Coordinate Compounds and Khmer Phrase Structure

John Haiman (Macalester College)
Noeurng Ourn (Regions Hospital, St. Paul)

In Khmer as in other isolating languages of SE Asia, lexical elaboration via both compounding and serialization is productive.. Our focus in the following discussion will be on symmetrical coordination and the place of symmetrical compounds in Khmer grammar.

There is a strong tendency to use coordinate near-synonym compounds rather than single words in Khmer, as in other SE Asian languages. These compounds may be verb-verb, as in:

cŋaɲ pisaa "tasty tasty"
doh khat "clean polish"
cruet criep "absorb absorb"
psaa pcoep "connect connect"
trəm "align" + trəw "correct"
priən "instruct" (<riən "learn") + prɑdaw "advise, edify"
khoh "wrong" + plaek "novel" = "different"
khookhəw "violent" + troskbɑt "treasonous"(1) = "violent"
kcəl "lazy" + crɑ'ooh "sluggish" (< 'ooh "drag") = "slothful"
lŋɔŋ "ignorant" + klaw "ineducable" = "dumb"
pluɯ "light" + tlaa "bright, clear" = "bright and clear"
crəəh rəəh "choose"+ "choose"
ciəh "detour around" + wiəŋ "move aside" = "avoid"
sɑɑm "appropriate" + srɑɑp "adapt" = "appropriate"
saak "test, try" + lbɑɑŋ "survey, test" (<*lɑɑŋ)(2) = "try, test"
wiəc "dishonest, crooked" + wee "deviant" = "dishonest"
loən luɯ "resound, sound"
soh "all over" + saaj "spread out, diffuse" (>psaaj "propagate")
= "far and wide"
swət rɔɲɔɔm "wither wilt"
kriəm "dry, desiccated"+ krɑm "clot, bruised" = "sad"
krɔp 'cover" + dɑndɑp "cover to keep warm" = "cover (e.g. as clouds cover sky)"

puən "hide" + sɑmnɑm "silence" (transitive verb <sŋɑm "silent") = "hide"

liət "extend" + sɑnthəŋ "stretch" = "extend"

pruəj baarɔm "anxious, worried"

klaŋ "strong" + klaa "courage" (> klahaan "brave") = "proud and strong"

khooc "damage" + khaat "lose the use of" = "damage"

phɑɑl "produce, utility" + prɑjooc "profit, use" = "profitability"

təŋ cɑŋ'iət "tight constrained"

bak "break (e.g. handle off something)" + baek "break (body of something)" = "break"

'ɑn "decline, be sickly, worsen" + thɑɑj "back up, reverse, diminish" = "decline, fade"

biət biən jiəjii "bother, annoy"

plah "change" + pdoo "change"

sŋat sŋiəm "silent mute".

prah "get away" + caak "leave" = "escape"

crah srɑlah "clean, clear"

sokdom rɔmniə "happiness"

They may (less frequently) be Noun- Noun:

bɑmnɑɑŋ pratnaa "goal" + "intention"

sok kseem ksaan "peace" + "peace"

baep jaaŋ "kind" + " sort"

krət krɑɑm "law" + "law"

lɛətaa wɔl "vine" + " vine"

'ɑmnɑɑ 'ɑkun "thanks" + "thanks"

kmaŋ sɑtrəw "enemy" + "enemy".

Some V-V compounds remain verbs (and some even function as conjunctions or prepositions). But often the compound V-V becomes nominalized ipso facto(3):

tok phəh "worry afraid" > "suffering"

tok saok "worry, weep" > "suffering"

paɲɲaa wiəŋ wiəj "intelligent wise" > "intelligence"

saarɛəh sɑmkhan "important important" > "importance"

tɛək "connect" + tɔɔŋ "connect" > "connection"
bət "hide" + puən "hide" > "(the game of) "hide and seek"

Though the number of such compounds seems to be indeterminate, not all combinations are equally frequent, or even possible. For example, the words mhoop (<hoop "eat") "any food that goes with rice", 'ahaa, "food", sbiəŋ "provisions, rations, stock, supplies", cɑmnəj "food" (<cəj "eat") all mean "food" and can be combined but only in the lexicalized formulas

mhoop 'ahaa
mhoop cɑmnəj
sbiəŋ 'ahaa
cɑmnəj 'ahaa

Similarly, priəp, thiəp, and prɑdooc all mean "compare" and can be combined only into the compounds

priəp thiəp
priəp prɑdooc

Not all coordinate compounds are binomial. Sometimes three or more synonyms occur together:

prɑhaən "disrespectful" + kaoŋ "crooked" + kaac "coarse, nasty, savage" = "excessive"

Sometimes, the coordination of two words will itself function as an element in a superordinate compound, so that in principle, the rule N --> NN is recursive:
cumŋɯɯ "illness" +[chɯɯ "ill" + tkat "???"(4)] = "illness"

The last example (which is not unique)(5) consists of a conjunction of a word with some distinct version of itself, and is therefore a particularly useful demonstration of the semantic redundancy of symmetrical compounds. Motivations for this kind of compounding are the subject of this paper.

Some reasons for the proliferation of these compounds are fairly obvious. The two conjoined synonyms may not in fact always be entirely synonymous, and so their combination is thus slightly richer or different in meaning from either one on its own. Examples are:

cŋoj "fragrant" + cŋaɲ "delicious"
trɑbak "snap at" + kham "bite"
bak "break (e.g. handle off)" + baek "break (e.g. into two or more pieces)"

A frequent correlate of compounding is that while single words have a concrete meaning, their combination or reduplication can have a figurative or metaphorical one:

muk "face", but muk + moət "mouth" = "face, reputation"
kpoh "high", but kpoŋ "???" + kpoh = "exalted".

In some cases, the combination may be quite idiomatic:

noəm "lead, guide" + 'aoj "give" = "cause"
cɑɑŋ "tie, bind" + cam "wait for, guard" = "remember".

There remain, however, many cases where coordination seems to provide little semantic information. Among the reasons for this compounding, we recognize two which are commonly cited: elegance and intensification. In addition, we would like to suggest two kinds of "trade-off": syntactically, compounding may provide a compensation for the absence of cues to syntactic category or constituency, and lexically, it may compensate for the absence of phonological bulk.

A) Elegance

The coordinate compound structure occurs in ordinary speech, but is associated with the flowery diction of a higher register or verbal art, thus occurring particularly often in proverbs:

[koəh rumləəŋ] pnom 'aoj rɔliəj sruəl ciəŋ kaa [plah
[dig up loosen] mountain so fall easy exceed that [change
pdoo][ca'rət ni'saj] rɑbɑh mnuh
change][attitude talent] of person
"It is easier to move a mountain than to change one's habits"

kaa 'ɑtooh   'aoj kɯɯ ciə kaa [sɑɑŋ sək]   dael   kuə
that not fault give :   be that repay revenge which must happy
[treek 'ɑɑ]   ciə tii bɑm phot
happy   be the uttermost
"Pardon is the most glorious revenge."

Notice the particular exuberance of such compounds in a
formulaic New Year's greeting:

soom 'aoj baan seeckdəj sok   [cɑmroŋ cɑmraən]
please give get matter   peace   ???   plenty
[prɑkɑɑp daoj] sokphiəp
 with   through peace
bɑriboo [prah   caak]   [cumŋɯɯ [chɯɯ tkat]]
enough   escape escape   illness   ill   ???
[crah srɑlah] 'ɑmpii
clear clean   from
[tok saok] [muəmaw](6) [kdaw krɑhaaj](7)
sad weep Frustrated Hot Burn
haəj soom [baan tɔtuəl]
 and please get receive
nəw phiəp [sŋɑp rumŋoəp] sawmɑɑŋ soom 'aoj [prɑkɑɑp nɯŋ]
with aspect calm pacify   distress   please give with   with
phiəp [sokdom rɔmiə]   sɑntəphiəp ni'rɔn tɑɑ   təw!
aspect happiness happiness peace   forever continue go

"I wish you peace and plenty, health and deliverance from
sickness, freedom from suffering , frustration, and irritation, and
may you receive surcease from distress; I wish you this with
happiness and peace forever."

Elegant style (sɑmnuən wɔhaa) or spice (krɯəŋ krəw) is in fact

one explicit motivation for compounding, and we specifically invoke this as the sole or principal motivation in examples like:

təŋ caŋ'iət "tight"
swət rɔŋoom "wilt"
ciəh wiəŋ "avoid"

as well as in every single coordinate compound in the New Year's greeting above. Closely related to the notion of elegance in this structure is the aesthetic notion of formal symmetry. The formal symmetry of the elements in a compound may be more important than their semantics. As evidence for this, we note that in addition to meaningful synonym compounds, Khmer has a huge number of symmetrical reduplications or twin forms of the razzle-dazzle type (more accurately, of the alliterating spic 'n' span type, cf. Ourn & Haiman ms. and the references there, particularly Ratliff 1992 for comparable structures in Hmong, and Vonvipanond 1992 for Thai). Here one or both of the elements in the compound may be entirely meaningless 'servant words' (bo'ri'waɑ sap), glossed "???" as in

kpoŋ "???" + kpoh "tall, high" = "exalted"
lbəj "famous" + lbaɲ "???" = "famous"
kruɯəŋ "spice" + krəw "???" = "spicy"
təəh "block" + tɛɛŋ "???" = "block someone's view"
prae "turn" + pruəl "???" = "turn, change"
ptuɲ "???" + ptoəl "against" = "bear witness against"
tiəŋ "accurate" + toət "??kick" = "accurate"
pruɯm "???" + pruɯj "lovable" = "lovable"
khuɯt "??approach" + khɑm "strive" = "strive"
cɑmroŋ "???" + cɑmraən "lots" = "plenty".

It is hard to say these servant words are meaningless: but it is also hard to say what they mean. It is notable also that although there are hundreds of these compounds in Khmer, there is no productive method of generating the meaningless element of each pair: they are words and verbal artistry consists exclusively in remembering them, not in making up new ones.

B) Intensification (twəə 'aoj klaŋ "make it strong")

      Where the synonymy of the coordinated words is nearly perfect, the use of two is sometimes more emphatic than using only one. Some examples of purely iconic strengthening via (quasi) repetition include:

psaa "join" + pcoəh "join" = "weld, join"
wiəc "dishonest" + wee "deviant" = "really dishonest"
saak "try" + lbɑɑŋ "try" = "really try".

In fact, it is sometimes unclear if the structure is a symmetrical compound or a head-modifier construction whose second word functions as an intensifier (wisɛh kun niəm) of the first as in "pitch black":

puən "hide" + sɑmŋɑm "silence" = "really hide"
nɯɯŋ "motionless" + tkɑl "still" = "very still"
kdaw "hot" + krɑhaaj "burn" = "burning hot"
srɑh "fresh" + bɑmprɔɔŋ "fresh" = "really fresh"

But sometimes, compounding seems to serve neither of these commonly acknowledged functions. It is as if the language, like the melody of "Twinkle Twinkle Little Star" loved to say everything twice, and this is the phenomenon we are trying to account for by suggesting that compounding may also have some compensatory functions.

C) Substitute for Phrase Marking?

Syntactically, compounding creates structures [XX] instead of [X]. We suggest, tentatively, that this bracketing provides some sorely needed structure in an otherwise strikingly unstructured string of words. Crucially, we make reference to two other perhaps functionally related facts. The first is that there are few cues to phrase structure constituency or even to category labelling in general, and those function words that exist (node building words like the relative pronoun dael) are highly optional. The

second is that simple repetition of a word (saying it literally twice) is surprisingly not an available option to signal intensification: repetition unambiguously and obligatorily marks plurality or repetition.

There are two overarching principles of hierarchical Khmer PS:

    a) modifiers follow heads (e.g. N +A in NP)
    b) operators precede operands (e.g. P + NP, VO)

but the situation is confused by a number of factors.

Within NP, the second constituent (whether N or A) should modify the first, but de facto A+N word order exists in cases like koon "offspring" + Noun, where koon has come to mean and act like the adjective "small". Phrases like koon cruuk not only mean "child (of the) pig (type)" but also "small pig".

Within NP, numerals must precede classifiers. Since the latter are themselves Nouns, there is de facto modifier + head order in phrases like pii tŋaj "two days", which coexists with expressions like cruuk pii "two pigs". Minimal contrast pairs like maoŋ pii "two o'clock" and pii maoŋ "two hours" abound. Under other circumstances which we are unable to characterize as yet, Numeral may precede simple non-classifying Noun:

nəw muəj kɑnlaeŋ
in   One place
"in place, without moving"

Kɲom cɑŋ təw tae muəj kɑnlaeŋ ponnooh
I   Want go only one place   That-many
"There's only one place I want to go."

Cf. Jəəŋ nəw ptɛəh tae muəj
    We live   house only one
"we live in the same house"

Node building words are optional. The syntax of one of these, the genitive marker rɔbɔh "(thing)> of" has been described in the literature: the particle acts as a compensatory diacritic (Haiman 1985:60-7) to signal constituency that is not directly signalled by the order and adjacency of the words. That is, it is optional when the possessor is adjacent to the possessee and is obligatory otherwise (Huffman 1970:73):

koon (rɔbɔh) kluən
child of   Self
"my/her own child"
?koon tooc ___ kluən
"her small child"
*koon tooc pii ____ kluən
 child small two   Self
"her two small children"
*cɑŋkoom dɑ mut ___ klaa
 fang   Very sharp tiger
"the extremely sharp fangs of the tiger"
*tpoəl sɑɑŋ khaaŋ ____ koon
cheek both side   Child
"both of the child's cheeks".

Although the behavior of other node-building words may be similar, we have not been able to demonstrate the same kind of pattern with all of them.

1. The modifier-maker jaaŋ

The noun jaaŋ "kind, sort" is not only a synonym of baep , with which it appears in synonym compounds. It also has at least an incipient purely syntactic function which baep does not have, in that it marks the following word or phrase as attributive rather than predicative. In principle, cruuk l'ɑɑ means "the pig (is) good" and can function either as a sentence or as a noun phrase "the good pig". Conversely, cruuk jaaŋ l'ɑɑ can only be a NP "the good pig". Again in principle, jaaŋ may cooccur only with stative verbs. Sometimes in this function it seems to signal

intensity and is synonymous with the intensive particle dɑ "very":

cɑŋkoom dɑ/jaaŋ mut
fang   very/kind sharp
"extremely sharp fangs"
tiərɔɔk dɑ/jaaŋ prɯm prɯj
infant   very/kind lovely
"extremely lovely infant"

It would seem jaaŋ would be extremely useful in distinguishing attribution from predication as well as for making a distinction between adjectives and active verbs. But generally, the distribution of jaaŋ as we have observed it in texts does not correspond to such "syntactic" or "derivational" functions:

koon ['ɑh   sɑmnaəc saəc] jaaŋ kɑk'ak
child exhaust laughter laugh kind uproarious
"the child laughs and laughs, uproariously"

Wiəl tumniəp   trəw   [ləc lɔŋ]   jaaŋ dɑmnɑm
field lowlands   undergo flood   kind   damage
"fields sustain severe damage from floods"

prɛəh ti'nakɔɔ bɑɲceeɲ soo saeŋ jaaŋ [klaŋ klaa]
lord sun   emit   rays kind strong clear
"the sun shines forth bright and strong"

'Aoj chɛh prəj jaaŋ [sɑnthoo sɑnthəw] laəŋ thaem tiət
to   light forest kind ablaze   up   add   more
"to set the forest ablaze"

prɛəh 'atət bɑɲceeɲ rɔsməj jaaŋ klaŋ
lord sun emit   rays Kind strong
"the sun shines bright"

cɑracɑɑ prɑprət təw jaaŋ lumbaak
traffic act   go kind difficult
"traffic moves with difficulty"

kee   baan [khooc khaat] [phɑɑl prɑjooc]   jaaŋ craən
they   get damage lose   profit usefulness   kind lots
"they undergo severe damage to their profits"

miən krooh tnak jaaŋ tɲɔn
exist   accident   kind heavy
"there are serious accidents"

kɲom baan dəŋ   nuɯŋ jɔl   nəw [mnooh sɑɲceetnaa]
I   can   know and   understand acc. Deep emotion
nəj kaa [tɛək   tɔɔng] nuh jaaŋ cbah nah
of nom. connect connect that kind clear very
"I can hear very clearly the deep emotion of that bond."

Leeŋ lbaeŋ kuuk wah ciəmuəj koon jaaŋ riikriəj taam tuursap
play   Game peekaboo with child kind joyful follow phone
"joyfully play peekaboo with her child over the phone"

mdaaj bɑnluɯɯ thaa "kuuk" Jaaŋ wɛɛŋ
mother pronounce say kuuk   kind long
"The mother says "kuuk" in a long drawn-out fashion."

2. The relative pronoun dael "ever"

The relative pronoun dael marks the following phrase as a (subordinate) relative clause rather than a predication on the same level as the preceding NP. In principle, it can be used to make distinctions like the following:

cruuk ŋoəp
pig   die
"The pig died."
cruuk dael ŋoəp
pig   which die
"the pig that died"

Again, this does not seem to correspond with its distribution in texts. Nor does it seem to function as a compensatory diacritic.

Note that in most of the following structures the RC occurs adjacent to its head:

mnuh dael miən[cɑmnɛh dəŋ]   tok dooc ciə miən pɔnluɯ
person that have knowledge know ? like be exist illumination
"a person that has knowledge is like one enlightened"

siəwphəw dael ciə wɔnna' kam rɑbɑh nɛək nipɔɔn
book   that be masterwork of   person writer
"a book that is the masterpiece of a writer"

nipɔɔn dael miən kee   cmooh [lbəj lbəj]
writer that have they name   famous famous
"writers that are famous"

prɑpheet siəwphəw dael miən kləmsaa l'ɑɑ
kind   book   that have essence good

skoəl ruup mnuh   dael lǝǝk jɔɔk mɔɔk ni'jiəj
know form person that raise bring come speak
"come to know the person in question"

nɛək dael miən [muk niətii] trəw cam tii
person that have task duty   must guard place
"the people who have the task of guarding the place"

In the sentences below, the dael cannot be omitted. Note that the RC is sometimes but not always separated from its (bolded) head by intervening material:

tɔptɔl nɯŋ **cɑŋkoom** mut rɑbɑh klaa [dael trɑbak kham koon]
contend with fang   sharp of tiger that snap Bite child
"contend against the sharp fangs of the tiger that snap and bite her offspring"

priən prɑmaɲ [dael baan dot cool prəj rɔɔk cap sat]
hunter hunt   That get set afire forest seek catch animal
"the hunter that sets a forest fire while trying to trap animals"

kaət 'ɑmpəə tocca'rət kumnɯt [khookhəw troskbɑt]
arise bad character   thought violent aggressive
dael 'aac   noəm mɔɔk nəw krooh tnak
that can   lead come at   accident
pseeŋ pseeŋ daoj   mɯn dəŋ   kluən
different   through not   know self
"a bad nature and violent thoughts arise that can lead to mishaps
of various sorts through the lack of self-knowledge".

Lbaeŋ nih ciə **lbaeŋ cɑmbɑɑŋ muəj** dael jəəŋ kuə cɑɑŋcam
game this be game principal one   that we should remember
"This is a principal game that is worth remembering."

Koon kcəj proh srəj tooc crɑləŋ srɑh ɲɔɲəm
child new male female small cute fresh smile
dael mdaaj pɲuət tək
that mother wash water
"The small cute infant boy or girl that the mother washes"

In each of the following examples, the brackets indicate the RC.
Note that the relative pronoun is not present in the original, but
can be supplied:

kmiən nɔɔnaa mnɛɛk [___ 'aac cuəj sɑŋkrooh]
not exist anyone person   can help save life
"there is nobody who can help save their lives"

miən tuuk muəj muəj [___ caev cool mɔɔk taam ptɛəh]
exist boat one   paddle enter come follow house
"an occasional boat paddles right up to and can enter the houses"

mnuh [___ cɛh dəŋ 'aac cuəj miəto' phuum]
person   know know can help motherland village
"a person who knows how to help their native country"

pɔnlɯɯ [___ nɯt nəw coəp nɯŋ kluən]
illumination   stay live adhere with self
"the illumination that stays by one's side"

mnuh kwak [___daə wɯlwɔl stiəp rɔɔk pləw nuh laəj]
person blind   walk in error grope seek road that at all
"a blind person that walks in error and gropes blindly for the path"

phiəsaa muəj [___ pseeŋ pii phiəsaa kmae]
language one   different from language Khmer
"a language other than Khmer"

We do not know the stylistic or syntactic principles that govern the distribution of this word. Other function words include ciə "be" and nəw "direct object" whose elementary distribution is even more puzzling to us. What they share with the other words we have described above is that they are disconcertingly optional so that they provide limited guidance in parsing.

The Parts of Speech are not clearly defined. Like other Sinospheric languages, Khmer lacks a category distinction between Verbs and Adjectives; prepositions and adverbs are identical with either nouns or transitive verbs; most conjunctions are verbs; the complementizer for STKP verbs is the verb thaa "say". So, the only solid part of speech distinction seems to be the distinction between nouns and verbs. But this distinction is also a relatively fragile one.

Roots which seem to be predominantly Verbs can surface as nouns (e.g. coən 'step on" can be "a step", in phrases like cap pdaəm taŋ pii ___ tiəp təw mun "begin from the lowest step "); less often (?) Roots which are nouns can surface as verbs (e.g. cmooh "name" as in cmooh kɲom "my name" can be "be named" in phrases like kɲom cmooh "I am called, my name is ..."). There is derivational morphology: the prefixes kaa- and kdəj "matter" can be added to any verbal root to create a derived noun. Almost as productive and with the same function, is the infix -Vm(n)-. Even derivational morphology can be ignored or seems to fail to give unambiguous signals about the characterization of the word. In particular, it seems that the derivational function of the nominalizing infix is not fixed. For example: in wiəl

t(umn)iəp "low fields", the qualifying word is either the adjective/verb tiəp "low" or a derived noun "lowness"; in klaa k(ɑmn)aac "fierce tiger", the qualifiying word may be either the adjective/verb kaac "fierce" or its derived nominal. Conversely, derivational morphology may be optional. When they occur, the nominalizing affixes kaa- seeckdəj- and phiəp- invariably create derived nouns. But they do not seem to be required for verbs to behave as nouns. For example, in phrases like daəmbəj twəə pɔɔrniə "in order to make describe", the final verb acts as a nominal object "description" of the transitive verb twəə "make". *The nominalizing prefix kaa- "ness" is available (as in kaa-*pɔɔrniə trəw chɔɔ ləə lɛəkkhənɛə' sɑmkhan "the description must stand on the important characteristics"). It is equally available (and is sometimes used) in V-V compounds which are supposedly already derived nominalizations.

Our Hypothesis: in addition to signaling intensification and providing some aesthetic spice, doubling may serve a possible cognitive function, compensating for that of "node-building" function words. It creates brackets and groups words into larger constituents. In slowing down the rate of speech it may provide the listener with more time to make sense of the stream of speech. If [XY] can be either an endocentric or an exocentric construction, the expansions [XXY] or [XYY] or [XXYY] serve to provide a breather for the listener. (As a mental exercise, try humming or whistling "Twinkle Twinkle Little Star" without the note pairs.)

Notably, however, the coordinate compounds investigated here are not simple repetitions of the same word, but pairs of synonyms. The language is also rich in symmetrical reduplicative compounds like spic 'n' span which may serve the same function. Why no simple repetition?

Total repetition is frequent in Khmer: in fact there is an orthographic shorthand device for indicating it. But it can't be used either for intensification or for elegance.

D) Simple repetition is not available.

Simple repetition of a given word is not available as a device since it has a different semantic function. Khmer does not allow structures like

a mad mad mad mad world
the green green grass of home

to signal intensification, or to be used without any meaning whatever. Instead, repetition of the same morpheme is possible but instead of bearing the iconic function of intensification repetition always carries other equally iconic functions: a plural, repetitive, continuative, or distributive meaning:

mien srɑlaɲ dooc dooc knie
exist love   Like like   each other
All love (their children) the same."

Cf. Mdaaj dooc knie nɯŋ 'ɔwpuk peeɲ cət
    Mother like each other with father full heart
"The mother together with the father is proud and happy."

Mien ceek 'ɑmpəw   nɯŋ daəm chəə pseeŋ pseeŋ
exist banana sugar cane with trunk tree different different
"There are banana trees, sugar cane, and various other trees"

Cf. Phiesaa muəj pseeŋ pii phiəsaa kmae
    Language one different from language Khmer
"a language other than Khmer".

Other examples of the distributive sense of repetition include:

prɑləm sraaŋ sraaŋ
dawn   Get light
it gets lighter bit by bit

wiə kəəm kəəm
crawl creep
crawl inch by inch

khɑm bɑntɑɑ dɑmnaə muəj muəj
strive continue journey one one
struggle onward one (sc.step) at a time

səh pukee pukee
student good
good students

daəm chəə kpoh kpoh
trunk wood tall
tall trees

cmooh lbəj "a famous name"
cmooh lbəj lbəj "famous names"

Our Second Hypothesis: compounding may be used in Khmer, as it may be used in Mandarin and other languages of the region, as a relatively unsystematic means of building up phonological bulk. It may thus be functionally related to infixation (cf. Haiman 1998) in Mon-Khmer, and to the variety of devices in Mundaric languages (Zide & Anderson, this volume) which have in common only the result that they "add one mora" to a (typically monosyllabic) root.

Some Comparative Remarks:

Clearly Khmer is similar to Vietnamese, Thai, Mandarin, Lahu, Hmong, and other SE Asian languages in both its syntactic underspecification and in the exuberance of its lexical elaboration (Riddle 1990, 1992, Stahlke & Riddle 1992). The isolating nature of SE Asian languages may be responsible for creating paratactic target structures in general (Riddle 1990) and these do in fact occur in non-contiguous isolating languages spoken in Africa. But the most common explanation for semantically redundant

symmetrical compounding is aesthetic (Haas 1964, Matisoff 1973, Johns & Strecker 1987, Riddle 1992). While greater length is almost universally an icon of politeness and refinement, the particular penchant for making things longer by saying them twice seems to be common to SE Asian languages and Khmer is typical of these languages in this respect.

As in Mandarin, symmetrical compounding may create figurative as opposed to concrete meanings (Li 1990, Liu 1992, apud Duanmu 1999). As in Mandarin and other languages, compounding may serve a derivational function, making nouns out of verbs (ibid.).

Duanmu is skeptical of these as motivations for symmetrical compounding. He also rejects the well-known Karlgren hypothesis that compounding in Mandarin was motivated by the drive to reduce intolerable homonymy as that language lost tonal and consonantal distinctions that were better preserved in other Chinese languages (creating compounds by the funny-haha vs. Funny-queer mechanism). He proposes instead that compounding in Mandarin is obligatory where stress andrhythmic requirements call for disyllables, and is blocked where those requirements call for monosyllables. Unlike Mandarin, Khmer does not seem to require compounds or disyllabic forms for rhythmic reasons. It is never absolutely ungrammatical to have repetition, nor to avoid it, as Duanmu convincingly shows it sometimes is in Mandarin. Therefore, Duanmu's explanation for the distribution and prevalence of disyllables and compounds in Mandarin cannot apply to Khmer.

Finally: Unlike Hmong, at least (Riddle 1990), Khmer cannot reduplicate for intensification. While one can say in Hmong (and in English)

koj txawj txawj ua paj ntaub
you know know do embroidery
"You really know how to do embroidery"

this metalinguistic use of simple repetition is spoken for or otherwise allocated in Khmer, and that fact may contribute to an even greater exuberance of symmetrical and reduplicative compounds in that language than elsewhere in SE Asia.

To say this is not to deny that the primary function of synonym compounds may be aesthetic — or to deny that, once established, the pattern may have simply taken on a life of its own.

Endnotes

1.Itself a compound of trusa "opposite side" + kbɑt "ally oneself".

2.Source of prɑlɑɑŋ "test, trial, examine"

3.This is particularly noticeable when opposites are conjoined: khoh "wrong" + trəw "correct" = "morality", and so forth.

4.Tkat is presently meaningless, an example of a non-symmetrical "servant word", cf. Noeurng & Haiman, ms.

5.Other examples include baoh cumhiən chiən "take step + step", and 'ɑh sɑmnaəc saəc "exhaust laughter+ laugh", where the verb is coordinated with its own cognate accusative construction. Related is the compounding of doublets: peel "time (<Skt)" + weeliə "time (<Pali)".

6.Neither part of this word is now meaningful, but the formal symmetry of the syllables strongly suggests compounding as in English words like **higgledy piggledy**, cf. Ourn & Haiman ms.

7.It is unclear whether the second element is a coordinate member of this compound or an intensifier.

References

Duanmu, S. 1999. Stress and the development of disyllabic words in Chinese. Diachronica 16: 1-35

Haas, M. 1964. Thai-English student's dictionary. Stanford University Press.

Haiman, J. 1985. Natural syntax. CUP.

Haiman, J. 1998. Possible origins of infixation in Khmer. Studies in Language 22:597-617.

Huffman, F. 1970. Modern spoken Cambodian. Yale University Press.

Johns, B. & D. Strecker, 1987. Lexical and phonological sources of Hmong elaborate expressions. Linguistics of the TB area, 10:2.

Matisoff, J. 1973. The grammar of Lahu. University of California Press.

Ourn, N. & J. Haiman. Ms. Reduplicating compounds in Khmer and other languages. Studies in Language, to appear.

Ratliff, M. 1992. Meaningful tone in White Hmong. Northern Illinois University Press.

Riddle, E. 1990. Parataxis as a target structure in Hmong. (23rd Intl. Conf. On ST Linguistics). University of Texas at Arlington.

Riddle, E. 1992. Lexical elaboration in White Hmong (25th Int;. Conf. On ST Linguistics). University of California at Berkeley.

Riddle, E. & H. Stahlke, 1992. Linguistic typology and Sinospheric languages. Papers from the 1st Annual Meeting of the SEA Linguistics Society