

Preliminary proposal to encode Sirmauri in Unicode

Anshuman Pandey
pandey@umich.edu

March 28, 2018

1 Introduction

This is a preliminary proposal to encode the ‘Sirmauri’ script in Unicode. Sirmauri is a Brahmic script that belongs to the Takri family. It has sufficient differences from the primary Chambali style of Takri to be considered a distinctive script. The name of the script derives from the region of Sirmaur, a former independent state that is now part of Himachal Pradesh, India. It was used for writing Sirmauri (ISO 639-3: srx), a Western Pahari language. The script of Sirmaur is similar to that used in Jaunsar for writing Jaunsari (ISO 639-3: jns), also a Pahari language. These two scripts may be unified in the Sirmauri encoding.

The Sirmauri language, as well as all other Pahari languages, are now written in the Devanagari script. However, before the 1950s the Sirmauri script was the formal writing system for the language of Sirmaur. The same applies to other scripts of the Takri family, each of which is known by the region in which it was used, eg. the Jaunsari variety is known by the name of the Jaunsar region, now in Uttarakhand state in India. The proper name for the script is ‘Sirmauri’ and this name should be assigned to the script in Unicode. Such usage will not cause confusion about which script is currently used for the Sirmauri language.

Usage of the Sirmauri script declined after implementation of Devanagari as the formal orthography. Very little information about the script exists in English literature. A description of it and some specimens are given in the *Linguistic Study of India*, vol. IX (1916), which are also provided here. Documents in the script are preserved in India. Contact has been made with experts and a formal proposal is forthcoming.

1.1 Vowel letters

There are 7 independent vowel letters:

- 3 SIRMAURI LETTER A
- 3) SIRMAURI LETTER AA
- 6 SIRMAURI LETTER I
- 3 SIRMAURI LETTER U

ॐ	SIRMAURI LETTER E
ॐ	SIRMAURI LETTER AI
ॐ)	SIRMAURI LETTER O

Notes on vowel letters:

- The Sirmauri letter ॐ A is unique among related scripts.
- The form ॐ) is shown in fig. 3 as a variant of ॐ) SIRMAURI LETTER AA in the Jaunsari style. It is to be treated as a glyphic variant.
- Distinctive letters do not exist for \bar{i} and \bar{u} . These are represented using ॐ SIRMAURI LETTER I and ॐ SIRMAURI LETTER U, respectively.
- The independent forms of i and e may be written using the letter ॐ instead of ॐ SIRMAURI LETTER I or ॐ SIRMAURI LETTER E (see fig. 1). This letter may be treated as a glyph variant.
- The ॐ SIRMAURI LETTER AI is written as ॐ in the Jaunsari style, which may be analyzed as a sequence of ॐ SIRMAURI LETTER I and ॐ SIRMAURI VOWEL SIGN E (see fig. 1).

1.2 Vowel Signs

There are 7 dependent vowel signs:

ॐ)	SIRMAURI VOWEL SIGN AA
ॐ	SIRMAURI VOWEL SIGN I
ॐ	SIRMAURI VOWEL SIGN U
ॐ	SIRMAURI VOWEL SIGN E
ॐ	SIRMAURI VOWEL SIGN AI
ॐ)	SIRMAURI VOWEL SIGN O
ॐ)	SIRMAURI VOWEL SIGN AU

Notes on vowel signs:

- The ॐ SIRMAURI VOWEL SIGN I and ॐ SIRMAURI VOWEL SIGN U are used, respectively, for both short and long forms of i and u .
- The SIRMAURI VOWEL SIGN E .. SIRMAURI VOWEL SIGN AU are atomic signs. They are not to be decomposed. The top components of SIRMAURI VOWEL SIGN O and SIRMAURI VOWEL SIGN AU are detached from the stem. In rapid writing they may ‘float’, eg. positioned above either the ॐ stem or the base

letter. In formal typography they should be positioned above the stem. Similarly, SIRMAURI VOWEL SIGN E and SIRMAURI VOWEL SIGN AI should be positioned at the right edge of the base letter.

1.3 Consonants

There are 28 consonant letters:

ⱱ	SIRMAURI LETTER KA
Ⱳ	SIRMAURI LETTER KHA
ⱳ	SIRMAURI LETTER GA
ⱴ	SIRMAURI LETTER GHA
Ⱶ	SIRMAURI LETTER CA
ⱶ	SIRMAURI LETTER CHA
ⱷ	SIRMAURI LETTER JA
ⱸ	SIRMAURI LETTER JHA
ⱹ	SIRMAURI LETTER TTA
ⱺ	SIRMAURI LETTER TTHA
ⱻ	SIRMAURI LETTER DDA
ⱼ	SIRMAURI LETTER DDHA
ⱽ	SIRMAURI LETTER NNA
Ȿ	SIRMAURI LETTER TA
Ɀ	SIRMAURI LETTER THA
Ⳁ	SIRMAURI LETTER DA
ⳁ	SIRMAURI LETTER DHA
Ⳃ	SIRMAURI LETTER NA
ⳃ	SIRMAURI LETTER PA
Ⳅ	SIRMAURI LETTER PHA
ⳅ	SIRMAURI LETTER BA
Ⳇ	SIRMAURI LETTER BHA
ⳇ	SIRMAURI LETTER MA
Ⳉ	SIRMAURI LETTER YA

᳚	SIRMAURI LETTER RA
᳛	SIRMAURI LETTER LA
᳜	SIRMAURI LETTER SA
᳝	SIRMAURI LETTER HA

Notes on consonants:

- The SIRMAURI LETTER KHA is written as **᳛** instead of **᳚** in the Jaunsari style. This form is to be treated as a glyphic variant.
- The representative shape of **᳞** SIRMAURI LETTER JHA is based upon the Jaunsari form. In the proper Sirmauri style *jha* is not distinguished from **᳟** SIRMAURI LETTER JA.
- The retroflex flap *ra* is represented by writing ◌᳚ SIRMAURI SIGN NUKTA with **᳜** SIRMAURI LETTER DDA, eg. **᳜᳚**.
- The representative shape of the **᳠** SIRMAURI LETTER NNA is based upon the Jaunsari form. In the proper Sirmauri style *na* is not distinguished from **᳚** SIRMAURI LETTER NA.
- The SIRMAURI LETTER THA is written as **᳡** instead of **᳢** in the Jaunsari style.
- The consonant *va* is written as **᳣**, which to be produced using the sequence **᳢** SIRMAURI LETTER BA + ◌᳚ SIRMAURI SIGN NUKTA.
- The nasal letters *na* and *ṅa* (and *ṇa* in proper Sirmauri) are represented using **᳚** SIRMAURI LETTER NA.
- The consonant *śa* assimilates with *sa* in the Pahari languages and is written using **᳜** SIRMAURI LETTER SA. It may be represented as **᳣**, which is to be produced using the sequence **᳢** SIRMAURI LETTER BA + ◌᳚ SIRMAURI SIGN NUKTA.
- The consonant *ṣa* assimilates with *kha* in Pahari, therefore it is often represented using **᳚** SIRMAURI LETTER KHA.

1.4 Various signs

◌᳚	SIRMAURI SIGN ANUSVARA
◌᳚	SIRMAURI SIGN NUKTA
◌᳚	SIRMAURI SIGN VIRAMA

The ◌᳚ SIRMAURI SIGN ANUSVARA is used for indicating nasalization.

The ◌᳚ SIRMAURI SIGN NUKTA is used for extending the consonant repertoire to represent a sound that does not have a distinctive letter in the script.

A *virāma* is not attested for Sirmauri in the materials from the *Linguistic Survey of India*. However, the ◌̣ SIRMAURI SIGN VIRAMA has been added to the tentative repertoire as it may be needed for indicating the absence of the inherent vowel in a consonant letter.

2 References

Grierson, George A. 1916a. *The Linguistic Survey of India*. Vol. IX. Indo-Aryan Family. Central Group. Part I. Specimens of Western Hindī and Pañjābī. Calcutta: Office of the Superintendent of Government Printing, India.

———. 1916b. *The Linguistic Survey of India*. Vol. IX. Indo-Aryan Family. Central Group. Part IV. Specimens of Pahārī Languages and Gujuri. Calcutta: Office of the Superintendent of Government Printing, India.

Jensen, Hans. 1969. *Die Schrift: In Vergangenheit und Gegenwart*. Reprint der 3. Auflage. Berlin: Deutscher Verlag der Wissenschaften.

Pandey, Anshuman. 2009. “Proposal to Encode the Takri Script in ISO/IEC 10646” (L2/09-424).
<http://www.unicode.org/L2/L2009/09424-takri.pdf>

3 Acknowledgments

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). Any views, findings, conclusions or recommendations expressed in this publication do not necessarily reflect those of the National Endowment of the Humanities.

	1185	1186	1187	1188
0	3 11850	ॣ 11860	᳚ 11870	◌̇ 11880
1	3᳚ 11851		᳚ 11871	◌̇ 11881
2	᳚ 11852	᳚ 11862	᳚ 11872	◌̇ 11882
3	᳚ 11853	᳚ 11863	᳚ 11873	
4	᳚ 11854	᳚ 11864		
5	᳚ 11855	᳚ 11865		
6	᳚ 11856	᳚ 11866		
7	᳚ 11857	᳚ 11867	᳚ 11877	
8	᳚ 11858	᳚ 11868	᳚ 11878	
9	᳚ 11859	᳚ 11869	᳚ 11879	
A	᳚ 1185A	᳚ 1186A	᳚ 1187A	
B	᳚ 1185B	᳚ 1186B	᳚ 1187B	
C		᳚ 1186C	᳚ 1187C	
D	᳚ 1185D	᳚ 1186D	᳚ 1187D	
E	᳚ 1185E	᳚ 1186E	᳚ 1187E	
F	᳚ 1185F	᳚ 1186F	᳚ 1187F	

Independent vowels

11850	3	SIRMAURI LETTER A
11851	3᳚	SIRMAURI LETTER AA
11852	᳚	SIRMAURI LETTER I
11853	᳚	SIRMAURI LETTER U
11854	᳚	SIRMAURI LETTER E
11855	᳚	SIRMAURI LETTER AI
11856	᳚	SIRMAURI LETTER O
11857	᳚	SIRMAURI LETTER AU

Consonants

11858	᳚	SIRMAURI LETTER KA
11859	᳚	SIRMAURI LETTER KHA
1185A	᳚	SIRMAURI LETTER GA
1185B	᳚	SIRMAURI LETTER GHA
1185C		<reserved>
1185D	᳚	SIRMAURI LETTER CA
1185E	᳚	SIRMAURI LETTER CHA
1185F	᳚	SIRMAURI LETTER JA
11860	᳚	SIRMAURI LETTER JHA
11861		<reserved>
11862	᳚	SIRMAURI LETTER TTA
11863	᳚	SIRMAURI LETTER TTHA
11864	᳚	SIRMAURI LETTER DDA
11865	᳚	SIRMAURI LETTER DDHA
11866	᳚	SIRMAURI LETTER NNA
11867	᳚	SIRMAURI LETTER TA
11868	᳚	SIRMAURI LETTER THA
11869	᳚	SIRMAURI LETTER DA
1186A	᳚	SIRMAURI LETTER DHA
1186B	᳚	SIRMAURI LETTER NA
1186C	᳚	SIRMAURI LETTER PA
1186D	᳚	SIRMAURI LETTER PHA
1186E	᳚	SIRMAURI LETTER BA
1186F	᳚	SIRMAURI LETTER BHA
11870	᳚	SIRMAURI LETTER MA
11871	᳚	SIRMAURI LETTER YA
11872	᳚	SIRMAURI LETTER RA
11873	᳚	SIRMAURI LETTER LA
11874		<reserved>
11875		<reserved>
11876		<reserved>
11877	᳚	SIRMAURI LETTER SA
11878	᳚	SIRMAURI LETTER HA

Dependent vowel signs

11879	᳚	SIRMAURI VOWEL SIGN AA
1187A	᳚	SIRMAURI VOWEL SIGN I
1187B	᳚	SIRMAURI VOWEL SIGN U
1187C	᳚	SIRMAURI VOWEL SIGN E
1187D	᳚	SIRMAURI VOWEL SIGN AI
1187E	᳚	SIRMAURI VOWEL SIGN O
1187F	᳚	SIRMAURI VOWEL SIGN AU

Various signs

11880	◌̇	SIRMAURI SIGN ANUSVARA
11881	◌̇	SIRMAURI SIGN NUKTA
11882	᳚	SIRMAURI SIGN VIRAMA

The Sirmauri Alphabet.

a	3	da	ε
ā	3)	dha	ε̄
i, i & ḍ	6 or 5	na	∩
u & ū	⊙	pa	5
ē	3 or 5	pha	5
ai	⊙	ba	9
ō & ȳ	⊙	bha	9
au	⊙	ma	∩
ka	x	ya	x
kha	5	ra	∩
ga	∩	la	ε
gha	5	wa	ε̄
cha	∩	śa	ε̄
chha	∩	sha	5
ja	∩	sa	ε̄
jha	∩	ha	ε̄
ṭa	∩	kā	x
ṭha	6	ki, kī	x
ḍa	5	ku, kū	x
ḍha	5	ke, kē	x
ṇa	∩	kai	x
ta	∩	ko, kō	x
tha	5	kau	x

Figure 1: Chart of the Sirmauri form of Takri (from Grierson 1916b: 459).

WESTERN PAHĀRĪ (SIRMAURĪ).

DHĀRTHĪ DIALECT.

SPECIMEN I.

3xṛ ʃāṭhē ɖ ʃāṭhē ʃāṭhē "
 xāṭhē ʃāṭhē ʃāṭhē ʃāṭhē ʃāṭhē ʃāṭhē "
 ʃāṭhē ʃāṭhē ʃāṭhē ʃāṭhē ʃāṭhē "
 ʃāṭhē ʃāṭhē ʃāṭhē ʃāṭhē ʃāṭhē "
 ʃāṭhē ʃāṭhē ʃāṭhē ʃāṭhē ʃāṭhē ʃāṭhē
 ʃāṭhē ʃāṭhē ʃāṭhē ʃāṭhē ʃāṭhē ʃāṭhē
 ʃāṭhē ʃāṭhē ʃāṭhē ʃāṭhē ʃāṭhē ʃāṭhē
 ʃāṭhē ʃāṭhē ʃāṭhē ʃāṭhē ʃāṭhē ʃāṭhē
 ʃāṭhē ʃāṭhē ʃāṭhē ʃāṭhē ʃāṭhē ʃāṭhē
 ʃāṭhē ʃāṭhē ʃāṭhē ʃāṭhē ʃāṭhē ʃāṭhē
 ʃāṭhē ʃāṭhē ʃāṭhē ʃāṭhē ʃāṭhē ʃāṭhē

Figure 2: Specimen of the Sirmauri form of Takri (from Grierson 1916b: 468).

JAUNSĀRĪ (SIRMAURĪ) ALPHABET.

३	a	⤴ X)	kau	३	tha
३), ३))	ā	⤴ X)	kā	३	da
३	i	५	ka	५	dha
३	ī	५	kha	५	na
३	u	३)	ga	५	pa
३	ū	३)	gha	५	pha
३, ३	ě, ē	—	na	३	ba
३)	ai	५	cha	३	bha
३)	ǎ, ǎ	३	chha	३	ma
३)	au	३)	ja	५	ya
.	~	५	jha	३	ra
५	ka	—	ña	३	la
३)	kū	✓	tā	३	wa
५)	ki	✓ ३	thā	५	śa
५)	kī	३ ३	ḍa	५	śha
५)	ku	३	ra	५	sa
५)	kū	३	dha	३	ha
५)	kě, kē	३	rha		
५)	kai	३)	ṛa		
५)	kǎ, kṓ	३)	ta		

When used before consonants of their own class, nasals are, as usual, represented by Anusvāra (◌̣). As ṛi and ṛī occur only in this position, they have no special character

Figure 3: Chart of the Jaunsari form of Takri (from Grierson 1916b: 384).

WESTERN PAHĀRĪ (JAUNŚĀRĪ).

STANDARD DIALECT.

DISTRICT DERA DUN.

SPECIMEN I.

ॐ नमो भगवते वासुदेवाय ॥ नमो भगवते वासुदेवाय ॥ नमो भगवते वासुदेवाय ॥
 नमो भगवते वासुदेवाय ॥ नमो भगवते वासुदेवाय ॥ नमो भगवते वासुदेवाय ॥
 नमो भगवते वासुदेवाय ॥ नमो भगवते वासुदेवाय ॥ नमो भगवते वासुदेवाय ॥
 नमो भगवते वासुदेवाय ॥ नमो भगवते वासुदेवाय ॥ नमो भगवते वासुदेवाय ॥
 नमो भगवते वासुदेवाय ॥ नमो भगवते वासुदेवाय ॥ नमो भगवते वासुदेवाय ॥
 नमो भगवते वासुदेवाय ॥ नमो भगवते वासुदेवाय ॥ नमो भगवते वासुदेवाय ॥
 नमो भगवते वासुदेवाय ॥ नमो भगवते वासुदेवाय ॥ नमो भगवते वासुदेवाय ॥
 नमो भगवते वासुदेवाय ॥ नमो भगवते वासुदेवाय ॥ नमो भगवते वासुदेवाय ॥
 नमो भगवते वासुदेवाय ॥ नमो भगवते वासुदेवाय ॥ नमो भगवते वासुदेवाय ॥
 नमो भगवते वासुदेवाय ॥ नमो भगवते वासुदेवाय ॥ नमो भगवते वासुदेवाय ॥
 नमो भगवते वासुदेवाय ॥ नमो भगवते वासुदेवाय ॥ नमो भगवते वासुदेवाय ॥
 नमो भगवते वासुदेवाय ॥ नमो भगवते वासुदेवाय ॥ नमो भगवते वासुदेवाय ॥
 नमो भगवते वासुदेवाय ॥ नमो भगवते वासुदेवाय ॥ नमो भगवते वासुदेवाय ॥
 नमो भगवते वासुदेवाय ॥ नमो भगवते वासुदेवाय ॥ नमो भगवते वासुदेवाय ॥
 नमो भगवते वासुदेवाय ॥ नमो भगवते वासुदेवाय ॥ नमो भगवते वासुदेवाय ॥

Figure 4: Specimen of the Jaunsari form of Takri (from Grierson 1916b: 401).

Lautwert	Śāradā 804	Kasch- mirī	Ṭākri		Laṇḍā		Multani	Gur- mukhi
			Jaun- sari	Cha- meāli	Khu- dāwādi	Sindhī- Schrift		
a	𑆑	𑆑	𑆑	𑆑	𑆑	𑆑	𑆑	𑆑
i	𑆒	𑆒	𑆒	𑆒	𑆒	𑆒	𑆒	𑆒
u	𑆓	𑆓	𑆓	𑆓	𑆓	𑆓	𑆓	𑆓
e	𑆔	𑆔	𑆔	𑆔	𑆔	𑆔	𑆔	𑆔
o	𑆕	𑆕	𑆕	𑆕	𑆕	𑆕	𑆕	𑆕
ā	𑆖	𑆖	𑆖	𑆖	𑆖	𑆖	𑆖	𑆖
ka	𑆗	𑆗	𑆗	𑆗	𑆗	𑆗	𑆗	𑆗
kha	𑆘	𑆘	𑆘	𑆘	𑆘	𑆘	𑆘	𑆘
ga	𑆙	𑆙	𑆙	𑆙	𑆙	𑆙	𑆙	𑆙
gha	𑆚	𑆚	𑆚	𑆚	𑆚	𑆚	𑆚	𑆚
ṅa	𑆛	𑆛	𑆛	𑆛	𑆛	𑆛	𑆛	𑆛
ṅha	𑆜	𑆜	𑆜	𑆜	𑆜	𑆜	𑆜	𑆜
ṣa	𑆝	𑆝	𑆝	𑆝	𑆝	𑆝	𑆝	𑆝
ṣha	𑆞	𑆞	𑆞	𑆞	𑆞	𑆞	𑆞	𑆞
ṭa	𑆟	𑆟	𑆟	𑆟	𑆟	𑆟	𑆟	𑆟
ṭha	𑆠	𑆠	𑆠	𑆠	𑆠	𑆠	𑆠	𑆠
ḍa	𑆡	𑆡	𑆡	𑆡	𑆡	𑆡	𑆡	𑆡
ḍha	𑆢	𑆢	𑆢	𑆢	𑆢	𑆢	𑆢	𑆢
ṇa	𑆣	𑆣	𑆣	𑆣	𑆣	𑆣	𑆣	𑆣
ṇha	𑆤	𑆤	𑆤	𑆤	𑆤	𑆤	𑆤	𑆤
ṭa	𑆥	𑆥	𑆥	𑆥	𑆥	𑆥	𑆥	𑆥
ṭha	𑆦	𑆦	𑆦	𑆦	𑆦	𑆦	𑆦	𑆦
ḍa	𑆧	𑆧	𑆧	𑆧	𑆧	𑆧	𑆧	𑆧
ḍha	𑆨	𑆨	𑆨	𑆨	𑆨	𑆨	𑆨	𑆨
pa	𑆩	𑆩	𑆩	𑆩	𑆩	𑆩	𑆩	𑆩
pha	𑆪	𑆪	𑆪	𑆪	𑆪	𑆪	𑆪	𑆪
ba	𑆫	𑆫	𑆫	𑆫	𑆫	𑆫	𑆫	𑆫
bha	𑆬	𑆬	𑆬	𑆬	𑆬	𑆬	𑆬	𑆬
ma	𑆭	𑆭	𑆭	𑆭	𑆭	𑆭	𑆭	𑆭
ya	𑆮	𑆮	𑆮	𑆮	𑆮	𑆮	𑆮	𑆮
ra	𑆯	𑆯	𑆯	𑆯	𑆯	𑆯	𑆯	𑆯
la	𑆰	𑆰	𑆰	𑆰	𑆰	𑆰	𑆰	𑆰
va	𑆱	𑆱	𑆱	𑆱	𑆱	𑆱	𑆱	𑆱
śa	𑆲	𑆲	𑆲	𑆲	𑆲	𑆲	𑆲	𑆲
ṣa	𑆳	𑆳	𑆳	𑆳	𑆳	𑆳	𑆳	𑆳
sa	𑆴	𑆴	𑆴	𑆴	𑆴	𑆴	𑆴	𑆴
ha	𑆵	𑆵	𑆵	𑆵	𑆵	𑆵	𑆵	𑆵

Figure 5: Comparison of Sirmauri (Jaunsari) with related scripts (from Jensen 1969: 366).