Register and tone in Eastern Cham: evidence from a word game[*]

Marc BRUNELLE
Cornell University, USA

Introduction

Eastern Cham (also Phan Rang Cham) is an Austronesian language spoken on the south-central coast of Vietnam. While Proto-Chamic was originally sesquisyllabic and had no word-level intonational contrasts, Eastern Cham, possibly through contact with neighboring Mon-Khmer languages, has become almost entirely monosyllabic and has developed phonological registers. Recently, it has further been described as incipiently tonal by a few phoneticians and historical linguists (Hoàng 1987, Phu et al. 1992, Thurgood 1996, Thurgood 1999). Although monosyllabicization and the development of register are uncontroversial, the phonological status of Eastern Cham 'tones' remains unclear. Claims about tonogenesis seem to stem from misinterpretations of the term 'tone' in descriptions where it was not meant to be equated with 'contrastive tone' (Blood 1967, Moussay 1971). Despite the fact that the authors of the only acoustic study of pitch in Eastern Cham have been careful not to draw firm conclusions about the phonemic status of the pitch contours resulting from the interaction of registers and codas (Phu et al. 1992), it is generally held that Eastern Cham has an incipient, if not full-fledged, tonal system.

Additional evidence is needed to determine whether Eastern Cham pitch contours are phonemic or not. One way to obtain such evidence is to carry out further acoustic research (Brunelle, *in progress*). Another type of evidence comes from phonological processes. In this squib, I present data from a word game that informs us about the phonological status of Eastern Cham registers and tones. The reliability of external evidence and, more specifically of word games, for phonological analyses is well established (Campbell 1986, Hombert 1986),

Hombert going as far as teaching new word games to his consultants in order to discover the underpinnings of the phonology of their languages. I have adopted a more conservative approach: the word game described here is already known by all members of the speech community.

1. Registers

Before introducing the word game and showing its relevance to the questions raised above, I briefly present the diachronic developments that gave rise to the phonological registers of Eastern Cham. This will shed light on their phonological structure.

1.1 Origin of register

The formation of register in Cham is parallel to developments in several Mon-Khmer languages (Ferlus 1979, Gregerson 1973, Huffman 1976, Thongkum 1988). Proto-Chamic originally had a voicing contrast in onset position. This contrast was lost, but the resulting neutralization was compensated for by the development of two phonological registers, i.e. two bundles of phonological features realized on vowels. The vowels following original voiceless stops came to be realized with high pitch and clear phonation (High register), while the vowels following previously voiced stops acquired low pitch and breathy phonation (Low register) (Phu et al. 1992). Low register vowels also seem to have a longer duration and a lower first formant than their high-register counterparts (Brunelle, *in progress*). Figure 1 exemplifies the formation of register in Cham. In my transcription, the low register is marked by a subscript dot under the onset and the high register is unmarked, following Moussay (1971).

	Proto-Chamic		Ancient Cham	Gloss
	*păk	>	păk	'at'
	*băk	>	p̣ăk	'full'

Figure 1. Registrogenesis

This relatively simple scenario was rapidly complicated by the left-to-right spreading of the low register through sonorants in polysyllabic words (Figure 2). Originally, the register of syllables with an onset sonorant was non-contrastive and predictable from preceding syllables, but the situation has now become more complex due to the monosyllabicization of a large majority of words in Colloquial Eastern Cham. As shown in Figure 2, the loss of the first syllable has triggered the extension of the register contrast to sonorants. In the current system,

words with any onset except a series of implosive stops and preglottalized glides (ɓ, ɗ, ʃ, ʔj, ʔw) exhibit register contrast.[1]

Proto-Chamic		Ancient Cham		Mod. E. Cham	Gloss
*naw	>	naw	[naw]	> naw	'to go'
*danaw	>	tanaw	[taṇaw]	> ṇaw	'pond, lake'

Figure 2. Register spreading and monosyllabicization

Keeping in mind that Eastern Cham register originates from onsets but is now realized on vowels, there are three logical analyses of its phonological status in modern Eastern Cham; 1) it could still be a property of onsets; 2) it could have become a feature of the vowel; 3) it could have been delinked from both onsets and vowels, effectively turning into a suprasegmental element that could be analyzed as a form of tone. The word game I discuss below strongly suggests that the first possibility is the right analysis.

1.2 Evidence from a word game

Crucial evidence on the phonological status of Eastern Cham register is revealed by a word game called "ɗom ḳac" '*speak – inverted*' or 'inverted speech'. In this word game, variants of which are very common throughout Southeast Asia, the rimes and onsets of a pair or a short string of words are reversed, which results in a new phrase that has a comical and often sexually-explicit meaning.

I first asked my consultants to give me well-known examples of the word game without providing them with expressions to permute. Examples (a)-(e) illustrate the word game with pairs of same-register words. Note that in (b)-(e), the disyllabic forms of written Cham (WC)[2] are realized as monosyllables in colloquial Eastern Cham (CC) and that, in some cases, as in (b)-(c) and (h), (k) and (l), the rime is slightly modified, which does not affect the humorous impact of the game.

a. WC naw puh   nuh paw
    CC **naw puh**   **nuh paw**
    go - dry rice field   plant - trap
    go to the dry rice field   *set a trap*

---

[1]Contrary to plain voiced stops, implosives did not undergo devoicing. Preglottalized glides block register-spreading.
[2]The Written Cham forms are listed for reference only. They are not crucial to my argument.

b. WC cɔh takuh   cuh ʔakɔʔ
    CC cɔh kuh   cuh kɔʔ
    dig out - mouse   light up - head
    hunt mice   *set someone's head on fire*

c. WC laʔu thɛh   ʔɛ̆ʔ thu
    CC ʔu   thɛh   ʔɛ̆ʔ thu
    coconut - split   faeces - dry
    split coconut   *dry faeces*

d. WC ɲɛh thɔh   ɲɔh ʔathɛh
    CC ɲɛh thɔh   ɲɔh thɛh
    to elbow - only   fornicate with - horse
    hit in the air with the elbow   *fornicates with a horse*

e. WC haj ʔɛ̆ʔ puc   huc ʔɛ̆ʔ tapaj
    CC haj ʔɛ̆ʔ puc   huc ʔɛ̆ʔ paj
    excl. - faeces - kind of racoon   suck - faeces - rabbit
    ha! racoon faeces   *suck rabbit faeces*

Examples (a)-(e) do not reveal anything about the issues at stake here. They simply show that onsets and rimes can be permuted. In examples (f)-(l), however, the two words of the original phrase have different registers. We see in these examples that when rimes and onsets are inverted, register always stays with the onset. This supports the claim that the two elements are not phonemically distinct.

f. WC ŋuj paʔ (talăj paʔ)   ŋă̆ʔ puj
    CC ŋuj paʔ   ŋă̆ʔ puj
    wear - chest band   make – happy
    wear a chest band   *make happy*

g. WC çah lɔh   çɔh kalah (kalah nih)
    CC çah lɔh   çɔh lah
    clear (forest) - chukrasia   break - lean on each other
                                          (buffaloes)
    pull out chukrasia plants   *push each other down* (buffaloes)

h. WC t̪ăj saʔaj   t̪aj ʔi
    CC t̪ăj ʔaj   t̪aj ʔi
    younger sibling - older sibling   rock - rice basket
                                          (used as a hammock for babies)
    siblings   *rock a baby in a basket*

i. WC naw p̥ac   nac kap̥aw
    CC naw p̥ac   nac p̥aw
    go - study   give birth - buffalo
    go study   *give birth to a buffalo /*
                                        female buffalo

j. WC p̥u klɔh   p̥ɔh klu
    CC p̥u klɔh   p̥ɔh klu
    congee - cut, separate   fruit - testicle
    congee with small noodles   *testicle*

k. WC p̥u ʔathăw   ʔap̥aw thu
    CC p̥u thăw   p̥aw   thu
    congee - dog   snail - dry
    congee for dogs   *dry snail*

l. WC k̥aj klɔŋ   k̥ɔ̆ŋ klaj
    CC k̥aj klɔŋ   k̥ɔ̆ŋ klaj
    club   rutting - penis
    club   *erect penis*

        In order to rule out the possibility that the word game is lexically frozen
and reflects an earlier stage of the language or that the absence of cases where
register stays with the vowel is accidental, I provided the subjects with pairs of
input words and asked them to play the word game with them, regardless of the
semantic well-formedness of their output. Once again, register always stayed with
onsets. I also suggested outcomes in which register goes with the vowel, but they
were systematically rejected. A few examples are given in (m)-(o).

m. WC naw ç̥alan
    CC naw k̥lan   nan k̥law   * n̥an klaw
    to go - road   meaningless   meaningless
    go on the road

n. WC wăn k̥laj
    CC wăn k̥laj   waj k̥lăn   * w̥aj klăn
    forget - forest   meaningless   meaningless
    get lost

o. WC mɔt pḁtaj
    CC mɔt t̥aj   maj t̥ɔt   * m̥aj tɔt
    harvest (VN) - rice   meaningless   meaningless
    harvest rice

Examples (f)-(o) clearly show that register can never be dissociated from the onset. Therefore, register is neither suprasegmental nor vocalic and the realization of register features such as pitch, voice quality and duration on the vowel is not a phonological property of the vowel itself, but is rather a matter of phonetic implementation of phonological representations.

2. Tones

I now review the literature on the alleged tones of Eastern Cham and introduce evidence from the same word game to show that although there are coda-conditioned pitch variations in this language, they cannot be analyzed as phonemic.

2.1 Claims about tonogenesis in Eastern Cham

Scholars agree that each of the two registers exhibits allophonic pitch variation conditioned by final consonants: codas affect the pitch contour of vowels. However, they disagree on the number and phonological status of these allotones (i.e. melodic allophones). Thus, according to Blood (1967: p.29): 'Before final stops and h the register of non-low pitch is higher than in syllables ending in other consonants or silence'. On the contrary, Phu et al. (1992), the only authors to provide experimental evidence, claim that it is the low register rather than the high one that is split into two tones. Moussay (1971) argues for a more symmetrical four-allotone system where each register has an allotone in front of laryngeals and another allotone elsewhere. Hoàng (1987) goes even further by proposing five allotones.

Some authors have recently suggested that the predictable allotones of Eastern Cham have become (or are becoming) distinctive due to the loss or phonological reanalysis of some codas (Hoàng 1987, Phu et al. 1992, Thurgood 1996, Thurgood 1999). There is no consensus on the description of the pitch variants and on the phonological status and number of Cham tones. Regardless of the details, the crucial point here is that tone can become distinctive if and only if some codas, more specifically laryngeals, are deleted or lose their segmental status. Blood (1967), Moussay (1971) and Bùi (1996) assume that codas are preserved and therefore argue for a two-register system in which each register has a few predictable allotones. Hoàng (1987), on the other hand, seems to propose an *analysis with four distinctive tones (but five allotones) based on the tacit* assumption that the glottal stop has become a tonal element. Also considering the possibility that the glottal stop has become suprasegmental, i.e. that it has become an integral part of a laryngealized tone, Phu et al. (1992) put forward a three-tone analysis in which the high register has only one distinctive tone, while the low register might have two. They are, however, more cautious than Hoàng in that their tonal analysis holds only "if one treats glottal finals as part of the internal stuff of a given tone […]" (p.41). The status of the laryngeal coda [-h] in their analysis is

unclear, but at any rate, their claim that coda "h has been lost altogether" (p.35) is unsupported, as evidenced by the investigation of a corpus of wordlists and short interviews recorded with 43 speakers of Eastern Cham (Brunelle, *in progress*). This investigation clearly shows that the laryngeal codas [-h] and [-ʔ] are not deleted, be it in careful or colloquial speech. The only word in which a laryngeal coda is sometimes dropped is the preposition [pǎʔ] 'at', a high frequency function word.

If register were already a suprasegmental property, i.e. a type of tone (Thurgood 1996), an additional tone split conditioned by the loss or weakening of laryngeal codas would not be surprising. It would simply increase the number of elements of a pre-existing phonological category. However, the word game discussed above shows that register is a property of onsets, which rules out such a straightforward account. Therefore, knowing that register does not behave like a suprasegmental feature and that [-ʔ] and [-h] are still realized, there are three possible analyses of the status of laryngeal codas and allotones:

1) Laryngeal codas are still segmental and are phonologically independent of register, although they condition register allophony.
2) Laryngeal codas are phonologically independent of register, but they have become suprasegmental, i.e. they have triggered tonogenesis and become a part of the tones that they have conditioned.
3) Laryngeal codas have triggered tonogenesis and become suprasegmental. However, in contrast with 2), registers and tones are not independent; they have merged into complex prosodic units that are linked to onsets.

As shown in the next section, additional evidence from the word game supports the first analysis.

2.2 Evidence from the word game

If the third analysis of section 2.1 were correct, the tones and their laryngeal accompaniments would have merged with registers into unified phonemic elements and would be predicted to behave as such in the word game. The following examples show that this position is untenable. In examples (p)-(r) laryngeal codas and registers are manipulated independently ((p)-(q) are repetitions of (f) and (j) respectively). This is good evidence that they have not merged into laryngealized tones.

p.	WC ŋuj p̬aʔ (talǎj p̬aʔ)	ŋǎʔ p̬uj
	CC **ŋuj p̬aʔ**	ŋǎʔ **p̬uj**
	wear - chest band	make - happy
	wear a chest band	*make happy*

q. WC pu klɔh   pɔh klu
    CC pu klɔh   pɔh klu
    congee - cut, separate   fruit - testicle
    congee with small noodles   *testicle*

r. WC pilɔʔ liʔu   plu ʔɔʔ
    CC plɔʔ ʔu   plu ʔɔʔ
    nut - coconut   jar – pearl
    coconut   *pearl jar*

Registers and laryngeal codas are therefore independent. This is further supported by the fact that when playing the word game with an input phrase, its allotones are not transposed in their original form on the output phrase, as we would expect if they had become phonemic. On the contrary, the allotones of the output phrase are fully predictable from the interaction of the codas and the onsets (including their register) of its constituent words. In examples (s)-(t), register conditions the tonal height at vowel onset (low and high registers condition low and high pitch respectively) while codas condition the contour throughout the vowel (a syllable closed by /-ʔ/ or /-h/ has a rising contour, an open syllable, a level one).

s. WC pilɔʔ liʔu   plu ʔɔʔ
    CC plɔʔ ʔu   plu ʔɔʔ
    low rising – high level   **low level – high rising**
    nut - coconut   jar – pearl
    coconut   *jar of pearls*

t. WC pu klɔh   pɔh klu
    CC pu klɔh   pɔh klu
    low level – high rising   **low rising – high level**
    congee - cut, separate   fruit - testicle
    congee with small noodles   *testicle*

Now that we have established that laryngeal codas did not merge with registers, we still have to address the issue of their phonological status. Are they still codas or have they become suprasegmental (i.e. tonal)? The answer to this question can be inferred from the fact that the word game never allows the separation of a nucleus and a coda belonging to the same rime. Since consultants did not spontaneously produce examples of the word game in which laryngeal codas are separated from their nuclei, it is likely that laryngeals are not floating tonal elements, but behave like all other codas. However, to make sure that the absence of floating codas is not a mere coincidence, I coined phrases like (u)-(v), in which laryngeal codas are moved out of their rimes (second column) and in which both registers and laryngeal codas are inversed without manipulating the

onset and the vowel (third column). The consultants rejected them as ill-formed, despite being aware that the outputs were not supposed to have a meaning and that they could produce meaningless forms themselves. This leads me to conclude that final laryngeals are integral parts of the indivisible rimes, i.e. codas.

u. WC thɔ̃w plo?
    CC thɔ̃w plo?   *thɔ̃w? plo   *thɔ̃w? plo
    dog – mould grey   meaningless   meaningless
    grey dog

v. WC pilɔ? li?u
    CC plɔ? ?u   *plɔ ?u?   *plɔ ?u?
    nut - coconut   meaningless   meaningless
    coconut

However, there is another possible explanation for the fact that final laryngeals are never separated from the rest of the rime. One could claim that laryngeals are suprasegmental, i.e. tonal elements, but that the word game works in such a way that tones cannot be delinked from their original tone-bearing unit (the vowel). There is no independent evidence in favor of this alternative explanation, but a partial answer comes from the fact that several Southeast Asian languages have word games similar to "ɗom kac", and that they usually allow tones to be moved independently from segmental material, while prohibiting the division of the rime. The most relevant example is the Vietnamese word game *nói lái*, in which tones, onsets and rimes can be reversed, but rime dislocation is prohibited.[3] For example, the word *giải phóng* 'liberation', with a curve and a rising tones, can be turned into *dái phỏng* 'burnt testicle', with identical segmental material but reversed tones.[4] The fact that the Cham, who know and play the Vietnamese language game, never move laryngeal codas out of their rimes in their own word game argues against a tonal account of these codas. However, it is the absence of any positive evidence in favor of a suprasegmental analysis of final laryngeals that ultimately convinces us to adopt the more conservative and partly supported segmental analysis.

We must therefore reject claims that laryngeal codas have lost their segmental status and that Eastern Cham is tonal or incipiently tonal. The data presented above show that laryngeal codas should instead be treated as segments and register as a melodic property of onsets.

---

[3]Rare cases of rime dislocation have been reported by Maken and Nguyễn (2005).
[4]In Vietnamese, *gi* and *d* have the same orthographic value.

3. Conclusion

The word game described in this squib sheds light on the phonological status of Eastern Cham register. It proves that the registers of Eastern Cham are still a phonological property of onsets and have not evolved into suprasegmental elements. Moreover, there is no evidence that laryngeal codas have been reanalyzed as "part of the internal stuff of a given tone" (Phu et al., 1992) or that they behave differently from other codas, which seriously challenges claims that Eastern Cham is tonal. We thus have to conclude that registers are still relatively conservative phonologically in that they are still a feature of consonants. However, Eastern Cham has evolved significantly and is no longer typical of register languages in two respects. While many Southeast Asian languages restrict register contrast to onset stops, register spreading and monosyllabicization have extended it to onset sonorants in Eastern Cham (Brunelle, *in progress*). Perhaps even more importantly, Eastern Cham register is phonetically implemented on rimes and one of its most salient phonetic correlate is pitch, whereas other register languages often emphasize voice or vowel quality distinctions. As a consequence, the language superficially sounds tonal, although the word game reveals that its pitch distinctions cannot be analyzed as tones.

REFERENCES

Blood, David Livingston. 1967. "Phonological units in Cham." *AL* 9:15-32.

Bùi, Khánh Thế. 1996. *Ngữ Phap Tiếng Chăm.* Hà Nội: Nhà Xuất Bản Giáo Dục.

Brunelle, Marc. In progress. "Phonological register in Eastern Cham: a phonetic and sociolinguistic approach." Ph.D. dissertation, Cornell University.

Campbell, Lyle. 1986. "Testing phonology in the field." In Jeri J. Jager John J. Ohala, ed., *Experimental Phonology*, pp. 163-173.

Ferlus, Michel. 1979. "Formation des registres et mutations consonantiques dans les langues Mon-Khmer." *MKS* VIII:1-76.

Gregerson, Kenneth J. 1973. "Tongue-root and register in Mon-Khmer." In Philip N. Jenner, Laurence Thompson and Stanley Starosta, eds., *AS*, pp. 323-369.

Hoàng, Thị Châu. 1987. "Hệ thống thanh điệu tiếng Chàm và các kí hiệu." *NN*:31-35.

Hombert, Jean-Marie. 1986. "Word games: some implications for analysis of tone and other phonological constructs." In Jeri J. Jager John J. Ohala, ed., *Experimental Phonology*, pp. 175-186.

Huffman, Franklin. 1976. "The register problem in fifteen Mon-Khmer languages." *Oceanic Linguistics Special Publication Austroasiatic Studies*, part 1:575-589.

Maken, Marlys A. and Nguyễn Thị Hành. 2005. *Phonological Constituents.* Paper presented at the LSA annual meeting, Oakland.

Moussay, Gérard. 1971. *Dictionnaire Cam-Vietnamien-Français*. Phan Rang: Trung-tâm Væn hoá Chæm.

Phu, Van Han, Jerold Edmondson and Kenneth Gregerson. 1992. "Eastern Cham as a tone language." *MKS* 20:31-43.

Thongkum, Therapan L. 1988. "Phonation types in Mon-Khmer languages." In Osamu Fujimura, ed., *Vocal Fold Physiology: Vocal Physiology, Voice Production, Mechanisms and Functions*, pp. 319-334.

Thurgood, Graham. 1996. "Language contact and the directionality of internal drift: the development of tones and registers in Chamic." *Language* 72: 1-31.

Thurgood, Graham. 1999. *From Ancient Cham to Modern Dialects: Two Thousand Years of Language Contact and Change*: Oceanic Linguistics Special Publication; no. 28. Honolulu: University of Hawai'i Press.

Department of Linguistics
Cornell University
USA
<mb236@cornell.edu>