PROTO-AUSTROASIATIC CREAKY VOICE

Gérard DIFFLOTH
Cornell University

Vowels are, by their phonetic nature, more fluid and likely to change than consonants. They are free to evolve continuously along several dimensions of the vowel space, unhindered by the articulatory limitations and discontinuities which restrain the consonants.

In the Mon-Khmer family, this inherent fluidity is compounded by the presence of very large vowel systems which are under constant pressure to become simpler, only to be replenished by later changes occurring elsewhere in the phonology. In this way, the occurrence of rich vowel systems, which is a feature of the Mainland-Southeast-Asian linguistic area, tends to survive itself through the centuries, constituting a linguistic era as well.

For these two reasons, the history of Mon-Khmer vocalism remains unknown to this day. An initial attempt (Shorto 1976) had to appeal to "variation" on a scale so large as to raise some doubts. The hard fact remains that the number of Proto-Mon-Khmer vowel correspondences is far too great to permit reconstruction, even with the kind of very rich vowel system we have become used to.

It may well be that a major dimension of the phonology has remained unnoticed until now.

In what follows, I will present evidence supporting the idea that Proto-Austroasiatic had a two-way contrast between creaky voice and clear voice in the vowels of major syllables. This contrast would not be predictable from the quality and length of the vowels, or from the nature of neighboring consonants. In other words, Proto-Austroasiatic was already a register language.

The evidence shows that such a contrast, or phonetic modifications of it, existed in Proto-Katuic, Proto-Pearic, and Proto-Vietic. Proto-Munda is also reconstructed by N. Zide (1976; 1986) as having glottalised vs. plain vowels, irrespective of final consonants. The presumption that something like this existed also in Proto-Austroasiatic is therefore quite strong; however, the etymological evidence at that level presents problems which are not explained at the moment.

It should be noted here that this older creaky vs. clear register contrast is historically unrelated to the breathy vs. clear register systems which developed later on, and independently, in Monic, Lamet, Waic, Khmeric and much of Katuic (due to devoicing of initials), and in Pacoh and most of North Bahnaric (due to vowel

quality shifts). As we shall see, Pearic is unique in cumulating both kinds of register systems.

KATUIC

In Katuic, there are two languages, Ong and Talan, both in the East-Katuic sub-branch, which indicate that we must reconstruct creaky vs. plain rimes at the proto-branch level. Both languages are spoken in Southern Laos and are unfortunately very poorly known.

For Ong, Ferlus (1974) showed that a glottal feature of some sort was needed to explain certain developments in final consonants, and what he called interchangeably "deux tons (ou deux registres)" or "ton glottal" in major vowels. He proposed that this feature was not historically ancient but was conditioned by proto-voiceless initial consonants, and that it became contrastive as a result of devoicing of initial stops. On this point the evidence is weak; there are many counter-examples, and Ferlus concluded "it seems that the process of formation of the glottal tone did not function as well as it should have" (p.117, my translation).

Recently, I have been working on an East Katuic language originally spoken in Talan.[1] This ethnic group is sometimes included under the Lao district name /taʔò:y/, though the language is not the same as Ta-oih.[2]

In Talan, as in Ong and in the same etyma, there is a full series of post-glottalised continuants in word final position: -mʔ -nʔ -ɲʔ -ŋʔ -lʔ -yʔ -wʔ , contrasting with plain ones: -m -n -ɲ -ŋ -l -y -w. There is also a full series of creaky[3] final rimes, noted here: -'m -'n -'ɲ -'ŋ -'l -'y -'w, as well as -'r -'s and -'h which also contrast with plain -r, -s and -h. Postglottalisation and creakiness do not co-occur.

As in Ong, some of the Talan post-glottalised final continuants historically come from final *stops: /kamʔ/ "to bite", /ramʔ/ "to accept", /lɛ:mʔ/ "sister-in-law", /co:mʔ/ "to enclose, gather together (fire)", /siamʔ/ "a bat", are examples of words reconstructed with Proto-Katuic final *-p; /manʔ/ "eye", /takɔ:nʔ/ "monitor lizard", /ro:nʔ/ "mucus", /se:nʔ/ "sap of plant", /həmpianʔ/ "a ring", /prinʔ/ "banana", /cɛnʔ/ "to stab" are examples of final *'t; and /ŋɔ:yʔ/ "to drink", /kayʔ/ "to pick up (fruits)", /to:yʔ/ "to steal", /sa:yʔ/ "to bail water" and /ŋkuayʔ/ "jaw's harp" are examples of words with final *-c in Proto-Katuic and even Proto-Mon-Khmer.

At the same time, there are words with Proto-Katuic *-p, *-t, *-c which keep these finals in Talan to this day. Examples for *-p: /kahe:p/ "centipede", /ʔəp/ "to cook (soup)", /kɨ:p/

"cave", /kabip/ "to keep mouth shut"; for *-t: /mət/ "gall-
bladder", /ʒɔ:t/ "sambhar deer", /mo:t/ "to enter", /cit/ "ten",
/kace:t/ "to die", /so:t/ "blind"; for *-c: /sɛc/ "flesh", /mic/
"ant", /kərdɔ:c/ "to tickle", /huc/ "to sip", /po:c/ "to uproot".
There are thus two sets of correspondances for the finals *-p,
*-t, and *-c.

Most[4] of the remaining postglottalized final continuants
historically come from plain final *nasals preceded by a short
vowel; Examples of -*m: /təm?/ "right (side)", /kalam?/ "hundred",
/rəm?/ "torch"; examples of *-n: /pran?/ "bait", /ban?/ "to take
care of"; examples of *-ɲ: /pɛɲ? / "full", /tɛɲ? / "to hammer";
examples of *-ŋ: /mpiŋ?/ "spider", /triŋ?/ "hornbill", /baŋ?/
"bamboo-shoot", /moŋ?/ "be alive", /raŋ?/ "dried up".

And there are also words with Proto-Katuic *-m, *-n, *-ɲ,
*-ŋ preceded by short vowels which keep these finals as plain
ones in Talan today: examples of *-m: /?um/ *to winnow", /hum/
to bathe", /kam/ "arrow"; examples of *-n: /kan/ "female", /pan/
"thousand"; examples of *-ɲ: /pɛɲ/ "to shoot" /ndrɛɲ/ "body-
louse"; examples of *-ŋ: /duŋ/ "house", /ntuŋ/ "eel", /saŋ/ "to
hear", /kləŋ/ "many". Here too, there are two sets of correspon-
dances for final *-m, *-n, *-ɲ and *-ŋ, when preceded by short
vowels.

Final *nasals preceded by long vowels show again two develop-
ments: in one set of etyma we find creaky rimes; with *-m: /ha:'m/
"blood", /ŋa:'m/ "sweet", /lɔ:'m/ "liver", /pho:'m/ "fragrant";
with *-n: /hɔ:'n/ "to grow", /ma:'n/ "to mold, model", /ko:'n/
"a large rodent"; with *-ɲ: /ta:'ɲ/ "to weave", /ma:'ɲ/ "to
borrow"; with *-ŋ: /?a:'ŋ/ "bright (light)", /de:'ŋ/ "winnowing
basket", /kəntrɔ:'ŋ/ "stairs", /ma:'ŋ/ "young boy".

In the other set, final *nasals preceded by long vowels
remain unchanged; with *-m: /plɒ:m/ "land-leech", /do:m/ "ripe",
/krɤ:m/ "under"; with *-n: /mo:n/ "pimple", /bɒ:n/ "get", /ce:n/
"cooked"; with *-ɲ: /ko:ɲ/ "male", /dɨ:ɲ/ "a long time", /tra:ɲ/
"a type of forest"; with *-ŋ: /pra:ŋ/ "roof-beam", /sɒ:ŋ/ "five",
/kale:ŋ/ "male bird", /dɒ:ŋ/ "to take".

The same thing can be said of words with final *-y, *-r,
*-1, *-s and *-h. Some etyma show up with creaky rimes: /rɔ:'y/
"a fly", /mo:'y/ "one",/ba:'r/ "two", /pənto:'r/ "star", /hua'r/
"stream", /ɲo:'l/ "anteater", /pre:'l/ "mango", /mɔ:'s/ "mosquito",
/ria's/ "root", /bɔ'h/ "ashes", /ca'h/ "to clear land".

Others show up with plain rimes: /talo:y/ "earthworm, /hɒ:y/
"to know", /kato:r/ "ear", /par/ "to fly", /tu:r/ "cobra", /bo:1/
"drunk", /tapo:1/ "seven", /ko:s/ "to shave", /ŋkɒ:s/ "porcupine",
/toh/ "breast", /mɔh/ "nose", /tapah/ "to slap".

With final *-k, there are also two reflexes: -? and -k,
regardless of the length of the preceding vowel: /nta:?/ "tongue",

/tri:?/ "buffalo", /rʊ̃ʔ/ "a toad", /palʊ:ʔ/ "tusk, ivory"; vs.
/ʔa:k/ "a crow", /sʊk/ "hair", /te:k/ "to lead", /kərdʊ:k/ "lazy".

And finally, with open syllables, ie. Proto-Katuic *-ø finals,
some etyma have creakiness, others do not. Phonetically speaking,
the creaky long open vowels of Talan are interrupted, about two
thirds of the way through their duration, by a brief period of
creakiness, after which the same vowel continues, unstressed and
partly devoiced.[5] This type of rime, noted here: -v'v: for con-
venience, is clearly distinct from rimes with a final -ʔ, noted
here: -v:ʔ and -vʔ. Examples with creakiness: /cɔ'ɔ/ "dog",
/pɛ'ɛ/ "three", /hərli'i/ "thorn", /hənta'a/ "hand-span", /tale'e/
"big river"; and without: /tapa:/ "turtle", /ti:/ "hand", /ri:/
"banian", /srɛ:/ "rice-field", /plʊ:/ "head".

Even the rare *-w finals have the same property: /kəta:'w/
"sugar cane" vs. /hənta:w/ "at night".

Therefore, the two sets of correspondances, A and B, occur
with all *finals and are probably not an original feature of the
final consonant itself, but may belong historically to the whole
rime, or to the main vowel:

	*-p *-t *-c *-k	*-m *-n *-ɲ *-ŋ	*-y *-r *-l *-w *-s *-h *-
A: short V: all V: long VV:	-m? -n? -y? -?	-m? -n? -ɲ? -ŋ? -'m -'n -'ɲ -'ŋ	-'y -'r -'l -'w -'s -'h -ʋ
B:	-p -t -c -k	-m -n -ɲ -ŋ	-y -r -l -w -s -h -ʋ

The situation in Ong appears to be practically the same,
except that Ferlus did not notice vowel length as a conditioning
factor; cognates to Talan words of series A regularly have post-
glottalistion or the "glottal tone" in Ong, and cognates to Talan
series B words do not.

In lexical frequency, A and B are about equally distributed:
in a Talan lexicon of over 2,500 items, words of series A slightly
outnumber (51.5%) those of series B (48.5%). This is a useful
indication that we are dealing with a major feature of the phono-
logy, not one which could be conditioned by a particular segment,
a combination of segments, or a minor sub-class of vowels or con-
sonants.

Historically, Talan, like Ong and all other Katuic languages
except Katu, has undergone devoicing of initial *voiced stops;[6]
often, this process creates register contrasts, as in Kuy, Souei,
Bru and Sô for example. In Ferlus's hypothesis, Series A words
in Talan would have *voiceless initials, and series B, *voiced
ones.

From a collection of about 600 Proto-Katuic etyma, the great majority, 79.5%, of Talan words of series A indeed have Proto-Katuic *voiceless initials; typically, they have the kind of initials which give rise to clear phonation in languages like Kuy (Prasert, 1978) or Bru (Theraphan and See, 1980). The remaining cases, 20.5% of series A words with Proto-Katuic *voiced initials, would be exceptions that would have to be explained.

However, Talan etyma of series B show almost exactly the same skewing, and in the same direction: 77.1% have *voiceless proto-initials, and only 22.9% have *voiced initials which give rise to breathy vowels in Kuy for example. In this case, exceptions would be the overwhelming majority, and the "regular" cases would be rather exceptional.

To put it another way, the proportion of A/B in the total set of etyma, 57%, is almost exactly the same as the proportion of A/B among the *voiceless subset, 58.6%, and of A/B among the *voiced subset, 55%. Thus, the distinction A vs. B is insensitive to the contrast *voiceless vs. *voiced, and Ferlus' hypothesis must be abandoned. Some of the examples quoted above, eg. /pɛɲʔ/ "full" which has a *voiced Proto-Katuic initial, while /pɛɲ/ "to shoot" has a *voiceless one, illustrate the point[7].

Another possibility would be the historical conditioning of series A and B by certain classes of proto-vowel qualities. A priori, the odds are against this because both series A and B occur with all the 21 distinct vowels of Talan. But the matter is not simple.

Pacoh, the nearest relative to Talan with abundant and good materials to work with, has a register contrast related to vowel height; see for example the long vowel system (v = clear vs. $\underset{..}{v}$ = breathy, corresponding to "tense" vs. "lax" in the terminology of Watson and Cubuat, 1979):

$$
\begin{array}{ccc}
\underset{..}{\text{i:}} & \underset{..}{\text{ɨ:}} & \underset{..}{\text{u:}} \\[4pt]
\underset{..}{\text{e:}} & \underset{..}{\text{ə:}} & \underset{..}{\text{o:}} \\[4pt]
\underline{\text{e:}} & \underline{\text{ə:}} & \underline{\text{o:}} \\[4pt]
\underline{\text{ɛ:}} & & \underline{\text{ɔ:}} \\[4pt]
 & \underline{\text{a:}} &
\end{array}
$$

The registers of Pacoh are a recent innovation conditioned by *vowel distinctions in a Proto-Katuic system without registers (Diffloth, 1982), but with numerous, probably 21, proto-vowels.

Comparison shows that there is no general correlation between the registers of Pacoh and the A and B series of Talan, although there are systematic gaps along these dimensions, in the vowel-quality correspondances between the two languages.

With a collection of 700 Pacoh-Talan cognates, there are 29 vowel correspondances, not counting those for which there are only two or three examples. This number seems too large for a proto-vowel system without registers, but not quite large enough for a proto-system where every vowel (long, short, diphthong) would exist in two registers, or series.

For some Pacoh vowels there is a single correspondance with Talan; eg. for Pacoh "lax" /i:/, 13 cognates have Talan /i:/ in series B (with 2 exceptions); and for Pacoh "tense" /e:/, 29 cognates have Talan /e:/ also in series B (with 1 exception). But for some other Pacoh vowels, there are two equal correspondances; eg. Pacoh "tense" /o:/ corresponds to Talan /o:/ with 32 cognates in series A, and 44 in series B; and Pacoh "tense" short /a/ corresponds to Talan short /a/ with 41 cognates in series A, and 36 in series B.

A full dicussion of this question would require the presentation of a voluminous etymological lexicon of Katuic. The short answer, in my opinion, is that Proto-Katuic should now be reconstructed with a full set of proto-vowels (as eg. in Diffloth, 1982), but occuring with two distinct registers, creaky and clear. These registers correspond mostly to Talan's A and B series respectively. For some proto-vowel qualities, Talan has lost that register distinction, and this accounts for the single Pacoh-Talan correspondances; for some other vowels, Talan has kept it, accounting for the double correspondances shown above. This would also explain the uneasy number of Pacoh-Talan vowel correspondances, 29 at least.

PEARIC

Huffman (1985) has given a phonemic analysis of Chong spoken in Chanthaburi Province, where a clear vs. breathy vowel distinction intersects a glottalised vs. plain final consonant contrast, giving four possibilities.

I think it is possible to describe Chong as having four registers which contain several phonetic features including phonation shifts, pitch contours and vowel breaking: R1 has clear voice with mid-falling pitch, R2 has breathy voice with mid-low falling pitch, R3 begins with clear voice and ends with tight voice or a glottal catch, and has high-falling pitch, and R4 has breathy voice followed by creaky voice, with mid-low-rising-falling pitch. All four registers are found with almost every final consonant and most vowels.

For historical purposes, Huffman's analysis, which separates out a glottal element from the clear/breathy voice contrast is also useful, especially if his final glottal stop is understood as a feature of the main vowel, not as a portion of the final

consonant.

	Non-glottalised	Glottalised
Clear	v̱ (R1: 44.8%)	v̱' (R3: 16%)
Breathy	v̤ (R2: 30.6%)	ṿ̰ (R4: 8.6%)

Percentages are taken from my earlier collection of about
1,200 Chong words. The predominance of the non-glottalised
registers (75.4%) is actually misleading because it includes the
numerous Khmer borrowings into Chong which are not glottalised[8].
But the ratio of glottalised to non-glottalised words in this
sample is roughly the same in the Clear (26.3%) and in the Breathy
(21.9%) categories. If we knew how to exclude Khmer borrowings,
this ratio would be much higher in both categories.

Other Pearic languages which I have recorded[9], Samrê and
Song ("Chong" of Trat province), also have a four-way contrast
corresponding in most etyma with the four registers of Chong.
We can thus reconstruct four registers at the proto-Pearic level.

Historically, the breathy registers, R2 and R4, of Chong
are due to Pre-Pearic voiced initial consonants, eg. R2: /tak/
"a trap, to trap", /tḁ:k/ "water"[10], /kṳ:m/ "to winnow", /lɔ̤:m/
"liver", /cə̤y/ "resin"[11], /kəma̤:/ "needle", /pri̤:/ "forest",
/plṳ:/ "thigh"; R4: /lɨ̰ŋ/ "deep", /pənɨ̰m/ "ant-hill", /ŋa̰:m/
"sweet", /mo̰:y/ "one", /pa̰:y/ "two", /cɔ̰:p/ "horsefly", /pra̰:y/
"thread", /kəla̰ʔ/ "thorn", /kəma̰ʔ/ "rain"[12], but the glottalised
vs. non-glottalised distinction cannot be explained historically
at the moment[13].

Here also there are systematic gaps: eg. Chong words with
final -h are never found with the glottalised registers R3 and
R5; these two registers can never be preceded by Chong ʔ-, h-,
b-, or d-; Chong long i: and uP are never found in the glottalised
registers R3 and R5 with final velars, although they exist with
other finals, or when short.

There are also a few cases of Chong register ablaut in-
volving glottalisation due to derivational morphology: /khe̱:t/
(R1) "to comb" vs. /kəne̱':t/ (R3) "a comb", /kɛ̤:p/ (R2) "to
roast fish held with sticks" vs. /kəngɛ̰:p/ (R4) "roasting sticks
for fish", /pɛ̤:k/ (R2) "to laugh" vs. /kənɛ̰:k/ (R4) "a joke".
By contrast, Talan never has any alternation of its A and B
series, even though it has a good deal of morphology.

Obviously, there are great differences between the glottalisa-
tion phenomena of Pearic and the *creakiness of Katuic, and also
a great time-depth separating the two; but in both branches a
glottalised or creaky register vs. a plain one can be reconstructed
at the Proto-branch level.

VIETIC

The theory proposed by Haudricourt to explain the origin of Vietnamese tones (Haudricourt, 1954) accounts for many, but not all, known etymologies. It can be summarised[14] in a diagram:

Finals:	*voiced continuants	*stops, *-ʔ	*voiceless fricatives
Initials:			
*voiceless	v (ngang)	v̓ (sǎc)	v̤ (hỏi)
*voiced	v̀ (huyên)	ɣ (nặng)	ṽ (ngã)

Gage (1985, publicly presented in 1980) points out, among other things, that a number of VN words have tones of the ngang-huyền category, even though Mon-Khmer comparison (in this case, Riang data) indicates a final *-ʔ which should result in a tone of the sǎc-nặng category; eg. si "banian tree", sâu "deep", mơ "dream", mày (or mì) "thou". After a survey of the situation in most branches of Mon-Khmer, he concludes that the evidence of Vietnamese tones and Chong data, as he puts it "is not dead against" a Proto-Mon-Khmer contrast[15] between final *-ʔ and *-∅. With this modification, Haudricourt's theory would actually be reinforced. I will propose below a different solution.

Gage also noted, in passing, Chong words having glottalisatic together with a final continuant which line up with Vietnamese words having a sǎc tone; eg. Chong[16]: /chi̱ʼ:n/ VN: chin "cooked", Chong: /pho̱ʼ:n/ VN: bốn "four", Chong: /kəŋa:y/ Middle-VN: ngái "far", Chong: /kəya̱ʼ:y/ VN: gió "wind". But he does not propose any explanation.

This remark actually opens Pandora's box, because the words in question have well-known Mon-Khmer etymologies where the final consonants (*-n, *-n, *-y and *-l respectively) can be reconstructed without question. These are voiced continuants which, in Haudricourt's theory, should give rise in Vietnamese to a ngang tone, not to sǎc.

In fact, there is a good number of such Vietnamese words having reliably reconstructed MK final *-nasals, *-r, *-l and *-y, and VN tones which are not explained by Haudricourt's theory, not only in sǎc, but in the sǎc-nặng category as a whole.

Examples with the sǎc tone: kám "rice-chaff" (*-m), sấm "thunder" (*-m), nhím "porcupine" (*-m), bắn "to shoot" (*-ɲ), rắn "snake" (*-ɲ), dǎng "bitter" (*-ŋ), trứng "egg" (*-ŋ), muối "salt" (*-y), mối "termite" (*-r), gối "knee" (*-l).

Examples with the nặng tone: bụŋ "belly" (*-ŋ), mụn "pimple" (*-n), rận "body-louse" (*-ɲ), miệng "mouth" (*-ŋ), sợi "thread" (*-y), chọi "animals) to fight" (*-l), nhẹ "light (weight)" (*-l)

The problem is not a peculiarity of the Vietnamese language: Mường cognates to these words show that the tonal discrepancies go back at least to Proto-Viet-Mường.

In the Mường Khoi dialect of Central Mường[17] (Soko-lovskaia and Tài, 1987) there are five tones which historically correspond quite regularly to Vietnamese ones in the following manner:

Mường Khói	Tone 1 (high level, 55)		: VN huyền
	Tone 2 (Mid-high falling, 43) merger	:	VN ngang and ngã
	Tone 3 (Mid-low rising, 24 or 23)		: VN sắc
	Tone 4 (Rising, 35)		: VN hỏi
	Tone 5 (Mid falling, 32 or 31)		: VN nặng

The words shown above conform to the pattern, eg. Mường Khói:

- bon 3 "four", khẳm 3 "thunder", ɲim 3 "porcupine", bặɲ 3 "to shoot", sặɲ 3 "snake, dăŋ 3 "bitter", tləŋ 3 "egg", pɔy 3 "salt", moy 3 "termite";
- and also: khɨəy 5 "thread", cɔy 5 "fight", ɲɛl 5 "light" etc.

The same thing is true of the Mường Khền dialect (Barker, 1966); the tone historically corresponding to VN sắc with non-stop finals is described as "low rising constricted" (here: 24?), abd the tone corresponding to VN nặng is described as "high-mid level"[18] (here: 44?); so we find:

- pon 24? "four", cin 24? "ripe", xẳm 24? "thunder", pặɲ 24? "to shoot", thặɲ 24? "snake", tăŋ 24? "bitter", tləŋ 24? "egg", bɔy 24? "salt", mol 24? "termite";
- and also: xẳɲ 44? "body louse", cɔy 44? "fight", ɲɛl 44? "light" etc.

The phonetics of these tones in Mường Khền are worth noting since they both include some form of glottalisation; however, the significance of this detail is diminished by the fact that the high-mid level tone (44?) is the result of a a merger of the VN nặng and ngã tones[19]. Even then, it is clear that we need to introduce a feature of this sort in the proto-language, and to revise Haudricourt's theory, possibly by adding "glottalised continuants" to the list of finals which condition the sắc-nặng tone category in VN.

The really important language in this context would be Arem, about which, unfortunately, very little is known for the moment[20]. According to Nguyễn Van Tai (pers. comm. 1988) Arem is related to Vietnamese, but lies outside the Viet-Mường sub-branch of Vietic[20]. It is apparently the only Vietic language to be without tones, though it may have two phonation types; but it has a full set of

post-glottalised final continuants: -m?, -n?, -ɲ?, -ŋ?, as well
as -w?, -y?, -l? (Arem has no final -r)[22]. Awaiting the publi-
cation of work on this important language by Vietnamese and
Soviet linguists, we can guess that Arem will probably confirm
the antiquity of the glottalisation feature in the Vietic branch
It would also show glottalisation to be originally independent
from the tone systems which developed later.

Returning now to the question raised by Gage about open
finals, we can see that the problem is exactly of the same type
as with final continuants: among the etyma with known Proto-MK
final ,*-?, there are two correspondances: some have VN tones of
the sắc-nặng category, eg. cá "fish", lá "leaf", trái "fruit",
gạo "rice", while others have VN tones of the ngang-huyền catego
eg. ba-ba "turtle", (chiêm) bao "rêver", tay "hand", nhà "house"
ngày "day". Using the evidence of a final *-? in Khmu, Haudri-
court's theory accounts for the first set, but not for the
second[23]. Here again, the Mường data agrees fully with Viet-
namese, and shows this double correspondance to be old.

The evidence for a Proto-Mon-Khmer *-? is just as solid foɪ
the first set as it is for the second; and since reconstructing
a "glottalised" glottal stop seems out of the question, we coulɗ
reconstruct this "glottalisation" as coming from a feature of tɦ
vowel, not of the final. If Proto-Vietic vowels had, like P-Kaʈ
and P-Pearic, a creaky vs. clear contrast, this would be sufficɪ
to account for the double correspondance.

Proto-Vietic may well have unconditionally lost this origiɪ
PMK *-?, acquiring open syllables; but it would have retained tɦ
original *creaky vs. *clear phonological contrast in some form,
which then played a determining role in the tonogenesis of open
syllables, conditioning the two series. In the case of final
continuants, *creakiness would become glottalisation which also
conditioned two tone categories, as discussed above.

We can now propose an enriched version of Haudricourt's
theory, one which accounts for a greater number of Mon-Khmer
etymologies:

Register:	*clear *creaky	(?)	(?)
Finals:	*voiced continuants: open syllables (<*-?)	*stops	*voicele fricativ
Initials:			
*voiceless	ngang	sắc	sắc	hỏi
*voiced	huyền	nặng	nặng	ngã

AUSTROASIATIC

We now have evidence for reconstructing something like a creaky vs. clear voice contrast in four branches of Austroasiatic, but not for Proto-AA itself. At that level, the evidence is unclear and insufficient.

The are sets of etyma where creakiness is found uniformly:

	Katuic:	Pearic:		Vietic:		
gloss:	Talan	Chong	Song	Vietnamese	M.Khói	M.Khền
"four"	pua'n	pho':n	pho':n	bôn	bon 3	pon 24?
"chaff"	ŋka:'m	kə̠ŋa̠':m	–	cám	–	–
"wind"	kaya:'1(^{24})	kəya̠':y	khya̠':1	gió	sɔ 3	sɔ 24?
"dog"	cɔ'ɔ	(cho̠:)	chɔ'ɔ	chó	cɔ 3	cɔ 24?
"leaf"	hala'a	kəla̠?	khĭa̠'a	lá	la 3	lá 24?
"sweet"	ŋa:'m	ŋa̠:m	ŋa̠:m	–*	–	–
"thread"	pra:'y	–	pra:y	sợi	khĭəy 5	–
"one"	mo:'y	mo̠:y	mo̠:y	mọt	moc 5	moc 44
"*thorn*"	hərli'i	kəla̠?	səla̠:	–	–	–
"mosquito"	mɔ:'s	–	mɔ̠':h	muỗi	mɔy 5	mɔy 44?
"termite"	kamua'r	–	–	môi	mol 3	mol 24?
"new"	tamɛ'ɛ	–	–	mối	məy 3	məy 24?

There are other sets where *clear voice agrees:

	Katuic:	Pearic:		Vietic:		
gloss:	Talan	Chong	Song	Vietnamese	M.Khói	M.Khền
"hand"	ti:	ti̠:	ti̠:	tay	say 2	thay 3
"tortoise"	tapa:	kəpha̠:	ləpha̠:	ba-ba	ba2-ba2	–
"child"	–	khe̠:n	khe̠:n	con	kɔn 2	kɔn 3
"nephew"[25]	mɒɒn	–	–	–	–	–
"river"	–	khrɔ̠:ŋ	khrɔ̠:ŋ	sông	khɔŋ 2	xoŋ 3
"arrow"	kam	kha̠m	kha̠m	–	–	–
"year"	–	nĭm	nĭm	năm	năm 2	năm 3
"spicy"	har	ha̠y	ha̠l	–	–	–
"hair"	sɒk	(kə)su̠k	su̠k	tóc	săk 3	thăk 45
"gall"	mʌt	kəma̠t	–	mật	mə̆t 5	mə̆t 44
"nose"	mɔh	mu̠h	mu̠h	mũi	muy 2	muy 44?
"deep"	jru: Baradat:	chrû	srû	sâu	khu 2	xu 3
"pestle"	ndra:y	kəhĭ:	(ləhĭ':)	cáy	khăy 1	xăy 21
"foot"	ɜɣ:ŋ	–	sĭɲ	(chân	cə̆n 2	cə̆n 3)
"gong"	kɒɒŋ	kɔ̠:ŋ	–	–	–	–
"sleep"	bĭc	pĭc	pĭc	–	· –	–

But there are many other sets where these features disagree in one branch or another:

	Katuic:	Pearic:		Vietic:		
gloss:	Talan	Chong	Song	Vietnamese	M.Khoi	M.Khen
"two"	ba:'r	pa:y	pa:	(hai	hal 2	hal 3)
"rain"	–	kəma?	kəmaa	(mữa	mɨə 2	mɨə 3)
"cooked"	(ce:n)	chḭ':n	chḭ':n	chín	–	cin 24?
"bird"	–	chḭ':m	chḭ':m	(chim	cim 2	cim 3)
"weep"	ɲa:'m	(ya:m	ya:m)	–	ɲam 5	ɲam 44?
"a fly"	rɔ:'y	rɔ̰:y	rɔ̰:y	(ruôi	hroi 1	ruəy 21)
"eye"	(man?)	mə̰t	mạt	mắt	mắt 5	mắt 44

A long interim period evidently separates Proto-Austroasiat
from each of the proto-branch levels. I have mentioned a few
systematic distributional differences between the creakiness of
Proto-Katuic, Proto-Pearic and Proto-Vietic; these differences ɛ
probably the result of distinct innovations which occured durinɡ
intermediate periods. If we can discover more of these early
innovations, and identify inter-branch borrowings, we may be abɭ
to explain the inconclusive etymological evidence shown above aɪ
reconstruct creakiness at the PAA level; for the moment, it is
just a presumption.

NOTES

1. Taseng ta-eun, Kong Kaleum, Muang Ta-oy, Saravan Provinɩ

2. The speakers themselves report important differences
between their speech and that of the /ta?ua's/, who are the proɟ
Ta-oih or "High Ta-oih"; the Talan group covers a dozen villageꜱ
around Talan, and is also referred to in Lao as /kata:ŋ ka:ŋ/,
"Middle Katang", although the language differs from Katang, or
"Real Katang". This expression indicates their intermediate
geographic position between "High Ta-oih" and "Real Katang". Tɭ
language has no real autonym other than the village name which ⸲
will use here. *

3. Phonetically, what I have noted as "' + continuant" has
a fairly broad range of variable realisations differing mostly
in the timing of creakiness. It can vary from a glottal catch
occuring in the middle of the final consonant, eg. [-m'm], [-1'ː
to a stretch of creaky voice occuring during the second half of
the preceding vowel. With -'r, -'s and -'h, only this second
type of realisation is heard. Creakiness corresponds to what
Ferlus (1974) calls the Ong "glottal tone". In any case, this
set of creaky final continuants is in phonological contrast wit
the postglottalised, and with the plain final continuants.

4. Lao borrowings, which are not as numerous here as in so
other Katuic languages, are pronounced without tones; but the
postglottalisation which accompanies certain tones of the Pakse
dialect of Lao is preserved in Talan, and this gives rise to a

number of Talan post-glottalised finals, especially -w$^{?}$ and -y$^{?}$, as well as -m?, -n? and -ŋ?. The rare cases of Talan -l? are found mostly in Expressives.

5. See note 1 above.

6. The voiced b- and d- initials found in Talan today are reconstructed as implosive *ɓ- and *ɗ- in Proto-Katuic (Diffloth, 1982), and were not subject to devoicing.

7. compare with: Bru (Theraphan and See, 1980): /pan/ "full" vs. /pan/ "to shoot"; and Kuy (Prasert 1978): /phɛɲ/ "full" vs. /paɲ/ "to shoot".

8. It also includes the numerous Thai borrowings which are likewise unglottalised except for a very few exceptions eg. Chong: /wən/ "day". In Song ("Chong" of Trat province) however, many Thai borrowings (from the Trat dialect of Central Thai) have a glottal catch, corresponding mostly to the C2 and the DL historical Tai tone categories: Song /pɛ'ɛ/ "to lose", /ma'ak/ "betel", /kɔ'ɔk/ "thick".

9. In Samrê, pitch contrasts are more audible than phonation differences, but the four-way distinction is maintained. *I wish to thank here Theraphan L. Thongkum of Chulalongkorn University for the untiring help she has given me through several seasons of field work on Song, Samrê and several dialects of Chong, at different times from 1981 to the present.*

10. These two examples show that Pre-Pearic, like Khmeric, had already merged the older implosive *ɓ- *ɗ- series with the ordinary voiced stop series *b- *d-.

11. Song: /sal/ "resin", from Pre-Pearic *ja'r; cf. Khmer /coa/, Surin Khmer /ɑar/ "resin, sap", Bahnar /jar/ "pitch of tree", Pacoh /car/ "to catch birds with sap", Ta-ang (Palaung): /jər/, Ra-ang (Palaung): /can/ "birdlime trap" etc.; Proto-Mon-Khmer *jar.

12. In Chong, R4 with *-∅ finals gives a short vowel with a final -?; but Song and Samrê show creakiness in the middle a vowel which is kept long and without final consonant; eg. Song: /səlḁ:/ "thorn", /kəmḁ:/ "rain".

13. Ferlus (1980) proposed to explain the glottal feature of Pearic vowels by a second round of devoicing, the first one affecting the consonants themselves. However, he only provides three words to support this idea, and in his "Chong-Pear" notation he does not indicate all the phonation contrasts, showing only two registers instead of four. The examples are "four", "two" and "forest"; in my notation: /pho':n/ (R3), /pa:y/ (R4) and /pri:/ (R2); there is no account of R1, counter-examples abound, and he is forced to reconstruct *mbri: "forest" where

the initial *m- has no support in the rest of Austroasiatic.
Gage (1980) is also unconvinced.

14. For brevity, I am not including the tone correspondances
for original disyllables; as Ferlus (1982) has shown, this has
an effect on the *voicing features, and therefore on tonogenesis,
but it does not affect the major tone categories such as ngang-
huyền or sắc-nặng, which interest us here.

15. I would propose, considering the Proto-Monic, Proto-
Senoic, Proto-Waic, Angkuic, Lamet and Khmuic (especially Mal-
Prai) evidence, that all such words had a Proto-Mon-Khmer final
*-ʔ, and this final was unconditionally lost in Proto-Vietic.

16. These are my notations of the Southern dialect of Chong,
from /ʔamphəə makhǎam/, Chanthaburi province.

17. Commune of An Nghĩa, Lạc Sơn district, Hà Sơn Bình
province.

18. Thompson specifies: "high mid level (ending in glottal
stop unless the syllable ends in another stop)" (Thompson 1976,
p.119).

19. The ngã tone of Vietnamese also has an element of
creakiness or even a glottal catch, but this has no historical
relation to what is discussed here; typologically however, the
proposed *creaky feature was probably a sound of this kind.

20. My information on Arem comes from personal communications
with Nguyễn Văn Tai and Nguyễn Văn Lợi (Linguistic Institute,
Hanoi) and Trân Trí Dôi (Hanoi University) who brought my atten-
tion to this language furing my recent visit to Hanoi (Dec 88),
and with Ilya Peyros (Far-Eastern Institute, Moscow) who informed
me about the work of the late Sokolovskaia during his visit to
Cornell. I take this occasion to thank them for their coopera-
tion and generosity.

21. I use the term "Vietic" for designating the branch of
Mon-Khmer which includes the Viet-Mường sub-branch on the one
hand, and several other related languages such as Arem, Sach,
Poọng (Không Kheng), Cuôi (Uy Lô), Thavửng etc., on the other.

22. Though my information is fragmentary and notations are
not definitive, it is possible to see that Arem cognates to the
above Việt-Mường forms are postglottalised: puən? "four",
krɨm? "thunder", pěɲ? "to shoot", uthěɲ? "snake", bɔy? "salt",
ənrěɲ? "body louse.

23. His original finding that Khmu final glottal stop
corresponds to VN sắc-nặng tones (Khmu: /kaʔ/ "fish", /hlaʔ/
"leaf", /pleʔ/ "fruit", /rŋkoʔ/ "rice grain"), leaves out con-
trary cases where the same Khmu final glottal stop (/tmpaʔ/
"tortoise", /hmpoʔ/ "rever", /tiʔ/ "hand", /sŋiʔ/ "day")

corresponds to VN ngang-huyèn tones.

24. Not found in Talan itself but in a neighbouring dialect of Katang.

25. From PKat *k-rm-ɔ:n, derived from "child" by infixation.

REFERENCES

Barker, M.E., 1966, "Vietnamese-Muong tone correspondances" in: *Studies in comparative Austroasiatic linguistics*, N.H. Zide ed., Mouton, The Hague.

Diffloth, G., 1982, "Registres, dévoisement, timbres vocaliques: leur histoire en Katouique", *Mon-Khmer Studies*, XI: 47-82.

Ferlus, M., 1974, "La langue Ong, mutations consonantiques et transphonologisations" *ASEMI* 5.1.

_____, 1979, "Formation des registres et mutations consonantiques dans les langues mon-khmer", *Mon-Khmer Studies*, VIII: 1-76.

_____, 1982, "Spirantisation des obstruantes médiales et formation du systeme consonantique du vietnamien", *Cahiers de Linguistigue Asie Orientale*, XI: 83-106, Paris.

Gage, W.W., 1985, "Glottal stops and Vietnamese tonogenesis" in: *For Gordon F. Fairbanks*, ; V.Z. Acson and R.L. Leed eds., Oceanic Linguistics Special Publication No.20, University of Hawaii Press, Honolulu

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

Huffman, F.E., 1985, "The phonology of Chong", in: *Southeast Asian Linguistic Studies presented to Andre-G. Haudricourt*, S. Ratankul, D. Thomas, S. Premsirat eds., Mahidol University, Bangkok.

Prasert, Srivises, 1978, *Kui (Suai) - Thai - English dictionary*. Chulalongkorn University, Language Institute; Bangkok. 434 + 54 pp.

Shorto, H.L., 1976, "The vocalism of Proto-Mon-Khmer" in: *Austroasiatic Studies*, P.N. Jenner, L.C. Thompson and S. Starosta eds., Oceanic Linguistics Special Publication No.13, University of Hawaii Press, Honolulu.

Sokolovskaia, N.K. and Nguyên Văn Tài, 1987, [The Mường language; Materials from the Sovieto-Vietnamese linguistic expedition of 1979]. In Russian.

Theraphan L. Thongkum and See Puengpa, 1980, A *Bruu-Thai-English Dictionary*, Chulalongkorn University, Bangkok, 614 pp.

Thompson, L.C., 1976, "Proto-Viet-Muong Phonology", in: *Austroasiatic Studies*, P.N. Jenner, L.C. Thompson and S. Starosta eds., Oceanic Linguistics Special Publication No.13, University of Hawaii Press, Honolulu.

Watson, R. and S., and Cubuat, 1979, *Pacoh Dictionary* Summer Institute of Linguistics, Huntington Beach, 447 pp.

Zide, N., 1976, "3 and 4 in South Munda" in: *Austroasiatic number systems*, G. Diffloth and N. Zide eds. Linguistics. 174. Mouton.

_____, 1986, "A Koraput Munda laryngeal as a conditioning factor for s-loss in Sora-Juray-Gorum" in: A *Festschrift for Hoenigswald*, M. Cardona and N. Zide eds., Narr pubs.

9 February 1989

Dept. of Modern Languages and Linguistics, Merrill Hall, Cornell University, Ithaca, NY 14853, USA