On the Origins of Sinitic

Scott DeLancey
University of Oregon

A striking fact about the Sinitic branch of Sino-Tibetan is that, while its lexicon, phonological structure and some reconstructible morphology clearly link it genetically to Tibeto-Burman, its basic morphosyntactic profile is the isolating SVO type characteristic of mainland Southeast Asia rather than the agglutinating SOV structure characteristic of Tibeto-Burman. This results from the history of Sinitic as a more typical Tibeto-Burman SOV language which was came to be used as a lingua franca by the Tai-Kadai, Hmong-Mien, and Austroasiatic "Bai Yue" populations of the south even before their incorporation into the Chinese Empire. The resulting creoloid syntax (Ansaldo and Matthews 2001) has remained robust in China and mainland Southeast Asia.

Evidence for early contact with Tai-Kadai languages is shared vocabulary (representing loans in both directions) in Proto-Tai-Kadai and Sinitic, matching in tone class. The oldest stratum is of Old Chinese date or older (Li 1945; Manomaivibool 1975), which implies that it dates from first contact between Sinitic and TK. Assuming Ostipirat's (2000) reconstruction of PTK to 4,000 years ago, the date correlates with Xia. Presumably Sinitic was still Tibeto-Burman-like, and TK had yet to diverge dramatically from Austronesian (Sagart 2005). The dramatic creolization which produced the grammatical structure of Sinitic, Tai, Kadai, and other mainland Southeast Asian languages originated in this original contact.

1   The problem

The essential problem in the formation of Chinese is that it has strong lexical, phonological, and grammatical connections both with the Tibeto-Burman languages to the west and with the Southeast Asian languages to the south, especially Tai-Kadai, but also Hmong-Mien and Mon-Khmer. While basic vocabulary and some reconstructible morphology clearly link Sinitic genetically to Tibeto-Burman, its basic morphosyntactic profile is the isolating SVO type characteristic of mainland Southeast Asia rather than the agglutinating SOV structure characteristic of Tibeto-Burman. This occasionally even leads to doubts about the genetic relationship of Sinitic and Tibeto-Burman (Beckwith

2002, 2006), but few linguists doubt that the history of Sinitic is of a Tibeto-Burman language, with the SOV syntax characteristic of that family, which was adopted by a population speaking Kadai, Austronesian and/or Austroasiatic, and quite possibly other, languages[1].

The evidence which requires explanation falls into four broad categories: lexical correspondences among Chinese and one or more other languages or families, morphological correspondences between Chinese and Tibeto-Burman, and the striking similarities in both syntactic and phonological structure between Chinese and the mainland Southeast Asian families. The difficulty is that there is significant evidence linking Chinese with several different language groups, including Tai-Kadai, Hmong-Mien, Austroasiatic, and Austronesian, but it cannot be genetically related to all or even several of them. Most of what Chinese shares with most of these languages must thus have resulted from language contact. The fundamental problem of Sinitic linguistics is to unravel the various linguistic threads which make up Old Chinese and its predecessors and understand how they came to be woven together into the language which we know.

The idea that many of the features of Old Chinese came about as a result of contact between the language of Shang (and/or possibly Zhou) conquerors and an indigenous population speaking a language or languages related to those currently spoken in mainland Southeast Asia is old (Terrien de la Couperie 1887). My purpose in this paper is to propose a more explicit model of this contact, and to argue that it offers a path toward solutions to several long-standing problems. Little of the content of this paper is novel. In sections 2 and 3 I will briefly summarize the kinds of evidence which a hypothesis of Sinitic origins needs to account for. In Section 4 I will suggest a hypothesis which I hope will account for at least some of it, invoking a model of language contact and interaction which has not really been applied previously to the problem of Chinese.

2   Lexical and phonological correlations

Chinese shows striking points of correspondence with Tai-Kadai, Hmong-Mien, and Vietnamese in both lexicon and phonological structure. In this section I will very briefly review some of the reasons why these congruences do not argue for a genetic relationship among any of these languages.

---

[1] We do not need to deal here with the status of Tibeto-Burman as a genetic unit, i.e the question of whether Sinitic represents one of two primary branches of the family, is a branch coordinate with higher-level branches like Bodic, or, as van Driem (1997, 2008) suggests, is a actually subordinate unit. However, the conventional view of the family as Sinitic + Tibeto-Burman is based on traits which Sinitic shares with the Bai Yue rather than the Tibeto-Burman languages. Therefore the explanation offered here for those correlations can be seen as undermining the basis of that view.

2.1 Issues in Lexical Comparison

Since the earliest days of serious linguistic study of Chinese, scholars have noted the huge amount of vocabulary shared between Chinese and neighboring languages. A great deal of this was obviously borrowed from Chinese, which throughout historic times has been the major cultural force in East Asia. But there is also a very substantial body of vocabulary shared with Kadai, Hmong-Mien, and/or Vietnamese, which is much older than that, and it is not easy to determine whether such shared forms are common inheritance or borrowing, and in the latter (more likely) case, borrowing in what direction.

Thus, aside from the Tibeto-Burman languages (Matisoff 2003), Chinese has been linked with Tai-Kadai (Wulff 1934, Nishida 1975, FK Li 1945, 1976, Manomaivibool 1975, 1976a, b, inter alia), Austroasiatic (Norman and Mei 1976), Hmong-Mien (Downer 1963, 1971, Wang 1986, Haudricourt and Strecker 1991), and Austronesian (Sagart 1994, 1995, 1999).[2] On the one hand, all of these proposals are supported by serious lexical comparisons, and some sort of historical connection with Tai-Kadai, Hmong-Mien and Vietic within Austroasiatic is established both by shared lexicon and by the astonishing correspondence in phonological typology (see below). But these three groups are not evidently related, so Sinitic can hardly be genetically related to all of them, much less to all of them and Tibeto-Burman as well, except at some very hypothetical, very high, unrecoverable level. Thus it has long been clear that some of the evidence which has been adduced to argue for genetic relations among these languages in fact reflects sustained intense contact among unrelated languages (Terrien de la Couperie 1887, Matisoff 1973, LaPolla 2001, inter alia).

The interpretation of the lexical evidence has sometimes been confused by unrealistic notions of when and how borrowing can take place, in particular simplistic ideas that lexical borrowing only occurs from a more dominant into a smaller and less "advanced" population, or under some kind of necessity. For example, Manomaivibool (1975:364), discussion shared Tai-Chinese vocabulary, says "It seems implausible that Tai *had to* borrow that many items of such common vocabulary from Chinese" (emphasis added). But without a more thorough understanding of the social, economic, and political context of Shang and Zhou era China, it is simply impossible to distinguish borrowings from cognates purely on how easily one can imagine a motivation for borrowing a word with a particular meaning.[3] I will suggest a model of language contact which makes room for exactly the sort of unsystematic lexical mixture which we find in Chinese.

---

[2] I ignore more far-fetched suggestions of connections to Uralic, Indo-European, North Caucasian, and Na-Dene.

[3] Consider the following French borrowings into English: *family, dinner, supper, soup, easy*, *difficult*, *quiet*, *silent*, *noisy*, *lake, river, mountain, valley, forest, marsh, flower, village, city, language, story, color, attack, defend, protect, accept, argue, agree, beautiful, stupid, count, real, false, very*, *front.* It would be hard to argue that any of these, or thousands of similar items, "had" to be borrowed for cultural reasons.

It has been suggested (FK Li 1976, Manomaivibool 1975, 1976a, b, Nishida 1975, 1976) that if a Sino-Tai form can be reconstructed for Proto-Tai-Kadai, this is evidence for genetic relationship between Sinitic and T-K, presumably on the grounds that PTK is too old to have been contemporary with any stage of Chinese, so that there would be no time at which borrowing could have taken place. But there is no logic to this argument – whether we imagine the common vocabulary to reflect a common proto-language or to represent borrowings, in either case PT-K or something ancestral to it, and Old Chinese or something ancestral to it, must have been contemporaneous. Noting this fact does not constitute an argument for one hypothesis or the other. What is important is that Li and other scholars consider the oldest layer of shared Tai-Chinese vocabulary (which certainly represents loans in both directions, not only from Chinese to Tai) to be of at leats Old Chinese date, so that this common lexicon probably dates from the earliest contact.

2.2 The Southeast Asian phonological profile

The most impressive correspondence between Sinitic and the Southeast Asian Tai-Kadai, Hmong-Mien, and Vietic languages is in their phonological structure. All share the stereotypical monosyllabic morpheme structure and elaborate tone systems. The most striking, and puzzling, fact about this congruence is the perfect correspondence of the tone systems (Wulff 1934, Haudricourt 1954a, b, FK Li 1945, 1976, Matisoff 1973, Ostapirat 2000, Ratliff 2010). Sinitic, Tai-Kadai, Hmong-Mien, and Vietnamese all have a four-tone system, with a three-way distinction on "smooth", i.e. open or sonorant-final syllables, and all "checked", i.e. obstruent-final, syllables manifesting a distinct fourth tone. Each of the other three shares with Sinitic (and to some extent with each other) a substantial body of shared vocabulary which shows regular correspondence in tone class. In all of the languages tones originated out of final laryngeal features, so that the original correspondence is in the type of rime: obstruent coda, coda *-h (sometimes < *-s), final *-?, and "smooth" syllables with none of these (Haudricourt 1954a, b, 1961/1972, Mei 1970, 1980). The shared vocabulary which shows these correspondences must have been borrowed[4] at a stage when both the donor and recipient languages still retained these final laryngeal distinctions, and had not yet developed phonemic tone; if we imagine that these items were borrowed with phonemic tone, it becomes impossible to explain the regularity of the correspondences. (For a very clear exposition of this argument see Ratliff 2010:187-93). The languages must have still been in close contact when they underwent a shared tonogenetic episode in which these laryngeal distinctions were reinterpreted as tonal, as they were still centuries later when they all shared in the "Great Tone Split" conditioned by mergers of initial consonant series.

---

[4] Even on the hypothesis that some or all of these languages might be genetically related at some level, there is no question that a great deal of the most obvious shared lexicon represents borrowing in one direction or the other.

The monosyllabic pattern is not really characteristic of Austroasiatic, or even of Kradai, and the Sinitic developments do have parallels in the phonological development of other Sino-Tibetan groups. So Sagart is probably right in attributing the original locus of monosyllabic structure to Chinese:

> From a typological point of view, Old Chinese was more similar to modern East Asian languages like Gyarong, Khmer or Atayal than to its daughter language Middle Chinese: its morphemes were nontonal and not strictly monosyllabic; its morphology was essentially derivational, and largely prefixing; but it also made use of infixes and suffixes. At some point between Old Chinese and Middle Chinese, and for unknown reasons, a cascade of changes caused the language to move away from this model. Its affixing morphology began to freeze; its loosely attached prefixes were lost, while other affixes clustered with root segments and were reinterpreted as root material. A new morphemic canon tending toward strict monosyllabism, with a great variety of initial and final consonant clusters, emerged. Further shifts saw the reduction of initial clusters, this resulting in a more complex inventory of initial consonants, and in new vowel contrasts. Final clusters were also reduced and the inventory of final consonants restricted to resonants and stops, this leading to the emergence of tones. Thus the classical 'Indochinese' typology common in its major features to Middle Chinese, Vietnamese, Miao-Yao, Tai, Burmese etc., was born. (Sagart 1999:13)

3 Grammatical evidence for Sino-Tibetan

The strongest evidence for the genetic affiliation of Sinitic with the Tibeto-Burman languages is grammatical, specifically correspondences in personal pronouns and in some reconstructable verb morphology.

3.1 The pronouns

An important argument for the Sino-Tibetan affiliation of Chinese has always been the correspondence of the 1st and 2nd person pronominal roots. We will look here at the 1st person forms 吾 (*nga* in Baxter's reconstruction) and 我 (*ngaX < *ngaj?*), which neatly match the reconstructed PTB root *ηa (Matisoff 2003).[5] Sagart has argued on the basis of its late appearance in Shang and Zhou inscriptions that the 1st person *ηa root is a secondary development in Chinese (1999:142-4), and a late borrowing from there into the rest of Tibeto-Burman (145-6). Instead of the well-established *ηa, Sagart proposes that the PTB 1st person pronoun was the stop-initial *ka* which occurs as the primary 1st person

---

[5] In fact the OC forms very specifically resemble the Tibetan 1st person pronominals *nga* and *nged*.

root in three geographically marginal branches of the family, northern Qiangic, Kuki-Chin, and "a few languages of eastern Nepal and neighboring areas", i.e. Kiranti in Nepal and Western Himalayan in northwest India. He suggests that this distribution reflects a spread of *ŋa, ultimately from Chinese, through the contiguous central TB area, leaving only the few branches on the edges of the TB area untouched. These three (actually four, plus a few strays in Nepal and Arunachal Pradesh) then retain what Sagart takes to be the original Sino-Tibetan 1st person root *ka.

This proposal cannot be correct, since the nasal root is found in the 1st person agreement suffix which is reconstructable for Proto-Tibeto-Burman (Bauman 1975, Sun 1983, van Driem 1993, DeLancey 1989, 2010, inter alia), and thus long predates its first appearance in the Chinese inscriptions.[6] Since a late borrowing from a Tibeto-Burman source into Chinese does not seem likely here, we have to recognize this root as dating back to their common ancestor. But it is likely that the *ka root may also be ancient; it now appears that it was a possessive or oblique form contrasting with the nominative *ŋa (Jacques 2007, DeLancey 2011b). What we see in the languages where this form has replaced the original nominative *ŋa, is the replacement of the original finite construction with an innovative finite form based on a nominalization, which thus takes a genitive rather than a nominative "subject" (DeLancey 2011c).

3.2   Ancient morphology

While we find no inflectional morphology recorded in any form of Chinese, the fossils of pre-Chinese prefixes and suffixes can be found in the phonological alternations of semantically and graphically related words. Two morphological constructions which are securely reconstructible for both Tibeto-Burman and Sinitic are a causative prefix *s- (Conrady 1896, Mei 1980, 2008, Dai 2001) and a nominalizing *-s suffix (Downer 1959, Forrest 1960, Mei 1980, Mazo 2002).

The *s- causative is retained in Written Tibetan and a handful of other languages, though in many it is no longer productive:

Tibetan	*log* 'return (intransitive)', *slog* 'turn (transitive)'
Boro	*gab* 'cry', *səgab* 'make s.o. cry'
Tarong	*ip*^{55} 'sleep', *səip* 'cause to sleep'

In most modern TB languages, we find the prefix reflected in devoicing of the initial consonant:

Tibetan	*nub* 'sink', *snub* 'destroy, abolish'
Zaiwa	*nop* 'sink in mud', *n̥op* 'make s.t. get bogged in mud'

---

[6] The *ka 1st person root which Sagart suggests is the original Tibeto-Burman form is in fact old, but as an oblique alternant to the nasal root (Jacques 2007, DeLancey 2011).

Boro   *gi* 'afraid', *si-gi* 'frighten'
Newar   *gya-* 'afraid', *khya-* 'frighten'

And we find the same in Old Chinese (Mei 2008):

见 xiàn 'be visible'   < *gians
见 jiàn 'see'   < *kians < *s-k < *s-g

别 píe 'leave, separate (intr.)'   < *bjät < *brjat
别 pìe 'discriminate, distinguish'   < *pjät < * prjat < *s-p < *s-b

墨 mò 'ink'   < *mək
黑 hēi 'black'   < *s-mək

Dai (2001) demonstrates that this construction is ancient in Tibeto-Burman; based on this and the abundant evidence for it in Chinese, Mei (2008) suggests that it is a defining feature of Sino-Tibetan languages; note that it is preserved, at least in fossil form, in some branches which have lost almost all other inherited morphology, e.g. Bodo-Garo.
     The nominalizing *-s* suffix is reflected in Old Chinese tonal alternations (Downer 1959, Forrest 1960); which correspond neatly to the suffix which is preserved in Written Tibetan (Mei 1980):

Chinese   Tibetan

量 liáng 'to measure'   < *liaŋ   'grang 'to count'
量 liàng 'a measure'   < *liaŋs   grangs 'a number'

织 zhī 'to weave'   < *tjək   'thag 'to weave'
织 zhì 'woven goods'   < *tjəks   thags 'web, woven stuff'

Morphological comparisons like this are the sine qua non of comparative linguistics, and without some strong argument discrediting comparisons like these, we can take this evidence as conclusively establishing the genetic relationship of Sinitic with the rest of Sino-Tibetan.

3.3   The South East Asian syntactic profile
Mainland Southeast Asia is well-known for its striking areal linguistic typology, characterized both by the elaborate and congruent tone systems discussed above (this is not shared by most Mon-Khmer languages) and by radically isolating SVO morphosyntax. Indeed the examples put forward to illustrate isolating typology are

always languages from this area; aside from modern European-based creole languages, few if any other languages in the world are as resolutely free of any sort of inflectional morphology. In this respect Chinese clearly sorts with the Southeast Asian rather than the Tibeto-Burman languages, which are characteristically agglutinative, SOV, and often morphological very complex.

There is no serious question that the Southeast Asian syntactic profile in Chinese is a secondary development:

> From the fact that we can clearly see changes in the word order of these three languages [Sinitic, Karen, and Bai] over time, and cannot see such changes in the Tibeto-Burman languages other than Bai and Karen, we assume that it was Bai, Karen and Chinese that changed rather than all the other Tibeto-Burman languages. (LaPolla 2003:28)

A few scholars see this development as internal to Sinitic:

> The new linguistic standard of the Han dynasty … typologically characterized by its incipient isolating morphology, and its emergent tonal and monosyllabic phonology, gradually spread to all parts of the empire, north and south, and this same typology further spread to all non-Chinese languages spoken in territories under Chinese rule after the Han: all of Miao-Yao, Viet-Muong (but not the rest of Mon-Khmer), all of Kam-Tai, some south-eastern Tibeto-Burman languages including Lolo-Burmese (but not Tibetan, Qiang, Gyarong, etc.). (Sagart 1999:8)

But most scholars, from Terrien de la Couperie on, see the shift in Sinitic as due to influence from neighboring languages to the south; Egerod (1976:59) points out that since SVO order is inherited in Thai, "Chinese was largely a recipient rather than a donor in the early times … it is Chinese which borrows a new word order"[7] (see also Benedict 1976). Indeed, all of the Southeast Asian groups have SVO syntax as far back as we can trace. And there are ample traces of earlier SOV patterning in Old Chinese (Cheng 1983) and even in modern Mandarin, as detailed in Wu Fuxiang's contribution to this conference (2011).

4.   The Formation of Chinese

There is no question that the formation of Chinese involved contact with neighboring languages, definitely including Austroasiatic, Tai-Kadai, and Hmong-Mien, and very possibly others which have completely disappeared. There has been a certain tension on

---

[7] Egerod links this claim to a putative shift in transitivity type, but this part of his argument is unnecessary; it is sufficient that a Bai Yue-type substratum was present to contribute the basic grammatical pattern.

the question of what sorts of contact might be involved. Traditionally there seem to be two basic possibilities: contact between adjacent languages, i.e. imagining Proto-Sinitic, Proto-Hmong-Mien, etc. as spoken in adjacent states, or super-substratum influence, i.e. an "elite dominance" model in which Proto-Sinitic formed in a state consisting of immigrant Tibeto-Burman conquerors interacting with indigenous Proto-Tai-Kadai or Proto-Hmong-Mien subjects. Both of these have important deficiencies; in this section I will develop an alternative model which is better suited to explain the kinds of data we have been considering.

The territory where Sinitic languages are spoken was an area of substantial linguistic diversity from prehistoric times (Terrien de la Couperie 1887, Pulleyblank 1983, 1995, Ballard 1984, LaPolla 2001, inter alia). We are particularly concerned with the "Bai Yue" 百越 languages, which seem to have been of mixed provenance, including both Austroasiatic and pre-Kadai and pre-Hmong-Mien languages (JZ Li 1994, Meacham 1996, LaPolla 2001). The Yue people and kingdom to the south are a long-term presence in Chinese history, but the first explicit reference to the Hundred Yue is in the Qin era *Annals of Lü Buwei*:

> For the most part, there are no rulers to the south of the Yang and Han rivers, in the confederation of the Hundred Yue tribes [lit. 百越之际 'among the Hundred Yue'], in the territories of Bikaizhu, Fufeng, and Yumi, and in the states of Fulou, Yangyu, and Huandou. (Lü et. al. 2000:112 / Book 20/1.3)

This term is important because it makes clear that the reference of Yue is multiethnic (Luo 1990:268):

> Leaving aside the Austronesian question, it seems highly likely that the peoples called Yue at various times by the Han Chinese spoke Austroasiatic languages, early forms of Hmong/Mien, Tai-Kadai languages and perhaps languages in other families now extinct. (Meacham 1996:98)

The question is, how does Chinese come to share large bodies of vocabulary, and characteristic phonological and morphosyntactic typological profiles, with these languages?

The similarity of the isolating Chinese-Southeast Asian morphosyntactic profile to creole languages has been noted for some time[8]. But history gives us no reason to suppose that Chinese at any stage was ever a true creole, in the traditional sense of a language which develops from a grammarless pidgin. Ansaldo and Matthews (2001)

---

[8] I don't know who first made this observation; I heard it in the 1970's from David Strecker.

suggest the term "creoloid" to refer to such a language, which shows the constellation of typological features traditionally associated with creole languages but does not have the history of a creole:

> Perhaps a more constructive way to see the "prototypical creole" traits is that languages which have been subject to intensive contact involving several typologically distant varieties will tend to show some combinations (or subset) of these features. (Ansaldo & Matthews 2001:317)

Such languages arise in conditions of intense contact, when for whatever reason some significant portion of the language community are second-language rather than native speakers (McWhorter 2007). This kind of development has occurred, and continues to occur, repeatedly in Tibeto-Burman (DeLancey 2010a, to appear a, b), and it is clear that Sinitic has the same kind of history.

The broad account which I suggest here is the familiar picture of a contact situation between western invaders speaking a TB tongue and locals speaking languages affiliated with one or more of the attested mainland Southeast Asian stocks. But it is not enough to simply say "contact" and pretend that we have explained anything. In this view of Sinitic we have a very specific outcome, with Sino-Tibetan lexical and grammatical core, heavy Bai Yue, especially Kadai, lexical influence, creoloid syntax based more on Bai Yue than on Sino-Tibetan patterns, and innovative phonological structure. This did not come about through people overhearing each other's languages on market day, or learning a few phrases for doing business; we have to imagine a situation of widespread bi- or multilingualism. This would be the case in a scenario in which Chinese or pre-Chinese speakers conquered a Bai Yue population, as happened as the kingdoms of Chu and then Yue were incorporated into Qin China. But this does not automatically explain the extent of the influence which we find on the whole language. Ballard's (1984) "Mother Soup" metaphor captures the problem but doesn't solve it. More importantly, the most important contact evidence predates the assimilation of the southern kingdoms into imperial China.

Instead, I propose that the features which so dramatically distinguish Sinitic from other Tibeto-Burman branches reflect the use of Proto-Sinitic as a lingua franca, used widely by non-Chinese (by whatever definition) outside of the actual administrative control of the Chinese state. As we have noted, the term Bai Yue refers to the multiethnic and multilingual situation in the south. One can imagine the utility of a vehicular lingua franca even without reference to the Chinese state and its influence; by the time the Chinese state is present on the historical stage, some version of its language would be a likely candidate. Thus, with the increasing power and prestige of Zhou, perhaps even Shang, China, a pidiginized version of its Tibeto-Burman language became a lingua franca throughout the region. Cheng (1983) speaks of "two sublanguages

coexisting in early archaic Chinese", an earlier SOV stratum and an innovative SVO syntax. This would, essentially, be "pure" Sino-Tibetan Chinese with SOV syntax, and innovative "foreigner" Chinese, spoken with the SVO pattern of the Bai Yue languages. Ultimately the widespread lingua franca version of Proto-Sinitic replaced the original everywhere.

REFERENCES

CAAAL = Computational Analysis of African and Asian Languages
ICSTLL = International Conference on Sino-Tibetan Languages and Linguistics
LTBA = Linguistics of the Tibeto-Burman Area

Ansaldo, Umberto, and Stephen Matthews. 2001. Typical creoles and simple languages: The case of Sinitic. Linguistic Typology 5.2–3:311–326.
Ballard, W.L. 1984. The Mother Soup: A South China recipe for tonometamorphogenesis. CAAAL 22:43-70.
Beckwith, Cristopher. 2002. The Sino-Tibetan problem. Medieval Tibeto-Burman Languages, ed. by C. Beckwith,113-57. Leiden: Brill.
_____. 2006. Introduction: Toward a Tibeto-Burman theory. Medieval Tibeto-Burman Languages II, , ed. by C. Beckwith, 1-38. Leiden: Brill.
Benedict, Paul. 1976. Sino-Tibetan: Another look. Journal of the American Oriental Society 96.2:167-197.
Cheng, Tsai Fa. 1983. SOV sentence patterns in Early Archaic Chinese.16th ICSTLL, Seattle.
Conrady, August. 1896. Eine indochinesische causative-denominativ-Bildung und ihr zusammenhang mit den Tonaccenten. Leipzig: Harrasowitz.
Dai, Qingxia. 2001. Zang-Mian yu zu yuyan shidong fanchou de lishi bianhua (The historical development of the causative category in Tibeto-Burmese languages). Journal of Chinese Linguistics 29.1: 1-10.
DeLancey, Scott. 2010a. Language replacement and the spread of Tibeto-Burman. Journal of the Southeast Asian Linguistics Society 3.1: 40-55.
_____. 2010b. Towards a history of verb agreement in Tibeto-Burman. Himalayan Linguistics Journal 9.1: 1-38.
_____. 2011a. Nocte and Jinghpaw: Morphological correspondences. North East Indian Linguistics 3, ed. by Gwendolyn Hyslop, S. Morey, and M. Post, 61-75. New Delhi: Foundation/Cambridge University Press India.
_____. 2011b. Agreement prefixes in Tibeto-Burman. Himalayan Linguistics Journal 10.1:1-35.
_____. 2011c. Finite structures from clausal nominalization in Tibeto-Burman languages. Nominalization in Asian Languages: Diachronic and Typological Perspectives,

ed. by Foong Ha Yap, K. Grunow-Hårsta & J. Wrona, 343-62 [Typological Studies in Language 96]. Philadelphia: Benjamins.

_____. to appear a. On the origins of Bodo-Garo. North East Indian Linguistics 4, ed. by Gwendolyn Hyslop, S. Morey, and M. Post. New Delhi: Foundation/Cambridge University Press India.

_____. to appear b. Creolization in the divergence of Tibeto-Burman. Papers from the 16th Himalayan Languages Symposium, ed. by Nathan Hill.

Downer, Gordon. 1959. Derivation by Tone-Change in Classical Chinese. Bulletin of the School of Oriental and African Studies 22.1/3:258-290.

_____. 1963. Chinese, Thai, and Miao-Yao. Linguistic Comparison in South East Asia and the Pacific, ed. by H.L. Shorto, 133-139.

_____. 1971. The further relationships of the Miao-Yao languages. 4th ICSTLL, Bloomington.

Driem, George van. 1993. The Proto-Tibeto-Burman verbal agreement system. Bulletin of the School of Oriental and African Studies 56.2: 292-334.

_____. 1997. Sino-Bodic. *BSOAS* 60.3: 455-488.

_____. 2008. To which language family does Chinese belong, or what's in a name. Past Human Migrations in East Asia: Matching Archaeology, Linguistics and Genetics, ed. by A. Sanchez-Mazas et.al., 219-53.

Egerod, Søren. 1976. Benedict's Austro-Tai hypothesis and the traditional view on Sino-Thai relationship. CAAAL 6:51-59.

Forrest, R. A. D. 1960. Les occlusives finales en chinois archaïque. Bulletin de la Société Linguistique de Paris 55:228-239.

Hashimoto, Mantaro. 1976c. Genetic Relationship, Diffusion and Typological Similarities of East & Southeast Asian Languages: Papers for the 1st Japan-US Joint Seminar on East & Southeast Asian Linguistics. Tokyo: Japan Society for the Promotion of Science.

Haudricourt, André. 1954a. De l'origine des tons en vietnamien. Journal Asiatique 242:69-82.

_____. 1954b. Comment reconstruire le chinois archaïque. Word 10.2-3:351-64.

_____. 1961. Bipartition et tripartion des systèmes de tons dans quelques langues d'Extrême-Orient. Bulletin de la Société de Linguistique de Paris56: 163-80.

_____. 1972. Two-way and three-way splitting of tonal systems in some Far Eastern languages. Tai phonetics and phonology, ed. by Jimmy Harris and Richard Noss, 58-86. Bangkok: Central Institute of English Language. (English translation of Haudricourt 1961).

_____, and David Strecker. 1991. Hmong-Mien (Miao-Yao) loans in Chinese. T'oung Pao n.s. 77.4-5: 335-41.

Henderson, Eugenie. 1965. The topography of certain phonetic and morphological characteristics of Southeast Asian languages. Lingua 15. 400-434.

Jacques, Guillaume. 2007. A shared suppletive pattern in the pronominal systems of Chang Naga and Southern Qiang. Cahiers de linguistique Asie Orientale 36.1:61-78.

LaPolla, Randy. 2001. The role of migration and language contact in the development of the Sino-Tibetan language family. Areal Diffusion and Genetic Inheritance, ed. by Alexandra Aikhenvald and R.M.W. Dixon. Oxford: Oxford University Press.

_____. 2003. Overview of Sino-Tibetan morphosyntax. Thurgood and LaPolla (eds.), 22-42.

Li, Fang-kuei. 1945. Some old Chinese loan words in the Tai languages. Harvard Journal of Asiatic Studies 8: 333-42.

_____. 1971. Shang gu yin yanjiu. Tsing Hua Journal of Chinese Studies 9.1-2: 1-61.

_____. 1976. Sino-Tai. Hashimoto (ed.), 230-39.

_____. 1983. Archaic Chinese. Keightley (ed.), 393-408.

Li, Jingzhong. 1994. Yuyan yanbian lun. Guangzhou: Guangzhou Chubanshe.

Lü, Buwei, John Knoblock and Jeffrey Riegel. 2000. The Annals of Lü Buwei: A Complete Translation and Study. Stanford: Stanford University Press.

Luo, Xianglin. 1990. Zhongguo Minzu Shi. Taibei : Zhonghua Wenhua Chuban Shi Ye She.

Manomaivibool, Prapin. 1975. A Study of Sino-Thai Lexical Correspondences. Ph.D. dissertation, University of Washington.

_____. 1976a. Thai and Chinese – are they genetically related? CAAAL 6:11-31.

_____. 1976b. Layers of Chinese loanwords in Thai. Tai linguistics in honor of Fang-Kuei Li, , ed. by Thomas Gething, J. Harris and P. Kullavanijaya, 179-184. Bangkok: Chulalongkorn University Press.

Matisoff, James. 1973. Tonogenesis in Southeast Asia. Consonant Types and Tone, ed. by Larry Hyman, 73-95. (Southern California Occasional Papers in Linguistics 1). Los Angeles: Linguistics Program, University of California Los Angeles.

_____. 2003. Handbook of Tibeto-Burman: System and philosophy of Sino-Tibetan reconstruction. [University of California Publications in Linguistics 135]. Berkeley; Los Angeles: University of California Press.

Mazo, Olga. 2002. Reflexes of the suffix *-s in Classical Tibetan and Old Chinese. *Tibet, Past and Present*, ed. by Henk Blezer, 455-472. Leiden: Brill.

McWhorter, John. 2007. Language Interrupted: Signs of Non-Native Acquisition in Standard Language Grammars. Oxford: Oxford University Press.

Meacham, William. 1996. Defining the Hundred Yue. Bulletin of the Indo-Pacific Prehistory Association 15: 93–100.

Mei, Tsu-lin. 1970. Tones and prosody in Middle Chinese and the origin of the rising tone. Harvard Journal of Asiatic Studies 30: 86-110.

_____. 1980. Sisheng bieyi zhong de shijian cengci. Zhongguo Yuwen 1980.6: 427-43.

    . 2008. Shanggu Hanyu dongci zhuoqing bieyi de laiyuan (The original differences between verbs with voiced and voiceless initials in Ancient Chinese). Minzu Yuwen 2008.3:3-19.

Nishida, Tatsuo. 1975. Common Tai and Archaic Chinese. Studia Phonologica 9:1-12.

    . 1976. Some problems of the morpheme stock in Sino-Tibetan: A preliminary observation. Hashimoto, (ed.), 30-38.

Norman, Jerry, and Mei Tsulin. 1976. The Austroasiatics in ancient south China: Some lexical evidence. Monumenta Serica 32:274-301.

Ostapirat, Weera. 2000. Proto-Kra. LTBA 23.1:1-251.

Pulleyblank, E. G. 1983. The Chinese and their neighbors in prehistoric and early historic times. Keightley (ed.), 411-466.

    . 1995. Historical and prehistorical relationships of Chinese. The Ancestry of the Chinese Language, ed. by Willam S-Y. Wang, 145-94. [Journal of Chinese Linguistics Monograph Series no. 8]. Berkeley: Project on Linguistic Analysis, University of California.

Ratliff, Martha. 2010. Hmong-Mien Language History. (Studies in Language Change 8). Canberra: Pacific Linguistics.

Sagart, Laurent. 1994a. Proto-Austronesian and Old Chinese evidence for Sino-Austronesian. Oceanic Linguistics 33.2:271-308.

    . 1994b. Questions of method in Chinese-Tibeto-Burman comparison. 27th ICSTLL, Paris.

    . 1995. Some remarks on the ancestry of Chinese. Wang (ed.) 195-223.

    . 1999. The Roots of Old Chinese. Philadelphia: Benjamins.

Terrien de la Couperie. 1887. The Languages of China Before the Chinese. London: David Nutt.

Wang, Fushi. 1986. A preliminary investigation of the genetic affiliation of the Miao-Yao languages. International Symposium on the Minority Nationalities of China, Santa Barbara.

Wu, Anqi. 2008. Shiqian huanan diqu de yuyan anchu. Minzu Yuwen 2008.3:21-35.

Wu, Fuxiang. 2011. Hanyu zhong jige you tese de jyushi (Language-specific syntactic constructions in Chinese). presented at the 23rd North American Conference on Chinese Linguistics, Eugene, OR, June 18, 2011.

Wulff, Kurt. 1934. Chinesisch und Tai: Sprachvergleichende Untersuchungen. (DanskeVidenskabernes Selskab, Historisk-filologiske Meddelelser 20.3). København: Levin & Munksgaard.