

PROPOSAL FOR SINHALA SCRIPT ROOT ZONE LABEL GENERATION RULES

Publication Date: 23 November 2018

Prepared By: IDN Program, ICANN Org

| Public Comment Proceeding | |
|---------------------------|------------------|
| Open Date: | 2 October 2018 |
| Close Date: | 14 November 2018 |
| Staff Report Due Date: | 28 November 2018 |

| Important Information Links |
|---|
| Announcement |
| Public Comment Proceeding |
| View Comments Submitted |

Staff Contact: Pitinan Kooarmornpatana **Email:** pitinan.koo@icann.org

Section I: General Overview and Next Steps

The Sinhala script community has formed the Sinhala script Generation Panel (GP), which in turn has developed a *Proposal for the Sinhala Script Root Zone Label Generation Rules (LGR)* ([Proposal](#) and [Supporting Documentation](#)). As per the [LGR Procedure](#), this proposal is being posted for public comment to allow those who have not participated in the Sinhala script GP to make their views known. Based on the feedback, the Sinhala script GP will finalize the proposal for submission for integration into the Label Generation Rules for the Root Zone (RZ-LGR).

Section II: Contributors

At the time this report was prepared, a total of three (3) community submissions had been posted to the forum. The contributors, both individuals and organizations/groups, are listed below in chronological order by posting date with initials noted. To the extent that quotations are used in the foregoing narrative (Section III), such citations will reference the contributor's initials.

Organizations and Groups:

| Name | Submitted by | Initials |
|------------------------------------|----------------------|----------|
| Myanmar Script Generation Panel | Thin Zar Phyo | MMGP |
| Neo-Brahmi Script Generation Panel | Udaya Narayana Singh | NBGP |

Individuals:

| Name | Affiliation (if provided) | Initials |
|-----------|---------------------------|----------|
| Liang Hai | | LH |

Section III: Summary of Comments

***General Disclaimer:** This section intends to summarize broadly and comprehensively the comments submitted to this public comment proceeding but does not address every specific position stated by each contributor. The preparer recommends that readers interested in specific aspects of any of the summarized comments, or the full context of others, refer directly to the specific contributions at the link referenced above (View Comments Submitted).*

MMGP reviewed the the Sinhala LGR and makes following comments:

MMGP1. MMGP shares a list of Sinhala-Myanmar confusable code points being considered by MMGP and comments that it is useful for the readers if the both the Malayalam and Myanmar LGR proposals share the same list.

NBGP reviewed the the Sinhala LGR and makes following comments:

NBGP1. NBGP agrees with the Sinhala GP analysis for Devanagari ः (U+0903) and Sinhala ೃ(U+0D83) that they are confusable code points.

NBGP2. NBGP observes that the dependent signs which cannot form any valid labels, e.g. e.g. Gujarati ૃ: (0A83) and Sinhala ೃ: (0D83), were not listed as confusable code points in Gujarati LGR proposal, while Sinhala LGR proposal does. As it is not the normative part, the discrepancy is not an issue.

NBGP3. NBGP strongly suggests that the Telugu-Sinhala and Kannada-Sinhala cross script variant code point, as listed in the Telugu and Kannada LGRs should be included in the normative part of Sinhala LGR.

LH reviewed the Tamil proposal and makes the following comments:

LH1. LH suggests that the following points or sections should be revised.

- (1) In section 3.2, the relationship between Sinhala and Myanmar script should be different from the relationship between Sinhala and Neo-Brahmi scripts
- (2) In section 3.3.1, the text regarding conjunct characters, ೆ (jna), modifiers, special symbols should be reviewed
- (3) In 3.3.4, LH suggests replacing "... represents all the nasals" with "represents a general nasal sound" or "represents a context-dependent nasal sound"
- (4) In section 5.2, Code Point Repertoire, LH suggests:
 - The usage of these f-sound graphemes should be discussed
 - The reason should be given for why U+0DF2 ೆaa SINHALA VOWEL SIGN DIGA GAETTA-PILLA is included but its independent form U+0D8E ೆaa SINHALA LETTER IRUUYANNA is excluded
 - U+0D8E, U+0D8F, U+0D90, U+0DDF, and U+0DF3, should be listed as the standard Sanskrit alphabet even though they are not used for the Sinhala language
 - Stronger case is needed for the exclusion of U+0D9E ೆ SINHALA LETTER KANTAJA NAASIKYAYA and U+0DA6 ೆ SINHALA LETTER SANYAKA JAYANNA
- (5) In section 5.5, the systematic necessity of ZWJ in the Sinhala encoding should be more prominent as the exclusion of ZWJ affects a great number of common words
- (6) In section 7, the scope of work should be clearly mentioned

(7) In section 7, the term “Indic Syllabic Category” can cause confusion with the Unicode character property of the same name. It should be noted that this is not the Unicode property term.

LH2. LH raises the following questions and discussion points.

- (1) In section 3.2, it is unclear whether Sanskrit is included in the analysis scope
- (2) In section 3.3.2, Table 2, the pronunciations of vocalic I and vocalic II is not listed in the table but only the names
- (3) In section 3.3.3, the requirement of ZWJ forming a conjunct label, especially with the special symbols, should be discussed
- (4) In section 3.3.6, LH has asked to note the rationale which constraints Sannjakas to not be followed by a Halanta
- (5) In section 6.1, LH considers the criteria for identifying in-script variants as strict. The Sinhala GP and NBGP should have the same criteria for in-script variants analysis
- (6) In section 6.1, the j ෙ (U+0D95) and ෙ (U+0DB9 U+0DCA) mapping is already disallowed by the akshar formation rule
- (7) In section 7, on Whole Label Evaluation (WLE) rules, it is unclear whether it is necessary to split J from C when the argument for disabling H and X after J is weak. Also the attestation of visarga following a prenasalized stop already exists, according to section 5.6.5, so it should not be restricted in the rule

Section IV: Analysis of Comments

General Disclaimer: This section intends to provide an analysis and evaluation of the comments submitted along with explanations regarding the basis for any recommendations provided within the analysis.

These comments are being submitted to the Sinhala Generation Panel for its consideration and incorporation (as required) in the final version of the proposals.