

FROM MALAYIC TO SINITIC: THE RESTRUCTURING OF TSAT UNDER INTENSE CONTACT

Graham Thurgood

California State University, Chico

<*gthurgood@csuchico.edu*>

Fengxiang “Frank” Li

California State University, Chico

<*FLi@csuchico.edu*>

1 Background

Hainan Cham [=Tsat], an Austronesian language of Hainan, is one of the clearest examples in the literature of a language restructuring under intense contact. Some two thousand years ago, traders speaking a Malayo-Chamic language set up trading posts on the coast of modern day Vietnam. Interaction and intermarriage with speakers of Bahnar led to the total restructuring of the language; under intense contact with Bahnar, it became Chamic, a language that differed strikingly from its Malayo-Chamic ancestor in phonology, morphology, lexicon, and syntax (constructions). For roughly a thousand years, this newly restructured Chamic language—the language of the Champa Federation—existed as an only moderately differentiated dialect continuum along the coastline of southern Vietnam, with a small trading post on Hainan Island.

The next major restructuring occurred after the northern capital fell to the Vietnamese in 982. This event led the Northern Cham to split into two groups: the bulk of the merchant class (including many Muslims—Huihui, in Chinese) fled to Hainan becoming the *Utsat* (etymologically, *u* ‘people classifier’ + *Tsat*, **cam* ‘Cham’; note that, a thousand years ago, all Chamic speakers were more than likely called Chams), while the bulk of the non-merchant class fled to south and, in many cases, inland, becoming the Northern Roglai (etymologically, *ra* ‘people’+ *glai* ‘forest’). It is worth noting that the Northern Roglai are the Chamic group reputed to have the royal treasures from the northern capital.

The arrival of the Northern Cham traders on Hainan—an event noted in the Chinese Dynastic records—led to another complete restructuring of the language, this time under the influence of the monosyllabic and tonal Hlai languages (Tai-Kadai), the monosyllabic and tonal Min dialects of Chinese, and, more recently, under the quite intense influence of Mandarin.

2 Restructuring the Phonology

The phonological restructuring was significant, although the initial steps were simple enough. Malayo-Chamic had penultimate stress, but like some dialects of modern Malay, when the penultimate vowel was shwa, it more than likely had final stress. Under influence from Bahnar speakers (and possibly other Mon-Khmer groups), stress switched to final position. This, combined with continued interaction with Bahnar speakers whose languages

were sesquisyllabic (weaker presyllable + stressed final syllable; in terms of stress, iambic) led to the change from disyllabic to sesquisyllabic forms.

While the four unstressed first syllable vowels remained unchanged, the stressed main syllable vowels proliferated. The inherited vowels went from seven to nine (four monophthongs and three diphthongs), with the splitting of the two high vowels into a diphthongized stressed variant and an unstressed, undiphthongized short variant. A number of vowels were borrowed into Chamic from Bahnar sources, although usually one or more inherited forms appear to have first developed the vowel phonetically, with this outlier providing a model for the borrowing. Finally, a length distinction developed, apparently triggered initially by the lowering of the inherited shwa to a short /a/, thus providing a length contrast with the inherited /a/. Continued contact with Bahnar reduced the four-way contrast in the first syllable in many dialects of Chamic, making the structure more sesquisyllabic. (for more on vowels, see Thurgood 1998, 1999)

Another minor change was the proliferation of glottalized stops, again apparently first through changes in inherited forms with native material (see Greenberg 1970; Thurgood 1999:87-94). This opened the way for borrowings, although given the intensity of the contact, the glottalized stops most likely would have been borrowed in any case.

Last, but not least, was the development of a register system, that is, contrasting voice qualities, typically a two-way between breathy voiced vowels (< earlier voiced obstruents) in contrast with modal or clear register. This simple register distinction may be reconstructable to PC. There is some question whether this register distinction developed under Mon-Khmer/Bahnaric influence or not; our initial assessment was that it did, but others such as Sidwell (p.c.) argue that Bahnar was not registral at that point. Be that as it may, register was widespread and certainly existed in the history of Hainan Cham. Other Chamic languages have gone on to elaborate their own systems in various ways (Thurgood 1996).

The final restructuring of Hainan Cham phonology takes place after their arrival on Hainan and had come into more intense contact both with Hlai speakers and with Min speakers, that is, sometime after 982. The two salient features are the much accelerated movement from sesquisyllabic to monosyllabic and the development of tones, but there was also simplification in the vowel system.

The increased monosyllabification sometimes came about through collapsing the two syllables into one. If the medial syllable began with *-h-, the monosyllabification was completed before the arrival on Hainan, perhaps as far back as the PC stage: as nothing has been distorted by doing so, we have used Malay to represent the Malayo-Chamic stage: Malay *tahun* ‘year’ cf. PC *thũn > Hainan Cham *t^hun*³³ (note that in some dialects of Chamic the reflexes of *th- actually are pronounced as clusters, not aspirated stops). Collapse of forms with medial *-l- or *-r- into monosyllables postdates the Hainan Cham arrival on Hainan, as several Northern Rglai sources still have two syllable forms, but otherwise the developmental pattern parallels that for the medial *-h-: Malay *bulan* ‘moon’ cf. PC *bila:n > Hainan Cham *p^hian*¹¹. Finally, where it was not possible to collapse the two syllables into one, a process with an intermediate stage still seen in Rade but completed in Hainan Cham: Malay *basah* ‘wet’ cf. PC *basah > Rade *msah*, Hainan Cham *sa*⁵⁵.

The most discussed development, however, is the development of a tone system, a tone system that parallels the tone systems of the other languages of Hainan, including the

Hlai dialects and Min dialects that Hainan Cham speakers had contact with. The external motivation for the tonogenesis was contact; the internal paths are transparent: forms ending in *-h have a 55 tone (high level); forms ending in a glottal stop have a 42 falling tone if the form began with a PC voiced obstruent (which led to breathiness, which determined the tone class), but with a 24 rising tone, if it did not; and, the remaining forms have a 11 tone (low level tone), if the form began with a PC voiced obstruent (which led to breathiness, which determined the low tone class), and a 33 (mid level) tone, if it did not. The actual picture is slightly more complicated; in PC disyllabic forms with two syllables, if the initial of the first syllable was a PC voiced obstruent, the breathiness from that obstruent spread through the medial and this breathiness determined the tone class of the second syllable. (Haudricourt 1984, Maddieson and Pang 1993, Thurgood 1996)

3 Lexicon

The vocabulary was restructured first in Champa and then again on Hainan. Half of the vocabulary, including much of the core vocabulary, is Bahnaric. In fact, there are often doublets, with one form inherited, the other borrowed. Also a significant number of body part words are borrowed e.g. *cadiŋ ‘finger’, *sua ‘dead skin’ (*kulit ‘skin’ is inherited). Other borrowings included kinship terms, adjectives, nouns, verbs, and so on. In fact, the borrowing is so massive that the language was sometimes thought to be Mon-Khmer, rather than Austronesian.

There is also the occasional borrowing of a grammatical morpheme, such as the negative imperative marker. More significant for the morphology is the iambic stress pattern; prefixes were reduced and then lost.

The arrival on Hainan led first to the borrowing of a small amount of Hlai vocabulary (and Hlai borrowed some from Chamic, most notably the word *nam ‘six’). Contact with Chinese, in contrast, led to massive lexical borrowing, including and grammatical morphemes (discussed in the next section). Zheng (1997:54) writes that, of some 2428 lexical items, roughly 20% are of Chinese origin: 21% of the nouns, 14% of the verbs, 31% of the adjectives, 31% of the classifiers, and several pronouns.

Much of the original Austronesian vocabulary is now gone, some lost to Bahnaric (Mon-Khmer) borrowings and now some to Chinese.

4 Constructions

The Hainan Cham came to Hainan speaking a language with limited grammatical morphology and with constructions marked with a grammatical morpheme plus word order. All of this is being rapidly restructured under contact with Chinese; we suspect that two aspects of this restructuring are the effects of the mass media and the results of near-universal schooling. Much of this is discussed in Thurgood and Li (2003), but it can be illustrated briefly here.

4.1 Genitives with Full Noun Phrases

In Northern Roglai all full NPs are postposed. In Hainan Cham, all genitive full NPs are preposed, as in Chinese, with the construction marked by *sa*³³ even in the most colloquial, least-Sinicized texts.

- (1) Northern Roglai: Nh GENNP
 gaʔ sa:k (Lee 1966:65)
 roof house
 ‘the roof of the house’
 Hainan Cham (all) GENNP sa³³ Nh
- (2) ...piaɪ³³ sa³³ za:ŋ³². (Zheng 1997:95)
 ...village GEN person
 ...cūn de rén
 ...cūn de rén
 ‘people of the village...’

In the most Sinicized texts, the native genitive marker *sa*³³ has been replaced by *tɪ*³, borrowed from Chinese.

- (3) Hainan Cham (Mandarinized) GENNP *tɪ*³³ HeadNP
 tan³³ k^hua⁵⁵ tɪ³³ si¹¹ha:u²¹, (Zheng 1997:4.1.3)
 arrive daybreak GEN after
 dào tiānliàng de shíhòu ...
 dào tiānliàng shíhòu, ...
 ‘At daybreak, ...’

4.2 Genitives with Pronouns

Genitives with pronouns show a mixed pattern in Hainan Cham. In more colloquial texts, the pronominal genitive is simply preposed without a genitive marker; in the more Sinicized texts it tends to be preposed with a genitive marker, as in Chinese. In Chinese, some variation in the use of a genitive marker occurs, apparently correlating with the transparency of situation being coded.

- Northern Roglai: Nh GENPr
 (4) sa:k hã (Lee 1966:65)
 house you
 ‘your house’
 Hainan Cham (colloquial): Nh GENPr
- (5) ko²⁴ ʔbu²⁴ nau³³ sa²⁴. (Zheng 1997:92)
 head.hair she messy
 tóufa tā luàn
 tā de tóufa luàn.
 ‘Her hair is messy.’
 Hainan Cham (Chinese influenced, with *sa*³³)
- (6) nau³³ sa³³ ko²⁴ ʔbu²⁴ sa²⁴. (Zheng 1997:97)
 she GEN head.hair messy
 tā de tóufa luàn
 tā de tóufa luàn.

‘Her hair is messy.’

4.3 Demonstratives and Head Nouns

As with genitive pronouns, both demonstratives and adjectives are postposed (without a genitive marker) in Northern Roglai and colloquial Hainan Cham, but preposed (with and without a genitive marker) in Chinese-influenced Hainan Cham and Mandarin. The tendency is for the genitive marker to show up in the Chinese-influenced Hainan Cham, a construction that reflects Mandarin influence.

- Northern Roglai:
 (7) sa:k gheŋ ʔunĩ (Lee 1966:65)
 house big this
 ‘this big house’
- Hainan Cham:
 (8) ʔai³³ ni³³ sat²⁴ ʔan³³ (Zheng 1997:84)
 water this truly cold
 shuĩ zhè zhēn lěng
 zhè shuĩ zhēn lěng.
 ‘This water is very cold.’
 Hainan Cham (Chinese influenced) (this + GEN) + clf
- (9) ni³³ sa³³ ta¹¹ p^han³² pi¹¹ kiau³³ lu³³. (Zheng 1997:75)
 this GEN one clf CM much
 zhè de yī fèn bǐjiào duō
 zhè yī fèn bǐjiào duō.
 ‘This portion is bigger.’

The alternation is found in the speech of the same speaker and correlates with the text type: the borrowed patterns are found in texts that describe more recent phenomena, whereas the native patterns are used in texts of traditional stories.

4.4 Adjectives and Head Nouns

The pattern for adjectives is the same. The N. Roglai has postposed adjectives as does the colloquial Hainan Cham, while the Mandarinized Hainan Cham has preposed adjectives, both with and without a genitive marker, calquing the Chinese.

- (10) N. Roglai: postposed adjectives
 sa:k gheŋ ʔunĩ (Lee 1966:65)
 house big this
 ‘this big house’
- (11) Hainan Cham: postposed adjectives
 na¹¹tsun³³ pion^ʔ24 poi^ʔ24: (Zheng 1997:1.1.9)
 bird big say
 niǎo dà shuō

dà niǎo shuō:
‘The big bird said:...’

Hainan Cham (Mandarin-influenced) preposed adjectives

- (12) ... kiʊ³³ san³³ (Zheng 1997:2.1.1)
... old village
... jiù cūn
‘... the old village’
preposed with sa³³, a calque on Mandarin *de*
- (13) na:iʔ³² sa³³ saŋ³³ huat²⁴, (Zheng 1997:2.1.10)
good GEN life
hǎo de shēnghuó
‘(the) good life’

4.5 Comparative Constructions

There are two distinct Hainan Cham comparative patterns: the native pattern is inherited from Chamic; the other is borrowed from Chinese. As Zheng writes (1997:75), the native construction is: X - Adj - CM/ST, with the adjective preceding the comparative marker (*la:uʔ*³² ‘CM; pass’) and the standard, while the Chinese pattern is: X - CM/ST - Adj, with the adjectives following the comparative marker (*pi*¹¹ ‘CM’ < Chinese) and the standard.

Hainan Cham (colloquial):

- (14) nau³³ ma⁴³ la:uʔ³² ha³³. (Zheng 1997:75)
he fat CM you
tā pàng bǐ nǐ
tā bǐ nǐ pàng.
‘He is fatter than you.’
Hainan Cham influenced by Chinese
- (15) kau³³ pi¹¹ ha³³ tsat²⁴ tso³³ kiə³³ sun³³. (Zheng 1997:75)
I CM you short three inch
wǒ bǐ nǐ ǎi sān cùn
wǒ bǐ nǐ ǎi sān cùn.
‘I am three inches shorter than you.’

4.6 Adverbs and Conjunctions from Chinese

As you might recall, some 45% of adverbs, prepositions (like *cong* ‘from’), and conjunctions are borrowed from Mandarin (Zheng 1997:54). In the case of these, what has been borrowed is not a grammatical word, but instead a construction along with the grammatical word that marks it. The rules of usage seem strikingly like those in Chinese.

- Adverbs: (the examples given here are intensifiers)
- (16) t̪^ha:i³³ ʔdiaʔ²⁴. p̪^hai³³ sianʔ²¹ na:iʔ³². (Zheng 1997:76)
very hot extremely good
tài rè fēicháng hǎo
tài rè fēicháng hǎo

‘very hot’ ‘extremely good’

- Conjunctions:
 (17) zi³³ pa³³ ziu³³ ha:i³³. (Zheng 1997:84)
 both hungry and tired
 yòu è yòu lèi
 yòu è yòu lèi.
 ‘Both hungry and tired.’
- (18) zi¹¹ ko¹¹ khiə⁴³ t^ha:i³³ ʔdai²⁴, kau³³ sau⁴³ pu³³ na:u³² lə³³.
 if tomorrow very hot, I then NEG go PERF
 rúguǒ míngtiān tài rè, wǒ jiù bú qù le (Zheng 1997:85)
 rúguǒ míngtiān tài rè, wǒ jiù bú qù le
 ‘If tomorrow is very hot, I won’t go.’

In discussing, not Hainan Cham, but the geographically distant and genetically distinct Tai-Kadai language Mulam, Zheng Guoqiao (1988:173) wrote that in Mulam the degree and quantity adverbs are all borrowed from Han Chinese and that these were subject to the same syntactic rules as in Han Chinese. Whether or not, it is literally true, it is instructive that a good scholar would make such a statement. The Hainan Cham parallels seem striking.

4.7 Other Chinese-influenced Constructions

All sorts of other constructions have been borrowed. In (19) below are three separate constructions showing Chinese syntactic influence. The first, indicated by the initial italicize characters with underline, is the extension of the prehead modification of the *sa*³³ construction to produce a prehead relative clause.

- (19) ʔdi⁵⁵ nan³³ sa³³ mo³³ si¹¹ mai³³ sa³³. (Zheng 1997:73)
 lie.down that GEN cow be female GEN
 tǎng nà de huángniú shì mǔ de
 tǎngzhe de nà tóu huángniú shì mǔ de.
 ‘The yellow cow lying down is female.’

The second is the use of the Chinese borrowing *si*¹¹ to mark the equative construction; the Chamic languages seem to use simple juxtaposition. And the third is the use of a postposed *sa*³³ as a nominalization in the *mai*³³ *sa*³³ ‘female’. All three reflect Chinese influence.

5 Observations

The successive stages in the restructuring of Hainan Cham, first in Champa and now on Hainan, reflect instances of intense contact. One suspects, in fact, that were it not for their identity as Hui, Muslims, Hainan Cham might very well have disappeared by now. As it is, the language has retained little more than words from its origins—the phonology is gone, the constructions are gone, and, although we did not discuss it here, even the rhetorical structures are becoming Chinese.

While the internal paths that brought about these changes are relatively transparent, the directionality and the impetus are provided by social, not linguistic pressures. Encroaching bilingualism with a powerful dominant language (along with schooling and social mobility), are among the most prominent factors that lead to the massive borrowing and drastic structural shift. Much of the language is gone; the rest will follow in another generation. What we fail to record now will be irretrievably lost.

References

- Greenberg, Joseph H. 1970. Some generalizations concerning glottalic consonants, especially implosives. *IJAL* 36:25-45.
- Haudricourt, André-G. 1984. Tones of some languages in Hainan. *Minzu Yuwen* 4:17-25.
- Lee, Ernest Wilson. 1966. Proto-Chamic phonologic word and vocabulary. Ph.D. dissertation. Indiana University. 67-3690.
- Maddieson, Ian and Keng-Fong Pang. 1993. Tone in Utsat [Hainan Cham]. *Tonality in Austronesian Languages*. Edited by Jerry Edmondson and Ken Gregerson. Oceanic Linguistics Special Publication No. 24. Honolulu, Hawaii: University of Hawaii Press. Pp. 75-89.
- Thurgood, Graham. 1999. *From Ancient Cham to modern dialects: Two thousand years of language contact and change*. Oceanic Special Publication No. 28. Honolulu: University of Hawai'i Press.
- _____. 1996. Language contact and the directionality of internal 'drift': the development of tones and registers in Chamic. *Language* 71.1:1-31.
- _____. 1998. The Austronesian and Mon-Khmer components in the proto-Chamic vowel system. *Papers in Southeast Asian Linguistics No. 15: Further Chamic Studies*. Edited by David Thomas. Pacific Linguistics. Series A-89. Canberra: The Australian National University. Pp. 61-90.
- Thurgood, Graham and Fengxiang Li. 2003. Contact induced variation and syntactic change in the Tsat of Hainan. *Language variation: Papers on variation in the Sinosphere and the Indosphere in honour of James A. Matisoff*. Pacific Linguistics. Pp. 185-200.
- Zheng, Guoqiao. 1988. The influences of Han on the Mulam language. In Jerold A. Edmondson and David B. Solnit, eds., *Comparative Kadai: Linguistic Studies Beyond Tai*. Summer Institute of Linguistics. Pp. 167-177.
- Zheng, Yiqing. 1986. A further discussion of the position of Huihui speech and its genetic relationship. *Minzu Yuwen* 6:37-43.
- _____. 1997. *Huihui Yu Yanjiu [A Study of Tsat]*. Shanghai Yuandong Chuban She [Shanghai Far East Publishers].