

## Named Sequences of Latin Letters for the Transliteration of Indic Vocalic Letters:

### Comment on L2/13-204

N. Ganesan ([naa.ganesan@gmail.com](mailto:naa.ganesan@gmail.com))

Sri. Anshuman Pandey, in L2/13-204, requests encoding of eight Latin characters in “Latin Extended-D” block. He notes: “The author understands that the Unicode Technical Committee tends presently towards disapproval of proposals for encoding precomposed characters for Latin, but the matter deserves a formal review as the proposed characters are part of a widely-used, legacy scholarly encoding and an existing ISO standard.”

In case these 8 letters do not get new atomic code points in Latin Extended-D block, these letters can be represented in Named Sequences List as is usually done in Brahmi derived scripts of India (Tamil, Bengali, Gurmukhi, Sinhala, ...) and South East Asia (e.g., Khmer). Gurmukhi named sequences list: <http://www.unicode.org/Public/6.3.0/ucd/NamedSequencesProv.txt>

<http://www.unicode.org/Public/6.3.0/ucd/NamedSequences.txt>

TAMIL CONSONANT K; 0B95 0BCD etc.,  
 BENGALI LETTER KHINYA;0995 09CD 09B7  
 SINHALA CONSONANT SIGN YANSAYA;0DCA 200D 0DBA etc.,  
 ARABIC SEQUENCE YEH WITH HAMZA ABOVE WITH ALEF;0626 0627 etc.,

NAMED SEQUENCES LIST example for transliteration of Indic Vocalic Letters:

LATIN SMALL LETTER R WITH RING BELOW AND MACRON  $\bar{r}$   
 <r U+0072 LATIN SMALL LETTER R, ◌ U+0325 COMBINING RING BELOW, ◌ U+0304  
 COMBINING MACRON>

A section with the title, # Additions for Lithuanian. Provisional 2006-05-18, Approved 2007-10-19 is seen,

<http://www.unicode.org/Public/6.3.0/ucd/NamedSequences.txt>

# Additions for Lithuanian. Provisional 2006-05-18, Approved 2007-10-19

LATIN CAPITAL LETTER A WITH OGONEK AND ACUTE;0104 0301  
 LATIN SMALL LETTER A WITH OGONEK AND ACUTE;0105 0301 etc.,

A similar section can be added to NamedSequences.txt with the title, # Additions for Indic. Once these 8 letters are in Named Sequences List, Notes on Usage can be added as “Indic Vocalic Letters” as in third column, Table 2. Examples of Named Character Sequences in:

<http://www.unicode.org/reports/tr34/>

**Summary:** If new atomic code points are not approved, Indic Vocalic Letters can be represented by a Section in NamedSequences.txt. The Unicode code points for the Named Sequences, if approved, can be taken from (see Table 1 in the next page):

<http://homepage.ntlworld.com/stone-catend/trinoun1.gif>

1.	Latin basic vowels with diacritics <b>prescribed</b> in ISO 15919	U+
Ä	LATIN CAPITAL LETTER A WITH MACRON AND TILDE	0100 + 0303
ä	LATIN SMALL LETTER A WITH MACRON AND TILDE	0101 + 0303
Ï	LATIN CAPITAL LETTER I WITH MACRON AND TILDE	012A + 0303
ï	LATIN SMALL LETTER I WITH MACRON AND TILDE	012B + 0303
Û	LATIN CAPITAL LETTER U WITH MACRON AND TILDE	016A + 0303
ü	LATIN SMALL LETTER U WITH MACRON AND TILDE	016B + 0303
Ŕ	LATIN CAPITAL LETTER R WITH RING BELOW	0052 + 0325
ŗ	LATIN SMALL LETTER R WITH RING BELOW	0072 + 0325
Ŗ	LATIN CAPITAL LETTER R WITH RING BELOW AND MACRON	0052 + 0325 + 0304
ŗ̄	LATIN SMALL LETTER R WITH RING BELOW AND MACRON	0072 + 0325 + 0304
Ł	LATIN CAPITAL LETTER L WITH RING BELOW	004C + 0325
ł	LATIN SMALL LETTER L WITH RING BELOW	006C + 0325
Ł̄	LATIN CAPITAL LETTER L WITH RING BELOW AND MACRON	004C + 0325 + 0304
ł̄	LATIN SMALL LETTER L WITH RING BELOW AND MACRON	006C + 0325 + 0304
Ë	LATIN CAPITAL LETTER E WITH MACRON AND TILDE	0112 + 0303
ë	LATIN SMALL LETTER E WITH MACRON AND TILDE	0113 + 0303
Ö	LATIN CAPITAL LETTER O WITH MACRON AND TILDE	014C + 0303
ö	LATIN SMALL LETTER O WITH MACRON AND TILDE	014D + 0303

Table 1: Code Points for Named Sequences for Indic Vocalic letters (Source: ISO 15919)  
<http://homepage.ntlworld.com/stone-catend/trinoun1.gif>

Reference: Anshuman Pandey, Proposal to Encode Latin Letters for the Transliteration of Indic Vocalic Letters, L2/13-204.