

**ISO/IEC JTC 1/SC 2/WG 2
PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS
FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646**

A. Administrative

1	Title:	Proposal to add Six characters in the Devanagari block for representation of Kashmiri language in Devanagari script.
2	Requester's name:	Ministry of Communications & IT, Government of India Through: Bureau of Indian Standards
3	Requester type (Member body/Liaison/Individual contribution):	National Member Body
4	Submission date:	14-10-2009
5	Requester's reference (if applicable):	
6	Choose one of the following: This is a complete proposal: or, More information will be provided later:	Yes If required more information may be provided later.

B. Technical – General

1	Choose one of the following:	
	a. This proposal is for a new script (set of characters): Proposed name of script:	No
	b. The proposal is for addition of character(s) to an existing block: Name of the existing block:	Yes DEVANAGARI
2	Number of characters in proposal:	Six
3	Proposed category: A-Contemporary 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	A
4	Proposed Level of Implementation (1, 2 or 3): Is a rationale provided for the choice? If Yes, reference:	1
5	Is a repertoire including character names provided? a. If YES, are the names in accordance with the "character naming guidelines" in Annex L of P&P document? b. Are the character shapes attached in a legible form suitable for review?	Yes
6	Who will provide the appropriate computerized font (ordered preference: True Type, or PostScript format) for publishing the standard? If available now, identify source(s) for the font (include address, e-mail, ftp-site, etc.) and indicate the tools used:	Government of India

7	References: Are references (to other character sets, dictionaries, descriptive texts etc.) provided? Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?	Yes
8	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)?	No
9	Additional Information:	

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 User Groups
3	Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included? Reference:	20 Lakhs approximately
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:	Yes Kashmir region of India as well as Kashmiri spread across India & world
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:	Yes
7	Should the proposed characters be kept together in a contiguous range (rather than being scattered)?	Preferably these should be placed contiguous
8	Can any of the proposed characters be considered a presentation form of an existing character or character sequence? If YES, is a rationale for its inclusion provided? If YES, reference:	No
9	Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters? If YES, is a rationale for its inclusion provided? If YES, reference:	No

10	Can any of the proposed character(s) be considered to be similar (in appearance or function) to an existing character? If YES, is a rationale for its inclusion provided? If YES, reference:	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:	No
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)	No
13	Does the proposal contain any Ideographic compatibility character(s)? If YES, is the equivalent corresponding unified ideographic character(s) identified? If YES, reference:	No

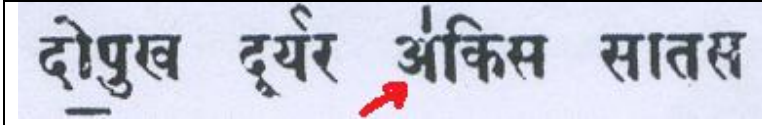
Kashmiri in Devanagari script

Kashmiri is mainly written in Perso-Arabic. Devanagari script is also used for writing Kashmiri. Central Hindi Directorate (CHD), Ministry of Human Resources Development, Government of India is the recommending authority of *Parivardhit Devanagari* (Extended Devanagari Set for representation of the other Indian languages in Devanagari). Department of Information Technology and CHD in consultation with all the stakeholders of the Kashmiri language which includes Experts from CIIL Mysore, Central Hindi Directorate (CHD), Commission for Scientific & Technical Terminology (Ministry of MHRD, Govt. of India), C-DAC and other Kashmiri-Devanagari experts finalized the additional characters which need to be encoded in the ISO 10646 and Unicode. The experts opined that there is need to encode 10 characters in the Unicode/ ISO 10646 for representation of the Kashmiri language. These characters are shown in the table below:

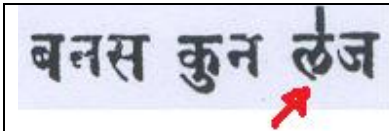
	Glyph	Name
1	अु	Kashmiri Devanagari Vowel High Central Unrounded U
2	अू	Kashmiri Devanagari Vowel High Central Unrounded UU
3	उ̣	Kashmiri Devanagari Vowel Sign High Central Unrounded U
4	उ̤	Kashmiri Devanagari Vowel Sign High Central Unrounded UU
5	अं	Kashmiri Devanagari Vowel Mid Central Unrounded O
6	अॉ	Kashmiri Devanagari Vowel Mid Central Unrounded OO
7	ँ̣	Kashmiri Devanagari Vowel Sign Mid Central Unrounded O
8	ँ̤	Kashmiri Devanagari Vowel Sign Mid Central Unrounded OO
9	औ	Kashmiri Devanagari Vowel Back Low Rounded O
10	औ̣	Kashmiri Devanagari Vowel Sign Back Low Rounded O

The characters shown at Serial No. 1-4 are already proposed through document no. ISO/IEC JTC1/SC2/WG2 N3480 and L2/08-250 by Mr. Praveen Satpute and Mr. Micheal Everson. Government of India has already recommended the encoding of these four characters through letter no. 13(6)/2008 -HCC(TDIL) dated August 11th, 2009 submitted to Unicode Consortium. The reference document no. is L2/09-311, as placed at Annexure-I.

The characters shown at Serial No. 5, अ is Devanagari Vowel Mid Central Unrounded O and character at Serial No. 7, ँ is the Vowel Sign for the same. This character is very frequently used in the Kashmiri Devanagari text. Following are some published examples of the usage of these characters.



Example 1 - Usage of "Devanagari Vowel Mid Central Unrounded O" taken from "Ramayan Kashmiri Bhasha Mein" shown with red arrow.

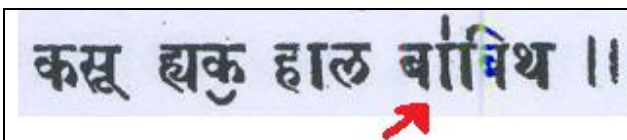


Example 2 - Usage of "Devanagari Vowel Sign Mid Central Unrounded O" taken from "Ramayan Kashmiri Bhasha Mein" shown with red arrow.

The characters shown at Serial No. 6, औ is "Devanagari Vowel Mid Central Unrounded OO" and character at Serial No. 8, ऌ is the Vowel Sign for the same. This character is very frequently used in the Kashmiri Devanagari text. Following are some published examples of the usage of these characters.

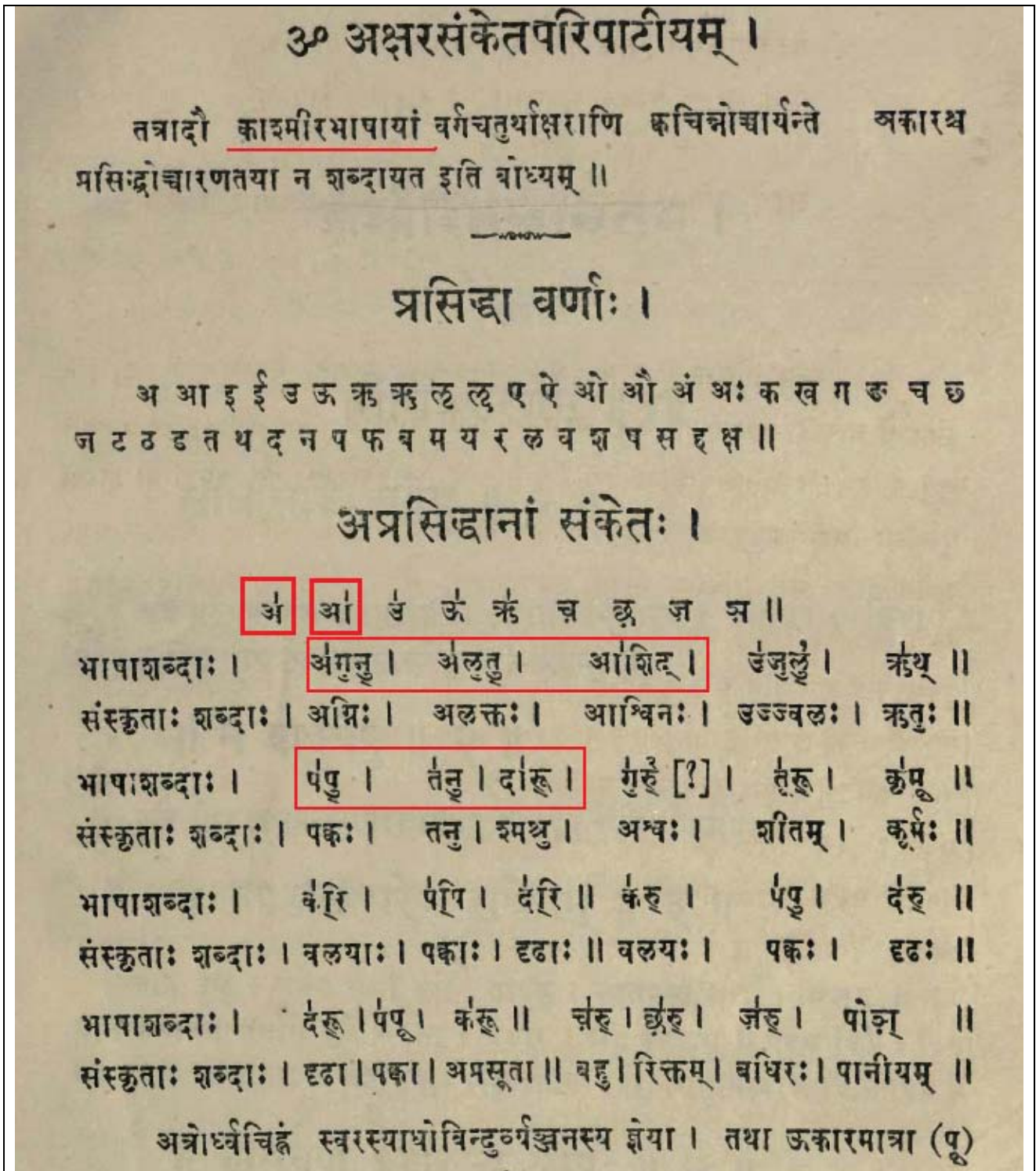


Example 3 - Usage of "Devanagari Vowel Mid Central Unrounded OO" taken from "Ramayan Kashmiri Bhasha Mein" shown with red arrow.



Example 4 - Usage of "Devanagari Vowel Sign Mid Central Unrounded OO" taken from "Ramayan -Kashmiri Bhasha Mein" shown with red arrow.

The complete text of the reference page may be seen at Annexure -II. There are many examples of usage of these characters in different publications as shown below:



Usage Example 5: The Kashmiri Vowels shown in "KAshmlra-shabdAmRtaM" written by Shri Ishwar Kaul and edited by Mr. George Grierson, published in 1889.

THE KĀSHMĪRĪ ALPHABET IN THE ROMAN CHARACTER WITH

A. VOWELS

Letter	Nāgarī Equivalent	Persian Equivalent	As in
a	अ (medial only)	ا	āh, अह, آه
ā	आ (medial only)	ا	andār, अन्दर, آندَر
ab	अब	ا, ا	abal, अबल, اَبَل
ad	अद, अ	ا, ا	ad ^a run, अदरुन, اَدْرُن; ganzār, गंज़र, گَنْزَر
ās	अस, अ	ا, ا	ās ⁱ , असि, آس; kār, करि, کَر
āb	आब, आ	آ	āb, आब, آب; kāth, काठ, کَاث
ai	ऐ, ई	ای, اے	aith, ऐठ, آیتھ; mail, मेल, مِل
au	औ, औ	او	gauv, गौव, گَو
ē	ए	ا	bañ ^a rāwun, बंजरावुन, بَنْزَرَاوُن
ō	य (after a consonant) or य (after certain consonants)	ا	vēth, वथ, وِث; bēñē, बंन, بَنْن. Initial य, or य after a vowel, is yē, as in yēmīs, यमिस, يَمِيس; biyē, बिय, بِي
ō	य (after a consonant) or य (after certain consonants)	ا	vēth ^a run, वदरुन, وِثْرُن
ō	ditto	ا	tēth ^a , तद, تَد
ē	इ (never initial)	ي	mēl, मेल, مِل
i	इ " "	ا	ās ⁱ , असि, آس
i	इ " "	ا	nishē, निश, نِش
ī	ई " "	ای	shīn, शीन, شِین
o	उ (medial only)	ا	d ^o n ^a , दनु, دَنْ
ō	व (after a consonant)	ا	dōd, दद, دَد
ō	व " "	ا	gōn ^a , गनु, گَنْ
ō	व " "	ا	shōp ^a , शूप, شُپ
o	अ, अ	ا	ogun, अगुन, اُگُن; bod ^a , बड, بَد
ō	ओ, ओ	و, و	ōr, ओर, اور; pōsh, पोश, پُوش
ô	ओ, ओ	و, و	ōs ^a , ओस, اوس; mōl ^a , मोल, مُول
ō	आ, अ	آ (initial and medial)	ōs, आस, آس; mōl ⁱ , मालि, مَالِ
u	य (never initial)	omitted	kor ^a , कर्, گَر
u	य " "	ا	kus, कुस, کُس
ū	य " "	او	śūr, शूर, چُور
u	य or अ " "	omitted, or ا	kūr ^a , कूर, کَر; abād ^a r ^a , अबदर, اَبَدَر
ü	अ, अ	ا, ا	ün ^a , अऊ, اُن; būd ^a , बड, بَد
ū	ऊ, अ	ي	sūty, सूत्य, سِيَت

The nasalization indicated in the Nāgarī character by *anunāsika* (ँ) is represented in the roman character by the sign ~ over the nasalized vowel. Thus अँ, ā, and so on. This does not affect the alphabetical order. *Anusvāra* (ं) is represented by m or n according to its sound. In Kāshmirī words it occurs only as a *compendium scripturæ* for a nasal before a consonant of the same class. Thus, खँडु for खँडु mōṇḍ^a.

Other common words examples for the proposed characters:

अ
ं

Arm -> नर

bird -> चर

unbroken -> अर

butter -> थन्य

आ
ी

mother -> माज

(she) was -> आस्य

last -> आखरी

to unknown -> गारन

boatman -> हान्ज़

The characters shown at Serial No. 9, औ is "Devanagari Vowel Back Low Rounded O" and character at Serial No. 10, ौ is the Vowel Sign for the same. This character and its vowel sign are recommended by the experts to represent Kashmiri sound close to pronunciation of International Phonetic symbol ɔ. Presently it is represented by the consonantal alphabet व (va). All the Devanagari vowels have vowel signs (*matras*) but the Devanagari consonantal alphabet letter व (va) does not have any vowel sign. Hence the experts have recommended new signs (glyphs) for this sound and its vowel symbol.

Word Examples:

Hindi word सुख (Sukh) is pronounced as sɔkh in Kashmiri, if the present व (va) is used for writing this word it will be written as स्वख and a person not knowing Kashmiri will pronounce it as "svakh", which is not correct. But when it is written using the new vowel sign it will be pronounced correctly i.e. सुख -> sukh -> sɔkh -> सौख.

Other example words:

खुश -> khush -> khɔsh -> खौश (happy)

पैर -> khush -> khɔsh -> खौर (foot)

दूध -> doodh -> dɔd -> दौद (milk)

Kashmiri words like द्रक्ख (vomit), न्योथ (thumb) etc can not be written at all with the help of व (va) as it is not having any vowel sign (*matra*). But these may be correctly represented using new signs.

vomit -> द्रौख -> drɔkh

thumb -> न्यौठ -> nyɔth

The printed usage examples for these characters are not presently available as the experts have recommended its usage in all future publications of Kashmiri language text in Devanagari script. CHD has also recommended usage of these characters and these will be shortly recommended in the official publication of Central Hindi Directorate "*Devanagari Lipi Tatha Hindi Vartani Ka Mankikaran*".



Swaran Lata
Director & HoD
Human Centred Computer Division

L2/09-311

भारत सरकार

GOVERNMENT OF INDIA

संचार और सूचना प्रौद्योगिकी मंत्रालय

MINISTRY OF COMMUNICATIONS AND INFORMATION TECHNOLOGY

सूचना प्रौद्योगिकी विभाग

DEPARTMENT OF INFORMATION TECHNOLOGY

इलेक्ट्रॉनिक्स निकेतन

ELECTRONICS NIKETAN

6, सी.जी.ओ. कॉम्प्लेक्स/6, C.G.O. COMPLEX

नई दिल्ली/New Delhi-110003

दिनांक/Date:

Email: slata@mit.gov.in
Telephone: 011-24301242

दूरभाष/Tele :

फैक्स/फैक्स :

D.O.No. 13/8/2008-HCC(TDIL)

August 11th, 2009

Subject: Encoding of additional characters in Unicode for representation of Kashmiri.

Dear Dr Davis,

This is with reference to the encoding of the additional characters for representation of Kashmiri in the Unicode Standard. The issue was discussed by the experts in detail, during various meetings at DIT and Central Hindi Directorate (CHD), New Delhi.

Department of IT, Govt of India has finalized the additional characters required for Kashmiri in consultation with CHD. There is need to encode 10 characters in the Unicode for the same. Out of these 10 characters, 4 characters are already proposed through Document L2/08-250 (ISO/IEC JTC1/SC2/WG2 N3480). Department of IT recommends the encoding of these four characters in the Devanagari block of Unicode Standard with the GOI recommended names as shown in the attached document. For the remaining six characters DIT will be submitting the formal proposal by September end.

Thanking You,

Yours sincerely,

(Swaran Lata)

Dr. Mark E. Davis,
President
The Unicode Consortium
P.O. Box 391476
Mountain View, CA 94039-1476
USA

Kashmiri Vowels

	Code	Glyph	Name
1		अ	Devanagari Vowel High Central Unrounded U
2		ऊ	Devanagari Vowel High Central Unrounded UU
3		उ	Devanagari Vowel Sign High Central Unrounded U
4		ू	Devanagari Vowel Sign High Central Unrounded UU

रामायण कश्मीरी भाषा में

Edited and Published
BY
P. DAMOODHAR KOUL
REVISED BY
P. SHYAMLAL BIRA

SAMBAT BIKRAMI 2000

दामोदर कोल मैनेजर के प्रबन्ध से
काशमीर स्टैंडार्ड प्रेस श्रीनगर में छपी

मु वद बुज वदनु सत्य गोय रंग बेरंग मु वद बुज वदनु सत्य शीशस प्यवुय संग
 वो^०लुख यलि बुरजु त्रोवुख खासु मलमल पकान गय त्रनुवय अज राहि जंगल
 ति याज बुछ शेहर कयव लोकव रिवान दाय दो^०पुख क्या सनुवनस मंज कति रटन जाय
 दिलस प्यठ दाग रोट बुजुल्यव गुलालव दो^०पुख दूरर अकिस सातस नु चालव
 गुलाबस सोसुनुक ह्यव रंग साम्पुज ति डीशित हाल मसवल ह्यचनु काम्पुज
 सपुज सारी प्रजलबुज गुल अवार फोलन त्यलि यलि दर्शुन दिन दुवार
 पकान गय तथ कोहस प्यठ आल्यहिमुत रंग बदल गव ज्यंठ पोशस कारतिक कुंग
 सु कीकी शीनु छट चंज मंजहूर्य गयि तेज बनस कुन लंज खलायिकन पोह पनस रेज
 चटुख मंजिल रटुख मंज बन खोटुख पान खलुख पथ फीर सारी आय रिवान
 तिथय गय तिम डन्डक बन मंज रटुख जाय जन्म केछर तु कर्मस क्या लु परवाय
 खबर यलि बूज कौसल्यायि कोत गव बननि लंज राजु पोत्रस कुन चु कन थव

—००००००—

राजा राम चन्द्र जी के फिराक में कौसल्या का विलाप करना
 कौसल्यायि हुन्दि गोबुरो ॥ करुयो गूर गूर ॥ परायो राम रामा ॥ करुयो गूर गूर
 कोतू गोहं चु त्राविथ ॥ कस ह्यकु हाल बाविथ ॥ अनी कुस मनु नाविथ ॥ करुयो०
 लगुयो पोत छाये ॥ हिय करथस बु हाये ॥ नारस बुट बु लाये ॥ करुयो० ॥ कनन छय
 कनु वाजे ॥ श्री कृष्ण महाराजे ॥ जगतुक लुख चु राजे ॥ करुयो० ॥ फेरुहय अन्दि अन्दी
 लागय पोशि गन्दी ॥ जामु चान्य सुनु सुन्दी ॥ करुयो० ॥ लोल चोन बुज म्य आमो ॥
 छारुथो शेहरु गामो ॥ व्ययि गूछहम चु रामो ॥ करुयो० ॥ नेरुहा बाजारी ॥ ह्यडुनम
 लख सारी ॥ पादन लगय पारी ॥ करुयो० ॥ नेरुयो दारिपती ॥ लागुयो कारिपती ॥
 तारि दिल गोम यती ॥ करुयो० ॥ म्य कमो शाफ आसी ॥ तिम ति कोनु कांसि
 कासी ॥ चु गोहम वनु वासी ॥ करुयो० ॥ च्य पर्यथम बुरजु जामय ॥ बु छारथ शेहरु
 गामय ॥ परुयो राम राम ॥ कारुयो० ॥ लुलि मंज ललु नावथ ॥ जिगरस मंज बु सावथ ॥
 बुज तिनो कांसि हावथ ॥ करुयो० ॥ नेरुयो शामु लटे बार न्याज छिय च्य मटे ॥ गाशरु
 लालु तटे ॥ करुयो० ॥ दूरर नो बु चालय ॥ कस करथस हवालय ॥ लाजथस मायि

38-01/2008-प.पा./जुलाई/09



भारत सरकार

GOVERNMENT OF INDIA

केंद्रीय हिंदी निदेशालय

CENTRAL HINDI DIRECTORATE

मानव संसाधन विकास मंत्रालय

(MINISTRY OF HUMAN RESOURCE DEVELOPMENT)

(उच्चतर शिक्षा विभाग)

(DEPARTMENT OF HIGHER EDUCATION)

पश्चिमी खंड-7, रामकृष्णपुरम, WEST BLOCK-VII, RAMA KRISHNA PURAM

नई दिल्ली/NEW DELHI-110066

दिनांक/DATED :

तार : राजभाषा

Tele : Rajbhasha

फोन/Phone : 26178454, 26105211

26105213, 26105214, 26105243

26105312, 26105245, 26105526

फैक्स/Fax : 26100758, 26103160

26102575

www.hindinideshalaya.nic.in

8.7.2009

सेवा में

निदेशक

संचार और सूचना प्रौद्योगिकी मंत्रालय

सूचना प्रौद्योगिकी विभाग

इलेक्ट्रॉनिक्स निकेतन

6, सी.जी.ओ. कॉम्प्लेक्स

नई दिल्ली-110003

महोदय,

केंद्रीय हिंदी निदेशालय द्वारा आयोजित बैठक में कश्मीरी भाषा के विशेषज्ञों ने जो चिह्न निर्धारित किए हैं उसकी एक प्रति आपको सूचना हेतु भेजी जा रही है।

शुभकामनाओं सहित,

भवदीया

TDR
JX(MJ)

17/8/09

संलग्न - कश्मीरी के लिए निर्धारित नागरी चिह्न

(डॉ. शशि भारद्वाज)
उपनिदेशक

147-हि० अ०
15/7/09
58/HOD(MC)/09
17/8/09

Please file Unicode for Kashmiri-Devanagari

TDIL (Section)

Manoj
18/8/2009

कश्मीरी भाषा के लिए देवनागरी लिपि चिह्न

स्वर - अ आ इ ई उ ऊ ऋ ए
 ऌ ड ढ ण ङ

(ऐ) औ (औ) (अं) (अः)

(अ) ऌ (ँ) (ँ) (ँ) (ँ)

कश्मीरी के विशेष स्वर - अ आ अ औ ऌ औ
 ड ढ ण ङ

(कश्मीरी के विशेष स्वरों का कोशीप क्रम 'ऊ' के पश्चात् होगा।)

व्यंजन - क ख ग (घ) (ङ)
 च छ ज (झ) झ
 ट ठ ड (ढ) (ण)
 त थ द (ध) न
 प फ ब (भ) म
 य र ल व (ड़) (ढ़)
 श ष स ज ह
 (क्ष) (ज) (त्र) (थ)

नोट : 1. कोष्ठक में रखे गए वर्ण आगत अथवा तत्सम शब्दों के लिए ही सुरक्षित हैं, मूल कश्मीरी शब्दों में ये वर्ण प्रयुक्त नहीं होते।

2. कश्मीर मूल शब्दों में दोष मराण ध्वनिश्रुत का अभाव है। तत्सम तथा आगत शब्दों में इनका प्रयोग संभव है।

3. अ

(श्री वी. एन. वेताब)

(प्रो. मोहनलाल स)

18/6/2009
 (जे. शास्त्रीशेखर तोषवानी)