7 *The tones from Proto-Chamic to Tsat [Hainan Cham]: insights from Zheng 1997 and summer 2004 fieldwork*[1]

Graham Thurgood and Ela Thurgood

The Language, the People, The History
The Cham of Hainan, termed Huihui (that is, Muslim) by the Chinese, mainly live just within the Yanglan township within the Sanya municipality on Hainan, in the villages of Huihui and Huixi. These people call themselves u^{33}tsa:nʔ^{32}, an autonym composed of U 'people' + Tsat 'Cham'(< Proto-Chamic [PC] *cam 'Cham'), and their language tsa:nʔ^{32} 'Tsat' (< PC *cam 'Cham').[2]

Historical reconstruction of Chamic (Thurgood 1999:224-227) makes it clear that the Tsat represent an offshoot of the Northern Cham of the Champa Kingdom. The first mention of the Kingdom of Champa, according to Coedès (1968:42), around 190 to 193 AD, but the first reference to what was the Tsat was undoubtedly in the Chinese dynastic records, which seem to have referred to the Tsat twice, once in 986 and once in 988. These dates follow not too long after the fall of the Vietnamese, in their 'Push to the South' sacked the capital in Indrapura in 982. This fall of the capital in 982 accounts for the refugees mentioned in the Chinese dynastic records of 986 (*History of the Song Dynasty* (960-1279), which records in 986 the arrival of some Cham in Hainan from Zhancheng [=Cham City] (Zheng 1986:37)). Another group is recorded in 988 in Guangzhou (Canton). The capital then moved to the south, but in 1491, the southern capital at Vijaya also fell resulting in another influx of refugees in Hainan.

The first modern account and the first account containing linguistic data, however, is H. Stübel's short note on the language found in his 1937 work entitled *Die Li-Stämme der Insel Hainan: ein Beitrag zur Volkskunde Süd-Chinas, unter Mitwirkung von P. Meriggi,* which, despite showing no indication of the tones and being limited to a relatively small number of forms provided the basis for Benedict's 1941 identification of the language as Chamic. Then, beginning with work by Zheng and Ouyang in Ya county in 1956, who ran across the language in the course of their monumental work on the Li languages (1983b), there has been sporadic modern work on the language. Initially busy

---

[1] This paper was originally given at the SEALS XV (Southeast Asian Linguistics Society) conference, in Canberra 20-22, April 2005. It has been significantly improved by the comments of participants, in particular, Phil Rose, Justin Watkins, Harold Koch, Paul Sidwell, Koichi Honda, Pittayawat Pittayapom, and Charatdao Intratat.

[2] As Goschnick (1977:106) notes, other Chamic subgroups have also used 'Cham' in their name: the Cham Raglai (the Roglai; from ra 'people' + glai 'forest'), the Cham Jarai (the Jarai), the Cham Kur (Cham + kŭr 'Khmer', the Western Cham of Cambodia and Southern Vietnam), and the Cham Ro > Chru (from Cham + rɔ 'remnant').

Thurgood, G. and Thurgood, E. "The Tones from Proto-Chamic to Tsat [Hainan Cham]: Insights from Zheng 1997 and from Summer 2004 fieldwork". In Grant, A. and Sidwell, P. editors, *Chamic and Beyond: Studies in mainland Austronesian languages.*
PL-569:247-271. Pacific Linguistics, The Australian National University, 2005. DOI:*10.15144/PL-569.247*

with other work, they collected a small amount of data and left its analysis and further investigation for a later date. The couple returned in the spring of 1981 to the Yanglan Township of Ya County to carry out a detailed investigation of the language, collecting some 3000 words and some 300 sentences. The work by Zheng and Ouyang attracted much attention from scholars both inside and outside China, particularly because, despite being a Chamic language (that is, in the Malayo-Chamic subgroup of Austronesian, see Figure 1) it was fully tonal.

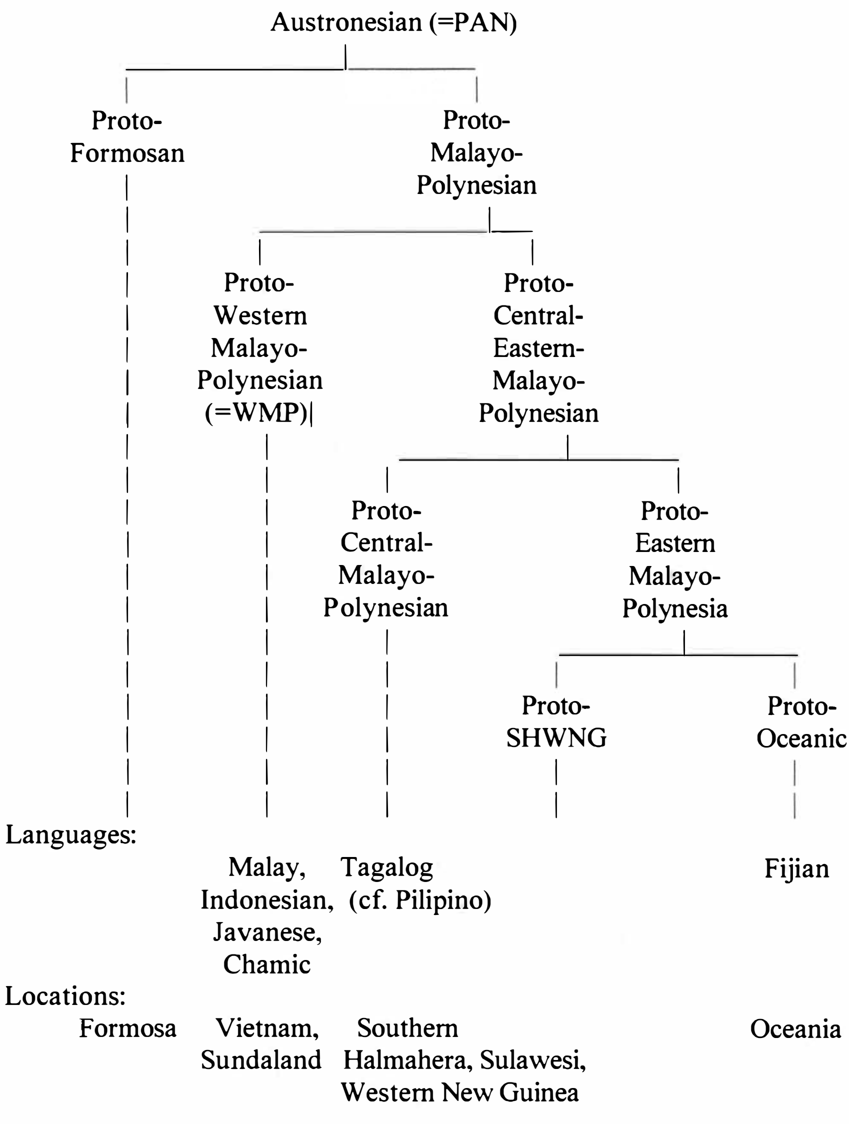


Figure 1: *Proto-Austronesian family tree (Blust c. 1980)*

Although there was initially some debate over the status of Tsat within China, foreign scholars recognized it immediately as Chamic. What made it of particular interest in general was its complete typological restructuring under the influence of the languages of Hainan. Several scholars wrote about the syntactic restructuring (Ni Dabai 1988ab, 1990ab; Thurgood and Li, forthcoming, 2003), but the major focus was on the development of a full tonal system from a completely atonal source. This development caught the interest of a number of scholars (Benedict 1984, Haudricourt 1984, Maddieson and Pang 1993, Thurgood 1992, 1993, 1996, 1999). The work on Tsat tonogenesis was done exclusively on the basis of several early articles by Ouyang and Zheng (1983a), Zheng (1986, 1997), and Ni (1988ab, 1990ab). With the publication of Zheng's 1997 grammar, however, the data base has been expanded significantly; that expanded database coupled with the results of our fieldwork in the summer of 2004 makes it possible to paint a more detailed, richer picture of Tsat tonogenesis.

However, first, a few comments will be made on one of the important precursors to tonogenesis, the development of Tsat monosyllables from Malayo-Chamic disyllables. Contact with the iambic Austroasiatic patterned languages along the coast of Vietnam caused a switch in stress from penultimate to iambic, ultimately resulting in a reduction first to iambic and then to monosyllabic Chamic forms, an explanation that looks likely even for the pattern of Table 1, for which the disyllabic form can no longer be recovered from the data of the modern languages.

Table 1 shows the reduction of pre-Proto-Chamic forms with a medial *-h- to monosyllabic after the loss of the first syllable vowel. PAn is Proto-Austronesian, while Written Cham is the oldest written records of Cham; Malay is included simply for comparison.

PAn	Malay	PChamic	Wr. Cham	Tsat	gloss
*taqun	tahun	*thŭn	thun	t^hun^{33}	'year'
*puqun	pohon	*phŭn	phun	p^hun^{33}	'plant'
*paqit	pahit	*phiəʔ	---	$p^hiʔ^{24}$	'bitter'
*paqat	pahat	*pha:ʔ	phaʔ	$p^haʔ^{24}$	'chisel'
*paqa	paha	*pha	phā	p^ha^{33}	'leg, thigh'
*daqiS	dahi	*dhəi	dhei	t^hay^{33}	'forehead'

Table 1. *Monosyllables from disyllables with medial *-h-.*

In a parallel way, Table 2 shows the reduction of pre-Proto-Chamic forms with a medial *-l- or *-r- to monosyllabic after the loss of the first syllable vowel. Note that the medial *-l- or *-r- is retained in Tsat as an –i- glide.

PAN	Malay	PChamic	Wr. Cham	Tsat	gloss
*baqeRu	baharu	*bahrəu	barăhau	p^hi^{11}	'new'
*qabaRa	---	*bara	bara	p^hia^{11}	'shoulder'
*bulan	bulan	*bila:n	bulan	p^hian^{11}	'moon'
*bulu	bulu	*biləu	bulău	$p^hiɤ^{11}$	'body hair'

Table 2. *Monosyllables from disyllables with medial liquids.*

In the remaining cases, as Table 3 shows, the initial syllable is simply lost without any trace in Tsat. A glance at the Chru and Rade columns reveals the path of development: first the vowel of the initial syllable was reduced and then the whole syllable was lost.

PAN	PChamic	Wr. Cham	Chru	Rade	Tsat	gloss
*baseq	*basah	basah	pəsah	msah	saʔ^{43}	'wet; damp'
*qubi	*hubəy	hubei	həbəi	hbei	p^{h}ay^{11}	'taro; yam'
*quzan	*huja:n	hujan	həjăn	hjan	sa:n^{11}	'rain'
*qumah	*huma	humā	ləma	hma	ma^{33}	'dry field'
*lapaR	*lapa	lapa	ləpa	epa	pa^{33}	'hungry'
*lima	*lima	limə	ləma	ema	ma^{33}	'five'
*m-uda	*muda	medā	məda	mda	t^{h}a^{11}	'young; unripe'
*mamaq	*mumăh	memeh	bəmah	mmah	maʔ^{43}	'chew'
*pajay	*paday	padai	pədai	mdie	t^{h}a:yʔ^{24}	'rice (paddy)'
*panaq	*panah	paneh	pənah	mnah	naʔ^{43}	'(shoot) bow'
*taliS	*taləy	talei	tələi	klei	lay^{33}	'rope; string'
*tangan	*taŋa:n	taŋin	təngăn	kngan	ŋa:n^{33}	'hand'

Table 3. *Monosyllables with no Tsat evidence of an original initial syllable.*

This reduction of disyllables to monosyllables did not complete the restructuring of the Tsat word along the lines of the languages of Hainan—most of the finals were to be lost, but the transition to monosyllabic was complete.

The Modern Tones

The notational system used for tones in this work represents an adaptation of the tonal system found in Zheng (1997:24-25), which is itself an adaptation of Ouyang and Zheng (1983a);[3] however, it differs only in minor details, the most obvious one being that all glottal stops are overtly marked. Nonetheless, the Zheng 1997 analysis is consistent with the instrumental data obtained during our fieldwork on Hainan in the summer of 2004.[4]

Procedure

The acoustic analysis of Tsat is based on recordings during fieldwork in the summer of 2004 in Hainan consisting of words produced in citation form by six speakers (three

[3] Tones are indicated using Chao tone numbers (1930), a five point scale with 1 the lowest and 5 the highest; the first number indicates the starting point, the second indicates the ending point. Thus, a high level tone would be 55, starting high and remaining high, while a 24 tone would indicate starting slightly below the mid point and rising slightly above the mid point.

[4] Our fieldwork in the summer of 2004 was supported in part by National Science Foundation Grant NSF #606232 Endangered Languages in China. During this time Professor Sun Hongkai arranged for us to be accompanied by two distinguished Chinese scholars, Professor Ouyang Jueya and Professor Jiang Di.

female subjects (F1, F2, F3) and three male subjects (M1, M2, M3)), with ages form the early 20s to the mid 60s. All were fluent in Hainan Cham [Tsat]; all also knew Hainanese, the Min dialect of Hainan, and all knew Mandarin. Each speaker repeated each word three times. The data was recorded in a quiet room on a laptop computer using the SoundEdit software and a head mounted Telex H-831 mic. The analyses were performed using Macquirer software. The recordings were digitized at a sampling rate of 11, 025 Hz.

Prior analyses

All analyses of Tsat treat the language as essentially having five phonemic tones. Zheng (1997), for instance, posits five basic phonemic tones, plus some allophonic variation conditioned by the existence of a final stop—usually a glottal stop, see Table 4.[5] Likewise, historical analyses of the origins of Tsat tones based on Ouyang and Zheng (1983a)—Haudricourt (1984), Benedict (1984), Maddieson and Pang (1993), Thurgood (1992, 1993, 1996, 1999)—treat also Tsat as having five etymological tones: three level tones and two contour tones in checked syllables, but none of these studies contains a fully adequate treatment of the allophonic variants. However, none of the works had access to Zheng (1997); Zheng supplemented with instrumental analyses of our own fieldwork allows a much more complete picture of the Tsat tones and allotones, both synchronically and diachronically.

	/11/	/24/	/33/	/43/	/55/
'live' finals	[11~21] sa^{11} 'tea'		[33] sa^{33} 'one'		[55~45] sa^{55} 'wet'
glottal finals	[21] ta:nʔ^{21} 'intestine'	[24] saʔ^{24} 'cook'	[32] sa:yʔ^{32} 'lay bricks'	[43] saʔ^{43} 'ladder'	
stop finals (borrowings)		[24] tsat^{24} 'narrow'		[43] tsat^{43} 'photo'	

Table 4: *The five phonemic tones (< Zheng (1997), modified)*

A handful of additional forms exists with patterns that fall outside of Table 4. Without exception, these forms represent loanwords not fully assimilated into the older Tsat phonological patterns; in fact, in many such cases it is hard to distinguish code-switching from borrowing. In some instances, these historically-interesting aberrancies give clues about historical contact patterns and when this is so, it is usually commented on. However, before historical inferences are drawn about inheritance or contact on the bases of forms that pattern irregularly, a fuller description of the regularly-patterning forms is in

[5] Organizationally, this table differs from the corresponding table in Zheng in the fact that the second row here consists of syllables with final glottal stops and, notationally, in that the final glottal stops are written as such.

order.

The tone system of Tsat is one in which the diachronic origins are still reflected in the modern distribution of the phonetic, and, thus, phonemic tones: to use Thai terms, the 'live' syllables, that is, open syllables and syllables ending in nasals have one set of tones, syllables, and dead syllables, that is, checked syllables, have another set of tones. The live syllables co-occur with the so-called level tones, tones 11, 33, and 55; the contour tones co-occur with stopped syllables.

The 'level' tones. Tsat has three level tones: 11, 33, and 55. Figures 1 and 2, show the level tones for the two youngest speakers.

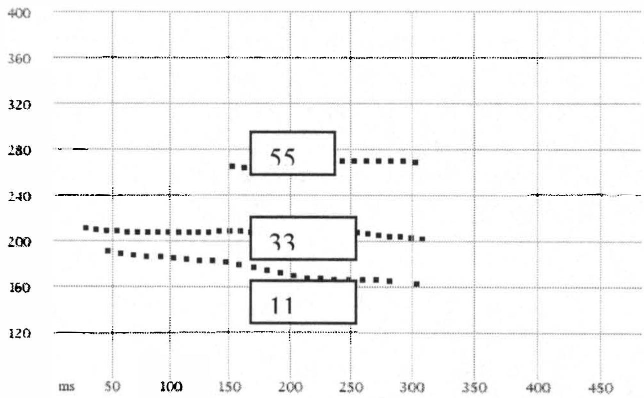


Figure 1: *The three level tones (speaker F1)*

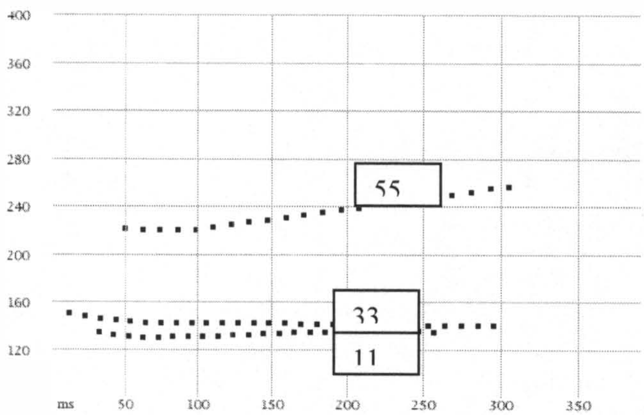


Figure 2: *The three level tones (speaker M1)*

These two speakers have different fundamental frequency (F0) ranges, as measured from the highest point of the 55 tone to the lowest point of the 11 tone. For the female speaker (F1), the highest point is ca. 320 Hz and the lowest point is ca. 168; for the male speaker

(M1), the highest point is ca. 255 Hz and the lowest point is ca. 137 Hz. Thus, the pitch range for the female speaker is 152 Hz, while for the male speaker it is 118 Hz. It is within these ranges that both the level tones and the contour tones are found.

Tone 55.
The pitch value of the 55 tone is strikingly high, separating it clearly from all the other tones (see Figures 1 and 2). In Figure 1, the 55 tone is essentially level, while in Figure 2 the 55 tone is rising, in this instance by 40 Hz. The distinctiveness of this extra-high tone, which at times has a falsetto quality to it, is immediately obvious, and has been commented on by all observers.[6] Ni Dabai (1988a) labeled it 55, but explicitly notes that it can be either 55 or 45, a characterization that matches Maddieson's instrumental work in Maddieson and Pang (1993:80), in which their figure shows it as initially rising.[7]

A historical note is in order: With the exception of one word in the data, the word zoʔ^{55} 'powder' listed in Ni (1988a:19), the 55 tone is restricted to non-stopped syllables.

Tone 33.
The pitch of the 33 tone is described as mid level although in relative terms it is sometimes a little lower in comparison with the extra-high pitch of the 55 tone. Ouyang and Zheng suggest that phonetically it might be described as 22, but leave it as 33 for reasons having to do as much with notation as phonetics; similarly, Maddieson and Pang (1993:79) note that the onset of the 33 tone is quite close to the onset of the 11 tone. As Figures 1 and 2 show the 33 is a level tone, but at times towards the end it drifts downward. When it drifts down, it would not be readily distinguishable from the checked 32 variant except by its final glottal stop.

The 32 pattern is the checked variant of the 33 tone. It is worth noting that all the syllable finals with a 32 (or 21) tone have a glottal catch when they end with a final *-n* or *-ŋ*. See the discussion of preploded final nasals below. The 32 tone variant will be discussed further with the contour tones.

Tone 11.
All authors note that the low level tone frequently is phonetically more a 21 than a 11 (see Figure 1). Maddieson's figure shows this tone as 21 and Ni's notes have it as 21.[8] Similarly, Ouyang and Zheng (1983a:32) note, even when this tone does not end in a glottal stop, the pitch is often 21. Our own recordings show low tone items with a final glottal stop as consistently falling, while there is variation between level and falling in those without a final glottal stop. When this non-checked low tone drifts down toward the end, it is largely distinguishable from the checked phonetic variant of the 11 tone, labeled

---

[6] Rose (1997:19) comments similarly on a super-high tone in Pakphanang Thai, noting that it is sometimes falsetto, convex in shape, and very salient. The Tsat 55 differs in that it is only sometimes convex.

[7] It is not clear to us whether the initially rising onset has any particular significance. Tones labeled 55 elsewhere also sometimes are actually more of a rising tone. For instance, in instrumental work we did on Jiamao, a Tai-Kadai (Kra-Dai) language of Hainan, the so-called 55 tone turned out upon instrumental analysis to be 45 or even 35. Whether this represents a change, notational convention, or phonetic variation is unclear.

[8] Although Ni notes the fall occurs with non-checked tone 11 forms, he observes that the fall is particularly noticeable when the form ends in a glottal stop.

21 by Zheng, only by the presence of a final glottal stop. The checked 21 variant will be further discussed with the contour tones.

The contour tones.

The remaining tones have contours: three falling: 43, 32 (the checked allophone of 33), and 21 (the checked allophone of 11); and, one rising, 24. Maddieson and Pang (1993) note quite correctly that the contour tones are always associated with checked syllables. By checked syllables, Maddieson and Pang mean Zheng's rising 24 and her falling 43, 32, and 21. Although it is not obvious from the transcriptional conventions used, Zheng and Ouyang were fully aware that these tones had a glottal final. The Zheng/Ouyang transcriptional system systematically distinguishes glottal final syllables—the labels 24, 43, 32, and 21 indicate the existence of both a contour and a final stop: if the segment does not already end in a *-p, -t,* or *-k,* the existence of a glottal stop is implicit. During our fieldwork together Ouyang mentioned more than once that it was unnecessary to mark final glottal stops, as they were implicit in the tone numbers.

It is important in terms of tonogenesis to note that the rising tone and the three falling tones differ in more than just pitch contour. In fact, other cues may be more salient for distinguishing the rising from the falling tones than simple pitch differences. The rising tone, which rises somewhat slowly, occurs with phonetically long vowels; the three falling contour tones, which fall somewhat abruptly, occur with phonetically shorter vowels. The patterns suggest that at the point where the contour pitch developed, the finals in question (*-ʔp, *-ʔt, *-ʔk; *-ʔn, *-ŋʔ; *-yʔ, *-wʔ) had co-articulated finals, perhaps accounting for why the effect of the glottal final of the rising tone differs from the effect of the glottal final of these falling tones. It has been noted in the tonogenesis literature that final glottal stops are sometimes related to rising contours and sometimes to falling (cf. Thurgood 2002).

Maddieson and Pang suggest on the basis of the limited data then available to them that there is a single rising tone and a single falling tone, both of which occur only in checked syllables. Essentially, this position is borne out by our study, with two qualifications: First, the claim that there is only one falling tone must be interpreted as phonemic rather than phonetic as there are three phonetically-distinct falling pitch patterns-—43, 32, and 21; these can, however, be easily phonemicized, leaving only one phonemically distinct falling tone. The six speakers we recorded all had three distinct phonetic patterns—although not for all words and not at all times, but no minimal pairs exist and the pitch patterns are only a part of a cluster of features that distinguish the various pitch patterns involved. Second, at times, all of the non-checked so-called level tones show some contour; more specifically, the so-called 55 tone is often phonetically a 45, something obvious in both Maddieson's figure and our own instrumental examination, the so-called 11 tone is often actually 21, something again in Maddieson's figure and our instrumental examination, and even the 33 tone at times noticeably drifts downward.

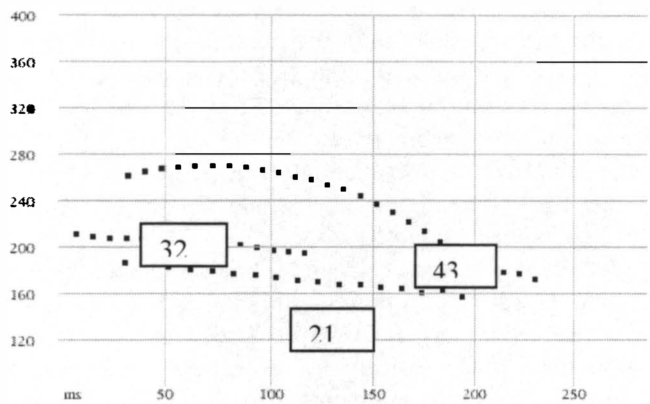


Figure 3: *Falling tones (speaker F1)*

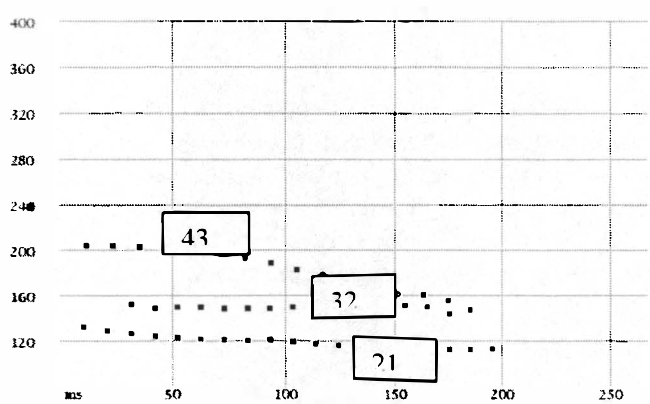


Figure 4: *Falling tones (speaker M1)*

Tone 43.
Tone 43 is high-falling, ending in a glottal stop. Maddieson and Pang (1993) described it as 42, Ni (1988ab, 1990ab) has it as 42, and Zheng (1997) has it as 43. In our data, it varies between 42 and 43 (see Figures 3 and 4). For example, in Figure 3 tone 43 looks like 42; this particular token has a rise in frequency at the beginning of the vowel followed first by a gradual fall, and then by a relatively steep fall of 60 Hz—an overall fall of roughly 100 Hz. In Figure 4, we have 43, with a level plateau through approximately the first half of the vowel followed by a gradual decline through the second half of the vowel—an overall fall of a little over 40 Hz.

Two historical notes are in order. Forms in 43 inherited from Proto-Chamic [PC] had a PC final stop, and all 43 forms ending in *-t* or *-k* are borrowings.

Tones 32 and 21

Tones 32 and 21 are squashed into the narrow range occupied by level tones 33 and 11. For the female speaker (F1), the two falling tones are between ca. 210 Hz and 165 Hz (Figure 3); and, for the male speaker (M1), the two falling tones are between ca 150 Hz and 130 Hz (Figure 4). As a result, tones 33 and 32 are often in the same pitch range, and tones 11 and 21 are often in the same pitch range. For M1, neither the pitch shape nor the fundamental frequency differentiates between tones 33 and 32; for F1 neither the pitch shape nor the fundamental frequency differentiates between tones 11 and 21. Clearly, there are other acoustic cues that differentiate between tones 33 and 32 and between 21 and 11. These cues are provided by the final glottal stop.

The final glottal stops in tones 32 and 21 are secondarily derived either from syllables with PC final nasals or from the PC diphthongs *-ay or *-aw. Although the 32 and the 21 tones consistently show a falling contour, this by itself would not always distinguish them from 33 and 11, respectively, as at times these also manifest a falling pitch pattern. Our recordings and measurements show the items without a final glottal stop varying between a level and a falling pitch pattern, while those with a final glottal stop are consistently falling and accompanied by creakiness (discussed below).

A historical note: While at times all six of our speakers keep these three falling pitch patterns distinct, on occasion some of them do not; further, the historical evidence indicates that at least for some speakers in some words the 32 and the 21 have begun to merge. This is an area for further investigation.

Tone 24.

Tone 24 is the mid-rising tone, ending in a glottal stop. As Maddieson notes in Maddieson and Pang (1993:80), the sources indicate that all rising tones occur in checked syllables, which Maddieson treats as a 24 tone. In the most recent work, Zheng (1997:24) also has a single 24 tone found in checked syllables. While it is not particularly clear from their notational system, it is clear from conversations with both Zheng and Ouyang, that they too regard the 24 as ending in a glottal stop. For examples, see Figures 5 and 6.

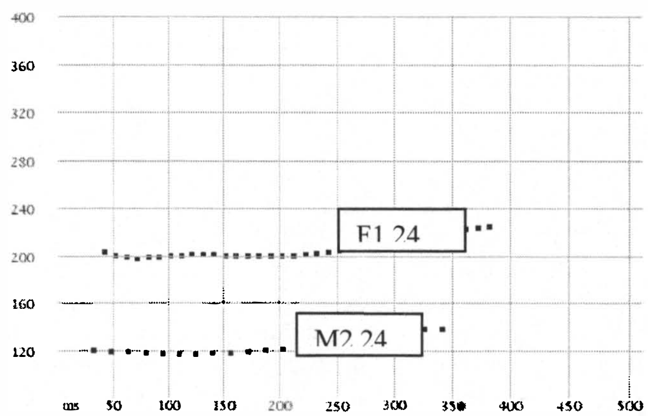


Figure 5: *The 24 rising tone (two tokens; speakers F1 and M2)*

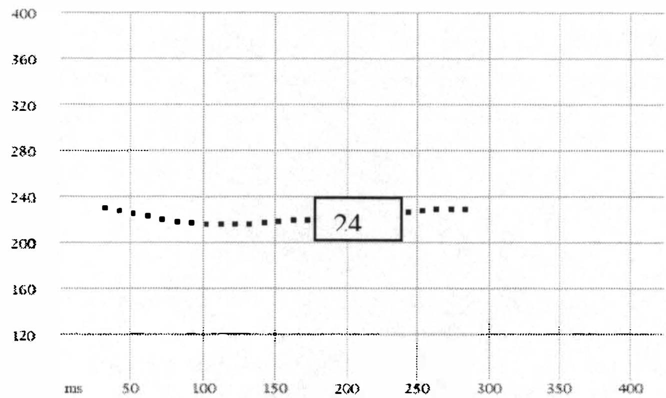


Figure 6: *The 24 rising tone (speaker F2)*

In our data, tone 24 either displays a level plateau through approximately the first half of the vowel followed by a rise in frequency (Figure 5) or it falls for the first third of its duration, flattens out for the second third of its duration, and goes back up for the last third of its duration (Figure 6). The tone itself is easily distinguished, although at times it patterns enough like the unchecked 33 tone that the glottal final is needed to distinguish it.

On inherited PC forms, the final glottal stop is the remnant of old final stops. All 24 forms ending in *-t* or *-k* are borrowings.

Creaky voice

All contour tones correlate with a creaky voice quality, which in Tsat is always associated with a final glottal stop. The non-modal phonation types, including creaky voice, are described by their unique spectral properties, including periodicity; overall acoustic intensity; level of intensity of the first harmonic (H1), level of intensity of the second harmonic (H2), and level of intensity of the highest harmonic both in the first formant (F1) and in the second formant (F2). In this study, the voice quality effects of a glottal stop on the preceding vowel were analyzed on the FFT power spectra with superimposed LPC spectra using a bandwidth of 43 Hz and a window with 256 points. We concentrated on the creaky phonation of the /a/ vowel, because intensity differences shown by FFT spectra are best shown for low vowels, for which the frequency location of F1 is far enough from H1 not to influence the amplitude of H1 (Jessen and Roux 2002; Ladefoged 2003). The voice quality effects induced by a glottal stop were analyzed on the second FFT spectrum computed in the middle of the vowel. For each vowel token, a number of amplitudes were measured from the spectra: the amplitude of the first harmonic (H1) of the fundamental frequency, the amplitude of the second harmonic (H2), and the amplitudes of the first formant (F1) and of the second formant (F2).

Table 5 gives the list of words in which the /a/ vowel was analyzed. The transcription is a modification of the notational system of Zheng (1997).

[a̰]			[a]		
pa:ʔ^{24}	(PC *pa:t)	'four'	pa^{33}	(PC *kapa:s)	'cotton'
p^{h}a:ʔ^{24}	(PC *p^{h}a:t)	'chisel'	p^{h}a^{11}		'bland'
ta:ʔ^{24}	(PC *rəta:k)	'beans, peas'	ta^{33}	(PC *ʔata:s)	'far'
taʔ^{43}		'pillow'	ta^{11}	(PC *buta)	'blind'
t^{h}aʔ^{43}		'white gourd'	t^{h}a^{33}		'speech'

Table 5: */a:ʔ/ and /a/*

In terms of spectral characteristics the biggest difference between creaky voice and modal voice lies in the comparison of the intensity values of higher frequencies to the intensity values of lower frequencies; this is the case in Tsat. For the female speakers (Figure 7), when we compare H1 with the highest harmonic in the first *formant*, for creaky voice, the highest harmonic in the first formant is 11.8 dB above H1 (H1-F1 = -11.8 dB), while for modal voice F1 is only 2.8 dB above H1 (H1-F1 = -2.8 dB). The difference between creaky and modal is even more marked when we compare H1 with the highest harmonic in the second formant. For creaky voice, F2 is 7.7 dB above H1 (H1-F2 = -7.7 dB). In contrast, for modal voice, F2 is 3.2 dB below H1 (H1-F2 = 3.2 dB). For the male speakers (Figure 8), for creaky voice, when we compare H1 with the highest harmonic in the first formant, the highest harmonic in the first formant is 11.6 dB above H1 (H1-F1 = -11.6 dB), while for modal voice F1 is 6.3 above H1 (H1-F2 = -6.3 dB). Similarly, for creaky voice, when we compare H1 with the highest harmonic in F2, F2 is 9.3 dB above H1 (H1-F2 = -9.3 dB), while for modal voice, F2 is 6.2 dB above H1 (H1-F2 = -6.2 dB).

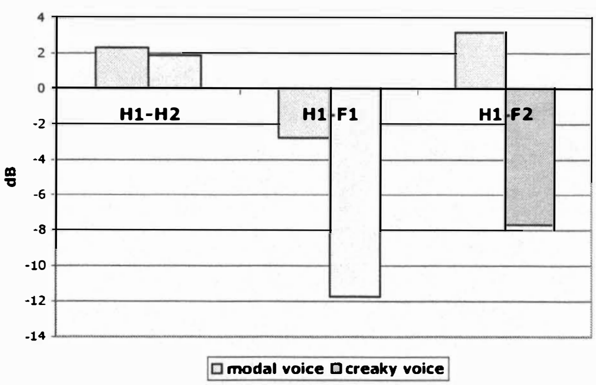


Figure 7: *Differences in H1-H2, H1-F1, and H1-F2 for female speakers*

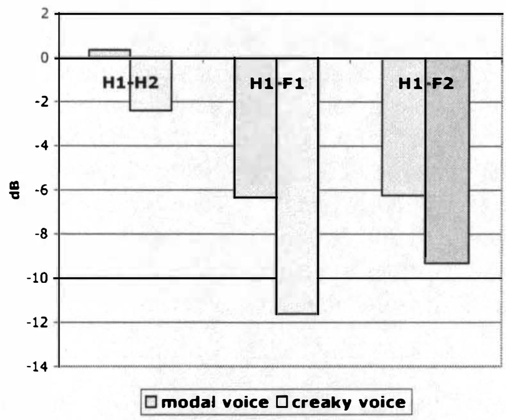


Figure 8: *Differences in H1-H2, H1-F1, and H1-F2 for male speakers*

During our recording sessions we often noticed, not only that the /a:yʔ/ was creaky, but also that /a:yʔ/ was sometimes pronounced as [a:ʔ]. To examine both features, we compared /a:y/ words with a final glottal stop to /ay/ words without a following glottal stop. For the words used, see Table 6.

pa:yʔ^{32} (PC *tapay) 'wine, liquor'	pay^{33} (PC *tampɛy) 'to winnow'
ta:yʔ^{32} (PC *hatay) 'liver'	pay^{33} (PC *lumpɛy) 'to dream'
ta:yʔ^{32} (PC *matay) 'to die'	t^{h}ay^{11} (PC *ʔadɛy) 'younger sibling'
t^{h}a:yʔ^{32} (PC *paday) 'paddy'	t^{h}ay^{33} (PC *ʔadhɛ̃y) 'forehead'

Table 6: */a:yʔ/ versus /ay/*

Figure 9 presents the distribution of [a:yʔ] and [a:ʔ] in 72 forms produced by the six subjects (4 tokens of /a:yʔ/ repeated 3 times by 6 speakers = 72 tokens of /a:yʔ/). The two older male speakers (M3 and M2) pronounce the /a:yʔ/ diphthong variably as either [a:yʔ] or [a:ʔ] (with slight preference for [a:yʔ] for M2). The remaining four speakers favor [a:ʔ] over [a:yʔ], with three of the four speakers (M1, F3, and F2) only producing [a:yʔ] once, interestingly all in the same word [t^{h}a:ʔ^{32}] 'paddy'. The youngest subject (F1) did not produce [a:yʔ] at all.

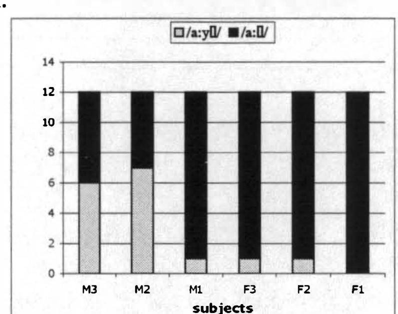


Figure 9: *Frequency of [a:yʔ] versus [a:ʔ] among the six Tsat speakers*

The characteristics of the Tsat creaky /a:yʔ/ can be seen in the spectrogram and waveform of /t^ha:yʔ^{32}/ 'paddy' in Figure 10. For comparison, the spectrogram of modal /ay/ in /t^hay^{33}/ 'forehead' is also given. The spectrogram of /t^ha:yʔ^{32}/ 'paddy' shows the sequence of diphthong + glottal stop realized as a modal vowel (between 150 and 209 milliseconds) which then becomes creaky (between 209 and 240 milliseconds) before turning into a glide; thus phonetically [t^haa̰yʔ]. The waveform corresponding to the portion of the spectrogram between 270 and 320 milliseconds encompasses both modal and creaky phases. The phonetic transcriptions carefully positioned above the spectrograms indicate the approximate location of different acoustic events.

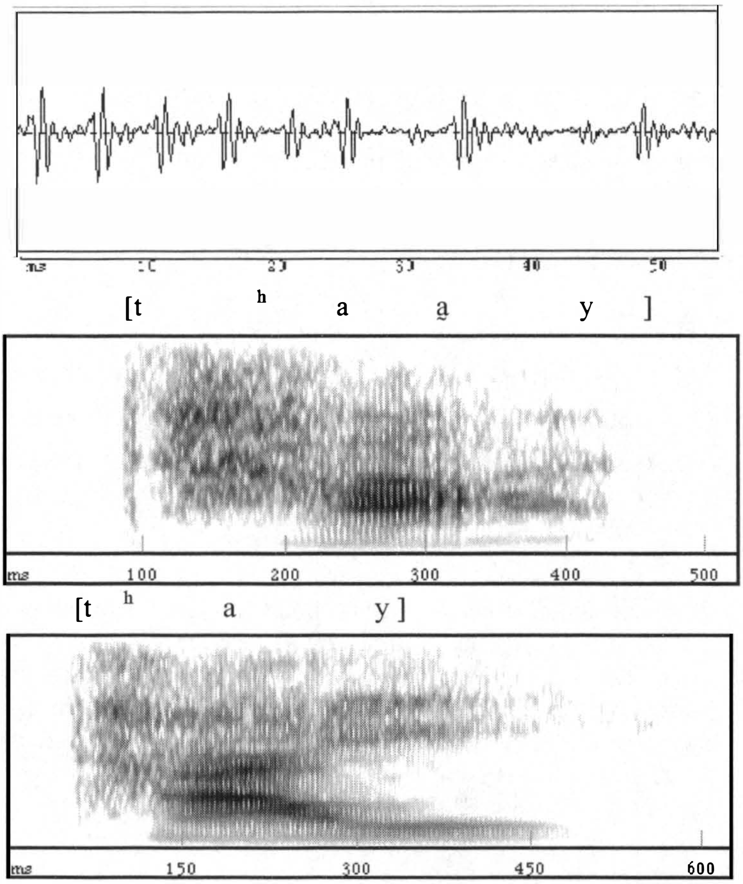


Figure 10: *Waveform and spectrogram of [t^haa̰y^{32}] 'paddy' and spectrogram of [t^hay^{33}] 'forehead' (speaker F2)*

The pitch periods of the creaky phase are irregular in terms of their duration and considerably longer than those of the modal phase. They are also relatively infrequent compared to the pitch periods of the modal phase. The increased length of the pitch periods indicates the lowered fundamental frequency values of the creaky [a̰] vowel.

A different phonetic realization of /a:yʔ/ is illustrated in Figure 11. In this case, /a:yʔ/ is pronounced as [aʔa̰]. The waveform and the corresponding spectrogram encompass first a modal phase (between 20 and approximately 125 ms), followed by a glottal stop, and then a creaky phase (between 150 and approximately 250 ms).

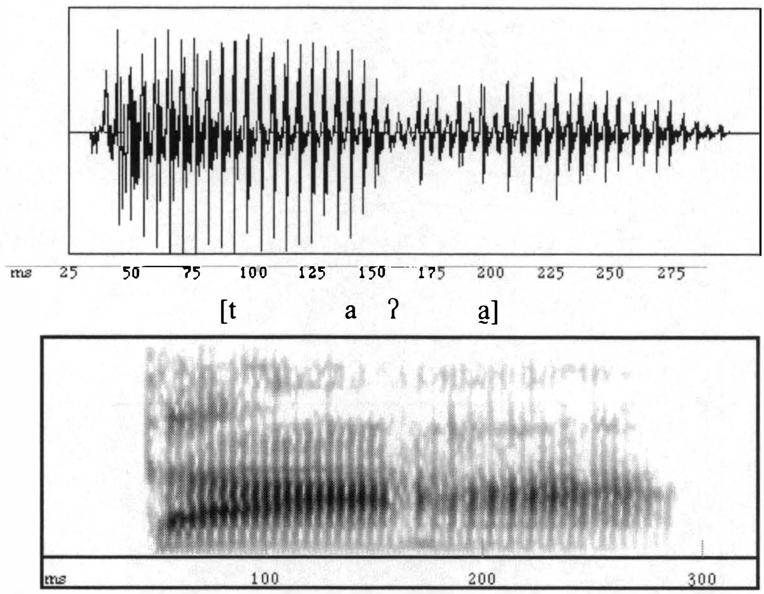
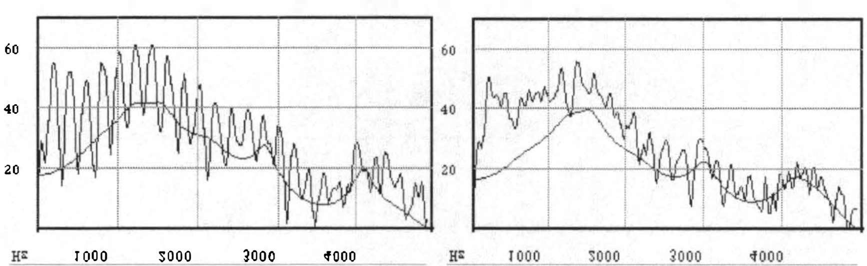


Figure 11: *Waveform and spectrogram of [taʔa̰ʔ^{32}] 'liver' (speaker F1)*

The FFT spectra with the superimposed LPC spectra of the modal phase and the creaky phase of /a/ are given in Figures 12-13. A shallower spectral tilt and a lower first harmonic (H1) clearly differentiate the creaky phase from the modal phase.



Figures 12-13: *FFT spectra with LPC spectra of modal and creaky /a/ in [taʔa̰] 'liver' (speaker F1)*

Creakiness in /a:wʔ/ was analyzed on the basis of the words given in Table 7. Table 7 also gives the words with /aw/ without a following glottal stop, the spectrograms of which were compared with the spectrograms of /a:wʔ/.

/a:wʔ/		/aw/	
ta:wʔ^{21} (PC *pataw)	'master, lord'	taw^{11} (PC *katɔw)	'louse'
ta:wʔ^{21} (borrowing)	'10 liters, clf.'	taw^{11} (PC *kukɔw)	'fingernail'
t^{h}a:wʔ^{21}	'hide something'		
t^{h}a:wʔ^{21}	'avoid (rain)'		

Table 7: */a:wʔ/ versus /aw/*

In Figure 14, /a:wʔ/ pronounced as [a:w̰ʔ] is illustrated and contrasted with modal /aw/ pronounced as [aw]. In the spectrogram of /t^{h}a:w̰ʔ^{21}/ 'hide something', the pitch periods of the glide have a greater distance between the vertical striations than the pitch periods of the glide in /taw^{11}/ 'louse'.

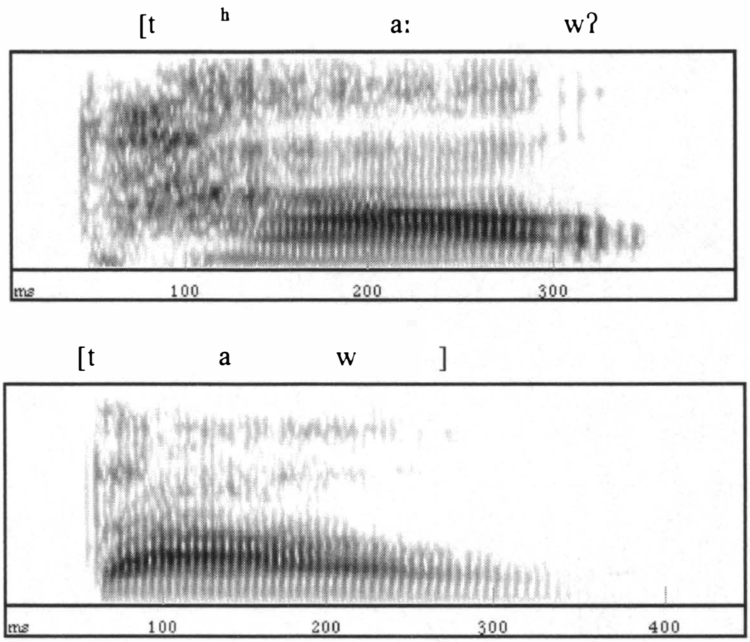


Figure 14: *Spectrograms of [t^{h}a:w̰ʔ^{21}] 'hide' and [taw^{11}] 'louse' (speaker F1)*

In Figure 15, /a:wʔ/ pronounced as [aa̰ʔ] is illustrated. The waveform and the corresponding spectrogram of /ta:wʔ^{21}/ 'master, lord' show a modal phase of /a/ followed by a creaky phase that culminates in the final glottal closure.

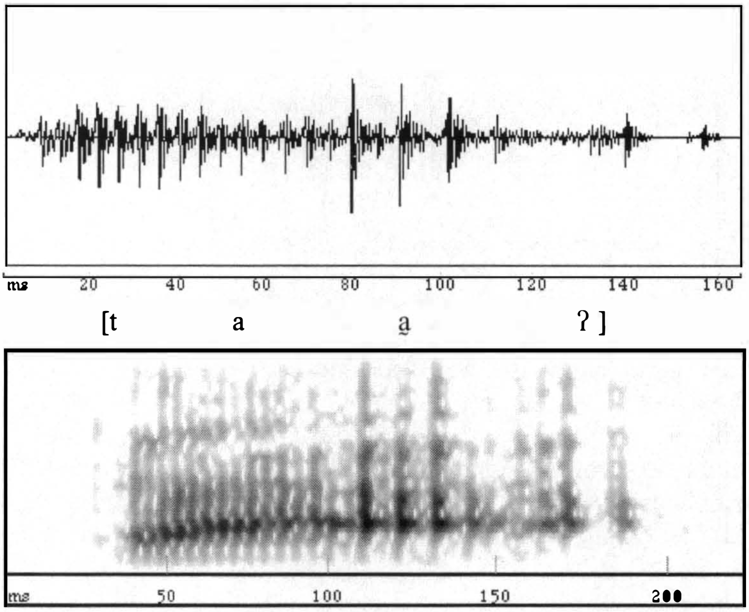


Figure 15: *Waveform and spectrogram of /ta̰:wʔ²¹/ 'master, lord' (speaker F3)*

The /a:wʔ/ can also be pronounced as [a̰:ʔ] with creakiness spreading over the whole vowel. Figure 16 presents the waveform of the creaky vowel [a̰] in /ta:wʔ²¹/ 'master, lord' taken over a 58 millisecond interval centered around the middle of the vowel. As comparison, Figure 17 presents the modal [a] vowel in /ta³³/ 'far' also taken over a 58 millisecond interval centered around the middle of the vowel. Creaky voice in Figure 16 is readily differentiated from the modal voice in Figure 17 by its irregularly spaced glottal pulses and reduced acoustic intensity relative to modal voice.

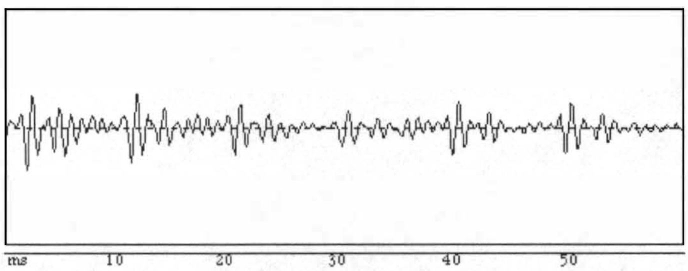


Figure 16: *Waveform of /a/ in /ta:wʔ²¹/ 'master, lord' (speaker M3)*

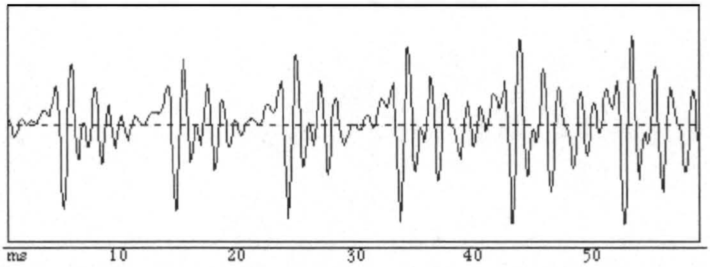


Figure 17: *Waveform of /a/ in /ta^{33}/ 'far' (speaker M3)*

Figures (18-19) show the FFT spectra with the superimposed LPC spectra, measured in the middle of /a/ in /ta:wʔ^{21}/ 'master, lord' and /ta^{33}/ 'far'. The creaky vowel [a] is characterized by a bigger increase in intensity as one moves from H1 to H2, to F1, and to F2 for the creaky vowel (Figure 18). In contrast, the modal voice shows a relatively small increase in intensity moving from H1 to H2 and a drop in intensity moving from F1 to F2. Also, the first harmonic is of lower frequency for the creaky [a̰] than for the modal [a].

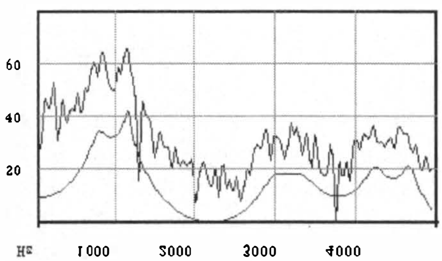


Figure 18: *Creaky [a̰] (speaker M3)*

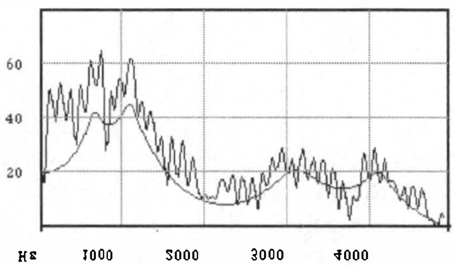


Figure 19: *Modal [a] (speaker M3)*

In Figures 20-21, for clarity [a̰] as the phonetic realization of /a:ʔ/ is referred to as creaky voice1; and, [a̰] as the phonetic realization of /a:wʔ/ is referred to as creaky voice2. The figures show that in terms of spectral characteristics, /a:wʔ/ pronounced as [a̰] patterns with the creaky [a̰] earlier discussed.

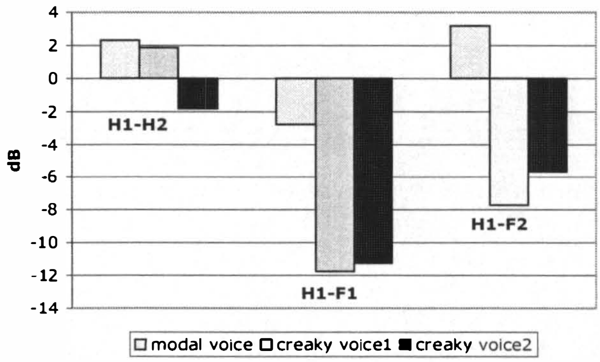


Figure 20: *Differences in H1-H2, H1-F1, and H1-F2 amplitude for female speakers*

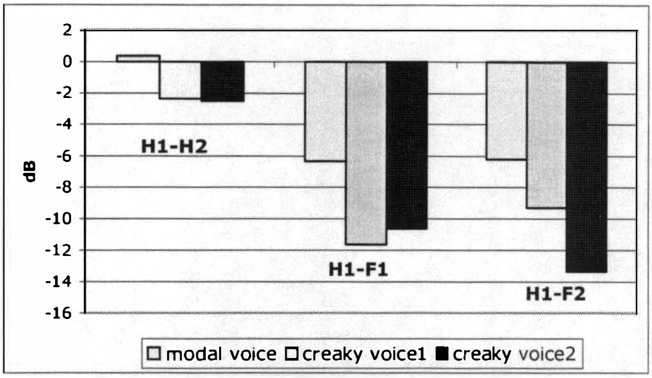


Figure 21: *Differences in H1-H2, H1-F1, and H1-F2 amplitude for male speakers*

Summary of *-a:y? and *-a:w?.

PC *-ay and *-aw developed glottal stops within the history of Tsat. The modern Tsat reflexes are -a:y? and -a:w?, respectively, but both exhibit considerable interspeaker and cross-speaker variation. In all cases, the glottal stop is consistently associated with creaky voice quality. This creaky voice quality begins at the back: sometimes only the final glide is creaky, sometimes the creaky voice ranges forward enough so that the second half of the long [a:] is creaky, and sometimes the whole vowel plus the glide is creaky. Also, in some cases, the final glide is voiceless. And, in some cases, presumably masked by the creaky voice and the voicelessness, the final glide has disappeared completely, leaving only the long monophthong with creaky voice. In summary, the two PC diphthongs *-ay and *-aw developed final glottal stops; in some cases the final glides are manifested as voiceless; and, in other manifestations, the glides have dropped, leaving only a long, monophthong behind. Thus, PC *-ay and *-aw have developed into Tsat -a:y? and -a:w? and this Tsat pair is in the process of developing into the long monophthong -a:?.

Tonogenesis

The history of Tsat tones correlates directly with the interaction of finals, initials, and phonation types. Of these, only the finals are still represented as such in Tsat; the initials and finals, however, are present in the reconstruction of PC, while the phonation types are clear from an examination of the other Chamic dialects.

The three so-called level tones have straightforward origins (Table 8). Items ending in a PC *-h have a 55 tone. Open syllables or syllables ending in a simple nasal, not a preploded final nasal, split into tones 11 and 33. The condition for the split is straightforward: syllables with a PC voiced obstruent initial developed breathy voice, which was accompanied by lower pitch; if in disyllabic items the initial of the first syllable was voiced, this led to breathy voice—not retained in Tsat, then the breathy voice spread to the next syllable, and then the breathy voice produced the low pitch, that is, the 11 tone.

final *-h:	[55~45]
*bah >	p^ha^{55}
*pah >	pa^{55}
other initials:	[33]
*ma >	ma^{33}
*pa >	pa^{33}
voiced obstruent initials:	[11~21]
*dapa (spreading)	pa^{11}
*ba (same syllable)	p^ha^{11}

Table 8: *Pitch patterns of PC *-h, and of open and nasal-final syllables*

The PC final stops *-p, *-t, *-k, and *-ʔ are manifested in Tsat as tones 43 and 24, with the 43 tone emerging if the syllable initial was a voiced obstruent and the 24 tone emerging otherwise, depending upon whether the initial was a voiced obstruent or not; again, if in disyllabic items the initial of the first syllable was voiced, this led to breathy voice, the breathy voice spread to the next syllable, and the breathy voice produced the low pitch, that is, the 43 tone. (see Table 9 below).

other initials:	[24]
*mak >	$maʔ^{24}$
*pak >	$paʔ^{24}$
voiced obstruent initials:	[43]
*bak > (same syllable)	$p^haʔ^{43}$
*dapak > (spreading)	$paʔ^{43}$
*dapay > (spreading from first-syllable)	$p^ha:yʔ^{21}$ (> $p^ha:yʔ^{32}$)

Table 9: *Pitch patterns of PC final stops*

Preploded final nasals.
Another source of final glottal stops is from syllable-final preploded nasals, an earlier feature of Tsat subgroup that was almost gone by the time Ouyang and Zheng (1983a) began working on the language in the 1980s. However, Ouyang and Zheng still managed to record a handful of forms (see Thurgood 1999:165):

tatn^{33} la:n^{11}	'section'	---
tsiakŋ^{33} lai^{11}	'where'	---
t^{h}okŋ^{33}	'knife'	PC ^{x}*dhɔŋ
t^{h}atn^{33}	'extinguish'	PC *padam

Table 10: *The attested forms with preploded final nasals*

The PC nasal finals *-am, *-an, *-aŋ, *-ɔŋ and the resonant finals *-al, and *-ar initially developed into preploded finals and then into final glottal stops. The *-al and *-ar, of course, first went to *-an before becoming preploded final nasals. The development of preploded final nasals dates back to the P-Roglai-Tsat subgroup of Chamic (for discussion of preploded final nasals see Thurgood 1999:164-177), but the subsequent developments described here are unique to Tsat.

The data on Tsat and on Northern Roglai indicates that all final nasals first developed into preploded final stops followed by homorganic nasals and then in most contexts these highly-marked codas were simplified. In fact, except after short ɔ and the short -a- in *-am, *-an, *-aŋ, *-al > *-an, and *-ar > *-an the complex coda was simplified (Thurgood 1999:164-177), leaving *-kŋ and *-tn (the final -m merged with final -n). The *-k- and *-t- then became glottal stops, hence the modern forms. The forms with these glottal finals have developed into 21 and 32 tones, with the split again depending on the presence or absence of an earlier voiced obstruent initial, respectively.

	Proto-Chamic	Tsat	gloss
other initials:	*cam	tsa:nʔ^{32}	'Tsat'
	*masam	sa:nʔ^{32}	'sour; vinegar'
	*klam	kianʔ^{32}	'dark; afternoon'
voiced obstruents:	*dar	t^{h}a:nʔ^{21}	'encircle'
	*padam	t^{h}a:nʔ^{21}	'extinguish'
	*hadaŋ	t^{h}a:ŋʔ^{32}	'charcoal'
voiced obstruent initial spreading	*gunam	na:nʔ^{21}	'cloud'
	*dalam	la:nʔ^{21}	'deep; inside'
	*gatal	ta:nʔ^{21}	'itchy'

Table 11: *Tones from PC forms with glottalized stops with final nasals, *-l, and *-r*

A third internal source was the development of final glottal stops from the epenthesis of the PC diphthongs *-ay and *-aw.

other obstruent initials	[32]
*pay >	pa:yʔ^{32}
voiced obstruent initials	[21]
*bay >	p^{h}a:yʔ^{21} (> p^{h}a:yʔ^{32})

Table 12: *Pitch patterns from PC forms *-ay > -a:yʔ and *-aw > -a:wʔ*

Table 13 provides examples of these developments, along with the PC reconstructions. Some of these forms pre-date PC, but those marked with #* only date back as far as PC.

	Proto-Chamic	Tsat	gloss
other initials	*maray	za:yʔ^{32}	'come'
	*matay	ta:yʔ^{32}	'die'
	*kakay	ka:yʔ^{32}	'foot; leg'
	*tapay	pa:yʔ^{32}	'rice wine'
	*hatay	ta:yʔ^{32}	'liver; heart'
	*haway	va:yʔ^{32}	'rattan'
	*naw	na:wʔ^{32}	'go; walk'
	#*paʔdaw	(kia^{33})ʔda:wʔ^{32}	'warm, hot'
voiced obstruents:	#*gay	k^{h}a:yʔ^{21}	'walking stick'
	*paday	t^{h}a:yʔ^{21}(>^{32})	'rice (paddy)'
	*glay	t^{h}a:yʔ^{21}(>^{32})	'forest; wild'
voiced obstruent initial spreading	*gatal	ta:nʔ^{21}	'itchy'

Table 13: *Examples of tones with glottal stops from PC *-ay and *-aw*

The fourth source of glottal finals is borrowings. That study is in progress; while it is easy to spot many of the borrowings, it is often quite difficult trying to determine where they are borrowed from and when. Presentation and evaluation of these forms will have to be left to another time.

Conclusions

The addition of the data from Zheng (1997) and from the summer 2004 fieldwork gives us a much clearer, much richer picture of the Tsat reflexes of Proto-Chamic and thus of Tsat tonogenesis. The three distinguishable falling tones are of particular interest. The Tsat reflexes of proto-Chamic, although not presented here, require no significant adjustments of Proto-Chamic and are straightforward, making the segmental origins of the various phonemic and subphonemic tones non-problematic. In fact, the relationships, although richer and more detailed, remain remarkably clear.

Several other points of interest emerged. First, it is speculated that the association found in the literature between final glottal stops and falling, rather than rising, tones might be, upon closer inspection an association between final glottal stops co-articulated with an oral closure of some kind, rather than simply a glottal stop. This would certainly account for the Tsat data, and, if co-articulation of oral final stops with glottal stops is as widespread in Southeast Asia as it now appears to be, it would account for many of the reported instances of glottal stops associated with falling contours.

Second, the variation in the Tsat data found in the reflexes of PC *-ay and *-aw show a path from diphthong to diphthong with a final glottal stop, to a monophthong with a final glottal stop. Here, the data are rich enough to posit a plausible path of change for the developments to have followed.

Finally, there is the extra high 55 tone, which is of wider interest largely because of its apparent rarity.

References

Benedict, Paul K. 1941. A Cham colony on the island of Hainan. *Harvard Journal of Asiatic Studies* 4:129-34.

-----. 1984. Austro-Tai parallel: A tonal Cham colony on Hainan. *Computational Analyses of Asian & African Languages* 22:83-86.

Burling, Robbins. 1966. The addition of final stops in the history of Maru. *Language* 42.3:581-586.

Chao, Yuen Ren. 1930. A system of tone letters. *Le Maître Phonetique.* Troisème série. 30.24-27.

Coedès, Georges. 1968. *The Indianized states of Southeast Asia.* Ed. Walter F. Vella. Trans. Susan Brown Cowing. Kuala Lumpur: U. of Malaya Press.

Goschnick, Hella. 1977. Haroi clauses. *Papers in South East Asian Linguistics No. 4: Chamic Studies.* Edited by David Thomas, Ernest W. Lee, and Nguyen Dang Liem. Pacific Linguistics Series A, No. 48:105-124.

Haudricourt, André-G. 1954. De l'origine des tons viêtnamien. *Journal Asiatique* 242:69-82.

-----. 1984. Tones of some languages in Hainan. *Minzu Yuwen* 4:17-25. Also published in *Bulletin de la Société de Linguistique de Paris* as "La tonologie des langues de Hainan." 79.1:385-394.

Jessen, Michael and Justus C. Roux. 2002. Voice quality differences associated with stops and clicks in Xhosa. *Journal of Phonetics* 30, 1-52.

Ladefoged, Peter. 2003. *Phonetic Data Analysis. An Introduction to Fieldwork and Instrumental Techniques.* Blackwell Publishing Ltd.

Maddieson, Ian and Keng-Fong Pang. 1993. Tone in Utsat. *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.

Ni, Dabai. 1988a. The genealogical affiliation of the language of the Hui people in Sanya Hainan. *Minzu Yuwen* 2:18-25.

-----. 1988b. The Kam-Tai of China and the Hui of Hainan. *Bulletin of the Central Institute of Minorities* 3:54-65.

-----. 1990a. The origins of the tones of the Kam-Tai languages. ms.

-----. 1990b. The Sanya (= Utsat) language of Hainan island: a living specimen of a linguistic typological shift. ms.

Ouyang, Jueya and Zheng Yiqing. 1980. *A brief description of Li (Hainan).* Chinese Minority People's Language, Basic Description Series. Beijing.

-----. 1983a. The Huihui speech (Tsat) of the Hui nationality in Yaxian, Hainan. *Minzu Yuwen* 1:30-40.

-----. 1983b. Survey of the Li (=Hlai) languages. Beijing.

Rose, Phil. 1997. A seven-tone dialect in Southern Thai with Super-High: Pakphanang tonal acoustics and physiological inferences. In *Southeast Asian Linguistics Studies in Honour of Vichin Panupong*, edited by Arthur S. Abramson. Chulalongkorn University Press. Pp. 191-208.

Stübel, Hans. 1937. *Die Li-Stämme der Insel Hainan: ein Beitrag zur Volkskunde Süd-Chinas, unter Mitwirkung von P. Meriggi.* Bern: Klinkhardt and Biermanm Verlag. (Reproduced 1976, Taipei, Orient Cultural Services, 2 volumes. (*Asian folklore and social life monographs 83*).

Thurgood, Graham. 1992. From atonal to tonal in Utsat (a Chamic language of Hainan). *Proceedings of the Eighteenth Annual Meeting of the Berkeley Linguistics Society. February 14-17, 1992.* Special Session on the Typology of Tone Languages. Edited by Laura A. Buszard-Welcher, Jonathan Evans, David Peterson, Lionel Wee, and William Weigel. Pp. 145-156.

-----. 1993. Phan Rang Cham and Utsat: tonogenetic themes and variants. *Tonality in Austronesian Languages.* Edited by Jerry Edmondson and Ken Gregerson. Oceanic Linguistics Special Publication No. 24. Honolulu, Hawaii: University of Hawaii Press. Pp. 91-106.

-----. 1996. Language contact and the directionality of internal 'drift': the development of tones and registers in Chamic. *Language* 71.1:1-31.

-----. 1999. *From Ancient Cham to modern dialects: Two thousand years of language contact and change. With an appendix of Chamic reconstructions and loanwords.* Oceanic Linguistics Special Publications, no. 28. June 1999, 6 x 9. ISBN: 0-8248-2131-9. Honolulu: University of Hawai'i Press. 403pp.

-----. 2002. Vietnamese and tonogenesis: Revising the model and the analysis. *Diachronica* XIX.2:333-363.

Thurgood, Graham and Fengxiang Li. forthcoming. From Malayic to Sinitic: The Restructuring of Tsat under Intense Contact. *Proceedings of the Southeast Asian Linguistics Society XII*, University of Arizona, Tempe, Arizona. 13 pp.

-----. 2003. Contact induced variation and syntactic change in the Tsat of Hainan. In Language variation: *Papers on variation and change in the Sinophere and the Indosphere in honour of James A. Matisoff.* Edited by David Bradley, Randy LaPolla, Boyd Michailovsky, and Graham Thurgood. Pacific Linguistics. Research School of Pacific and Asian Studies., Pp. 285-200.

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 Publisher]

Zhangmianyu Yuyin he Cihui (= ZMYYC) Tibeto-Burman phonology and lexicon. 1991. China Social Sciences Press.

Thurgood, G. and Thurgood, E. "The Tones from Proto-Chamic to Tsat [Hainan Cham]: Insights from Zheng 1997 and from Summer 2004 fieldwork". In Grant, A. and Sidwell, P. editors, *Chamic and Beyond: Studies in mainland Austronesian languages.*
PL-569:247-271. Pacific Linguistics, The Australian National University, 2005. DOI:*10.15144/PL-569.247*