The syntax of nouns and noun phrases in dated pre-Angkorian inscriptions*

Chhany SAK-HUMPHRY
University of Hawaii at Manoa

1. Introduction
1.1 Background

Any diachronic analysis of Khmer must be based on the extensive corpus of inscriptions.\(^1\) According to Jacob (1960: 351; 1965: 143; \(*1991\)), conventionally recognized periods in the development of Khmer are ‘Old Khmer’, ‘Middle Khmer’, and ‘Modern Khmer’. The Old Khmer period includes the pre-Angkorian Khmer of the earliest inscriptions to A.D. 802 and Angkorian Khmer, attested from 802 to 1431 (the fall of Angkor). ‘Middle Khmer’ covers the period of transition from Old to Modern Khmer, from approximately 1431 to 1800, and marks the culmination of various complex changes in both the phonology and grammar. ‘Modern Khmer’ is considered to extend from about 1800 to the present. All three periods are, of course, historical fictions, or rather historical-linguistic constructs, in the sense that there were no abrupt demarcations between one stage of the language and the next.

1.2 Previous analyses of pre-Angkorian syntax

Whilst a number of scholars have contributed significantly to—and indeed laid the foundations of—the historical study of Old Khmer, there are so far no syntactic analyses of Old Khmer based on a generative framework.\(^2\) Jenner

* This work is a revision of my master’s thesis in linguistics at the University of Hawaii, Manoa.

I would like to express my sincere gratitude to Professor Stanley Starosta, who is the founder of lexicase dependency grammar. He introduced me to this theory and has provided continuous support and scholarly comments on this manuscript. My warmhearted appreciation also goes to Professor Philip N. Jenner, who taught me Old Khmer, and who has supported and encouraged me to pursue a career in linguistics. I also would like to thank Gary Y. Takeuchi for his careful review and editing of the English text.

\(^1\) The Founan period is a historical period. The inscriptions from this period are in Sanskrit; the vernacular language of Founan is unknown.

published a series of articles on Old Khmer syntax, and we possess now a
diachronic survey of Khmer function words by Jacob.\textsuperscript{3}

Texts of most of the pre-Angkorian inscriptions can be found in Cœdes’
eight-volume *Inscriptions du Cambodge* (1937–66) or scattered throughout the
*Bulletin de l’Ecole Française d’Extème-Orient* [BEFEO]

1.3 Primary sources

The pre-Angkorian corpus can be divided into dated and undated
inscriptions.\textsuperscript{4} The data used in this investigation of the syntax of nouns and noun
phrases of the pre-Angkorian dated inscriptions are not drawn from the original
texts, but rather from the transliterated versions found in Jenner’s (1983–84)
unpublished work. Sixty-six texts have been taken from two volumes of Jenner’s
work and are listed in Appendix A.

Working with data from an ancient language imposes serious limitations on
the analysis. The compilation and interpretation of the texts, the primary data for
the analysis, is itself a major analytical task. As Jenner put it: “The linguistic
archaeologist must first consider the orthographic shape of each item; he must then
translate that shape into phonological terms, analyze its morphology, determine its
syntactic function, identify it with later or cognate forms, and eventually set up a
hypothesis regarding its meaning” (Jenner *CPAK* 1981–82: iv). All this work must
be done without a co-operative native-speaking consultant whose intuition can be
accessed to confirm or discount a particular hypothesis.

1.4 Goals and objective of this study

The purpose of this study is twofold: (1) to apply lexicase dependency
theory to the analysis of certain grammatical aspects of nouns and noun phrases of
dated pre-Angkorian inscriptions to determine whether the grammatical
properties of these inscriptions can be insightfully described and illuminated
within this formal and explicit theory, and (2) to identify any areas in which the
data prove to be incompatible with the claims made by the theory, thereby
possibly necessitating a modification of the theory itself.\textsuperscript{5} It is to be expected that
in the absence of a native-speaking consultant, a narrowly constrained universal
linguistic theory will provide assistance in choosing among alternative hypotheses
about the correct analysis of particular constructions. At the same time, data from
a natural text from another non-Indo-European language will be helpful in
developing a theory that can seriously claim to be universal.

\textsuperscript{3} The latter was not available to me at the time of writing; Jenner’s contribution to the
London Shorto Festschrift concerns Angkorian Khmer. Pou contributed a survey of some basic
OKhm. constructions to this journal (1980).

\textsuperscript{4} Jenner says that “the dated inscriptions are those which contain their own internal dating”
(Jenner 1982).

\textsuperscript{5} See Tesnière 1959 for a general reference on dependency grammar and Starosta 1988 for
an introduction to the lexicase implementation of this theory.
1.5 Methodology

The corpus of this study is composed of the texts presented in Philip N. Jenner’s unpublished work, "Textes Vieux Khmers faisant partie du Corpus des Inscriptions du pays khmer" (Tome 1.1 et 2, Tome 1.3, 1983–4). I began analyzing the texts of the earliest of the dated inscriptions, which had already been segmented into their constituent sentences by Jenner’s translations of each text in French. Each example cited in this thesis is labeled with a number identifying its location in these texts. Each word in the texts was checked against Jenner’s pre-Angkorian lexicon and labeled in terms of the categories made available by the lexicase theory, and an explicit lexicase dependency representation was drawn for each sentence. These representations conform to the constraints imposed by the lexicase theory. The purpose of this task was not to check on Jenner’s translation of the text, but rather to investigate the structure of the sentences of the pre-Angkorian dated inscriptions within the lexicase framework. However, this study indicates places where the requirements of the theory suggested alternative translations or analyses that differ from those proposed by Jenner.

1.6 The Lexicase model

This investigation of nouns and noun phrases of pre-Angkorian dated inscriptions is formulated within the lexicase dependency grammar framework developed over the last twenty years, primarily by Stanley Starosta and was carried out in accordance with the principles and constraints of this theory.

A lexicase grammar is a grammar of words. It has no deep structure, no transformations, and no phrase structure rules. It represents the structure of a sentence solely in terms of a network of dependency relations obtaining among pairs of words in the sentence. Co-reference relationships are described in terms of coindexing words. The syntactic and semantic properties of words are characterized in terms of contextual and non-contextual features of lexical entries (Starosta [forthcoming]). This theory attempts to capture cross-linguistic generalizations, and makes a claim about human language in general.

A lexicase grammar is a set of generalizations about the internal compositions, external distributions, and lexical relationships of the words in the language (Starosta 1988: 2). The relationship among lexicase rules can be represented in terms of the flow chart in Appendix B.

The analysis presented here does not include the set of formalized rules that would be required by a complete grammar. It does, however, provide fully specified dependency representations of a broad range of example sentences relevant to a description of the nouns and noun phrases of the Old Khmer dated inscriptions. For each representation, the functional label of each binary dependency relationship is specified in terms of an indexed contextual feature in the matrix of the regent term of the relation. Such labels include case relations [CRs], case forms [CFs], predication [prdc], or finiteness [fint] of the dependents.
This study is divided into the following sections:

Section 1 — Introduction
Section 2 — Overview of Dated pre-Angkorian syntax
Section 3 — Noun phrases with a single noun and no dependents
Section 4 — Noun phrases with a single noun and one / multiple dependents
Section 5 — Multiple dependent constructions; possessive, locative, equative
and prepositional phrases
Section 6 — Conclusion
Appendices
References

2. Overview of dated pre-Angkorian syntax

The focus of this analysis is on noun and noun phrase structures. Because so many examples of the constructions in the data presented here are non-sentential, and since this is an accurate reflection of the language of Old Khmer inscriptions, a brief description of the overall structure of sentence patterns of the dated pre-Angkorian inscriptions is needed to set the foundation of this work. This overall description provides a sketch of the general syntactic properties of the Old Khmer language, many of which are relevant to a description of the constructions which depend on N[noun]s in N[noun]P[hrase] structures. In this study, the term 'Old Khmer' [OKhm.] is used to mean 'Old Khmer dated pre-Angkorian'.

2.1 Basic sentence patterns

A basic sentence structure has only one clause. A sentence is any phrase that has a word marked [+prdc] as its lexical head, that is, a verb, or a prdc-inflected preposition or a noun. The focus of this section is on the analysis of sentences with and without verbs. The outline of this presentation is: (1) a brief description of verbless sentences and their subcategorization, (2) verbal sentences, and (3) conjoined sentences.

2.1.1 Verbless sentences

This section examines sentences without verbs as the head of the constructions. In principle, lexicase allows two types of verbless sentence constructions: (1) the nominal predicate where the regent noun \( N_2 \) is the head of the construction, bears the feature [+prdc], and has \( N_1 \) as its nominative/patient dependent; and (2) the predicate prepositional phrase where the preposition \( P_1 \) itself bears the [+prdc] feature.

The Old Khmer pre-Angkorian dated inscriptions have only one example of an independent clause with a nominal predicate NP (see example 8 of section 4.2.2.1) and no examples of prepositional predicate phrases as described above. Instead, most of the verbless sentences occur as relative clauses in NPs composed minimally of two nouns, a regent noun \( N_1 \) and a nominal dependent \( N_2 \), which is
the head of a verbless relative clause and which bears the feature [+prdc]. As is the case in relative clauses generally, the non-verbal relative clause lacks one of its NPs, in this case its subject (+Nom, PAT); this missing NP, in turn, is interpreted as coreferential with the regent N1. The significance of this nominal predicate is that the noun N2 is a relative clause of the NP-predicate clause pattern described in general terms above. A schematic representation of this pattern is shown below. A more detailed analysis of this nominal predicate relative clause construction can be found in section 4.4.4.

Nominal predicate in the relative clause construction.

Here [S [NP N] N2[+prdc] ] is the form which the verbless relative clause would be expected to take if it were not functioning as an attribute of a regent N1. The implied subject N[index] is missing and is coreferential with the regent N1. The relative clause has the predicate N2 as head, and N2 is the dependent of N1.

The following is an additional schematic example:

1. (K.54:12)

kñum 'amnoy sināhv pragat
slaves gift hermit pious
‘Slaves the gift of pious hermit’
The missing but implied \( \text{kñum} \) is the nominative patient of the regent predicate noun \( \text{amnøy} \) and is interpreted as coreferential with the regent NP \( \text{kñum} \); in turn \( \text{kñum} \) is the regent of the relative clause \( \text{amnøy sināhv pragat} \).

2.1.2 Verbal Sentences

A verbal sentence has as its construction head a verb that bears the \([+\text{prdc}]\) predicate lexical feature. Due to the nature of the data, some of the verbal S types are only attested in relative clause constructions. The word order of the S constituents is subject–verb–object. The verb is subcategorized into transitive and intransitive types. Old Khmer transitive verbal sentences will be examined next.

2.1.2.1 Transitive verbal sentences

In the transitive clause the V \([+\text{trns}]\) is the head of the construction. It implies an Agent subject \( N_1 \) indicated by the features \([1(+N)], [1(+\text{AGT}], [1(+\text{Nom}])\), and a Patient object \( N_2 \) indicated by the features \([3(+N)], [3(+\text{PAT}], [3(+\text{Acc}]]\). This is an example of an Old Khmer transitive clause:

2. (K.493:19)

\[
\text{poñ bhā vinaya ktīn krapi canmat} 1 \\
\text{Sir Bhā Vinaya owe carabao uncastrated} 1
\]

'Sir Bhā Vinaya owed one uncastrated carabao'
In the lexicase dependency representation, a vertical line indicates the head of the construction, the slanted lines indicate the dependents which are written two steps below the level of the head word to its right or its left. Each constituent word is indexed to establish the dependency link to one other word. The verb *ktiṅ* [+trns] is the head of the clause, the NP *pōṅ bhā vinaya* is to the left, and the NP *kraṅj canmat* 1 is to the right.

The grammatical functions of this tree are shown by linking the contextual features on the head of V to the index of the dependent attributes. The [1[+Nom]] and [1[+AGT]] imply that the verb *ktiṅ* expects an [Index] dependent to be marked by the ‘nominative’ case form [+Nom], and interprets it as bearing an ‘agent’ case relation [AGT]. The [4[+PAT]] and [4[+Acc]] imply that the head word *ktiṅ* requires its [4Index] dependent to bear an ‘accusative’ case form [+Acc], and interprets it as having a ‘patient’ case relation [PAT]. Thus the transitive verb *ktiṅ* identifies *pōṅ* [Index] as its subject [AGT, +Nom] and *kraṅj* [4Index] as its object [PAT, +Acc]. The subject *pōṅ* is a title [+titl] noun and has the proper [+prpr] noun *bhā vinaya* as its predicate [+prdc] dependent attribute. The patient object *kraṅj* is the regent of the predicate classifier [+clsf] noun *canmat*, and the predicate number noun 1.

The following shows that the grammatical representations of this verbal sentence are stated not in terms of layers of constituents, but in terms of pairwise dependency relations between a dominant regent *ktiṅ* and dependent adjuncts *pōṅ* and *kraṅj*; regent *pōṅ*, adjunct *bhā vinaya*; regent *kraṅj* and adjunct *canmat*; and regent *canmat* and adjunct number 1.

The transitive verbs can be subcategorized in terms of whether they take locational or non–locational dependents.

### 2.1.2.1.1 Non–locational transitive verbal sentence

Previous examples have demonstrated where verbs act as case markers and expect [+[AGT, +Nom], [+]PAT, +Acc] and even [+]LOC as dependents. In this section the focus is on verbs that do not require LOC dependents.

Example 3 shows the non–locational transitive verb *dār* with an adverb *ukk* as its dependent.

3. (K79:19)

\[
\begin{align*}
\text{pōṅ mano dār} & \quad \text{canleq ukk yugala yau 1} \\
\text{Sir} & \quad \text{Mano demand cloth in addition double unit 1}
\end{align*}
\]

‘Sir Mano also demanded [a certain amount of] cloth in addition, [namely] 1 yau of double [cloth].’
In this sentence, the transitive verb dār ‘demand’ is the head of the clause. It has poñ as its nominative agent [+Nom, AGT], canlek as its accusative patient [+Acc, PAT], ukk as its adverb, and yugala as its other accusative patient [+Acc, PAT]. This construction is interesting because the transitive verb dār has two consecutive accusative patients, in violation of the lexicase 1/Sent constraint. Usually yugala (the second accusative patient) is the predicate dependent of the word canlek (the first accusative patient), but because of the presence of the adverb, it must be analyzed as a separate constituent in this example.

### 2.1.2.1.2 Locational transitive verbal sentences

In Old Khmer there are some ditransitive verbs like oy : ‘to give’, jāhv : ‘to acquire’, dār : ‘to demand’, and jon : ‘to offer’, which bear the semantic features [+goal] or [+source] and obligatorily expect one of their dependents to bear the [+Lctn] case form and [+LOC] case relation as complement (allowing for zero anaphora).

The tree in example 4 illustrates a locational transitive sentence where the verb oy [+goal] implies a LOC dependent.

In example 4, the transitive verb oy [+goal] implies a LOC dependent, has a coordinate NP with the conjunction dañ as its dependent attribute.

4. (K.18:3)

| mratān bhāskarapāḷa oy sre dañ kñumma |
|------------------------|---------------------|
| Lord Bhāskarapāḷa give ricefield and slaves |

ta\(^6\) vrah kamratān 'añ śrīsankarāṇārayana to V. K. 'Añ Śrīsankarāṇarayaṇa

‘The Lord Bhāskarapāḷa gave the ricefield and slaves to the V. K. 'Añ Śrīsankarāṇarayaṇa.’

\(^6\) Jenner interprets ta as ‘subordinating conjunction, normally optional.” (Jenner, personal communication, February 17, 1992)
In this example the verb oy, the head of the construction, has three dependents, the nominative NP mraitān bāskarapāla [+Nom, AGT], the coordinate accusative NP sre dañ knumma, with dañ as the coordinating conjunction [+Cnjc], and the locational PP ta₁ vrah, where the noun vrah bears the LOC complement to the verb oy.

2.1.2.2 Intransitive verbal sentences

In the intransitive clause construction, the V as head of the clause, requires only a patient subject [PAT, +Nom]. The intransitive verbs are subcategorized into the copula and non-copula classes.

2.1.2.2.1 Intransitive sentences with copula verbs

A copula is an intransitive verb that takes a patient subject and a predicate noun, stative verb, or predicative PP as its complement (Starosta 1991). Here we are looking at the nominal predicate dependent of the copula verb gui.7

In example 5, the copula verb gui equates the patient subject neh, which is a demonstrative noun, to its human predicate object knum vrah.

7 Jenner interprets gi as “a weak demonstrative pronoun commonly standing before a zero copula so that its proper function became blurred by the Middle Khmer period.” (Jenner, personal communication, February 17, 1992)
5. (K.388C:7)

\[
\text{neh gui kñum vrah}
\]

these be slaves Vrah

'These are the slaves of the Vrah.'

The copula verb *gui* [+copl] has a demonstrative noun *neh* [+dmns] as patient subject and a predicate NP *kñum* [+humn] as dependent attribute.

2.1.2.2.2 Intransitive clauses with non–copula verbs

The intransitive non–copula verb is examined next. The intransitive non–copula verb *dana* has a nominative patient pronoun *ge*.

6. (K.49:17)

\[
ge \quad \text{dana}
\]

they be punished

'They shall be punished.'

The intransitive verb *dana* is the head of the clause where the pronoun *ge* [+Nom, PAT] is its subject.

This type of construction can be further subcategorized into verbs requiring or not requiring locational nouns as dependent attributes.
2.1.2.2.2.1 Intransitive verb with location noun as dependent

An example of an intransitive verb with a locational noun dependent is shown in example 7.

7. (K.561:21)

ge  dau  'avיכinarakk
they  go  'Avיכinarakk

'They go to the 'Avיכinarakk.'

This intransitive clause has the movement verb *dau* as the head, the pronoun *ge* as subject, and the locational noun 'avיכinarakk as LOC complement. The intransitive verb *dau* requires a LOC complement as dependent. Its LOC complement can be either a [+lctn] location noun, as in this case, or a relator noun as demonstrated in the following section.

2.1.2.2.2.2 Intransitive verb with relator noun as dependent

Example 8 shows the intransitive verb *lanlyan* 'bur' is the head of the clause. In this section the relator noun is the dependent of a verb in the sentence.

8. (K.341N:11–2)

ge  lanlyan kaмluи niraya nu gotra phon
they burn in Hell with kinsmen all

'They [shall] burn in Hell together with their kinsmen.'
Here the relator noun *kamluṅ* is the dependent attribute of the intransitive verb *lanlyañ*. *Kamluṅ* has the localistic feature [+ntrr], which characterizes the specific kind of relationship associated between the regent verb *lanlyañ* 'burn' and the dependent noun *niraya* 'hell'.

### 2.2 Conjoined sentences

This study assumes that full sentences in Old Khmer could be conjoined by using the coordinate conjunctions *don* or *dañ* or *droñ* or *drañ* [+Cnjc]. Following lexicase dependency grammar principles, the coordinate conjunction is the lexical head of the conjoined constructions, not the S or the NP. These conjunctions presumably were able to join two complete independent sentences. So far, probably because of the nature of the texts, examples of full sentences joined with a coordinate conjunction have not been found. As stated previously, the absence of such examples is an accurate reflection of the language of the inscriptions, in that most of them are in the form of NPs. Coordinate NPs with and without a conjunction are discussed in section 4.1.

#### 2.2.1 Fulls with *don*/*dañ* gap

Example 9 illustrates a full sentence with the missing coordinate conjunction *don*/*dañ*.

9. (K.49:17)

```
ge cer 'ājña vrañ kamṛtañ 'aṅ ge dañḍa
they disobey order V. K. 'A. they be punished
```

'They [thereby] disobey the order of the V.K.'A. [and] they shall be punished.'
Example 9 is potentially two independent sentences. The verb cer [+V, +trns] is the head of the transitive verbal sentence where ge [+Nom, AGT] is the nominative agent, and 'ajña [+Acc, PAT], is the accusative patient. The intransitive verbal sentence ge dānda has ge [+Nom, PAT] as the nominative patient of the intransitive verb dānda.

These sentences can be considered coordinate rather than independent sentences because both share the same subject ge, as it is this ge that does not obey the order of the V.K.'A. These two verbal sentences are analyzed here as being conjoined with the missing coordinate conjunction [dor/dan]. This is necessary because of the lexicase dependency grammar requirement that every phrase, including sentences, have a word that acts as the head of the whole construction. This is not a very intuitively satisfying analysis, but no alternative is made available by the theory at this point in time.

Before closing this section on an overview of Dated pre-Angkorian Inscriptions syntax, I would like to illustrate a complex verbal sentence from the data.

10. (K.49:13)

'amupal kñum tmur krapı sre dănṛn
all slaves cattle buffaloes ricefields plantation

guí ta man ge pu cah 'añ oy ta vrah
be of what they Elder Lord Our give to Vraḥ

don kñum pradāna 'nak ta psañ
and slaves to be given by people who contribute

ta gui ukkra gui tel prasiddha
that who be also be which convey

All these slaves, cattle, buffaloes, ricefields and plantations—those things which Our Elder Lords have given to the shining one—as well as the slaves presented by persons also contributing to [this pious work] are conveyed.
The copula verb gui [8ndex] is the main verb of this sentence and has 'ampal as its subject and ta₄ man as its predicate. The noun 'ampal is the regent of the predicate relative clauses kǔm, tmur, krap, sre, damrini. The predicate PP ta₄ man has ta₄ [9ndex] as complementizer of the relative clauses ge pu caḥ 'añ oy ta₁ vraḥ don kǔm pradāna 'nak ta₄ psam ta₄ gui ukkra. The coordinate
conjunction don [17ndex] conjoins the verbal relative clauses ge pu cah ‘aň oy taŋ vrâh and khûm pradâna ‘nak taŋ psam taŋ gui ukkra. If the verb gui [8ndex] and the gui [27ndex] are conjoined with the missing conjunction [doň] then the head of the S is the missing [doň].

The predicate relative clauses with ‘ampal as the regent list all gifts that are involved in the transaction. The second relative clause with taŋ man as regent gives the description and information about the source of these gifts and its recipients.

3. Noun phrases with a single noun and no dependents

A noun phrase is headed by a noun. In nominal attribution, the head noun may have zero, one, or more dependent modifiers.

The purpose here is to analyze the structure of nominal attribution in noun phrases, by establishing: (1) the grammatical classification of Pre–Angkorian nouns, (2) the classification of noun dependency relationships, and (3) the multiple noun dependency relationships.

In this section noun phrases without reference to their dependent attributes will be examined. Nouns in Old Khmer, as in many other Southeast Asian languages, do not require a determiner as attributes.

Based on syntactic function and morphological criteria, eight pre–Angkorian noun classes can be distinguished.

Except for relative nouns, nouns that can occur without attributes are: pronouns, demonstrative nouns, common nouns, number nouns, classifier nouns, and proper nouns.

Noun phrase modifiers are optional in all seven types of head nouns; they also depend upon the context of the sentences. In this section, analysis focuses only on zero–attributes. Cases where attributions are allowed in the structure are analyzed in Chapters IV and V.

3.1 Pronouns

It is not clear whether pronouns constitute a syntactically definable subclass of nouns in Old Khmer pre–Angkorian language. In this thesis the term ‘pronoun’ is used as a notional label to refer to a small set of nouns used as substitutes for other nouns, whose reference is set or constant for a given discourse. They are marked by the feature [+prnn] here.

Following the criteria of X–bar syntax, pronouns in lexicase are identified as a subset of nouns since they are lexical heads of NP. All personal pronouns such as ‘aň ‘I, me, my, our’, and ge ‘he, she, they’ are lexically definite, and bear the features [+N, +prnn, +dfnt]. They will be examined first.
3.1.1 ‘aṅ and ge as personal pronouns

The personal pronouns in Old Khmer are: ‘aṅ for first person and ge for third person. These pronouns have many syntactical functions based upon the dependency relationship to their regent as shown in the following examples.

In examples 1 and 2, the pronouns ‘aṅ and ge function as the COR ‘possessor’ to their regent nouns.

1. (K.44:8)

\[\begin{array}{c}
tān & ‘aṅ & kloṅ & raṅko \\
\text{retainer} & \text{our} & \text{Commissioner} & \text{Rice} \\
\text{+N} & \text{COR} \\
\end{array}\]

‘our retainer the Commissioner of Rice’

2. (K.451S:3)

\[\begin{array}{c}
jmaṅ & ge \\
\text{names} & \text{3rd} \\
\text{+N} & \text{COR} \\
\end{array}\]

‘their names’

In examples 3 and 4, ‘aṅ and ge function as the Locus LOC complement to their regents respectively.

3. (K.54:16)

\[\begin{array}{c}
spid dik tān prājñasan ‘ay ta₁ ‘aṅ \\
\text{libation retainer Prājñasan at to me} \\
\text{+N} & \text{LOC} \\
\end{array}\]

‘the libation of the retainer Prajñasen to me’

4. (K.493:23)

\[\begin{array}{c}
nu man gui saṅ kara ta₁ ge \\
\text{with which be pay back tax to him} \\
\text{+V} & \text{LOC} \\
\end{array}\]

‘with which to pay back the tax to him’
In example 5, the pronoun \( ge \) functions as nominative patient to the directional verb \( dau \).

5. (K.561:21)

\[
\begin{align*}
\text{ge} & \quad \text{dau} & \quad \text{'avicinarakk} \\
\text{they} & \quad \text{go} & \quad \text{Avici Hell}
\end{align*}
\]

‘they [shall] go to Avici Hell’

In summary, in the no–dependent construction, the personal pronoun ‘añ or \( ge \) can be the dependent attribute of a noun, verb, or preposition. \( Ge \) or ‘añ can function as the subject or Agent, object or Patient or the localistic LOC complement of a verb. \( Ge \) or ‘añ can bear [+Nom] or [+Acc] case forms, or [COR], [PAT] or [LOC] case relations to its regent.

The pronouns ‘añ or \( ge \) can be singular or plural and are not dependent upon the syntactic relationship to their regent, but this plurality is based upon the semantics of the regent and the context.

3.1.2 Derived pronouns

Observation indicates that the language of the pre–Angkorian period has so few true pronouns that titles and kinship terms seem to do double duty. Based upon the position of these nouns, which are attributes of the head NPs, they can be identified as derived pronouns, analyzed as carrying the features [+titl, +prnn, +dfnt]. Derived pronouns are title nouns that have no dependents and always bear the Correspondent case relation optionally allowed by their regents. The differences between a title noun and a derived pronoun are that a title can be: (1)
the regent of an equative predicate proper NP, (2) the regent of an equative predicate number NP or (3) the regent of an equative predicate title NP, while a pronoun does not occur in any of these functions.

Examples of derived pronouns are presented in this paper according to their functions in relation to their regents.

6. (K.505:13)

\[
\text{vā} \quad \text{men} \quad \text{kantai} \quad \text{vā} \quad \text{ku} \quad \text{lah}
\]

male slave Men wife his female slave Lah

‘the male slave Men, his wife the female slave Lah’

\[\text{In example 6, the first vā has the noun men as its dependent, and the second noun vā is a derived pronoun with no attribute. The antecedent of this derived pronoun vā is vā men, and both cooccur in the same phrase. The derived pronoun vā functions as COR to its regent kantai.}\]

7. (K.54:14–15)

\[
\text{'amnōy} \quad \text{somakīrtri} \quad \text{ta₁} \quad \text{vrah} \quad \text{ku} \quad \text{kdok}
\]

gift Somakīrtri to him female slave Kdok

\[\text{In example 7, the derived pronoun vrah is the attribute of the complementizer ta₁ and functions as LOC to the noun 'amnōy, the regent of PP.}\]

3.2 Demonstrative nouns neh and noh

Data indicate the demonstrative nouns neh and noh appear in two different grammatical functions. The non-predicate neh₁ or noh₁ functions as nominative patient subject [+Nom, PAT] to its regent verb; and the predicate neh₂ or noh₂ functions as a [+prdc] complement to its regent noun. This section focuses on the demonstrative nouns neh₁ or noh₁ which have no dependents.
In example 8, *neh₁* is the nominative subject [+Nom, PAT] of the copula verb *gi ~ gui*.

8. (K.388C:15)

\[
\begin{align*}
\text{neh₁} & \rightarrow \text{gi} \rightarrow \text{amnøy} \rightarrow \text{upādhāya} \\
\text{these} & \rightarrow \text{be} \rightarrow \text{gift} \rightarrow \text{Upādhāya} \\
\end{align*}
\]

‘these are the gifts of Upādhāya’

3.3 Common nouns

The following examples show the various syntactic relationship between the common nouns (with no-attributes) and their regents.

In example 9, *kantai* functions as predicate to its regent noun *kñum*. In example 10, *kñum* functions as accusative patient to its regent verb *oy*. In example 11, *pitr* functions as LOC to its regent verb *oy*.

9. (K.18:6)

\[
\begin{align*}
kñum & \quad \text{slave} \\
kantai & \quad \text{female} \\
ku & \quad \text{female slave} \\
\text{kmer} & \quad \text{Kmer} \\
1 & \quad \text{1} \\
kon & \quad \text{child} \\
4 & \quad \\
\end{align*}
\]

‘the female slaves: one female slave Kmer [and] four children’

10. (K.600:1)

\[
\begin{align*}
pōn & \quad \text{Sir} \\
oy & \quad \text{give} \\
ai & \quad \text{slaves} \\
ta & \quad \text{at} \\
kpoñ & \quad \text{Kpoñ} \\
kamratān & \quad \text{Kpōñ K} \\
\text{aṅ} & \quad \text{‘A} \\
\end{align*}
\]

‘Sir Uy gave slaves to the Kpoñ K. ‘A’
11. (K.561:21)

kñum man poñ candrānna oy ta pitṛ
slaves whom Sir Candrānna give to dead +V LOC

‘the slaves whom Sir Candrānna has given to the dead’

3.4 Number nouns

Number nouns can be predicate dependent attributes of countable nouns and classifier nouns. The analysis of number nouns is presented in section 4.4.4.3.

Example 12 shows the number noun ponna functions as predicate to its regent ‘anak, and the common noun ‘aṅgana functions as COR to the same regent.

12. (K.1004:8)

'anak  'aṅgana ponna
people court four

‘four people of the court’

3.5 Classifier nouns

For Old Khmer, it is an unusual pattern for a number noun to be the regent of a regular classifier noun. However, this is the way they appear in data (detail see section 4.4.4.3).

13. (K.388C:5)

tmur tap dnem
cow 10 yoke

‘10 yokes of cows’
The regular classifier *dnem* [+clsf] usually has a number noun as its dependent attribute. However, in this example the number *tap* is the head of the NP and the regent of the classifier noun *dnem*. *Dnem* has no dependent.

3.6 Proper nouns

Proper nouns in Old Khmer Dated Pre–Angkorian are divided into locational nouns and non–locational nouns. Locational proper nouns consist of place names as well as temporal nouns, i.e., the days of the week and months of the year. These nouns, which are marked with the localistic location feature [+lctn] and can serve as immediate LOC complements of words requiring [+lctn] dependents are discussed in section 4.4.2 of Chapter IV.

The next section focuses only on non–locational proper nouns. Non–locational proper nouns can be subcategorized into deity or non–deity proper nouns. The non–deity proper nouns are subdivided into human or non–human proper nouns. Deity proper nouns are names of the Vraḥ, human proper nouns are everyone’s names, including slaves, and non–human proper nouns are names of institutions, administrations, or things.

In general, most proper nouns are predicate attributes of the regent title nouns or kinship nouns. In some cases, when the title is mentioned once at the beginning of a paragraph, the proper name is not preceded by the title for the second or third references.

Example 14 has a human proper noun *jñanaparakasa* [+prdc] as dependent of the regent title noun *poñ*.

14. (K.561:13–14)

<table>
<thead>
<tr>
<th>satra</th>
<th>poñ</th>
<th>jñanaparakasa</th>
</tr>
</thead>
<tbody>
<tr>
<td>oblation</td>
<td>Sir</td>
<td>Jñanaparakasa</td>
</tr>
<tr>
<td></td>
<td>+N</td>
<td>+prdc</td>
</tr>
</tbody>
</table>

‘the oblation of Sir Jñanarakasa’

The category of compound nouns is defined morphologically rather than syntactically. A detailed investigation of pre–Angkorian word formation has not
been done; however nouns tentatively identified during the course of the analysis are described in this section.

Examples 15 and 16 show that the compound nouns 'nak sre and tnai vrah have different syntactic functions [COR] and [PAT], respectively, in relation to their regents nouns.

15. (K.600:E:3)

\[
\text{jma\text{h} 'nak sre names ricefield workers} +N \text{COR}
\]

'names of the ricefield workers'

16. (K.557/600:N:3)

\[
\text{ta}_4 \text{ cuh tnai vrah that which keep record days} +V \text{ +Acc PAT}
\]

'that which keeps record of the holy days'

In example 17, the title compound noun kura\text{k jme\text{n}} functions as the LOC complement to its regent verb j\text{h}v.

17. (K.943:20)

\[
\text{sre phalad\text{a} man j\text{h}v ta ricefield Phalada which [he] acquire from}
\]

The Phalada ricefield, which [he] acquired from the lieutenant governor and Sir Vidy\text{a}sakti
 Examples 1-17 show that pronouns, demonstrative nouns, common nouns, number nouns, classifier nouns, and proper nouns can all occur without attributes and have various syntactic functions: case forms (Acc, Nom, or Lctn) and either case relations (AGT, PAT, LOC, COR), or predicate (prdc), based upon the dependency relationship with their regents.

4. Noun phrases with a single noun and one / multiple dependents

4.1 Introduction

In lexicase representation each phrase must have a lexical head. The lexical head of the construction is written under a vertical line, and the dependents are written below the level of the head word on a line slanted either to the right or to the left. Based on my analysis of Old Khmer as a language with no determiner or adjective, and on the structure of Old Khmer noun phrases, the slant is always to the right; that is, Old Khmer noun phrases are right-branching. This is true of other syntactic constructions as well, with the exception of subjects of verbs, which slant to the left. The grammatical functions in a lexicase representation are shown by chaining a contextual feature on the head to the index of the appropriate dependent. A classification of the dependent modifier of nouns is presented first. Next, types of relationships that can obtain between a head noun and a dependent modifier are examined and described.

4.1.1 Classification of NP–Internal dependency relationships

A noun phrase can have as its immediate dependent modifier a sentence, a prepositional phrase, or another NP. These three categories are illustrated in the following tree diagrams:

Diagram #1              Diagram #2              Diagram #3

\[ \text{NP} \quad \text{NP} \quad \text{NP} \]
\[ \mid N_1 \quad \mid N_1 \quad \mid N_1 \]
\[ S \quad PP \quad NP \]

Noun phrases with Ss as dependent attributes (diagram #1) are examined in section 4.2. Noun phrases with PPs as dependent attributes (diagram #2) are investigated in section 4.3. Noun phrases with other NPs as dependent attributes (diagram #3) are studied in section 4.4.

4.1.2 Conjoined noun phrases

Noun phrases can be conjoined by using the coordinate conjunctions doni, dañ, droñ or drañ, as shown in diagram #4.
Diagram #4

\[ \begin{array}{c}
N_1 \quad NP \\
\downarrow \quad \downarrow \\
doh \quad +Cnjc \\
\end{array} \quad \begin{array}{c}
N_2 \quad NP \\
\downarrow \quad \downarrow \\
dan \quad +Cnjc \\
\end{array} \]

In diagram #4, in accordance with strict dependency principles, the coordinate conjunction *doh* is the lexical head, not *N_1* or *N_2*.

4.1.2.1 *Conjoined noun phrases with an overt coordinating conjunction*

Example 1 shows a coordinate conjunction NP with *dan* as its lexical head.

1. (K.18:21)

\[ \begin{array}{c}
sre \\
\downarrow \\
\text{ricefield} \\
\end{array} \quad \begin{array}{c}
dan \\
\downarrow \\
\text{and} \\
\end{array} \quad \begin{array}{c}
knum \\
\downarrow \\
 slavess \end{array} \]

'ricefields and slaves'

Example 2 illustrates two possessive noun phrases *punya mratâñ* and *kloñ me* conjoined with the coordinate conjunction *dan* as their lexical head.

2. (K.18:23)

\[ \begin{array}{c}
punya \\
\downarrow \\
\text{pious work} \\
\end{array} \quad \begin{array}{c}
mratâñ \\
\downarrow \\
\text{lord} \\
\end{array} \quad \begin{array}{c}
dan \quad +Cnjc \\
\downarrow \\
\text{and} \\
\end{array} \quad \begin{array}{c}
kloñ \\
\downarrow \\
\text{his} \\
\end{array} \quad \begin{array}{c}
me \\
\downarrow \\
\text{mother} \\
\end{array} \]

'the pious work of the lord and his mother the baroness'
Example 3 exhibits two coordinate conjunctions doni. The doni’s are the regents of the conjoined title possessive noun phrases poñ rudrabhava, poñ rudrantakila, and poñ puspananda.

3. (K.30:6-7)

sre 'ây travañ vo pradâna poñ rudrabhava don poñ ricefields at Travañ Vo gift Sir Rudrabhava and Sir rudrantakila don poñ puspananda Rudrantakila and Sir Puspananda

‘ricefields at Travañ Vo the gift of Sir Rudrabhava and Sir Rudrantakila and Sir Puspananda’

Example 4 shows NP coordination within the PP exocentric construction. The coordinate conjunction doni is the lexical regent of the Locus noun vrañ.

4. (k127:14)

kîm 'amnōy mratañ īśvaravindu ta vrañ kamratañ 'añ slaves gift Lord Īśvaravindu to V. K. 'A. suvarñalîṅga doni vrañ kamratañ 'añ maniśiva Suvarñalîṅga and V. K. 'A Maniśiva

‘slaves the gift of the Lord Īśvaravindu to V. K. 'A. Suvarñalîṅga and V. K. 'A Maniśiva’