Proposal for the Generation Panel for the Khmer Script Label Generation Ruleset for the Root Zone

1   General information

Khmer script is an Abugida system used to write Khmer language and some other languages spoken in Cambodia. It is also used to write Pali in the Buddhist liturgy of Cambodia and Thailand. Historically, it derives from the Pallava script, a variant of the Grantha alphabet, which descended from the Brahmi script, was used in southern India and South East Asia during the 5th and 6th centuries AD[1].

 Khmer is written from left to right, without space between words, except to mark phrase or sentence boundaries. Originally there were 35 consonant characters, but modern Khmer uses only 33. Consonants are written in clusters, with additional consonants written in reduced form under the main consonant. Each such character represents a consonant sound along with an inherent vowel â or ô.

There are some independent vowel characters, but vowel sounds are more commonly represented as dependent vowels, which are marks accompanying a consonant character, and indicating the vowel sound to be pronounced after that consonant (or consonant cluster). Most dependent vowels have two pronunciations, based on the inherent vowel of the consonant they attach with. There are additional diacritical marks used for further modifying the pronunciation, numerals and punctuation marks[2].

1.1   Script for which the panel is to be established

ISO 15924 – Code: Khmr
ISO 15924 – Number: 355
ISO 15924 - English Name: Khmer
Native Name: ខ្មែរ

Khmer Unicode range: U+1780 – U+17FF
The Khmer Generation Panel will consider the subset of code points within this range which have been shortlisted in Maximal Starting Repertoire version 2 (MSR-2).

1.2   Principal languages using that script

Khmer script is used to write Khmer language, which is the official language of Cambodia. In addition, it is also used to write a few other languages spoken in Cambodia, as listed in the table.

Language	ISO 639-3 Code(s)	Countries	Local Names of the Script
Khmer	khm	Cambodia, Laos, Thailand, Vietnam	ខ្មែរ
Tampuan	tpu	Northeast border area, Central Rattanakiri province, Cambodia	ទំព័ន
Mnong	Cmo	Mondol Kiri province, Cambodia	ក្នោង

---

[1] Source: http://en.wikipedia.org/wiki/Khmer_alphabet.
[2] ibid.

Kuy	kdt	North central Cambodia	គួយ
Kru'ng	krr	RattanaKiri and Mondol Kiri provinces, Cambodia	ក្រៀង
Jarai	jra	RattanaKiri province, Cambodia	ចារាយ

Note: There are few other languages spoken in Cambodia, but with small speaking population. They are not listed in this table, as it is not known to the panel at this time if these are still actively being written and using Khmer script. The panel may look at these languages during the analysis.

1.3   Geographic territories or countries with significant user communities for the script

Khmer script is used widely in Cambodia and Southern Vietnam and by the Cambodian expat community in France, Canada, Australia, USA and other countries.

1.4   Are there related scripts? If so, which ones?

Thai and Lao scripts are derived from Khmer script. Though Khmer script is visually not similar to these scripts, the panel will do a more detailed analysis to determine any similarity.

2   Proposed initial composition of panel

The panel is composed of individuals actively involved in research and development of Khmer language computing and the maintenance of the domain name system in Cambodia. The panel has also invited community members working on Khmer language. The panelist come from government, academia, private sector, ccTLD registry and the community.

2.1   List prospective chair and members

Name	Role	Designation	Organization	Relevant Experience
Dr. Sopheap Seng	Chair	President	National Institute of Posts, Telecoms and ICT (NIPTICT)	- Research Khmer Language Processing
Mr. Rapid SUN	Secretary	Director	Center of Research and Development, NIPTICT	- Supervise Khmer Language Processing project such as Khmer OCR, Machine Translation, Text to Speech, Automatic Speech Recognition
Mr. Daro Chin	Member	Director	Department of ICT, Telecom Cambodia	- Manage nameserver of ccTLD .kh

Mrs. Yatal Lim	Member	Chief	KH-NIC Office, Telecom Regulatory of Cambodia	- Administrator of ccTLD .kh domain name
Mr. Chea Sok Huor	Member	Unicode Expert	Pan Localization Cambodia	- Project Manager of Pan Localization Cambodia
Mr. Makara Than	Member	Deputy Director	Center of Research and Development, NIPTICT	- Assist Khmer Language Processing Project
Mr. Hong Danh	Member	Unicode Expert	Freelance/ Community	- Develop Khmer Unicode Keyboard layout - Develop Khmer Unicode keyboard for android, ios - Research on Khmer OCR
Mr. Ra An	Member	Director	Department of Application and Content, General Department of ICT	- Work on ICT policy and software standard
Mr. Khemara Mok	Member	Director	Department of E-Government, General Department of ICT	- Work on e-Government and all government websites project
Mr. RANGSEY KEN	Member	Professor/ linguist	Royal University of Phnom Penh	- Expert in linguistic and Khmer literature

Note: The panel will invite additional member as required.

3 Work plan

3.1 Suggested Timeline, listing of significant milestones

The generation panel will start the work from June 2015 and intends to finalize the proposal by March 2016. The panel intends to form working groups on sub-topics, which will do the preparation and present details to the complete panel. The full panel will then finalize these details. The full panel will meet fortnightly for this purpose. As the members of the panel are in Phnom Penh, these meetings will be face to face and organized by NIPTICT.

The panel will hold an open public consultation at the end of the work before submitting the proposal to ICANN. The purpose of public consultation is to get feedback on the work from the larger community and experts. Experts will be invited from IT companies, internet service providers, universities and media. The feedback from this consultation will be used to finalize of the proposal for submission.

Activity	Description	Start Date	Duration
Develop Principles	Principles to be used to determine valid code points, variants and labels	10 June 2015	1 month
Determine Code Points	Select the code points from MSR which are needed for Root Zone LGR	10 July 2015	2 months
Determine (any) Variants	From the codes points selected, determine if the end-user may confuse two code points	10 September 2015	1 month
Determine Label Rules	Determine if there are any label level constrains on the use of selected code points	10 November 2015	1 month
Hold Public Consultation	Hold a workshop on the work accomplished by the generation panel to get feedback from the community and experts.	Early December 2015	1 Day
Write Proposal and Create XML	Write up the Root Zone LGR proposal, including references to each code point included, why variants needed and details of label rules developed + XML file	10 December 2015	2 months
Submit	Get Public Comments, Finalize and Submit	10 February 2016	1 month

3.2 Anticipated need for logistic support, such as mailing list?

Though most of the work done by the panel will be discussed in face to face interactions, the panel will use the email list provided by ICANN for sharing administrative information and discussions. The panel outputs will also be updated on its wiki page, provided by ICANN.

The panel anticipates additional support from ICANN for the following:
- Public consultation organized in Cambodia
- Support of expert in the area of DNS/IDN to travel to Cambodia to interact with the Panel
- Travel support to ICANN meeting for panel representative to interact with the community

3.3 Does the panel expect to call on ICANN provided advisors?

The panel anticipates that it will require a DNS/IDN expert to interact with the Panel.