

# CREATIVE FORCES IN KHMER

Noeurng Oum

*Regions Hospital, St. Paul*

John Haiman

*Macalester College, St. Paul*

*<Haiman@Macalester.edu>*

## 1 An important characteristic of Khmer (*leakkhana' piseh*)

The forces of erosion run rampant in spoken Khmer with unexampled viciousness. In casual registers (or perhaps routinely in the speech of younger speakers) the unstressed initial syllable CVN- or CrV- is typically reduced to Cə, C, or zero (cf. Huffman 1970, *passim*).

Speaker N.	Speaker P.	
<i>kawndaal</i>	<i>kədaa</i>	“middle”
<i>krawdaah</i>	<i>kədaa</i>	“paper”
<i>bawnthaem</i>	<i>(p)thaem</i>	“add”
<i>bawntaaw</i>	<i>taaw</i>	“continue(transitive)”
<i>bawbaaw</i>	<i>baaw</i>	“rice gruel, porridge”

For speaker P., All consonant clusters are simplified in at least the following ways:

- a) [h] is lost after affricates
- b) stops are elided before stops and nasals
- c) affricates are simplified to fricatives before obstruents
- d) [r] is elided after stops

Speaker N.	Speaker P.	
<i>chiem</i>	<i>ciəm</i>	“blood”
<i>pteah</i>	<i>tea</i>	“house”
<i>knjom</i>	<i>njom</i>	“I”
<i>kmuej</i>	<i>muej</i>	“nephew”
<i>ckae</i>	<i>skae</i>	“dog”
<i>cngawl</i>	<i>sngawl</i>	“wonder”
<i>cmooh</i>	<i>smue</i>	“name”
<i>kawntraj</i>	<i>kətaj</i>	“scissors”
<i>trawlawp</i>	<i>tawlawp</i>	“return”

(Change (c) is perhaps a Vietnamese-influenced dialect: all the other changes are general.)

In all dialects, final [r] survives only in the orthography; final written [s] in all but the most formal speaking styles is lenited to [h]; final [h] is often entirely elided; and final stops /p/, /t/, /k/, all unreleased, are approaching near acoustic identity with each other and with the glottal stop.

Written	Spoken	
<i>kmaer</i>	<i>kmae</i>	“Khmer”
<i>psaar</i>	<i>psaa</i>	“market”
<i>cas</i>	<i>cah</i>	“old”
<i>nih</i>	<i>ni</i>	“this”
<i>kawmsawt</i>	<i>kawmsaw’</i>	“wretched”
<i>peek</i>	<i>pee’</i>	“too (much)”
<i>kooraup</i>	<i>koorau’</i>	“salute”

In a small number of common words, final consonants are optionally elided in the conservative pronunciation of speaker N. and are no longer part of the word at all for speaker P.

Speaker N.	Speaker P,	
<i>jaau(k)</i>	<i>jaau</i>	“get, take”
<i>meeu(l)</i>	<i>meeu</i>	“look”
<i>maau(k)</i>	<i>maau</i>	“come”

Channeled as it is, the Principle of Least Effort in Khmer is not only alive, but well on the way towards reducing the canonical structure of the sesquisyllabic word from CVN+C(r)VC to an open monosyllable CV(‘).

As Bloomfield (1933:370 et passim) among others clearly recognized, an enormous number of sound changes that have been attested are compatible with the Principle of Least Effort. Indeed, the method of reconstruction using the principle of *lectio difficilior* as our best bet for the ancestral form (the PLE in reverse, as it were) silently encourages the assumption that most changes are of this type, which is one reason why our reconstructions of PIE look so unpronounceable. Part of the reason why the PLE of Zipf and phoneticians like Passy is nevertheless not more generally recognized as a linguistically significant tendency is that no language has yet been reduced to anything like silence. But the true reason for the fact that languages continue to be vocal, we suggest, is not that the PLE is invalid, nor that sound change is inhibited or reversed in all but the most extreme cases (Bloomfield, 395-6; Bolinger 1975: 438) but that there exist creative forces which are forever building up phonetic structure at the same time that sound change is wearing it down. Although these forces are less systematic than those of sound change and analogy, they must be active, or Khmer speakers at least would soon find themselves saying very little. And we believe that in Khmer, some of these forces are very much in the open.

## 2 Another characteristic feature of Khmer: infixation

One of these changes, John argued in an earlier report, is analogy itself. The productive nominalizing infix -Vm(n)- ~ -VN-, attested in nouns like *c-awm-rieng* “song” (< *crieng* “sing”). *c-awm-hang* “(monk’s) food” (< *chang* “(monk) eat”) may be the result of a kind of backformation. Given alternating pronunciations [C(VN)CVC] for the same etymon, the elided syllable coda of the unstressed syllable may have been reinterpreted as a meaningful morpheme, and then inserted into words where it had no etymological pedigree (Haiman

1998). This process may have played a considerable part in preserving the sesquisyllabic word in Khmer, as opposed to its loss in related Mon-Khmer languages like Vietnamese.

Another closely related change, however, may be that of purely DECORATIVE infixation. While in the majority of cases, the infix -Vm(n)- ~ -VN- can indeed be analyzed as a derivational morpheme of some kind, there exist a number of other cases where this infix seems to have little cognitive meaning, or perhaps none whatsoever. In another paper, we have called these cases of “syntactic backsliding” inasmuch as what looks to be a “deverbal noun” is in fact syntactically acting exactly like the verb from which it is presumably “derived” (Haiman & Ourn 2003). But it may be that cases of this sort are not as perverse as the label “backsliding” may suggest, if the infixation has no cognitive function to begin with. Noeurng’s intuitions about the meanings of the infixes in examples such as these are practically ineffable. They may mean SOMETHING but whatever that something is, it cannot be characterized either syntactically (as a nominalizing morpheme, for example) or semantically:

*kmaoc bejsaac*      *k-awmn-aac*  
ghost spirit      vicious-ness  
“vicious ghosts and evil spirits”

*neak c-um-ngww*  
person sick-ness  
“sick person; a patient”

*koo nji*      *s-awm-kaaum*  
cow female      skinny-ness  
“skinny female cow”

*ktaaum*      *l-um-haau*      *muej*  
hut      empti-ness      one  
“a hut without walls”

*knong ptej*      *s-awm-ŋam*      *nej*      *rietrej*  
in surface silence      of      night  
“in the silent surface of the night”

We hesitate to say that infixation in these cases is purely decorative, partly because unsystematic differences of meaning are often associated with infixation, and partly because Noeurng does not feel it to be *sawmnuən vauhaa* “elegant style”, which *bawnthaem Ibaoj* “adds flavor”, an institutionalized and hence recognizable Khmer stylistic category on which speakers agree.

### 3 A third characteristic feature of Khmer: Symmetrical Compounds

But we will now present other cases which can be explicitly labelled in this way. In particular, these are cases of decorative REPETITION. We have suggested (Ourn & Haiman 2000, Haiman & Ourn 2002) that Khmer is a language which “likes to say

everything (at least) twice”. In those earlier articles, we focussed on two kinds of such at least partially decorative repetition:

- a) genuine compound forms (*samah*): these are synonym pairs like “cease and desist”: e.g. *lwen rauhah* “quick fast”;
- b) fake compound forms: these are typically alliterated twin forms like “spic ‘n’ span” which consist of at most one meaningful root and a meaningless “servant word” (*bo’ri’waa sap*) :e.g. *rauneeñ raunoon* “dangle”, whose first member is meaningless.

Both real and fake compounds exhibit both type and token frequency. A typical page of literary Khmer will have four or five genuine compound forms. Fake compounds do not seem to occur so often in texts, but are well-represented in the total lexicon. In a card index file that John has been keeping since we began working together, 24 out of the 202 entries beginning with [rau..], 8 out of 168 entries beginning with [caw..], and 12 out of 191 entries beginning with [tr..] are servant word compounds. Nor are combinations of both real and fake compounds excluded: triplets like *lwen rauhah rauhuen* “quick fast schmast” are not too unusual.

While the explicit motivation for some of these {A+B} compounds may have been partly that of elegance, there is always the (sometimes remote?) possibility that the pairing actually meant something. For example, perhaps in genuine compounds, A and B are not totally synonymous: in that case, their conjunction C may therefore mean something new. Or, in the case of the alliterative twin forms, perhaps neither A nor B by themselves mean anything (any more?), and meaning then arises only from their conjunction. Either way, the conjunction is motivated by factors other than the purely aesthetic.

#### 4 Compounds which “add flavor”

We would like to present a third series of A+B forms where there is no trace of any semantic difference between A and B, which are judged to be purely synonymous, and in which the repetition quite explicitly has none of the iconic functions (marking plurality or iterativity or emphasis) that repetition typically has not only in Khmer but in languages generally. The general formula for these pairs (which seem not to have a specific label in the Khmer grammatical tradition) is this: a verb is paired with a light-verb version of its cognate accusative construction: **have a dream+dream**, and so forth. (Curiously, although genuine cognate accusative constructions do exist, we have not yet encountered cases like **dream a dream+dream**.) Among the most frequent examples of this construction are conjuncts like:

<i>baoh c-um-hien chien</i>	“take a step step”
<i>baek k-umn-wt kwt</i>	“open thought think”
<i>miən c-awmn- eh ceh</i>	“have knowledge know”
<i>ciə awmn-aoj aoj</i>	“be a gift give”
<i>cie c-umn-uum cuun</i>	“be offering offer”
<i>awh s-awmn- aəuc saeuc</i>	“exhaust laughter laugh”

It will be noted that the nominalization is characteristically formed by the infix -Vm(n)-, described earlier. Although there are some variations on this pattern it is remarkable that the order of conjuncts is largely fixed as above. The morphologically elaborated cognate accusative conjunct precedes the monolexemic equivalent, in a stubborn and consistent violation of a presumable typological universal, the law of increasing members (Behaghel 1932, Malkiel 1959). But the most remarkable thing about these compounds is that they seem not to mean anything qua compounds, and are (in Noeung's opinion) interchangeable with the monolexemic root from which they are derived. In support of this claim we will do more than list some of the examples we have noted, and present them instead embedded in part of the context where we encountered them in written Khmer.

*tok naa          {ciə s-awmn-aen          saen}      pnoo*  
 while            be offering                  offer    grave  
 "while making an offering at the grave.."

*coh      kmae    jeeuŋ    {mien p-um-nie pie}    kam 'wej kaaw baan cie wetuniə*  
 maybe Khmer we      have bad deed do ill action some ?    cause    misfortune  
 "perhaps we Khmer have committed some evil action to cause this misfortune.."

*klaaj    {cie c-um-looh clooh}                  prawkaek knie*  
 become be    conflict    fight                  argue          each other  
 ".. came to be in conflict.." (Note the additional genuine synonym compound here.)

*kumnaau          {miən k-awm-poh kpoh}                  dawl mleh*  
 pile                  have height high                  until so much  
 "the pile was high up to this level."

*pkaaj preuk    {bawnjceenj p-aun-lww plww}                  ceunjcaej laeuŋ leeu meek*  
 star morning    emit illumination illumine                  bright          up          on sky  
 "the morning star was bright up in the sky.."  
 (Again, note the additional genuine compound)

*jeeuŋ dael          {mien c-um-nwe cwə }                  ceak neuŋ kooraup preah put*  
 we                  who have    belief believe                  clear and    salute    Lord Buddha  
 "we who believe and salute the Buddha.." (Again, a possible synonym compound)

*knjom pum dael kheeuŋ koet {awh s-awmn-aeuc saeuc}*  
 I          never          see          him          exhaust laughter laugh  
 "I never see him', (he) laughed..."

*{mien c-awmn-ah cah} cieng    knjom 2 rww 3 cnam*  
 have age                  old than    me          2 or    3 Years  
 "..was two or three years older than me.."

Variations on this pattern occur:

*dael {mien deuŋ d-awmn-eung} 'wəj plaek klah tee*  
 which have know knowledge any different other at all  
 “.. (is there) any news at all out of the ordinary....” (B occurs in the middle of A)

*{mien tae troem t-um-roem} dawl tii daw*  
 have only watchful watchfulness till destination  
 “..be watchful till we reach our destination..” (B occurs in the middle of A)

*(t-awmn-aaw taaw} teuŋ jeeuŋ neung chup laeng praeu*  
 continuation continue go we will stop quit use  
 “..continue on till we stop using ...” (No light verb with A)

*daoŋ nweŋ hawt nəŋ {d-awmn-aeu t-m-aeu} ceeuŋ*  
 through tired exhausted with traveling traveler foot  
 “through exhaustion from having walked..” (No light verb with A, B also nominalized)

In some cases (very few that we have noticed so far) the construction has become partially opaque through sound change, and so is no longer perceived as a repetition at all. We are currently unable to offer a detailed synchronic parsing of examples like the following:

*mien teevaadaa teep - {rak reaksaa} vault*  
 exist angel angel - guardian take-care temple  
 “there is a guardian angel watching over the temple..”

(Here, the orthography still reveals that [rak] “guardian” derives from and was presumably at one point identical with the following word [reaksaa].)

*{Ibej rauntww lww} soh saaj*  
 famous thunder hear bright expand  
 “(his) shining reputation grew..”

(Here, Noeurng is confident that [Ibej], which is pronounced [lbww] in some dialects, is a nominalization of [lww].)

Examples of this sort may provide a preview of the next plausible stage in the development of compounds of this sort. If they do, Khmer may offer an example of a kind of evolution which is well attested — indeed may be standard — in biology (Mayr 2002:38). The most common and harmless mutations (whether of genes or of larger structures) are replications,  $A > AA$ . By a later possible development  $AA > Aa$ . The novel (paralogous) form “a” is free to deviate not only in form, but in function from the original (orthologous) form “A” of which it was once a clone.

## 5 Discussion.

Sometimes, languages seem to include morphological material which seems to function for no other reason than to provide bulk. Consider the following paradigm from French:

en janvier  
 en février  
 en mars  
 en avril  
 ...  
 Au mois d'aout (not en aout)

This seems to be quite well attested particularly in SE Asian languages (Matisoff 1978, passim; 1982:74-76 et passim; Anderson & Zide 2002), but may not be restricted to them. It may be that Benveniste's famous "enlargements" of the PIE root, most of which are still unglossable (Benveniste 1935:chapter 9), or the final consonant of the triliteral root in Semitic postulated by Diamond 1959, are akin to the etymologically illegitimate tacked-on bits and pieces which Anderson & Zide have postulated as required to satisfy a "bimoraic root constraint" in Mon-Khmer. In presenting the data that we have here, we are conscious of simply confirming Karlgren's hypothesis for compounding in Mandarin (Karlgren 1923 [1962]) with data from an unrelated language of the same linguistic alliance. Our difference is one of functional motivation. We suggest that in Khmer at least compounding is not motivated primarily by the need to restore phonological bulk, but by a more aesthetic or playful drive for elegance—what Miller 1973 has called "galumphing". That galumphing produces extra structure and that this extra structure may serve to reduce ambiguity, may be unintended consequences of a drive whose origins have nothing to do with cognition and much to do with art.

It is very unlikely, however, that these bits and pieces whether they are added by speakers who are driven to be understood, or speakers who just want to have fun, came from nowhere. In presenting the data we have considered here, we are suggesting merely two more possible sources (exaptation of elided sounds, and decorative repetition) for such material in a language which needs all the bits and pieces it can get.

## References

- Anderson, G. & N. Zide. 2002. Issues in Proto-Munda and Proto-Austroasiatic nominal derivation: The bimoraic constraint. *SEALS 10*: 39-54.
- Behaghel, O, 1932. *Deutsche Syntax. Bd. 4*. Carl Winter.
- Benveniste, E. 1935 [1984]. *Origines de la formation des noms en indoeuropeen*. Maisonneuve.
- Bloomfield, L. 1933. *Language*. Holt, Rhinehart, Winston.
- Bolinger, D. 1975. *Aspects of Language*. 2<sup>nd</sup> edition. Harcourt, Brace, Jovanovich.
- Diamond, A.S. 1959. *The history and origin of language*, New York: Philosophical Society.
- Haiman, J. 1998. Possible origins of infixation in Khmer. *Studies in Language* 22:597-617.
- Haiman, J. & N. Ourn, 2002. Coordinate compounds and Khmer phrase structure *SEALS 10*: 147-66.

- Haiman, J. & N. Ourn. 2003. Nouns, verbs, and syntactic backsliding in Khmer. *Studies in Language* 27(3):451-74
- Huffman, F. 1970. *Modern spoken Cambodian*. Yale University Press.
- Karlgren, B. 1962. *Sound and symbol in Chinese*. London: Oxford University Press.
- Malkiel, J. 1959. Studies in irreversible binomials. *Lingua* 8:113-60.
- Matisoff, J. 1978. *Variational semantics in Tibeto-Burman*. Philadelphia: Institute for the Study of Human Issues.
- Matisoff, J. 1982. *The grammar of Lahu*. 2<sup>nd</sup> edition. University of California Press.
- Mayr, E. 2002. *What evolution is*. New York: Basic Books.
- Miller, S. 1973. Ends, means, and galumphing: Some leitmotifs of play. *American Anthropologist* 75(1):87-98.
- Ourn, N. & J. Haiman. 2000. Symmetrical compounds in Khmer. *Studies in Language* 24:483-514.

Note: Khmer specialists will be outraged by the transcription, a practical orthography which we have been developing over the last several years. They will, however, also be able to translate it back into Huffman's system without too much trouble.

For non-specialists:

- a) word-initial glottal stop is transcribed only before another consonant. (e.g. {aoj} = [ʔaoj] "give")
- b) The graph {e} after the graphs {i,o,u, e, w} is schwa (e.g. {moen} = [moən])
- c) The graph {e} before {a} is epsilon (e.g. {neak} = [nɛak] "person")
- d) The graph {a} after the graph {e} is schwa (see above)
- e) Elsewhere the sound schwa is represented by {eu} (e.g. {peut} = [pət] "true")
- f) The graph {w} is a high back unrounded vowel when it appears right after a consonant (e.g. {kwt} = [kit] "think")
- g) The graph {aw} is the default vowel in the first register (e.g. {bawt} = [bat] "form")
- h) The graph {au} is the default vowel in the second register (e.g. {raut} = [rət] "run")
- i) {eeu} is long schwa (e.g. {meeul} = [mɛ:l] "look")
- j) {aaw} and {aau} are the long default vowels (e.g. {baawng} = [bɑ:ŋ] "older sibling")
- k) In all other cases, length is represented by doubling (e.g. {baan} = {ba:n} "get")
- l) {ng} is the velar nasal [ŋ]
- m) {nj} is the palatal nasal (e.g. {knjom} = [kɲom] "1sg.")
- n) {v} is a bilabial approximant with phonetic values [w] or [v] (e.g. {vie} = [wiə] "third person non-respectful").



