ISO/IEC JTC1/SC2/WG2 N3649

L2/09-192 2009-05-04

Universal Multiple-Octet Coded Character Set International Organization for Standardization Organisation Internationale de Normalisation Международная организация по стандартизации

Doc Type: Working Group Document

Title: Preliminary proposal for encoding the Ranjana script in the SMP of the UCS

Source: Michael Everson

Status: Individual Contribution

Action: For consideration by JTC1/SC2/WG2 and UTC

Date: 2009-05-04

Varieties of the Rañjana script (called Lañdza or Lañtsa in Tibetan) have been used to writing Sanskrit, Tibetan, and Newari in an area ranging from Nepal to Tibet. The script is a descendent of the ancient Brahmic script from India, and has the same basic behaviours (inherent vowel, vowel-sign modifiers, consonant conjunct formation). There are very many texts in Rañjana. It is mostly used for printing Hindu and Buddhist scriptures and literature in Sanskrit and Buddhist Hybrid Sanskrit used by the Newar community. Tarthang Tulku in America recently published an edition of the entire 8,000 verse Prajñāpāramitā in parallel Rañjana and Tibetan scripts. The Rañjana script continues to be taught in Tibetan and Bhutanese monasteries—and books on how to write the script continue to be published.

Rañjana is also in current use for printing "high status" documents (wedding invitations, certificates, etc) in Nepal in the Newar language and for Newar language book titles.

Since Rañjana is visually and structurally similar to the Lañtsa & Wartu scripts used for Buddhist Sanskrit documents in Tibet (China), Bhutan, Mongolia, Nepal, Sikkim & Ladakh (India) it has been considered would be practical to merge these two scripts (Lañtsa and Wartu) with Rañjana for encoding purposes.

- **2. Structure.** Rañjana is of the Brahmic type: the consonant letters contain an inherent vowel. Consonant clusters are written with conjunct characters, and a visible vowel killer shows the deletion of the inherent vowel when there is no conjunct.
- 3. Independent vowels. The full set of Sanskrit independent vowels exist, though not all of these are used when Rañjana is used to write Newar: $\mathbb{Z}[a,\mathbb{Z}[a]]$ $a,\mathbb{Z}[a]$ $a,\mathbb{Z}[a]$

5. Consonant clusters are used in a manner similar to that employed in the Devanagari script. There are many excellent manuals available which show the conjunct glyphs deriving from these combinations, and any font design should for a consistent style base itself on one of them. Of particular interest are a number of voiceless consonants which are represented by consonant clusters. These are listed below. It is expected that named-sequences for these should be added to the standard to give guidance to users in Nepal.

- **6. Collating order.** The standard Brahmic sorting order applies to Ranjana.
- 7. Character names. Consonant letter names use the typical Brahmic transliteration used in the UCS.
- **9. Linebreaking.** Letters and digits behave as in Devanagari.
- **8. Digits and punctuation.** A set of digits exists. Alongside a single and double danda punctuation, Rañjana has many other signs.

10. Unicode Character Properties.

```
11800; RANJANA LETTER A; Lo; 0; L;;;;; N;;;;;
11801; RANJANA LETTER AA; Lo; 0; L;;;;; N;;;;;
11802; RANJANA LETTER I; Lo; 0; L;;;;; N;;;;
11803; RANJANA LETTER II; Lo; 0; L;;;;; N;;;;;
11804; RANJANA LETTER U; Lo; 0; L;;;;; N;;;;
11805; RANJANA LETTER UU; Lo; 0; L;;;;; N;;;;
11806; RANJANA LETTER VOCALIC R; Lo; 0; L;;;;; N;;;;;
11807; RANJANA LETTER VOCALIC RR; Lo; 0; L;;;;; N;;;;
11808; RANJANA LETTER VOCALIC L; Lo; 0; L;;;;; N;;;;
11809; RANJANA LETTER VOCALIC LL; Lo; 0; L;;;;; N;;;;;
1180A; RANJANA LETTER E; Lo; 0; L;;;;; N;;;;;
1180B; RANJANA LETTER AI; Lo; 0; L;;;;; N;;;;
1180C; RANJANA LETTER O; Lo; 0; L;;;;; N;;;;
1180D; RANJANA LETTER AU; Lo; 0; L;;;;; N;;;;;
1180E; RANJANA LETTER KA; Lo; 0; L;;;;; N;;;;
1180F; RANJANA LETTER KHA; Lo; 0; L;;;;; N;;;;;
11810; RANJANA LETTER GA; Lo; 0; L;;;;; N;;;;
11811; RANJANA LETTER GHA; Lo; 0; L;;;;; N;;;;;
11812; RANJANA LETTER NGA; Lo; 0; L;;;;; N;;;;;
11813; RANJANA LETTER CA; Lo; 0; L;;;;; N;;;;;
```

```
11814; RANJANA LETTER CHA; Lo; 0; L;;;;; N;;;;;
11815; RANJANA LETTER JA; Lo; 0; L;;;;; N;;;;
11816; RANJANA LETTER JHA; Lo; 0; L;;;;; N;;;;;
11817; RANJANA LETTER NYA; Lo; 0; L;;;;; N;;;;;
11818; RANJANA LETTER TTA; Lo; 0; L;;;;; N;;;;;
11819; RANJANA LETTER TTHA; Lo; 0; L;;;;; N;;;;;
1181A; RANJANA LETTER DDA; Lo; 0; L;;;;; N;;;;;
1181B; RANJANA LETTER DDHA; Lo; 0; L;;;;; N;;;;;
1181C; RANJANA LETTER NNA; Lo; 0; L;;;;; N;;;;;
1181D; RANJANA LETTER TA; Lo; 0; L;;;;; N;;;;
1181E; RANJANA LETTER THA; Lo; 0; L;;;;; N;;;;;
1181F; RANJANA LETTER DA; Lo; 0; L;;;;; N;;;;;
11820; RANJANA LETTER DHA; Lo; 0; L;;;;; N;;;;;
11821; RANJANA LETTER NA; Lo; 0; L;;;;; N;;;;;
11822; RANJANA LETTER PA; Lo; 0; L;;;;; N;;;;;
11823; RANJANA LETTER PHA; Lo; 0; L;;;;; N;;;;;
11824; RANJANA LETTER BA; Lo; 0; L;;;;; N;;;;;
11825; RANJANA LETTER BHA; Lo; 0; L;;;;; N;;;;;
11826; RANJANA LETTER MA; Lo; 0; L;;;;; N;;;;
11827; RANJANA LETTER YA; Lo; 0; L;;;;; N;;;;
11828; RANJANA LETTER RA; Lo; 0; L;;;;; N;;;;;
11829; RANJANA LETTER LA; Lo; 0; L;;;;; N;;;;;
1182A; RANJANA LETTER WA; Lo; 0; L;;;;; N;;;;
1182B; RANJANA LETTER SHA; Lo; 0; L;;;;; N;;;;
1182C; RANJANA LETTER SSA; Lo; 0; L;;;;; N;;;;;
1182D; RANJANA LETTER SA; Lo; 0; L;;;;; N;;;;;
1182E; RANJANA LETTER HA; Lo; 0; L;;;;; N;;;;
1182F; RANJANA SIGN CANDRABINDU; Mn; 230; NSM; ;; ;; N; ;; ;;
11830; RANJANA SIGN ANUSVARA; Mn; 230; NSM;;;;; N;;;;;
11831; RANJANA SIGN VISARGA; Mc; 230; L;;;;; N;;;;
11832; RANJANA VOWEL SIGN AA; Mn; 0; NSM;;;;; N;;;;
11833; RANJANA VOWEL SIGN I; Mn; 0; NSM;;;;; N;;;;;
11834; RANJANA VOWEL SIGN II; Mn; 0; NSM;;;;; N;;;;;
11835; RANJANA VOWEL SIGN U; Mn; 220; NSM;;;;; N;;;;
11836; RANJANA VOWEL SIGN UU; Mn; 220; NSM; ;;;; N;;;;
11837; RANJANA VOWEL SIGN VOCALIC R; Mn; 220; NSM; ;;;; N;;;;;
11838; RANJANA VOWEL SIGN VOCALIC RR; Mn; 220; NSM;;;;; N;;;;;
11839; RANJANA VOWEL SIGN VOCALIC L; Mn; 220; NSM;;;;; N;;;;;
1183A; RANJANA VOWEL SIGN VOCALIC LL; Mn; 220; NSM;;;;; N;;;;
1183B; RANJANA VOWEL SIGN E; Mc; 0; L;;;;; N;;;;;
1183C; RANJANA VOWEL SIGN AI; Mn; 0; NSM;;;;; N;;;;;
1183D; RANJANA VOWEL SIGN O; Mn; 0; NSM;;;;; N;;;;
1183E; RANJANA VOWEL SIGN AU; Mn; 0; NSM;;;;; N;;;;
1183F; RANJANA VIRAMA; Mn; 9; NSM;;;;; N;;;;
11840; RANJANA DIGIT ZERO; Nd; 0; L;; 0; 0; 0; N;;;;;
11841; RANJANA DIGIT ONE; Nd; 0; L;; 1; 1; 1; N;;;;;
11842; RANJANA DIGIT TWO; Nd; 0; L;; 2; 2; 2; N;;;;;
11843; RANJANA DIGIT THREE; Nd; 0; L; ; 3; 3; 3; N; ; ; ; ;
11844; RANJANA DIGIT FOUR; Nd; 0; L; ; 4; 4; 4; N; ; ; ;
11845; RANJANA DIGIT FIVE; Nd; 0; L; ; 5; 5; 5; N; ; ; ; ;
11846; RANJANA DIGIT SIX; Nd; 0; L;; 6; 6; 6; N;;;;;
11847; RANJANA DIGIT SEVEN; Nd; 0; L;; 7; 7; 7; N;;;;;
11848; RANJANA DIGIT EIGHT; Nd; 0; L; ; 8; 8; 8; N; ; ; ; ;
11849; RANJANA DIGIT NINE; Nd; 0; L; ; 9; 9; 9; N; ; ; ;
1184B; RANJANA OM; Lo; 0; L;;;;; N;;;;
1184E; RANJANA YIG MGO; Po; 0; L;;;;; N;;;;
1184C; RANJANA DANDA; Po; 0; L;;;;; N;;;;
1184D; RANJANA DOUBLE DANDA; Po; 0; L;;;;; N;;;;;
1184F; RANJANA VIRAM; Po; 0; L;;;;; N;;;;;
```

11. Bibliography.

Šhākya, Rabison. 2002. *Nepal Lipi Varṇamālā* = *Alphabet of the Nepalese script*. Patan: Motirāj Śhākya. ISBN 99933-34-36-7

Shakyavansha, Hemraj. 1985. Nepal Lipi Samgraha = Nepalese Alphabets. Kathmandu: [s.n.].

Acknowledgements.

This project was made possible in part by a grant from the U.S. National Endowment for the Humanities, which funded the Universal Scripts Project (part of the Script Encoding Initiative at UC Berkeley) in respect of Rañjana and Pracalit encoding. Any views, findings, conclusions or recommendations expressed in this publication do not necessarily reflect those of the National Endowment of the Humanities.

98]

प्य

Consonants mixed with Ya

स्ति ए इंग्निस्पि Kya Khya Gya Ghya Nya Chya Chhya Jya Jhhya Yanya क्य क्य ग्य ह्य इन्य इन्य इन्य क्य

म्य

Figure 1. Example of conjuncts in -ya from Shakyavansha 1985.

[??

प्र

ऋ

Consonants mixed with Ra रकार संयोग व्यञ्जनवर्ग

Kra Khra Gra Ghra Nra Chra Chhra Jra Jhra Yanra ञ् म् ङ् স্প छ त्र प्र घ ख ऋ

Tra Thra Dra Dhra Nra Tra Thra Dra Dhra Nra

न् द्र ध्र थ्र ढ्र Ų त्र ठॅ ड्र ट्र Yra Rra Lra Wra

Pra Phra Bra Bhra Ma 3 यू म्र

Shra Shra Sra Hra Kshra Trra Jnra

त्र ज् त्र ह्र भ

Figure 2. Example of conjuncts in *-ra* from Shakyavansha 1985.

(3]

Consonants mixed with La लकार संयोग व्यञ्जनवग





Figure 3. Example of conjuncts in -la from Shakyavansha 1985.

14]

Consonants mixed with Wa

सिद्धिद्वि

Twa ThwaDwaDhwaNwa Twa ThwaDwaDhwaNwa

ट्व ट्व ड्व ड्व एव त्व ध्व ट्व ध्व न्व

स्याद्वस्य सुमुख्द

Pwa Phwa Bwa Bhwa Mwa Ywa Rwa LwaWwa प्रव पन उन पन उन पन उन उन उन उन उन

ह्य यहाँ विक् स्थान

Shwa Shwa Swa Hwa Kshwa Trwa Jnwa হৰ তা হৰ হ তথ স্থা হৰ

Figure 4. Example of conjuncts in -wa from Shakyavansha 1985.

?x]

Consonants mixed with Na नकार संयोग व्यञ्जनवग

Kna Khna Gna Ghna Nna Chna Chhna Jna Jhna भन छ न इत हत् ङन च्त ग्न ख्न क्त

Tna Thna Dna Dhna Nna Tna Thna Dna Dhna Nna द्न रन त्न ध्न ध्न ন্ন ढ्न ठन इन टन

Pna Phna Bna Bhna Mna Yna Rna Lna Wna **न्**न यन ल्न म्न ब्न भ्न फ्न प्न

Shna Shna **S**na **H**na Kshna स्न

इन

Figure 5. Example of conjuncts in *-na* from Shakyavansha 1985.

ष्टन

श्न

16] १६

Consonants mixed with Ma मकार संयोग व्यञ्जनवर्ण

KmaKhmaGmaGhmaNmaChmaChhmaJmaJhma

छ्म ङ्म क्स

६न

14

७म

च्रम

भम

उम

TmaThmaDmaDhmaNmaTma ThmaDmaDhmaNma न्म ह्म ध्स द्म ध्म एम त्म ड्म

Pina Phma Bma Bhma Mma Yma Rma Lma Wma च्म र्म रम ∓Ħ प्म पम जम **¥4**

Shma Shma Sma Hma Kahma इम ह्य स्म र्म

Figure 6. Example of conjuncts in *-ma* from Shakyavansha 1985.

१७]

पूर्वतो वर्गीय पश्चम व्यञ्जनवर्ण

Read State To A Stat

प्राच्चा स्वास्त्र स्वास्



Figure 7. Example of conjuncts with initial homorganic nasals \dot{n} -, \tilde{n} -, n-, n-, and m-, from Shakyavansha 1985.

18]

Consonants mixed with Sa सकार संयोग व्यक्तनवर्ण

स्यायास्यान्त्रव्यान्त्रव

Ska Skha Sga Sgha Sna Scha Schha Sja Sjha Syan ভদ্ধ ভদ্ধ ভদ্ধ ভদ্ধ হন্ত হন্দ হন্দ

वाव्ययायात्रम्यस्य

Sta Stha Sda Sdha Sna Sta Stha Sda Sdha Sna ष्ट ष्ट ब्ह ब्हा स्त स्थ स्द स्व स्न



Figure 8. Example of conjuncts with initial *s*- from Shakyavansha 1985.

For reading practice अभ्यास निम्ति परलप्पेगु शिचा

Figure 9. Example text in Rañjana script from Shakyavansha 1985.

Consonants mixed with Vowels from Ka to Jna स्वरवर्ण संयोग ब्यञ्जनवर्ण क- ज्ञ पर्यन्त

Ka Ka Ki Kee Ku Koo Kri Kree Ke Kai Ko Kau Kan Kah



Figure 10. Example showing the letters *ka*, *kha*, *ga*, and *gha* with vowel signs, from Shakyavansha 1985.

A. Administrative

1. Title

Proposal for encoding the Ranjana script in the SMP of the UCS

2. Requester's name

Michael Everson

3. Requester type (Member body/Liaison/Individual contribution)

Individual contribution.

4. Submission date

2009-05-04

- 5. Requester's reference (if applicable)
- 6. Choose one of the following:

6a. This is a complete proposal

No.

6b. More information will be provided later

Yes.

B. Technical - General

1. Choose one of the following:

1a. This proposal is for a new script (set of characters)

Ves

1b. Proposed name of script

Rañjana.

1c. The proposal is for addition of character(s) to an existing block

No.

1d. Name of the existing block

2. Number of characters in proposal

79.

3. Proposed category (A-Contemporary; B.1-Specialized (small collection); B.2-Specialized (large collection); C-Major extinct; D-Attested extinct; E-Minor extinct; F-Archaic Hieroglyphic or Ideographic; G-Obscure or questionable usage symbols)

Category A.

4a. Is a repertoire including character names provided?

Yes

4b. If YES, are the names in accordance with the "character naming guidelines" in Annex L of P&P document?

Yes

4c. Are the character shapes attached in a legible form suitable for review?

Yes.

5a. Who will provide the appropriate computerized font (ordered preference: True Type, or PostScript format) for publishing the standard?

Michael Everson.

5b. If available now, identify source(s) for the font (include address, e-mail, ftp-site, etc.) and indicate the tools used:

Michael Everson, FontLab.

6a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?

Yes

6b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached? **Yes.**

7. Does the proposal address other aspects of character data processing (if applicable) such as input, presentation, sorting, searching, indexing, transliteration etc. (if yes please enclose information)?

Yes.

8. Submitters are invited to provide any additional information about Properties of the proposed Character(s) or Script that will assist in correct understanding of and correct linguistic processing of the proposed character(s) or script. Examples of such properties are: Casing information, Numeric information, Currency information, Display behaviour information such as line breaks, widths etc., Combining behaviour, Spacing behaviour, Directional behaviour, Default Collation behaviour, relevance in Mark Up contexts, Compatibility equivalence and other Unicode normalization related information. See the Unicode standard at http://www.unicode.org for such information on other scripts. Also see Unicode Character Database http://www.unicode.org/Public/UNIDATA/UnicodeCharacterDatabase.html and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.

See above.

C. Technical – Justification

1. Has this proposal for addition of character(s) been submitted before? If YES, explain.

Nο.

2a. Has contact been made to members of the user community (for example: National Body, user groups of the script or characters, other experts, etc.)?

Yes.

2b. If YES, with whom?

2c. If YES, available relevant documents

3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included?

People living in Nepal, India, Tibet. Scholars.

4a. The context of use for the proposed characters (type of use; common or rare)

Common.

4b. Reference

5a. Are the proposed characters in current use by the user community?

Yes

5b. If YES, where?

In Nepal and elsewhere.

6a. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP?

Ves.

6b. If YES, is a rationale provided?

Yes.

6c. If YES, reference

Contemporary use and accordance with the Roadmap.

7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?

Yes.

8a. Can any of the proposed characters be considered a presentation form of an existing character or character sequence?

No.

8b. If YES, is a rationale for its inclusion provided?

8c. If YES, reference

9a. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters?

No.

9b. If YES, is a rationale for its inclusion provided?

9c. If YES, reference

10a. Can any of the proposed character(s) be considered to be similar (in appearance or function) to an existing character?

No.

10b. If YES, is a rationale for its inclusion provided?

10c. If YES, reference

11a. Does the proposal include use of combining characters and/or use of composite sequences (see clauses 4.12 and 4.14 in ISO/IEC 10646-1: 2000)?

No.

11b. If YES, is a rationale for such use provided?

11c. If YES, reference

11d. Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?

No.

11e. If YES, reference

12a. Does the proposal contain characters with any special properties such as control function or similar semantics?

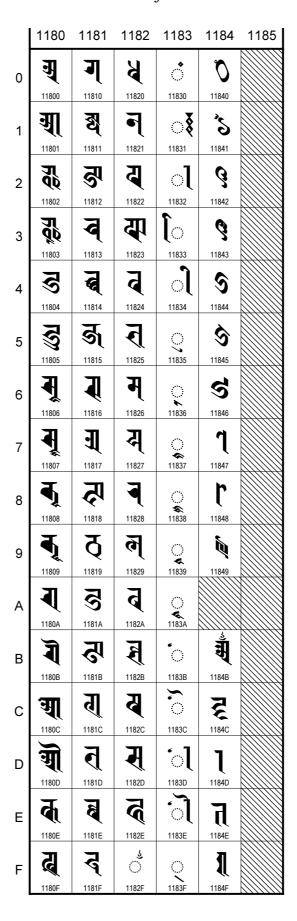
No.

12b. If YES, describe in detail (include attachment if necessary)

13a. Does the proposal contain any Ideographic compatibility character(s)?

No.

13b. If YES, is the equivalent corresponding unified ideographic character(s) identified?



Date: 2009-05-04

Independent vowels

- 11800 U RANJANA LETTER A
 11801 U RANJANA LETTER AA
 11802 U RANJANA LETTER I
 11803 U RANJANA LETTER II
 11804 U RANJANA LETTER U
 11805 U RANJANA LETTER U
- 11806 ◀ RANJANA LETTER VOCALIC R 11807 ◀ RANJANA LETTER VOCALIC RR 11808 ◀ RANJANA LETTER VOCALIC L 11809 ◀ RANJANA LETTER VOCALIC LL 1180A ◀ RANJANA LETTER E
- 1180B T RANJANA LETTER AI 1180C T RANJANA LETTER O 1180D RANJANA LETTER AU

Consonants

- 1181D d RANJANA LETTER TA 1181E d RANJANA LETTER THA 1181F d RANJANA LETTER DA
- 11820 d RANJANA LETTER DHA 11821 d RANJANA LETTER NA
- 11822 **4** RANJANA LETTER PA 11823 **4** RANJANA LETTER PHA 11824 **4** RANJANA LETTER BA
- 11825 T RANJANA LETTER BA 11826 T RANJANA LETTER MA
- 11827 RANJANA LETTER YA
- 11828 RANJANA LETTER RA 11829 RANJANA LETTER LA
- 1182A **4** RANJANA LETTER WA 1182B **4** RANJANA LETTER SHA
- 1182C RANJANA LETTER SSA 1182D RANJANA LETTER SA
- 1182E 🐧 RANJANA LETTER HA

Various signs

Dependent vowel signs

11832 OL RANJANA VOWEL SIGN AA 11833 OL RANJANA VOWEL SIGN I 11834 OL RANJANA VOWEL SIGN II 11835 Q RANJANA VOWEL SIGN U
11836 Q RANJANA VOWEL SIGN UU
11837 Q RANJANA VOWEL SIGN VOCALIC R
11838 Q RANJANA VOWEL SIGN VOCALIC R
11839 Q RANJANA VOWEL SIGN VOCALIC L
1183A Q RANJANA VOWEL SIGN VOCALIC LL
1183B Q RANJANA VOWEL SIGN E
1183C Q RANJANA VOWEL SIGN AI
1183D Q RANJANA VOWEL SIGN O
1183E Q RANJANA VOWEL SIGN O

Various signs

1183F 🍳 RANJANA VIRAMA

Digits

11840 Nanjana digit zero 11841 Ś RANJANA DIGIT ONE Ġ 11842 RANJANA DIGIT TWO **9** RANJANA DIGIT THREE 11843 **\$** RANJANA DIGIT FOUR 11844 11845 🥱 RANJANA DIGIT FIVE 11846 **ತ** RANJANA DIGIT SIX 11847 **1** RANJANA DIGIT SEVEN 11848 C RANJANA DIGIT EIGHT 11849 RANJANA DIGIT NINE

Punctuation

Date: 2009-05-04

- 1184B RANJANA OM
 1184C RANJANA YIG MGO
 1184D RANJANA DANDA
 1184E RANJANA DOUBLE DANDA
- 1184F 🌓 RANJANA VIRAM