

Title: Preliminary Proposal to Encode the Jenticha Script in ISO/IEC 10646
Author: Anshuman Pandey (pandey@umich.edu)
Status: Individual Contribution
Action: For consideration by UTC and WG2
Date: 2010-11-05

1 Introduction

This is a preliminary proposal to encode the Jenticha script in the Universal Character Set (ISO/IEC 10646). Jenticha is used in Sikkim, India, for writing the Sunuwar (Kõits) language. This document provides a basic description of the writing system, a preliminary code chart and names list, character properties, and a few specimens. Additional research on the script is required and is being performed by the present author.

There is another script used in Nepal for writing Sunuwar, which is called ‘Tikamuli’. A preliminary proposal (N3963) to encode Tikamuli in the UCS has been submitted by the present author to the Unicode Technical Committee (UTC) and ISO/IEC JTC1/SC2/WG2 for review.

2 Background

Jenticha is a script used for writing Sunuwar [suz], a Tibeto-Burman language of the Kiranti family that is spoken in eastern Nepal and the state of Sikkim in India. Sunuwar is spoken by forty thousand persons in Nepal and it was recognized as an official language of Sikkim in 1996. Jenticha is a recognized script in Sikkim. The Devanagari script is the most commonly used script for writing Sunuwar in both Nepal and Sikkim.

Sunuwar (सुनुवार *sunuvāra*), also spelled ‘Sunwar’, is also known by the name ‘Mukhiya’ (मुखिया *mukhiyā*); both are Nepali exonyms for the language and its speakers. The autonym is कौइच *kōica*, which is normalized as ‘Kõits’. The language is called कौइच लो *kōica lo* ‘Kõits Lo’ or किराँती-कौइच *kirāṭī-kōica* ‘Kiranti-Kõits’. The name ‘Kõits’ is also romanized as ‘Koinch’, ‘Koincha’, ‘Koints’, etc.

The Jenticha script is named after Krishna Bahadur Jenticha (1923–1990).¹ It is believed that K. B. Jenticha developed the script in 1942. It is known indigenously as जैतिच ब्रेसे *jētica brese*, ‘Jenticha Brese’. The script is also called कौइच ब्रेसे *kōic brese* ‘Kõits Brese’, or the ‘Kõits script’; ‘brese’, or script, is also romanized as ‘brehs’, ie. ‘Koinch Brehs’. A comparison of Jenticha letters in Dörte Borchers (2008) and Lāla-Śyāṃkārelu Rāpacā (2009) suggest that the script has evolved since its creation (see Figure 2 and Figure 5).

The script is used at present in Sikkim. The Sunuwar edition of the *Sikkim Herald* is written in Jenticha.² There is also active scholarly study of the Jenticha script (cf. Rāpacā 2009).

A digitized font for Jenticha named ‘Kirat1’ was developed by Shyan Kirat Rai. It is used by Rāpacā in the charts and examples shown in figures 2–4. Another font has been designed by the present author. It is used in throughout this document and in the code chart in Figure 1.

¹ The name is also given as ‘Karna Bahadur Jentich’. ² Lāla-Śyāṃkārelu Rāpacā, personal correspondence, November 2010.

3 Writing System

Structure Jentich is a semi-alphabetic writing system. There are independent letters for vowels and consonants. Consonants bear an inherent vowel. The script is written from left to right.

Character Repertoire The script consists of 44 characters: 6 vowels, 25 consonants, 3 signs, and 10 digits. A tentative code chart and names list is given in Figure 1.

Vowels There are 6 vowel letters: 𑄀 A, 𑄁 AA, 𑄂 I, 𑄃 U, 𑄄 E, 𑄅 O.

Vowel Length Vowel lengthening is indicated by writing 𑄆: VOWEL LENGTH MARK after a vowel letter. eg. 𑄂: = *ī*. This length mark may only be written with I, U, E, and O; the long form of 𑄀 A is the independent vowel 𑄁 AA.

Diphthongs Diphthongs are indicating using a sequence of vowel letters, eg. 𑄀𑄂 (AA + I) *ai*.

Consonants There are 25 consonant letters, including a glottal stop (see Figure 1). Each consonant letter, except for the glottal stop, bears the inherent vowel 𑄀 A [ə]. For these consonants, the inherent vowel is changed by writing a vowel letter after the letter, eg. 𑄂 *ka*, 𑄂𑄁 *kā*, etc. The inherent vowel may be silenced by writing 𑄇 NULL VOWEL beneath the consonant, eg. 𑄂𑄇 *k*. Geminated consonants are not indicated orthographically.

The table in Figure 6 shows an older orthography for writing aspirated sounds. It appears that aspiration was represented by writing 𑄈 HA after the letter for the unaspirated consonant, eg. 𑄀𑄈 (PA + HA) [p^h]. The version of Jenticha shown in Rāpacā (2009) has a unique letter for this sound: 𑄉 PHA.

Null Vowel The 𑄇 NULL VOWEL behaves similar to VIRAMA in Brahmic scripts. It is always rendered visibly in Jenticha.

Nasalization Nasalization is indicated by writing 𑄊 NASAL SIGN above a vowel or consonant.

Digits The script has digits for representing numbers of the decimal system: 𑄋 ZERO, 𑄌 ONE, 𑄍 TWO, 𑄎 THREE, 𑄏 FOUR, 𑄐 FIVE, 𑄑 SIX, 𑄒 SEVEN, 𑄓 EIGHT, 𑄔 NINE.

Punctuation There are no native punctuation marks.

Collation The sort order follows the encoding order.

4 Relationship to Other Scripts

Jenticha has no genetic relationship to other writing systems; however, some letters resemble those of other scripts. Some characters are derived from the Sirijonga (Limbu) script, used for writing Limbu [lif], also a Kiranti language. For example, Jenticha 𑄁 AA is similar to 𑄁 U+1900 LIMBU VOWEL-CARRIER LETTER; 𑄃 TTA is similar to 𑄃 U+1916 LIMBU LETTER RA; 𑄄 TTHA is similar to 𑄄 U+190F LIMBU LETTER NA; 𑄅 SHA is similar to 𑄅 U+1919 LIMBU LETTER SHA; etc. Others are similar to Latin characters, eg. 𑄅 O resembles ‘O’.

5 Character Properties

Character properties for Jenticha are given below. The use of code points starting at U+11C00 is arbitrary.

```

11C00;JENTICHA LETTER A;Lo;0;L;;;;N;;;;;
11C01;JENTICHA LETTER AA;Lo;0;L;;;;N;;;;;
11C02;JENTICHA LETTER I;Lo;0;L;;;;N;;;;;
11C03;JENTICHA LETTER U;Lo;0;L;;;;N;;;;;
11C04;JENTICHA LETTER E;Lo;0;L;;;;N;;;;;
11C05;JENTICHA LETTER O;Lo;0;L;;;;N;;;;;
11C06;JENTICHA LETTER KA;Lo;0;L;;;;N;;;;;
11C07;JENTICHA LETTER KHA;Lo;0;L;;;;N;;;;;
11C08;JENTICHA LETTER GA;Lo;0;L;;;;N;;;;;
11C09;JENTICHA LETTER NGA;Lo;0;L;;;;N;;;;;
11C0A;JENTICHA LETTER CA;Lo;0;L;;;;N;;;;;
11C0B;JENTICHA LETTER JA;Lo;0;L;;;;N;;;;;
11C0C;JENTICHA LETTER TTA;Lo;0;L;;;;N;;;;;
11C0D;JENTICHA LETTER TTHA;Lo;0;L;;;;N;;;;;
11C0E;JENTICHA LETTER TA;Lo;0;L;;;;N;;;;;
11C0F;JENTICHA LETTER THA;Lo;0;L;;;;N;;;;;
11C10;JENTICHA LETTER DA;Lo;0;L;;;;N;;;;;
11C11;JENTICHA LETTER NA;Lo;0;L;;;;N;;;;;
11C12;JENTICHA LETTER PA;Lo;0;L;;;;N;;;;;
11C13;JENTICHA LETTER PHA;Lo;0;L;;;;N;;;;;
11C14;JENTICHA LETTER BA;Lo;0;L;;;;N;;;;;
11C15;JENTICHA LETTER MA;Lo;0;L;;;;N;;;;;
11C16;JENTICHA LETTER YA;Lo;0;L;;;;N;;;;;
11C17;JENTICHA LETTER RA;Lo;0;L;;;;N;;;;;
11C18;JENTICHA LETTER LA;Lo;0;L;;;;N;;;;;
11C19;JENTICHA LETTER VA;Lo;0;L;;;;N;;;;;
11C1A;JENTICHA LETTER VHA;Lo;0;L;;;;N;;;;;
11C1B;JENTICHA LETTER SHA;Lo;0;L;;;;N;;;;;
11C1C;JENTICHA LETTER SA;Lo;0;L;;;;N;;;;;
11C1D;JENTICHA LETTER HA;Lo;0;L;;;;N;;;;;
11C1E;JENTICHA LETTER GLOTTAL STOP;Lo;0;L;;;;N;;;;;
11C1F;JENTICHA NASAL SIGN;Mn;0;NSM;;;;N;;;;;
11C20;JENTICHA VOWEL LENGTH MARK;Mc;0;L;;;;N;;;;;
11C21;JENTICHA NULL VOWEL;Mn;9;NSM;;;;N;;;;;
11C26;JENTICHA DIGIT ZERO;Nd;0;L;0;0;0;N;;;;;
11C27;JENTICHA DIGIT ONE;Nd;0;L;1;1;1;N;;;;;
11C28;JENTICHA DIGIT TWO;Nd;0;L;2;2;2;N;;;;;
11C29;JENTICHA DIGIT THREE;Nd;0;L;3;3;3;N;;;;;
11C2A;JENTICHA DIGIT FOUR;Nd;0;L;4;4;4;N;;;;;
11C2B;JENTICHA DIGIT FIVE;Nd;0;L;5;5;5;N;;;;;
11C2C;JENTICHA DIGIT SIX;Nd;0;L;6;6;6;N;;;;;
11C2D;JENTICHA DIGIT SEVEN;Nd;0;L;7;7;7;N;;;;;
11C2E;JENTICHA DIGIT EIGHT;Nd;0;L;8;8;8;N;;;;;
11C2F;JENTICHA DIGIT NINE;Nd;0;L;9;9;9;N;;;;;

```

6 Issues

Script Name The script is tentatively assigned the name ‘Jenticha’. It is also known by the generic name ‘Koits Brese’ (‘Koincha Brehs’). The final name will be determined through correspondence with specialists and potential users.

Characters The Jenticha chart from Borchers shows the letter ೞ, which is given the transliteration व्हो *who* (see Figure 5).³ This character may represent the bilabial implosive [ɓ], which is likely given its placement in the chart after *va*. It is included in the character repertoire and tentatively named *vɦA*. However, Borchers writes that in Devanagari orthography for Sunuwar, [ɓ] is written as अ॒व.⁴ Additional research will assist in determining the actual value of ೞ.

Character Properties The ◌: VOWEL LENGTH MARK is given the property *gc=Mc* (Mark, spacing). It may be written after the ◌̃ NASAL SIGN, which is *gc=Mn* (Mark, nonspacing). Given the issues with rendering VISARGA (*gc=Mc*) after ANUSVARA OF CANDRABINDU (*gc=Mn*) in Devanagari and other Indic scripts, it may be practical to define the Jenticha VOWEL LENGTH MARK as *gc=Lo* (Letter, other).

7 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.
- Pandey, Anshuman. 2010. “Preliminary Proposal to Encode the Tikamuli Script in ISO/IEC 10646” (N3963). November 5, 2010. <http://std.dkuug.dk/JTC1/SC2/WG2/docs/n3963.pdf>
- रापचा, लाल-श्याँकारेलु [Rāpacā, Lāla-Śyāmkārelu]. 2009. नेपालको आदिवासी किराँती जाति सेरोफेरो : ओल्लो, माझ र पल्लो किराँत. Kathmandau.

8 Acknowledgments

I would like to extend my gratitude to Lāla-Śyāmkārelu Rāpacā (Kathmandu, Nepal) for providing me with details about the history and orthography of the Jenticha script. I would also like to thank Krishna Sunuwar (Kathmandu, Nepal) for giving me information about the scripts used for writing Sunuwar in Sikkim and Nepal.

³ Borchers 2008: 47.

⁴ Borchers 2008: 34.

	11C0	11C1	11C2
0	𑄀 11C00	𑄁 11C10	𑄂 11C20
1	𑄃 11C01	𑄄 11C11	𑄅 11C21
2	𑄆 11C02	𑄇 11C12	
3	𑄈 11C03	𑄉 11C13	
4	𑄊 11C04	𑄋 11C14	
5	𑄌 11C05	𑄍 11C15	
6	𑄎 11C06	𑄏 11C16	𑄐 11C26
7	𑄑 11C07	𑄒 11C17	𑄓 11C27
8	𑄔 11C08	𑄕 11C18	𑄖 11C28
9	𑄗 11C09	𑄘 11C19	𑄙 11C29
A	𑄚 11C0A	𑄛 11C1A	𑄜 11C2A
B	𑄞 11C0B	𑄟 11C1B	𑄠 11C2B
C	𑄢 11C0C	𑄣 11C1C	𑄤 11C2C
D	𑄧 11C0D	𑄨 11C1D	𑄩 11C2D
E	𑄭 11C0E	𑄮 11C1E	𑄯 11C2E
F	𑄱 11C0F	𑄲 11C1F	𑄳 11C2F

Vowels

- 11C00 𑄀 JENTICHA LETTER A
- 11C01 𑄃 JENTICHA LETTER AA
- 11C02 𑄆 JENTICHA LETTER I
- 11C03 𑄈 JENTICHA LETTER U
- 11C04 𑄊 JENTICHA LETTER E
- 11C05 𑄌 O JENTICHA LETTER O

Consonants

- 11C06 𑄎 JENTICHA LETTER KA
- 11C07 𑄑 JENTICHA LETTER KHA
- 11C08 𑄔 JENTICHA LETTER GA
- 11C09 𑄗 JENTICHA LETTER NGA
- 11C0A 𑄚 < JENTICHA LETTER CA
- 11C0B 𑄞 JENTICHA LETTER JA
- 11C0C 𑄢 JENTICHA LETTER TTA
- 11C0D 𑄧 Z JENTICHA LETTER TTHA
- 11C0E 𑄭 𑄲 JENTICHA LETTER TA
- 11C0F 𑄱 𑄲 JENTICHA LETTER THA
- 11C10 𑄁 JENTICHA LETTER DA
- 11C11 𑄄 𑄅 JENTICHA LETTER NA
- 11C12 𑄇 𑄈 JENTICHA LETTER PA
- 11C13 𑄉 𑄊 JENTICHA LETTER PHA
- 11C14 𑄋 𑄌 JENTICHA LETTER BA
- 11C15 𑄍 𑄎 JENTICHA LETTER MA
- 11C16 𑄏 𑄐 JENTICHA LETTER YA
- 11C17 𑄒 N JENTICHA LETTER RA
- 11C18 𑄕 𑄖 JENTICHA LETTER LA
- 11C19 𑄘 𑄙 JENTICHA LETTER VA
- 11C1A 𑄛 𑄜 JENTICHA LETTER VHA
- 11C1B 𑄟 U JENTICHA LETTER SHA
- 11C1C 𑄣 𑄤 JENTICHA LETTER SA
- 11C1D 𑄧 𑄨 JENTICHA LETTER HA
- 11C1E 𑄭 𑄮 JENTICHA LETTER GLOTTAL STOP

Various signs

- 11C1F 𑄲 JENTICHA NASAL SIGN
- 11C20 𑄂 JENTICHA VOWEL LENGTH MARK
- 11C21 𑄅 JENTICHA NULL VOWEL

Digits

- 11C26 𑄐 JENTICHA DIGIT ZERO
- 11C27 𑄓 JENTICHA DIGIT ONE
- 11C28 𑄖 JENTICHA DIGIT TWO
- 11C29 𑄙 JENTICHA DIGIT THREE
- 11C2A 𑄜 JENTICHA DIGIT FOUR
- 11C2B 𑄠 JENTICHA DIGIT FIVE
- 11C2C 𑄤 JENTICHA DIGIT SIX
- 11C2D 𑄩 JENTICHA DIGIT SEVEN
- 11C2E 𑄯 JENTICHA DIGIT EIGHT
- 11C2F 𑄳 JENTICHA DIGIT NINE

Figure 1: Tentative code chart and nameslist for Jenticha.

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

Consonants व्यन्जन वर्ण

क	ख	ग	ङ	च
/k,क/	/k ^h ,ख/	/g,ग/	/ŋ,ङ/	/t,अ/
ज	झ	ञ		
/c,ज/	/dʒ,झ/	/ɟ,ञ/		
ट	ठ	ड	ड	
/ʈ,ट/	/ʈ ^h ,ठ/	/ɖ,ड/	/ɖ,ड/	
त	थ	द	न	
/t,त/	/t ^h ,थ/	/d,द/	/n,न/	
प	फ	ब	म	
/p,प/	/p ^h ,फ/	/b,ब/	/m,म/	
र	ल	व		
/r,र/	/l,ल/	/w,व/		
स	श			
/s,स/	/ʃ,श/			

Vowels स्वर वर्ण

इ	उ
/i,इ/	/u,उ/
अ	आ*
/ə,अ/	/ɑ,आ/
ए	ऐ
/ɛ,ए/	/ā,आ/

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

(स्रोत: रापचा ई.२००१ए २००२डी)

Figure 2: Consonants and vowels of Jenticha (from Rāpacā 2009: 143).

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

०	१२३४ sum /sum/ 'zero' शून्य
१	कख ka /kā/ 'one' एक
२	जझनिसि nizsi /nizsi/ 'two' दुई
३	जझसआN /sā/ 'three' तीन
४	जझसआ le /le/ 'four' चार
५	ङग nga /ŋə/ 'five' पाँच
६	कखरaku /raku/ 'six' ६
७	जझत tsani /cəni/ 'seven' सात
८	जझत sasi /səsi/ 'eight' आठ
९	जझत yAN /yā/ 'nine' नौ
१०	जझत gau /gəu/ 'ten' दश

Figure 3: Digits of Jenticha (from Rāpacā 2009: 144).

तालिका घ

किराँती-कौँच ब्रे:समा लेखन नमूना ट्रान्स्लिटरेसन सहित

ललमिच ममिब
ləlmits məmibə
<ललमिच ममिब:>

गोम तनातन

gom tənātən
<गोम् तनातन्>

रागीआ ललमिच रवाकलि

rāgiā ləlmits rəwākəli
<रागीआ ललमिच रवाकलि>

रेउम् तुषो यो ममइ

rēum tūṣo yo məmɪ
<रेउम् तुषो यो ममइ>

फश्यीम् हेमशो यो ममइ

p^həsim hēmṣoyo məmɪ
<फश्यीम् हेमशो यो ममइ>

नाम की'शो या ममइ

nām k'īšo yo məmɪ
<नाम की'शो या ममइ>

गोम तशोनु

gom təṣonu
<गोम् तशोनु>

मेको ललमिच रवा

meko ləlmits rəwā
<मेको ललमिच रवा>

बोइस्शोन् बाअब

bois̄ṣon bāʔbə
<बोइस्शोन् बाअब:>

सी'शोन् बाअब

s'īṣon bāʔbə
<सी'शोन् बाअब:>

का पाचआ ची'बपिकी

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

मेकेन् बेर्नमी

meken bərnimi
<मेकेन् बेर्नमी>

मेकेन् ना थेर्नमी

meken nā thernimi

<मेकेन् ना थेर्नमी>

मेकेन् ड्जूनिमी

meken dz'unimi
<मेकेन् ड्जूनिमी>

मिनु

minu
<मिनु>

कुम्शो पश्या शयील् पाइनिमी

kumṣo pəṣyā ṣyil pāinimi
<कुम्शो पश्या शयील् पाइनिमी>

श्येण मेको रागीन्आ

ṣyeṇ meko rāgin ā
<श्येण मेको रागीन्आ>

वेक् पाचआ

wek pāts ā
<वेक् पाचआ>

त्स'िब अंकलि

ts'ib ānkəli
<ची:बआन्कलि>

आँके रागीमि गेनाइयो ललमिच ममिब

āke rāgimi genāiyo ləlmits məmibə
<आँके रागीमि गेनाइयो ललमिच ममिब:>

(Source: Rapacha (2001a: 24))

Figure 4: Sunuwar phrases given in Jenticha, Latin, and Devanagari transliteration (adapted from Rāpacā 2009: 145–146).

Sunwar	Deva nā- garī	Roman	Sunwar	Deva- nāgarī	Roman/ Arabic
၂	अ	a	၂	उ ऊ	u, ū
၃	ब	b	၃	व	v
၄	च	c	၄	ळो	vho
၅	द	d	၅	य	y
၆	ए	e, e	၆	०	0
၇	ग	g	၇	१	1
၈	ह	h	၈	२	2
၉	इ ई	i, ī	၉	३	3
၁၀	ज	j	၁၀	४	4
၁၁	क	k	၁၁	५	5
၁၂	ल	l	၁၂	६	6
၁၃	म	m	၁၃	၇	7
၁၄	न	n	၁၄	၈	8
၁၅	ओ	o	၁၅	၉	9
၁၆	प	p	၁၆	၁၀	10
၁၇	र ड	r, ḍ			
၁၈	स	s			
၁၉	त ट	t, ṭ			

Figure 5: A chart of Jenticha (from Borchers 2008: 47).

ꣳꣳꣳ	= मूँह = बेला = TIME	—	1
ꣳꣳꣳꣳ	= लूज्ज-सेकन्ड = SECOND		2
ꣳꣳꣳꣳꣳ	= नौहँरस = मिनिट = MINUTE		3
ꣳꣳꣳꣳꣳꣳ	= तुलि = घन्ट = HOUR		4
ꣳꣳꣳꣳꣳꣳꣳ	= नौहिन = दिन = DAY		5
ꣳꣳꣳꣳꣳꣳꣳꣳ	= नुँफो = हफता = WEEK		6
ꣳꣳꣳꣳꣳꣳꣳꣳꣳ	= सिंन = मँना = MONTH		7
ꣳꣳꣳꣳꣳꣳꣳꣳꣳꣳ	= मुँडि = साल = YEAR		8
ꣳꣳꣳꣳꣳꣳꣳꣳꣳꣳꣳ	= तिरि = युग = CENTURY		9

Figure 6: A table showing names for chronological and calendrical terminology in Sunuwar (with Hindi equivalents) written in Jenticha and Devanagari (from Borchers 2008: 48).