

Reviewing the prehistoric linguistic relationships of the Tai–Kadai language family and its putative linguistic affiliations: a meta-analysis study and abbreviation edition

Hugo Yu-Hsieu Lee¹

Jack Clontz²

Abstract

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 ethno-linguistic 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 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 Baiyue (or Hundred Yue) languages and some of its subdivisions, including proto–Yaht and proto–Vietnamese in the prehistory of these language groups.

Keywords: Linguistic prehistory, the Austric language family, the Austro–Tai hypothesis, Baiyue, the Tai–Kadai language family

¹Lecturer, Ph.D., Graduate School of Language and Communication, Founder and Director of M.A. Program in Chinese for Professional Development and Ph.D. Program in Chinese and Communication, National Institute of Development Administration (NIDA), Bangkok City, Thailand.

The corresponding author can be reached at YL15@umail.iu.edu

²Professor, Ph.D., Professor Emeritus, Kyoai Gakuen College Maebashi, Japan.

Background

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.

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, 1973, 1984). 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 and Yunnan and the island of Hainan 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).

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 inquiries 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, lexeis, 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 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.

Context review for peopling of mainland and Pacific Asia

In an attempt to capitalize this topic, this study should begin with research that charts recent works exploring the trajectory of humans based on DNA and mitochondrial genome. In the past decade one of the intriguing claims is that East and Southeast Asians share a common origin or source of DNA. That is to say, one major wave of migration into Asia from India through the southern route, in contrast to two waves of migration into Asia, is suggested by HUGO Pan-Asian SNP Consortium (2009), mapping genetic diversity in Asia. After adopting Affymetrix GeneChip Human Mapping 50K Xba Array to genotype 1,928 individuals from 73 Asian (Southeast Asian and

East Asian) populations scattered across 10 Asian countries at 54,794 autosomal SNPs collecting 2000 some samples and 50,000 some single nucleotide polymorphic markers by 90 scientists, their report shows that an influx of Southeast Asian individuals contributed genetically to most peopling of East Asia, because the haplotype diversity increases from north to South and all the East Asian haplotypes are present in extant Southeast Asian populations. It is logical to argue that the older population with more anomalies has been in one place longer than the younger population, because the older population has had more opportunities to allow the genetic stock to diverge whereas the younger genetic lines might have had less time to make genetic anomalies happen. Thus, this DNA study's findings conclude that East Asians have genetically originated and spread out from Southeast Asia, and Southeast Asia is the gene pools for East Asia with minor contributions from central–South Asian populations.

Inferred by research findings based on HUGO Pan-Asian SNP Consortium (2009), it is unlikely for linguists to deny that many of proto-modern Asian languages can be reconstructed and proven to be genetically related. If so, a hypothesized common ancestral language derived from a higher super-stock language family at a greater time-depth might have contained numerous and widespread populations, diversification of genes and language variations than within a single language family. It is most likely that the major migratory wave through southern route to contribute to peopling of Asia also brought the greater super-family of language ancestral to many, if not all, proto-modern Asian languages including proto-Tai-Kadai language family. At this stage it is not clear to hypothesize what languages these Asian-initial settlers speak as they gradually inhabited what is now Asia. However, among linguists who accept the Greater Austric Hypothesis (e.g., John D. Bengtson) albeit not well-received by all linguists, there is no difficulty to reach an agreement to what proto-languages these southern routers spoke.

Introduction for the descent of Tai–Kadai languages from the Austric languages

Those who hold the view that the Tai–Kadai languages are descendants 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, La Vaughn Hayes (1992, 1997, 1999, 2000, 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, John 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.

Introduction for Tai-Kadai languages co-founding the Austro-Tai language family

Hypothetically descended from the Austric language family, proto-Tai-Kadai speaking people were founding members of their Austro-Tai language family. Comparative linguistic studies appear to suggest that the proto-Tai-Kadai language speaking people might have been in a genetic relation with the

Austronesian language family, either an Austronesian sister or a descent. By presenting suggestive evidence, Western linguists indicate that Tai–Kadai language families are related to or a branch of the Austronesian language family. A similar statement made by genetic studies also confirms that O1 Y–DNA haplogroup can be found in both the Austronesian and the Tai–Kadai (See Y–DNA Human Migration).

As early as in 1901 Schlegel maintains that Malay and Siamese/Thai are genetically related to one another, confirming the Austro–Tai hypothesis. The most commonly accepted classification scheme to expand Austric super family's phylum is that Austronesian and Tai–Kadai families are sister relationship, termed Austro–Tai, as two daughters of the Austric parent, because many of the Austro–Tai cognates including similarities in forms (morphologically similar) and meanings (semantically similar) are found in their core vocabulary which are by and large resistant to change and borrowing albeit frequent or infrequent language contacts (Benedict, 1942, 1975). Ostapirat (2005) presents fifty core vocabulary found in all five branches of Tai–Kadai, and demonstrated half of their words, for example, "I", "you", "child", "head", "hand", "eye", "tooth", "eat", "water", "fire", "moon", and "this", can be regularly corresponded to sounds in proto–Austronesian.

However, after examining reconstructions in subgroups of Tai–Kadai, Thurgood (1994) denied a genetic relationship between Tai–Kadai and Austronesian but claims a relationship of loan words from language contact. Other linguists who disagree with Paul Benedict's sister relationship argument suggest that proto–Tai–Kadai language speakers are a back–migration from Taiwan/Formosa to what is now southeast coast of mainland China (Sagart, 2002) or a later migration from the Philippines to Hainan Island during the Austronesian expansion (Blust, 1996). In particular, Sagart (2004) hypothesizes that the proto–Kadai language is originated from Austronesian languages. That is to say, proto–Kadai language speakers migrated to Mainland China from Taiwan/Formosa. After migration from Taiwan/Formosa to mainland China, proto–Tai–Kadai speakers might have been underwent ethnic fusion by absorbing and inter–marrying with co–inhabited pre–existing populations and underwent sinicized or un–sinicized process.

Introduction for linguistic affiliation between Tai-Kadai language family and Austronesian language family

More likely derived from the Austric language family contributed to peopling of Asia with a major migratory wave and later became a proposed affinity with the Austro-Tai language family, it is conjectural that proto-Tai-Kadai language family is in a genetic relationship (Benedict, 1942, 1975) or a contact relationship (Thurgood, 1994) with the Austronesian language family. It is thus important for this review article to point out some leading hypotheses and speculation regarding the location of homeland of the Austronesian language family.

In what is today subcontinent of Southeast Asia, there used to be a larger landmass (twice the size of modern India) named Sundaland to cover present-day Indonesia, Malaysia, Thailand, Philippines, Vietnam, and Bangladesh (Maps of Sundaland Shelf, n.d.). Sundaland might be a candidate for what Plato speaks of, a sinking civilization called Atlantis, because it was sunk around 9000 BC and it matches what insular Southeast Asians' legend called Mu, which is similar to the sinking of Atlantis. The new hypothesis, arguably the Sundaland is the cradle of humans and the homeland of Austronesian, originally hypothesized by Oppenheimer (1998) and supported by Richards et al's (2008) full mitochondrial genome study, is not unlikely to draw criticism and debates from anthropologists, archaeologists and linguists who support the old consensus (Bellwood, 2004; Bellwood & Dizon, 2005) termed out-of-Taiwan hypothesis. Albeit the domestication of rice might not come from China but in the Malay Peninsula about 9,000 years ago (Oppenheimer, 1998) and stone-age people on southeastern shore of Sundaland expanded across the island chains about 12,000 years ago (Richards et al, 2008), Austronesians may have been formed by an elite group that arrived later to have power over already peopled insular Southeast Asia and Pacific Islands. According to the Genographic Project (1999-2004), National Geographic Society spent five years from 1999 to 2004 to map the genome DNA of modern humans throughout the world, indicating that modern human migration flowed into Sundaland from north to south rather than from south to north as proposed by Oppenheimer (1998).

The leading hypothesis, albeit some linguists take this view as definitive, of the homeland of the Austronesian language family remarks that Austronesians are rice-growing agriculturalists who expanded from Taiwan/Formosa about 4,000 years ago and hypothetically displaced the indigenous hunter-gathers or foragers in insular Southeast Asia and Pacific Islands (Bellwood, 2004). Concord with the pulse-pause scenario after constructing a phylogeny of 400 languages, Gray and Greenhill's study (2009) places the homeland of Austronesian in Taiwan around 5230 years ago and suggests a series of expansion pulses and settlement pauses associated with scientific (i.e., canoes) and societal modernization.

Introduction for Baiyue languages as an outgrowth from Tai-Kadai languages

Supposedly deriving from the Austric language family spoken by initial Asian settlers and genetically relating to later developed Austronesian language family (called Austri-Tai hypothesis), proto-Tai-Kadai language family members most probably underwent a transition, fusion and formation from prototypes toward more modern Tai-Kadai languages during Baiyue periods. Leaving aside the far-fetched homeland issue of Austronesian languages and their linguistic affiliation with proto-Tai-Kadai languages, it is a consensus among linguists and historians that Tai-Kadai languages are descents of ancient Baiyue languages and are heavily influenced by neighboring languages in a contact relationship such as Sino-Tibetan and Hmong-Mien, borrowing many loanwords and being sinicized. The ancient Baiyue or Hundred Yue or Yue or Viet is a loose-fitting terminology employed in referring to the partly sinicized or un-sinicized non-Chinese populations who inhabited what is now southern China and northern Vietnam between the first millennium BC and the first millennium AD (William, 1996).

After the unification of China by the emperor Qin Shi Huang or emperor Qin about 210 BC, Yue peoples in former Wu and Yue states turned out to be incorporated into Chinese empire. Yue people underwent sinicization by imperial military power, regular settlement of Chinese migrants and Chinese refugees. However, the difficulty of logistics and malarial climate in southern China made the signification and displacement of Yue peoples a slow process.

Most Yue peoples were eventually sinicized, but Tai–Kadai Yue people or Kam–Tai (Daic) including Zhuang, Dai, Ki, Thai, Lao, and Shan retained their ethnic identities and migrated to Austroasiatic (more specially, Mon–Khmer) speaking areas in the old Indo–China or mainland Southeast Asia (see William, 1996, for defining Baiyue, emphasis in original).

The current state of knowledge on Yue speech is limited to incomplete references and loanwords in ancient and medieval Han–Chinese, but scholars agree that there was ethno–linguistic diversity during Baiyue period in present–day southern China. There is disagreement about the languages ancient Yue or Viet people speak, with candidates drawn from the non–Sinitic language families still represented in areas of modern southern China, for example, Tai–Kadai, Miao–Yao or Hmong–Mien and Austroasiatic. Evidence presented by Jerry Norman and Mei Tsu–Lin (1976) indicate that some Yue or Viet people spoke an Austroasiatic language. Zhengzhang (1990) maintains that some Yue or Viet people spoke an early form of Tai–Kadai because Tai affiliations in naming places in ancient Wu and Yue states can be found, but do not indicate that all Yue or Viet people were exclusively Tai–Kadai speakers. Wei (1982) presents a correspondence in lyrics and rhythm between the Yue–Ren–Ge or Song of the Yue boatman recorded in the sixth century BC and Done Zhuang language—a member of Tai–Kadai language family whose speakers still live in what is modern Guanxi province, Southern China. Zhuang speaking people are believed to be ancestral to modern Thai people living in Thailand by some Chinese anthropologists and historians (i.e., Honggui, Fan, 2000). Zhengzhang (1991) suggests a possible interpretation of the Yue–Ren–Ge or Song of the Yue boatman written in Thai ideology–graph/letters dating from the late 13th century as the closest approximation existing to the original language spoken by Baiyue people, but his claim remains debatable and is subjected to further study.

A recent phylogenetic tree study conducted by Lin (n.d.) in Mackay Memorial Hospital, Taiwan, reports that Minnan or Southern–Min and Hakka who are ancient Minyue (A branch of Baiyue or Yue or Viet people) decedents of settlers in Taiwan from the southeast coast of China, (i.e., Fujian) during the last few centuries merge together and cluster with Thai–Chinese and

Singapore Chinese, followed by a cluster with Thais who are descents of ancient Luoyue (A branch of Baiyue or Yue or Viet people). The highest frequency of three-to five-locus haplotype of A33–Cw10–B58–DRB1*03–DQB*02 have been found in Minnan or Southern–Min, Hakka in Taiwan, Thai–Chinese or Chinese–Thai in Thailand and Singapore Chinese in Singapore. A2–B46–DRB1*09 (3%) as the second most common three locus haplotype found among Minnan or Southern–Min and Hakka people in Taiwan has been found in Singapore Chinese (7.7%), Thai–Chinese or Chinese–Thai (2.4%), and Thais (4.7%). A2–Cw1–B46–DRB1*09–DQB1*0303 as the second most common five-locus haplotype found among Minnan or Southern–Min and Hakka people in Taiwan has also been found in Thais in Thailand (2.1%), Black Thais (8.3%), and Dai Lui Thais (5.1%). Lin's study corroborates the truth that Minnan or Southern–Min, Hakka descendants in Taiwan, and Thais in Thailand are descendants of Yue in what is now Southern–China, and their HLA haplotype have been preserved as a Yue genetic marker through thousands years during Yue people's fusion with neighboring ethnic groups.

In agreement with the genetic study undertaken by Lin, historical cognate studies indicate a language contact relationship instead of a genetic relationship (because lack of regular sound correspondences in a vast number of core vocabulary) among the Tai–Kadai language family and Minan or Holo/Southern–Min languages and Hakka as a branch of Cantonese. Before listing some lexical parallels involving descendants of Baiyue languages and Tai–Kadai languages, this review article disclaims any author's endorsement on proto- etymology reconstruction. The following list of lexical parallels between the Tai–Kadai language family, Minan or Southern–Min, and Cantonese should not be seen as definitive etymologies but cited as references only. Some vocabulary found in Minan or Holo/Southern–Min listed by Lu (n.d.) can also be found in Tai–Kadai language family members, for example, "alnni" (this), "dah/doh" (where), "tay" (to kill; die), "vah/bah" (fish; meat), za bou/cha "po" (men, modifier before noun) in Minan or Holo/Southern–Min is a cognate with Tai–Kadai's "bou za/po chai" (men; modifier after noun). Jin–Fang, (1990) lists cognates between Zhuang as a member of the Tai–Kadai language family and Cantonese, for instance, "wa" (speak), "mei" (don't; negation), "ka/kab" or "ga" (bite), "gong" (get confused), "ni" (this), etc.

Some of the above listed cognates (shared among Baiyue languages, Minnan or Holo/Southern–Min and Hakka) cannot find sound correspondences in Sino–Tibetan and Han–Chinese, indicating their origins of Baiyue or proto–Baiyue instead of Sino–Tibetan. However, for some cognates shared by Baiyue or Viet and Minnan or Holo/Southern–Min, sound correspondences can be found in Sino–Tibetan and Han–Chinese, indicating their fusion and mutual exchange of vocabulary in a contact relationship between Baiyue and Sino–Tibetan speakers. Albeit corresponded cognates or etymologies can be found between Tai–Kadai languages and Minnan or Holo/Southern–Min, it seems likely that whether Tai–Kadai or Mon–Khmer languages were dominant in Yue areas during Baiyu period in modern southern–China is an open question.

Synthesis for Tai–Kadai languages

It would seem wise to not examine homeland and dispersal of the Tai–Kadia language family until its prehistoric linguistic scenario is outlined from initial peopling of Asia, Austric languages, Austro–Tai hypothesis, to Baiyue or Yue or Viet languages. Tai–Kadai languages share comparable tones, syllables, grammatical structures and short of inflection with Han–Chinese and Miao–Yao languages, but these are believed to be in a language contact relationship instead of arguing for a common descent (Enfield, 2005). Tai–Kadai languages contain a vast number of cognates with ancient and medieval Sinitic or Han–Chinese. For example, 900 sets of phonological and semantic correspondences are found by Xing (1999). However, these cognates with Sino–Tibetan are rarely and infrequently found in all branches of Tai–Kadai language family members and do not include basic vocabulary, indicating that they are loanwords in a language contact relationship with Sino–Tibetan (Ostapirat, 2005) despite Xing (1999) insists on a common descent hypothesis.

It is believed by scholars in comparative linguistic research that proto–Tai–Kadai language speakers might be migrants from Taiwan/Formosa during the period of Austronesian’s pulse–pause expansion. There is a consensus among some linguists (excluding Chinese historians), albeit disagreement on their genetic relationship between Austroenesian and Tai–Kadai, that the homeland

of the Tai–Kadai language family is in Taiwan or Hainan. By the same token, the Austro–Tai hypothesis suggests a common origin for the Austronesian language family and the Tai–Kadai language family whose hypothesized location of homeland is geographically close to Taiwan/Formosa. Because the diversity of Tai–Kadai languages is found on the Hainan Island (see Robert Blust's (1999), hypothesis on migration from Philippines during the Austronesian expansion, emphasis in original). Both genetic verification and Tai–Kadai speaking people's oral traditions put their hypothesized homeland on the Hainan Island. The claims to place the homeland of the Tai–Kadai language family on what is now Southern China or on major adjacent islands and coastal Fujian or Guangdong as the Neolithic Lonshang culture of 3000 BCE to 2000 BCE are dismissed by William (1996), because no historical references can show that any Austronesian languages were spoken in mainland China.

In disagreement with western linguistics, Chinese historians place the homeland of Tai–Kadai on southern–China instead of Taiwan or Hainan. Some Chinese scholars used to support the view to place the cradle of Tai–Kadai speaking people in the common border of today's Guanxi, Yunan of China and northern Vietnam. A newer theory in the past decade regarding the homeland of the Tai–Kadai language family claims that Guanxi province of modern China is the cradle of Tai–Kadai language speaking people rather than Yunan province in China (Fan, 2000). A large number of Tai people, known as the Zhuang, still live in Guangxi Zhuang Autonomous Region of Guanxi province in China today. Tai (not Kadai) speaking people settled in what is today Dien Bien Phu in modern Vietnam about 700 AD (Khun Borom Legend) (Fan, 2000). Commencing from Dien Bien Phu in modern Vietnam, Tais began to disperse into northern highlands and founded the cities of Luang Prabang and Chiang Saen. About 800 AD, a Tai chief named Simhanavati led Tais to contact with the Austroasiatic (more specifically, Mon, Khmer or Cambodian) speaking people of mainland Southeast Asia and founded the city of Chiang Saen (Simhanavati Legend) (Fan, 2000). Keeping away from annihilation by China led the Tai dispersal to further south from Southern China to found the Lao, Thai and Shan people of the old Indochina or modern mainland Southeast Asia, and migrated to as far as Assam, India.

Bengtson, J. D. (1992). A case for the Austric affiliation of Ainu. In Shevoroshkin (1992), p.364. [One–page abstract derived from unpublished notes.]

----- (1996). Nihali and Ainu. *Mother Tongue*, 2, pp. 51–55.

----- (1997). Some comments on Ilia Peiros’ “Nihali and Austroasiatic”. *Mother Tongue* 3: 47–50.

----- (1998). Review of Patrie (1982). *Mother Tongue*, 4, pp. 111–113.

Blust, R. (1996). Beyond the Austronesian homeland: The Austric hypothesis and its implications for archaeology. In: Prehistoric Settlement of the Pacific, ed. by Ward H. Goodenough, DIANE Publishing Co, Collingdale PA, 1996, pp. 117–137. (Transactions of the American Philosophical Society 86.5. (Philadelphia: American Philosophical Society).

Enfield, N. J. (2005). Areal linguistics and mainland Southeast Asia. *Annual Review of Anthropology*, 34, pp. 181–206.

Evolution of Human Languages. (n.d.). An international project on the linguistic prehistory of humanity coordinated by the Santa Fe Institute. *Retrieved from* <http://ehl.santafe.edu/>

Fan, H. 范宏貴. (2000). *Tong Gen Sheng de Minzu: Zhuang Tai Gezu Yuanyuan yu Wenhua* 同根生的民族——壯泰各族淵源與文化. Beijing: Guangmin Ribao Chubanshe.

Forrest, R.A.D. (1948). *The Chinese language*, London: Faber and Faber.

----- (1973). *The Chinese language* (3rd edition, revised and enlarged), London: Faber and Faber.

Gray, R. D., Drummond, A. J., & Greenhill, S. J. (2009). Language phylogenies reveal expansion pulses and pauses in Pacific settlement, *Science* 23: 323 pp. 479–483 (DOI: 10.1126/science.1166858). *Retrieved from* <http://www.sciencemag.org/content/323/5913/479>.

Hayes, La Vaughn H. (1992). On the track of Austric, Part I: Introduction, *Mon–Khmer Studies*, 21, pp. 143–77.

----- (1997). “On the track of Austric, Part II: Consonant mutation in Early Austroasiatic.” *Mon–Khmer Studies*, 27, pp. 13–41.

----- (1999). On the track of Austric, Part III: Basic vocabulary correspondence. *Mon–Khmer Studies*, 29, pp. 1–34.

----- (2000). The Austric Denti–alveolar sibilants. *Mother Tongue*, 5, pp. 3–14.

----- (2001). On the origin of affricates in Austric. *Mother Tongue*, 6, pp. 95–117.

----- (n.d.). Austric glossary. Retrieved from <http://home.att.net/~lvhayes/Langling/langpg3.htm>.

Li, J. F. (1990). 粵語中的壯侗語族語言底層初析. *Journal of Minzu University of China*, 6, pp. 71–76. China Academic Journal Eletronic Publishing House.

Lu, Ekki. (n.d.). Baiyue vocabulary in holo. Retrieved from <http://ws.twl.ncku.edu.tw/hak-chia/1/lou-ek-ki/kongpah-oat.htm>

Mansfield Library at the University of Montana (n.d.). Austric language, Retrieved from <http://www.lib.umt.edu/default.php>

Maps of Sundaland Shelf (n.d.). The information below relates to the Minerva cruise “Palaces and Pagodas.” Retrieved from <http://petetherockman.com/Sundaland.html>

M. Lin. (n.d.). The origin of Minnan & Hakka, the so-called “Taiwanese”, inferred by HLA study, Retrieved from <http://www.wufi.org.tw/eng/linmalie.htm>

Norman, J., & Mei, T.L. (1976). The Austroasiatics in ancient south China: Some lexical evidence. *Monumenta Serica*, 32, pp. 274–301.

Oppenheimer, S. (1998). *Eden in the East: The drowned continent of Southeast Asia*, Phoenix (Orion).

Ostapirat, W. (2005). Kra–Dai and Austronesian: Notes on phonological correspondences and vocabulary distribution. Laurent Sagart, Roger Blench & Alicia Sanchez–Mazas, eds. *The Peopling of East Asia: Putting together archaeology, linguistics and genetics*. London: Routledge Curzon, pp. 107–131.

Richards, M. B. et al. (2008). Climate change and postglacial human dispersals in Southeast Asia. *Molecular Biology and Evolution*, 25, pp. 1209–1218.

Sagart, L. (2002). Sino–Tibeto–Austronesian: An updated and improved argument. (PDF) Paper presented at The *Ninth International Conference on Austronesian Linguistics (ICAL9)*. 8–11 January 2002. Canberra, Australia.

----- (2004). The higher phylogeny of Austronesian and the position of Tai–Kadai, *Oceanic Linguistics*, 43, pp. 411–440.

Schmidt, W. (1906). Die Mon—Khmer—Völker, ein Bindeglied zwischen Völkern Zentralasiens und Austronesiens [The Mon–Khmer Peoples, a Link between the Peoples of Central Asia and Austronesia]. *Archiv für Anthropologie, Braunschweig, new series*, 5, pp. 59–109.

The Genographic Project (1999–2004). Retrieved from <https://genographic.nationalgeographic.com/genographic/atlas.html>.

The HUGO Pan–Asian SNP Consortium. (2009). Mapping human genetic diversity in Asia, *Science*, 326 (5959), pp. 1541–1545. Retrieved from <http://www.sciencemag.org/content/326/5959/1541.short>

Thurgood, G. (1994). Tai–Kadai and Austronesian: The nature of the historical relationship. *Oceanic Linguistics*, 33 (2), pp. 345–368.

William, M. (1996). Defining the hundred Yue, *Bulletin of the Indo–Pacific Prehistory Association*, 15, pp. 93–100.

Wei, X.W. (1982). On the language of the Hundred Yue people, In *Papers on the History of the Hundred Yue People*, Beijing: Chinese Social Sciences Press (in Chinese).

Xing, G.W. 刑公畹 (1999).《汉台语比较手册》, 商务印书馆.

Y–DNA Human Migration. (n.d.). Retrieved from http://www.kerchner.com/images/dna/ydnav_migrationmap_%28FTDNA2006%29.jpg

Zhengzhang, S.F. (1990). Some Kam–Tai words in place names of the ancient states of We and Yue. *Minzu Yuwen*, 6, pp. 16–18 (in Chinese).

----- (1991). Decipherment of Yue–Ren–Ge (Song of the Yue boatman), *Cahiers de Linguistique—Asie Orientale*, 20 (2), pp. 159–168.