Proposal to Encode the Jenticha Script in ISO/IEC 10646

Anshuman Pandey
University of Michigan
Department of History
Ann Arbor, Michigan, U.S.A.
pandey@umich.edu

May 31, 2011

1 Introduction

This is a preliminary proposal to encode the Jenticha script in the Universal Character Set (ISO/IEC 10646). It replaces the preliminary proposal submitted for encoding the script (N3962 L2/10-466R). Changes made to the proposal include the use of English names for signs, the addition of a new combining sign, and more general detail about the proposed characters. Code points have been revised to reflect changes to the allocation for the script in the Unicode Roadmap.

2 Background

2.1 Overview

Jenticha is a script used for writing Sunuwar (ISO 639-3: suz), a Kiranti language of the Tibeto-Burman family that is spoken in Sikkim, India and eastern Nepal. Sunuwar is spoken by 40,000 persons in Nepal. The number of speakers of the language in Sikkim is unknown because data for Sunuwar is absent from the 2001 Census of India. However, Sunuwar was recognized as an official language of Sikkim in 1996. Jenticha is a recognized script in Sikkim. Devanagari is also used for writing Sunuwar and is the most commonly used script for the language.

Sunuwar (सुनुवार *sunuvāra*), alternately spelled 'Sunwar', is also known by the name 'Mukhiya' (मुखिया *mukhiyā*). Both are Nepali exonyms for the language and its speakers. The autonym is कोँइच *kõica*, which is normalized as 'Kõits'. The language is called कोँइच लो *kõica lo* 'Kõits Lo' or किराँती-कोँइच *kirā̃tī-kõica* 'Kiranti-Kõits'. The name 'Kõits' is also romanized as 'Koinch', 'Koincha', 'Koints', etc.

The script was developed by Krishna Bahadur Jenticha (1926–1991) in 1942. It is known indigenously as जैँःतिच ब्रेःसे *jẽtica brese*, 'Jenticha Brese'. The script is also called कोँइच ब्रेःसे *kõica brese* 'Kõits Brese', or the 'Kõits script'; 'brese', or script, is also romanized as 'brehs' or 'blehs', ie. 'Koinch Brehs'. The script is referred to here as 'Jenticha'. It has no genetic relationship to other writing systems. However, some of its letters resemble those of other scripts, such as Limbu and Latin.

There are two major phases in the development of Jenticha. The earliest script developed by K. B. Jenticha is a pure alphabet (see figures 3, 4, and 5). The modern script is a revision of the original and differs from its precedessor in structure, letter forms, and character repertoire. Modern Jenticha possesses features of an alphasyllabic script, namely that each consonant letter has an inherent vowel.

2.2   Usage

Jenticha is currently used in writing and print for various purposes:

- *Educational Materials*   Primary educational materials for teaching Sunuwar and Jenticha are published by Kõic Bu in Sikkim. Examples of script primers are ꝥꞔ桷 桷ꞔ *Kirãtī-Kõica Blesethekā* (2003) and 桷ꞔ *Kõicā The:si Khyõpat* (2004); see Figure 11. The Department of Education of the Government of Sikkim is planning to prepare textbooks in the script for primary-level education.

- *Literary Works*   Anthologies of Sunuwar poetry are printed in the script.

- *Historical Records*   Histories of the Sunuwar communities, known as *vaṃśāvalī*, were written in Jenticha beginning in the 1940s.

- *Newspapers*   A Sunuwar ('Mukhiya') edition of the *Sikkim Herald* (桷ꞔ) is produced in Jenticha. The newspaper was first published in April 2002 and maintains an active circulation. The first edition was hand-written by Lāl-Śyāṃkārelu Rāpacā, with subsequent editions written by Balarām Sunuwār (see Figure 9). More recent editions are printed in digitized type (see Figure 10).

- *Linguistic Works*   The script is used in grammars of Sunuwar, ie. *Elementary Grammar of Kiranti-Sunuwar* by Kamalādevī Ṅāvacā-Mukhiyā, Raghuvīra Rujicā-Mukhiya, and Lāla Rāpacā (Sikkim: Kõica Bu, 2003).

- *Government Records*   Proceedings of the Sikkim legislative assembly are translated into Sunuwar using the script.

- *Academic Works*   The script is used in several scholarly works on the Sunuwar language and it is also the focus of scholarly attention (cf. Rāpacā 2009).

- *Ephemera*   The Sikkim Sunuwar Mukhia Koinchbu (SSMK) produces and distributes calendars using Jenticha (see Figure 13 and 14).

- *Digital Resources*   Jenticha has been adapted for use in digital media. A digitized font for the script named 'Kirat1' was developed by Shyan Kirat Rai. It is based upon a non-Unicode encoding and is mapped to Latin letters. The font is used for publishing the *Sikkim Herald* (see Figure 10) and by Rāpacā in the charts and examples shown in figures 7–12. Another effort is the Jenticha font used in this document. This font was designed by the present author and will be further developed to support the Unicode implementation for Jenticha.

3   Implementation Details

3.1   Script Name

The script is named 'Jenticha'.

3.2 Character Repertoire

The proposed characters are intended for representing historical and current forms of Jenticha. The set consists of 45 characters: 6 vowel letters, 25 consonant letters, 4 signs, and 10 digits. A code chart and names list is given in Figure 1.

3.3 Structure

Modern Jenticha is an alphabetic writing system with alphasyllabic features. The early version of the script is a pure alphabet. Both forms of the script are written from left to right. The modern form of Jenticha is a superset that encompasses the character repertoire and orthography of the early script. The proposed script is based upon modern Jenticha.

Jenticha should be implemented according to the *virāma* model used for Brahmi-based scripts. This encoding model offers the most appropriate means for representing consonant conjuncts written in the script.

3.4 Vowels

Vowels are represented using individual letters, of which there are six:

�	A	/a/	अ		ꚋ	I	/i/	इ		±	E	/e/	ए
	AA	/aː/	आ		ꛯ	U	/u/	उ		O	O	/o/	ओ

In the early form of the alphabet, the letter ꗓ A was used for writing /a/ and /aː/ (see figures 3 and 5). The letter ꗙ AA was later introduced as a unique character for /aː/.

The diphthongs of Sunuwar are written as digraphs using vowel letters:

/ai/	ꗓꚋ	ꗓ A + ꚋ I	ऐ		/eu/	±ꛯ	± E + ꛯ U	एउ		/oi/	Oꚋ	O O + ꚋ I	ओइ
/au/	ꗓꛯ	ꗓ A + ꛯ U	औ		/ui/	ꛯꚋ	ꛯ U + ꚋ I	उइ					

3.5 Vowel Length

Vowel length is indicated using the sign ◌ː VOWEL LENGTH MARK, which is called *gyosh* in Sunuwar. The length mark is used in modern Jenticha. Vowel length is not distinguished in writing in the older alphabetic script.

3.6 Virāma

The Jenticha ◌ VIRAMA behaves similar to the ◌ VIRAMA of Devanagari. It is used in modern Jenticha and is not attested in early forms of the script.

3.7 Consonants

There are 25 consonant letters:

�	KA	/k/	क		୨	PA	/p/	प
⊓	KHA	/kʰ/	ख		α	PHA	/pʰ/	फ
੫	GA	/g/	ग		ɟ	BA	/b/	ब
୪	NGA	/ŋ/	ङ		Л	MA	/m/	म
כ	GLOTTAL STOP	/ʔ/	अ्		ଅ	YA	/y/	य
‹	CA	/tʃ/	च		N	RA	/r/, [l]	र
⩛	JA	/dʒ/	ज		⅄	LA	/l/, [l]	ल
⅄	TTA	/ʈ/	ट		N	VA	/v/	व
ⱬ	TTHA	/ʈʰ/	ठ		∪	SHA	/ʃ/	श, ष
∂	TA	/t/	त		⅄	SA	/s/	स
ઝ	THA	/tʰ/	थ		ᚻ	HA	/h/	ह
⊔	DA	/d/	द, ड		Ӿ	HHA	/ɦ/	व्हो, ह्व
�	NA	/n/	न, ण					

In modern Jenticha, each consonant letter, except for כ GLOTTAL STOP, bears the inherent vowel /a/, which corresponds to the letter ୨ A. The inherent vowel is changed by writing a vowel letter after the consonant, eg. � /ka/, �Ⴈ /ka:/, etc. It is silenced by writing ○ VIRAMA under a consonant, eg. <� KA + ○> → <�> /k/.

In early Jenticha, consonant letters are purely alphabetic and do not have an inherent vowel. The sequence <� KA + ୨ A> represents both /ka/ and /ka:/.

3.8   Consonant Conjuncts

Consonant conjuncts are represented differently in the early and modern forms of Jenticha. In early Jenticha, conjuncts are written as a linear sequence of consonant letters without VIRAMA or special rendering. In modern orthography, the default representation for conjuncts is with a visible VIRAMA. The encoded representation is $<C_1 + $ ○ VIRAMA $ + C_2>$.

In modern Jenticha orthography there are alternate methods for representing conjuncts. It may be necessary to represent both default and alternate forms in plain text. In order to do so, the [zw] U+200D ZERO WIDTH JOINER should be used to force the alternate form: $<C_1 + $ VIRAMA $ + $ [zw] ZWJ $ + C_2>$.

1. *Half-Forms*   The following consonant letters may be written using half-forms (pre-base forms) when they appear as C_1 in a conjunct:

	Regular form	Default bare form	Half-form	Devanagari half-form
CA	‹	‹	‹	च
MA	Л	Л	Л	म
SA	⅄	⅄	1	स

2. *Above-Base Forms*   In modern orthography, the letter N RA has an alternate contextual form when it occurs as C_1 in conjuncts. The default behavior for RA as C_1 is with visible *virāma*: N‹ *rca*. The alternate is the form ○̇ *sangkirs*, which is written above C_2: ‹̇ *rca*. It corresponds to *repha* in Devanagari. The encoded representation for *sangkirs* is <RA + VIRAMA + [zw] ZWJ + C>.

4

3.9 Geminate Consonants

There are no formal rules for writing geminate consonants. In early Jenticha, a consonant letter is written twice. In the modern script, geminate consonants should be treated as a conjunct with the first consonant marked with a visible VIRAMA.

3.10 Nasalization

Nasalization is indicated by $\overset{\star}{\circ}$ CANDRABINDU, which is called *taslathenk* in Sunuwar. This sign corresponds to $\overset{\circ}{\circ}$ CANDRABINDU in Devanagari.

3.11 Aspiration

Early Jenticha did not have unique letters for the aspirated consonants /k^h/, /ṭ^h/, /t^h/, /p^h/ (see Figure 5). Aspiration was indicated by writing ⊦ HA after an unaspirated consonant letter, eg. ⟨ℐ PA + ⊦ HA⟩ → ⟨ℐ⊦⟩ /p^h/. The modern script has independent letters for the four aspirated consonants.

3.12 Trill Mark

The $\overset{\circ}{\circ}$ TRILL MARK represents [ḷ], which is an allophone of /l/ when it occurs as C_2 in a consonant conjunct with a velar or labial consonant as C_1. There is no distinct Jenticha letter for this sound. The mark is called *sangrums* and is used only in modern Jenticha. Use of the TRILL MARK is optional. The word ⱦⱡⱦⱡ *bḷese* may be written using either ⋈ RA or ⋋ LA; eg. ⱦ⋈ⱦⱡ *brese*, ⱦ⋋ⱦⱡ *blese*. In Devanagari orthography, [ḷ] is represented using र RA, ल LA, or ड DDA. Valid sequences are:

ⱦ	KA	+	◌	→	ⱦ	/kḷ/	क्ड, क्र, क्ल	ℐ	PA	+ ◌ →	ℐ	/pḷ/	प्ड, प्र, प्ल
⋂	KHA	+	◌	→	⋂	/k^{h}ḷ/	ख्ड, ख्र, ख्ल	⍺	PHA	+ ◌ →	⍺	/p^{h}ḷ/	फ्ड, फ्र, फ्ल
�	GA	+	◌	→	Ⴡ	/gḷ/	ग्ड, ग्र, ग्ल	ⱦ	BA	+ ◌ →	ⱦ	/bḷ/	ब्ड, ब्र, ब्ल

3.13 Digits

Digits represent decimal numbers: Ѳ ZERO, ⎮ ONE, ⱦ TWO, ⱦ THREE, ▯ FOUR, ⊟ FIVE, ◿ SIX, �圓 SEVEN, Χ EIGHT, ⊦ NINE.

3.14 Punctuation

There are no script-specific punctuation marks. Latin marks, such as periods and commas, are used.

3.15 Collation

The sort order follows the encoded order.

3.16 Line-Breaking

All characters have line-breaking rules similar to that of Devanagari.

3.17   Ordering

A letter with two combining marks is common in Jenticha, eg. <◌̐> CANDRABINDU and VOWEL LENGTH MARK. In such cases, CANDRABINDU should be ordered first. Rendering engines should recognize sequences of combining marks as valid.

3.18   Character Properties

```
116D0;JENTICHA LETTER A;Lo;0;L;;;;;N;;;;;
116D1;JENTICHA LETTER AA;Lo;0;L;;;;;N;;;;;
116D2;JENTICHA LETTER I;Lo;0;L;;;;;N;;;;;
116D3;JENTICHA LETTER U;Lo;0;L;;;;;N;;;;;
116D4;JENTICHA LETTER E;Lo;0;L;;;;;N;;;;;
116D5;JENTICHA LETTER O;Lo;0;L;;;;;N;;;;;
116D6;JENTICHA LETTER KA;Lo;0;L;;;;;N;;;;;
116D7;JENTICHA LETTER KHA;Lo;0;L;;;;;N;;;;;
116D8;JENTICHA LETTER GA;Lo;0;L;;;;;N;;;;;
116D9;JENTICHA LETTER NGA;Lo;0;L;;;;;N;;;;;
116DA;JENTICHA LETTER CA;Lo;0;L;;;;;N;;;;;
116DB;JENTICHA LETTER JA;Lo;0;L;;;;;N;;;;;
116DC;JENTICHA LETTER TTA;Lo;0;L;;;;;N;;;;;
116DD;JENTICHA LETTER TTHA;Lo;0;L;;;;;N;;;;;
116DE;JENTICHA LETTER TA;Lo;0;L;;;;;N;;;;;
116DF;JENTICHA LETTER THA;Lo;0;L;;;;;N;;;;;
116E0;JENTICHA LETTER DA;Lo;0;L;;;;;N;;;;;
116E1;JENTICHA LETTER NA;Lo;0;L;;;;;N;;;;;
116E2;JENTICHA LETTER PA;Lo;0;L;;;;;N;;;;;
116E3;JENTICHA LETTER PHA;Lo;0;L;;;;;N;;;;;
116E4;JENTICHA LETTER BA;Lo;0;L;;;;;N;;;;;
116E5;JENTICHA LETTER MA;Lo;0;L;;;;;N;;;;;
116E6;JENTICHA LETTER YA;Lo;0;L;;;;;N;;;;;
116E7;JENTICHA LETTER RA;Lo;0;L;;;;;N;;;;;
116E8;JENTICHA LETTER LA;Lo;0;L;;;;;N;;;;;
116E9;JENTICHA LETTER VA;Lo;0;L;;;;;N;;;;;
116EA;JENTICHA LETTER SHA;Lo;0;L;;;;;N;;;;;
116EB;JENTICHA LETTER SA;Lo;0;L;;;;;N;;;;;
116EC;JENTICHA LETTER HA;Lo;0;L;;;;;N;;;;;
116ED;JENTICHA LETTER HHA;Lo;0;L;;;;;N;;;;;
116EE;JENTICHA LETTER GLOTTAL STOP;Lo;0;L;;;;;N;;;;;
116EF;JENTICHA SIGN VIRAMA;Mn;9;NSM;;;;;N;;;;;
116F0;JENTICHA SIGN CANDRABINDU;Mn;0;NSM;;;;;N;;;;;
116F1;JENTICHA VOWEL LENGTH MARK;Mc;0;L;;;;;N;;;;;
116F2;JENTICHA TRILL MARK;Mn;0;NSM;;;;;N;;;;;
116F6;JENTICHA DIGIT ZERO;Nd;0;L;;0;0;0;N;;;;;
116F7;JENTICHA DIGIT ONE;Nd;0;L;;1;1;1;N;;;;;
116F8;JENTICHA DIGIT TWO;Nd;0;L;;2;2;2;N;;;;;
116F9;JENTICHA DIGIT THREE;Nd;0;L;;3;3;3;N;;;;;
116FA;JENTICHA DIGIT FOUR;Nd;0;L;;4;4;4;N;;;;;
116FB;JENTICHA DIGIT FIVE;Nd;0;L;;5;5;5;N;;;;;
116FC;JENTICHA DIGIT SIX;Nd;0;L;;6;6;6;N;;;;;
116FD;JENTICHA DIGIT SEVEN;Nd;0;L;;7;7;7;N;;;;;
116FE;JENTICHA DIGIT EIGHT;Nd;0;L;;8;8;8;N;;;;;
116FF;JENTICHA DIGIT NINE;Nd;0;L;;9;9;9;N;;;;;
```

3.19   'Confusable' Characters

Below is a basic list of Jenticha characters that are 'confusable' with those encoded in other script blocks:

6

```
116D1 JENTICHA LETTER AA          ; 1900 LIMBU VOWEL-CARRIER LETTER
116D5 JENTICHA LETTER O           ; 004F LATIN CAPITAL LETTER O
116DC JENTICHA LETTER TTA         ; 1916 LIMBU LETTER RA
116DD JENTICHA LETTER TTHA        ; 190F LIMBU LETTER NA
116DF JENTICHA LETTER THA         ; 190C LIMBU LETTER THA
116E3 JENTICHA LETTER PHA         ; 03B1 GREEK SMALL LETTER ALPHA
116EB JENTICHA LETTER SA          ; 1919 LIMBU LETTER SHA
116EF JENTICHA SIGN VIRAMA        ; 0332 COMBINING LOW LINE
116F1 JENTICHA VOWEL LENGTH MARK  ; 0903 DEVANAGARI SIGN VISARGA
116F7 JENTICHA DIGIT ONE          ; 0964 DEVANAGARI DANDA
116FE JENTICHA DIGIT EIGHT        ; 0058 LATIN CAPITAL LETTER X
```

4   References

Borchers, Dörte. 2008. *A Grammar of Sunwar: Descriptive Grammar, Paradigms, Texts and Glossary*. Tibetan Studies Library; Languages of the Greater Himalayan Region, vol. 5/7. Leiden: Brill.

kiratsunuwar.org.np. (n.d.) "Summary of K. B. Sunuwar". Written by Shankar Sunuwar. Source: Purna Bahadur Sunuwar. `http://kiratsunuwar.org.np/index.php?option=com_content&view=article& id=8:summary-of-kb-sunuwar&catid=3:articles&Itemid=19`. Accessed January 2011.

Mukhia, Kamaladevi, and Raghubir Mukhia, and Lal Rapaca. 2004. ꗞꗞ [= *Elementary Reader of Sunuwar*]. 1st ed. Sikkim: Kõica Bu.

Pandey, Anshuman. 2011a. "Preliminary Proposal to Encode the Tikamuli Script in ISO/IEC 10646" (N3963 L2/10-465R). January 25, 2011. `http://std.dkuug.dk/JTC1/SC2/WG2/docs/n3963.pdf`

———. 2011b. "Preliminary Proposal to Encode the Jenticha Script in ISO/IEC 10646" (N3962 L2/10-466R). January 25, 2011. `http://std.dkuug.dk/JTC1/SC2/WG2/docs/n3962.pdf`

रापचा, लाल-श्याँकारेलु [Rāpacā, Lāl-Śyāṃkārelu]. 2009. नेपालको आदिवासी किराँती जाति सेरोफेरो : ओल्लो, माझ र पल्लो किराँत. काठमांडौ : किरांतविज्ञान अध्ययन संस्थान.

Subba, Jash Raj. 2008. *The Evolution of Man and the Modern Society in Mountainous Sikkim*. Delhi: Gyan Publishing House.

5   Acknowledgments

	116D	116E	116F
0	Ɠ 116D0	Ⱑ 116E0	⸫ 116F0
1	Ⱑ 116D1	η 116E1	◌⸲ 116F1
2	† 116D2	Ƴ 116E2	◌̂ 116F2
3	Ⱬ 116D3	α 116E3	
4	± 116D4	ƒ 116E4	
5	O 116D5	Ⱉ 116E5	
6	Ƭ 116D6	ℓ 116E6	θ 116F6
7	Π 116D7	Ν 116E7	I 116F7
8	Ƴ 116D8	⅄ 116E8	≢ 116F8
9	ß 116D9	Ν 116E9	ƚ 116F9
A	< 116DA	U 116EA	▯ 116FA
B	ⱴ 116DB	ⱦ 116EB	⊟ 116FB
C	ⱬ 116DC	ⱨ 116EC	⊿ 116FC
D	Ɀ 116DD	Ⱶ 116ED	Ⅲ 116FD
E	∂ 116DE	Ƅ 116EE	Χ 116FE
F	ⱬ 116DF	◌̲ 116EF	�янь 116FF

Vowels

116D0   Ɠ   JENTICHA LETTER A
116D1   Ⱑ   JENTICHA LETTER AA
116D2   †   JENTICHA LETTER I
116D3   Ⱬ   JENTICHA LETTER U
116D4   ±   JENTICHA LETTER E
116D5   O   JENTICHA LETTER O

Consonants

116D6   Ƭ   JENTICHA LETTER KA
116D7   Π   JENTICHA LETTER KHA
116D8   Ƴ   JENTICHA LETTER GA
116D9   ß   JENTICHA LETTER NGA
116DA   <   JENTICHA LETTER CA
116DB   ⱴ   JENTICHA LETTER JA
116DC   ⱬ   JENTICHA LETTER TTA
116DD   Ɀ   JENTICHA LETTER TTHA
116DE   ∂   JENTICHA LETTER TA
116DF   ⱬ   JENTICHA LETTER THA
116E0   Ⱑ   JENTICHA LETTER DA
116E1   η   JENTICHA LETTER NA
116E2   Ƴ   JENTICHA LETTER PA
116E3   α   JENTICHA LETTER PHA
116E4   ƒ   JENTICHA LETTER BA
116E5   Ⱉ   JENTICHA LETTER MA
116E6   ℓ   JENTICHA LETTER YA
116E7   Ν   JENTICHA LETTER RA
116E8   ⅄   JENTICHA LETTER LA
116E9   Ν   JENTICHA LETTER VA
116EA   U   JENTICHA LETTER SHA
116EB   ⱦ   JENTICHA LETTER SA
116EC   ⱨ   JENTICHA LETTER HA
116ED   Ⱶ   JENTICHA LETTER HHA
116EE   Ƅ   JENTICHA LETTER GLOTTAL STOP

Various signs

116EF   ◌̲   JENTICHA SIGN VIRAMA
        = sangmilu, halant
116F0   ◌̇   JENTICHA SIGN CANDRABINDU
        = taslathenk
116F1   ◌⸲   JENTICHA VOWEL LENGTH MARK
        = gyosh, dirghata
116F2   ◌̂   JENTICHA TRILL MARK
        = sangrums

Digits

116F6   θ   JENTICHA DIGIT ZERO
116F7   I   JENTICHA DIGIT ONE
116F8   ≢   JENTICHA DIGIT TWO
116F9   ƚ   JENTICHA DIGIT THREE
116FA   ▯   JENTICHA DIGIT FOUR
116FB   ⊟   JENTICHA DIGIT FIVE
116FC   ⊿   JENTICHA DIGIT SIX
116FD   Ⅲ   JENTICHA DIGIT SEVEN
116FE   Χ   JENTICHA DIGIT EIGHT
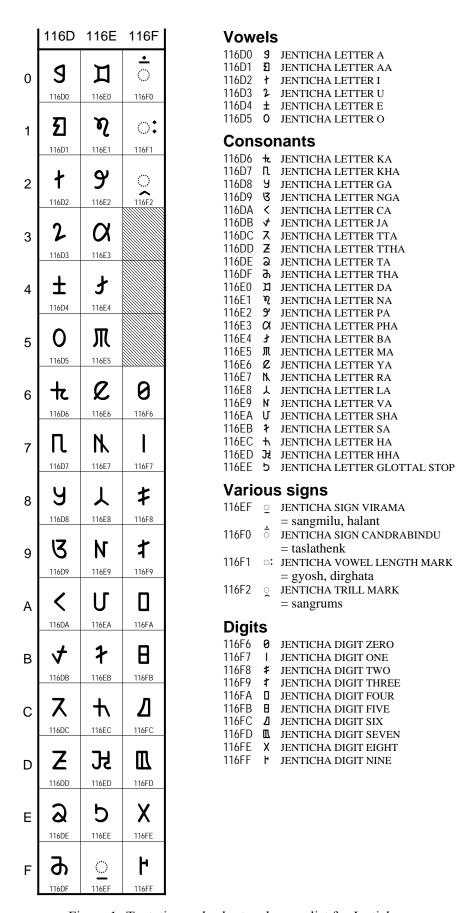116FF   ⱨ   JENTICHA DIGIT NINE

Figure 1: Tentative code chart and nameslist for Jenticha.

Figure 2: Krishna Bahadur Jenticha (Sunuwar), the inventor of the Jenticha script (from kirat-sunuwar.org.np).
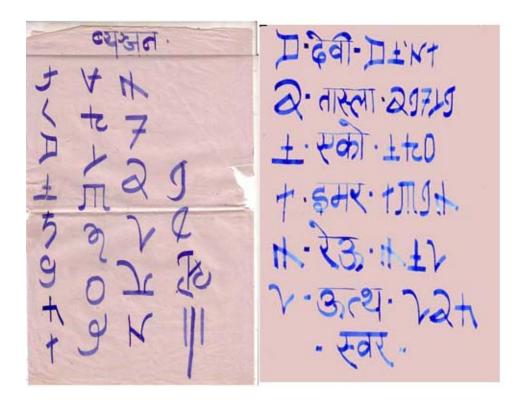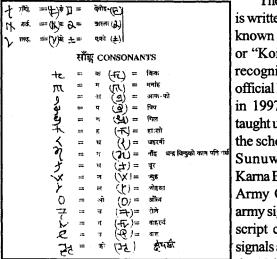

Figure 3: Early Jenticha letterforms (from kiratsunuwar.org.np). Notice the use of ꣯ A for writing both the inherent vowel /a/ and /a/; eg. in the words ꣯ꣳ /tasla/ and ꣷ /imar/.

Sunwar	Devanāgarī	Roman	Sunwar	Devanāgarī	Roman/ Arabic
	अ	a		उ ऊ	u, ū
	ब	b		व	v
	च	c		व्हो	vho
	द	d		य	y
	ए ॒	e, e		०	0
	ग	g		१	1
	ह	h		२	2
	इ ई	i, ī		३	3
	ज	j		४	4
	क	k		५	5
	ल	l		६	6
	म	m		७	7
	न	n		८	8
	ओ	o		९	9
	प	p		१०	10
	र ड	r, ḍ			
	स	s			
	त ट	t, ṭ			

Figure 4: A chart of early Jenticha (from Borchers 2008: 47).

10

Figure 5: A table showing names for chronological and calendrical terminology in Sunuwar (with Hindi equivalents) written in early Jenticha and Devanagari (from Borchers 2008: 48).

7.6 Sunuwar Script (Koinch Bress):

The Sunuwar language is written in Sunuwar script known as "Koinch Bress" or "Koinch Bnu" since its recognition as one of the official languages of Sikkim in 1997. Now, it is being taught upto primary level in the schools of Sikkim. The Sunuwars believe that Karna Bahadur Jehntich, an Army Officer engaged in army signals developed the script combining wireless signals and Brahmi script in the 19th century. The script has three vowels and 16 consonants. A considerable effort is yet to be put in for further development of writing systems in this script (personal interview with many Sunuwars; Mukhia, R.B. in Lipi Sangalo (ed) by Limboo, R.B. 1998: 113-139)

Figure 6: A description of the early Jenticha script in a book about Sikkim (from Subba 2008: 95).

किराँती-कोँइच ब्रे:स (जेँ:तिचा लिपि) को मोडिफाइड र रिफर्म्ड भर्सन नमूना ट्रान्स्लिटरेशन सहित

Consonants व्यन्जन वर्ण

꠰	ꠤ	ꠥ	ꠦ	ꠧ
/k,क/	/kʰ,ख/	/g,ग/	/ŋ,ङ/	/ʔ,अ॒/
꠨		꠩		
/c,च/		/dz,ज/		
꠪		꠫		
/T,ट/		/Tʰ,ठ/		
꠬	꠭	꠮		꠯
/t,त/	/th,थ/	/d,द/		/n,न/
꠰	꠱	꠲		꠳
/p,प/	/pʰ,फ/	/b,ब/		/m,म/
꠴	꠵	꠶		꠷
/y,य/	/r,र/	/l,ल/		/w,व/
	꠸	꠹		꠺
	/s,स/	/š,श/		/h,ह/

Vowels स्वर वर्ण

꠻			2
/i,इ/			/u,उ/
		꠼	0(◌)*
		/ə,अ/	/o,आ/
꠽			꠾
/ε,ए/			/ā,आ/

----

*means optional and /ε, ā, i, u/ are also subject to have such lengthening phonetically only and all 6 vowels have nasal differences

(स्रोतः रापचा ई.२००१ए २००२डी)

Figure 7: Consonants and vowels of modern Jenticha (from Rāpacā 2009: 143).

किराँती-कोँइच ब्रे:स (जेँ:तिचा लिपि) को अङ्क लेखन विधिमा कुनै परिमार्जन नगरिएको सबै भन्दा बायाँ

0	꠰2꠷	*sum*	/sum/	'zero'	शुन्य
1	꠰꠾	*kA*	/kā/	'one'	एक
2	꠯꠻ꠧ꠸꠻	*niʔsi*	/niʔsi/	'two'	दुई
3	꠰꠾̈	*sAN*	/sā/	'three'	तीन
4	꠷꠽	*le*	/lε/	'four'	चार
5	ꠦ꠼	*nga*	/ŋə/	'five'	पाँच
6	꠺꠰2	raku	/rəku/	'six'	६
7	꠨ꠧ꠻	*tsani*	/cəni/	'seven'	सात
8	꠸꠸꠻	*sasi*	/səsi/	'eight'	आठ
9	꠴꠾̈	*yAN*	/yā/	'nine'	नौ
10	ꠥꠦ2	gau	/gəu/	'ten'	दश

Figure 8: Digits of modern Jenticha (from Rāpacā 2009: 144).

Figure 9: A hand-written edition of the *Sikkim Herald*, 2005.

Figure 10: A digitally typeset edition of the *Sikkim Herald*, 2009.

Figure 11: Excerpt from the *Elementary Reader of Sunuwar* showing example words beginning with ⅄ LA and Ν VA (from Mukhia, Mukhia, and Rapaca 2004: 40).

तालिका घ
किराँती-कोँइच ब्रे:समा लेखन नमूना ट्रान्सलिटरेसन सहित

�ꞧꞧꞏ ꞧꞧꞏ
ləlmits məmibə
<ललमिच् ममिब:>

ꞧꞏ ꞧꞏ
gom tənātən
<गोम् तनातन्>

ꞧꞏ ꞧꞏ ꞧꞏ
rāgiā ləlmits rəwākəli
<रागीआ ललमिच् रवाकलि>

ꞧꞏ ꞧꞏ ꞧꞏ ꞧꞏ
rɛum tupšo yo məməi
<रेउम् तुप्शो यो ममइ>

ꞧꞏ ꞧꞏ ꞧꞏ ꞧꞏ
pʰəsim hɛmšoyo məməi
<फश्यीम् हेम्शो यो ममइ>

ꞧꞏ ꞧꞏ ꞧꞏ ꞧꞏ
nām k´išo yo məməi
<नाम की'शो या ममइ>

ꞧꞏ ꞧꞏ
gom təšonu
<गोम् तशोनु>

ꞧꞏ ꞧꞏ ꞧꞏ
mɛko ləlmits rəwā
<मेको ललमिच् रवा>

ꞧꞏ ꞧꞏ
boisšon bāʔbə
<बोइस्शोन् बाअ्ब:>

ꞧꞏ ꞧꞏ
s´išon bāʔbə
<सी'शोन् बाअ्ब:>

ꞧꞏ ꞧꞏ ꞧꞏ
kā pāts ā ts´ibpiki
<का पाच्आ ची'बपिकी>

ꞧꞏ ꞧꞏ
mɛkɛn bɛrnimi
<मेकेन् बेर्नमी>

ꞧꞏ ꞧꞏ ꞧꞏ
mɛkɛn nā thɛrnimi
<मेकेन् ना थेर्नमी>

ꞧꞏ ꞧꞏ
mɛkɛn dz´unimi
<मेकेन् जू'निमी>

ꞧꞏ
minu
<मिनु>

ꞧꞏ ꞧꞏ ꞧꞏ ꞧꞏ ꞧꞏ ꞧꞏ
kumšo pəšyā šyil pāinimi
<कुम्शो पश्या श्यील् पाइनिमी>

ꞧꞏ ꞧꞏ ꞧꞏ
šyɛŋ mɛko rāgin ā
<श्यङ् मेको रागीन्आ>

ꞧꞏ ꞧꞏ
wɛk pāts ā
<वेक् पाच्आ>

ꞧꞏ ꞧꞏ
ts´ib ānkəli
<ची:बआन्कलि>

ꞧꞏ ꞧꞏ ꞧꞏ ꞧꞏ ꞧꞏ ꞧꞏ
ãkɛ rāgimi gɛnāiyo ləlmits məmibə
<आँके रागीमि गेनाइयो ललमिच् ममिब:>

(Source: Rapacha (2001a: 24)

Figure 12: Sunuwar phrases given in Jenticha, Latin, and Devanagari transliteration (adapted from Rāpacā 2009: 145–146).

Figure 13: An excerpt from a 2009 wall calendar in Jenticha published by the Sikkim Sunuwar Mukhia Koinchbu.

Figure 14: An excerpt from a 2010 wall calendar in Jenticha published by the Sikkim Sunuwar Mukhia Koinchbu.

ISO/IEC JTC 1/SC 2/WG 2
PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS
FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646[1]
Please fill all the sections A, B and C below.
Please read Principles and Procedures Document (P & P) from http://www.dkuug.dk/JTC1/SC2/WG2/docs/principles.html **for guidelines and details before filling this form.**
Please ensure you are using the latest Form from http://www.dkuug.dk/JTC1/SC2/WG2/docs/summaryform.html**.**
See also http://www.dkuug.dk/JTC1/SC2/WG2/docs/roadmaps.html **for latest** *Roadmaps.*

A. Administrative

1. **Title:** *Proposal to Encode the Jenticha Script in ISO/IEC 10646*
2. Requester's name: *Script Encoding Initiative (SEI) / Anshuman Pandey (pandey@umich.edu)*
3. Requester type (Member body/Liaison/Individual contribution): *Liaison contribution*
4. Submission date: *2011-05-31*
5. Requester's reference (if applicable):
6. Choose one of the following:
    This is a complete proposal: *Yes*
    (or) More information will be provided later:

B. Technical – General

1. Choose one of the following:
    a. This proposal is for a new script (set of characters): *Yes*
        Proposed name of script: *Jenticha*
    b. The proposal is for addition of character(s) to an existing block:
        Name of the existing block:
2. Number of characters in proposal: *45*
3. Proposed category (select one from below - see section 2.2 of P&P document):
A-Contemporary   X   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
4. Is a repertoire including character names provided? *Yes*
    a. If YES, are the names in accordance with the "character naming guidelines"
        in Annex L of P&P document? *Yes*
    b. Are the character shapes attached in a legible form suitable for review? *Yes*
5. Fonts related:
    a. Who will provide the appropriate computerized font to the Project Editor of 10646 for publishing the standard?
    Anshuman Pandey
    b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.):
    Anshuman Pandey (pandey@umich.edu)
6. References:
    a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided? *Yes*
    b. Are published examples of use (such as samples from newspapers, magazines, or other sources)
    of proposed characters attached? *Yes*
7. Special encoding issues:
    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. Additional Information:

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/reports/tr44/) and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.

---

[1] Form number: N3902-F (Original 1994-10-14; Revised 1995-01, 1995-04, 1996-04, 1996-08, 1999-03, 2001-05, 2001-09, 2003-11, 2005-01, 2005-09, 2005-10, 2007-03, 2008-05, 2009-11, 2011-03)

C. Technical - Justification

1. Has this proposal for addition of character(s) been submitted before? — *No*
    If YES explain

2. 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*
        If YES, with whom? — *Dr. Lal-Shyakarelu Rapacha*
        If YES, available relevant documents: — *See text of proposal*

3. Information on the user community for the proposed characters (for example:
    size, demographics, information technology use, or publishing use) is included? — *Yes*
    Reference: — *Size of user community is unknown. Script is used for print and digital publications.*

4. The context of use for the proposed characters (type of use; common or rare) — *Common*
    Reference: — *The Jenticha script is used for writing Sunuwar. See text of proposal for details.*

5. Are the proposed characters in current use by the user community? — *Yes*
    If YES, where? Reference: — *SIkkim (India) and Nepal. See text of proposal for details.*

6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely
    in the BMP? — *No*
            If YES, is a rationale provided?
            If YES, reference:

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

8. Can any of the proposed characters be considered a presentation form of an existing
    character or character sequence? — *No*
            If YES, is a rationale for its inclusion provided?
            If YES, reference:

9. Can any of the proposed characters be encoded using a composed character sequence of either
    existing characters or other proposed characters? — *No*
        If YES, is a rationale for its inclusion provided?
        If YES, reference:

10. Can any of the proposed character(s) be considered to be similar (in appearance or function)
    to an existing character? — *Yes*
        If YES, is a rationale for its inclusion provided? — *Yes*
        If YES, reference: — *See text of proposal.*

11. Does the proposal include use of combining characters and/or of composite sequences? — *Yes*
    If YES, is a rationale for such use provided? — *Yes*
        If YES, reference: — *Combining signs*
    Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?
        If YES, reference:

12. Does the proposal contain characters with any special properties such as
    control function or similar semantics? — *Yes*
        If YES, describe in detail (include attachment if necessary) — *Virama*
                        See text of proposal

13. Does the proposal contain any Ideographic compatibility characters? — *No*
    If YES, are the equivalent corresponding unified ideographic characters identified?
        If YES, reference: