THE STRUCTURE OF THE WORD IN OLD KHMER

Since the publication in 1954 of M. Cœdès' sixth and last volume of edited and translated Cambodian inscriptions,¹ almost every legible text of Old Khmer which has so far come to light has been readily available in transcribed form to the student of the Khmer language. It is a humble and restricted literature which has come down to us, as the praises of gods and kings and most of the important edicts were written in Sanskrit while the native tongue was found useful chiefly for recording (and ensuring that everyone understood) such matters as the duties of temple slaves, the revenues due to the foundations, and the pious deeds of the local gentry. The language lacks uniformity because the field covered by the inscriptions is both geographically and temporally extensive, spreading over the whole area of Cambodia and spanning the seventh to the fourteenth centuries.

The orthography used for Old Khmer came from outside and did not fit at all perfectly. The Sanskrit consonant-system provided sufficient symbols for the Khmer language but the implication in the consonant-symbol of an 'inherent' vowel to be pronounced after it (except when another consonant or a vowel-symbol or a 'killer' sign was joined to it in writing) was unsuitable for the writing of Khmer consonant clusters. The vowel-system of Sanskrit was inadequate for Khmer and resulted in a certain amount of confusion. It is, for example, difficult to know whether a subscript ơ or ơ is to be regarded as representing the first sound in a diphthong or a second consonant.

Both the inscriptions and the language in which they were written are usually classed as either 'pre-Angkorian' or 'Angkorian'. This convenient chronological and historic division is supported linguistically by the marked different in the vowel-system of the Angkorian dialect as compared with the language of all the pre-Angkorian inscriptions.

The aim of this study is to produce a system or systems based as far as possible on the internal evidence of the texts rather than on deductions drawn from a diachronic approach. The two dialects are discussed together because the same system of initial consonants and final consonants and the same type of syllable-structure operated in both. The different vowel-systems constitute the only fundamental distinction between them.

Simple initial consonants. Pre-Angkorian and Angkorian dialects

The initial consonant system may be presented as follows²:

² The distinction between phonological postulates and the transliterated characters (for which M. Cœdès' system is followed) is maintained throughout by using bold type for the former and italics for the latter.
<table>
<thead>
<tr>
<th>Phonology</th>
<th>Orthography</th>
</tr>
</thead>
<tbody>
<tr>
<td>Velar</td>
<td>k g ń</td>
</tr>
<tr>
<td>Palatal</td>
<td>c j ŋ</td>
</tr>
<tr>
<td>Dental</td>
<td>t d n</td>
</tr>
<tr>
<td>Labial</td>
<td>p b m</td>
</tr>
<tr>
<td>Semivowel</td>
<td>y r l v</td>
</tr>
<tr>
<td>Spirant</td>
<td>s h</td>
</tr>
<tr>
<td>Glottal stop</td>
<td>?</td>
</tr>
</tbody>
</table>

Orthographically two groups of consonants occur in the spelling of native Khmer words which are not represented in the above tables. They are the aspirated occlusives and the retroflex consonants. They have been rejected for different reasons as follows:

1. Aspirated occlusives. These do occur as initials in the writing of Khmer words but they are not regarded as representing simple initial consonants. They are used in two ways:

   (a) to represent a 2-place initial sequence,¹ *kh*, *gh*, *ch*, etc. The separateness of the occlusive and the aspirate is seen in the infixed forms: e.g. *p-an-hem*, *d-an-hum*.²

   (b) to represent in the Angkorian dialect aspiration occurring at the junction between certain 2-place initials, usually an unvoiced occlusive followed by a liquid or nasal: e.g. *khūm*, *thvān*, *phūi*.

   The pre-Angkorian *kūm*, *tloŋ*, and *plū* indicate that 2-place sequences, *kū*, *tl*, *pl*, were here involved, not 3-place sequences (*k + h + ŋ*, etc.)

2. Retroflex consonants. There are a few occurrences, in words which have a Khmer-like form, of retroflex consonant symbols, e.g. *ṭeh*, *ṇīn*, *kraṇān*.³ The instances are not sufficiently numerous, however, to justify the inclusion of retroflex consonants in the phonological table.

   Few instances of the use of palatal or retroflex sibilants have been noted in the spelling of Khmer words.

   Two items in the phonological table need further comment:

   1. *b* (*v*, *b*) and *v* (*v*, *hv*). Many words which in the modern language are spelt with initial *b*⁴ were written in Old Khmer with *v*:

---

² The hyphen is here used to help in demonstrating the infix. Elsewhere, in transliterated words it is used exactly as in Cœdès, *Inscriptions du Cambodge*; in the phonology it is used to indicate disyllables.
³ The word *moc* is in the Angkorian dialect often spelt *moc*, but this is an alternative for the spelling with dental *n*.
⁴ The fact that the modern pronunciation of this character is unvoiced does not seem relevant here; the important point is that modern *b* and *v* represent distinct initial consonants.
The structure of the word in Old Khmer

e.g. OK vave MK babe¹ goat
     valvel babil taperholder

There are on the other hand words which have v in both Old and Modern Khmer:

OK vā (male slave title) MK vā he, she, it, they
     (derogatory)
     svāy svāy mango
     lvā lvā fig

The decision to assume that two initial consonants, b and v, operated in Old Khmer (and not just b or v, represented by v) rests in the first instance on some slight evidence in Old Khmer orthography. Thus:

(a) the character b occurs as an initial in native Khmer words very occasionally; e.g. in one of the pre-Angkorian inscriptions of Hīn K’őn, K. 388, the word ber ‘two’ is written several times with initial b² (cf. modern bīr ‘two’).

(b) the digraph hv was used fairly frequently as an alternative spelling for v in some words which in the modern language have v.³

<table>
<thead>
<tr>
<th>e.g. OK</th>
<th>rval or rhval</th>
<th>MK</th>
<th>raval</th>
<th>busy</th>
</tr>
</thead>
<tbody>
<tr>
<td>veṅ or hveṅ</td>
<td></td>
<td>veṅ</td>
<td>long</td>
<td></td>
</tr>
<tr>
<td>cf. also hvek</td>
<td></td>
<td>vek</td>
<td>ladle</td>
<td></td>
</tr>
</tbody>
</table>

It would seem reasonable to suppose that the purpose of such a digraph was to distinguish two labial consonants. The modern language, in which v (OK v and hv) and b (OK b and v) represent two phonologically distinct initials, lends support to this assumption.

(2) Glottal stop. In many words the initial consonant character is followed by an initial-vowel symbol, written either below it in the conjunct consonant position or next along the line: e.g. ⁴ ṛaēt and ṛeēt ta-et

(alternative spellings of a slave name most probably), saap ‘to hate’, paem ‘sweet’, and the slave names kaol, laałc, saañ, saañt, etc.

It seems that an initial consonant is involved here, part of an initial sequence. It is proposed to assume that this was a glottal stop⁵ and that the initial-vowel symbol was used to represent glottal stop + vowel. The above-mentioned words would therefore be transcribed tʰet,¹ sʰap, pʰem kʰol, etc. The occurrence of the glottal stop can be the only differentiating feature between two words, as in tem ‘tree’ and tʰem (slave name).

¹ The same system of transliteration is used for Modern Khmer as for Old Khmer with the additions of the phonetic symbols, e, v and u used to represent the vowels
₂ The numeral occurs on many other pre-Angkorian inscriptions but is usually not written in word-form. In Angkorian Khmer the word-form is yvar.
³ The digraph hv is also used to represent a final consonant in Old Khmer which is, in Modern Khmer, ṣ; e.g. OK jḥv, MK jāv, ‘to barter’.
⁴ The modern script is used here to facilitate printing.
⁵ As in modern Khmer.
In a word in which the first character is an initial-vowel symbol, e.g. Qualifier 1 oxy, the glottal stop is regarded as operating as a simple initial (phonological transcription, ʔoıy).

Vowels
Here there are two separate systems to be considered.

(1) Vowel system of the pre-Angkorian dialect. The following vowels of the Sanskrit syllabary occur in the writing of native Khmer words:\footnote{\textsuperscript{2}}

\begin{align*}
a & \mapsto \text{ā} \\
i & \mapsto \text{i} \\
u & \mapsto \text{ū} \\
e & \mapsto \text{ai} \\
o & \mapsto \text{au}
\end{align*}

Of these vowel-symbols the following are not held to represent units of the Khmer system (either pre-Angkorian or Angkorian) and are therefore left out of the final table given below.

\textit{i}: This vowel-symbol is used in the writing of the following words: \textit{gi}, \textit{ci}, \textit{ni}, \textit{mi}, \textit{ri}, \textit{si}, \textit{ti}, \textit{gi}; \textit{kamvīc}, \textit{līnī}, \textit{chīnī}; \textit{damrīnī} (for \textit{admīnī}).

These occurrences are few enough to make one suspect, even without further examination, that the symbol \textit{i} does not represent a unit of the vowel-system. One is not surprised to find that for most of the words with zero final alternative spellings with short \textit{i} occur with equal frequency (\textit{gi}, \textit{ci}, \textit{ni}, \textit{mi}, \textit{si} have been noted). This would suggest that the vowel-length which was heard in these syllables and recorded in the writing was in fact a prosodic feature, occurring with zero final.

The final consonant may be held responsible for the use of the long vowel-symbol in the case of the three words \textit{kamvīc}, \textit{līnī}, and \textit{chīnī}. The utterance of a palatal consonant after a close front vowel might well be heard as a prolongation of the vowel.

\textit{ú}: Again, most instances occur in open syllables in words for which an alternative spelling, given in parentheses, has been noted: \textit{kū} (\textit{kū}), \textit{gū}, \textit{cū} (\textit{cu}), \textit{nū}, \textit{nū} (\textit{nu}), \textit{pū} (\textit{pu}), \textit{rū} (\textit{ru}), \textit{kū}, \textit{tamrū}, \textit{sū} (\textit{sru}), \textit{laddū}; \textit{cānkūy}, \textit{tpūr} (for \textit{tpur ?}); \textit{kañjū} (for \textit{kañje}, an isolated instance). \textit{ū} in \textit{cānkūy} remains unexplained.

\textit{ai} and \textit{au}: These vowels do not operate with final consonants. The exceptions to this statement are in words for which an alternative spelling either without the final or with a different vowel is more usual. Thus \textit{dai} and \textit{kantai} are alternatives of \textit{daın} and \textit{duın}, \textit{jauın}, and \textit{jauhın} are written in place of the more common \textit{dōn}, \textit{jon}, and \textit{jaın}. \textit{ai} and \textit{au} are therefore held to represent the vowel \textit{a} with final \textit{y} and \textit{v} respectively (\textit{ay} and \textit{av}).

The other Sanskrit vowel-signs are regarded as representing units of the Khmer system:

\footnote{\textsuperscript{2} For the phonological interpretation of both \textit{taet} and \textit{ta-et} as \textit{the}, see p.10.}

\textit{r} and \textit{l} do occur occasionally in Pre-Angkorian and Angkorian Khmer, e.g. \textit{tr} (MK \textit{tri} 'fish') but the examples are too few to need discussion here.