

# 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 कौँच *kōica*, which is normalized as ‘Kōits’. The language is called कौँच लो *kōica lo* ‘Kōits Lo’ or किराँती-कौँच *kirāñī-kōica* ‘Kiranti-Kōits’. The name ‘Kōits’ 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 जैँतिच ब्रेसे *jētica brese*, ‘Jenticha Brese’. The script is also called कौँच ब्रेसे *kōica brese* ‘Kōits Brese’, or the ‘Kōits 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 predecessor 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 Kōic Bu in Sikkim. Examples of script primers are  $\text{ཁྱེད་ཀྱི་ཆུང་མ་གྱི་སྐད་ཀྱི་རྒྱུ་རྒྱུ་}$  *Kirāñi-Kōica Blesethekā* (2003) and  $\text{ཁྱེད་ཀྱི་ཆུང་མ་གྱི་སྐད་ཀྱི་རྒྱུ་རྒྱུ་}$  *Kōicā The:si Khyōpat* (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 *vamśāvalī*, were written in Jenticha beginning in the 1940s.
- *Newspapers* A Sunuwar (‘Mukhiya’) edition of the *Sikkim Herald* ( $\text{སྐྱེད་ཀྱི་སྐད་ཀྱི་ཆུང་མ་གྱི་སྐད་ཀྱི་རྒྱུ་རྒྱུ་}$ ) 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-Śyāmkārelu 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 Nāvacā-Mukhiyā, Raghuvīra Rujicā-Mukhiya, and Lāla Rāpacā (Sikkim: Kōica 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	†	I	/i/	इ	±	E	/e/	ए
॒	AA	/a:/	आ	॒	U	/u/	उ	0	O	/o/	ओ

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

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

/ai/	॑†	॑ A + † I	ऐ	/eu/	±॒	± E + ॒ U	एउ	/oi/	0†	0 O + † I	ओइ
/au/	॒॑	॑ A + ॒ U	औ	/ui/	॒†	॒ U + † I	उइ				

### 3.5 Vowel Length

Vowel length is indicated using the sign ◌◌: 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:

𑄏	KA	/k/	क	𑄑	PA	/p/	प
𑄎	KHA	/k <sup>h</sup> /	ख	𑄒	PHA	/p <sup>h</sup> /	फ
𑄙	GA	/g/	ग	𑄔	BA	/b/	ब
𑄚	NGA	/ŋ/	ङ	𑄕	MA	/m/	म
𑄛	GLOTTAL STOP	/ʔ/	अ्	𑄖	YA	/y/	य
𑄜	CA	/tʃ/	च	𑄗	RA	/r/, [r]	र
𑄝	JA	/dʒ/	ज	𑄘	LA	/l/, [l]	ल
𑄞	TTA	/t̪/	ट	𑄙	VA	/v/	व
𑄟	TTHA	/t̪ <sup>h</sup> /	ठ	𑄚	SHA	/ʃ/	श, ष
𑄠	TA	/t/	त	𑄛	SA	/s/	स
𑄡	THA	/t <sup>h</sup> /	थ	𑄜	HA	/h/	ह
𑄢	DA	/d/	द, ङ	𑄝	HHA	/ɦ/	व्हो, ह्व
𑄣	NA	/n/	न, ण				

In modern Jenticha, each consonant letter, except for 𑄛 GLOTTAL STOP, bears the inherent vowel /a/, which corresponds to the letter 𑄑 A. The inherent vowel is changed by writing a vowel letter after the consonant, eg. 𑄏 /ka/, 𑄏𑄑 /ka:/. It is silenced by writing 𑄒 VIRAMA under a consonant, eg. <𑄏 KA + 𑄒> → <𑄏< /k/.

In early Jenticha, consonant letters are purely alphabetic and do not have an inherent vowel. The sequence <𑄏 KA + 𑄑 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<sub>1</sub> + 𑄒 VIRAMA + C<sub>2</sub>>.

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 𑄚 U+200D ZERO WIDTH JOINER should be used to force the alternate form: <C<sub>1</sub> + VIRAMA + 𑄚 ZWJ + C<sub>2</sub>>.

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

	Regular form	Default bare form	Half-form	Devanagari half-form
CA	<	<̣	<	च
MA	𑄕	𑄕̣	𑄕	फ
SA	𑄛	𑄛̣	𑄛	स

2. *Above-Base Forms* In modern orthography, the letter 𑄗 RA has an alternate contextual form when it occurs as C<sub>1</sub> in conjuncts. The default behavior for RA as C<sub>1</sub> is with visible *virāma*: 𑄗< rca. The alternate is the form 𑄗̣ *sangkirs*, which is written above C<sub>2</sub>: 𑄗̣ rca. It corresponds to *repha* in Devanagari. The encoded representation for *sangkirs* is <RA + VIRAMA + 𑄚 ZWJ + C>.

### 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  $\overset{\circ}{\text{C}}$  CANDRABINDU, which is called *taslathenk* in Sunuwar. This sign corresponds to  $\overset{\circ}{\text{C}}$  CANDRABINDU in Devanagari.

### 3.11 Aspiration

Early Jenticha did not have unique letters for the aspirated consonants /k<sup>h</sup>/, /t<sup>h</sup>/, /tʰ/, /p<sup>h</sup>/ (see Figure 5). Aspiration was indicated by writing 𑌵 HA after an unaspirated consonant letter, eg. <𑌵 PA + 𑌵 HA> → <𑌵𑌵> /p<sup>h</sup>/. The modern script has independent letters for the four aspirated consonants.

### 3.12 Trill Mark

The  $\overset{\circ}{\text{C}}$  TRILL MARK represents [l̥], which is an allophone of /l/ when it occurs as C<sub>2</sub> in a consonant conjunct with a velar or labial consonant as C<sub>1</sub>. 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 𑌵𑌵𑌵𑌵 *blese* may be written using either 𑌵 RA or 𑌵 LA; eg. 𑌵𑌵𑌵𑌵 *brese*, 𑌵𑌵𑌵𑌵 *blese*. In Devanagari orthography, [l̥] is represented using र RA, ल LA, or ड DDA. Valid sequences are:

𑌵	KA	+	$\overset{\circ}{\text{C}}$	→	𑌵	/k/	कड, क्र, क्ल	𑌵	PA	+	$\overset{\circ}{\text{C}}$	→	𑌵	/p/	पड, प्र, प्ल
𑌵	KHA	+	$\overset{\circ}{\text{C}}$	→	𑌵	/k <sup>h</sup> /	खड, ख्र, ख्ल	𑌵	PHA	+	$\overset{\circ}{\text{C}}$	→	𑌵	/p <sup>h</sup> /	फड, फ्र, फ्ल
𑌵	GA	+	$\overset{\circ}{\text{C}}$	→	𑌵	/g/	गड, ग्र, ग्ल	𑌵	BA	+	$\overset{\circ}{\text{C}}$	→	𑌵	/b/	बड, ब्र, ब्ल

### 3.13 Digits

Digits represent decimal numbers: 𑌵 ZERO, 𑌵 ONE, 𑌵 TWO, 𑌵 THREE, 𑌵 FOUR, 𑌵 FIVE, 𑌵 SIX, 𑌵 SEVEN, 𑌵 EIGHT, 𑌵 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. <ᳵᳶ> 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:

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](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 Raghbir Mukhia, and Lal Rapaca. 2004. ་འཕྲིན་ལྷོ་མཚོ་ལྷོ་མཚོ་ [= *Elementary Reader of Sunuwar*]. 1st ed. Sikkim: Kōica 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-श्यांकरेलु [Rāpacā, Lāl-Śyāmkārelu]. 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

I would like to extend my gratitude to Lāl-Śyāmkārelu Rāpacā of Kathmandu for providing details about the history and orthography of the Jenticha script. I would also like to thank Krishna B. Sunuwar of Kathmandu for information about the scripts used for writing Sunuwar in Sikkim and Nepal.

This project was made possible in part by a grant from the United States National Endowment for the Humanities, which funded the Universal Scripts Project (part of the Script Encoding Initiative at the University of California, Berkeley). Any views, findings, conclusions or recommendations expressed in this publication do not necessarily reflect those of the National Endowment of the Humanities.

	116D	116E	116F
0	𑌀 116D0	𑌁 116E0	𑌂 116F0
1	𑌃 116D1	𑌄 116E1	𑌅 116F1
2	𑌆 116D2	𑌇 116E2	𑌈 116F2
3	𑌉 116D3	𑌊 116E3	
4	𑌋 116D4	𑌌 116E4	
5	𑌍 116D5	𑌎 116E5	
6	𑌏 116D6	𑌐 116E6	𑌑 116F6
7	𑌒 116D7	𑌓 116E7	𑌔 116F7
8	𑌕 116D8	𑌖 116E8	𑌗 116F8
9	𑌘 116D9	𑌙 116E9	𑌚 116F9
A	𑌛 116DA	𑌜 116EA	𑌝 116FA
B	𑌞 116DB	𑌟 116EB	𑌠 116FB
C	𑌡 116DC	𑌢 116EC	𑌣 116FC
D	𑌦 116DD	𑌧 116ED	𑌨 116FD
E	𑌩 116DE	𑌪 116EE	𑌫 116FE
F	𑌭 116DF	𑌮 116EF	𑌯 116FF

**Vowels**

116D0	𑌀	JENTICHA LETTER A
116D1	𑌃	JENTICHA LETTER AA
116D2	𑌆	JENTICHA LETTER I
116D3	𑌉	JENTICHA LETTER U
116D4	𑌋	JENTICHA LETTER E
116D5	𑌍	JENTICHA LETTER O

**Consonants**

116D6	𑌏	JENTICHA LETTER KA
116D7	𑌒	JENTICHA LETTER KHA
116D8	𑌕	JENTICHA LETTER GA
116D9	𑌘	JENTICHA LETTER NGA
116DA	𑌛	JENTICHA LETTER CA
116DB	𑌞	JENTICHA LETTER JA
116DC	𑌡	JENTICHA LETTER TTA
116DD	𑌦	JENTICHA LETTER TTHA
116DE	𑌩	JENTICHA LETTER TA
116DF	𑌭	JENTICHA LETTER THA
116E0	𑌐	JENTICHA LETTER DA
116E1	𑌓	JENTICHA LETTER NA
116E2	𑌇	JENTICHA LETTER PA
116E3	𑌊	JENTICHA LETTER PHA
116E4	𑌌	JENTICHA LETTER BA
116E5	𑌎	JENTICHA LETTER MA
116E6	𑌑	JENTICHA LETTER YA
116E7	𑌓	JENTICHA LETTER RA
116E8	𑌖	JENTICHA LETTER LA
116E9	𑌙	JENTICHA LETTER VA
116EA	𑌜	JENTICHA LETTER SHA
116EB	𑌟	JENTICHA LETTER SA
116EC	𑌢	JENTICHA LETTER HA
116ED	𑌧	JENTICHA LETTER HHA
116EE	𑌪	JENTICHA LETTER GLOTTAL STOP

**Various signs**

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

**Digits**

116F6	𑌑	JENTICHA DIGIT ZERO
116F7	𑌔	JENTICHA DIGIT ONE
116F8	𑌗	JENTICHA DIGIT TWO
116F9	𑌚	JENTICHA DIGIT THREE
116FA	𑌝	JENTICHA DIGIT FOUR
116FB	𑌠	JENTICHA DIGIT FIVE
116FC	𑌣	JENTICHA DIGIT SIX
116FD	𑌨	JENTICHA DIGIT SEVEN
116FE	𑌫	JENTICHA DIGIT EIGHT
116FF	𑌯	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).

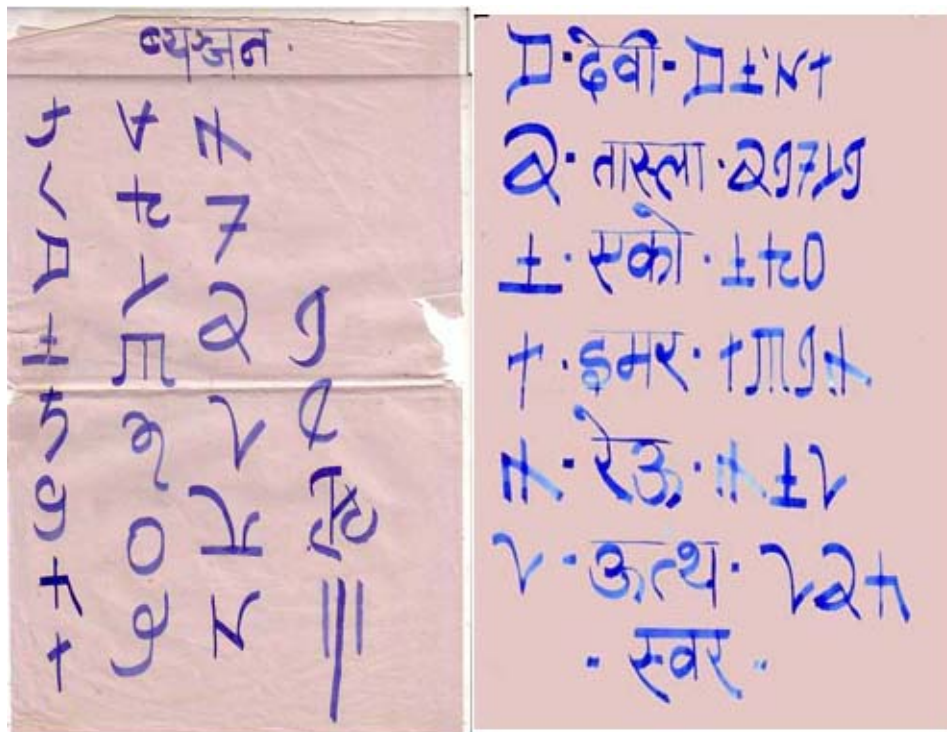


Figure 3: Early Jenticha letterforms (from kiratsunuwar.org.np). Notice the use of ॐ A for writing both the inherent vowel /a/ and /a/; eg. in the words व. व. व. /tasla/ and म. म. म. /imar/.

Sunwar	Deva nā- garī	Roman	Sunwar	Deva- nāgarī	Roman/ Arabic
𑌧	अ	a	𑌨	उ ऊ	u, ū
𑌩	ब	b	𑌪	व	v
𑌫	च	c	𑌬	ळो	vho
𑌭	द	d	𑌮	य	y
𑌰	ए े	e, e	𑌱	०	0
𑌲	ग	g	𑌳	१	1
𑌴	ह	h	𑌵	२	2
𑌷	इ ई	i, ī	𑌸	३	3
𑌹	ज	j	𑌺	४	4
𑌻	क	k	𑌼	५	5
𑌽	ल	l	𑌾	६	6
𑌿	म	m	𑌽	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).



## तालिका क

किराँती-कौइच ब्रे:स (जै:तिचा लिपि) को मोडिफाइड र रिफाइन्ड भर्सन नमूना ट्रान्सलिटरेशन सहित

### Consonants व्यन्जन वर्ण

क	ख	ग	ङ	च
/k,क/	/kʰ,ख/	/g,ग/	/ŋ,ङ/	/t,च/
क	ख	ज		
/c,क/		/dz,ज/		
स		ड		
/ʈ,स/		/ʈʰ,ड/		
त	थ	द	न	
/t,त/	/tʰ,थ/	/d,द/	/n,न/	
प	फ	ब	म	
/p,प/	/pʰ,फ/	/b,ब/	/m,म/	
र	ल		व	
/r,र/	/l,ल/		/w,व/	
स	श		ह	
/s,स/	/ʃ,श/		/h,ह/	

### Vowels स्वर वर्ण

इ	उ
/i,इ/	/u,उ/
	अ
	/ə,अ/
ए	ऐ
/e,ए/	/ā,आ/

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

(स्रोत: रापचा ई.२००१ए २००२डी)

नोट: किराँती-कौइचका यी वर्णहरू लघुतम युग्म जाँच (Minimal Pair Test- तुल- रापचा २००४, पनि) र प्रोडक्टिभिटी उत्पादकत्व कम मेजोरिटी (Productivity cum majority) सिद्धान्तका आधारमा इन्भेन्ट गरिएको छ।

॥१४३॥

### तालिका ख

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

०	१२३४५६७८९०	sum /sum/ 'zero' शुन्य
१	१२३४५६७८९०	ka /kā/ 'one' एक
२	१२३४५६७८९०	nirsi /nirsi/ 'two' दुई
३	१२३४५६७८९०	sān /sā/ 'three' तीन
४	१२३४५६७८९०	le /le/ 'four' चार
५	१२३४५६७८९०	nga /ŋə/ 'five' पाँच
६	१२३४५६७८९०	raku /raku/ 'six' छ
७	१२३४५६७८९०	tsani /cəni/ 'seven' सात
८	१२३४५६७८९०	sasi /səsi/ 'eight' आठ
९	१२३४५६७८९०	yān /yā/ 'nine' नौ
१०	१२३४५६७८९०	gau /gəu/ 'ten' दश

### तालिका ग

किराँती-कौइच (सुनुवार/मुखिया/राई) भाषा लेखनका लागि शुल्ज (ई. १९९५: २) द्वारा प्रस्तावित देवनागरी लिपि

क	ख	ग	(घ)	ङ
च	छ	ज	झ	(ञ)
ट	ठ	(ड)	(ढ)	(ण)
त	थ	द	(ध)	न
प	फ	ब	(भ)	म









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

तालिका घ

लेखन नमूना ट्रान्सलिटरेसन सहित  
 <मेकेन् ना थेरनमी>  
 मेकेन् ना थेरनमी

ललमिच ममिबः  
 ləlmits məmibə  
 <ललमिच ममिबः>

गोम तनातन्  
 gom tənātən  
 <गोम तनातन्>

रागीआ ललमिच रवाकलि  
 rāgiā ləlmits rəwākəli  
 <रागीआ ललमिच रवाकलि>

रेउम् तुषो यो ममइ  
 reum tupšo yo məməi  
 <रेउम् तुषो यो ममइ>

फश्यीम् हेमशो यो ममइ  
 pʰəsīm hēmšoyo məməi  
 <फश्यीम् हेमशो यो ममइ>

नाम की'शो या ममइ  
 nām k'īšo yo məməi  
 <नाम की'शो या ममइ>

गोम तशोनु  
 gom təšonu  
 <गोम तशोनु>

मेको ललमिच रवा  
 meko ləlmits rəwā  
 <मेको ललमिच रवा>

बोइस्शोन् बाअबः  
 boišson bāʔbə  
 <बोइस्शोन् बाअबः>

सी'शोन् बाअबः  
 s'īšon bāʔbə  
 <सी'शोन् बाअबः>

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

मेकेन् बेरनमी  
 meken bernimi  
 <मेकेन् बेरनमी>

मेकेन् ना थेरनमी  
 meken nā thernimi

<मेकेन् ना थेरनमी>

मेकेन् जू'निमी  
 meken dz'unimi  
 <मेकेन् जू'निमी>

मिनु  
 minu  
 <मिनु>

कुम्शो पश्या शयील् पाइनिमी  
 kumšo pəšyā šyil pāinimi  
 <कुम्शो पश्या शयील् पाइनिमी>

श्येण मेको रागीन्आ  
 šyeṅ meko rāgin ā  
 <श्येण मेको रागीन्आ>

वेक् पाच्आ  
 wek pāts ā  
 <वेक् पाच्आ>

चीःबआन्कलि  
 ts'ib ānkəli  
 <चीःबआन्कलि>

आँके रागीमि गेनाइयो ललमिच ममिबः  
 āke rāgimi genāiyo ləlmits məmibə  
 <आँके रागीमि गेनाइयो ललमिच ममिबः>

(Source: Rapacha (2001a: 24))

<मेकेन् ना थेरनमी>

मेकेन् जू'निमी  
 meken dz'unimi  
 <मेकेन् जू'निमी>

मिनु  
 minu  
 <मिनु>

कुम्शो पश्या शयील् पाइनिमी  
 kumšo pəšyā šyil pāinimi  
 <कुम्शो पश्या शयील् पाइनिमी>

श्येण मेको रागीन्आ  
 šyeṅ meko rāgin ā  
 <श्येण मेको रागीन्आ>

वेक् पाच्आ  
 wek pāts ā  
 <वेक् पाच्आ>

चीःबआन्कलि  
 ts'ib ānkəli  
 <चीःबआन्कलि>

आँके रागीमि गेनाइयो ललमिच ममिबः  
 āke rāgimi genāiyo ləlmits məmibə  
 <आँके रागीमि गेनाइयो ललमिच ममिबः>

(Source: Rāpacā 2009: 145–146)

स्रोतः दुमाकेतु (ई. २००६) वर्ष ३ अंक ३ पाना १५-२०  
 स्रोतः दुमाकेतु (ई. २००६) वर्ष ३ अंक ३ पाना १५-२०

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<sup>1</sup>**

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

**B. Technical – General**

1. Choose one of the following:		
a. This proposal is for a new script (set of characters):	<input checked="" type="checkbox"/> Yes	
Proposed name of script:	<i>Jenticha</i>	
b. The proposal is for addition of character(s) to an existing block:		
Name of the existing block:		
2. Number of characters in proposal:	<i>45</i>	
3. Proposed category (select one from below - see section 2.2 of P&P document):		
A-Contemporary <input checked="" type="checkbox"/>	B.1-Specialized (small collection) <input type="checkbox"/>	B.2-Specialized (large collection) <input type="checkbox"/>
C-Major extinct <input type="checkbox"/>	D-Attested extinct <input type="checkbox"/>	E-Minor extinct <input type="checkbox"/>
F-Archaic Hieroglyphic or Ideographic <input type="checkbox"/>	G-Obscure or questionable usage symbols <input type="checkbox"/>	
4. Is a repertoire including character names provided?	<input checked="" type="checkbox"/> Yes	
a. If YES, are the names in accordance with the "character naming guidelines" in Annex L of P&P document?	<input checked="" type="checkbox"/> Yes	
b. Are the character shapes attached in a legible form suitable for review?	<input checked="" type="checkbox"/> Yes	
5. Fonts related:		
a. Who will provide the appropriate computerized font to the Project Editor of 10646 for publishing the standard?	<i>Anshuman Pandey</i>	
b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.):	<i>Anshuman Pandey (pandey@umich.edu)</i>	
6. References:		
a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?	<input checked="" type="checkbox"/> Yes	
b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?	<input checked="" type="checkbox"/> 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)?	<input checked="" type="checkbox"/> 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.

<sup>1</sup> 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? If YES explain	No
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.)? If YES, with whom? If YES, available relevant documents:	Yes <i>Dr. Lal-Shyakarelu Rapacha</i> <i>See text of proposal</i>
3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included? Reference:	Yes <i>Size of user community is unknown. Script is used for print and digital publications.</i>
4. The context of use for the proposed characters (type of use; common or rare) Reference:	Common <i>The Jenticha script is used for writing Sunuwar. See text of proposal for details.</i>
5. Are the proposed characters in current use by the user community? If YES, where? Reference:	Yes <i>Sikkim (India) and Nepal. See text of proposal for details.</i>
6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP? If YES, is a rationale provided? If YES, reference:	No
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? If YES, is a rationale for its inclusion provided? If YES, reference:	No
9. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters? If YES, is a rationale for its inclusion provided? If YES, reference:	No
10. Can any of the proposed character(s) be considered to be similar (in appearance or function) to an existing character? If YES, is a rationale for its inclusion provided? If YES, reference:	Yes Yes <i>See text of proposal.</i>
11. Does the proposal include use of combining characters and/or use of composite sequences? If YES, is a rationale for such use provided? If YES, reference: Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided? If YES, reference:	Yes Yes <i>Combining signs</i>
12. Does the proposal contain characters with any special properties such as control function or similar semantics? If YES, describe in detail (include attachment if necessary)	Yes <i>Virama</i> <i>See text of proposal</i>
13. Does the proposal contain any Ideographic compatibility characters? If YES, are the equivalent corresponding unified ideographic characters identified? If YES, reference:	No