L2/25-224
2025-10-15

Proposal to encode Sirmauri in Unicode

Anshuman Pandey
pandey@umich.edu
pandey.github.io/unicode

Biswajit Mandal
biswajitmandal.bm90@gmail.com

October 15, 2025

1 Introduction

This proposal to encode the Sirmauri script in the Unicode standard supersedes the following documents:

- L2/18-085: "Preliminary proposal to encode Sirmauri in Unicode" (Pandey)
- L2/25-038: "Proposal to encode Sirmauri in Unicode" (Pandey & Mandal, October 2024)
- L2/25-134: "Proposal to encode Sirmauri in Unicode" (Pandey & Mandal, March 2025)

The following major changes have been made following Script Encoding Working Group reviews in November 2024, and July – October 2025:

- The vowel model has been simplified to a set of vowel-carrier letters and combining vowel signs
- The encoding model does not support conjunct formation and has no VIRAMA control character
- The SIRMAURI SIGN HALANTA is a simple combining sign and is always rendered visibly
- Addition of decomposition for SIRMAURI VOWEL SIGN O and SIRMAURI VOWEL SIGN AU
- Digits have been moved to the end of column U+1188 and column U+1189 has been unallocated

Sirmauri is allocated to U+11850..1188F on the Roadmap to the Supplementary Multilingual Plane (SMP).

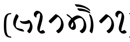
2 Proposal History

Sirmauri was first introduced to the Unicode Technical Committee (UTC) in L2/18-085. The aim was to establish the suitability of encoding Sirmauri, with initial evidence drawn from charts and specimens published in *The Linguistic Survey of India* by George A. Grierson (1916b) and *Die Schrift* by Hans Jensen (1969). However, Pandey did not want to pursue the encoding using only secondary sources; and at that time, he was unable to locate experts of the script or primary materials. For this reason, the proposal had been on hold. A few years ago, Mandal independently established contact with Sirmauri users in Himachal Pradesh, who provided manuscripts and feedback on the script. In 2024, Pandey and Mandal decided to merge their efforts in order to present a single, comprehensive proposal for encoding for the script in Unicode.

This proposal presents the form of Sirmauri as attested in actual primary sources, such as manuscripts containing accounting (*huṇḍī*), astrological (*sāñcā*), calendrical (*pañcāṅgam*), and other ritual texts, as well as

personal diaries and administrative records. These materials are more representative of conventional Sirmauri than the script shown by Grierson. Moreover, there are prominent differences between the two, as well as inaccuracies in Grierson's form (see § 5). The representative digital letterforms used in this proposal are based upon normalization of characters used in these primary sources, and were drawn by Pandey and vetted by Mandal. Most importantly, this proposal and the representative glyphs have been reviewed and approved by several current users of Sirmauri, who are acknowledged in § 10.

3 Background

Sirmauri () is a contemporary Brahmi-based script that is written from left to right. Until the middle of the 20th century, it was a commonly used script in Sirmaur (also spelled 'Sirmour' and 'Sirmur'), a princely state in pre-independence India, which is today split across the present district of Sirmaur in Himachal Pradesh and the Jaunsar-Bawar region of the Garwahl district of Uttarakand. It was used for writing Sirmauri (ISO 639-3: srx), a Western Pahari language, and a sub-regional variant was used for Jaunsari (ISO 639-3: jns). Similar to other Indic scripts, the appellation 'Sirmauri' for the script is derived from the name of the Sirmauri language, which in turn refers to the Sirmaur region. This is the common name for the script, rendered as सिरमौरी *sirmauri* in Devanagari, and used in conventional media (see fig. 39).

The script of Sirmaur is a distinctive script related to the Takri (U+11680) of Chamba; the Dogra (U+11800) of Jammu; and the common mercantile Mahajani (U+11150) of northern India. An origin story for the script relates that when the royal families of Sirmaur formed marital alliances with those of Rajasthan, the latter brought their accountants, clerks, and administrators, who, in turn, brought their native Mahajani script. When this Mahajani was combined with the local form of Takri, a distinctive Sirmauri script emerged.

Sirmauri was used as an official script from at least the 18th century for producing documents such as land grants (see fig. 6–12), revenue records, accounting books, etc., which were published in government gazetteers. It was also used in formal education. According to the *Sirmur State Gazetteer, 1904*, there was a school in Kufarmand of Pachchad, where the Sirmauri and Hindi languages were taught in the Sirmauri and Devanagari scripts, respectively. While usage of Sirmauri for official and educational purposes gave way to Devanagari in the 20th century, it continues to be used cultural and religious practices. The script is intimately connected in Himachal Pradesh with *sāñcā vidyā* — ritual, calendrical, and astrological practices — which is maintained in manuscripts known as *sāñcā grantha* (see fig. 13–17).

Western scholars have known about the Sirmauri script since the 18th century. However, there is very little coverage of both the Sirmauri language and script in Western secondary literature. George A. Grierson described the script and published specimens and charts of it in *The Linguistic Survey of India*, vol. IX (1916b); shown here in figs. 1–4. The letters of the Jaunsari variety appears in a chart in *Die Schrift* (1969) by Hans Jensen (fig. 5), but, those forms appear to be derived from those in Grierson.

Although Sirmauri has been continously used for centuries with a stable orthography, today it is an endangered script. It is kept alive by Sirmauri speakers, who actively read and write it, and who have maintained knowledge of it through their cultural practices and community traditions. While users continue to write the script, they also want to be able to type it on digital devices. They have begun to use digital fonts to create and exchange documents in their script (see fig. 40 for a poem by Sherjung Chauhan). They also express a strong desire to digitize their manuscript collections and to be able to represent the contents of these documents in digital plain text. Encoding Sirmauri in Unicode will enable its users to preserve, interchange, and create content in the script, and to expand its usage on digital platforms on par with other long-lived and active minority Indic scripts.

4 The Script

4.1 Vowel letters

There are 4 basic vowel letters in Sirmauri:

a *i, ī* *u, ū* *e*

The letters *a*, *i*, and *u* function as vowel carriers for representing independent vowels. The letter 𑚆 represents both *i* and *ī*; and 𑚇 represents both *u* and *ū*. But, vowel length may also be distinctively indicated, as described below, by combining dependent vowel signs with these base letters.

4.2 Vowel signs

The basic letters are accompanied by 7 dependent vowel signs:

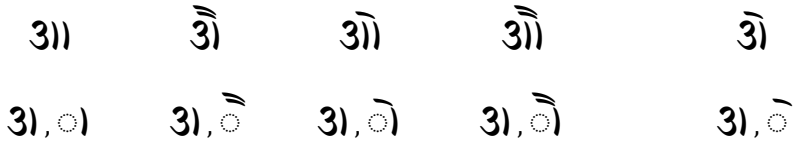
-ā *-i, -ī* *-u, -ū* *-e* *-ai* *-o* *-au*

Like other Indic scripts, there is no dependent sign for *-a*, the vowel inherent in each consonant letter. The sign ◌𑚈 represents *-i* and *-ī*; the sign ◌𑚉 represents *-u*, *-ū*.

4.3 Representation of independent long vowels and diphthongs

4.3.1 *ā, ai, o, au*; and alternate *e*

Independent forms of *ā*, *ai*, *o*, *au*, and an alternate *e* are represented by combining the respective vowel sign with the letter *a*.



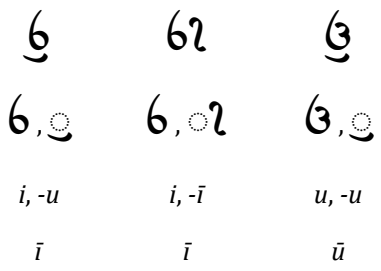
a, -ā *a, -ai* *a, -o* *a, -au* *a, -e*

ā *ai* *o* *au* *e*

Both the independent letter 𑚇 and the vowel-carrier form 𑚃𑚊 are commonly used for writing *e* (see fig. 19, 20).

4.3.2 *ī, ū*

Independent forms of *ī* and *ū* are written using vowel letters and combining vowel signs (see fig. 19, 20).



The variant form 𑐆𑐶 of *ī* is attested in one document present in the available materials (see fig. 35). The usage of two dots to distinguish *i* and *ī* is a common pattern in related scripts. It is also quite possible that the conventional sign 𑐺 may have evolved from writing the two dots as a single swash.

	i	*ī*
Sharada	𑆈	𑆉
Dogra	𑐆	𑐆𑐶
Takri	𑚆	𑚆

The usage of the sign 𑐲 *-i* for writing 𑐆𑐲 is also an orthographic convention in related scripts, eg. Khojki 𑈈𑈲 *ī*, which is palaeographically a combination of 𑈈 *i* and the sign 𑈲 *-ī*.

4.4 Consonant letters

There are 28 consonant letters, arranged according to the typical Brahmi paradigm. The script lacks distinctive letters for representing the nasals *ṅa* and *ña*, the sibilants *śa* and *ṣa*, and the semi-vowel *va*.

𑐎	𑐏	𑐐	𑐑	—	𑐔	𑐕	𑐖	𑐗	—
ka	*kha*	*ga*	*gha*	*ṅa*	*ca*	*cha*	*ja*	*jha*	*ña*
𑐘	𑐙	𑐚	𑐛	𑐜	𑐝	𑐞	𑐟	𑐠	𑐡
ṭa	*ṭha*	*ḍa*	*ḍha*	*ṇa*	*ta*	*tha*	*da*	*dha*	*na*
𑐢	𑐣	𑐤	𑐥	𑐦	𑐧	𑐨	𑐩	—	
pa	*pha*	*ba*	*bha*	*ma*	*ya*	*ra*	*la*	*va*	
—	—	𑐪	𑐫						
śa	*ṣa*	*sa*	*ha*						

Other consonant sounds are represented by combining a distinctive letter with the sign *nukta* (see § 4.6).

4.5 Variants

Naturally, there is variation in shapes of letters in hand-written documents, but distinctive variants are also attested (see fig. 32–33, 38):

e	*ca*	*ṭa*	*ḍha*	*ta*	*ya*	*ha*

Additionally, some documents show Devanagari-like forms for letters, such as *tha*, *pa*, *pha* (see fig. 34–35). Usage of these is likely due to contact with Devanagari, and the fluency of the writer with both scripts.

4.6 Modifier signs

The following modifier signs are used:

anusvāra	*visarga*	*halanta*	*nukta*

- *anusvāra* indicates vowel nasalization or is a proxy for a nasal consonant, eg. *kaṃ*
- *visarga* indicates post-vocalic aspiration and occurs in Sanskrit contexts, eg. *namaḥ* (see fig. 37)
- *halanta* silences the inherent vowel *a* of a consonant letter, eg. *k* (see fig. 21, 22), *kka*
- *nukta* is used for extending the script to represent sounds for which distinctive letters do not exist:

ḍa, nukta	*ḍha, nukta*	*sa, nukta*	*pa, nukta*	*ba, nukta*	*la, nukta*
ṛa	*ṛha*	*śa*	*va*	*va*	*ḷa*

4.7 Consonant clusters

Consonant clusters are conventionally represented by writing the sign *halanta* (*virāma*) under the consonant whose inherent vowel is silenced, eg. *k*. There are no atomic conjuncts. However, while rare, there is attestation in a single document of half or contextual forms of letters for representing conjuncts. Such usage appears to be borrowed from Devanagari.

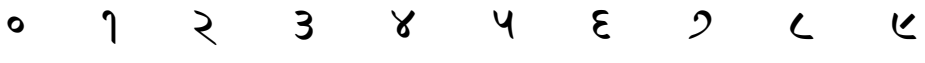
half *ma*	*ra-kāra*

- The half *ma* occurs in the cluster *mma* in a single document (see fig. 34). This conjunct representation of a cluster is idiosyncratic.

- The *ra-kāra* occurs in a few documents (see fig. 38). It is a contextual form of *ra* used in a cluster-medial position, eg. *kra*, *pra*. The usage of this form indicates a strong orthographic influence from Devanagari. Other contextual forms of *ra*, such as the cluster-initial *repha* are not attested.

4.8 Digits

There is a full set of decimal digits (see fig. 24, 25):



0 1 2 3 4 5 6 7 8 9

4.9 Auspicious sign

An auspicious sign known as *ekam* is used in Sirmauri (see fig. 26, 36), which is related to signs known as *añjī* in other northern and eastern Indic scripts:

ekam

4.10 Punctuation

Several signs of punctuation are used (see examples in figs. 27–30, 36):

daṇḍa double *daṇḍa* two dots four dots

One document shows rare usage of a Latin sign of punctation, the question mark '?' (see fig. 31).

Sirmauri documents often show a 'headline'. This is not a feature of the script, but of the writing surface. Ruled paper was often used, or lines were written on paper before starting with the text. Additional lines would be drawn as needed as the page is filled (see fig. 6 onwards).

5 Comparison with Grierson's form of Sirmauri

The following is a comparison of forms of Sirmauri and Jaunsari letters shown by Grierson (top) and digitized forms of letters used in the primary materials examined by Pandey and Mandal (below):

The letterforms in the two sets of sources have several similarities, but also some prominent differences, which are described below:

- *a* Grierson's **अ** *a* is mysterious. It does not occur in any of the manuscripts reviewed by the proposal authors, nor is it known to experts of the script, who report that they have never observed it in any Sirmauri materials, especially as a form for *a*. In Grierson, this **अ** form is the basis for **आ** *ā*, **ए** *e*, **ऐ** *ai*, **ओ** *o*, **औ** *au*. But, despite the consistency of this form across the independent vowel letters, the source for **अ** is unknown. It could have been an idiosyncrasy of the scribe of Grierson's text. Experts confirm that **आ** is the correct form for *a*. Accordingly, Grierson's forms would be analyzed as follows: **आ** *ā* = *a*; **ए** *e* and **ऐ** *ai* are invalid forms; **ओ** *o* = *e*; **औ** *au* = *ai*. Nonetheless, his recording of the Jaunsari form **आा** as *ā* is correct, as this form has the appropriate **आ** *a* as the base combined with the conventional vowel sign ◌ा for *ā*.

- *ā* The forms 𑐳𑐵 and 𑐳𑐶 are given for Jaunsari *ā*. Of these, 𑐳𑐶 is the correct form for *ā*. It is possible that 𑐳𑐵 *a* was used for *ā* in some contexts, just as 𑐴 is used for both *i* and *ī*, and 𑐺 for *u* and *ū*.

- *i*, *ī* The 𑐴 is the paleographical form for *i*, also used for *ī*. The 𑐾is the letter *e*, which along with 𑐬 *ya*, is used as a proxy for *i*.

- *e* The source for 𑐳 is unknown. It is *a*, functioning as a vowel carrier, with the vowel sign *e*. Similarly, 𑐽 is 𑐬 *ya* bearing the vowel sign *e*. The 𑐾is the letter for *e*. The correct vowel carrier form of *e* is 𑐳𑐾.

- *ai* The Jaunsari 𑐵 *ai* is the letter 𑐴 *i* written with the vowel sign *e*. Apart from Grierson, this form is not attested in other materials.

- *o*, *au* The source for the forms 𑐳𑐺 *o* and 𑐳𑐻 *au* is unknown. Apart from Grierson, these forms are not attested in other materials.

- *ṭa* The 𑐇 is a form heavily influenced by Devanagari *ṭa*. It is more commonly used in Sirmauri for *ṭha*, while the representative form for *ṭa* is 𑐆.

- *bha* The 𑐯 is a reversed form of conventional 𑐰 *bha*.

- *va* The 𑐮𑐶 is not a distinct letter, but the letter *ba* written with a *nukta*.

- *śa* The 𑐭𑐶 is not a distinct letter, but the letter *sa* written with a *nukta*.

- *ṣa* The 𑐑 shown for *ṣa* is simply a duplication of *kha*, which is commonly used for writing the sibilant.

6 Principles of the Encoding

The Unicode encoding for Sirmauri is based upon the following principles:

6.1 Scope

The 'Sirmauri' block encodes the script known as 'Sirmauri'. The 'Jaunsari' variety is to be unified with Sirmauri.

6.2 Consonant letters

All attested consonant letters have been proposed for encoding. The conventional forms of letters are specified as the representative glyphs.

Distinctive variants of consonant letters, such as those shown in fig. 32–33 are to be considered glyphic variants and their display is to be managed using fonts.

6.3 Conjunct encoding model

The usage of consonant conjuncts is not conventional for Sirmauri. Non-final consonants in clusters are marked using the sign *halanta* (synonymous with *virāma*). Therefore, the typical Indic conjoining model that uses the *virāma* conjoiner is not necessary for Sirmauri. Instead, the clusters may be represented as follows:

- SIRMAURI SIGN HALANTA The sign ◌𑑂 *halanta* is written under a consonant to indicate that the inherent vowel is silenced. It is proposed for encoding as a common combining sign as the character SIRMAURI SIGN HALANTA , but it does not have any control properties or conjoining behavior for conjunct formation, and is always rendered visibly.

  Half-forms that appear in Sirmauri texts should be normalized using the consonant letter + an explicit *halanta*; for example, the singular occurrence of 𑐣 *mma* would be normalized as:

  𑐣𑑂𑐣 <𑐣 SIRMAURI LETTER MA , ◌𑑂 SIRMAURI SIGN HALANTA , 𑐣 SIRMAURI LETTER MA>

- SIRMAURI CONSONANT SIGN MEDIAL RA The contextual form ◌𑑃 *ra-kāra* of 𑐪 *ra* is proposed for encoding as a separate combining sign: SIRMAURI CONSONANT SIGN MEDIAL RA. This reduces the need for a conjoiner for representing clusters. Clusters with *-ra* would be represented as:

  𑐏𑑃 *kra* <𑐏 SIRMAURI LETTER KA , ◌𑑃 SIRMAURI CONSONANT SIGN MEDIAL RA>
  𑐥𑑃 *pra* <𑐥 SIRMAURI LETTER PA , ◌𑑃 SIRMAURI CONSONANT SIGN MEDIAL RA>

6.4 Dependent vowel signs

A full set of seven dependent vowel signs have been proposed for encoding.

- *Decompositions* Two signs have the following decompositions:

  ◌𑑋 SIRMAURI VOWEL SIGN O : ◌𑑌 SIRMAURI VOWEL SIGN E + ◌𑑋 SIRMAURI VOWEL SIGN AA
  ◌𑑍 SIRMAURI VOWEL SIGN AU : ◌𑑍 SIRMAURI VOWEL SIGN AI + ◌𑑋 SIRMAURI VOWEL SIGN AA

- *Character name for vowel sign for -i, -ī* The ◌𑒲 is used for both *-i* and *-ī*, but it is palaeographically- *ī*. For this reason, it has been assigned the name SIRMAURI VOWEL SIGN II.

6.5 Vowel letters

There are two possible models for encoding Sirmauri vowel letters, whicb may be called 'palaeographical' and 'pedagogical'. These are described below. Based on feedback from the user community, the vowel-carrier model has been selected, which also simplifies the overall encoding model and Unicode implementations.

1. **Palaeographical: Vowel-carrier model**: This approach encodes only the distinctive vowel letters and defines them as vowel carriers. Vowel letters would be created using a sequence of a base vowel letter + a dependent vowel sign.

a	𑒰	SIRMAURI LETTER A
i	𑒱	SIRMAURI LETTER I
u	𑒳	SIRMAURI LETTER U
e	𑒵	SIRMAURI LETTER E

These base vowel letters would be used in combination with dependent vowel signs to represent the full set of independent vowel letters:

a	𑒰	𑒰 SIRMAURI LETTER A
ā	𑒰𑒹	𑒰 SIRMAURI LETTER A , ◌𑒹 SIRMAURI VOWEL SIGN AA
i	𑒱	𑒱 SIRMAURI LETTER I
ī	𑒱𑒲	𑒱 SIRMAURI LETTER I , ◌𑒲 SIRMAURI VOWEL SIGN U
u	𑒳	𑒳 SIRMAURI LETTER U
ū	𑒳𑒲	𑒳 SIRMAURI LETTER U , ◌𑒲 SIRMAURI VOWEL SIGN U
e	𑒵	𑒵 SIRMAURI LETTER E
e	𑒰𑒺	𑒰 SIRMAURI LETTER A , ◌𑒺 SIRMAURI VOWEL SIGN E
ai	𑒰𑒻	𑒰 SIRMAURI LETTER A , ◌𑒻 SIRMAURI VOWEL SIGN AI
o	𑒰𑒼	𑒰 SIRMAURI LETTER A , ◌𑒼 SIRMAURI VOWEL SIGN O
au	𑒰𑒽	𑒰 SIRMAURI LETTER A , ◌𑒽 SIRMAURI VOWEL SIGN AU

This approach faciliates support for alternate representations of vowel letters without requiring the independent encoding of such characters:

ī	𑒱𑒲	𑒱 SIRMAURI LETTER I , ◌𑒲 SIRMAURI VOWEL SIGN II
ai	𑒶	𑒱 SIRMAURI LETTER I , ◌𑒺 SIRMAURI VOWEL SIGN E

• normalization of the Jaunsari 𑒶 *ai* in Grierson

2. **Pedagogical: Atomic character model** This approach encodes all conventional independent vowel letters as atomic characters, following the pattern of other northern Indic scripts.

a	𑐃	SIRMAURI LETTER A
ā	𑐃𑐵	SIRMAURI LETTER AA
i	𑐄	SIRMAURI LETTER I
ī	𑐄𑐵	SIRMAURI LETTER II
u	𑐅	SIRMAURI LETTER U
ū	𑐅𑐵	SIRMAURI LETTER UU
e	𑐇	SIRMAURI LETTER E
e	𑐃𑐶	SIRMAURI LETTER ALTERNATE E
ai	𑐃𑐷	SIRMAURI LETTER AI
o	𑐃𑐸	SIRMAURI LETTER O
au	𑐃𑐹	SIRMAURI LETTER AU

Accordingly, the following would be treated as glyphic variants, whose representation would be controlled using fonts by substituting the representative glyph with the variant. If evidence of their concurrent usage with the normative form is found, then they may be encoded as alternate characters.

ī	𑐄𑐲	glyphic variant of 𑐄𑐵 SIRMAURI LETTER II
ī	𑐄𑐴	glyphic variant of 𑐄𑐵 SIRMAURI LETTER II

6.6 Digits

A set of script-specific digits for Sirmauri is proposed for encoding.

6.7 Auspicious sign

The auspicious sign *ekam* is proposed for encoding as an independent character in the Sirmauri block.

6.8 Punctuation

No Sirmauri-specific punctuation signs are proposed for encoding. Existing signs in Unicode can be used to support punctuation commonly used in Sirmauri:

- *daṇḍā*-s The । *daṇḍā* and ॥ double *daṇḍā* are to be unified with the corresponding characters in the Devanagari block. Script extensions have been specified.

।	U+0964 DEVANAGARI DANDA
॥	U+0964 DEVANAGARI DOUBLE DANDA

- *Dotted marks* The commonly used ∶ and ∷ may be represented using : U+003A COLON.

6.9 Reservation of codepoints in the block

Space has been reserved for the following characters, in the event that distinctive forms are identified in the future:

- *Vowel signs* The reservation of space for *SIRMAURI VOWEL SIGN I and *SIRMAURI VOWEL SIGN UU is driven by a recent update for Khojki, which also has a merger of *i*, *ī* and *u*, *ū*. The Khojki 𑈬 is palaeographically *ī*, but it was encoded as U+11202 KHOJKI LETTER I instead of *KHOJKI LETTER II (see L2/11-021). A few years ago, a distinctive independent form 𑈭 for *i* was identified. But as 𑈬 was already named KHOJKI LETTER I, the 𑈭 was encoded as U+11240 KHOJKI LETTER SHORT I (see L2/21-104). Using character names for Sirmauri based on palaeography avoids such complications.

- *Cconsonants* The following letters have reserved codepoints in their usual positions in the consonant order: *SIRMAURI LETTER NGA, *SIRMAURI LETTER NYA, *SIRMAURI LETTER SHA, *SIRMAURI LETTER SSA.

7 Proposed Character Repertoire

The proposed Unicode repertoire for Sirmauri consists of 55 characters:

Category	Character	Character Name
Vowel letters (4)	𑊰	SIRMAURI LETTER A
	𑊱	SIRMAURI LETTER I
	𑊲	SIRMAURI LETTER U
	𑊳	SIRMAURI LETTER E
Dependent vowel signs (7)	◌𑊰	SIRMAURI VOWEL SIGN AA
	◌𑊱	SIRMAURI VOWEL SIGN II
	◌𑊲	SIRMAURI VOWEL SIGN U
	◌𑊳	SIRMAURI VOWEL SIGN E
	◌𑊴	SIRMAURI VOWEL SIGN AI
	◌𑊵	SIRMAURI VOWEL SIGN O
	◌𑊶	SIRMAURI VOWEL SIGN AU
Consonant letters (28)	𑊷	SIRMAURI LETTER KA
	𑊸	SIRMAURI LETTER KHA
	𑊹	SIRMAURI LETTER GA
	𑊺	SIRMAURI LETTER GHA
	𑊻	SIRMAURI LETTER CA
	𑊼	SIRMAURI LETTER CHA
	𑊽	SIRMAURI LETTER JA

	𑚪	SIRMAURI LETTER JHA
	𑚫	SIRMAURI LETTER TTA
	𑚬	SIRMAURI LETTER TTHA
	𑚭	SIRMAURI LETTER DDA
	𑚮	SIRMAURI LETTER DDHA
	𑚯	SIRMAURI LETTER NNA
	𑚰	SIRMAURI LETTER TA
	𑚱	SIRMAURI LETTER THA
	𑚲	SIRMAURI LETTER DA
	𑚳	SIRMAURI LETTER DHA
	𑚴	SIRMAURI LETTER NA
	𑚵	SIRMAURI LETTER PA
	𑚶	SIRMAURI LETTER PHA
	𑚷	SIRMAURI LETTER BA
	𑚸	SIRMAURI LETTER BHA
	𑚹	SIRMAURI LETTER MA
	𑚺	SIRMAURI LETTER YA
	𑚻	SIRMAURI LETTER RA
	𑚼	SIRMAURI LETTER LA
	𑚽	SIRMAURI LETTER SA
	𑚾	SIRMAURI LETTER HA
Various signs (4)	◌𑚶	SIRMAURI SIGN ANUSVARA
	◌𑚷	SIRMAURI SIGN VISARGA
	◌𑚸	SIRMAURI SIGN HALANTA
	◌𑚹	SIRMAURI SIGN NUKTA
Consonant sign (1)	◌𑚻	SIRMAURI CONSONANT SIGN MEDIAL RA
Auspicious sign (1)	𑚿	SIRMAURI EKAM
Digits (10)	𑛀	SIRMAURI DIGIT ZERO
	𑛁	SIRMAURI DIGIT ONE
	𑛂	SIRMAURI DIGIT TWO
	𑛃	SIRMAURI DIGIT THREE
	𑛄	SIRMAURI DIGIT FOUR

ᑋ	SIRMAURI DIGIT FIVE
ᑌ	SIRMAURI DIGIT SIX
ᑍ	SIRMAURI DIGIT SEVEN
ᑎ	SIRMAURI DIGIT EIGHT
ᑏ	SIRMAURI DIGIT NINE

8 Character Data

Character Properties: `UnicodeData.txt`

```
11850;SIRMAURI LETTER A;Lo;0;L;;;;;N;;;;;
11851;SIRMAURI LETTER I;Lo;0;L;;;;;N;;;;;
11852;SIRMAURI LETTER U;Lo;0;L;;;;;N;;;;;
11853;SIRMAURI LETTER E;Lo;0;L;;;;;N;;;;;
11854;SIRMAURI LETTER KA;Lo;0;L;;;;;N;;;;;
11855;SIRMAURI LETTER KHA;Lo;0;L;;;;;N;;;;;
11856;SIRMAURI LETTER GA;Lo;0;L;;;;;N;;;;;
11857;SIRMAURI LETTER GHA;Lo;0;L;;;;;N;;;;;
11858;<reserved>
11859;SIRMAURI LETTER CA;Lo;0;L;;;;;N;;;;;
1185A;SIRMAURI LETTER CHA;Lo;0;L;;;;;N;;;;;
1185B;SIRMAURI LETTER JA;Lo;0;L;;;;;N;;;;;
1185C;SIRMAURI LETTER JHA;Lo;0;L;;;;;N;;;;;
1185D;<reserved>
1185E;SIRMAURI LETTER TTA;Lo;0;L;;;;;N;;;;;
1185F;SIRMAURI LETTER TTHA;Lo;0;L;;;;;N;;;;;
11860;SIRMAURI LETTER DDA;Lo;0;L;;;;;N;;;;;
11861;SIRMAURI LETTER DDHA;Lo;0;L;;;;;N;;;;;
11862;SIRMAURI LETTER NNA;Lo;0;L;;;;;N;;;;;
11863;SIRMAURI LETTER TA;Lo;0;L;;;;;N;;;;;
11864;SIRMAURI LETTER THA;Lo;0;L;;;;;N;;;;;
11865;SIRMAURI LETTER DA;Lo;0;L;;;;;N;;;;;
11866;SIRMAURI LETTER DHA;Lo;0;L;;;;;N;;;;;
11867;SIRMAURI LETTER NA;Lo;0;L;;;;;N;;;;;
11868;SIRMAURI LETTER PA;Lo;0;L;;;;;N;;;;;
11869;SIRMAURI LETTER PHA;Lo;0;L;;;;;N;;;;;
1186A;SIRMAURI LETTER BA;Lo;0;L;;;;;N;;;;;
1186B;SIRMAURI LETTER BHA;Lo;0;L;;;;;N;;;;;
1186C;SIRMAURI LETTER MA;Lo;0;L;;;;;N;;;;;
1186D;SIRMAURI LETTER YA;Lo;0;L;;;;;N;;;;;
1186E;SIRMAURI LETTER RA;Lo;0;L;;;;;N;;;;;
1186F;SIRMAURI LETTER LA;Lo;0;L;;;;;N;;;;;
11870;<reserved>
11871;<reserved>
11872;<reserved>
```

```
11873;SIRMAURI LETTER SA;Lo;0;L;;;;;N;;;;;
11874;SIRMAURI LETTER HA;Lo;0;L;;;;;N;;;;;
11875;SIRMAURI VOWEL SIGN AA;Mc;0;L;;;;;N;;;;;
11876;<reserved>
11877;SIRMAURI VOWEL SIGN II;Mc;0;L;;;;;N;;;;;
11878;SIRMAURI VOWEL SIGN U;Mn;0;NSM;;;;;N;;;;;
11879;<reserved>
1187A;SIRMAURI VOWEL SIGN E;Mn;0;NSM;;;;;N;;;;;
1187B;SIRMAURI VOWEL SIGN AI;Mn;0;NSM;;;;;N;;;;;
1187C;SIRMAURI VOWEL SIGN O;Mc;0;L;1187A 11875;;;;N;;;;;
1187D;SIRMAURI VOWEL SIGN AU;Mc;0;L;1187B 11875;;;;N;;;;;
1187E;SIRMAURI SIGN ANUSVARA;Mn;0;NSM;;;;;N;;;;;
1187F;SIRMAURI SIGN VISARGA;Mc;0;L;;;;;N;;;;;
11880;SIRMAURI SIGN HALANTA;Mn;9;NSM;;;;;N;;;;;
11881;SIRMAURI SIGN NUKTA;Mn;0;NSM;;;;;N;;;;;
11882;SIRMAURI CONSONANT SIGN MEDIAL RA;Mn;0;NSM;;;;;N;;;;;
11883;SIRMAURI EKAM;So;0;L;;;;;N;;;;;
11884;<reserved>
11885;<reserved>
11886;SIRMAURI DIGIT ZERO;Nd;0;L;;0;0;0;N;;;;;
11887;SIRMAURI DIGIT ONE;Nd;0;L;;1;1;1;N;;;;;
11888;SIRMAURI DIGIT TWO;Nd;0;L;;2;2;2;N;;;;;
11889;SIRMAURI DIGIT THREE;Nd;0;L;;3;3;3;N;;;;;
1188A;SIRMAURI DIGIT FOUR;Nd;0;L;;4;4;4;N;;;;;
1188B;SIRMAURI DIGIT FIVE;Nd;0;L;;5;5;5;N;;;;;
1188C;SIRMAURI DIGIT SIX;Nd;0;L;;6;6;6;N;;;;;
1188D;SIRMAURI DIGIT SEVEN;Nd;0;L;;7;7;7;N;;;;;
1188E;SIRMAURI DIGIT EIGHT;Nd;0;L;;8;8;8;N;;;;;
1188F;SIRMAURI DIGIT NINE;Nd;0;L;;9;9;9;N;;;;;
```

8.1 Linebreaking Properties: `LineBreak.txt`

```
11850..11857   ; AL # Lo     [8] SIRMAURI LETTER A..SIRMAURI LETTER GHA
11859..1185C   ; AL # Lo     [4] SIRMAURI LETTER CA..SIRMAURI LETTER JHA
1185E..1186F   ; AL # Lo    [18] SIRMAURI LETTER TTA..SIRMAURI LETTER LA
11873..11874   ; AL # Lo     [2] SIRMAURI LETTER SA..SIRMAURI LETTER HA
11875          ; CM # Mc         SIRMAURI VOWEL SIGN AA
11877          ; CM # Mc         SIRMAURI VOWEL SIGN II
11878          ; CM # Mn         SIRMAURI VOWEL SIGN U
1187A..1187B   ; CM # Mn     [2] SIRMAURI VOWEL SIGN E..SIRMAURI VOWEL SIGN AI
1187C..1187D   ; CM # Mc     [2] SIRMAURI VOWEL SIGN O..SIRMAURI VOWEL SIGN AU
1187E          ; CM # Mn         SIRMAURI SIGN ANUSVARA
1187F          ; CM # Mc         SIRMAURI SIGN VISARGA
11880..11882   ; CM # Mn     [2] SIRMAURI SIGN HALANTA..SIRMAURI CONSONANT SIGN MEDIAL RA
11883          ; AL # Lo         SIRMAURI EKAM
11886..1188F   ; NU # Nd    [10] SIRMAURI DIGIT ZERO..SIRMAURI DIGIT NINE
```

8.2 Syllabic Categories: IndicSyllabicCategory.txt

```
# Indic_Syllabic_Category=Bindu
1187E         ; Bindu               # Mc        SIRMAURI SIGN ANUSVARA

# Indic_Syllabic_Category=Visarga
1187F         ; Visarga             # Mc        SIRMAURI SIGN VISARGA

# Indic_Syllabic_Category=Pure_Killer
11880         ; Pure_Killer         # Mn        SIRMAURI SIGN HALANTA

# Indic_Syllabic_Category=Vowel_Independent
11850..11853  ; Vowel_Independent   # Lo    [4] SIRMAURI LETTER A..SIRMAURI LETTER E

# Indic_Syllabic_Category=Vowel_Dependent
11875         ; Vowel_Dependent     # Mc        SIRMAURI VOWEL SIGN AA
11877         ; Vowel_Dependent     # Mc        SIRMAURI VOWEL SIGN II
11878         ; Vowel_Dependent     # Mn        SIRMAURI VOWEL SIGN U
1187A..1187B  ; Vowel_Dependent     # Mn    [2] SIRMAURI VOWEL SIGN E..SIRMAURI VOWEL SIGN AI
1187C..1187D  ; Vowel_Dependent     # Mc    [2] SIRMAURI VOWEL SIGN O..SIRMAURI VOWEL SIGN AU

# Indic_Syllabic_Category=Consonant
11854..11857  ; Consonant           # Lo    [4] SIRMAURI LETTER KA..SIRMAURI LETTER GHA
11859..1185C  ; Consonant           # Lo    [4] SIRMAURI LETTER CA..SIRMAURI LETTER JHA
1185E..1186F  ; Consonant           # Lo   [18] SIRMAURI LETTER TTA..SIRMAURI LETTER LA
11873..11874  ; Consonant           # Lo    [2] SIRMAURI LETTER SA..SIRMAURI LETTER HA

# Indic_Syllabic_Category=Consonant_Medial
11882         ; Consonant_Medial    # Mn        SIRMAURI CONSONANT SIGN MEDIAL RA

# Indic_Syllabic_Category=Number
11886..1188F  ; Number              # Nd   [10] SIRMAURI DIGIT ZERO..SIRMAURI DIGIT NINE
```

8.3 Positional Categories: IndicPositionalCategory.txt

```
# Indic_Positional_Category=Top
1187E         ; Top                 # Mn        SIRMAURI SIGN ANUSVARA
1187A..1187B  ; Top                 # Mn    [2] SIRMAURI VOWEL SIGN E..SIRMAURI VOWEL SIGN AI

# Indic_Positional_Category=Bottom
11878         ; Bottom              # Mn        SIRMAURI VOWEL SIGN U ..
11880..11881  ; Bottom              # Mn    [2] SIRMAURI SIGN HALANTA..SIRMAURI SIGN NUKTA
11882         ; Bottom              # Mn        SIRMAURI CONSONANT SIGN MEDIAL RA

# Indic_Positional_Category=Right
11875         ; Right               # Mc        SIRMAURI VOWEL SIGN AA
11877         ; Right               # Mc        SIRMAURI VOWEL SIGN II
1187F         ; Right               # Mc        SIRMAURI SIGN VISARGA
```

```
# Indic_Positional_Category=Top_And_Right
1187C..1187D  ;  Top_And_Right     # Mc    [2] SIRMAURI VOWEL SIGN O..SIRMAURI VOWEL SIGN AU
```

8.4 Script Extensions: `ScriptExtensions.txt`

```
0964 ; ... # Po DEVANGARI DANDA
0965 ; ... # Po DEVANGARI DOUBLE DANDA
```

9 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.

———. 1916b. *The Linguistic Survey of India*. Vol. IX. Indo-Aryan Family. Central Group. Part IV. Specimens of Pahāṛī 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. 2009a. "Proposal to encode the Sharada script in Unicode" (L2/09-074R2).
https://www.unicode.org/L2/L2009/09074r2-n3595-sharada.pdf

———. 2009b. "Proposal to Encode the Takri Script in ISO/IEC 10646" (L2/09-424).
http://www.unicode.org/L2/L2009/09424-takri.pdf

———. 2011a. "Final Proposal to Encode the Khojki Script in ISO/IEC 10646" (L2/11-021).
https://www.unicode.org/L2/L2011/11021-khojki.pdf

———. 2011b. "Proposal to Encode the Mahajani Script in ISO/IEC 10646" (L2/11-274).
https://www.unicode.org/L2/L2011/11274-n4126-mahajani.pdf

———. 2015. "Proposal to encode the Dogra script in Unicode" (L2/15-234R).
http://www.unicode.org/L2/L2015/15234r-dogra.pdf

———. 2018. "Preliminary proposal to encode Sirmauri in Unicode" (L2/18-085).
https://www.unicode.org/L2/L2018/18085-sirmauri.pdf

———. 2021. "Proposal to encode the Khojki letter SHORT I in Unicode" (L2/21-104).
https://www.unicode.org/L2/L2021/21104-khojki-letter-short-i.pdf

Pandey, Anshuman; Mandal, Biswajit. 2025a. "Proposal to encode Sirmauri in Unicode" (L2/25-038).
https://www.unicode.org/L2/L2025/25038-sirmauri.pdf

———. 2025b. "Proposal to encode Sirmauri in Unicode" (L2/25-134).
https://www.unicode.org/L2/L2025/25134-sirmauri.pdf

Punjab Government. 1907. *Sirmur State Gazetteer, 1904*. Part A. Lahore: The Punjab Government Press.

10 Acknowledgments

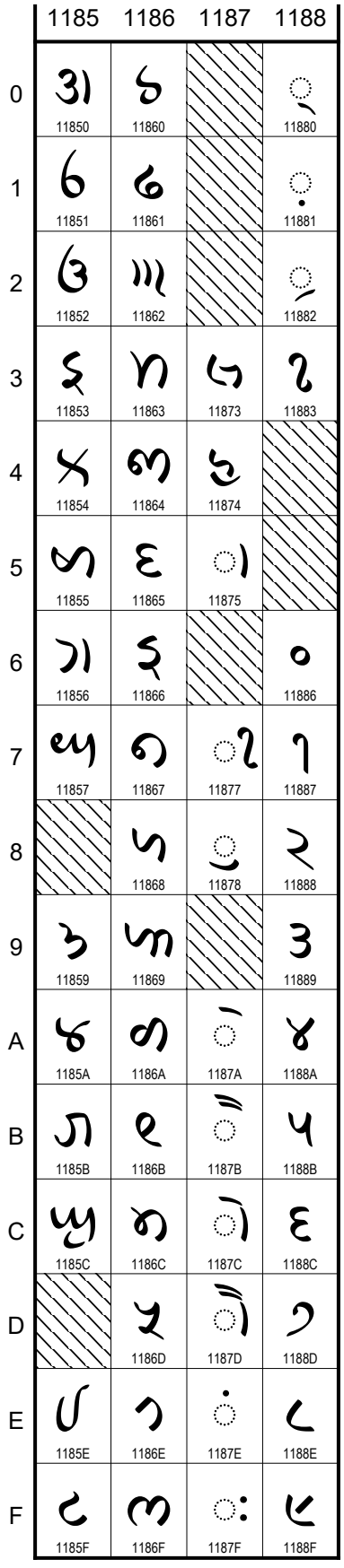
Our work to encode Sirmauri would not be possible without assistance from the following, to whom we express our heartfelt gratitude:

- Nikhil Bhardwaj (‘Project Pahari’: `http://x.com/project_pahari`)
- Sherjung Chauhan (Sirmaur)
- Bhupal Singh Chauhan (Nohradhar, Sirmaur)
- The family of Natiram Negi (Maleata, Jaunsar-Bawar)
- Gangaram Pabuch (Sirmaur)
- Yatin Pandit (Kullu)
- Dr. Om Prakash Sharma (Shimla)
- Dr. Dilip Singh Tilkan (Sirmaur)
- The family of Har Singh Tilkan (Shilla village, Sirmaur)
- The family of Ved Singh Tilkan (Shilla village, Sirmaur)

This project has been made possible in part by grants from:

- The U.S. National Endowment for the Humanities, which funded the Universal Scripts Project PR-253360-17 (part of the Script Encoding Initiative at the University of California, Berkeley)

- The Mellon Foundation to the Script Encoding Initiative at the University of California, Berkeley


1185 1186 1187 1188
0 11850 11860 11880
1 11851 11861 11881
2 11852 11862 11882
3 11853 11863 11873 11883
4 11854 11864 11874
5 11855 11865 11875
6 11856 11866 11886
7 11857 11867 11877 11887
8 11868 11878 11888
9 11859 11869 11889
A 1185A 1186A 1187A 1188A
B 1185B 1186B 1187B 1188B
C 1185C 1186C 1187C 1188C
D 1186D 1187D 1188D
E 1185E 1186E 1187E 1188E
F 1185F 1186F 1187F 1188F

Independent vowels

11850 𑐀 SIRMAURI LETTER A
11851 𑐁 SIRMAURI LETTER I
11852 𑐂 SIRMAURI LETTER U
11853 𑐃 SIRMAURI LETTER E

Consonants

11854 𑐄 SIRMAURI LETTER KA
11855 𑐅 SIRMAURI LETTER KHA
11856 𑐆 SIRMAURI LETTER GA
11857 𑐇 SIRMAURI LETTER GHA
11858 <reserved>
11859 𑐉 SIRMAURI LETTER CA
1185A 𑐊 SIRMAURI LETTER CHA
1185B 𑐋 SIRMAURI LETTER JA
1185C 𑐌 SIRMAURI LETTER JHA
1185D <reserved>
1185E 𑐎 SIRMAURI LETTER TTA
1185F 𑐏 SIRMAURI LETTER TTHA
11860 𑐐 SIRMAURI LETTER DDA
11861 𑐑 SIRMAURI LETTER DDHA
11862 𑐒 SIRMAURI LETTER NNA
11863 𑐓 SIRMAURI LETTER TA
11864 𑐔 SIRMAURI LETTER THA
11865 𑐕 SIRMAURI LETTER DA
11866 𑐖 SIRMAURI LETTER DHA
11867 𑐗 SIRMAURI LETTER NA
11868 𑐘 SIRMAURI LETTER PA
11869 𑐙 SIRMAURI LETTER PHA
1186A 𑐚 SIRMAURI LETTER BA
1186B 𑐛 SIRMAURI LETTER BHA
1186C 𑐜 SIRMAURI LETTER MA
1186D 𑐝 SIRMAURI LETTER YA
1186E 𑐞 SIRMAURI LETTER RA
1186F 𑐟 SIRMAURI LETTER LA
11870 <reserved>
11871 <reserved>
11872 <reserved>
11873 𑐣 SIRMAURI LETTER SA
11874 𑐤 SIRMAURI LETTER HA

Dependent vowel signs

11875 𑐥 SIRMAURI VOWEL SIGN AA
11876 <reserved>
11877 𑐧 SIRMAURI VOWEL SIGN II
11878 𑐨 SIRMAURI VOWEL SIGN U
11879 <reserved>
1187A 𑐪 SIRMAURI VOWEL SIGN E
1187B 𑐫 SIRMAURI VOWEL SIGN AI
1187C 𑐬 SIRMAURI VOWEL SIGN O
≡ 1187A 𑐪 11875 𑐥
1187D 𑐭 SIRMAURI VOWEL SIGN AU
≡ 1187B 𑐫 11875 𑐥

Various signs

1187E 𑐮 SIRMAURI SIGN ANUSVARA
1187F 𑐯 SIRMAURI SIGN VISARGA
11880 𑐰 SIRMAURI SIGN HALANTA
• inherent vowel silencer
• does not control conjunct formation
11881 𑐱 SIRMAURI SIGN NUKTA

Consonant sign

11882 𑐲 SIRMAURI CONSONANT SIGN MEDIAL RA
= ra-kara

Auspicious sign

11883 𑐳 SIRMAURI EKAM
= anji
11884 <reserved>
11885 <reserved>

Digits

11886 𑐴 SIRMAURI DIGIT ZERO
11887 𑐵 SIRMAURI DIGIT ONE
11888 𑐶 SIRMAURI DIGIT TWO
11889 𑐷 SIRMAURI DIGIT THREE
1188A 𑐸 SIRMAURI DIGIT FOUR
1188B 𑐹 SIRMAURI DIGIT FIVE
1188C 𑐺 SIRMAURI DIGIT SIX
1188D 𑐻 SIRMAURI DIGIT SEVEN
1188E 𑐼 SIRMAURI DIGIT EIGHT
1188F 𑐽 SIRMAURI DIGIT NINE

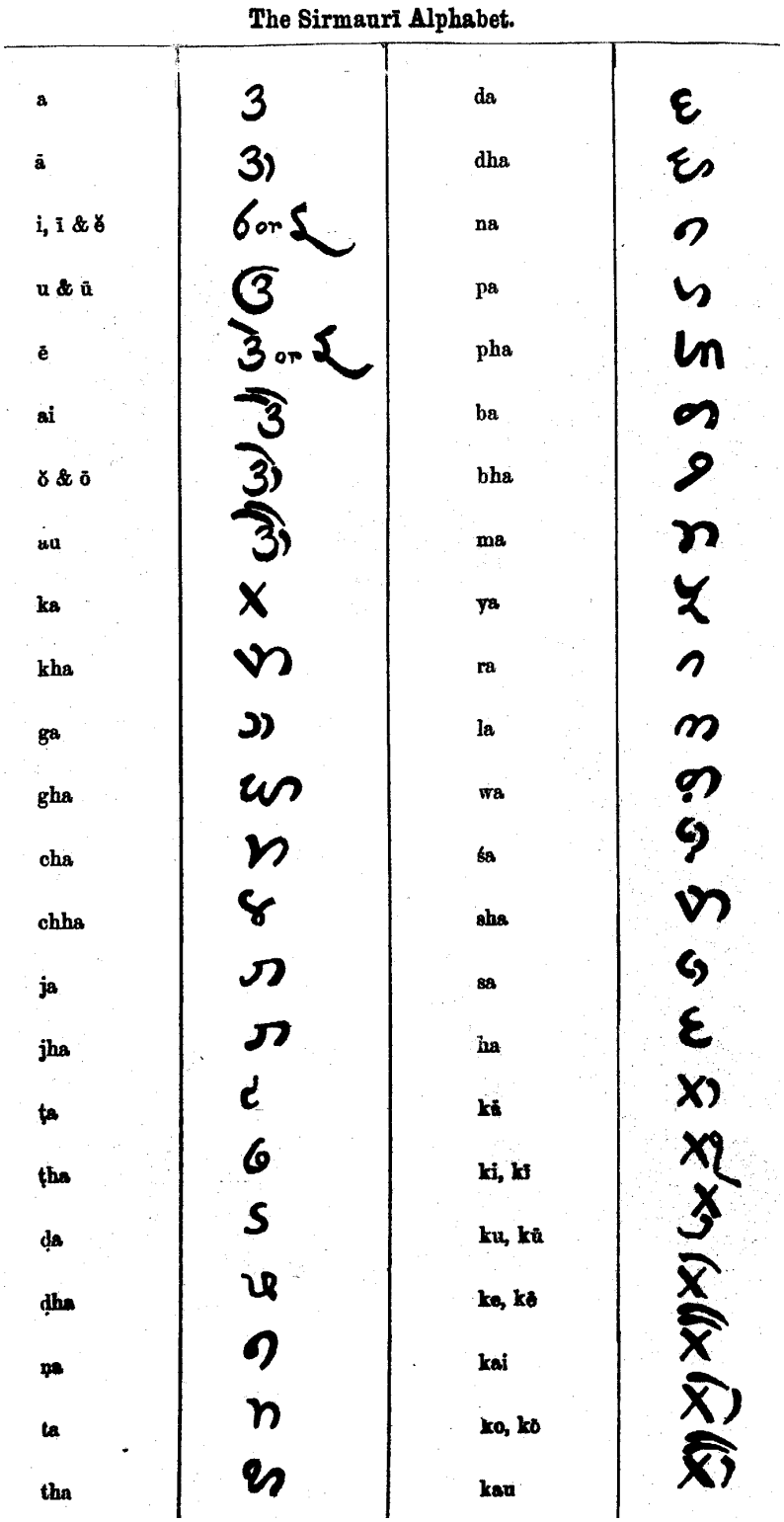


Figure 1: Chart of the Sirmauri form of Takri (from Grierson 1916b: 459).

WESTERN PAHĀṚĪ (SIRMAURĪ).

DHĀRṬHĪ DIALECT.

SPECIMEN I.

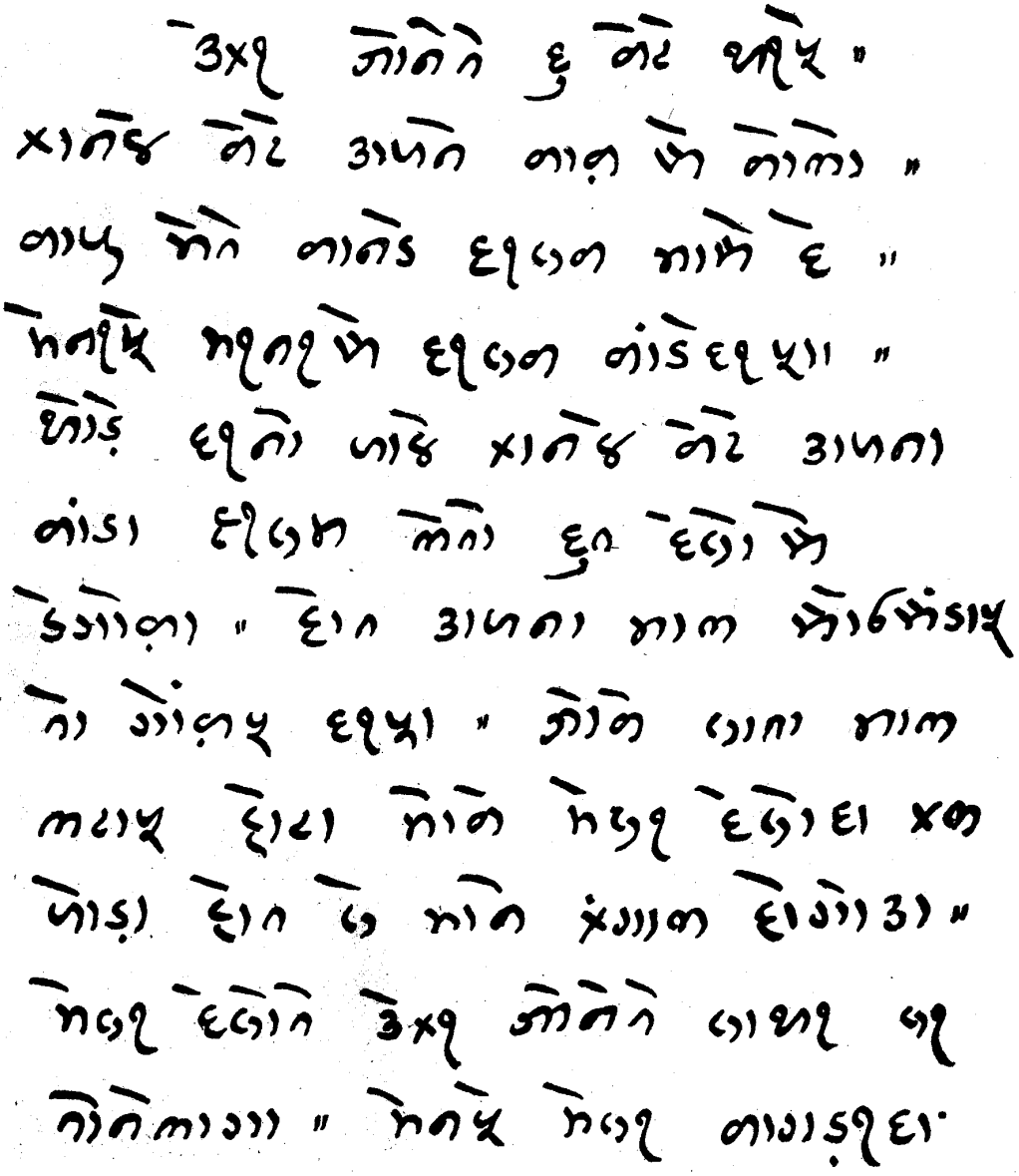


Figure 2: Specimen of the Sirmauri form of Takri (from Grierson 1916b: 468).

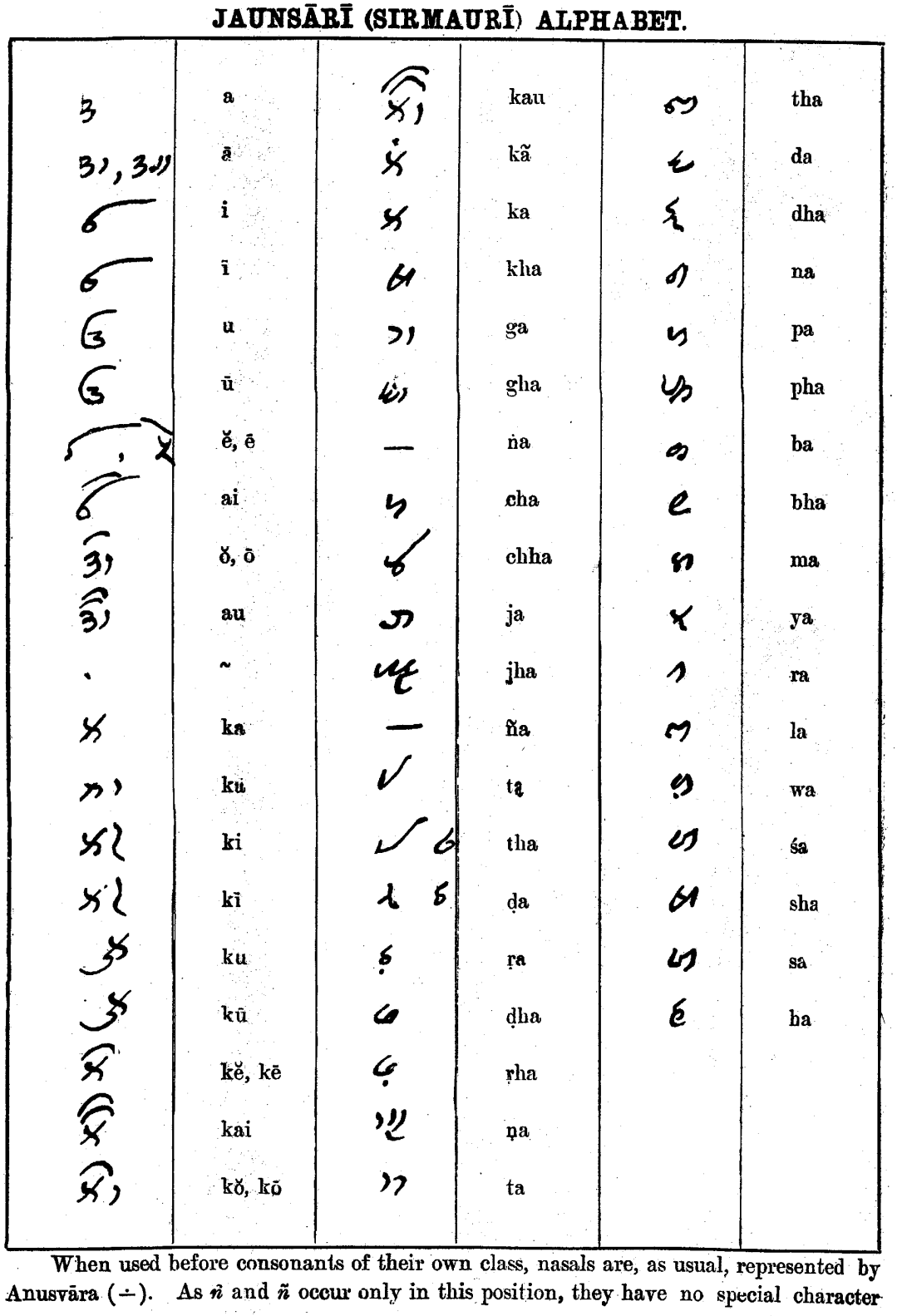


Figure 3: Chart of the Jaunsari form of Takri (from Grierson 1916b: 384).

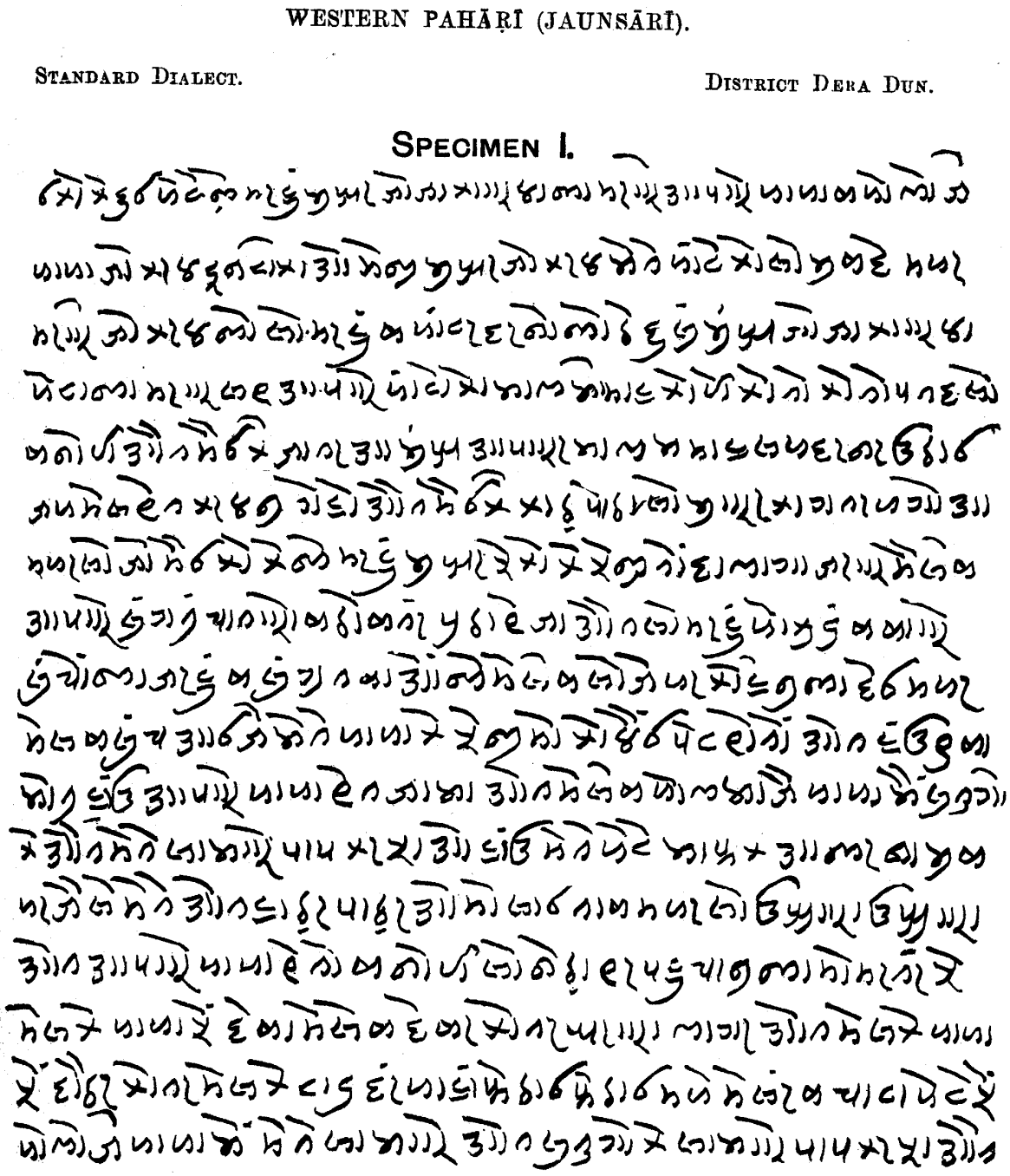
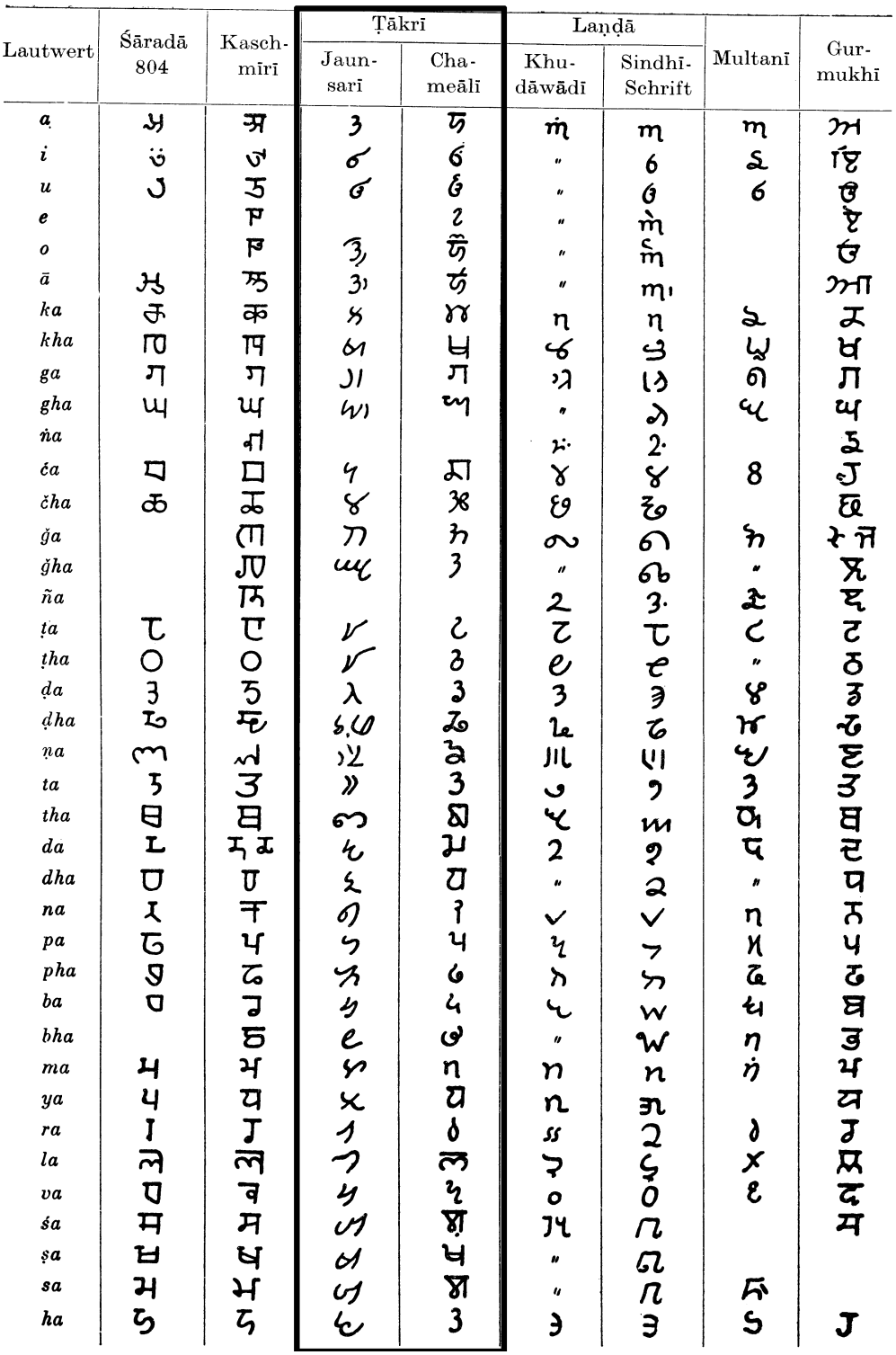


Figure 4: Specimen of the Jaunsari form of Takri (from Grierson 1916b: 401).



Lautwert	Śāradā 804	Kasch-mīrī	Ṭākrī		Laṇḍā		Multanī	Gur-mukhī
			Jaun-sarī	Cha-meālī	Khu-dāwādī	Sindhī-Schrift		
a								
i								
u								
e								
o								
ā								
ka								
kha								
ga								
gha								
ṅa								
ća								
čha								
ǧa								
ǧha								
ña								
ṭa								
ṭha								
ḍa								
ḍha								
ṇa								
ta								
tha								
da								
dha								
na								
pa								
pha								
ba								
bha								
ma								
ya								
ra								
la								
va								
śa								
ṣa								
sa								
ha								

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

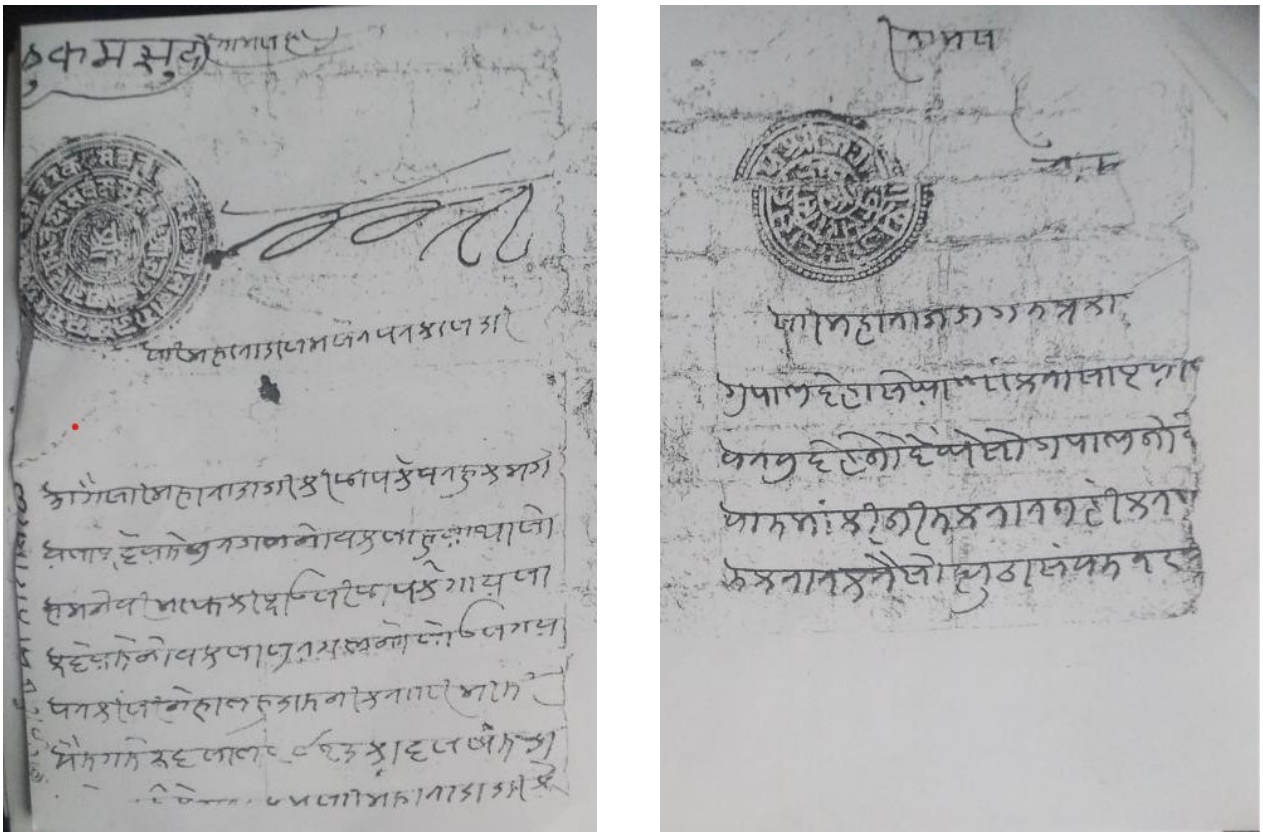


Figure 6: Documents recording land donations to the Shirgul Devta temple by *rājā* Shamsher Prakash (1770–1789) (left) and *rājā* Jagat Prakash (1856–1898) (right).

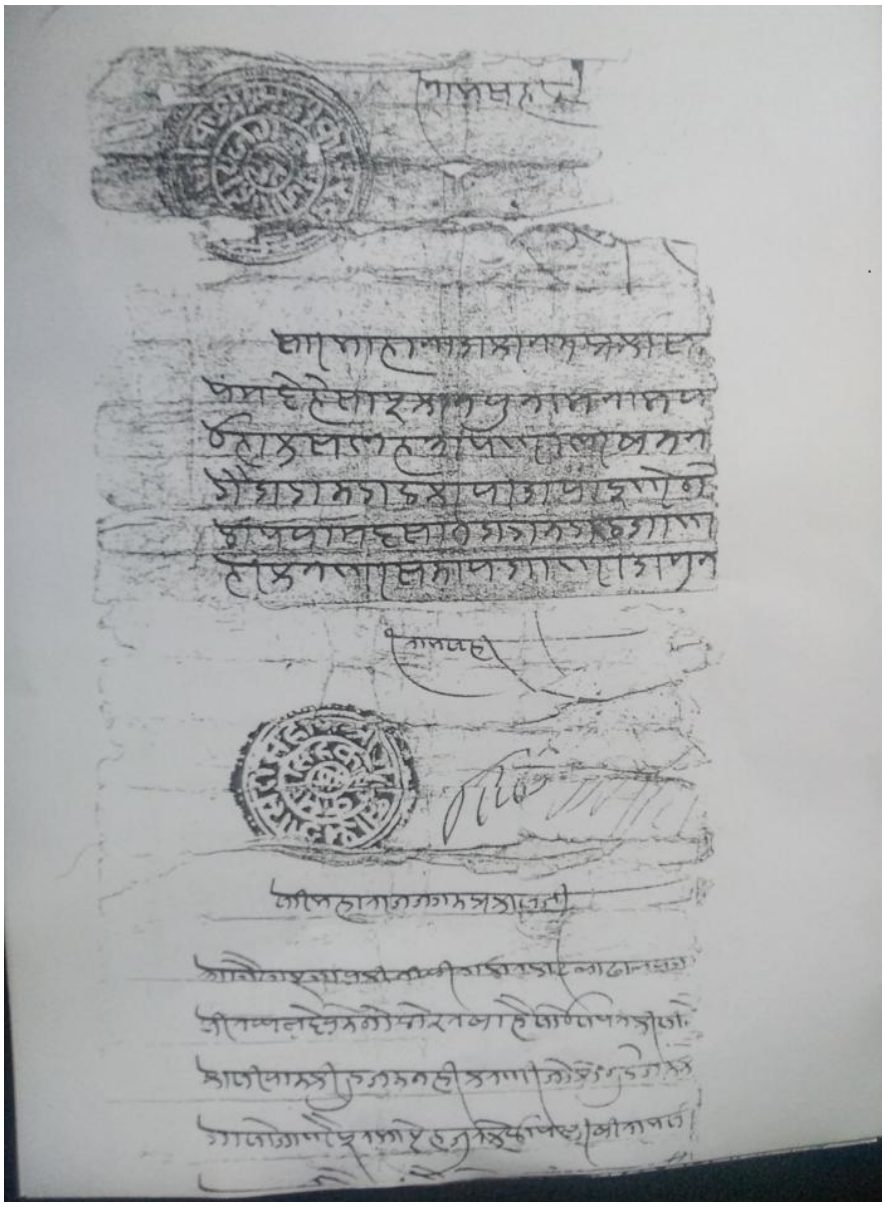


Figure 7: Record of land donation to the Shirgul Devta temple by *rājā* Karma Prakash (1804–1815).

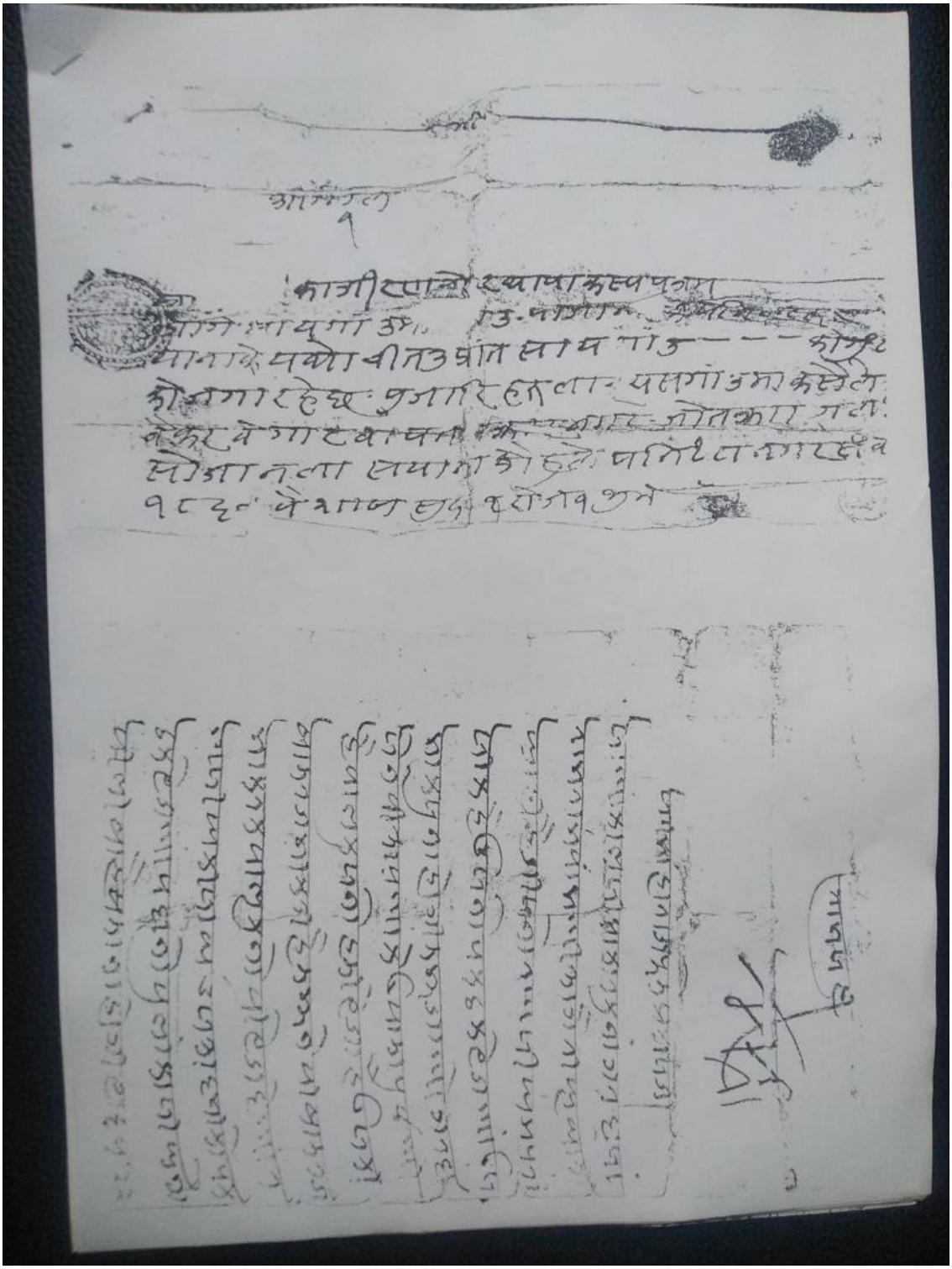


Figure 8: Record of land donation to the Shirgul Devta temple by *rājā* Fateh Prakash (1815–1850).

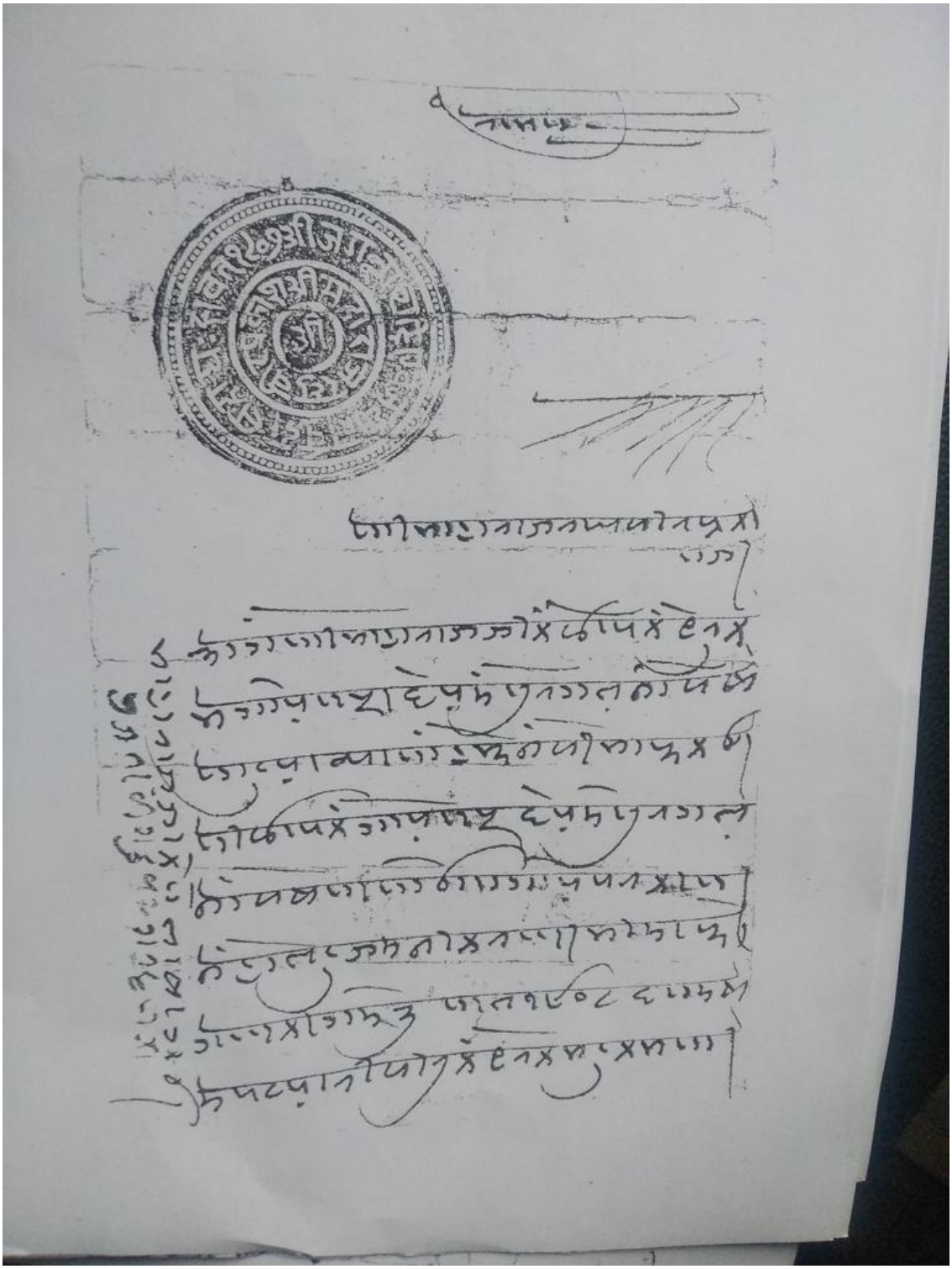


Figure 9: Record of land donation to the Shirgul Devta temple by Raghbir Prakash ll (1850–56).

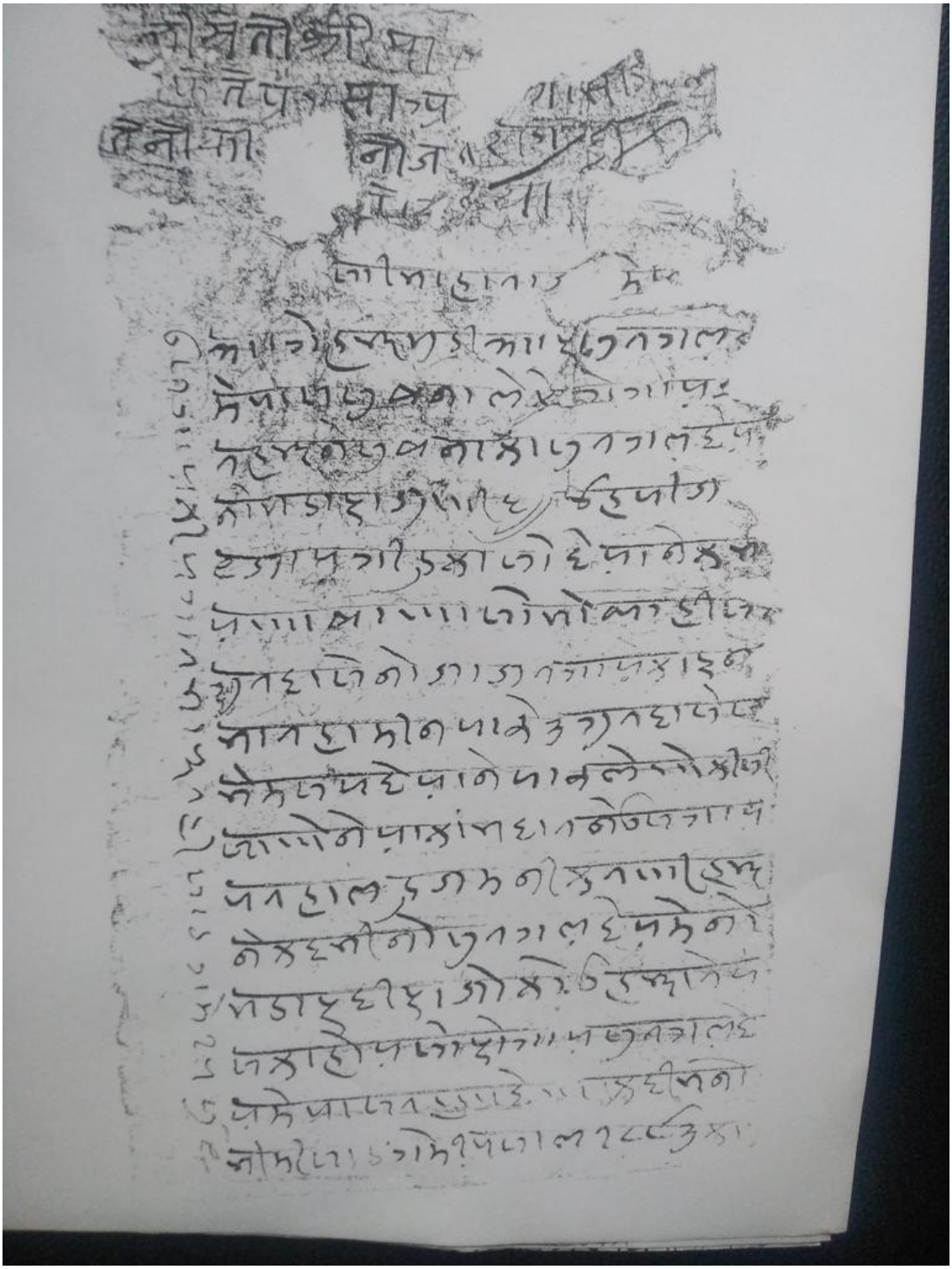


Figure 10: Record of land donation to the Shirgul Devta temple by *rājā* Raghbir Prakash ll (1850–56).

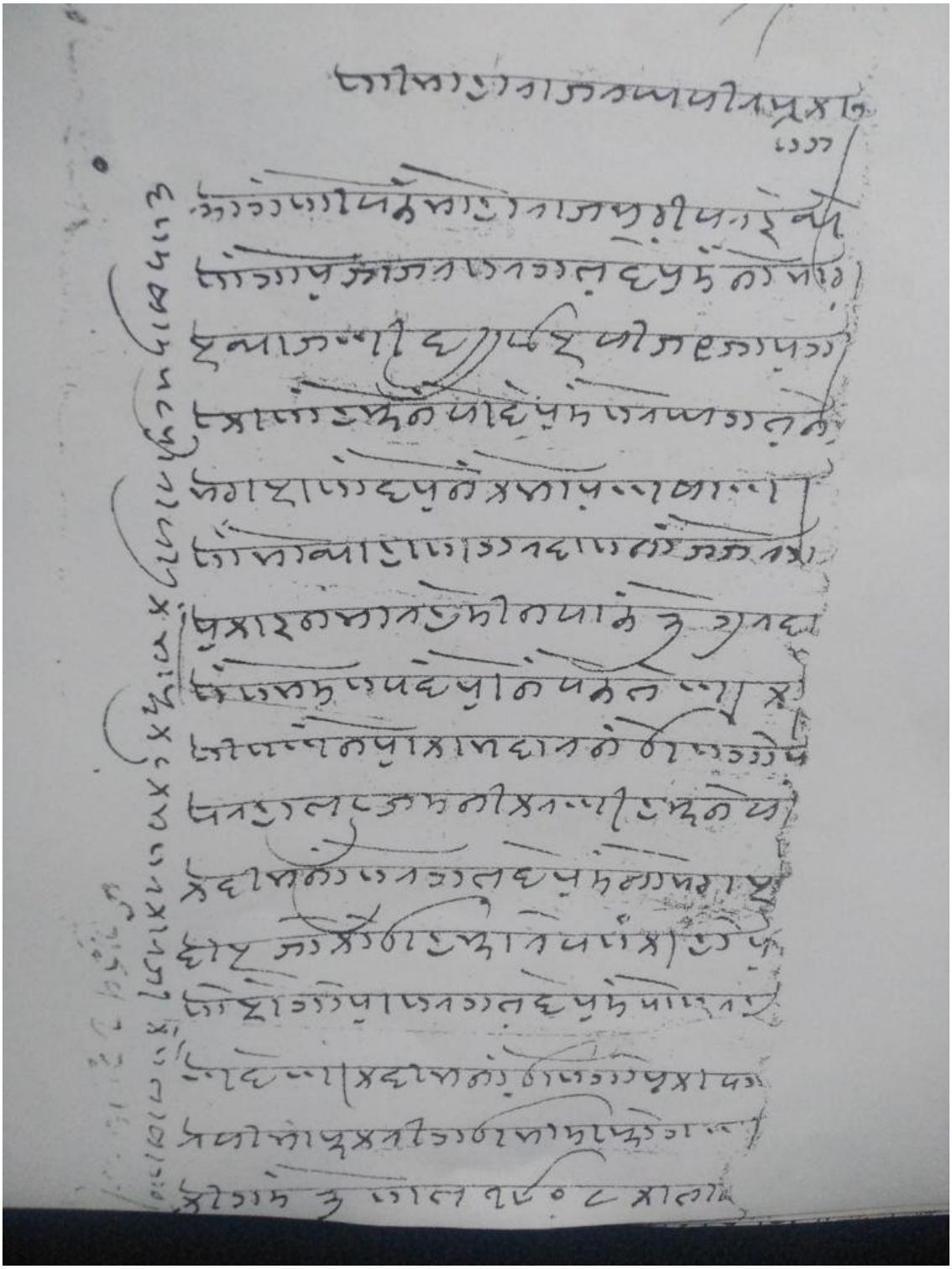


Figure 11: Record of land donation to the Shirgul Devta temple by *rājā* Raghbir Prakash ll (1850–56).

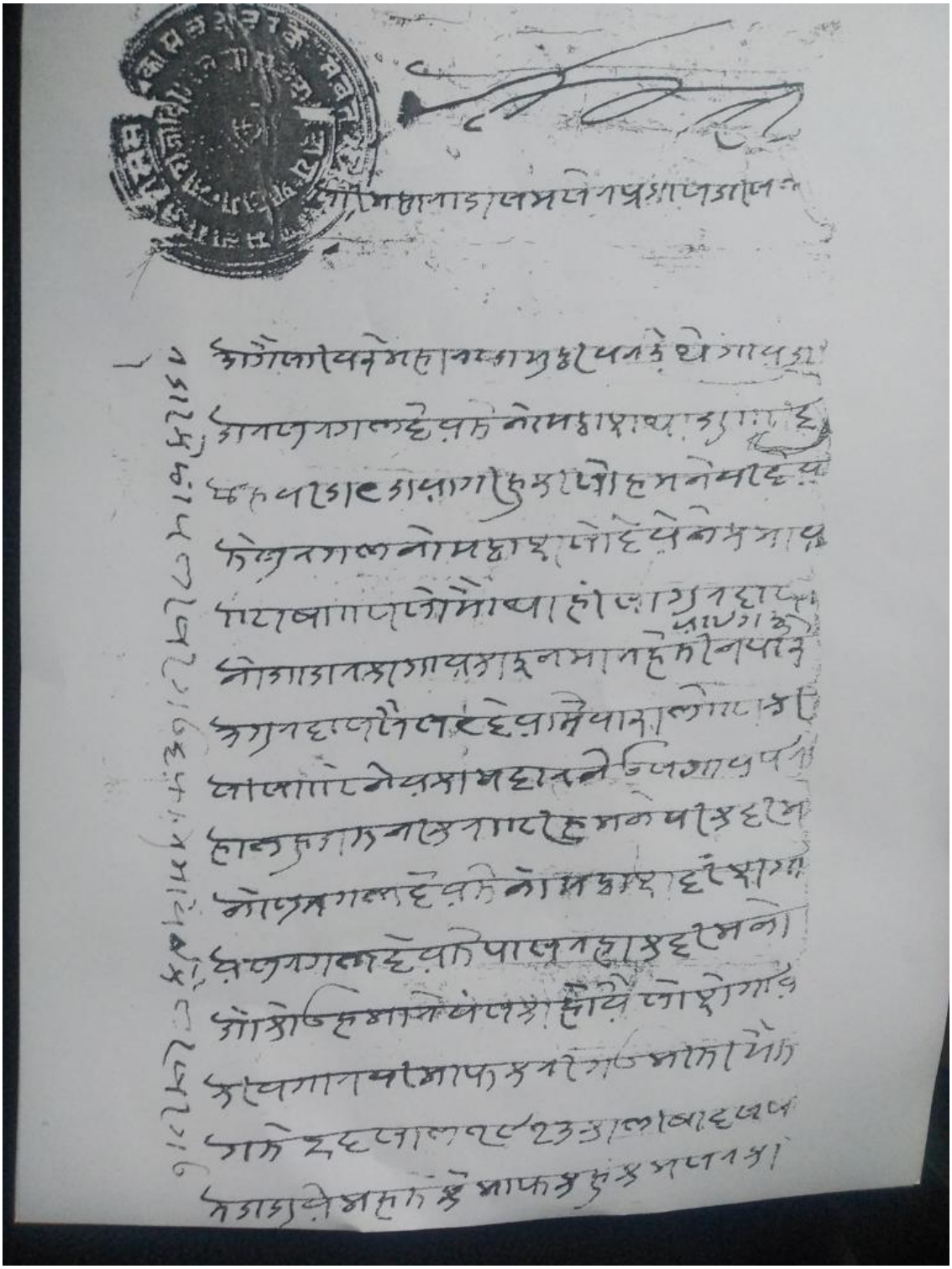


Figure 12: Record of land donation to the Shirgul Devta temple by *rājā* Shamsher Prakash (1856-98).

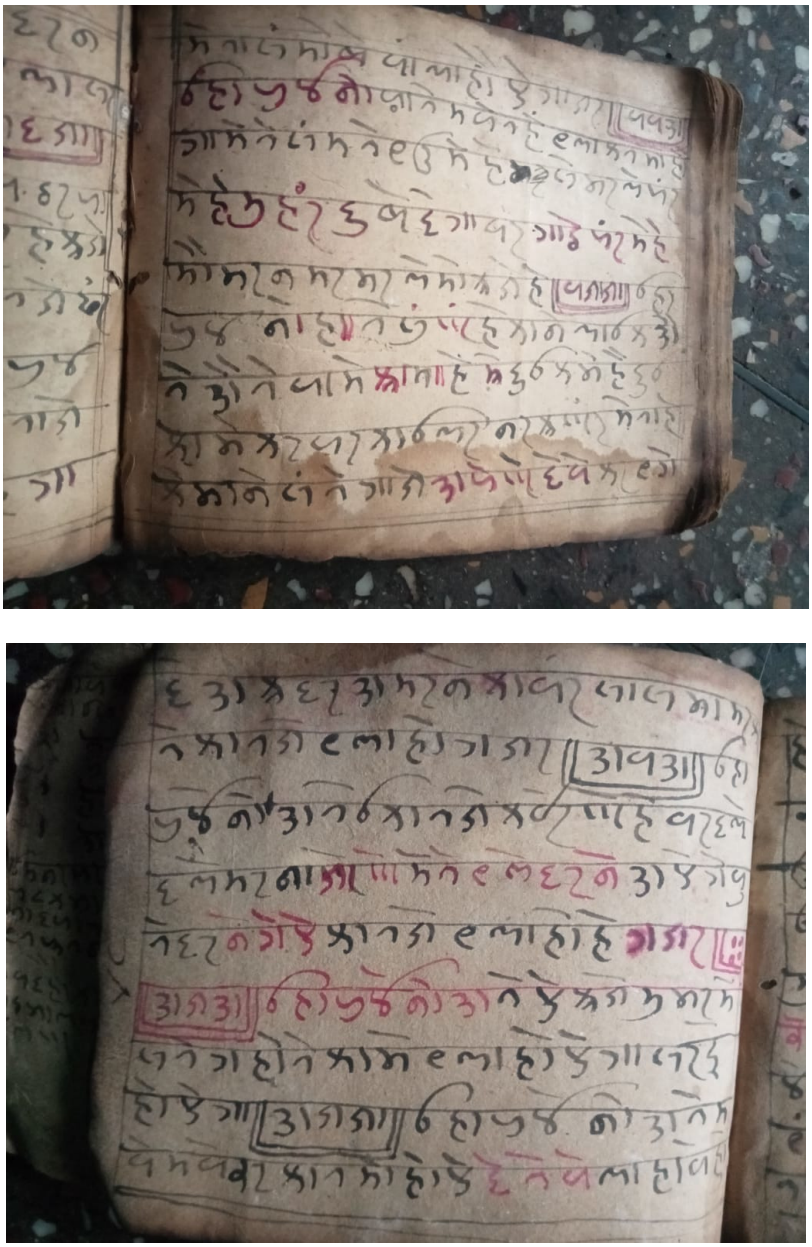


Figure 13: Sancha text written in the Jaunsari variety by Natiram Negi.

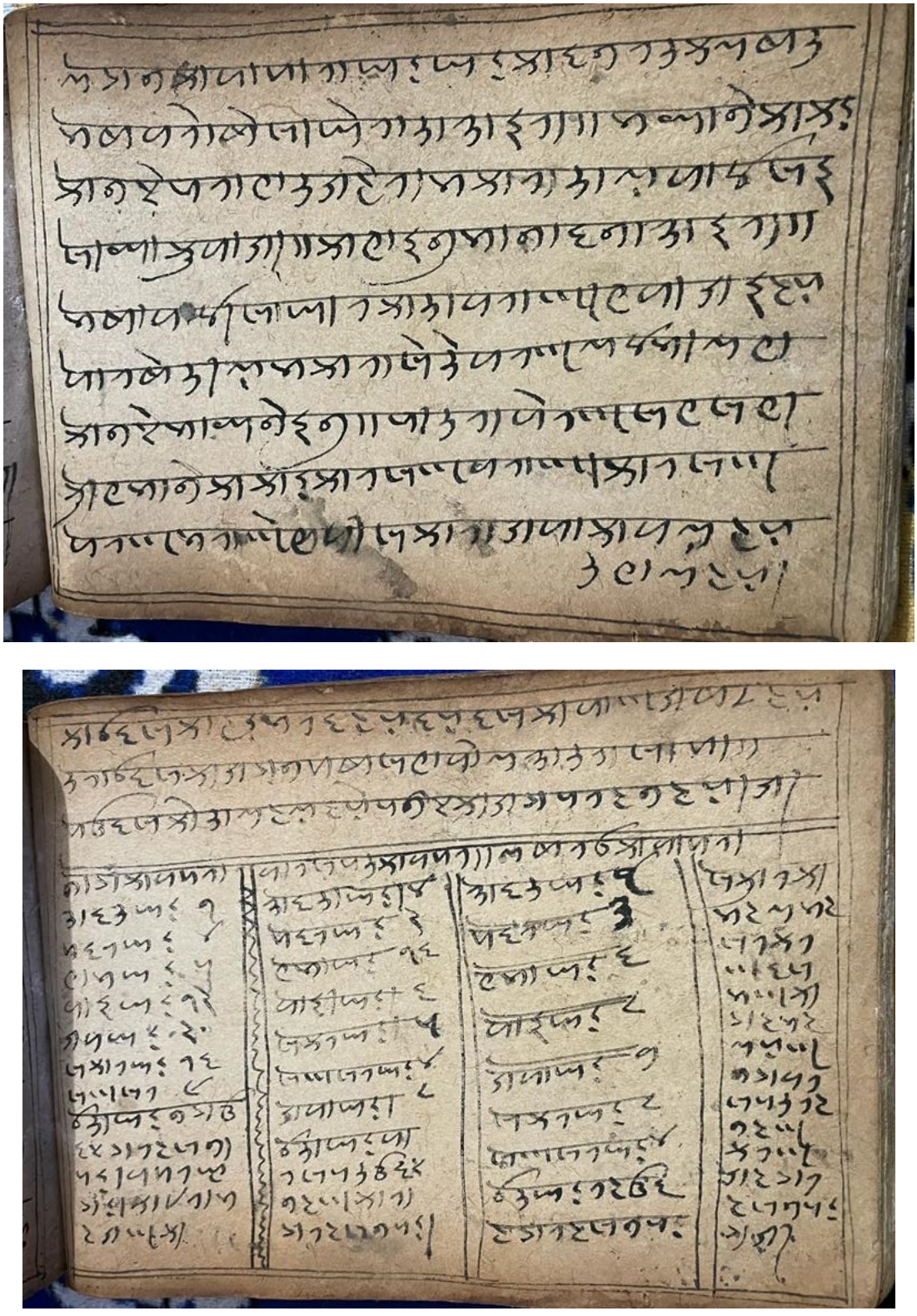


Figure 14: Sancha text written in Sirmauri by Gangaram Pabuch.

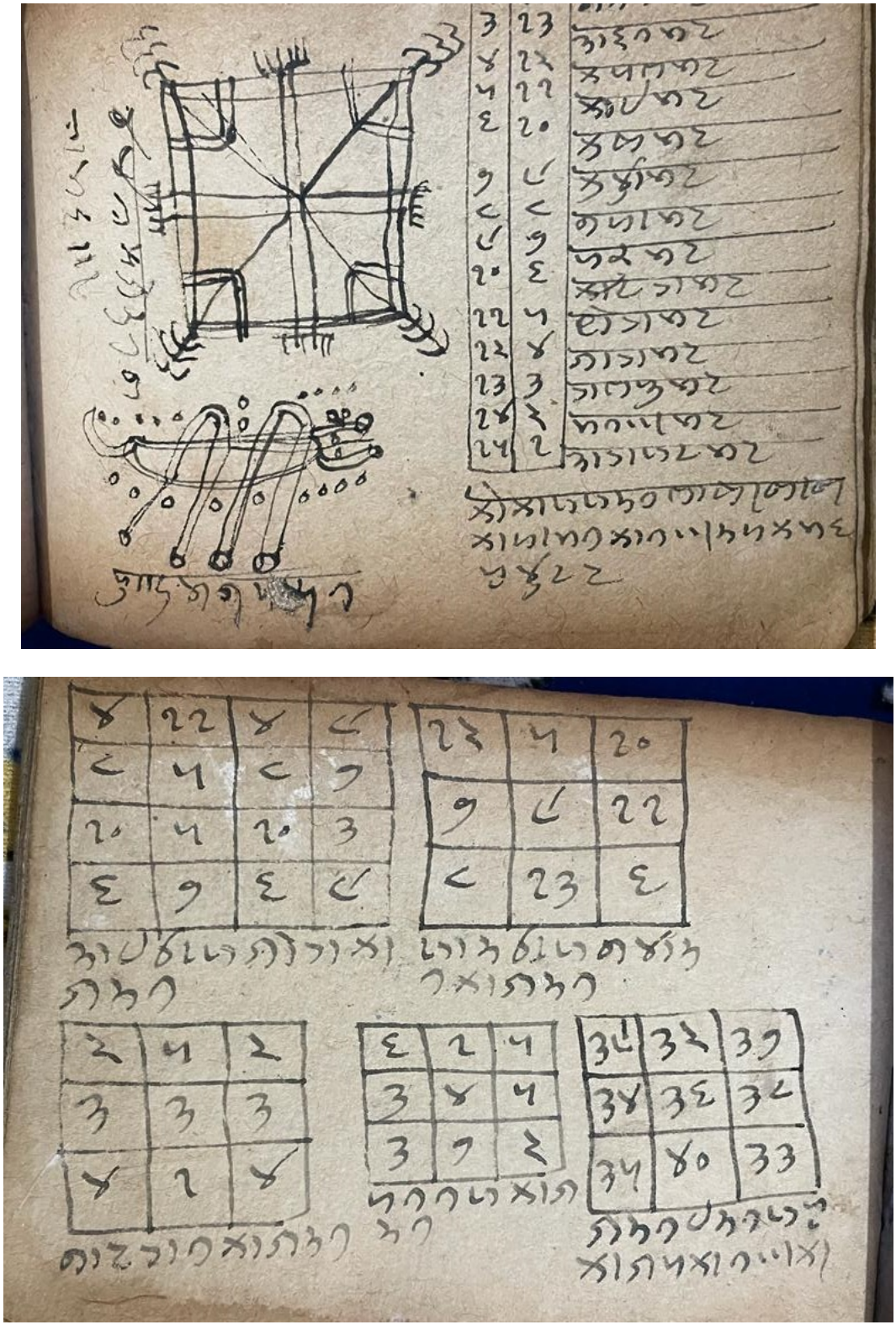


Figure 15: Sancha text written in Sirmauri by Gangaram Pabuch.

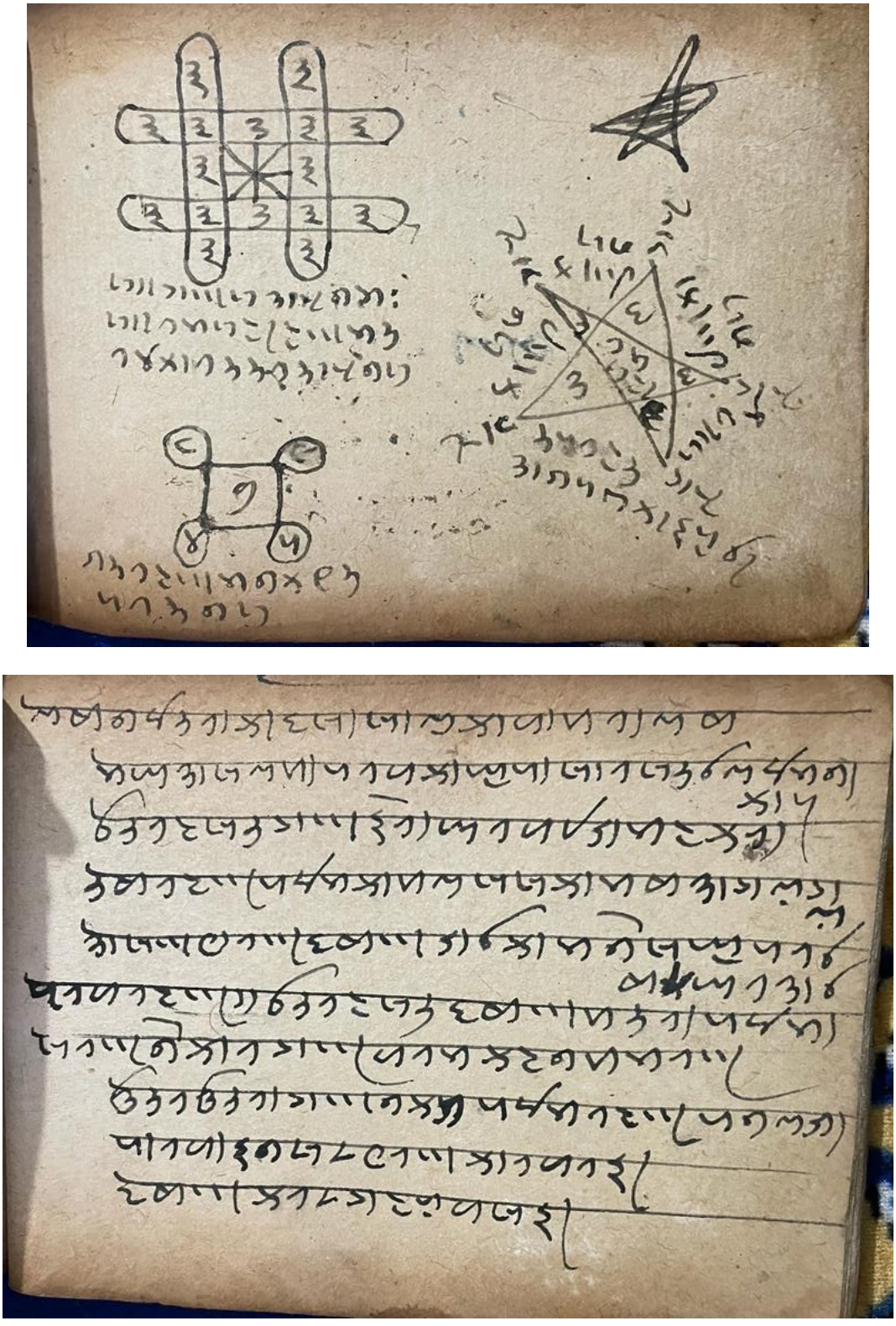


Figure 16: Sancha text written in Sirmauri by Gangaram Pabuch.

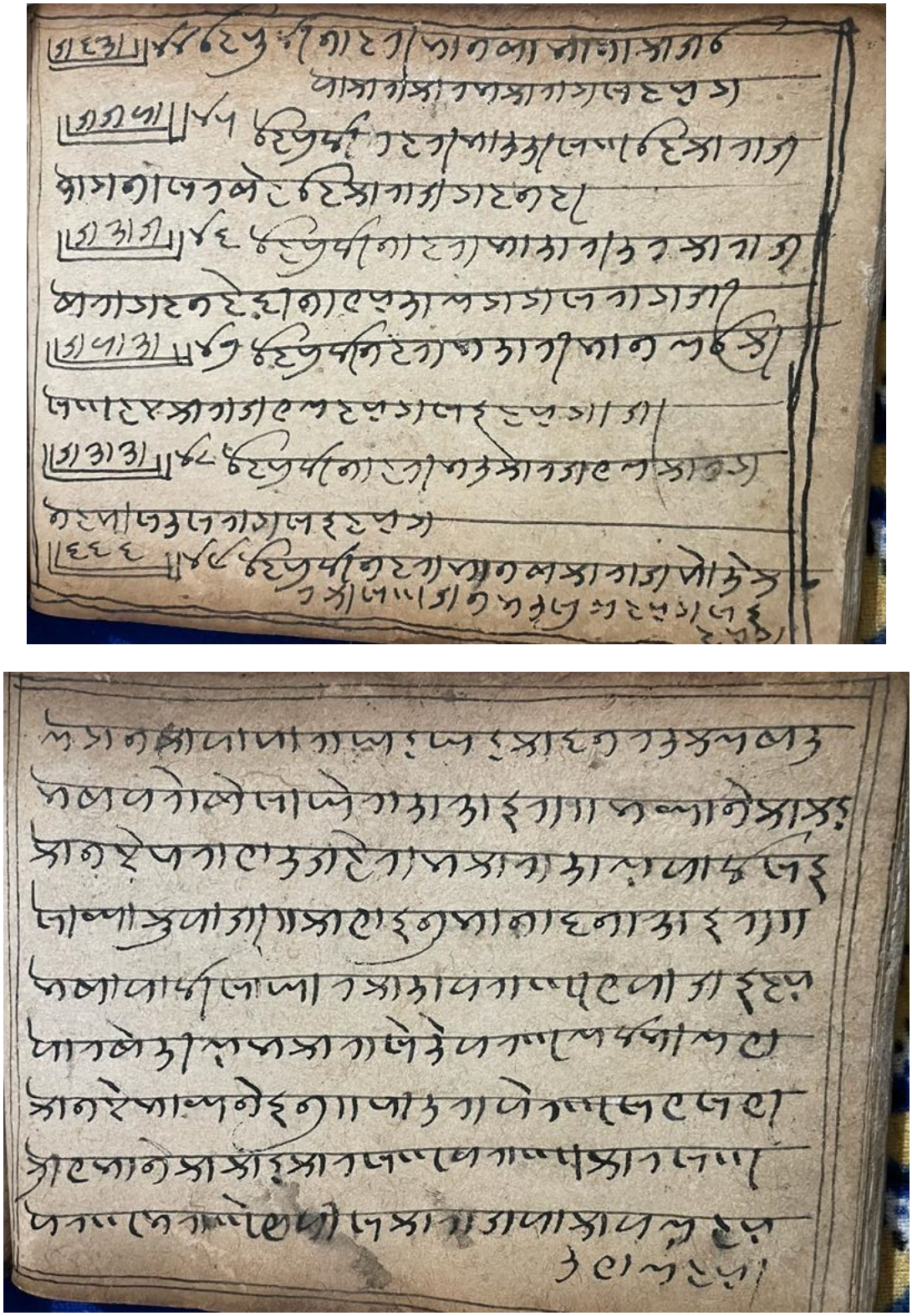


Figure 17: Sancha text written in Sirmauri by Gangaram Pabuch.

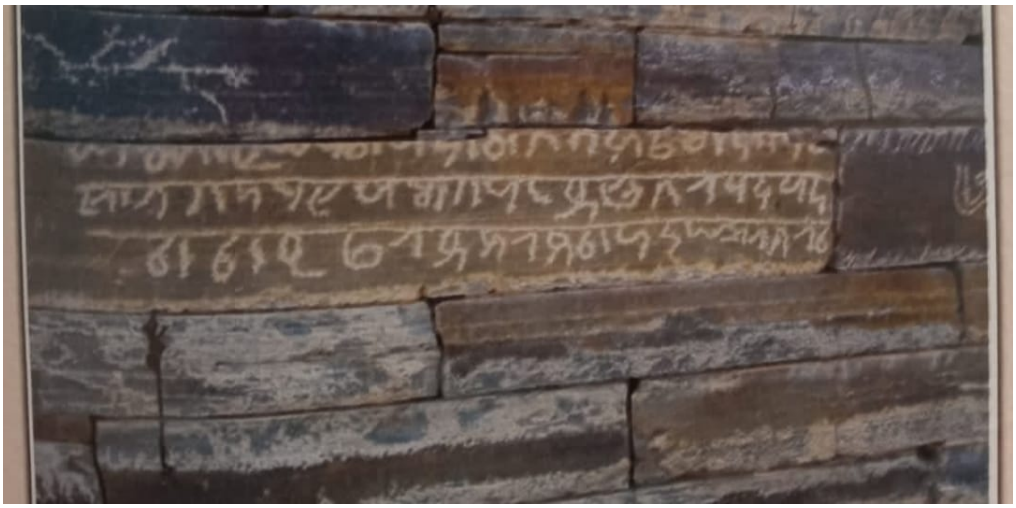


Figure 18: Inscription found on a tower (*killa*) adjacent to a house

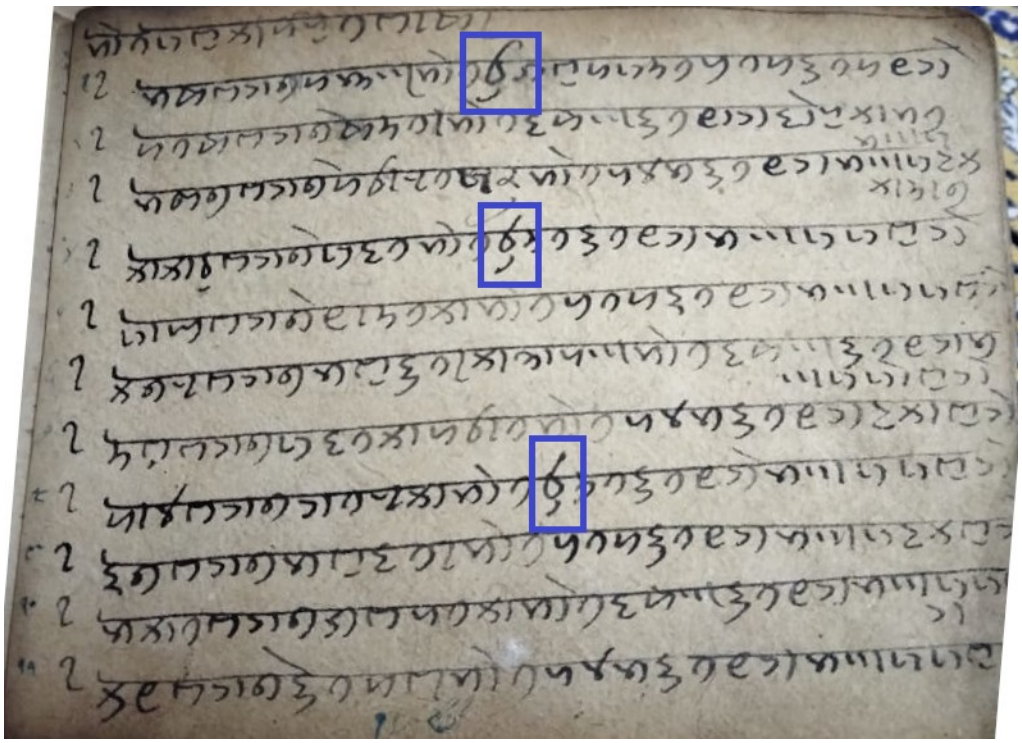


Figure 19: Examples of ੬ for writing *ī*.

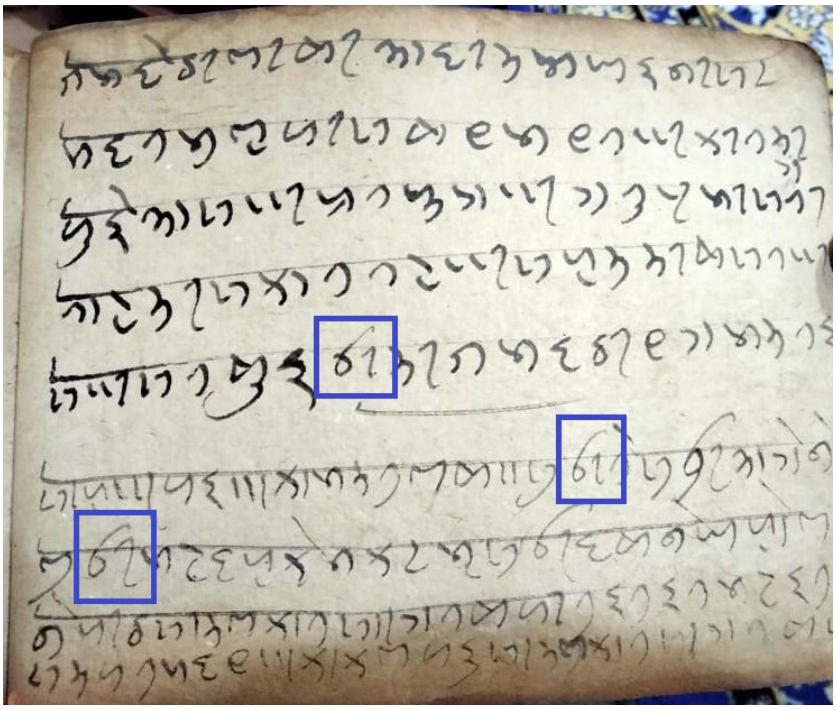


Figure 20: Example of 𑐏𑐲 for representing *ī*, illustrating usage of 𑐏 used as a vowel carrier bearing the vowel sign ◌𑐲 .

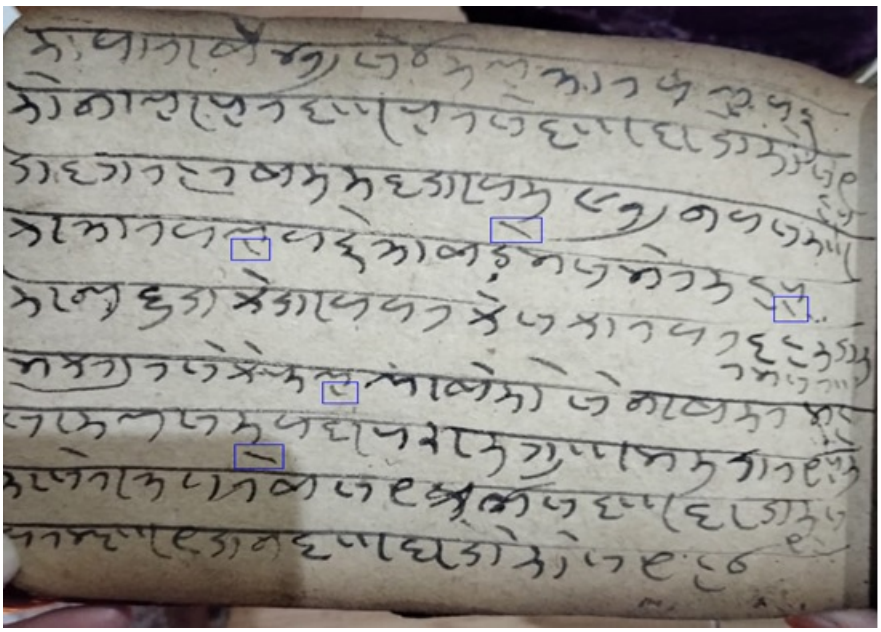


Figure 21: Usage of *halanta*.

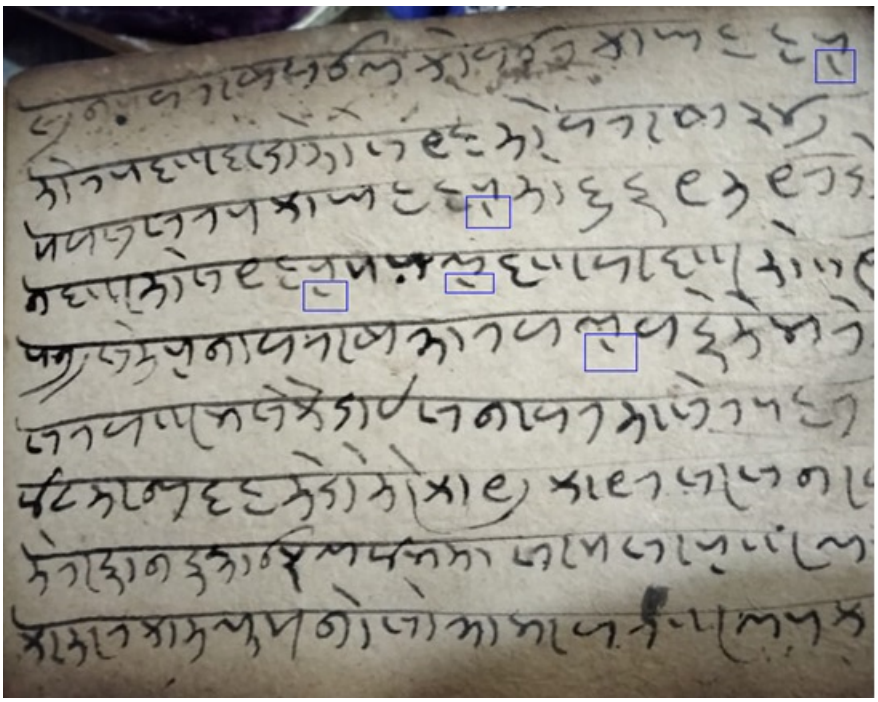


Figure 22: Additional examples of *halanta*.

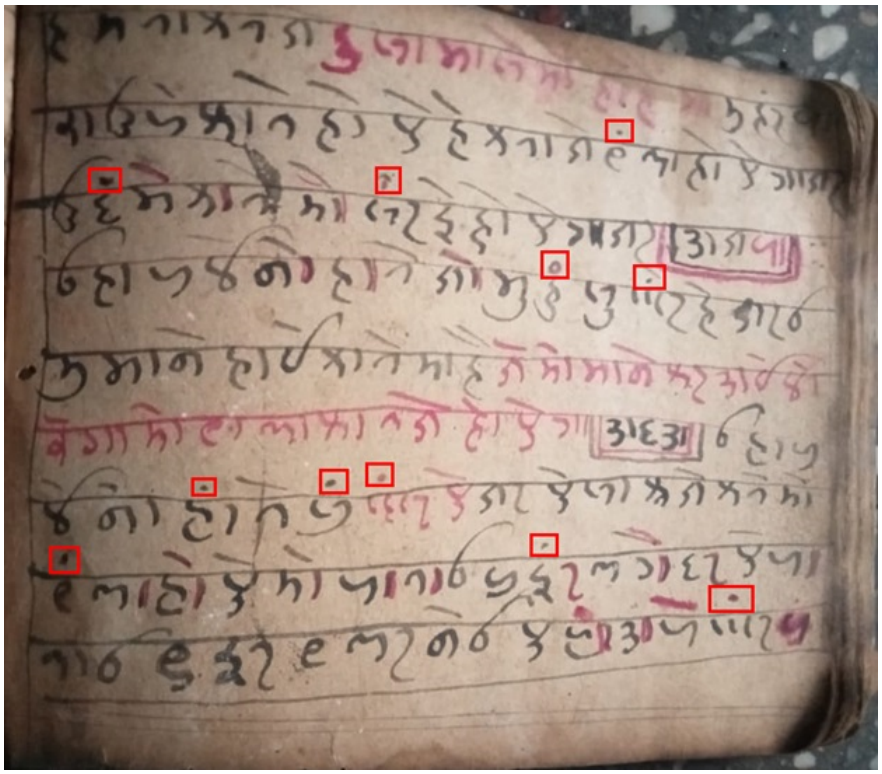


Figure 23: Usage of *anusvāra*

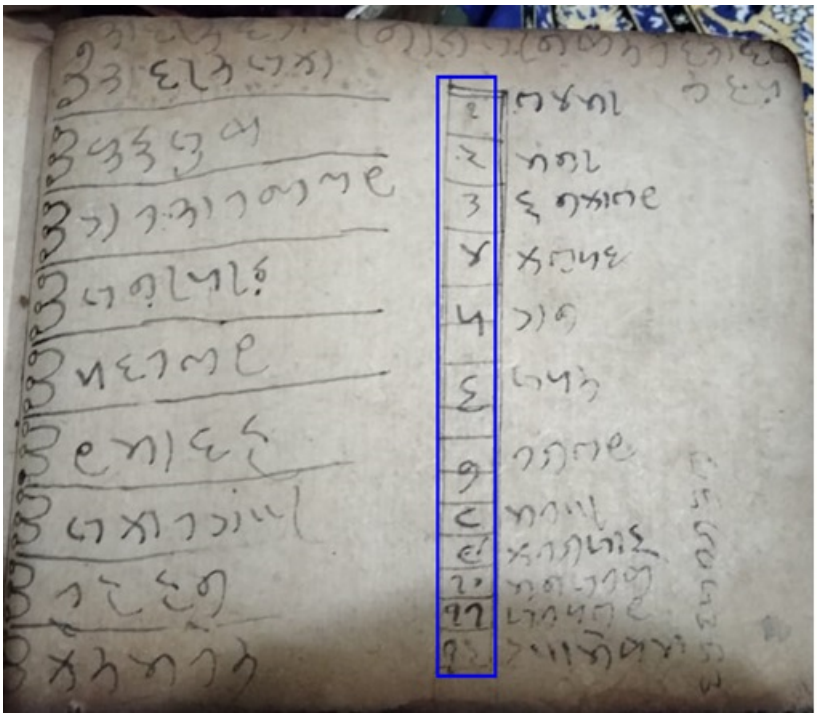


Figure 24: Examples of digits.

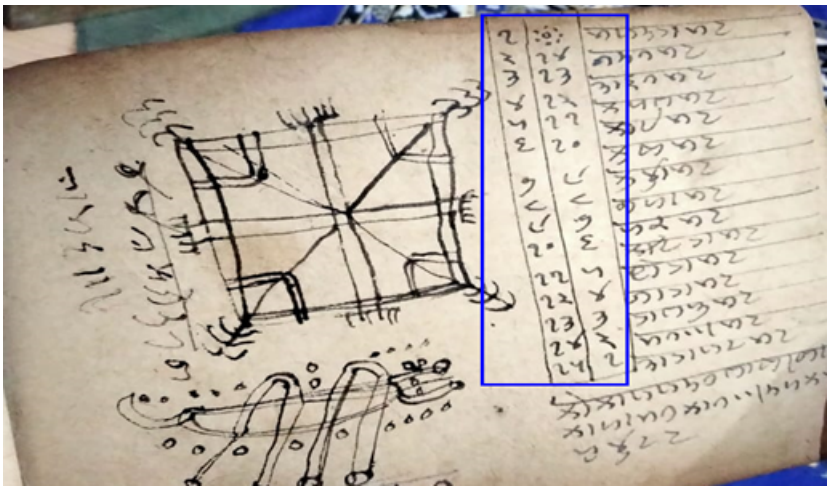


Figure 25: Examples of digits.

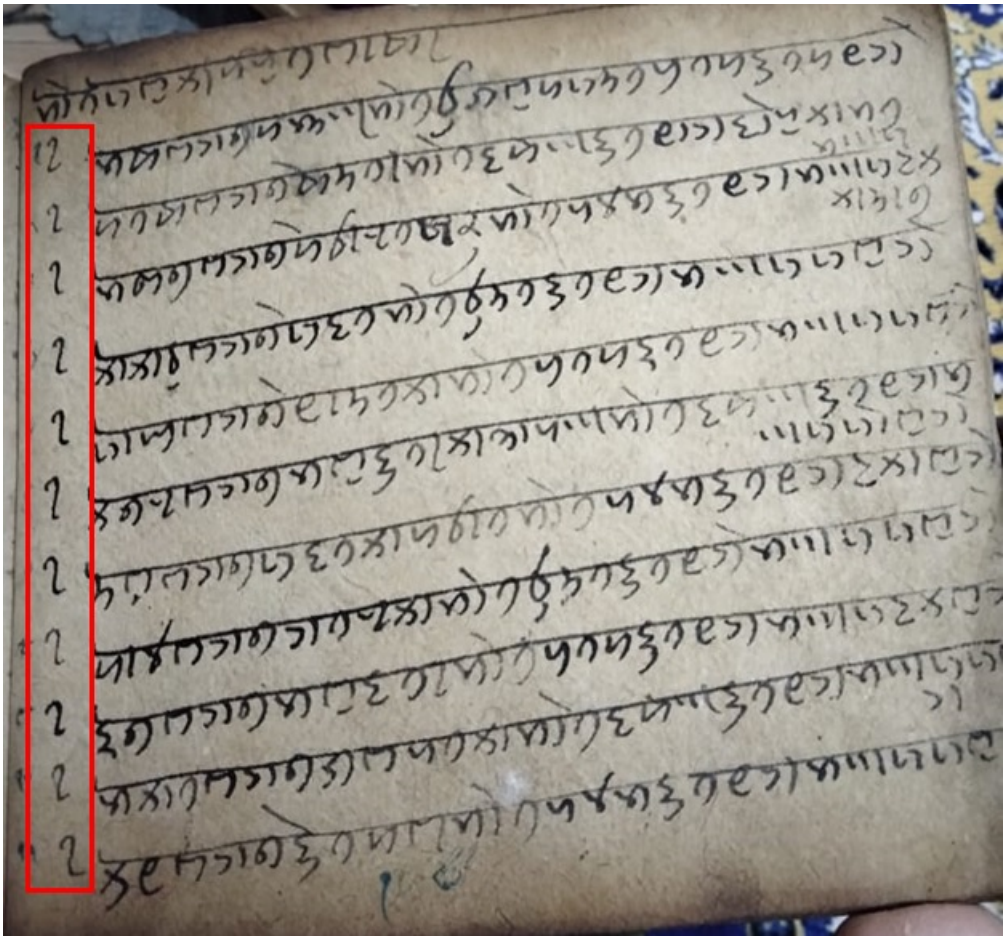


Figure 26: Usage of the auspicious sign 𑐷 *ekam*

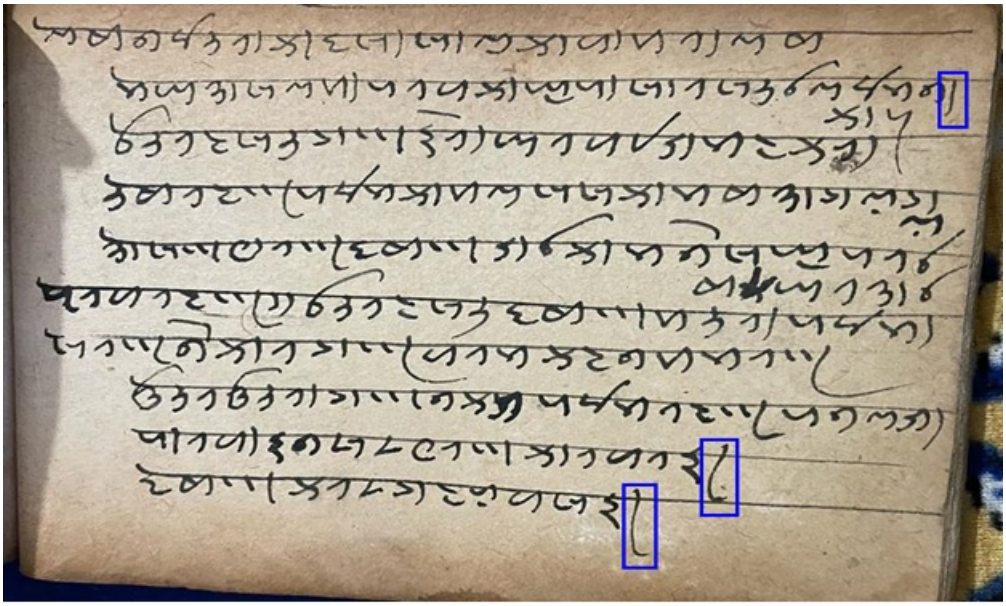


Figure 27: Usage of *daṇḍā*-s.

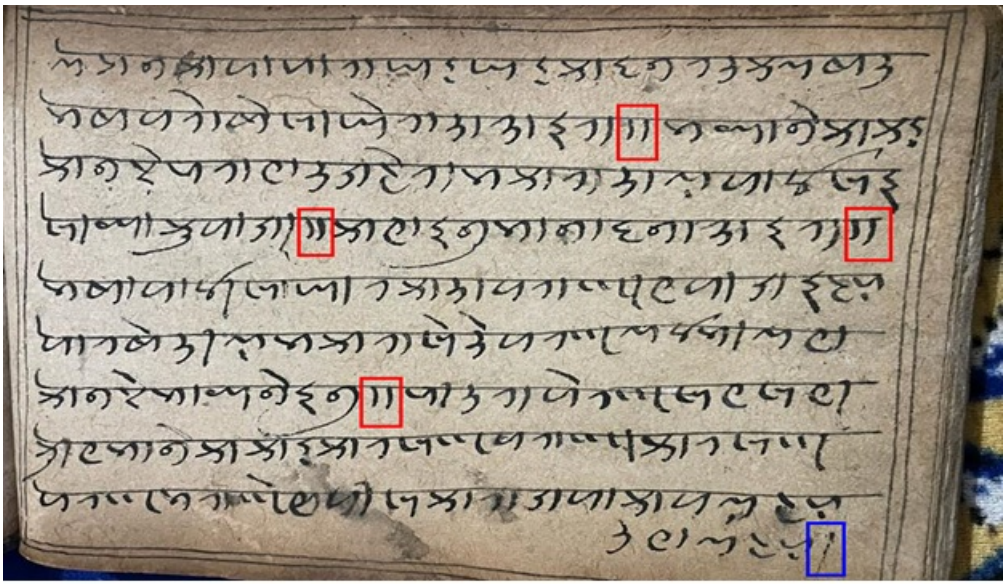


Figure 28: Single (blue) and double (red) *daṇḍā*-s.

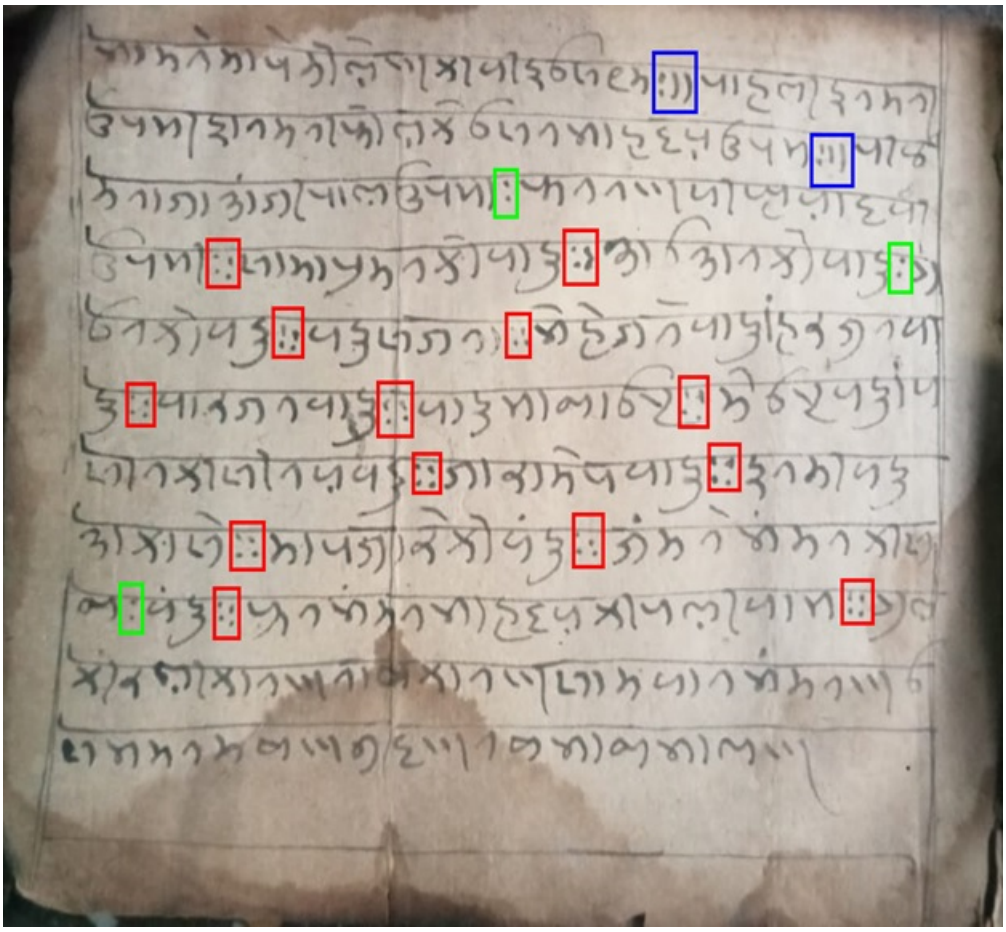


Figure 29: Various punctuation signs.

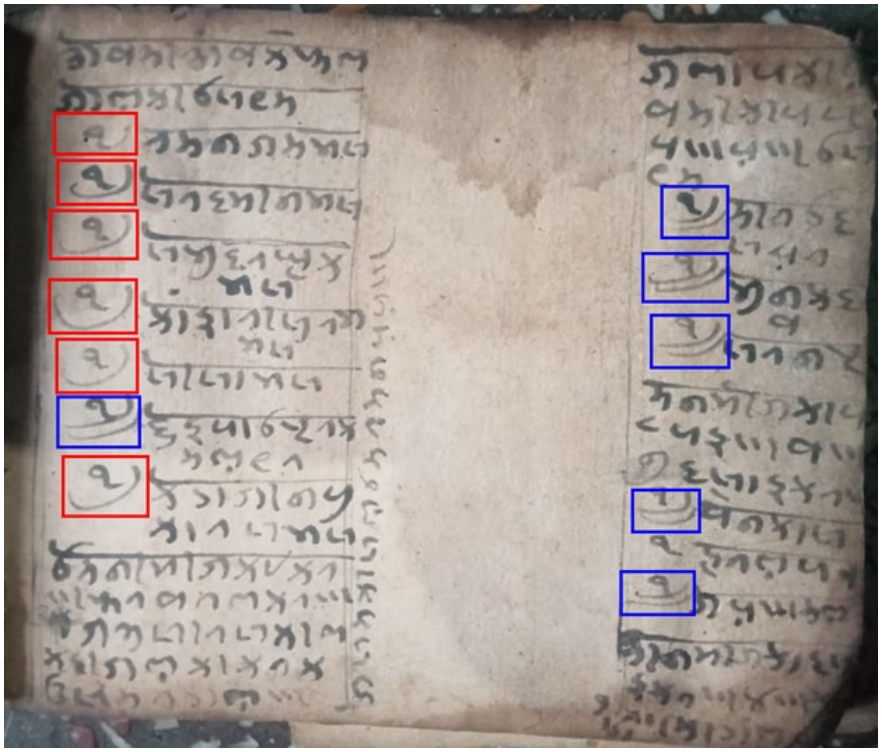


Figure 30: Section signs.

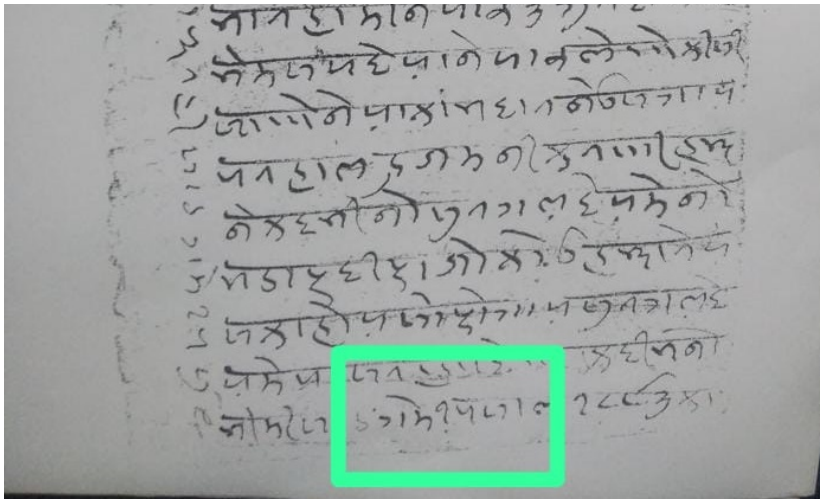


Figure 31: Usage of a Latin question mark in a Sirmauri document.

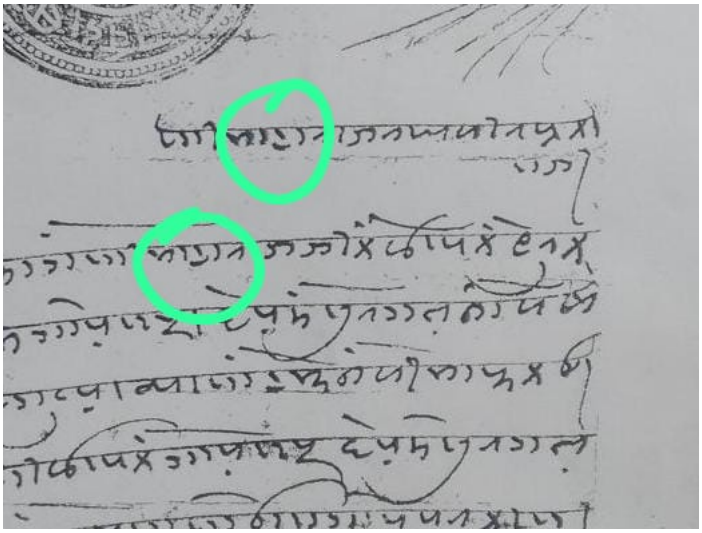


Figure 32: Document showing variant form 𑐴 of 𑐳 *ha*.

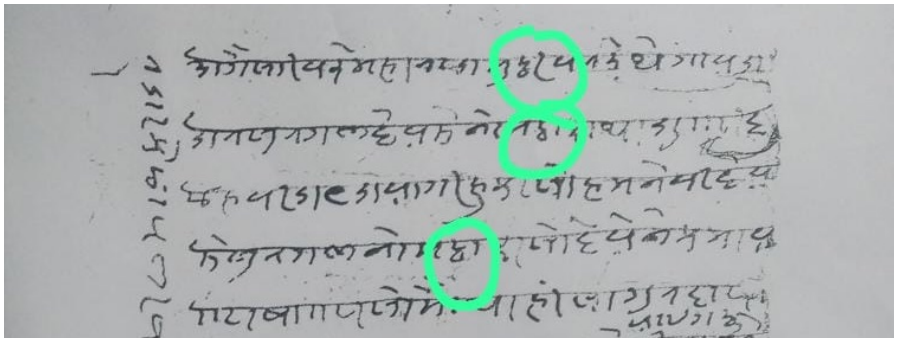


Figure 33: Document showing Jaunsari form 𑐸 of 𑐶 *ḍha*.

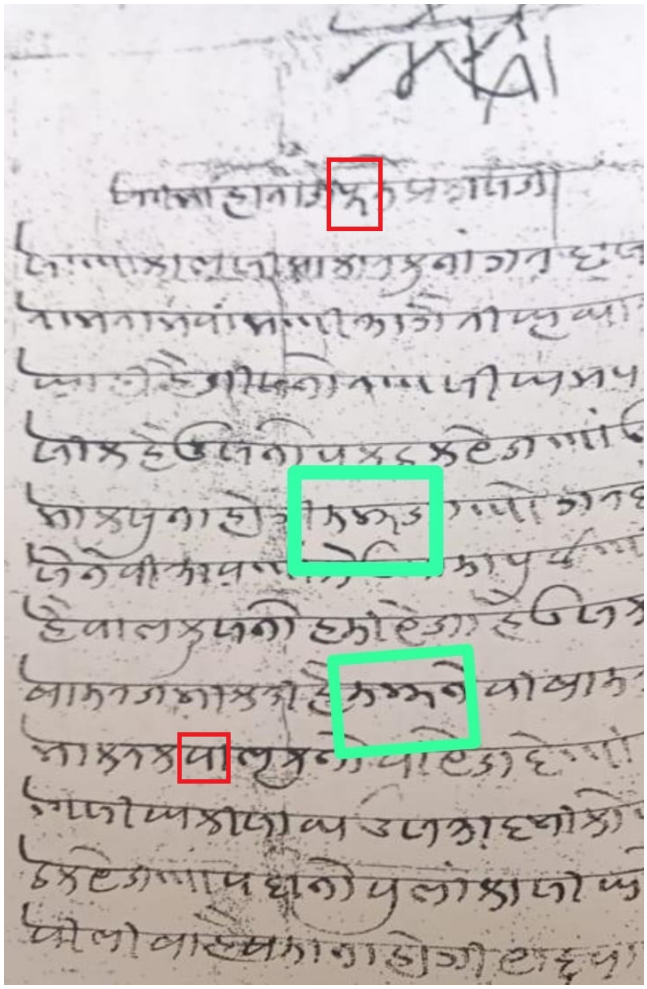


Figure 34: Document showing usage of the conjunct *mma* (green) and Devanagari-like forms for *pha* (red, top) and *pa* (red, bottom).

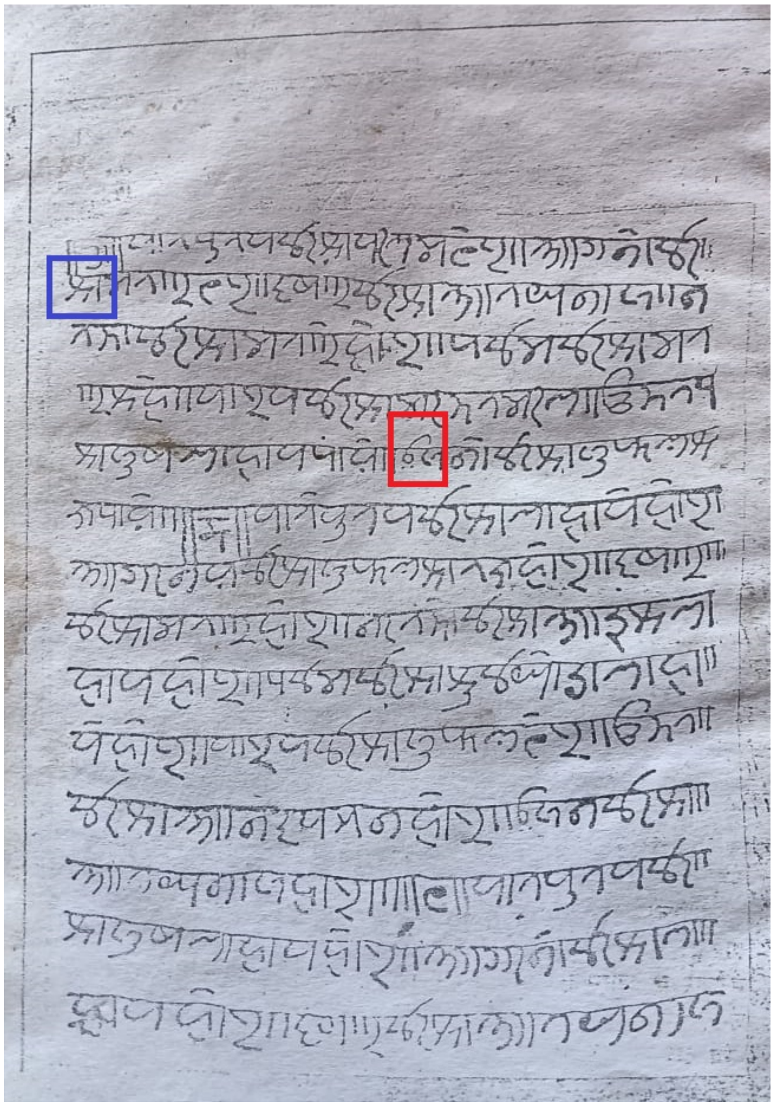


Figure 35: Alternate form of *ca* (blue); usage of variant for for *ī*.

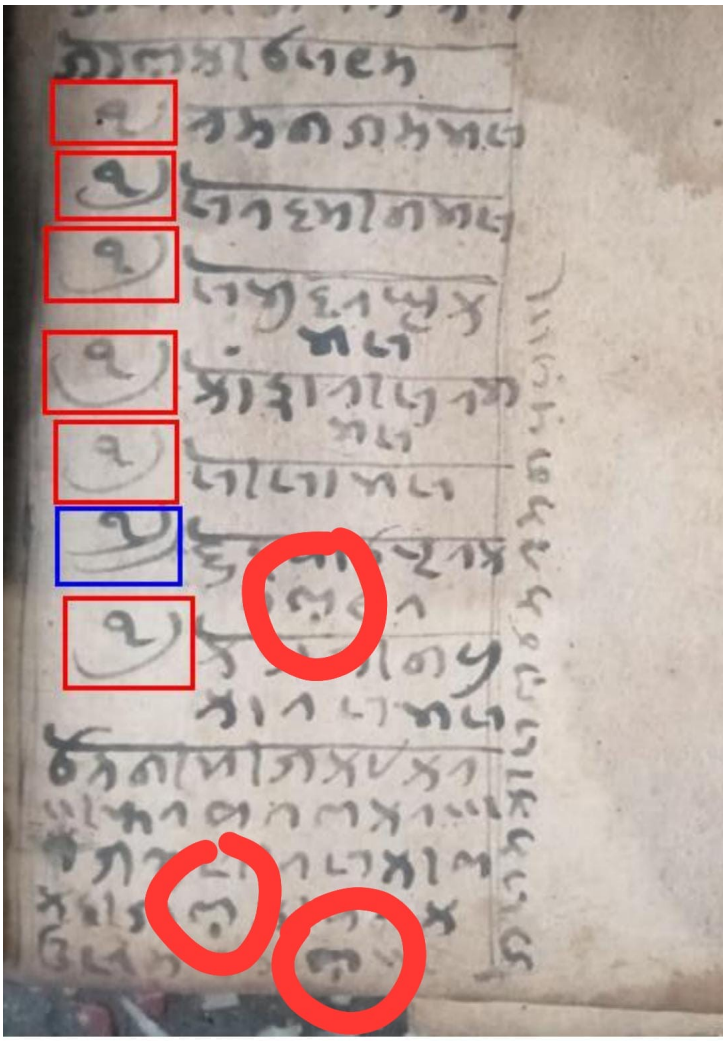


Figure 36: Usage of *ekam* + section marks; also *la* + *nukta* for writing *ḷa*.

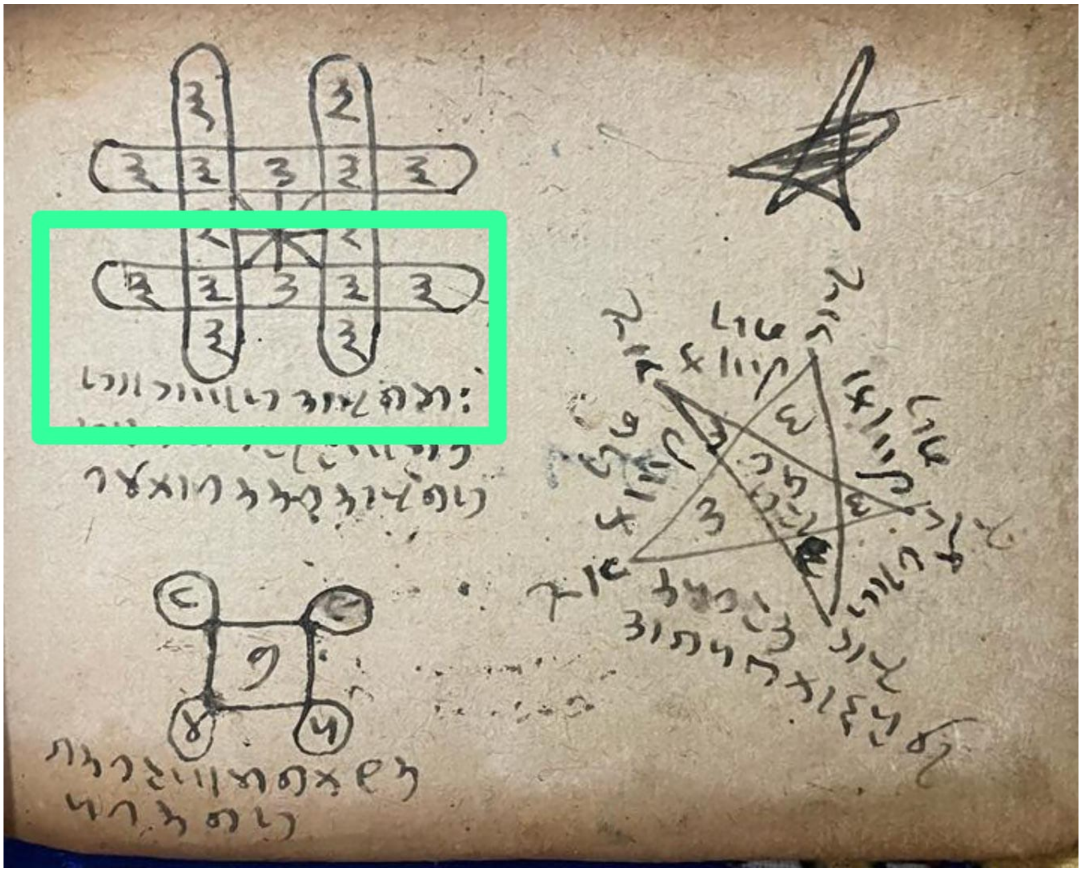


Figure 37: Usage of *visarga* in a Sirmauri document, in the phrase *sarī gaṇesa-āya namaḥ* [= *śrī gaṇeśāya namaḥ*], note the use of the variant form of *ya*.

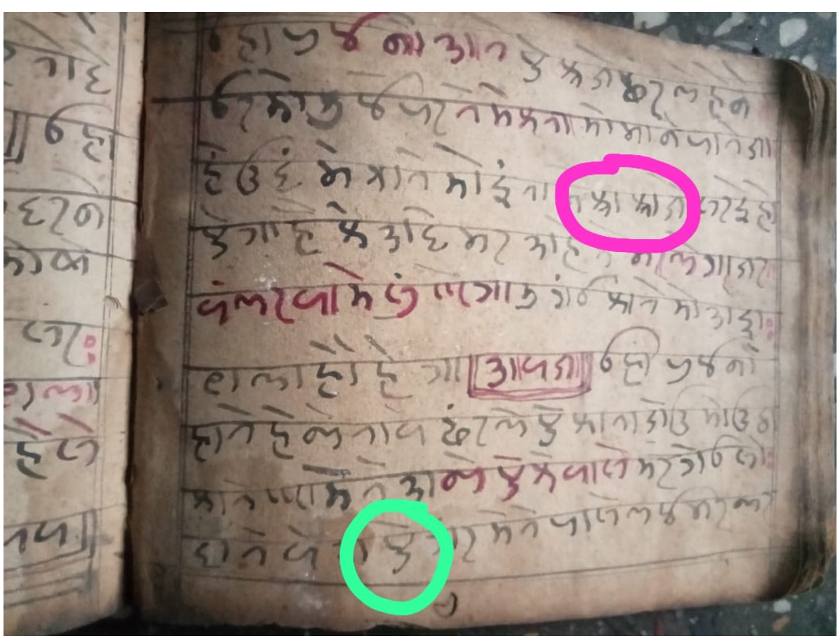


Figure 38: Jaunsari document showing usage of the conjunct *kra* (magenta), with *ra* represented using *ra-kāra*. Also shown is a Jaunsari-specific form of *ya* (green).

सिरमौरी लिपि के ज्ञाता गंगा राम का निधन

भास्कर न्यूज ◆ राजगढ़, 7 मई। सिरमौरी लिपि के एकमात्र ज्ञाता गंगाराम चौहान का शनिवार को उनके पैतृक गांव नौहराधार में निधन हो गया। वे 95 वर्ष के थे।

वे लंबे समय से बीमार चल रहे थे। सिरमौरी लिपि में हस्तलिखत जंतरी (सांचा), पाशा व दुलर्भ पुस्तकें भी उनके पास थी। गंगाराम चौहान के साथ सिरमौरी लिपि भी लुप्त हो गई। ज्ञात रहे कि कभी सिरमौर का राज-काज इसी भाषा में होता था। हिमाचल निर्माता डॉ. यशवंत सिंह परमार के सहयोगी रहे गंगाराम को वर्ष 1951 में सिरमौर जिला बोर्ड का सदस्य बनाया गया था।

इस दौरान उन्होंने क्षेत्र में शिक्षण संस्थाएं खोलने व क्षेत्र में सड़कों का जाल बिछाने में अहम भूमिका निभाई थी।

शोक जताया : राजगढ़ की स्वयंसेवी संस्था एजूकेशन सोसायटी विद्या निकेतन के सचिव वाई.डी. शर्मा, कृति संस्था फागू के सचिव एस. चौहान, मुंशीराम वर्मा के अलावा सोलन के लेखक मदन हिमाचली, डॉ. शंकर वशिष्ठ, डॉ. प्रेम लाल गौतम, रतनचंद निर्झर, अमर सिंह बरवाल और डॉ. परमार यूनिवर्सिटी के पीआरओ पीडी भारद्वाज ने भी गंगाराम चौहान के निधन पर शोक जताया है।

जारी संयुक्त बयान में उन्होंने कहा कि सिरमौर वासियों ने संस्कृति का पुजारी एवं युग पुरुष को खो दिया है,जिसकी भरपाई नहीं हो सकती। उन्होंने शोक संतप्त परिवार के प्रति अपनी गहरी संवेदना व्यक्त की है।

Dainik Bhaskar, Himachal Pradesh edition, February 26, 2005

गंगाराम के साथ लुप्त हो जाएगी सिरमौरी भाषा

■ भास्कर न्यूज

राजगढ़, 26 फरवरी। सिरमौर जिले की सिरमौरी भाषा लुप्त होने की कगार पर है। जिला सिरमौर की सिरमौरी भाषा जिसके एकमात्र ज्ञाता गंगाराम चौहान है। 92 वर्षीय गंगाराम चौहान इस भाषा की लिपि एवं भाषा की विरासत को संभाले हुए है।

गौरतलब है कि देश से कई भाषाएं एवं लिपियां बदलते परिवेश के साथ-साथ समाप्त होती जा रही हैं, उन्होंने कई नौजवानों को खेत्र की सांस्कृतिक धरोहर, सिरमौरी भाषा को सीखने व सहेजने की अपील की, लेकिन भाग दौड़ वाली इस दुनियां में किसी के पास समय नहीं है कि इस भाषा को जीवित रखने में उनकी मदद कर सके। सिरमौर रियासत के राजकाज की कभी यही भाषा एवं लिपि प्रचलित रही है। गंगाराम के घर में आज भी सदियों पुराना सांचा यानी जंत्री को देखा जा सकता है जो कि सिरमौरी लिपि में लिखा गया है। सिरमौरी में लिखे गए कई दस्तावेज जो कि किसी की समझ में नहीं आते उनके संग्रह में मौजूद मिलेंगे। राजगढ़ से 30 किमी दूर नोहरा गांव के गंगाराम चौहान आजादी के बाद जाने-माने नेता रहे व तत्कालीन मुख्यमंत्री डॉ. यशवंत सिंह परमार का भरपूर सहयोग दिया। 1951 में सिरमौर डिस्टक्ट बोर्ड के सदस्य मनोनीत होने के बाद उन्होंने क्षेत्र में पचासों स्कूल खुलवाए व शिक्षा का प्रचार किया। जहां आज भी उनको स्कूल खुलवाने के लिए बड़ी इज्जत से याद किया जाता है। शिक्षा के प्रति जागरूक चौहान ने अपने बच्चों को भी खूब तालीम दी है। लड़की पीएचडी, एक लड़का वकील और एक लड़का अध्यापक है। शिक्षा के साथ-साथ पूरा परिवार आध्यात्मिकता के लिए पूरे क्षेत्र में जाना जाता है।

पिछली यादों को दोहराते हुए वे कहते है कि क्षेत्र में जो आज सड़कों का जाल बिछा है, डॉ. परमार की बदौलत है। अगर वे सोलन से राजगढ़ को सड़क का निर्माण न करवाते तो आज तो लिंक रोड गांव-गांव को बनाए जा रहे है वे कैसे बनते। सैकड़ों लोगों के साथ गिरी पुल में सड़क निर्माण के उद्घाटन पर परमार ने खुद श्रमदान दिया था और उन्हीं के नाम पर वहां का नाम यशवंत नगर रखा गया है।

■ गंगाराम के घर सदियों पुरानी जंत्री मौजूद

■ उनके संग्रह में कई सिरमौरी दस्तावेज

Dainik Bhaskar, Himachal Pradesh edition, May 7, 2005

Figure 39: Newspaper articles regarding the passage of the Sirmauri language and script expert Gangaram Chauhan. Images courtesy of Srinidhi A.

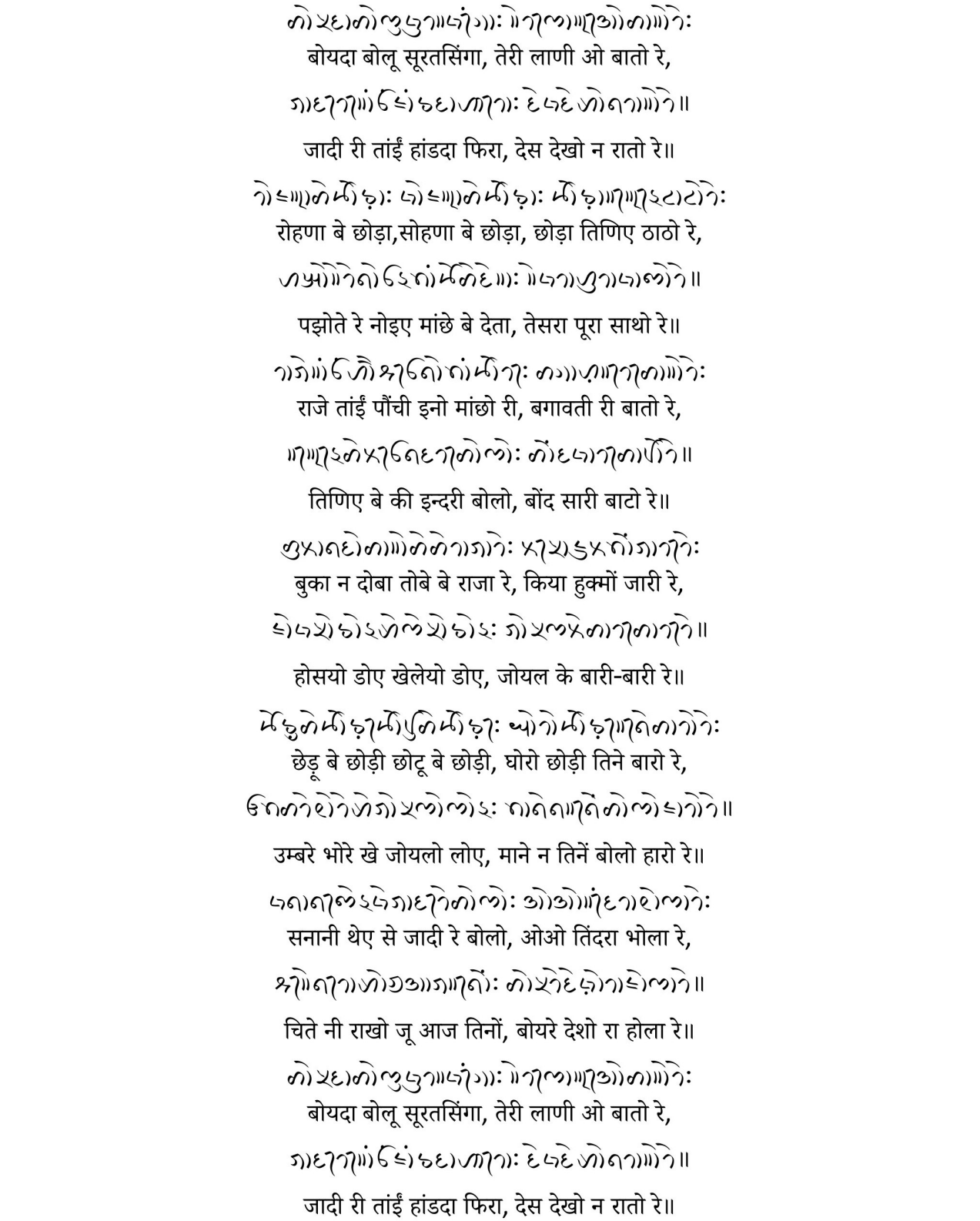


Figure 40: A poem by Sherjung Chauhan typed in both Sirmauri and Devanagari using digital fonts.

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://std.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://std.dkuug.dk/JTC1/SC2/WG2/docs/summaryform.html **.**
See also http://std.dkuug.dk/JTC1/SC2/WG2/docs/roadmaps.html **for latest *Roadmaps*.**

A. Administrative

1. **Title:** ***Proposal to encode Sirmauri in Unicode***
2. Requester's name: *Anshuman Pandey <pandey@umich.edu>*
   Biswajit Mandal <biswajitmandal.bm90@gmail.com>
3. Requester type (Member body/Liaison/Individual contribution): *Expert contribution*
4. Submission date: *2025-10-03*
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: *Sirmauri*
   b. The proposal is for addition of character(s) to an existing block:
      Name of the existing block:
2. Number of characters in proposal: *55*
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
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: N4502-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, 2012-01)

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? *Nikhil Bhardwaj, Sherjung Chauhan, Bhupal Singh Caouhan, The family of Natiram Negi, Gangaram Pabuch, Yatin Pandit, Dr. Om Prakash Sharma, Dr. Dilip Singh Tilkan, The family of Har Singh Tilkan, The family of Ved Singh Tilkan*
   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? *Yes*
   Reference: *See text of proposal*
4. The context of use for the proposed characters (type of use; common or rare) *Rare*
   Reference: *See text of proposal*
5. Are the proposed characters in current use by the user community? *Yes*
   If YES, where? Reference: *Used by Sirmauri language speakers for literary, religious, and cultural purposes, as well as for preservation of historical documents*
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, or could be confused with, an existing character? *No*
    If YES, is a rationale for its inclusion provided?
    If YES, reference:
11. Does the proposal include use of combining characters and/or use of composite sequences? *No*
    If YES, is a rationale for such use provided? *No*
    If YES, reference:
    Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided? *No*
    If YES, reference:
12. Does the proposal contain characters with any special properties such as control function or similar semantics? *No*
    If YES, describe in detail (include attachment if necessary)
13. Does the proposal contain any Ideographic compatibility characters? *No*
    If YES, are the equivalent corresponding unified ideographic characters identified?
    If YES, reference: