

# UC Berkeley

## Proposals from the Script Encoding Initiative

### Title

Proposal to encode the Dogra script in Unicode

### Permalink

<https://escholarship.org/uc/item/84s4p2zq>

### Author

Pandey, Anshuman

### Publication Date

2015-11-04

Peer reviewed

# Proposal to encode the Dogra script in Unicode

Anshuman Pandey  
Department of Linguistics  
University of California, Berkeley  
Berkeley, California, U.S.A.  
anshuman.pandey@berkeley.edu

November 4, 2015

## 1 Introduction

This is a proposal to encode the Dogra script in Unicode. Dogra was briefly described in “Proposal to Encode the Takri Script in ISO/IEC 10646” (L2/09-424) as one of several scripts to be unified in the ‘Takri’ block. The block was intended for representing several related scripts of Himachal Pradesh, Jammu, Kashmir, and Panjab by means of a generic character repertoire until additional research on individual scripts could be conducted. As described in L2/09-424, the representative glyphs for characters in the ‘Takri’ block are based upon the standard, printed form of the Chamba or Chambeali form of Takri. The Chamba form was selected because materials in the script were readily available. Although Dogra has its own standardized form, it was stated in L2/09-424 that there were insufficient materials available for developing a formal encoding. Dogra was, therefore, grouped with ‘Takri’ on the basis of historical typology and as a matter of practicality. My recent research upon various ‘Takri’ scripts has yielded additional sources and information that that support of the encoding of Dogra as an independent script,<sup>1</sup> especially because of its official status in the former State of Jammu and Kashmir and on account of the differences between it and the standard Chamba script (see tables 2 and 3). A step towards encoding Dogra was taken through the submission of a preliminary proposal (L2/15-213). The present document supersedes L2/15-213 and provides additional information and a number of new sources that support the encoding of Dogra in Unicode.

## 2 Background

The Dogra script is historically associated with the Dogri language (ISO 639-3: doi), but it is not actively used at present. In 1916, George A. Grierson wrote in the *Linguistic Survey of India* that “Ḍōgrā has an alphabet of its own, which is allied to the Ṭakrī alphabet current in the Punjab Himalayas” (1916: 638). Although it was used mostly for informal communication and commercial activities, Dogra was standardized in the 1860s. Grierson noted that “[s]ome thirty or forty years ago the then Mahārājā of Jammu and Kashmir caused to be invented a modified form of the current Ṭakrī so as to bring it more into line with Dēvanāgarī and Gurmukhī” and “[t]his improved Ḍōgrī is used for official documents, but it has not generally displaced the old Ṭakrī form of script” (*ibid*). This official form of Dogra is known as ‘Name Dogra Akkhar’ or

---

<sup>1</sup> The Jaunsari or Sirmauri script was also briefly described in L2/09-424 and unified with the Takri block. It is another potential candidate for separate encoding in Unicode.

the ‘New Dogra Script’. It was the official script of the State of Jammu and Kashmir during the reign of Maharaja Ranbir Singh (r. 1857–1885), who is the ruler referred to by Grierson. The New Dogra script was used for administrative purposes and upon currency, judicial and non-judicial stamp papers, postcards, and postage stamps. It was also used for literary activities. The first book published in the standard Dogra was a translation of a Sanskrit mathematical treatise by Bhāskarācārya titled *Līlavatī*. The work was commissioned by Ranbir Singh and printed in 1872 at Vidya Vilas Press, the first printing press in Jammu.

The old and standard forms of Dogra differ primarily in the shapes of letterforms and the representation of consonant clusters. Specific differences are described throughout this proposal. As mentioned by Grierson, standard Dogra was influenced by Devanagari and Gurmukhi. The motivation behind such a move is not clear, but it could be related to the idea of providing the Dogra script with an air of ‘authority’ by giving it a form that resembled scripts which were already endowed with ‘official’ and ‘literary’ attributes. The standardization of Dogra in the 1860s established two distinctive forms of the script. Throughout this document, the common written form is called ‘Old Dogra’ and the standard form is referred to as ‘New Dogra’.

The Dogra script is no longer used actively and the Dogri language is now generally written in Devanagari. Nonetheless, the script continues to attract the interest of philatelists who collect ephemera from Jammu and Kashmir. Although there is very little scholarship on the script, it was described briefly by the prominent Sanskrit scholar Fritz Staal in his *Stamps of Jammu & Kashmir* (New York: The Collectors Club, 1984). Information on Dogra is also curated on a philately website called “Collecting Kashmir”, maintained by Carol von der Lin.

### 3 Proposed Encoding

#### 3.1 Approach

The justification for encoding Dogra as a separate script in Unicode is based upon the following criteria: 1) graphical differences between Dogra, Chamba, and other forms of Takri; 2) usage of the script in an official capacity; 3) cultural and political recognition of it as a distinctive script; 4) differentiation into ‘old’ and ‘standard’ forms. Of these, the fourth point requires a means for handling the internal distinctions of the script. A separate block for Dogra would provide a means to maintain the script’s external identity from other related scripts and to also maintain the identity of styles internally.

#### 3.2 Structure

The Dogra script is a Brahmi-based alphasyllabary that is written from left to right. Independent and initial vowels are written using letters, while dependent vowels are expressed using combining signs. Consonant letters possess the inherent vowel *a*, which is changed by the attachment of a vowel sign. Consonant clusters are written in various ways, such as with visible *halanta*, as an atomic ligature, and with half-forms; *repha* is attested, but rarely used. There are no formal conventions regarding representation of conjuncts.

#### 3.3 Script name

The designation ‘Dogra’ has been assigned to the script block in Unicode. The script is referred to as both ‘Dogra’ and ‘Dogri’ in English secondary sources. The name ‘Dogra’ is preferred over ‘Dogri’ because it links the script with the Dogra dynasty of Jammu and Kashmir, under whose government it was established as the official script. Secondly, ‘Dogri’ suggests a connection with the language of the same name; however, the language is no longer written in the script. The script was officially referred to as *divigarta akṣara* or “Dogra

letters”, where ‘Dvigarta’ is presented as the Sanskritic root of the name ‘Dogra’. Another indigenous name for the script is *name dogra akkhar* or “New Dogra letters”.

### 3.4 Representative glyphs

The representative glyphs for Dogra characters are based primarily upon forms used in New Dogra. Characters from Old Dogra have been used when necessary, sourced primarily from figures 1 and 2. The font has been designed by the proposal author.

### 3.5 Vowel letters

There are 10 vowel letters:

𑂀	DOGRA LETTER A	𑂛	DOGRA LETTER UU
𑂁	DOGRA LETTER AA	𑂜	DOGRA LETTER E
𑂂	DOGRA LETTER I	𑂝	DOGRA LETTER AI
𑂃	DOGRA LETTER II	𑂞	DOGRA LETTER O
𑂄	DOGRA LETTER U	𑂟	DOGRA LETTER AU

In Old Dogra the values of the vowel letters 𑂂, 𑂃, and 𑂄 are variable. As shown in the accompanying sources, the letter 𑂂 was used for both *i* and *ī*; 𑂃 for *u* and *ū*; and 𑂄 for *u*, *ū*, *o*, and *au*. In New Dogra, the values of these three letters became fixed and new letters were introduced in order to provide distinctive representations for all basic vowels. It is apparent that the attempt to stabilize the vowel repertoire resulted in overlap and reassignment of several letters. The Old Dogra 𑂂 (*i*, *ī*) was retained solely for *ī* and 𑂃 (*u*, *ū*) was reassigned for *i*. A modified form of 𑂄 (*u*, *o*) came to represent *u*, while the form 𑂛 was introduced for *ū*, which is certainly an adaptation of the letter *ū* from Devanagari or Sharada. Distinct letters for 𑂞 *o* and 𑂟 *au* were provided by placing the corresponding dependent vowel signs upon the modified *u*.

	Old	New
DOGRA LETTER I	𑂂	𑂃
DOGRA LETTER II	—	𑂃
DOGRA LETTER U	𑂄	𑂛
DOGRA LETTER UU	—	𑂛
DOGRA LETTER O	𑂄	𑂞
DOGRA LETTER AU	𑂄	𑂟

The proposed vowels letters are based upon the repertoire of New Dogra. This approach completely accommodates vowels in the old and new forms of the script. Although the names of New Dogra vowel letters do

not strictly correlate with values of Old Dogra vowels, the proposed letters can be used for representing the old script. The 6 DOGRA LETTER I can be used for *u*; 6 DOGRA LETTER II can be used for *i*; and 3 DOGRA LETTER U can be used for *o* and *au*, the height and curvature of the left stroke may be considered a stylistic attribute.

### 3.6 Vowel Signs

There are 10 dependent vowel signs:

◌ <sup>v</sup>	DOGRA VOWEL SIGN AA	◌ <sup>◌</sup>	DOGRA VOWEL SIGN VOCALIC R
◌ <sup>f</sup>	DOGRA VOWEL SIGN I	◌ <sup>—</sup>	DOGRA VOWEL SIGN E
◌ <sup>ी</sup>	DOGRA VOWEL SIGN II	◌ <sup>==</sup>	DOGRA VOWEL SIGN AI
◌ <sup>◌</sup>	DOGRA VOWEL SIGN U	◌ <sup>^</sup>	DOGRA VOWEL SIGN O
◌ <sup>◌</sup>	DOGRA VOWEL SIGN UU	◌ <sup>^^</sup>	DOGRA VOWEL SIGN AU

The ◌<sup>◌</sup> DOGRA VOWEL SIGN VOCALIC R appears to have been introduced as part of New Dogra. It is used for representing words of Sanskrit origin that contain the vocalic sound *r*. Its shape is derived from either U+11190 SHARADA VOWEL SIGN VOCALIC R or ◌<sup>◌</sup> U+0943 DEVANAGARI VOWEL SIGN VOCALIC R. There is no corresponding independent vowel letter.

There is an alphabetic aspect to the representation of consonant-vowel syllables in written Dogra. The dependent vowel may be represented using the independent vowel letter as well as the vowel sign, eg.  $\mathfrak{a}\mathfrak{u}$  may be used in place of  $\mathfrak{a}\mathfrak{u}$  for the syllable *ke*.

There are differences in the shape of certain vowel letters between the old and new script. The typical Takri forms of the vowel signs ◌<sup>—</sup> -*u* and ◌<sup>◌</sup> -*u*, which also resemble corresponding forms in Gurmukhi, were replaced with forms resembling Devanagari signs ◌<sup>◌</sup> and ◌<sup>◌</sup>, respectively. The typical Takri style of the vowel sign ◌<sup>◌</sup> *o* was applied for writing ◌<sup>^^</sup> *au*, while a simplified of *o* in the form ◌<sup>^</sup> was used for *o*. The duplication of ◌<sup>◌</sup> *o* as ◌<sup>◌</sup> for writing *au* was eliminated. A comparison of vowel letters in the old and new scripts as they align with the proposed characters is as follows:

	Old	New
DOGRA VOWEL SIGN I	◌ <sup>f</sup>	◌ <sup>f</sup>
DOGRA VOWEL SIGN II	◌ <sup>ी</sup>	◌ <sup>ी</sup>
DOGRA VOWEL SIGN U	◌ <sup>◌</sup>	◌ <sup>◌</sup>
DOGRA VOWEL SIGN UU	◌ <sup>◌</sup>	◌ <sup>◌</sup>
DOGRA VOWEL SIGN O	◌ <sup>◌</sup>	◌ <sup>^</sup>
DOGRA VOWEL SIGN AU	◌ <sup>^^</sup>	◌ <sup>^^</sup>

### 3.7 Consonants

There are 34 consonant letters:

ਕ	DOGRA LETTER KA	ੳ	DOGRA LETTER DA
ਖ	DOGRA LETTER KHA	ਢ	DOGRA LETTER DHA
ਗ	DOGRA LETTER GA	ਠ	DOGRA LETTER NA
ਘ	DOGRA LETTER GHA	ਠ	DOGRA LETTER PA
ਙ	DOGRA LETTER NGA	ਠ	DOGRA LETTER PHA
ਚ	DOGRA LETTER CA	ੳ	DOGRA LETTER BA
ਛ	DOGRA LETTER CHA	ੳ	DOGRA LETTER BHA
ਜ	DOGRA LETTER JA	ਯ	DOGRA LETTER MA
ਯ	DOGRA LETTER JHA	ਯ	DOGRA LETTER YA
ਨ	DOGRA LETTER NYA	ਰ	DOGRA LETTER RA
ਟ	DOGRA LETTER TTA	ਲ	DOGRA LETTER LA
ਠ	DOGRA LETTER TTHA	ੳ	DOGRA LETTER VA
ਡ	DOGRA LETTER DDA	ਸ਼	DOGRA LETTER SHA
ਢ	DOGRA LETTER DDHA	ਖ	DOGRA LETTER SSA
ਨ	DOGRA LETTER NNA	ਠ	DOGRA LETTER SA
ਤ	DOGRA LETTER TA	ੳ	DOGRA LETTER HA
ਠ	DOGRA LETTER THA	ੳ	DOGRA LETTER RRA

The ਖ DOGRA LETTER SSA was introduced into New Dogra for representing the retroflex sibilant *ṣa*, which occurs in words of Sanskrit origin. It resembles ਖ U+111B0 SHARADA LETTER SSA.

There are glyphic variants of several consonant letters:

	Regular	Variant
DOGRA LETTER GA	ਗ	ਗ
DOGRA LETTER GHA	ਘ	ਘ
DOGRA LETTER CHA	ਛ	ਛ
DOGRA LETTER JHA	ਯ	ਯ
DOGRA LETTER THA	ਠ	ਠ

DOGRA LETTER DHA	𑂏	𑂐
DOGRA LETTER RA	𑂑	𑂒

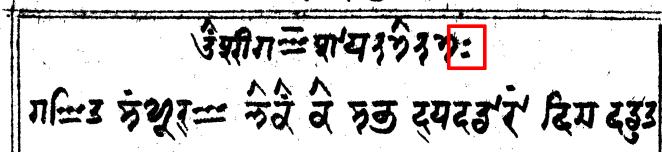
There is evidence of usage of different forms of the same letter in the available sources. For instance, both 𑂏 and 𑂑 are both used for CHA.

### 3.8 Anusvara

The ◌̣ DOGRA SIGN ANUSVARA is used for marking nasalization.

### 3.9 Visarga

The ◌̣: DOGRA SIGN VISARGA is used for indicating post-vocalic aspiration in words of Sanskrit origin. It is shown below in the word 𑂑𑂏: *namah* “homage to”:



### 3.10 Nukta

The ◌̣ DOGRA SIGN NUKTA is used for representing sounds that are not native to Dogri and related languages.

### 3.11 Virama

The ◌̣ DOGRA SIGN VIRAMA has two functions, similar to the corresponding character of Devanagari. It is used as a *halanta* for marking the absence of the inherent vowel of a consonant letter. It is also a control character that is used for producing conjuncts.

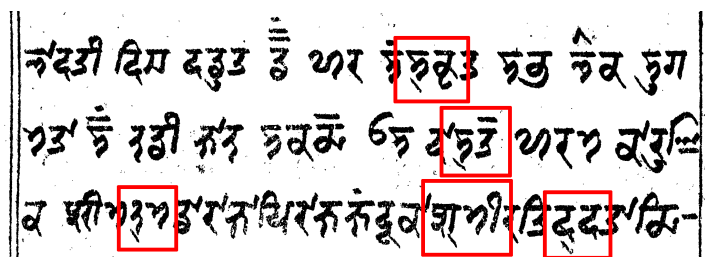
### 3.12 Consonant Conjuncts

Consonant clusters are rendered in three ways: 1) with visible *halanta* beneath the full form of each bare consonant; 2) with a half-form of the initial consonant followed by the full form of the following letter; and 3) as ligatures. There are no formal conventions for the written expression of conjuncts. The ◌̣ DOGRA SIGN VIRAMA is used for the encoded representation of conjuncts.

#### 3.12.1 Types of conjuncts

Some representations of Dogra conjuncts as found in the available sources are illustrated below:

- *Visible halanta* In both Old and New Dogra, the most common method of representing consonant clusters is by placing a *halanta* beneath each bare consonant:



- *Half-forms* Another method is to use half-forms of letters if the graphical structure provides such an opportunity:

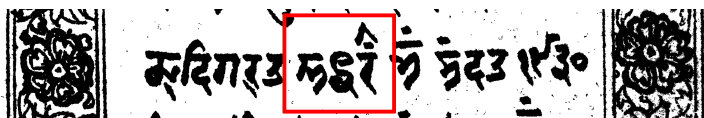


The above shows the conjunct श्र *śra* in the word श्री *śrī* (a Sanskrit honorific). The half form श्र is produced by dropping the right vertical bar of श SHA. The conjunct श्र is used consistently in New Dogra for representing the cluster *śra*; the form श्र with visible *halanta* beneath a full form of श SHA does not occur in the available sources.

- *Ligatures* There are two types of ligature. In the first, the shapes of individual letters are visible. In the below the boxed text contains the conjunct श्र *śra* from the word श्रंण *śāṅṅ*, a transcription of the English “stamp”.

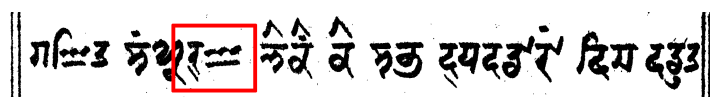


The second type is an atomic (*akhaṇḍa*) ligature whose form may be considered a distinctive letter. Such a ligature occurs in the excerpt below:



The boxed text above contains the atomic ligature क्ष *kṣa*, which occurs in the word अक्षर *akṣarom* “in letters” (oblique form of अक्षर *akṣara* “letter”). The conjunct क्ष *kṣa* is used consistently in New Dogra for representing the cluster /ksa/. Its shape may be influenced by Devanangari क्ष *kṣa*.

- *Forms of RA* Special conjunct forms of र RA do not generally occur, eg. *repha* or *vattu*. Cluster-initial RA is represented as a full form with *halanta*, as in र्ना *rṇa* in the word *sampūrṇa* below:



and non-initial RA is treated similarly, as in प्रा *pra* in the word *prabhu* below:



However, there is one instance in the available sources in which cluster-initial RA is represented as a *repha*. The occurrence is curious because the word *caturbhujā* “square”, in which *repha* occurs, appears three times in proximity, but is spelled differently in all three instances.

In the boxed text at the top, the RA in *caturbhujā* is written with *halanta*:  $\text{᳚᳚᳚᳚᳚}$  *catur.bhuja*. In the text at bottom left, RA occurs with its inherent vowel intact, without *halanta*:  $\text{᳚᳚᳚᳚᳚}$  *caturabhujā*. At bottom right, *ra* is represented as *repha*:  $\text{᳚᳚᳚᳚᳚}$  *caturbhujā*.

The usage of *repha* in Dogra is to be considered anomalous. There is no present need to support this form of RA in encoded representation.

### 3.12.2 Encoded representation of conjuncts

The Devanagari model for the representation of conjuncts is appropriate for Dogra. In encoded text, conjuncts are produced by placing the control character ̣ DOGRA SIGN VIRAMA after each non-initial consonant in a cluster, ie.  $\langle (C, \text{̣ VIRAMA})^*, C \rangle$ .

Representations of conjuncts rendered as atomic ligatures are:

*kṣa*  $\text{᳚᳚}$   $\langle \text{᳚ KA, ̣ VIRAMA, ᳚ SSA} \rangle$

*ṣṭa*  $\text{᳚᳚}$   $\langle \text{᳚ SA, ̣ VIRAMA, ᳚ TTA} \rangle$

Examples of conjuncts rendered using half-forms are:

*śra*  $\text{᳚᳚}$   $\langle \text{᳚ SHA, ̣ VIRAMA, ᳚ RA} \rangle$

Conjuncts displayed using visible *halanta* are:

*pra*  $\text{᳚᳚᳚}$   $\langle \text{᳚ PA, ̣ VIRAMA, ᳚ RA} \rangle$

*rṇa*  $\text{᳚᳚᳚}$   $\langle \text{᳚ RA, ̣ VIRAMA, ᳚ NNA} \rangle$

### 3.12.3 Modifying conjunct formation

As is the case in Devanagari and other Indic scripts, a ligature may be broken using the generic control character ZWJ U+200C ZERO WIDTH NON-JOINER

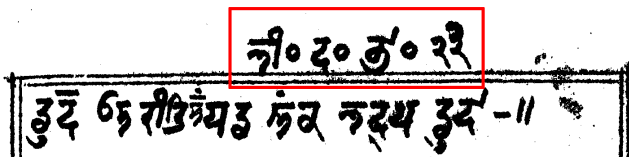
*kṣa* & <क KA, ् VIRAMA, ः SSA>

*kṣa* र् ः <क KA, ् VIRAMA, ZWJ ZERO WIDTH NON-JOINER, ः SSA>

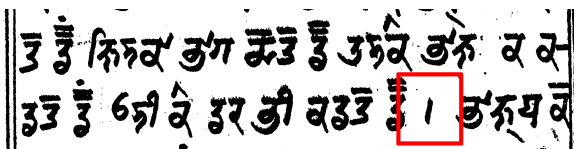
The rendering of a conjunct as a ligature or with visible VIRAMA is a matter of font design.

### 3.13 Punctuation

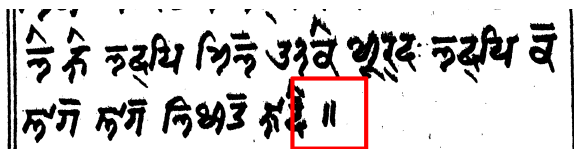
**Abbreviation sign** The ◦ DOGRA ABBREVIATION SIGN is used for indicating truncation of text. In the excerpt below, the highlighted text ली० ट० भ० (lī. va. bhā. 23) is an abbreviation for *līla vatī bhāga 23*.



**Danda-s** The *daṇḍā* and double *daṇḍā* are commonly used in Dogra documents for marking the ends of sentences and paragraphs. The below excerpt shows usage of the *daṇḍā*:

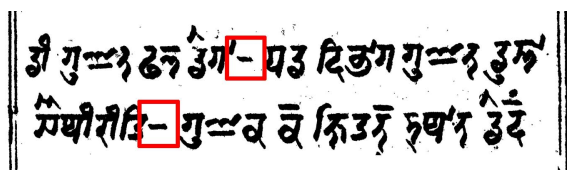


and the following shows use of the double *daṇḍā*:



These *daṇḍā*-s are not proposed for encoding as separate characters in the ‘Dogra’ block. They resemble | U+0964 DEVANAGARI DANDA and || U+0965 DEVANAGARI DOUBLE DANDA, and may be represented using these characters. The Devanagari *daṇḍā*-s are to be specified as extensions for Dogra.

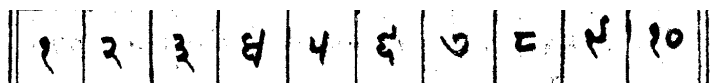
**Other signs** Other marks of punctuation, such as dashes, are also used:



A hyphen is used in some cases for indicating breaks in a word at end of line (see section 3.16).

### 3.14 Digits

Digits are commonly used in Dogra documents (see figure 11). The style of digits varies across written and printed sources. Some digits in Old Dogra sources resemble those of Takri (see figure 4), while those in new Dogra sources are similar to Devanagar. The excerpt below shows digits 1–9 and the number 10 as represented in the *Līlavatī* written in New Dogra:



The majority of the digits shown here resemble Devanagari forms; however, the exception is ੪ ‘four’. The form of Dogra ‘four’ does not resemble ੪ U+116C4 TAKRI DIGIT FOUR, which is nearly identical to the Devanagari digit four. Information about its origins are not known at this time, but it may be related to ॠ U+111D4 SHARADA DIGIT FOUR. The forms of ‘eight’ and ‘nine’ are well-known variant forms of the regular Devanagari digits.

Additional research into Dogra digits is required before they may be considered for encoding. The available sources suggest that there may have been a cultural understanding that Dogra possessed unique digits. For instance, on the bank notes shown in figure 20 the digits written in Devanagari are rather similar, but are nonetheless displayed twice in order to align the display of the currency value with each script used for the inscriptions.

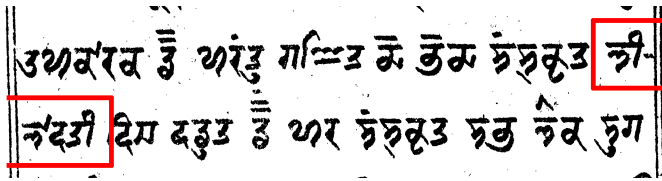
At present, script-specific digits are not proposed for inclusion in the Dogra repertoire. Devanagari digits should be used normatively and should be extended for usage with Dogra. If future research yields distinctive digits for Dogra, there is sufficient space in the block to accommodate their encoding.

### 3.15 Number forms

Fraction signs and currency marks are commonly used in Dogra sources (see figure 12). These may be represented using the fraction and currency signs already encoded in the ‘Common Indic Number Forms’ block (U+A830). These characters are specified as extensions for Dogra. The ‘Common Indic Number Forms’ are described briefly in *The Unicode Standard* (chapter 22.3 ‘Numerals’, pp. 770–771) and in more detail in the original encoding proposal (L2/07-354).

### 3.16 Linebreaking

Linebreaking occurs after an orthographic syllable. Breaks are generally not indicated in informal sources, but are marked in literary materials using hyphens.



### 3.17 Collation

The primary collating order for Dogra is as follows:

क A < क AA < ६ I < ६ II < उ U < उ UU < ए E < ए AI < ओ O < ओ AU <  
 क KA < क KHA < ग GA < ग GHA < ँ NGA < क CA < क CHA < न JA <  
 ग JHA < न NYA < ट TTA < ट TTHA < ड DDA < ड DDHA < ञ NNA < उ TA <  
 क THA < द DA < द DHA < र NA < प PA < फ PHA < ब BA < उ BHA <  
 म MA < य YA < र RA < ल LA < व VA < श SHA < स SSA < ङ SA < उ HA <  
 ड RRA < ◌ VOWEL SIGN AA < ◌ VOWEL SIGN I < ◌ VOWEL SIGN II <  
 ◌ VOWEL SIGN U < ◌ VOWEL SIGN UU < ◌ VOWEL SIGN VOCALIC R < ◌ VOWEL SIGN E <  
 ◌ VOWEL SIGN AI < ◌ VOWEL SIGN O < ◌ VOWEL SIGN AU

The following characters have secondary weights: ◌ SIGN ANUSVARA, ◌ SIGN VISARGA, ◌ SIGN VIRAMA, ◌ SIGN NUKTA.

## 4 Character data

### 4.1 Character properties

The properties for Dogra in the Unicode Character Database format are:

```

11800;DOGRA LETTER A;Lo;0;L;;;;;N;;;;;
11801;DOGRA LETTER AA;Lo;0;L;;;;;N;;;;;
11802;DOGRA LETTER I;Lo;0;L;;;;;N;;;;;
11804;DOGRA LETTER U;Lo;0;L;;;;;N;;;;;
11806;DOGRA LETTER E;Lo;0;L;;;;;N;;;;;
11807;DOGRA LETTER AI;Lo;0;L;;;;;N;;;;;
11808;DOGRA LETTER O;Lo;0;L;;;;;N;;;;;
1180A;DOGRA LETTER KA;Lo;0;L;;;;;N;;;;;
1180B;DOGRA LETTER KHA;Lo;0;L;;;;;N;;;;;
1180C;DOGRA LETTER GA;Lo;0;L;;;;;N;;;;;
1180D;DOGRA LETTER GHA;Lo;0;L;;;;;N;;;;;
1180E;DOGRA LETTER NGA;Lo;0;L;;;;;N;;;;;
1180F;DOGRA LETTER CA;Lo;0;L;;;;;N;;;;;
11810;DOGRA LETTER CHA;Lo;0;L;;;;;N;;;;;
11811;DOGRA LETTER JA;Lo;0;L;;;;;N;;;;;
11812;DOGRA LETTER JHA;Lo;0;L;;;;;N;;;;;
11813;DOGRA LETTER NYA;Lo;0;L;;;;;N;;;;;

```

```

11814;DOGRA LETTER TTA;Lo;0;L;;;;N;;;;;
11815;DOGRA LETTER TTHA;Lo;0;L;;;;N;;;;;
11816;DOGRA LETTER DDA;Lo;0;L;;;;N;;;;;
11817;DOGRA LETTER DDHA;Lo;0;L;;;;N;;;;;
11818;DOGRA LETTER NNA;Lo;0;L;;;;N;;;;;
11819;DOGRA LETTER TA;Lo;0;L;;;;N;;;;;
1181A;DOGRA LETTER THA;Lo;0;L;;;;N;;;;;
1181B;DOGRA LETTER DA;Lo;0;L;;;;N;;;;;
1181C;DOGRA LETTER DHA;Lo;0;L;;;;N;;;;;
1181D;DOGRA LETTER NA;Lo;0;L;;;;N;;;;;
1181E;DOGRA LETTER PA;Lo;0;L;;;;N;;;;;
1181F;DOGRA LETTER PHA;Lo;0;L;;;;N;;;;;
11820;DOGRA LETTER BA;Lo;0;L;;;;N;;;;;
11821;DOGRA LETTER BHA;Lo;0;L;;;;N;;;;;
11822;DOGRA LETTER MA;Lo;0;L;;;;N;;;;;
11823;DOGRA LETTER YA;Lo;0;L;;;;N;;;;;
11824;DOGRA LETTER RA;Lo;0;L;;;;N;;;;;
11825;DOGRA LETTER LA;Lo;0;L;;;;N;;;;;
11826;DOGRA LETTER VA;Lo;0;L;;;;N;;;;;
11827;DOGRA LETTER SHA;Lo;0;L;;;;N;;;;;
11828;DOGRA LETTER SSA;Lo;0;L;;;;N;;;;;
11829;DOGRA LETTER SA;Lo;0;L;;;;N;;;;;
1182A;DOGRA LETTER HA;Lo;0;L;;;;N;;;;;
1182B;DOGRA LETTER RRA;Lo;0;L;;;;N;;;;;
1182C;DOGRA VOWEL SIGN AA;Mc;0;L;;;;N;;;;;
1182D;DOGRA VOWEL SIGN I;Mc;0;L;;;;N;;;;;
1182E;DOGRA VOWEL SIGN II;Mc;0;L;;;;N;;;;;
1182F;DOGRA VOWEL SIGN U;Mn;0;NSM;;;;N;;;;;
11830;DOGRA VOWEL SIGN UU;Mn;0;NSM;;;;N;;;;;
11831;DOGRA VOWEL SIGN VOCALIC R;Mn;0;NSM;;;;N;;;;;
11832;DOGRA VOWEL SIGN E;Mn;0;NSM;;;;N;;;;;
11833;DOGRA VOWEL SIGN AI;Mn;0;NSM;;;;N;;;;;
11834;DOGRA VOWEL SIGN O;Mn;0;NSM;;;;N;;;;;
11835;DOGRA VOWEL SIGN AU;Mn;0;NSM;;;;N;;;;;
11836;DOGRA SIGN ANUSVARA;Mn;0;NSM;;;;N;;;;;
11837;DOGRA SIGN VISARGA;Mc;0;L;;;;N;;;;;
11838;DOGRA SIGN VIRAMA;Mn;9;NSM;;;;N;;;;;
11839;DOGRA SIGN NUKTA;Mn;7;NSM;;;;N;;;;;
1183A;DOGRA ABBREVIATION SIGN;Po;0;L;;;;N;;;;;

```

## 4.2 Linebreaking

Linebreaking properties given in the data format of LineBreak.txt:

```

11800..1182B; AL # Lo [44] DOGRA LETTER A .. DOGRA LETTER RRA
1182C..1182E; CM # Mc [3] DOGRA VOWEL SIGN AA .. DOGRA VOWEL SIGN II
1182F..11835; CM # Mn [7] DOGRA VOWEL SIGN U .. DOGRA VOWEL SIGN AU
11836; CM # Mn DOGRA SIGN ANUSVARA
11837; CM # Mc DOGRA SIGN VISARGA
11838..11839; CM # Mn [2] DOGRA SIGN VIRAMA .. DOGRA SIGN NUKTA
1183A; AL # Po DOGRA ABBREVIATION SIGN

```

## 4.3 Syllabic categories

Syllabic categories given in the format of IndicSyllabicCategory.txt:

```

# Indic_Syllabic_Category=Bindu
11836 ; Bindu # Mn DOGRA SIGN ANUSVARA

```

```

# Indic_Syllabic_Category=Visarga
11837          ; Visarga          # Mc          DOGRA SIGN VISARGA

# Indic_Syllabic_Category=Nukta
11839          ; Nukta            # Mn          DOGRA SIGN NUKTA

# Indic_Syllabic_Category=Virama
11838          ; Virama           # Mn          DOGRA SIGN VIRAMA

# Indic_Syllabic_Category=Vowel_Independent
11800..11808   ; Vowel_Independent # Lo   [3] DOGRA LETTER A..DOGRA LETTER AU

# Indic_Syllabic_Category=Vowel_Dependent
1182C..1182E   ; Vowel_Dependent  # Mc   [3] DOGRA VOWEL SIGN AA..DOGRA VOWEL SIGN I
1182F..11835   ; Vowel_Dependent  # Mn   [7] DOGRA VOWEL SIGN U..DOGRA VOWEL SIGN AU

# Indic_Syllabic_Category=Consonant
1180A..1182B   ; Consonant        # Lo   [34] DOGRA LETTER KA..DOGRA LETTER RRA

```

#### 4.4 Positional categories

Positional data for characters in the format of IndicPositionalCategory.txt:

```

# Indic_Positional_Category=Right
1182C          ; Right # Mc          DOGRA VOWEL SIGN A
1182E          ; Right # Mc          DOGRA VOWEL SIGN II
11837          ; Right # Mc          DOGRA SIGN VISARGA

# Indic_Positional_Category=Left
1182D          ; Left # Mc           DOGRA VOWEL SIGN I

# Indic_Positional_Category=Top
11836          ; Top # Mn            DOGRA SIGN ANUSVARA
11832..11835   ; Top # Mn            [4] DOGRA VOWEL SIGN E..DOGRA VOWEL SIGN AU

# Indic_Positional_Category=Bottom
1182F..11831   ; Bottom # Mn   [3] DOGRA VOWEL SIGN U..DOGRA VOWEL SIGN VOCALIC R
11838          ; Bottom # Mn          DOGRA SIGN VIRAMA
11839          ; Bottom # Mn          DOGRA SIGN NUKTA

```

#### 4.5 Script extensions

The following characters should be extended for usage with the Dogra script in ScriptExtensions.txt:

```

0964          ; # Po          DEVANAGARI DANDA
0965          ; # Po          DEVANAGARI DOUBLE DANDA

0966..096F    ; # Nd   [10] DEVANAGARI DIGIT ZERO..DEVANAGARI DIGIT NINE

A830..A835    ; # No   [6] NORTH INDIC FRACTION ONE QUARTER..
              ;       NORTH INDIC FRACTION THREE SIXTEENTHS
A836..A837    ; # So   [2] NORTH INDIC QUARTER MARK..NORTH INDIC PLACEHOLDER MARK
A838          ; # Sc          NORTH INDIC RUPEE MARK
A839          ; # So          NORTH INDIC QUANTITY MARK

```

## 4.6 ‘Confusable’ characters

Some Dogra characters that resemble characters encoded in other script blocks are:

DOGRA LETTER U	;	11187 SHARADA LETTER U
DOGRA LETTER UU	;	11188 SHARADA LETTER UU
DOGRA LETTER O	;	11684 TAKRI LETTER U
DOGRA LETTER KA	;	0A15 GURMUKHI LETTER KA
DOGRA LETTER CA	;	11196 SHARADA LETTER CA
DOGRA LETTER TTHA	;	1119C SHARADA LETTER TTHA
DOGRA LETTER TA	;	11699 TAKRI LETTER TA
DOGRA LETTER DHA	;	11680 TAKRI LETTER A
DOGRA LETTER PHA	;	1169F TAKRI LETTER PHA
DOGRA LETTER BHA	;	116A1 TAKRI LETTER BHA
DOGRA LETTER YA	;	116A3 TAKRI LETTER YA
DOGRA LETTER SHA	;	0936 DEVANAGARI LETTER SHA
DOGRA LETTER SSA	;	111B0 SHARADA LETTER SSA
DOGRA VOWEL SIGN U	;	0941 DEVANAGARI VOWEL SIGN U
DOGRA VOWEL SIGN UU	;	0942 DEVANAGARI VOWEL SIGN UU
DOGRA VOWEL SIGN E	;	116B2 TAKRI VOWEL SIGN E
DOGRA VOWEL SIGN AI	;	116B3 TAKRI VOWEL SIGN AI
DOGRA VOWEL SIGN O	;	116B4 TAKRI VOWEL SIGN O
DOGRA VOWEL SIGN AU	;	116B5 TAKRI VOWEL SIGN AU
DOGRA SIGN ANUSVARA	;	0902 DEVANAGARI SIGN ANUSVARA
DOGRA SIGN VIRAMA	;	094D DEVANAGARI SIGN VIRAMA
DOGRA SIGN NUKTA	;	093C DEVANAGARI SIGN NUKTA
DOGRA ABBREVIATION SIGN	;	0970 DEVANAGARI ABBREVIATION SIGN

## 5 References

Bhāskarācārya. 1872. *Līlavatī*. Translation into Dogri. Jammu: Vidya Vilas Press.

Grierson, George A. 1916. *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.

Leitner, Gottlieb William. 1882. “A Collection of Specimens of Commercial and Other Alphabets and Handwritings as also of Multiplication Tables Current in Various Parts of the Panjab, Sind and the North West Provinces”. Published as Appendix V of *History of Indigenous Education in the Punjab*. Lahore: Anjuman-i-Punjab Press.

Pandey, Anshuman. 2007. “Proposal to Encode North Indic Number Forms in ISO/IEC 10646” (L2/07-354). <http://www.unicode.org/L2/L2007/07354-north-indic.pdf>

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

———. 2015. “Preliminary Proposal to Encode the Dogra Script in Unicode” (L2/15-213). <http://www.unicode.org/L2/L2015/15213-dogra.pdf>

Staal, Fritz. 1984. *Stamps of Jammu & Kashmir*. New York: The Collectors Club.

von der Lin, Carol A. “Collection Kashmir”, “Philatelic Glossary: Dōgrī-Takrī Scripts”. <http://www.kashmirstamps.com/DogriGloss.html>

## **6 Acknowledgments**

I would like to thank Christopher Shackle (School of Oriental and African Studies, London) for reviewing this proposal.

This project was made possible in part through a Google Research Award, granted to Deborah Anderson for the Script Encoding Initiative, and a grant from the United States National Endowment for the Humanities (PR-50205-15), which funds 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 Google or the National Endowment for the Humanities.



	1180	1181	1182	1183	1184
0	𑂀 11800	𑂁 11810	𑂂 11820	𑂃 11830	
1	𑂄 11801	𑂅 11811	𑂆 11821	𑂇 11831	
2	𑂈 11802	𑂉 11812	𑂊 11822	𑂋 11832	
3	𑂌 11803	𑂍 11813	𑂎 11823	𑂏 11833	
4	𑂐 11804	𑂑 11814	𑂒 11824	𑂓 11834	
5	𑂔 11805	𑂕 11815	𑂖 11825	𑂗 11835	
6	𑂘 11806	𑂙 11816	𑂚 11826	𑂛 11836	
7	𑂜 11807	𑂝 11817	𑂞 11827	𑂟 11837	
8	𑂠 11808	𑂡 11818	𑂢 11828	𑂣 11838	
9	𑂤 11809	𑂥 11819	𑂦 11829	𑂧 11839	
A	𑂨 1180A	𑂩 1181A	𑂪 1182A	𑂫 1183A	
B	𑂬 1180B	𑂭 1181B	𑂮 1182B		
C	𑂯 1180C	𑂰 1181C	𑂱 1182C		
D	𑂲 1180D	𑂳 1181D	𑂴 1182D		
E	𑂵 1180E	𑂶 1181E	𑂷 1182E		
F	𑂸 1180F	𑂹 1181F	𑂺 1182F		

**Independent vowels**

11800	𑂀	DOGRA LETTER A
11801	𑂁	DOGRA LETTER AA
11802	𑂂	DOGRA LETTER I
11803	𑂃	DOGRA LETTER II
11804	𑂄	DOGRA LETTER U
11805	𑂅	DOGRA LETTER UU
11806	𑂆	DOGRA LETTER E
11807	𑂇	DOGRA LETTER AI
11808	𑂈	DOGRA LETTER O
11809	𑂉	DOGRA LETTER AU

**Consonants**

1180A	𑂊	DOGRA LETTER KA
1180B	𑂋	DOGRA LETTER KHA
1180C	𑂌	DOGRA LETTER GA
1180D	𑂍	DOGRA LETTER GHA
1180E	𑂎	DOGRA LETTER NGA
1180F	𑂏	DOGRA LETTER CA
11810	𑂐	DOGRA LETTER CHA
11811	𑂑	DOGRA LETTER JA
11812	𑂒	DOGRA LETTER JHA
11813	𑂓	DOGRA LETTER NYA
11814	𑂔	DOGRA LETTER TTA
11815	𑂕	DOGRA LETTER TTHA
11816	𑂖	DOGRA LETTER DDA
11817	𑂗	DOGRA LETTER DDHA
11818	𑂘	DOGRA LETTER NNA
11819	𑂙	DOGRA LETTER TA
1181A	𑂚	DOGRA LETTER THA
1181B	𑂛	DOGRA LETTER DA
1181C	𑂜	DOGRA LETTER DHA
1181D	𑂝	DOGRA LETTER NA
1181E	𑂞	DOGRA LETTER PA
1181F	𑂟	DOGRA LETTER PHA
11820	𑂠	DOGRA LETTER BA
11821	𑂡	DOGRA LETTER BHA
11822	𑂢	DOGRA LETTER MA
11823	𑂣	DOGRA LETTER YA
11824	𑂤	DOGRA LETTER RA
11825	𑂥	DOGRA LETTER LA
11826	𑂦	DOGRA LETTER VA
11827	𑂧	DOGRA LETTER SHA
11828	𑂨	DOGRA LETTER SSA
11829	𑂩	DOGRA LETTER SA
1182A	𑂪	DOGRA LETTER HA
1182B	𑂫	DOGRA LETTER RRA

**Dependent vowel signs**

1182C	𑂬	DOGRA VOWEL SIGN AA
1182D	𑂭	DOGRA VOWEL SIGN I
1182E	𑂮	DOGRA VOWEL SIGN II
1182F	𑂯	DOGRA VOWEL SIGN U
11830	𑂰	DOGRA VOWEL SIGN UU
11831	𑂱	DOGRA VOWEL SIGN VOCALIC R
11832	𑂲	DOGRA VOWEL SIGN E
11833	𑂳	DOGRA VOWEL SIGN AI
11834	𑂴	DOGRA VOWEL SIGN O
11835	𑂵	DOGRA VOWEL SIGN AU

**Various signs**

11836	𑂶	DOGRA SIGN ANUSVARA
11837	𑂷	DOGRA SIGN VISARGA
11838	𑂸	DOGRA SIGN VIRAMA
11839	𑂹	DOGRA SIGN NUKTA

**Punctuation**

1183A	𑂺	DOGRA ABBREVIATION SIGN
-------	---	-------------------------

Set No 2

Comparison of the alphabets of Nagri Gurumukhi with Different Kinds of Landa Sarafi & with Thakuri and Dogri

Roman equivalent	गुरुमुखी	नगरी	लण्डा I	लण्डा II	लण्डा III	पठिया लण्डा	पठिया लण्डा	पठिया लण्डा	पठिया लण्डा	पठिया लण्डा	पठिया लण्डा	पठिया लण्डा	पठिया लण्डा	पठिया लण्डा	पठिया लण्डा
	Gurumukhi	Nagri	Lande I	Lande II	Lande III	Nagri Land	Patiala Land	Landi Sarafi	Dehli Sarafi	Urdi eguriant letters	Takri	Dogri	Kiraki		
i	ੳ irā	इ	ੳ	ੳ	ੳ	ੳ	ੳ	ੳ	ੳ	ੳ	ੳ	ੳ			
e	ਅ airā	अ	ਅ	ਅ	ਅ	ਅ	ਅ	ਅ	ਅ	ਅ	ਅ	ਅ			
i	ੲ iri	इ=ए	ੲ	ੲ	ੲ	ੲ	ੲ	ੲ	ੲ	ੲ	ੲ	ੲ			
s	ਸ sasa	स=श	ਸ	ਸ	ਸ	ਸ	ਸ	ਸ	ਸ	ਸ	ਸ	ਸ			
h	ਹ hahā	ह	ਹ	ਹ	ਹ	ਹ	ਹ	ਹ	ਹ	ਹ	ਹ	ਹ			
k	ਕ kakkā	क	ਕ	ਕ	ਕ	ਕ	ਕ	ਕ	ਕ	ਕ	ਕ	ਕ			
kh	ਖ khakkhā	ख=ख	ਖ	ਖ	ਖ	ਖ	ਖ	ਖ	ਖ	ਖ	ਖ	ਖ			

g	ਗੰ gā	ग	ਗ	ਗ	ਗ	ਗ	ਗ	ਗ	ਗ	ਗ	ਗ	ਗ	
gh	ਗਘ gāghā	घ	ਘ	ਘ	ਘ	ਘ	ਘ	ਘ	ਘ	ਘ	ਘ	ਘ	
	ੳ ngā	ङ	ੳ	ੳ	ੳ	ੳ	ੳ	ੳ	ੳ	ੳ	ੳ	ੳ	
ch	ਚ chā	च	ਚ	ਚ	ਚ	ਚ	ਚ	ਚ	ਚ	ਚ	ਚ	ਚ	
chh	ਚਘ chāghā	छ	ਚ	ਚ	ਚ	ਚ	ਚ	ਚ	ਚ	ਚ	ਚ	ਚ	
j	ਜ jā	ज	ਜ	ਜ	ਜ	ਜ	ਜ	ਜ	ਜ	ਜ	ਜ	ਜ	
jh	ਜਘ jāghā	झ	ਜ	ਜ	ਜ	ਜ	ਜ	ਜ	ਜ	ਜ	ਜ	ਜ	
	ੲ njā	ञ	ੲ	ੲ	ੲ	ੲ	ੲ	ੲ	ੲ	ੲ	ੲ	ੲ	
t	ਟ tā	ट	ਟ	ਟ	ਟ	ਟ	ਟ	ਟ	ਟ	ਟ	ਟ	ਟ	

Figure 1: Comparison of Old Dogra with Takri and other scripts (from Leitner 1882: Set 2).

kh	ठ	thatha	ठ	ठ	ॠ	ॡ	ॢ	ॣ	।	॥	०	
d	ड	dadala	ड	ड	ॢ	ॣ	।	॥	०	॥	०	
th	ठ	thadla	ठ	ठ	ॢ	ॣ	।	॥	०	॥	०	
	ह	haza	ह	ह	ॢ	ॣ	।	॥	०	॥	०	
z	ज	jalta	ज	ज	ॢ	ॣ	।	॥	०	॥	०	
h	घ	ghatla	घ	घ	ॢ	ॣ	।	॥	०	॥	०	
z	ह	hadla	ह	ह	ॢ	ॣ	।	॥	०	॥	०	
z	घ	ghatla	घ	घ	ॢ	ॣ	।	॥	०	॥	०	
N	न	nanna	न	न	ॢ	ॣ	।	॥	०	॥	०	

P	प	pappa	प	प	प	ॢ	ॣ	।	॥	०	
ph	फ	phappa	फ	फ	ॢ	ॣ	।	॥	०	॥	
b	ब	babba	ब	ब	ॢ	ॣ	।	॥	०	॥	
bh	भ	bhabba	भ	भ	ॢ	ॣ	।	॥	०	॥	
m	म	manma	म	म	ॢ	ॣ	।	॥	०	॥	
y	य	yayya	य	य	ॢ	ॣ	।	॥	०	॥	
r	र	rara	र	र	ॢ	ॣ	।	॥	०	॥	
l	ल	lalla	ल	ल	ॢ	ॣ	।	॥	०	॥	
w	व	wana	व	व	ॢ	ॣ	।	॥	०	॥	
r	र	rara	र	र	ॢ	ॣ	।	॥	०	॥	

Figure 2: Comparison of Old Dogra with Takri and other scripts (from Leitner 1882: Set 2).

PAÑJĀBĪ.

DŌGRA DIALECT.

(JAMMU STATE.)

SPECIMEN I.

DŌGRA CHARACTER.

ਪੰਚ ਨਾਮਾ ਸੁ ਸੁ ਮੁਕੁਤ ਆਪ ਸਿ ਪਾਸ  
 ਨਾਰ ਸੁਪੁ ਪੰ ਪ ਵਾ ਸੁਯੋਲ ਨਰ ਤਪ ਧੰ  
 ਕੁਪਸੁਤਾ ਸੁਕੁ ਤਪ ਤੇਨ ਸੇਕਾ ਮੋਕੁਦ-  
 ਤਪ ਨਤ ਸੇਕਾ ਸੁ-ਸੁ ਤਸੁ ਸਿਨੇ ਸੁਲ  
 ਠਿਪਕਾ ਪਾਗ-ਸੁ ਸੁਕੁ ਆਸਿ ਸੁਦ ਮੋਦੁ  
 ਕੇਵੇ ਮੁਕੁ = ਕੁ ਸੁਪ-ਕੁਕੁ ਕੰ ਯ - ਕੁ  
 ਸੁਕੁ ਸੁਪੁ ਸੁਕੁ ਸੁਕੁ ਕੁਕੁ ਸੁਕੁ ਸੁਕੁ  
 ਸੁਕੁ ਸੁਕੁ ਸੁਕੁ ਸੁਕੁ ਕੁਕੁ ਸੁਕੁ-ਸੁਕੁ  
 ਸੁਕੁ ਕੁਕੁ ਸੁਕੁ ਸੁਕੁ ਕੁਕੁ ਸੁਕੁ ਸੁਕੁ  
 ਸੁਕੁ ਪਾਸ ਪਾਸ ਕੁਕੁ ਸੁਕੁ ਸੁਕੁ  
 ਤੁਕੁ ਕੁਕੁ ਤੁਕੁ ਕੁਕੁ ਸੁਕੁ ਸੁਕੁ ਸੁਕੁ  
 ਕੁਕੁ ਪਾਸ ਕੁਕੁ ਸੁਕੁ ਕੁਕੁ ਕੁਕੁ

Figure 3: Specimen of the Dogri language written in Old Dogra (from Grierson 1916: 760).

PAÑJĀBĪ.

ḌOGRĀ DIALECT.

(JAMMU STATE.)

SPECIMEN II.

ḌOGRĀ CHARACTER.

੧ ਤੰਠੁ ਭਾਲੁ ਆਤਪਠੰਠਿ ਸੁਠੁ ਠਠੁ  
 ਠਠਾਠੰਠਾ. ਸਠੁਠੁ ਠੁਠੁ ਠੁਠੁ ਠੁਠੁ  
 ਠਠਾਠੰਠਾ ਠੁਠੁਠੁ

੨ ਤੰਠੁ ਠੁਠੁ ਠੁਠੁ . ਸਠੁਠੁ ਠਠਾਠੰਠੁ  
 ਠੁਠੁ ਠੁਠੁ ਠੁਠੁ ਠੁਠੁ ਠੁਠੁ ਠੁਠੁ  
 ਠੁਠੁ ਠੁਠੁ ਠੁਠੁਠੁ

੩ ਤੰਠੁ ਠੁਠੁ ਠੁਠੁ ਠੁਠੁ ਠੁਠੁ  
 ਠਠਾਠੰਠੁ ਠੁਠੁ ਠੁਠੁ ਠੁਠੁ ਠੁਠੁ  
 ਠਠਾਠੰਠੁ ਠੁਠੁ ਠੁਠੁ

੪ ਤੰਠੁ ਠੁਠੁ ਠੁਠੁ ਠੁਠੁ ਠੁਠੁ  
 ਠੁਠੁ ਠੁਠੁ ਠੁਠੁ ਠੁਠੁ ਠੁਠੁ ਠੁਠੁ  
 ਠੁਠੁ ਠੁਠੁ

Figure 4: Another specimen of the Dogri language written in Old Dogra (from Grierson 1916: 772).

The following are the Dōgrā characters as used in the specimens :—

**Vowels.**

**INITIAL FORMS.**

क a, ऌ ā, ॐ i or ī, ॐ u or ū, ॐ or ॐ  
 e, ē or ai, ॐ ō or au, ॐ m or ~

**NON-INITIAL FORMS.**

ॐ ka, ॐ kā, ॐ ki or kī, ॐ or ॐ ku, ॐ kū, ॐ ke or kē,  
 ॐ kai, ॐ kō, ॐ kau, ॐ kam or kām.

NOTE.—Great carelessness is allowed in writing the vowels and the nasal sign. They are often omitted altogether. Long and short vowels are frequently interchanged. Initial vowels are often written in the place of non-initial long ones: Thus—

ॐ for ॐ dā; ॐ for ॐ tī. The letter e or ē is frequently written for i, and ō for u.

**Consonants.**

ॐ ka, ॐ kha, ॐ ga, ॐ gha, ॐ na;  
 ॐ cha, ॐ chha, ॐ ja, ॐ jha, ॐ ṅ;  
 ॐ ṭa, ॐ ṭha, ॐ ḍa, ॐ ḍha, ॐ or ॐ ṇa;  
 ॐ ta, ॐ tha, ॐ da, ॐ dha, ॐ na;  
 ॐ pa, ॐ pha, ॐ ba, ॐ bha, ॐ ma;  
 ॐ ya, ॐ ra, ॐ la, ॐ va, wa;  
 ॐ sha, ॐ sa, ॐ ha, ॐ ra.

NOTE.—That the same sign is employed for ja and ya, and for ba and va (or wa), respectively. There is really only one sibilant,—the letter sa. When it is necessary to represent the sound of the Persian ṣh, the character for chha is employed.

Figure 5: Chart of the Old Dogra script (from Grierson 1916: 641).

In order to facilitate comparison, I next give the current written forms of the letters of the Gurmukhi, Kāṅgrā, and Ḍōgrā alphabets.

Gurmukhi.	Kāṅgrā.	Ḍōgrā.		Gurmukhi.	Kāṅgrā.	Ḍōgrā.	
ਅ	ਯ	ਯ	'āirā'	ੳ	ੳ	ਭੁਭ	ḍa
ੲ	ਠ	ਠ	'īrā'	ਢ	ਢ	ਢੁਢ	ḍha
ਊ	ਊ	ਊ	'ūṛā'	ਘ	=	ਘੁਘ	ḡa
ਊ	ਊ	ਊ	ō	ਤ	ਤ	ਤੁ	ta
ਸ	ਸ	ਸ	sa	ਬ	ਬ	ਬੁਬ	tha
ਹ	ਹ	ਹ	ha	ਦ	ਦ	ਦੁਦ	da
ਕ	ਕ	ਕ	ka	ਪ	ਪ	ਪੁਪ	dha
ਖ	ਖ	ਖ	kha	ਨ	ਨ	ਨੁ	na
ਗ	ਗ	ਗ	ga	ਪ	ਪ	ਪੁ	pa
ਘ	ਘ	ਘ	gha	ਠ	ਠ	ਠੁ	pha
ਙ	ਙ	ਙ	ṅa	ਬ	ਬ	ਬੁ	ba
ਚ	ਚ	ਚ	cha	ਭ	ਭ	ਭੁ	bha
ਛ	ਛ	ਛ	chha	ਮ	ਮ	ਮੁ	ma
ਜ	ਜ	ਜ	ja	ਯ	...	ਯੁ	ya
ਝ	ਝ	ਝ	jha	ਰ	ਰ	ਰੁ	ra
ਞ	...	ਞ	ṅa	ਲ	ਲ	ਲੁ	la
ਟ	ਟ	ਟ	ṭa	ਵ	ਵ	ਵੁ	va
ਠ	ਠ	ਠ	ṭha	ਸ਼	ਸ਼	ਸ਼ੁਸ਼	ra

Figure 6: Comparison of Gurmukhi, Kangra, and Old Dogra scripts (from Grierson 1916: 642).



	Nagari	Dogri		Nagari	Dogri		Nagari	Dogri
a	अ	ਲ	ka	क	ਕ	ta	त	ਤ
ā	आ	ਲ਼	kha	ख	ਖ	tha	थ	ਥ
i	इ	ੳ	ga	ग	ਗ	da	द	ਦ
ī	ई	ੳ਼	gha	घ	ਘ	dha	ध	ਧ
u	उ	ੳ	nga	उ	ੳ	na	न	ਨ
ū	ऊ	ੳ਼						
e	ए	ੲ	ca	च	ਚ	pa	प	ਪ
ai	ऐ	ੲੲ	cha	छ	ਚ਼	pha	फ	ਫ
o	ओ	ੳੲ	ja	ज	ਜ	ba	ब	ਬ
au	औ	ੳੲੲ	jha	झ	ਝ	bha	भ	ਭ
am (nasalized vowel)	अं	ੳਂ	ña	ञ	ਞ	ma	म	ਮ
śa	श	ਸ਼	ṭa	ट	ਟ	ya	य	ਯ
ṣa	ष	ਸ਼	ṭha	ठ	ਠ	ra	र	ਰ
sa	स	ਸ	ḍa	ड	ਡ	la	ल	ਲ
ha	ह	ਹ	dha	ढ	ਢ	va	व	ਵ
			ṇa	ण	ਣ			

Figure 7: Comparison of New Dogra ('Dogri') and Devanagari vowel and consonant letters (from Staal 1984: 33).

	Nagari Dogri			Nagari Dogri	
kā	का	ਕਾ	gā	गा	ਗਾ
ki	कि	ਕਿ	gī	गि	ਗਿ
kī	की	ਕੀ	gi	गी	ਗੀ
ku	कु	ਕੁ	gu	गु	ਗੁ
kū	कू	ਕੂ	gū	गू	ਗੂ
ke	के	ਕੇ	ge	गे	ਗੇ
kai	कै	ਕੈ	gai	गै	ਗੈ
ko	को	ਕੋ	go	गो	ਗੋ
kau	कौ	ਕੌ	gau	गौ	ਗੌ
kaṃ	कं	ਕਂ	gaṃ	गं	ਗਂ
k	क	ਕ	g	ग	ਗ

Figure 8: Comparison of New Dogra ('Dogri') and Devanagari vowel signs (from Staal 1984: 34).

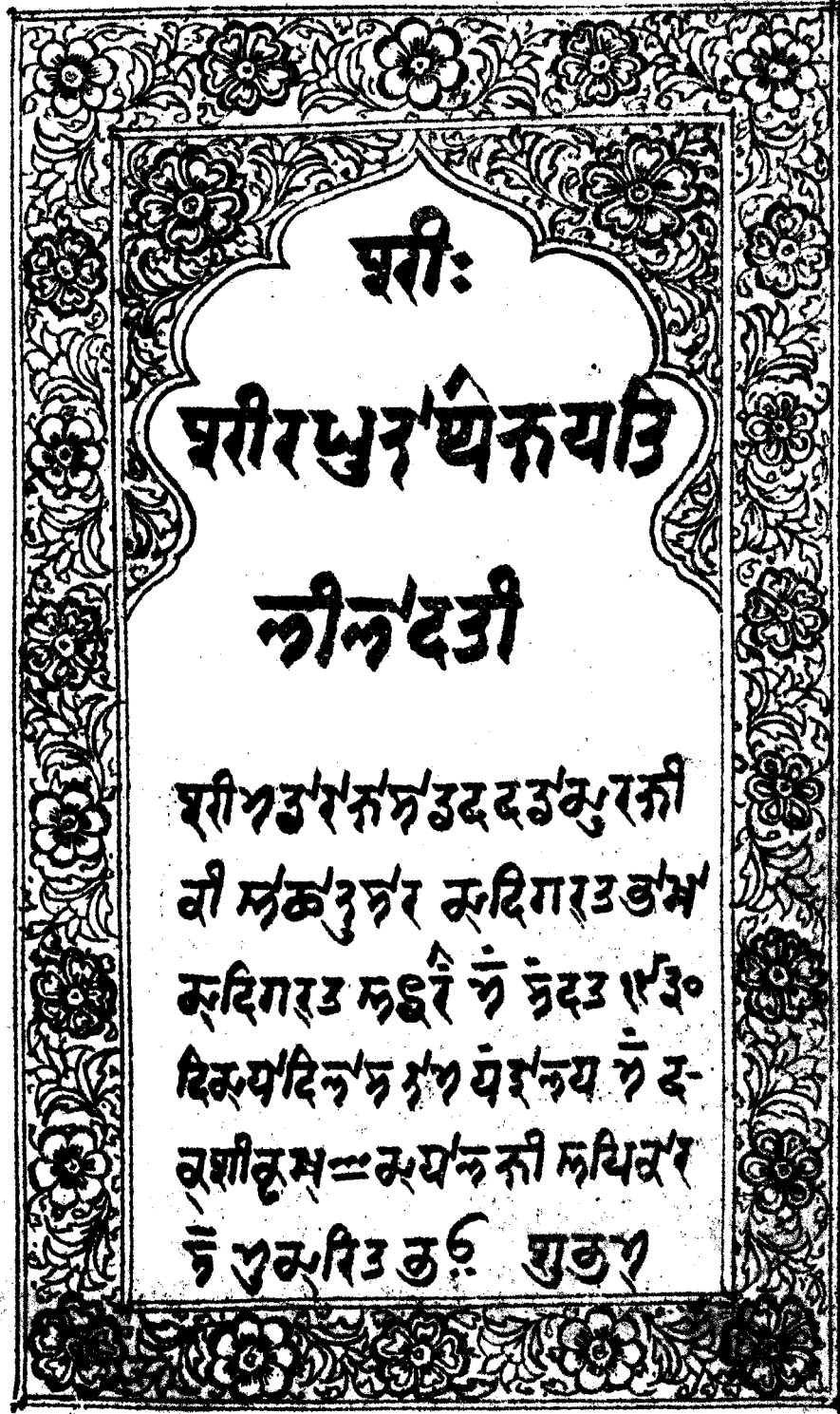


Figure 9: The first page of a Dogri translation of the *Līlavatī* by Bhāskarācārya, a Sanskrit treatise on mathematics, printed in the New Dogra script in 1872 (p. 1). Figure 10 contains the second page of this book.

ली.र.ड. २

उशीगः ॥ पायः ॥ १०० ॥  
 गः ॥ ३ ॥ अथ ॥ १ ॥ १०० ॥ १०० ॥ १०० ॥  
 उथरः ॥ ३ ॥ अथ ॥ १ ॥ १०० ॥ १०० ॥ १०० ॥  
 नदी ॥ ३ ॥ अथ ॥ १ ॥ १०० ॥ १०० ॥ १०० ॥  
 १३ ॥ ३ ॥ अथ ॥ १ ॥ १०० ॥ १०० ॥ १०० ॥  
 व ॥ ३ ॥ अथ ॥ १ ॥ १०० ॥ १०० ॥ १०० ॥  
 दीरः ॥ ३ ॥ अथ ॥ १ ॥ १०० ॥ १०० ॥ १०० ॥  
 १३ ॥ ३ ॥ अथ ॥ १ ॥ १०० ॥ १०० ॥ १०० ॥  
 १३ ॥ ३ ॥ अथ ॥ १ ॥ १०० ॥ १०० ॥ १०० ॥  
 १३ ॥ ३ ॥ अथ ॥ १ ॥ १०० ॥ १०० ॥ १०० ॥  
 १३ ॥ ३ ॥ अथ ॥ १ ॥ १०० ॥ १०० ॥ १०० ॥  
 १३ ॥ ३ ॥ अथ ॥ १ ॥ १०० ॥ १०० ॥ १०० ॥  
 १३ ॥ ३ ॥ अथ ॥ १ ॥ १०० ॥ १०० ॥ १०० ॥  
 १३ ॥ ३ ॥ अथ ॥ १ ॥ १०० ॥ १०० ॥ १०० ॥  
 १३ ॥ ३ ॥ अथ ॥ १ ॥ १०० ॥ १०० ॥ १०० ॥

Figure 10: Another page from the Lilavatī in New Dogra (p. 2).

ਜੀ.ਦ.ਭ.੩

**ਕ੍ਰਮ**

੧	੨	੩	੪	੫	੬	੭	੮	੯	੧੦
੨	੪	੬	੮	੧੦	੧੨	੧੪	੧੬	੧੮	੨੦
੩	੬	੯	੧੨	੧੫	੧੮	੨੧	੨੪	੨੭	੩੦
੪	੮	੧੨	੧੬	੨੦	੨੪	੨੮	੩੨	੩੬	੪੦
੫	੧੦	੧੫	੨੦	੨੫	੩੦	੩੫	੪੦	੪੫	੫੦
੬	੧੨	੧੮	੨੪	੩੦	੩੬	੪੨	੪੮	੫੪	੬੦
੭	੧੪	੨੧	੨੮	੩੫	੪੨	੪੯	੫੬	੬੩	੭੦
੮	੧੬	੨੪	੩੨	੪੦	੪੮	੫੬	੬੪	੭੨	੮੦
੯	੧੮	੨੭	੩੬	੪੫	੫੪	੬੩	੭੨	੮੧	੯੦
੧੦	੨੦	੩੦	੪੦	੫੦	੬੦	੭੦	੮੦	੯੦	੧੦੦
<b>ਨਿਰਵਾਰੀ ਨੁਮੀ</b>									
੧੧	੧੨	੧੩	੧੪	੧੫	੧੬	੧੭	੧੮	੧੯	੨੦
੨੨	੨੪	੨੬	੨੮	੩੦	੩੨	੩੪	੩੬	੩੮	੪੦
੩੩	੩੬	੩੯	੪੨	੪੫	੪੮	੫੧	੫੪	੫੭	੬੦

Figure 11: Page from the *Lilavati* showing the use of printed New Dogra digits (p. 3).

ੜੲੳ	ੳੲੲੳ	ੜੲੳ	ੳੲੲੳ
ੳੲੲੳੲੲ	)।	ੲੲੲੲ	≡)
ੲੲੲੲੲੲ	)॥	ੳੲੲੲੲੲ	≡)।
ੲੲੲੲੲੲ	)॥॥	ੳੲੲੲੲੲ	≡)॥
ੲੲੲੲ	- )	ੲੲੲੲੲੲ	≡)॥॥
ੳੲੲੲੲੲ	- )।	ੲੲੲੲੲ	)।
ੲੲੲੲੲੲ	- )॥	ੳੲੲੲੲੲੲ	)।

Figure 12: Usage of fractions for denoting currency in New Dogra (from Staal 1984: 75). The text appears to be typeset using metal fonts. The fraction signs are currency marks may be represented using characters from the ‘Common Indic Number Forms’ block (U+A830) in Unicode.

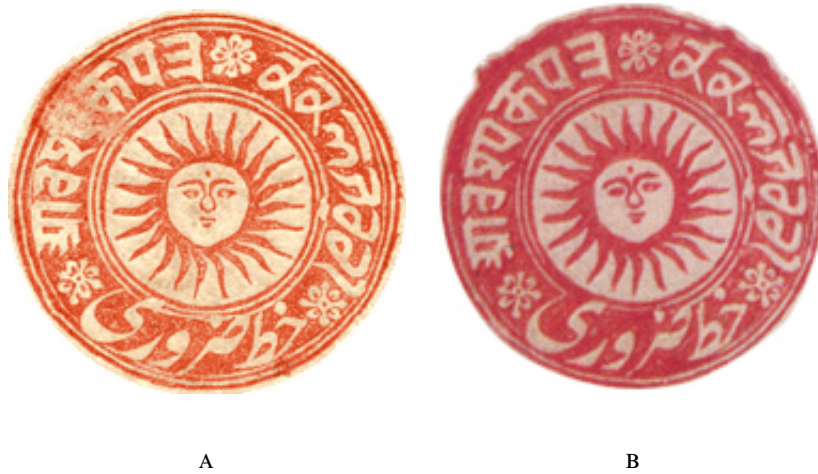
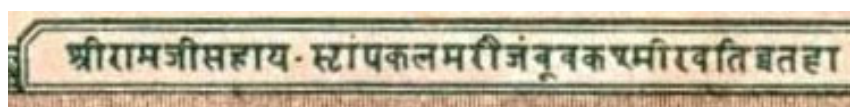


Figure 13: Two specimens of a trilingual stamp for expedited mail from Jammu State in Old Dogra, Persian, and Devanagari scripts. The Dogra text reads **ਕਕਲ ਜਰੂਰੀ** *kakal jarurī*. The Persian reads **خط ضروری** *khaṭ zarūrī*; the Devanagari reads **आवश्यक पत्र** *āvashyaka patra*.



Figure 14: Four bilingual regular postage stamps from Jammu State with text in the Old Dogra and Persian scripts. Stamps A, B, and D are half ‘anna’ stamps and C is a one ‘anna’ stamp. In all stamps, the Dogra text at the top reads **ਜਮੂ ਕਾਚਮੀਰ** *jamu kachamīr* ‘Jammu and Kashmir’. In stamps A, B, and D, the first line of the center is Persian **نیم آناه ۱۹۲۳ ۱۹۲۳** *[samvat] nim ānah* ‘half ānnā’; the second line of the center is Dogra **ਅਦਾ ਅਨਾ** *ada ana* ‘half ānnā’. In C, the first line of the center text is Persian **یک آناه ۱۹۲۳** *yak ānah* ‘one ānnā’; the second line of the center is Dogra **ੲਕ ਅਨਾ** *ek ana* ‘one ānnā’. In all four stamps, the third line of the center is the date 1923 *[samvat]* written in Devanagari-like digits १९२३. The Persian text at the bottom of the stamp reads **قلمرو سرکار جموں و کشمیر** *qalamrau sarkār jammūn o kaśmīr* ‘dominions of the ruler of Jammu and Kashmir.’



Detail of top, left (Devanagari)



Detail of top, right (New Dogra)



Detail of center (New Dogra, Devanagari, and Urdu)

Figure 15: The mast of a two rupee non-judicial stamp paper from Kashmir State.



Figure 16: The mast of a five rupee non-judicial stamp paper from Kashmir state containing text in New Dogra, Devanagari, and Arabic (Persian) scripts

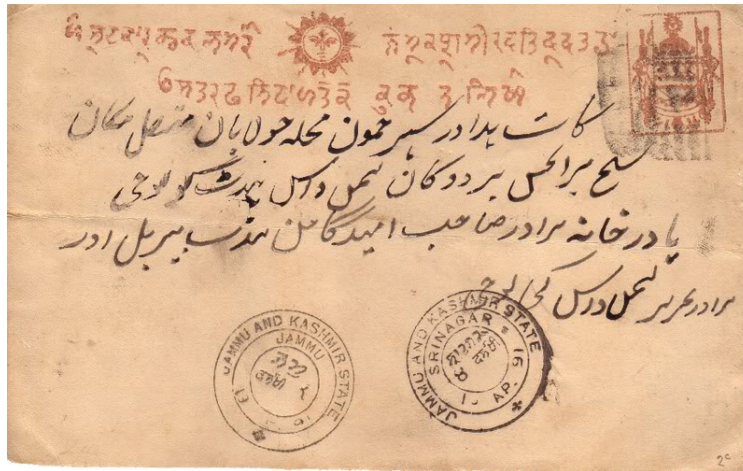


Figure 17: A one anna stamp from Jammu and Kashmir state with text in New Dogra, Devanagari, and Arabic (Urdu).





Figure 18: Telegraph stamps from Jammu and Kashmir state with text in Dogra, Devanagari, Urdu, and English. The Dogra denominations show usage of Common Indic fractions and the rupee mark.



*postakarda kalamrau jammu kasmir va tibbataha*  
 “Postcard of the realm of Jammu, Kashmir, and Tibet”

*isa tarapha siva pate ke kucha na likho*  
 “Do not write anything on this side except for the address”



*jamu*  
 “Jammu”

*bhaisakh (4 ?)*  
 “Vaisakh 4”



*sinagar*  
 “Srinagar”

*4 vaisakh*  
 “4 Vaisakh”

Figure 19: A state postcard from Jammu and Kashmir with post mark of 19 April 1891. The inscription at top is in New Dogra and the content is in Persian. The postmarks are in English and Dogra.



One rupee banknote



Twenty rupee banknote

Figure 20: Bank notes from Jammu and Kashmir State containing text in New Dogra, Devanagari, and Arabic (Persian) scripts.

	Chamba	Dogra		Chamba	Dogra
A	ਯ	ਯ	VOWEL SIGN A	<i>no dependent form</i>	
AA	ਯ	ਯ	VOWEL SIGN AA	ਯ	ਯ
I	ਏ	ਏ	VOWEL SIGN I	ਏ	ਏ
II	ਏ	—	VOWEL SIGN II	ਏ	ਏ
U	ਊ	ਊ	VOWEL SIGN U	ਊ	ਊ
UU	ਊ	—	VOWEL SIGN UU	ਊ	ਊ
E	ਏ	ਏ	VOWEL SIGN E	ਏ	ਏ
AI	ਏ	ਏ	VOWEL SIGN AI	ਏ	ਏ
O	ਓ	ਓ	VOWEL SIGN O	ਓ	ਓ
AU	ਊ	—	VOWEL SIGN AU	ਊ	ਊ

Table 2: Comparison of Chamba and Dogra vowel letters and signs.

	Chamba	Dogra		Chamba	Dogra
KA	क	क	DA	ड	ड
KHA	ख	ख	DHA	ध	ध
GA	ग	ग	NA	न	न
GHA	घ	घ	PA	प	प
NGA	ङ	ङ	PHA	फ	फ
CA	च	च	BA	ब	ब
CHA	छ	छ	BHA	भ	भ
JA	ज	ज	MA	म	म
JHA	झ	झ	YA	य	य
NYA	न्य	न्य	RA	र	र
TTA	ट	ट	LA	ल	ल
TTHA	ठ	ठ	VA	व	व
DDA	ड	ड	SHA	श	श
DDHA	ढ	ढ	SSA	—	स
NNA	न्न	न्न	SA	स	स
TA	त	त	HA	ह	ह
THA	थ	थ	RRA	र्र	र्र

Table 3: Comparison of Chamba and Dogra consonant letters.

**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://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:	<i>Proposal to encode the Dogra script in Unicode</i>
2. Requester's name:	<i>Anshuman Pandey / Script Encoding Initiative</i>
3. Requester type (Member body/Liaison/Individual contribution):	<i>Liaison contribution</i>
4. Submission date:	<i>4 November 2015</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>Dogra</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>59</i>
3. Proposed category (select one from below - see section 2.2 of P&P document):	
A-Contemporary <input 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 checked="" 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 type="checkbox"/>
b. Are the character shapes attached in a legible form suitable for review?	<input type="checkbox"/>
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</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: 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? 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>Christopher Shackle (SOAS, London)</i>
3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included? Reference:	No <i>Historical script, currently not used</i>
4. The context of use for the proposed characters (type of use; common or rare) Reference:	Common
5. Are the proposed characters in current use by the user community? If YES, where? Reference:	No
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:	N/A
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, or could be confused with, an existing character? If YES, is a rationale for its inclusion provided? If YES, reference:	No
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>See text of proposal</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