

Mon Khmer Word Order from a Crosslinguistic Perspective

Matthew S. Dryer
SUNY Buffalo

Mon-Khmer languages overwhelmingly exhibit word order properties that can be described as head-initial. If we ask why we find this strong tendency for consistent head-initial order among Mon-Khmer languages, a popular answer is that languages tend to be consistently head-initial or consistently head-final, and Mon-Khmer languages are simply instances of the former type. However, as I have argued elsewhere (Dryer 1988, 1991, 1992), the assumption that languages tend to be consistently head-initial or head-final is not true, that the order of a number of kinds of modifiers with respect to their heads does not exhibit any crosslinguistic correlation with the order of verb and object. I will discuss alternative explanations for the pattern we find in Mon-Khmer languages.¹

The table in (1) summarizes properties that can be said to characterize languages which are consistently head-final or consistently head-initial.

(1)

Head-final	Head-initial
verb occurs last, following subject, object, adpositional phrases, adjuncts, adverbs	verb precedes object, adpositional phrases, adjuncts, adverbs, though not necessarily the subject
postpositions	prepositions
noun occurs last in noun phrase	noun occurs first in noun phrase
modifiers in general precede the modified element	modifiers in general follow the modified element

The examples in (2) (from Jacob 1968) illustrate how Khmer conforms to the general characteristics of head-initial languages.

(2) Khmer

a. VO

ʔo:pùk thuy: t̩:
father make cupboard

- ‘Father is making a cupboard’ (p. 262)
- b. Verb - Nominal Adjunct
 pʔo:n tɣu sa:la:-riən (p. 262)
 younger.sibling go school
 ‘(My) younger brother/sister is going to school’
- c. Verb - Verbal Adjunct
 vi:ə tɣu rəhás
 3SG go quick
 ‘He goes quickly’ (p. 79)
- d. Noun-Adjective
 phtə̀əh tɔ:c
 house small
 ‘a small house’ (p. 60)
- e. Noun-Genitive (i.e. noun possessor)
 phtə̀əh ta:
 house old.man
 ‘the old man’s house’ (p. 263)
- f. Noun-Possessor (i.e. pronoun possessor)
 ʔo:pùk khɲom
 father 1SG
 ‘my father’ (p. 60)
- g. Noun-Numeral
 phtə̀əh pi:r
 house two
 ‘two houses’ (p. 62)
- h. Noun-Demonstrative
 bɔntùp nih
 room this
 ‘this room / these rooms’ (p. 64)
- i. Noun-Interrogative
 phtə̀əh na:
 house which
 ‘which house’ (p. 266)
- j. Adjective-Intensifier
 thmɣu nas
 new very
 ‘very new’

As shown by Dryer (1992), a number of the word order characteristics illustrated for Khmer in (2) are indeed typical of VO languages. These include the characteristics listed in (3).

(3) Typical VO characteristics (Greenberg 1963, Dryer 1992)

Noun-Genitive
Noun-Relative Clause
Prepositions
Verb-PP
Verb-Adverb
Adjective-Marker-Standard
etc.

However, some of the characteristics of Khmer illustrated in (2) are not typical of VO languages, despite their involving head-initial order. These include those listed in (4).

(4) Typical Mon-Khmer but NOT typical VO characteristics

Noun-Adjective
Noun-Demonstrative
Adjective-Intensifier

Below, I will illustrate each of these characteristics from other Mon-Khmer languages, and present data from a database containing data for over 750 languages showing that these are not typical VO characteristics.

The examples in (5) illustrate the noun-adjective order from a number of Mon-Khmer languages.

(5) a. Vietnamese

nhà nhỏ
house small
'a small house' (Thompson 1965: 221)

b. Palaung

rapyā kərīar
girl beautiful
'the beautiful girl' (Milne 1921: 38)

c. Khmu?

kòn lə̌?
man good

'a good man' (Premsrirat 1987: 30)

d. Katu

ao tam
shirt black
'a black shirt' (Costello 1969: 30)

e. Stiêng

nhi gêq
house small
'small house' (Miller 1976: 24)

f. Chrau

iêr mǎq
chicken large
'a large chicken' (Thomas 1971: 139)

g. Sre

caw-ʔùr hàn
woman beautiful
'the beautiful woman' (Manley 1972: 156)

h. Sedang

məngê 'mei
person evil
'evil person' (Smith 1979: 77)

i. Mon

hœʔ hnòk
house large
'a large house' (Bauer 1982: 333)

The data in (6) below illustrates the distribution of AN versus NA order among VO languages from my database. The data is organized in the following way (cf. Dryer 1989a, 1992). First, I have classified the languages into genetic groups, which I call genera, which are roughly comparable to the subfamilies of Indo-European, and the numbers in (6) represent the number of genera containing languages of each type. In addition, I group these genera into six large geographic areas and count the number of genera within each area.² For example, the '5' in the upper lefthand corner of (6) indicates that there are 5 genera in

Africa containing languages in my database which are VO and AN, while the '32' below it indicates that there are 32 genera in Africa containing languages in my database which are VO and NA. The larger figure for each area is enclosed in a box.

(6)	Afr	Eur	SEA	A-NG	NAm	SAm	Total
VO&AN	5	6	4	5	17	3	40
VO&NA	32	4	12	3	7	5	63

NA order is often thought to be typical of VO languages, but the data in (6) shows that both orders are common, and that AN order is more common than NA order in 3 of the 6 areas. The order VO&NA is somewhat more common overall (by 63 genera to 40), but this difference is completely attributable to the large number of VO&NA languages in Africa: 32, or more than half of the 63 VO&NA genera, are from Africa. Outside of Africa, VO&AN is slightly more common, by 35 genera to 31.

Now one might argue that NA order is still somewhat more common than AN order among VO languages and that this suggests that there is a weak correlation between VO order and NA order. However, it turns out that this weak preference for NA order is also found among OV languages, as demonstrated by the data in (7).

(7)	Afr	Eur	SEA	A-NG	NAm	SAm	Total
OV&AN	7	28	2	5	9	7	58
OV&NA	24	4	5	17	15	19	86

The data in (7) show that in 5 areas out of 6, OV&NA order is more common than OV&AN, that the latter order is more common only in Eurasia.

In short, while NA order may be slightly more common among VO languages, it is also somewhat more common among OV languages. Hence, the occurrence of NA order among Mon-Khmer languages cannot be attributed to a general tendency for VO languages to be NA.

The situation with the order of noun and demonstrative is somewhat similar. First, the examples in (8) illustrate how the normal order among Mon-Khmer languages is for the demonstrative to follow the noun.

(8) Noun-Demonstrative

a. Vietnamese

nhà này
building this
'this building' (Thompson 1965: 191)

b. Palaung

jū̄ār inān
basket this
'this basket' (Milne 1921: 47)

c. Khmu?

kó:n nê? k̄i:
child little this
'this little child' (Premsrirat 1987: 31)

d. Katu

tariiq gamak achek
buffalo big that
'that big buffalo' (Costello 1969: 34)

e. Stiêng

ctôc nêy
place that
'that place' (Miller 1976: 43)

f. Chrau

iêr heq
chicken this
'this chicken' (Thomas 1971: 139)

g. Sre

?aso nɛ
dog that
'that dog' (Manley 1972: 156)

h. Sedang

hài tá
day that
'that day' (Smith 1979: 78)

i. Mon

kwan te?
village that
'that village' (Bauer 1982: 321)

One Mon-Khmer language which is an exception to this pattern is Khasi, as illustrated in (9).

(9) Khasi : Dem-"Article"-Noun

uta u briew
MASC-that MASC person
'that man' (Nagaraja 1985: 13)

I assume that this is related to the geographical location of Khasi, in northeast India, outside the area in which NDem order is common and in an area in which most other languages are DemN.

The data in (10) from my database shows that both orders of demonstrative and noun are common among VO languages. Although the overall pattern is one whereby both orders are about equally common (54 versus 56), again the large number of VO&NDem languages in Africa is distorting and DemN order is actually more common outside of Africa (by 50 genera to 22).

(10)	Afr	Eur	SEA	A-NG	NAm	SAm	Total
VO&DemN	4	8	7	6	20	9	54
VO&NDem	34	1	12	0	7	2	56

The examples in (11) illustrate how what I call intensifiers, words modifying adjectives and indicating degree, typically follow the adjective in Mon-Khmer languages.

(11) Adjective-Intensifier

a. Khmer

thmɤu nas
new very
'very new' (Jacobs 1968)

b. Sre

mhar naŋ
fast very
'very fast' (Manley 1972: 219)

c. Sedang

lém rəlei
beautiful exceptionally
'exceptionally beautiful' (Smith 1979: 128)

d. Mon

khoh d'əŋ
good quite
'quite good' (Bauer 1982: 387)

For this pair of elements, the preference for head-initial order is less strong among Mon-Khmer languages. The examples in (22) to (26) illustrate how a number of Mon-Khmer languages exhibit some words of this sort following the adjective but others preceding. In the (a) examples, the intensifier precedes the adjective while in the (b) examples, it follows.

(12) Vietnamese

a. rất vui
very happy
'very happy' (Thompson 1965: 222)

b. đẹp lắm
beautiful very
'very beautiful' (Thompson 1965: 272)

(13) Palaung

a. gōp ũr
very tired
'very tired' (Milne 1921: 105)

b. dāng găt
large very
'very large' (Milne 1921: 104)

(14) Khmu?

- a. lək pɲkà
really shy
'really shy' (Premsrirat 1987: 42)
- b. kò:n ləʔ phó:t
too.much good too.much
'too good' (Premsrirat 1987: 68)

(15) Katu

- a. long gamak
very big
'very big' (Costello 1969: 31)
- b. gamak pablong
big very
'very big' (Costello 1969: 31)

(16) Chrau

- a. mo'yäh mǎq
very big
'very big' (Thomas 1971: 110)
- b. mǎq trôq
big extremely
'extremely big' (Thomas 1971: 110)

Despite the number of languages in which both orders of intensifier and adjective are found, I am not aware of any Mon-Khmer language in which the normal order is Intens-Adj. Among languages in which one order is dominant, the order Adj-Intens appears to be the dominant one.

Once again, crosslinguistic data shows that there is no crosslinguistic tendency for the intensifier to follow the adjective in VO languages. The data in (17) shows that both orders are about equally common among VO languages.

(17)	Afr	Eur	SEA	A-NG	NAm	SAm	Total
VO&IntensAdj	0	8	5	1	10	7	54
VO&AdjIntens	12	0	10	1	3	3	56

We can summarize what has been demonstrated to this point as follows. Mon-Khmer languages tend to be consistently

head-initial, as reflected by the common orders NA, NDem, and AdjIntens. A common assumption is that these characteristics are simply typical of VO languages, as reflections of head-initial order. I have shown, however, that these characteristics are by no means typical of VO languages, so that their occurrence in Mon-Khmer languages cannot simply be explained in terms of a supposed tendency for VO languages to be head-initial. In the remainder of this paper, I will examine some alternative hypotheses for why Mon-Khmer languages might exhibit consistent head-initial order.

Even though I have demonstrated that VO languages do not exhibit a general tendency to be head-initial, one might suggest that the VO languages which are not consistently head-initial are languages which were once OV and have retained certain head-final characteristics. Such an explanation is plausible for languages like Finnish, for example, which have demonstrably changed their order from OV to VO while still retaining a number of characteristics from the time when the language was OV. One might argue that the VO order in Mon-Khmer is very old and that Mon-Khmer languages are in some sense “purer” VO languages, without OV ancestry. The problem with this explanation is that there are a number of families or areas in which adjectives or demonstratives commonly precede the noun or intensifiers precede the adjective and in which an explanation in terms of OV ancestry seems unlikely. This includes the Austronesian family, where there is no evidence of OV ancestry and where there is considerable variation in the order of adjective and noun, sufficient variation that the prenominal position of the adjective in many languages must be relatively recent. In addition, there are two areas in the New World in which AN order and DemN order are quite common in which VO is the normal order as part of an areal phenomenon, namely Meso-America and the Pacific Northwest. The common AN order in these areas does not seem attributable to any OV ancestry.

A further hypothesis would be that there is a subtype of OV languages which are consistently head-final and a subtype of VO languages which are consistently head-initial, that both subtypes belong to a general type that involves crosscategorical consistency in head position in contrast to languages that do not. But while there may indeed be such a subtype for OV languages, illustrated by Japanese, closer examination of Mon-Khmer languages shows that they are often not entirely head-initial, that

many of them exhibit sporadic exceptions to the trend towards head-initial order, so that in fact we cannot say that Mon-Khmer languages are consistently head-initial without exception. Consider, for example, the order of numeral and noun. While some Mon-Khmer languages are NNum, as in (18), others place the numeral before the noun, as in (19).

(18) Noun-Numeral

a. Khmu?

kò:n pà:r kòn
child two CLSFR
'two children' (Premssirat 1987: 34)

b. Khmer

phèh pì:r
house two
'two houses' (Jacob 1968: 62)

c. Mon

nəm chu? pən nəm
CLSFR tree four CLSFR
'four trees' (Bauer 1982: 359)

(19) Numeral-Noun

a. Vietnamese

ba cái nhà
three CLSFR house
'three houses' (Thompson 1965: 198)

b. Katu

pe (panong) anuq
three (CLSFR) dog
'three dogs' (Costello 1969: 26)

c. Chau

du vanông gapu
one CLSFR buffalo
'one buffalo' (Thomas 1971: 133)

d. Sre

bàr nəm ʔaso ñ
two CLSFR dog that
'those two dogs' (Manley 1972: 156)

e. Sedang

môi to ket
one CLSFR frog
'one frog' (Smith 1979: 86)

Other languages, such as Palaung and Stiêng, illustrated in (20) and (21), employ two patterns with numerals, one without a classifier, in which the numeral precedes the noun, and a second with a classifier in which the numeral plus classifier follows the noun.

(20) Palaung: Num+Noun *or* Noun+Num+Clsfr

- | | |
|----------------|------------------------------|
| a. ār rŭ | b. br̄ng ũ tō |
| two village | horse one CLSFR |
| 'two villages' | 'one horse' (Milne 1921: 57) |

(21) Stiêng: Num+Noun *or* Noun+Num+Clsfr

- | | |
|---------------|------------------------------------|
| a. p̄y snam | b. nhi baar Δc |
| three year | house two CLSFR |
| 'three years' | 'two houses' (Miller 1976: 31, 32) |

Now one might argue that in constructions with a classifier, the classifier is really the head and the noun is itself a modifier, unlike the analysis commonly assumed for European languages. However, it seems unlikely that in some Mon-Khmer languages the classifier is head while in others the noun is head. If we claim that the classifier is consistently the head, then the languages in (18), in which the numeral plus classifier follows the noun would not be exhibiting a head-initial construction. Furthermore, if we restrict attention to the order of numeral and classifier, which is apparently always Num + Classifier in Mon-Khmer languages, we have a further problem, because the numeral is most plausibly analysed as a modifier of the classifier, in which case we have a further example of an order which is not head-initial.

A further example of a type of modifier that commonly precedes the noun is plural words (cf. Dryer 1989b), separate words indicating the plurality of the noun phrase, as in (22).

(22) a. Vietnamese

những chó
PLURAL dog
'dogs' (Thompson 1965: 179)

b. Stiêng

bol ôi au
PLUR morning here
'these mornings' (Miller 1976: 34)

c. Sedang

vai drôh
PLUR girl
'the girls' (Smith 1979: 92)

Such plural words typically precede the noun in Mon-Khmer languages, again contrary to the general trend towards consistent head-initial order.

And in various Mon-Khmer languages there are further sporadic instances of other sorts of prenominal modifiers. For example, in Mon, the quantifier meaning 'every' precedes the noun (in contrast to numerals which follow the noun), as in (23).

(23) Mon: NNum but QuantN

rəh ɲoa
every day
'every day' (Bauer 1982: 343)

In (24) is a summary chart of the Mon-Khmer languages in my database. In most of these languages, we find at least one instance in which a modifier precedes the word it modifies, thus showing that completely consistent head-initial order is apparently the exception.

(24) Summary chart ³

	Gen	Adj	Dem	Num	Art	Plur	Intens	other
Khasi	NG	NA	DemN	NumN	ArtN	PIN	AdjInt/ intadj	NOrd QuantN IntN
Palaung	NG	NA	NDem	NumN/ NNum			IntAdj/ AdjInt	
Khmu?	NG	NA	NDem	NNum			IntAdj/ AdjInt	
Vietna- mese	NG	NA	NDem	NumN	DefN	PIN	IntAdj/ AdjInt	
Katu	NG/ gn	NA /an	NDem	NumN/ nnum			IntAdj/ AdjInt	ClsfrN
Stiêng	NG	NA	NDem	NumN/ NNum		PIN		
Chrau	NG	NA	NDem	NumN			IntAdj/ AdjInt	NOrd
Sre	NG	NA	NDem	NumN			AdjInt	
Brao	NG	NA	NDem	NNum				
Sedang	NG	NA	NDem	NumN		PIN	AdjInt	NOrd QuantN
Khmer	NG	NA	NDem	NNum	NArt		AdjInt	NInt
Mon	NG	NA	NDem	NNum	NDef	NPI	AdjInt	QuantN

This suggests that while there may be a general trend towards head-initial order in Mon-Khmer, we cannot say that they are consistently head-initial, thus casting doubt on the idea that the instances in which they are head-initial can be explained in terms of some general rule of head-initial order.

So why do the Mon-Khmer languages exhibit a stronger tendency towards head-initial languages than VO languages elsewhere in the world? The simplest explanation would be that these orders in MK languages are simply accidental, that their occurrence does not reflect a greater pattern, any more than a random win on a slot machine. While much effort in modern linguistics is to find appealing explanations for specific phenomena in different languages, the possibility that many phenomena have no real explanation should not be overlooked. According to this view, the fact that the adjective and the demonstrative generally follow the noun in Mon-Khmer languages, and that intensifiers so often follow the adjective are simply among the many arbitrary properties found in any language, and any attempt to reduce them to some more abstract principle may simply be misguided.

Notes

¹ It should be noted that the pattern exhibited by Mon-Khmer languages is part of an overall pattern within Asia, in which Mon-Khmer languages exhibit head-initial characteristics, shared with Daic languages, and contrasting sharply with languages to the north and west, which exhibit head-final characteristics (e.g. Japanese, Korean, Mongolian, Turkic, Uralic, Indo-Iranian, Dravidian), with a belt of languages in between exhibiting mixed characteristics (Chinese, Tibeto-Burman, Miao-Yao). The properties of Mon-Khmer languages discussed here are thus part of a larger areal pattern that includes Daic languages, and one Mon-Khmer language, Khasi, spoken in northeast India, does not exhibit as clear a head-initial pattern. For this reason, the phenomenon being discussed is really an areal phenomenon rather than something specific to Mon-Khmer, though I will restrict attention here to Mon-Khmer.

² The abbreviations for the six areas are: Afr = Africa, Eur = Eurasia, SEA = Southeast Asia and Oceania, A-NG = Australia-New Guinea, NAM = North America, and SAM = South America. SEAsia & Oceania includes the languages of Southeast Asia, Sino-Tibetan, and Austronesian languages. Eurasia contains the remaining languages of Europe and Asia.

³ Key to summary chart in (24):

Gen	Genitive	Pl	Plural Word
Adj	Adjective	Intens	Intensifier
Dem	Demonstrative	Quant	Quantifier
Num	Numeral	Ord	Ordinal Numeral
Art	Article	Int	Interrogative Modifier
Def	Definite marker	Clsfr	Classifier

References

- Bauer, Christian Hartmut Richard. 1982. Morphology and Syntax of Spoken Mon. University of London Ph.D. dissertation.
- Costello, Nancy A. 1969. The Katu noun phrase. In Mon-Khmer Studies III, Publication No. 4 of the Linguistics Circle of Saigon, pp. 21-35. Summer Institute of Linguistics.
- Dryer, Matthew S. 1988. Object-verb order and adjective-noun order: dispelling a myth. *Lingua* 74: 77-109.

- Dryer, Matthew S. 1989a. Large linguistic areas and language sampling. *Studies in Language* 13: 257-292.
- Dryer, Matthew S. 1989b. Plural words. *Linguistics* 27: 865-895.
- Dryer, Matthew S. 1991. SVO languages and the OV:VO typology. *Journal of Linguistics* 27: 443-482.
- Dryer, Matthew S. 1992. The Greenbergian word order correlations. *Language* 68: 81-138.
- Greenberg, Joseph H. 1963. Some universals of grammar with particular reference to the order of meaningful elements. In Joseph Greenberg, ed., *Universals of Language*, pp. 73-113. Cambridge, Mass: MIT Press.
- Jacob, Judith M. 1968. *Introduction to Cambodian*. London: Oxford University Press.
- Keller, Charles E. 1976. *A Grammatical Sketch of Brao, A Mon-Khmer Language*. Work Papers 20, Supplement 1. Summer Institute of Linguistics, North Dakota: University of North Dakota.
- Manley, Timothy M. 1972. *Outline of the Sre Structure*. Honolulu: University of Hawaii Press.
- Miller, Vera Grace. 1976. *An Overview of Stieng Grammar*. Work Papers, Vol. 20, Supplement 3. University of North Dakota: Summer Institute of Linguistics.
- Milne, Mary Lewis. 1921. *An Elementary Palaung Grammar*. London: Oxford University Press.
- Nagaraja, K. S. 1985. *Khasi: A Descriptive Analysis*. Pune: Deccan College Post-Graduate & Research Institute.
- Premrirat, Suwilai. 1987. *A Khmu Grammar*. In *Papers in South-East Asian Linguistics* No. 10, Pacific Linguistics, Series A, No. 75. Canberra: Australian National University.
- Rabel, Lili. 1961. *Khasi, a Language of Assam*. Baton Rouge: Louisiana State University Press.
- Smalley, William A. 1961. *Outline of Khmu? Structure*. New Haven, Conn.: American Oriental Society.
- Smith, Kenneth. 1979. *Sedang Grammar: Phonological and Syntactic Structure*. Pacific Linguistics, Series C, No. 50. Australian National University.
- Thomas, David D. 1971. *Chrau Grammar*. Honolulu: University of Hawaii Press.
- Thompson, Laurence. 1965. *A Vietnamese Grammar*. Seattle: University of Washington Press.

¹ It should be noted that the pattern exhibited by Mon-Khmer languages is part of an overall pattern within Asia, in which Mon-Khmer languages exhibit head-initial characteristics, shared with Daic languages, and contrasting sharply with languages to the north and west, which exhibit head-final characteristics (e.g. Japanese, Korean, Mongolian, Turkic, Uralic, Indo-Iranian, Dravidian), with a belt of languages in between exhibiting mixed characteristics (Chinese, Tibeto-Burman, Miao-Yao). The properties of Mon-Khmer languages discussed here are thus part of a larger areal pattern that includes Daic languages, and one Mon-Khmer language, Khasi, spoken in northeast India, does not exhibit as clear a head-initial pattern. For this reason, the phenomenon being discussed is really an areal phenomenon rather than something specific to Mon-Khmer, though I will restrict attention here to Mon-Khmer.

² The abbreviations for the six areas are: Afr = Africa, Eur = Eurasia, SEA = Southeast Asia and Oceania, A-NG = Australia-New Guinea, NAM = North America, and SAm = South America. SEAsia & Oceania includes the languages of Southeast Asia, Sino-Tibetan, and Austronesian languages. Eurasia contains the remaining languages of Europe and Asia.

³ Key to summary chart:

Gen	Genitive	Pl	Plural Word
Adj	Adjective	Intens	Intensifier
Dem	Demonstrative	Quant	Quantifier
Num	Numeral	Ord	Ordinal Numeral
Art	Article	Int	Interrogative Modifier
Def	Definite marker	Clsfr	Classifier