

DATE: 2002-02-13

DOC TYPE: Expert contribution

TITLE: **Proposal to add missing Khmer characters**

SOURCE: CHEA Sok Huor, LAO Kim Leang, HARADA Shiro, Norbert KLEIN

PROJECT:

STATUS: Proposal

ACTION ID:

DUE DATE: --

DISTRIBUTION: Worldwide

MEDIUM: Paper

NO. OF PAGES:

A. Administrative

1. Title	Proposal to encode missing Khmer characters
2. Requester's name	CHEA Sok Huor, LAO Kim Leang, HARADA Shiro, Norbert KLEIN
3. Requester type	Expert request.
4. Submission date	2002-02-13
5. Requester's reference	ISO/IETC JTC 1/SC 2/WG 2 N2380R
6a. Completion	This is a complete proposal.
6b. More information to be provided?	Only as required for clarification.

B. Technical – General

1a. New script? Name?	No.
1b. Addition of characters to existing block? Name?	Yes. Khmer
2. Number of characters	11
3. Proposed category	

4. Proposed level of implementation and rationale	
5a. Character names included in proposal?	Yes
5b. Character names in accordance with guidelines?	Yes
5c. Character shapes reviewable?	Yes.
6a. Who will provide computerized font?	Paul Nelson, Microsoft
6b. Font currently available?	Yes
6c. Font format?	TrueType.
7a. Are references (to other character sets, dictionaries, descriptive texts, etc.) provided?	
7b. Are published examples (such as samples from newspapers, magazines, or other sources) of use of proposed characters attached?	
8. Does the proposal address other aspects of character data processing?	Character classifications are provided.

C. Technical – Justification

1. Contact with the user community?	Yes. National Body Contribution
2. Information on the user community?	Native.
3a. The context of use for the proposed characters?	
3b. Reference	
4a. Proposed characters in current use?	Yes.
4b. Where?	Cambodia.
5a. Characters should be encoded entirely in BMP?	Yes.
5b. Rationale	Daily use of characters in Cambodia.
6. Should characters be kept in a continuous range?	
7a. Can the characters be considered a presentation form of an existing character or character sequence?	No
7b. Where?	

7c. Reference	
8a. Can any of the characters be considered to be similar (in appearance or function) to an existing character?	
8b. Where?	
8c. Reference	
9a. Combining characters or use of composite sequences included?	
9b. List of composite sequences and their corresponding glyph images provided?	
10. Characters with any special properties such as control function, etc. included?	

D. SC2/WG2 Administrative

To be completed by SC2/WG2

1. Relevant SC 2/WG 2 document numbers:	
2. Status (list of meeting number and corresponding action or disposition)	
3. Additional contact to user communities, liaison organizations etc.	
4. Assigned category and assigned priority/time frame	
Other Comments	

In the process of originally encoding the Khmer script, some commonly used characters were not encoded. It is proposed that the characters listed be added to the current Khmer block to allow modern Khmer documents to be created.

Additional Diacritic Signs

17DD;KHMER SIGN ATTHACAN;Mn;0;NSM;;;;;N;;;;;



The ATTHACAN sign was introduced in the early 1960s to resolve ambiguity in spelling where two words are pronounced differently, but spelled the same. The use of this mark is defined in the Chuon Nath dictionary at Appendix A. The use of this sign never caught on and disappeared from use before the end of the 1960s. However, the presence of this sign in many books printed during the 1960s require the inclusion of this mark into the Khmer script encoding if those books are to be represented in digital format.

Numeric symbols for divination lore

The following numbers are used for divination, but are not used for calculating. The numeric values are 0 – 9 and are indicated in the properties below. The actual use of these marks is very limited. They seem to be only known to people who deal with divination. No documentation, other than what the KPP (Khmer Philological Project) has been able to gather on the shape of these is known.

17F0;KHMER SYMBOL LEK ATTAK SON;No;0;L;;;0;N;;;;;

០

17F1;KHMER SYMBOL LEK ATTAK MUOY;No;0;L;;;1;N;;;;;

១

17F2;KHMER SYMBOL LEK ATTAK PII;No;0;L;;;2;N;;;;;

២

17F3;KHMER SYMBOL LEK ATTAK BEI;No;0;L;;;3;N;;;;;

៣

17F4;KHMER SYMBOL LEK ATTAK BUON;No;0;L;;;4;N;;;;;

៤

17F5;KHMER SYMBOL LEK ATTAK PRAM;No;0;L;;;5;N;;;;;

៥

17F6;KHMER SYMBOL LEK ATTAK PRAM-MUOY;No;0;L;;;6;N;;;;;

៦

17F7;KHMER SYMBOL LEK ATTAK PRAM-PII;No;0;L;;;7;N;;;;;

N

17F8;KHMER SYMBOL LEK ATTAK PRAM-BEI;No;0;L;;;8;N;;;;;

/

17F9;KHMER SYMBOL LEK ATTAK PRAM-BUON;No;0;L;;;9;N;;;;;

W

APPENDIX A – Sample and definition of ATTHAKAN in Chuon Nath dictionary, volume 2.

១១០- ពាក្យពេលជាស្រះ: អ ដែលឥត តួ ។ ប្រកបហៅថា ពាក្យក្រព្រា ដូចជា កិ (កមនុស្សសត្វ), ដី (ដីស្មៅ), ល (លមើល), លី (សម្បទស) ជាដើមនេះ គណៈកម្មការដ៏បួនបានបង្កើតសញ្ញាថ្មី មួយហៅថា អឌ្ឍបន្ទំ មានសណ្ឋានស្រដៀងនឹងចំណិតសាច់ដូង ប៉ុន្តែដាច់ (្គ) នេះប្រើដាក់ពីលើ ដូចជា

ឈឹកិ, លើកដាក់លើកិ, ទូកដី, ថាមមើល, លើបលើ, កុកសិកិ, ដីសិ កុំឲ្យប្រឡងក្នុងការពាន ។