OBSERVATIONS ON THE SURIN DIALECT OF KHMER

PHILIP N. JENNER

0. INTRODUCTION

It is sometimes forgotten that the southern third of Northeast Thailand, a generally arid and economically depressed territory, falls within the zone of distribution of Khmer speech. One of the more fertile parts of this region, otherwise known as the Khorat Plateau, is the basin of the Mae Nam Mun, which rises in the Dong Phraya Yen chain west of Nakhon Ratchasima (Khorat) and flows east past Ubon to join the Mekong some 40 kilometres above Paksé. In its progress the river crosses Buriram, Surin, Sisaket and Ubon provinces, all bounded on the south by the Dangrek escarpment and northern Cambodia. Most of the inhabitants of the relatively densely populated lands watered by the Mun speak Khmer as their mother-tongue. Official figures are wanting for the total number of Khmer-speakers in the four provinces, and estimates range from an ultra-conservative 200,000 to a probably excessive 500,000.

Of movements and intercommunications between the Mun valley and the Cambodian hearland to the south we have little except inferential knowledge. It is tempting to conjecture that they could never have been important. During much of the Angkorian period the whole Khorat Plateau along with much of Central Thailand was under Khmer suzerainty. This circumstance cannot be used to prove that Khmer speech had a similar extension, for in reality the settlement of the Khorat Plateau is not known in any detail. There are nevertheless good archeological grounds for taking the confluence of the Mae Nam Mun with the Mekong as the earliest identifiable centre of Khmer power - providing Khmer was not also the common language of Fu-nan.2 However this may be, it is reasonable to suppose that Khmer-speaking rice-growers have been in occupation
of the lower Mun valley for a millennium or more. Control over most of
the Khorat Plateau was wrested from Angkor by Rāmādhhipati of Sukhotal
by the year 1350. In the wars culminating in the abandonment of Angkor
(1431-3) the territories on both sides of the Dangrek were repeatedly
ravaged and depopulated. Although it is not known how they were af-
fected, it is unlikely that the humble Khmer-speakers of the Mun came
through these events unscathed. At least it can be pointed out that
the usual direction of the mass deportations practised at the time was
into rather than away from the sparsely settled Northeast. It may even
be supposed that from the 15th century on the earlier Khmer population
was increased by important accessions of transportees from the Tonlé
Sap basin. In 1794 Thailand formally annexed the old Cambodian pro-
vinces of Battambang, Angkor, Mongkolborey, Sisophon and Khorat. While
this whole territory was under Thai administration there were presum-
ably ample opportunities for contacts and migrations between the Mun
valley and central Cambodia, notwithstanding the difficulty of communica-
tions across the Dangrek chain. Since the mid-15th century, however,
there has been no sustained motive for major contacts or movements, and
the historical connection of the Khmer dialects of the Mun basin with
modern standard Khmer has yet to be determined.

The town (/myyen/) of Surin, 3 seat of the province of the same name,
lies 54 kilometers south of the Mun, the same distance due north of the
Cambodian frontier, 42 kilometers by rail east of Buriram, and about
95 kilometers by rail west of Sisaket. During the spring of 1967 it was
my good fortune to meet on the University of Hawaii campus a young Thai
student, Miss Aree Somthawin, who had been born and raised in Surin, who
spoke Khmer as her first language, and who accepted with much good
grace my request for her services as an informant. In the following
paragraphs I sketch out the results of our collaboration over a period
of some four months, believing that the peripheral position of the Surin
dialect may justify the presentation of such tentative findings to
others in the field of Khmer or Mon-Khmer studies.

The Surin dialect is an unwritten vernacular which has for long, ap-
parently, followed a line of development independent of the Cambodian
mainstream.4 The mutual intelligibility of Surin Khmer and the standard
of Cambodia, which I had the opportunity to test, is limited. Its main
areas of divergence from standard Khmer are intonation, the vowel system,
and lexicon. Register5 appears to be entirely absent; whereas standard
Khmer has a maximum of 31 contrasting syllable nuclei, Surin Khmer has
only 23. The circumflex clause terminal seems to be characteristic.
The vocabulary contains a good many archaisms (/biel/ 'near'), region-
alisms (/kmaat/ 'first person singular pronoun, masculine'), and loans
from Thai (/talaat/ 'market').

1. SEGMENTAL PHONEMES

1.1. VOWELS

There are 9 simple (long) vowel phonemes, namely /iː, eː, eː; yː, aː, aː; uː, oː, oː/. These combine with shortness to yield 8 (or 9) shortened counterparts, namely /i, e, y, a, u, o, oː/; functional contrast between /ee/ and /e/ cannot be demonstrated by my data, and is only tentatively assumed for the sake of symmetry. The 3 simple (/i, y, u/) and 3 shortened (/i, y, u/) high vowels combine with /a/ to form 6 falling diphthongs. The total vowel inventory is shown in Table I. Note the absence of /aa, aː, eː, oː, oːe, and of /aːː, aːː, aːː/.

The phonetic quality of these syllable nuclei is not essentially different from that of standard Khmer and will not be detailed here. It may be worth mentioning, however, that as in the standard the phonetic length of the long nuclei is perceptibly decreased by voiceless finals while that of the shortened nuclei is increased by voiced finals. Again as in the standard, the shortened vowels tend to be lower than their long counterparts.

1.2. CONSONANTS

There are 17 consonant phonemes, /p, b, m; t, d, n; c, n; k, ɲ; q; w, j; r, l; s, h/, all of which occur initially. /b, d/ are preglottalised (implosive); /q/ represents [ʔ], often theoretical in word-initial position; /w, j/ are the labial and palatal semivowels respectively; /r/ is a voiced lingual flap (occasionally a trill) with alveolar contact in syllable-initial position.

<table>
<thead>
<tr>
<th>Front Unrounded</th>
<th>Central Unrounded</th>
<th>Back Rounded</th>
</tr>
</thead>
<tbody>
<tr>
<td>i</td>
<td>iː</td>
<td>y</td>
</tr>
<tr>
<td>ɨ</td>
<td>ɨː</td>
<td>ɨː</td>
</tr>
<tr>
<td>(e)</td>
<td>eː</td>
<td>e</td>
</tr>
<tr>
<td>e</td>
<td>eː</td>
<td>a</td>
</tr>
</tbody>
</table>

Table 1: The Syllable Nuclei
Table 2, which probably falls short of representing all possible combinations, shows the initial consonant clusters occurring in my data.

<table>
<thead>
<tr>
<th>initials</th>
<th>p</th>
<th>b</th>
<th>m</th>
<th>t</th>
<th>n</th>
<th>c</th>
<th>k</th>
<th>w</th>
<th>r</th>
<th>l</th>
<th>s</th>
</tr>
</thead>
<tbody>
<tr>
<td>p</td>
<td>x</td>
<td></td>
<td>x</td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>b</td>
<td></td>
<td>x</td>
<td>x</td>
<td></td>
<td>x</td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>m</td>
<td></td>
<td></td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>t</td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>d</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>n</td>
<td></td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td>x</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>c</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>p</td>
<td></td>
<td>x</td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>k</td>
<td></td>
<td></td>
<td>x</td>
<td>x</td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>q</td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>q</td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>w</td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>j</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>r</td>
<td></td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td>x</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>l</td>
<td></td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>s</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>h</td>
<td>x</td>
<td></td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td>x</td>
</tr>
</tbody>
</table>

Table 2: Complex Initials
The most common initial elements are /k, p, s, c, t/ while the most common post-initials are /r, l, h, m, n/, in that order. A few ternary clusters with prefix /m-/ 'one' occur: /mkhēr/ 'one month'. As in the standard, transition from the initial to the post-initial may be direct or indirect, i.e. over a phonetic bridge-sound. Direct transition is normal between an initial stop and post-initial /r, h/. Bridge-sounds include a short voiced vowel, normally [ə] but not seldom [i, ɔ]; a short voiceless vowel, here represented by [h]; and aspiration + [ə]. Before post-initial /b, d, q/ transition is indirect, nearly always by way of [ə]. After initial /s/ it is likewise over [ə] in deliberate speech but becomes direct in normal speech. Between stops other than /b, d, q/ aspiration + [ə] marks deliberate speech while normal speech shows only [h]. Before post-initial nasals and /w, l/ transition is indirect, with [h] being most common while [ə], with or without prior aspiration, is heard in deliberate speech. These phonetic details are exemplified by /ppuək/ [p̥p̥oək] 'cloud', /tbaŋ/ [t̥əoə^n̥] 'to weave', /cmas/ [č̥maː] 'cat', /knɔŋ/ [k̥nɔŋ] 'room', /pkuə/ [p̥k̥uə] 'thunder', /tnɔt/ [t̥nəoʔ] 'sugar palm'.

Final consonants, always simple, include all of the initials except /b, d, s/. Final stops are normally unreleased; as a result, contrast between /-k/ and /-q/ is so weak as to be usually lost: /srok/ 'village' is indifferently [srok - sroʔ] while /pluəq/ 'to taste' is indifferently [p̥luəq - p̥luək]. Final /l/, like its initial counterpart, is a voiced post-alveolar lateral; final /r/, unlike its initial counterpart, is a voiced retroflex lateral with frictionless [r] colouring of the prior vowel: /tliər/ [tɭi:əɾ] 'duck', /kemmoor/ [kəm'mɔ:ɾ] 'lime', /skoor/ [ʃkoːɾ] 'sugar'.

Transition from the syllable nuclei to the finals is direct or indirect, i.e. effected by means of glides. The latter include (a) labial, palatal, and neutral glides and (b) voiced homorganic occlusion before nasal finals. Labial glides are heard mainly after back rounded nuclei, their frequency decreasing with the openness of the nucleus: /ruup/ [ɾu̠p] 'image', /poot/ [p̥uːt] 'corn, maize', /tsoːm/ [t̥soːm] 'to be big'. Palatal glides are heard after front unrounded nuclei before velar finals, as in /peek/ [p̥eɛk] 'too much' and /pleen/ [p̥leɛn] 'music', and above all before palatal finals: /sac/ [saːʔ] 'meat', /qoc/ [ʔɔʔ] 'to kindle', /kheen/ [kʰeɛn] 'to see', /peeŋ/ [p̥eŋ] 'to be full'. Neutral glides occur most often before voiced finals: /kroːm/ [kɾo̠ːm] 'underside', /pui/ [p̥ui] 'poison'. They are also occasionally heard before voiceless finals, where they signal the length of the preceding nucleus: /croot/ [t̥ɾo̠ːt] 'to reap'. In the speech of my informant voiced homorganic occlusion was heard only before final /ŋ/:
/γυγή [jw·γη] 'first person plural pronoun', /ροοη [ro·γη] 'hall, building'. Direct transition from the nucleus to the final is usual in the environments not specified above: /κικ/ [κι·κ] 'to dig', /λεπ/ [λε·π] 'to swallow', /βσή [βση] 'to get', /χοοτ/ [χο·τ] 'to be stupid'.

Before a pause, particularly in citation forms, the lengthening of final /m, n, η, 1/ (though not of /-η/) is common following long as well as short nuclei: /κετ/ [κε·τ] 'onion', /μεν/ [μενη] 'fowl', /οη/ [οη] 'tip, end', /κβαλ/ [κβαλη] 'head', /κεεν/ [κεενη] 'much', /ρηγ/ [ρη·γ] 'to be dried up', /κομελ/ [κομελη] 'salt'.

2. SUPRASEGMENTALS

2.1. STRESS

There are 3 degrees of stress: primary or strong, secondary or moderate, and tertiary or weak. Primary and secondary stress are represented by the acute and grave respectively while tertiary stress is unmarked. As in the standard, monosyllables uttered in isolation take primary stress: /βάν [βάν] 'do it', /στόυ [στόυ] 'It's easy'; disyllables of native provenance take primary stress on the ultima: /καναερ/ 'ladder', /λαμβάκ/ 'trouble'. The same pattern obtains in disyllabic compounds: /κον κόν / 'children', /πλέω τνό / 'road, highway'. Polysyllables take primary stress on the ultima, secondary stress on the first syllable: /τόρεςάπ/ 'telephone', /μπέσωκολ/ 'engineer'. Within the phrase qualifying elements take primary stress, qualified elements secondary or tertiary stress: /τοκάι μην/ 'this day = today', /μπέ νήν/ 'last night', /νέκ πρεφάν/ 'Mr Prachan', /πλεθ κεν 1οκ/ 'your house', /βαβ τύκ ινιακ μην/ 'this evening's supper', /τεω νάσα / 'to go all over'.

2.2. INTONATION

My informant's speech showed 3 pitch levels, namely low, mid (normal), and high, hereafter number 1, 2, and 3 respectively. Four clause terminals stand out clearly: a rising pitch contour, /↑/, marking a question in the absence of an interrogative word; a sustained pitch, /→/, marking a momentary suspension of an utterance; a circumflex or rising-falling contour, /~↑/, marking questions with the interrogative /νε/ and the enclitic interrogative /κε/; and a falling or trailing contour, /↓/, marking the completion of a declarative utterance. These may be represented graphically as follows:
3. STRUCTURE OF THE WORD

As in the standard, the word may be defined in terms of the syllable. The latter may be expressed as (C)CV(VF), wherein F stands for a non-obligatory consonant final. Monosyllables have the shape CV(VF), exemplified by /iil/ 'place', /toq/ 'table', /bep/ 'manner'. Subdi syllables have the shape CCV(VF) and may be interpreted as phonemically monosyllabic but for the most part disyllabic on the phonetic level: /sdam/ [sa?dam] 'right (side)', /rmye/ [nomaw] 'itoh', /ppal/ [phniaw] 'visitor'. Dismyllables of native origin are either (1) compounds of monosyllables or subdi syllables or (2) derivatives by affixation, these last consisting of a monosyllabic or subdi syllabic main syllable together with an unstressed presyllable exhibiting one or the other of two structures: CvN- and Crv-. C in these cases represents any consonant except /m, n, p, q, v, j, h/ while the lower-case v represents a phonemically short, neutral vowel; N represents /m, n, q/ (in /n/ is here included in /n/) while the lower-case r is /r/.

Illustrating these various presyllables are /prenaw/ 'servant', /bentaw/ 'to blame', /prityw/ 'to meet', /timpew/ 'to be bald', /denrav/ 'elephant', /trepew/ 'swamp', /cenkaw/ 'chopsticks', /cemow/ 'nose', /kemmar/ 'lime', /krebav/ 'carabao', /qennaw/ 'well (puits)', /rantiw/ 'lightning', /lambaak/ 'trouble', /sennaw/ 'nest', /sremooc/ 'ant'. Paralleling the rotocalised presyllables the Surin dialect has at least one instance of a /mra/-, not found in the standard: /mratw/ 'chili pepper', corresponding to mdesa /mte/ [m"teh].

It is worth noting that the presyllable /qaw-/ was consistently pronounced by my informant either as a nasal with a faint vocalic onglide or as a syllabic nasal: /qempaw/ [\^m'praw - \^m'praw] 'sugar cane', /qencul/ [\^n'ko - \^n'ko] 'needle'. Of special interest is the dentalisation of the nasal finals of CvN- presyllables before main-syllable initial /c, r, i/: /berian/ 'to teach' (standard /bgrfien/), /conriw/ 'to salute, greet' (standard /cumriw/), /kentaw/ 'strength' (standard
4. SAMPLE TEXT

The following reproduces part of a breakfast-time dialogue, written and recorded by my informant, between two sisters the elder of whom manages the house, the younger being a teacher.

A 2ŋaaŋ níh bocŋ 1n 2thè 3qëj 1kliŋ +

day this elder-sibling future-marker do what pluraliser
What are you going to do today?

B 2thè 3qëj-qëj jaŋ 3ná→ 1mjáŋ 2bocŋ 1kco 2tëŋ
do what-what kind which one-kind elder-sibling then must
tëw 2taláat 3tëŋ 1pryk níh +
go-to market time morning this

Among other things, I have to go to market this morning.

A 2bocŋ 1n 2tàw 2talâat 3ná 2 tàew talâat
elder-sibling future-marker go-to market which go-to market
myet 3stìŋ + 2ryy talâat 3tëŋ +
edge river or market public

Which market are you going to, the one down by the river or the main one?

B 2ɗëŋ jaŋ 3ná 2tëw

know way which question-marker
How am I to know? (= I'm not sure.)
A

1 jyyo 2 kwah 3 qeej 1 klah +
we lack what pluraliser
What are we out of?

2 knom kyt thaa mee 1 kyet 2 tip qeej kempeh kempan 3 maw
I think say mother she buy what all sorts of some

1 heej teew mset 2 men +
completion-marker go-to yesterday
I thought Mama already bought plenty of things yesterday.

B

2 tip 1 qenliawk 2 kdo 1 non 2 sac somrap niisch nih 3 tlet +
buy greens and with meat use-for evening this more
I want to get some more vegetables and some meat for tonight.

1 bentli 2 qeej kyet qeac 1 neq 2 nyem 3 pnew 1 maw 2 ptleth +
maybe father he able linker take guest some house
Papa may be bringing a visitor home with him.

A

2 kommoj pliec tip 1 mreth 3 maw 3 nih

2 jup men 3 mee 1 kyet 2 prap thaa 1 wile 2 min miyen 3 liawk 1 qii
night last mother she tell say one not have sell at
talat 2 +
market
Mama said last night they didn't have any at the market.

1 ken 2 jyyo koo hap 1 neq 3 mot 2 heej +
thing our clause-connector near linker gone completion-marker
Ours are just about gone.
E 1bécnoh² → bōng  l¹n  2rûwe  3tîn.  2qoqj  bān +
if-so elder-sibling will seek buy give (= benefactive) get-to
In that case, I can get some for (her, us).

A 2bōng  l¹n  2têw talâet myet  3stîn →  2bōng  1koc
elder-sibling will go-to market edge river elder-sibling then
  2têng  2phān  1ptieh  3lîqûn +
must pass house La'ong

If you're going to the market by the river, you'll have to pass by
La'ong's house.

  2jûwe qensyy  3níh  2têw  qoqj kē phoon  3kso  2jēn nāa +
take book this go-to give her on-way then way what
How about taking this book to her on your way?

B 1min  2qêcêj →  3bān  2hēej +
not what get-to completion-marker
It's nothing (= no trouble). I can do it.

A 2kndm  l¹n  2tûk qensyy  1qii tron myet  2twîlēr  3níh  2hēej →
I will put book at place edge door this completion-marker

  2prûm  1tiiŋ  2cûtmaēj  3phoŋ +
along with note as-well
I'll leave it here by the door, along with a note.

  2kēe soom kēcêj  1môw  3mênēc  2hēej +
she ask borrow come one-month completion-marker
She asked to borrow it a month ago (now),

  2kôcēc  1min  2cînh  bān  3jûwe  2têw  qoqj  1kêe ~
but not know-how get-to take go-to give her
but I haven't been able to get it to her.

  2qee  bōng  3dèq  1qe~thaa  2qëw  1kyet  nōn  2nyâm  nîek
er elder-sibling know? say father he will take person

  3nāa  2môw +
what come

Um-m, who's papa bringing home, do you know?
I hear say person come from Bangkok. I understand (it's) someone from Bangkok.

I don't know what his name is.

If not be lawyer then be person

Well, if he's not a lawyer, I expect he's some kind of technician.

I suppose he person control make road agree?

I guess he (must be) in charge of building the road...

Road go-to Khorat agree? or kind what this completion you know, the road to Khorat - or something like that.

You may be right. We'll find out when he gets (they get) here.

I go all-right? Must fix self go-to hall learn now.

Well, I'm going... I've got to get ready for school now.

I'm going... I've got to get ready for school now.

What time will you be getting home tonight?

How am I to know? (= I'm not sure.)

I must test person study all on morning this I have to give a test to the students this morning.
able linker stay correct work test give finish first
I may stay on and finish correcting their work before coming home.

B 2kommoj 1seew 2nëew ingiec 3nëh nêe ~
don't very stay late very agree?
Don't be very late, will you?

2tëc toon hâq 1mow 2ptîeh 3iqëen 1toon 2nyt ñyt +
else must come-home come house alone in all-dark
Otherwise you'll have to come home all by yourself in the dark.

A 3cåh + 2knom nong pjàm hâq 3mow 2mun pràm mông +
yes I will try come-home come before five hour
All right, I'll do my best to get back before five.

3më [2qoøj ingiec 1ciën 2nuh 3teo ~
not give = let late more that emphasiser
I won't make it (let it be) any later than that.

B 2teew con +
go descend
Run along!

3tëc 2thea bocq thee qoøj niën jûr
else say elder-sibling make give = so-that young-lady late

3kàer +
work
Otherwise I'll make you late for work.

2cië 1qoøj 3nëe ~
be-well be-fine agree?
Bye-bye.
NOTES

1. Khmer also overlaps Cambodia's western frontier into Prachinburi (Krabin), Chanthaburi, and Trat provinces.


3. Surin is shown on AMS L509 (1:250,000) sheet ND 48-5 at grid co-ordinates UB 3946 or 14° 53' N x 103° 29' E. The town claims a population of a little over 147,000.

4. Radio Phnom-Penh is received in Surin but its broadcasts appear to be imperfectly understood. Newspapers, books, and films are not imported from Cambodia. Identification with the Khmer language community appears to be slight.

5. I use the term "register" here in its original Hendersonian sense, which emphasises contrasting oral and pharyngeal resonance and contrasting higher and lower pitch; see Eugénie J.A. Henderson, "The Main Features of Cambodian Pronunciation", in *BSOAS*, XIV (1952):1:149-74.