viz., the Austro-Asiatic or Austroasiatic and the Austronesian or Malayo-Polynesian language families. In this scheme, the Austroasiatic branch itself would contain three sub-branches: Munda, Mon-Khmer and Vietnamese. On the other hand, in this scheme, the Austronesian branch would have four sub-branches: Malayan, Melanesian, Micronesian and Polynesian.

Commencing circa 1996, Bengtson (1996, 1997, 1998) of the Association of the Study of Language in Prehistory proposed a Greater Austric macro-phylum or macro-family which would include Nihali (a tribal language spoken in central India) in the west, Munda spoken in southern India and the Mon-Khmer languages spoken in mainland southeast Asia in what is now Cambodia and Vietnam. On this view, Munda and the Mon-Khmer languages are closely related and the whole group would contain approximately 155 languages. Also included in this group of closely connected languages would be Hmong-Mien (Miao-Yao) spoken over a wide area from southern China to Hainan, Vietnam, Laos and Thailand in addition to the Tai-Kadai or Daic languages (of which Ethnologue lists 92 languages). Furthermore, included in this huge group of languages would be the Austronesian or Malayo-Polynesian language family with approximately 1,000 different languages spoken over an enormous area extending from Madagascar (Malagasy) through Malaysia (e.g., Dayak) and Indonesia (e.g., Javanese and Balinese), the Philippines (e.g., Tagalog, Cebuano, Ilokano and Ilongo), Polynesia (e.g., Maori, Marquesan and Samoan), the indigenous languages spoken by mountain aboriginal groups in Taiwan (Formosa), and perhaps even the endangered indigenous Ainu language spoken in northern Japan.

4. The Evidence for Tai-Kadai Languages as the Original Members of the Austro-Tai Language Family

On the hypothesis that the proto-Tai-Kadai languages descended from the Austric-Tai language family, it could be argued that the former languages could have been the original members of the Austro-Tai language family. Comparative linguistic studies appear to suggest that the proto-Tai-Kadai language speaking peoples might have been genetically related to the Austronesian language speakers. Thus, a putative relationship between the language families would suggest that the proto-Tai-Kadai language family was either intimately associated with a sister family group of languages or belonged to a language family group descended from the Austronesian language family. Some Western linguists have accordingly claimed that there is evidence indicating that the Tai-Kadai language families are related to or an actual branch of the Austronesian language family. Analogously, such a relationship has been suggested by the results of genetic inquiry in which it has been learned that the 01-YDNA haplogroup is found in both Austronesian and Tai-Kadai language family members (cf. the YDNA Human Migration data).

As we know, as early as 1901, Schlegel maintained that Malay and Siamese/Thai are genetically related and saw that this proposed relationship would tend to confirm the Austro-Tai hypothesis. The most commonly accepted classification scheme has been based on an expansion of the Austric super family or phylum such that the relationship between the Austronesian and Tai-Kadai families becomes a closer sister relationship. Thus, on this hypothesis, the sister relationship can be best described by applying the term "Austro-Tai" to the two daughter languages, which are thereby construed as being co-descendants of Austric parents.

The evidence for this hypothesis stems from the fact that many Austro-Tai cognates are morphologically and semantically similar in core lexis. This is significant because, as we know, core vocabulary is highly resistant to change or borrowing from non-related languages with which there are contacts of varying degrees of frequency (Benedict, 1942, 1975). Ostapirat (2005) presents fifty core items of vocabulary found in all five branches of Tai-

Kadai, and demonstrated half of these words—for example, the words signifying “I,” “you,” “child,” “head,” “hand,” “eye,” “tooth,” “eat,” “water,” “fire,” “moon,” and “this”—ordinarily correspond to items in the phonological systems of proto-Austronesian.

However, after examining reconstructions in subgroups of Tai-Kadai, Thurgood (1994) denied there was a genetic relationship between Tai-Kadai and Austronesian by instead claiming that the common vocabulary items actually are loan words taken over at periods of linguistic contact. Other linguists who disagree with Benedict’s sister relationship argument suggest that proto-Tai-Kadai language speakers are or were actually representatives of a return migration from Taiwan/Formosa to what is now the southeastern coast of mainland China (Sagart, 2002), or even had participated in a later migration from the Philippines to Hainan Island during the period of Austronesian expansion (Blust, 1996). In particular, Sagart (2004) hypothesizes that the proto-Kadai language originally derived from Austronesian languages. That is to say, proto-Kadai language speakers migrated to Mainland China from Taiwan/Formosa. After the migration from Taiwan/Formosa to mainland China, the proto-Tai-Kadai speakers might have undergone an ethnic fusing with pre-existing inhabitants through absorption, intermarriage, and co-inhabitation whether through a process of sinicization or even a process of de-sinicization.

5. The Evidence for Linguistic Affiliation between Tai-Kadai Languages and the Postulated Sundanese and Mu Austronesian Languages in Contradistinction to the Supposed Affiliation with Formosan/Taiwanese Languages

In the view of the researcher, it is more likely that members of the Austric language family contributed to the peopling of Asia through a major migratory wave and that this language family bears an affinity with the Austro-Tai language family. The relationship between the two language families has accordingly been conjectured as being on the order of a genetic relationship (Benedict, 1942, 1975), or—at a weaker level—a contact relationship (Thurgood, 1994). It is therefore important for this reviewer to canvass some of the leading hypotheses and conjectures regarding proposed homelands for speakers of constituent members of the Austronesian language family.

A number of researchers are convinced that today’s Southeast Asian subcontinent was once part of a larger landmass, a landmass doubles the size of today’s Indian subcontinent. This conjectured landmass has been called Sundaland and would have been the approximate size of today’s Indonesia, Malaysia, Thailand, Philippines, Vietnam, and Bangladesh (Maps of Sundaland Shelf, n.d.). Some have speculated that Sundaland could have been at least the counterpart to Plato’s Atlantis, a civilization allegedly engulfed by the seas. The Platonic Atlantis has been claimed to have sunk beneath the seas around 9,000 BCE and there are accounts of this alleged event in Plato’s dialogues *Timaeus* and *Critias*, both of which were written in 360 BCE. It is of interest that the Platonic account seems to match the Southeast Asian legend of Mu in which a flourishing civilization disappears into the depths of the sea as allegedly had Atlantis.

The Sundaland hypothesis has been most prominently postulated to be the Austronesian homeland by Oppenheimer (1998). Supposed support for Oppenheimer’s work is said to have been provided by Richard et al. (2008) on the basis of mitochondrial genome studies. Even so, these views are strongly opposed by other anthropologists, archaeologists and linguists who support the old consensus (Bellwood, 2004; Bellwood and Dizon, 2005) summarily described as the “out-of-Taiwan” hypothesis.

Other supposed evidence derives from the view that rice was not originally domesticated in China as hitherto thought, but rather in the Malay Peninsula about 9,000 years ago (Oppenheimer, 1998). On this view, rice was originally domesticated by Paleolithic peoples dwelling on the southeastern shore of Sundaland and who migrated via island chain by island chain approximately 12,000 years ago (Richards et al, 2008). If so, the Austronesians may

have constituted a dominant group which arrived later than the earlier inhabitants and assumed power over the indigenous peoples of insular Southeast Asia and the Pacific Islands. In the Genographic Project (1999-2004), the National Geographic Society took five years—from 1999 to 2004—to map the genome DNA of modern humans throughout the world. In the course of this project, it was indicated that the migrations of humans flowed from north to south into Sundaland rather than from south to north as proposed by Oppenheimer (1998).

Therefore, on the basis of these results, some linguistics conclude that it has been established that the original homeland of members of the Austronesian language family was Taiwan/Formosa where these peoples sustained themselves by the cultivation of rice crops. On this view, these rice growing agriculturalists migrated from Taiwan/Formosa about 4,000 years ago and hypothetically displaced the indigenous hunter gathers or foragers in insular Southeast Asia and the Pacific Islands (Bellwood, 2004). This hypothesis is in consonance with the pulse-pause scenario proposed by Gray et al (2009) on the basis of a phylogenetic reconstruction of some 400 languages. Accordingly, Gray et al place the homeland of the Austronesians in Taiwan around 5,230 years ago and suggest a series of expansion pulses and settlement pauses associated with technological advances (especially, canoes) and societal evolutionary modifications.

6. Evidence for the Hypothesis that the Baiyue Languages Are Outgrowths of Tai-Kadai Languages

The Austric-Tai hypothesis entails the view that the proto-Tai-Kadai language family was ultimately derived from the Austric language family through descent from the intermediary Austronesian language family. In this view, the earliest Asia settlers were genetically and linguistically related to speakers of Austronesian tongues. If so, the more modern Tai-Kadai languages have been constituted on the basis of a complex process of evolutionary transition, fusion, and derivation from prototypical linguistic forms in the course of the evolution of the Baiyue languages.

Even if we disregard the highly controversial hypothesis concerning the original homeland of speakers of Austronesian languages and the linguistic affiliation of these languages with proto-Tai-Kadai languages, linguists and historians have still reached a consensus to the effect that the Tai-Kadai languages are descendants of ancient Baiyue languages in addition to being heavily influenced by neighboring languages. It seems clear that speakers of Tai-Kadai languages were in contact with speakers of Sino-Tibetan Hmong-Mien languages, among others.

As we know, the imprecise or equivocal terms Baiyue or Hundred Yue(百越, *Bǎiyuè*), Yue (Yuè) and Viet (越, *Việt*) are used in order to mention the ancient Baiyue peoples who inhabited what is now southern China and northern Vietnam between the first millennium BCE and the first millennium with some of these peoples becoming sinicized while others were un-sinicized.

The king of the Chinese state of Qin, Qin Shi Huang (秦始皇) (259 BCE – 210 BCE), was able to unify China in 221 BCE and became the first emperor (始皇帝, *Shǐhuángdì*) of a unified China (William, 1996). After China was unified, the Yue peoples who had lived in former Wu and Yue states gradually became sinicized once their states were incorporated into a unified China. The process of sinicization involved a number of elements, including military coercion, permanent settlements of Han Chinese (漢人, *hànrén*) who were either migrants or refugees.

However, difficult logistical conditions and the intractability of a southern climate in which malarial infections were frequent militated against the rapid displacement or even sinicization of the Yue peoples. Nevertheless, the lion's share of the Yue peoples was eventually sinicized. On the other hand, the Tai-Kadai or Kam-Tai (Daic) peoples in the Yue regions tenaciously maintained their own identities. These peoples included the Zhuang, Dai, Ki,

Thai, Lao and Shan ethnic groups. Eventually, most of them migrated to areas where Austroasiatic, especially Mon-Khmer, languages were dominant in mainland Southeast Asia (see William, 1996, for the definition of “Baiyue”).

In the current state of scholarship, our knowledge of the spoken languages and dialects of the Yue group is confined to oblique references and known loanwords found in ancient and medieval Han Chinese texts. Nevertheless, there is a scholarly consensus that there was a high degree of ethnolinguistic diversity in the early Baiyue period in the area of what is now southern China.

Still, there is disagreement concerning what languages were actually spoken by the ancient Yue or Viet peoples. It has been argued that strong candidates would be members of non-Sinitic languages still spoken in certain areas of modern southern China. Examples of such proposed languages are Tai Kadai, Miao Yao or Hmong Mien and Austro-Asiatic. Evidence presented by Norman and Mei (1976) indicate that some Yue or Viet people spoke an Austro-Asiatic language. Zhengzhang (1990) maintains that some Yue or Viet people spoke an early form of Tai-Kadai because apparent Tai affiliations in place names can be found in areas in southern China where the ancient Wu and Yue states were found. Nevertheless, this does not suffice as evidence to show that all Yue and Viet peoples were speakers of Tai-Kadai tongues. Wei (1982) presents correspondences in lyrics and rhythm between the *Yue Ren Ge* or “Song of the Yue Boatman” recorded in the sixth century BC and the lexis and phonology of Eastern Zhuang, a member of the Tai-Kadai language family with native speakers still extant in the modern Guanxi province of southern China. Some Chinese anthropologists and historians (Fan, 2000) hold that Zhuang-speaking peoples were ancestors of the modern Thai people living in Thailand. Zhengzhang (1991) suggests that what may be a possible interpretation of the *Yue Ren Ge* written in Thai abugida characters dating from the late 13th century is the closest extant approximation to the original languages spoken by the Baiyue peoples, but this claim remains moot and must be further studied.

A recent phylogenetic tree study conducted by Lin (n.d.) of Mackay Memorial Hospital, Taiwan reports that the Minnan (閩南, *Minnan*, Southern Min) and Hakka (客家, Mandarin: *Kèjiā*; Cantonese: *Haakgaa*) peoples are descendants of the ancient Minyue (閩粵, *Mínyuè*) peoples from that part of the southeastern coast of China known as Fujian (福建, *Fújiàn*). In turn, these peoples were members of a branch of the larger family of Baiyue, Yue, or Viet peoples who migrated and settled in Taiwan. In the last few centuries, these peoples also migrated from southeastern China to what are now Thailand, Malaysia and Singapore and intermingled with Thais who were descendents of the ancient Luoyue (雒越, *Luòyuè*; Vietnamese: Lạc Việt) peoples.

From the point of view of genetics, it is significant that the highest frequency of three to five loci of haplotype A33Cw10B58DRB1*03DQB*02 have been found in Minnan and Hakka peoples in Taiwan, Sino-Thai in Thailand and Singapore Chinese in Singapore. The second most frequent haplotype with three loci, A2B46DRB1*09, at three percent is found among the Minnan and Hakka peoples in Taiwan with 7.7 percent being found in Singapore Chinese, 2.4 percent being found in Sino-Thais and 4.7 percent being found in Thais. A2Cw1B46DRB1*09DQB1*0303 as the second most common five loci haplotype found among Minnan and Hakka people in Taiwan has also been found in Thais (2.1 percent), Black Thais (8.3 percent), and Dai Lui Thais (5.1 percent) in Thailand.

Therefore, Lin’s study corroborates the conclusion that the Minnan and Hakka peoples in Taiwan and the Thais in Thailand are all descendents of the Yue peoples in what is now Southern China. This is shown in the fact that their HLA haplotype has been preserved as a Yue genetic marker over the course of thousands of years in which Yue peoples intermingled with neighboring ethnic groups.

In contrast to Lin’s genetics studies, cognate historical studies have strongly indicated that the Yue origin of these peoples only suggest linguistic contact rather than a stronger linguistic relationship of direct genetic descent. This is because of the absence of unequivocal phonological correspondences in respect to the corpus of core lexis being found when

comparing the Tai-Kadai language family and the Minnan and Hakka language groups as languages belonging to the overarching Cantonese group of languages.

Some lexical parallels are listed below, but the researcher does not endorse the Baiyue-Tai-Kadai lineal linkage on the basis of putative etymological reconstructions suggesting direct lineal descent. What follows is a suggestive list of what may be lexical parallels between members of the Tai-Kadai family of languages and Minnan and other languages in the Cantonese group.

Some items in the lexeis of Minnan or the Minnan spoken in Taiwan by peoples known as the Ho□h-ló (鶴佬) Hô-lo□k (河洛), or Hok-ló (福佬) peoples are listed by Lu (n.d.) insofar as they match items found in the lexeis of Tai-Kadai languages. E.g., we have alnni (this), dah/doh (where), tay (to kill; die), vah/bah (fish; meat), and za bou/cha po (men used as a modifier preceding nouns) in Minnan or Holo as a purported cognate with the Tai-Kadai bou za/po chai (men used as a postpositional modifier).

Moreover, Li (1990) lists cognates between Zhuang (a member of the Tai-Kadai language family) and Cantonese. Examples are as follows: wa (speak), mei (do not; negation), ka/kab or ga (bite), gong (to become confused), ni (this) and others. Some of these aforementioned cognates are found generally in Baiyue languages, Minnan, Holo and Hakka. In addition, phonological correspondences between Sino-Tibetan languages and dialects and Han Chinese seem to suggest a Baiyue or proto-Baiyue origin in lieu of a Sino-Tibetan origin. In view of the fact that phonological correspondences can be found in cognates shared by the Baiyue, Viet, Minnan, Holo languages and Sino-Tibetan languages and Han Chinese, some researchers have strongly suggested that in these case we have instances of linguistic fusion and mutual exchanges of lexeis in the context of contact relationships between speakers of Baiyue and Sino-Tibetan tongues. Even though it appears that etymological studies suggest that there are Tai-Kadai and Minnan and Holo correspondences and cognates, it still remains open whether the Tai-Kadai or the Mon-Khmer languages were dominant in Yue areas in what is now southern China during the Baiyue period.

7. The Evidence for the Origins of the Tai-Kadai Language Family

It would seem methodologically and theoretically sound not to examine the question of the geographical origin or origins and eventual dispersal of members of the Tai-Kadai language family until its prehistoric linguistic *mise-en-scène* and loci are at least known in outline on the basis of how Asia was peopled *ab initio* in addition to what evidence can be brought to bear on this question from considerations taken from inquiries into and evaluations of what we know of Austric languages, the Austro-Tai hypothesis and the proposed connections with the Baiyue or Yue or Viet languages. Tai-Kadai languages share similar tones, syllables, grammatical structures and evince at best a paucity of morphological inflection with Han-Chinese and the Miao-Yao languages. However, these similarities are thought to stem from language contact rather than from lineal descent from a common ancestor or ancestors (Enfield, 2005). Tai-Kadai lexeis contain a vast number of cognates with the lexeis of ancient and medieval Sinitic languages or Han-Chinese. For example, more than 900 sets of phonological and semantic correspondences are listed by Xing (1999). However, these Sino-Tibetan cognates are rarely and infrequently found in all branches of the Tai-Kadai language family members and do not include basic vocabulary items, indicating that they are loanwords deriving from language contact relationships with Sino-Tibetan (Ostapirat, 2005). However, Xing (1999) still insists that the large number of cognates provide clear evidence for the common descent hypothesis.

It is believed by scholars conducting comparative linguistic research that proto-Tai-Kadai language speakers might have migrated from Taiwan/Formosa during the period of the Austronesian pulse-pause expansion. There is a consensus among some linguists—even though Chinese historians dissent from this view—that the homeland of the Tai-Kadai language family is Taiwan or Hainan even though they may still disagree on the question of a genetic relationship between Austronesian and Tai-Kadai. By the same token, holders of the

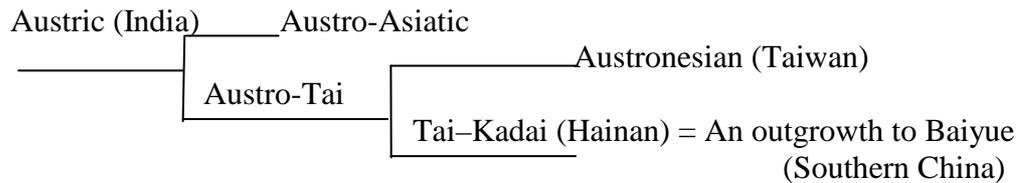
Austro-Tai hypothesis hypothesize that the Austronesian and Tai-Kadai language families share a common homeland in an area geographically close to Taiwan/Formosa. In view of the great diversity of Tai-Kadai languages spoken on the island of Hainan, Burst (1999) suggests that this indicates migration from what is now the Philippines during the period of Austronesian expansion. It is significant that both genetic evidence and Tai-Kadai oral traditions suggest that the island of Hainan. Both genetic verification and the oral traditions of Tai-Kadai speaking peoples indicate that their homeland is the island of Hainan. It has been claimed that the homeland of the Tai-Kadai language family speakers is what is now southern China, especially Fujian or Guangdong, or large adjacent islands whose inhabitants belonged to the Neolithic Longshang culture (龍山文化, Lóngshānwénhuà) of the period approximately extending from 3000 to 2000 BCE. However, these claims are dismissed by William (1996), because no historical references show that Austronesian languages were spoken in mainland China.

In contradistinction to Western linguists, Chinese historians place the homeland of members of the Tai-Kadai language family in southern China in lieu of Taiwan or Hainan. Some Chinese scholars formerly believed that the birthplace of the Tai-Kadai speaking peoples was in the borderland areas of today's Guangxi and Yunnan provinces and northern Vietnam. A newer theory propounded in the last past decade regarding the homeland of the Tai-Kadai language family claims that the Guanxi province of modern China is the birthplace of the Tai-Kadai language speaking peoples rather than Yunnan province (Fan, 2000). A large group of Tai people, known as the Zhuang, still live today in the Guangxi Zhuang Autonomous Region of Guanxi province in China. Tai (not Kadai) speaking peoples settled in what is now the area of Dien Bien Phu (Vietnamese: *Điện Biên Phủ*; French: *Điện Biên Phủ*) in northwestern Vietnam near present day Laos about 700 CE according to the Legend of Khun Borom (Fan, 2000). Branching out from the area of Dien Bien Phu, the Tai peoples began to disperse into the northern highlands and founded the cities of Luang Prabang and Chiang Saen. About 800 AD, a Tai chief named Simhanavati led the Tais to make contact with the Austroasiatic (more specifically, Mon-Khmer or Cambodian) speaking peoples of mainland Southeast Asia) and founded the city of Chiang Saen as recorded in the Legend of Simhanavati (Fan, 2000). Avoiding absorption or even annihilation by the Chinese, the Tais moved further away from southern China and eventually the Thai, Lao and Shan peoples adopted separate identities and areas of regional dominance. Eventually the Tai peoples spread throughout the former French Indo-China and as far west as Assam in what is now India.

8. Conclusions

The Tai-Kadai language family and its relationships with other language families are of tremendous importance in understanding the linguistic prehistory of modern southern China, mainland Southeast Asia, Asia and Oceania. The prehistoric linguistic relationships between the individual Tai-Kadai language family and their linguistic affiliations merits additional study with reconstructions based on genetic, archaeological, anthropological and linguistic evidence in order to generate testable hypotheses. Building on Benedict's model (1942, 1975), this review article has briefly sketched some of the major hypotheses developing concerning the prehistoric linguistic relationships between the members of the Tai-Kadai language family and their supposed linguistic affiliations with other—especially more capacious—language families. What the researcher has presented here is a skeleton outline of important hypotheses, though obviously there is no claim or pretension to definitive conclusions. What has been presented is of course subject to both criticism and future development and amplification.

In conclusion, finally, presented below is a diagrammatic view of the hypothesized prehistoric linguistic relationships between the Tai-Kadai language family and its affiliations (modified from Benedict, 1942, 1975):



Acknowledgement

This review study is funded by the Graduate School of Language and Communication at the National Institute of Development Administration (NIDA), Thailand. The review article is edited by Professor Dr. Jack Clontz. He is also recognized as the honorary second author of this article.

References

- [1] P. Bellwood, *First Farmers: The Origins of Agricultural Societies*, (2004), John Wiley & Sons.
- [2] P. Bellwood and E. Dizon, The Batanes Archaeological Project and the "Out Of Taiwan" Hypothesis for Austronesian Dispersal, 1, Taitung, Taiwan: *Journal of Austronesian Studies*, (2005), 1–32.
- [3] P. K. Benedict, Thai, Kadai and Indonesian: a new alignment in Southeast Asia, *American Anthropologist*, 44 (1942), 576-601.
- [4] _____, *Austro-Thai language and culture, with a glossary of roots*, (1975), New Haven: HRAF Press.
- [5] J. D. Bengtson, A Case for the Austric Affiliation of Ainu. In Shevoroshkin (1992), 364.
- [6] _____, Nihali and Ainu, *Mother Tongue*, 2 (1996), 51-55.
- [7] _____, Some comments on "Ilia Peiros" "Nihali" and "Austroasiatic," *Mother Tongue*, 3 (1997), 47-50.
- [8] _____, Review of Patrie (1982), *Mother Tongue*, 4 (1998), 111-113.
- [9] R. Blust, Beyond the Austronesian homeland: The Austric hypothesis and its implications for archaeology. In *Prehistoric Settlement of the Pacific*, ed. by H. Ward (1996).
- [10] N. J. Enfield, Areal linguistics and mainland Southeast Asia, *Annual Review of Anthropology*, 34 (2005), 181-206.
- [11] Evolution of Human Languages, An international project on the linguistic prehistory of humanity coordinated by the Santa Fe Institute (n.d.). Retrieved December 12, 2011, from <http://ehl.santafe.edu/>
- [12] H. G. Fan (范宏贵), Tong Gen Sheng de Minzu: Zhuang Tai Gezu Yuanyuan yu Wenhua (in Chinese) (同根生的民族——壮泰各族渊源与文化), (2000), Beijing: Guangmin Ribao Chubanshe.
- [13] K. A. D. Forrest, *The Chinese Language* (Second edition), (1965), London: Faber and Faber Ltd.
- [14] _____, *The Chinese Language* (Third edition, revised and enlarged), (1973), London: Faber and Faber Ltd.
- [15] R. D. Gray, Drummond, A. J. and Greenhill, S. J., Language phylogenies reveal expansion pulses and pauses in pacific settlement, *Science* 323 (5913) (2009), 479-483. Retrieved December 12, 2011, from <http://www.sciencemag.org/content/323/5913/479.abstract>
- [16] L.V. H. Hayes, On the track of Austric, Part I: introduction, *Mon-Khmer Studies*, 21 (1992), 143-77.

- [17] _____, On the track of Austric, Part II: Consonant mutation in early Austroasiatic, *Mon-Khmer Studies*, 27 (1997), 13-41.
- [18] _____, On the track of Austric, Part III: Basic vocabulary correspondence. *Mon-Khmer Studies*, 29 (1999), 1-34.
- [19] _____, The Austric Denti-alveolar sibilants, *Mother Tongue*, 5 (2000), 3-14.
- [20] _____, On the origin of affricates in Austric. *Mother Tongue*, 6 (2001), 95-117.
- [21] _____, Austric Glossary (n.d.). Retrieved December 12, 2011, from <http://home.att.net/~lvhayes/Langling/langpg3.htm>
- [22] J. F. Li, 粵語中的壯侗語族語言底層初析 (in Chinese), *Journal of Minzu University of China*, 6, (1990), 71-76, China Academic Journal Electronic Publishing House.
- [23] M. Lin, The origin of Minnan and Hakka, the so-called "Taiwanese", inferred by HL AStudy (n.d.). Retrieved January 31, 2012, from <http://www.wufi.org.tw/eng/linmalie.htm>
- [24] E. Lu, Baiyue vocabulary in holo (n.d.). Retrieved December 12, 2011, from <http://ws.twl.ncku.edu.tw/hak-chia/1/lou-ek-ki/kongpah-oat.htm>
- [25] Mansfield Library at the University of Montana, Austric language (n.d.). Retrieved December 12, 2011, from <http://www.lib.umt.edu/default.php>
- [26] Maps of Sundaland Shelf, The information below relates to the Minerva cruise "Palaces and Pagodas" (n.d.). Retrieved January 31, 2012, from <http://petetherockman.com/Sundaland.html>
- [27] J. Norman and T. L. Mei, The Austroasiatics in ancient south China: Some lexical evidence, *Monumenta Serica*, 32 (1976), 274-301.
- [28] S. Oppenheimer, *Eden in the East: The Drowned Continent of Southeast Asia*, Phoenix (Orion) (1998).
- [29] W. Ostapirat, Kra–Dai and Austronesian: Notes on phonological correspondences and vocabulary distribution. In L. Sagart, R. Blench, and A. Sanchez-Mazas, eds, *The Peopling of East Asia: Putting Together Archaeology, Linguistics and Genetics*, (2005), 107–131, London: Routledge Curzon.
- [30] M. B. Richards, et al, Climate change and postglacial human dispersals in Southeast Asia, *Molecular Biology and Evolution*, 25 (2008), 1209-1218.
- [31] L. Sagart, Sino-Tibeto-Austronesian: An updated and improved argument, Paper presented at *Ninth International Conference on Austronesian Linguistics (ICAL9)* (2002), 8–11, Canberra, Australia.
- [32] _____, The higher phylogeny of Austronesian and the position of Tai–Kadai, *Oceanic Linguistics*, 43 (2004), 411-440.
- [33] W. Schmidt, Die Mon-Khmer-Völker, ein Bindeglied zwischen Völkern Zentralasiens and Austronesiens [The Mon–Khmer Peoples, a Link between the Peoples of Central Asia and Austronesia]. *Archiv für Anthropologie, Braunschweig*, new series, 5 (1906), 59-109.
- [34] The Genographic Project (1999-2004). Retrieved December 12, 2011, from <https://genographic.nationalgeographic.com/genographic/atlas.html>
- [35] The HUGO Pan-Asian SNP Consortium, Mapping Human Genetic Diversity in Asia, *Science*, 326 (5959), (2009), 1541-1545. Retrieved January 31, 2012, from <http://www.sciencemag.org/content/326/5959/1541.short>
- [36] G. Thurgood, Tai-Kadai and Austronesian: The nature of the historical relationship, *Oceanic Linguistics*, 33 (2) (1994), 345-368.
- [37] X.W. Wei, On the language of the Hundred Yue people (in Chinese), *In papers on the history of the Hundred Yue People*, (1982), Beijing: Chinese Social Sciences Press.
- [38] M. William, Defining the hundred Yue, *Bulletin of the Indo-Pacific Prehistory Association*, 15 (1996), 93–100.
- [39] G. W. Xing (邢公畹), *汉台语比较手册* (in Chinese), (1999), 商务印书馆.

- [40] Y-DNA Human Migration, (n.d.). Retrieved December 12, 2011, from http://www.kerchne.com/images/dna/ydna_migrationmap_%28FTDNA2006%29.jpg
- [41] S. F. Zhengzhang, Some Kam-Tai words in place names of the ancient states of We and Yue. *Minzu Yuwen* (in Chinese), 6 (1990), 16-18.
- [42] _____, Decipherment of Yue-Ren-Ge (Song of the Yue boatman), *Cahiers de Linguistique Asie Orientale*, 20 (2) (1991), 159-168.