



8. INSROT – Indian Script to Roman Transliteration - (Revised)

Indian Scripts are phonetic and have similarity in alphabetic correspondence. Transliteration between Indian languages is simple, unambiguous and phonetically similar. However, there is need for a Scheme for transliteration from Indian Scripts to Roman Script. People in India have been using many different schemes for Hindi- Roman transliteration. ITRANS is one of the popular schemes. ITRANS provides choice and uses capital letters also. The present scheme is not suitable for searching the contents on the Internet. There is a need to evolve a Case-insensitive transliteration scheme to facilitate searching on web. INSROT (Indian Script Roman Transliteration) Table is proposed as a standard transliteration scheme. In the previous issue of the VishwaBharat@tdil newsletter the draft of the INSROT was published. Based on the feedback from user community some changes have been made in the proposed scheme. The use of capital letters is avoided in the modified scheme to facilitate case-insensitive search. Single quote (') that is on non-shift is used for this purpose. This is read as "prime". The revised transliteration table is given below :

Vowels

Syllabic Form	Intra Syllabic Form	Roman Equivalent
अ	-	a
आ	ा	aa
इ	ि	i
ई	ी	ii
उ	ु	u
ऊ	ू	uu
ऐ		e'
ए	े	e
ऐ	ै	ai
ओ	ो	o

औ	ौ	au
अँ	ँ	a'
आँ	ाँ	o'
ऋ	ृ	r'i
ॠ	ॄ	r'ii
ऌ		lr'i
ॡ		lr'ii
ं		om'
ँ		om_'
ः		o:
ः		o.
◌ (for Sindhi)		◌_

Consonants

	V̄A	V̄A	V̄A	VA	Nasals	
	Voiceless Unaspirated	Voiceless aspirated	Voiced unaspirated	Voiced aspirated		
Velars	क (ka)	ख (kha)	ग (ga)	घ (gha)	ङ (ng'a)	
Pre-palatals	च (cha)	छ (c'ha)	ज (ja)	झ (jha)	ञ (nj'a)	
Retro flexes	ट (t'a)	ठ (t'ha)	ड (d'a)	ढ (d'ha)	ण (n'a)	
Dentals	त (ta)	थ (tha)	द (da)	ध (dha)	न (na)	
Labials	प (pa)	फ (pha)	ब (ba)	भ (bha)	म (ma)	
Semi vowels	य (ya)		र (ra)	ल (la)	व (va)	
Sibilants		श (sha)	ष (s'a)	स (sa)		
Glottal			ह (ha)			
			क्ष (ks'a),	त्र (tra),	ज्ञ (jna),	ळ (l'a)

[Note : use h' for rendering as explicit ह rather an aspirated consonant, e.g., त्ह (th'a), ज्ह (jh'a), ब्ह (bh'a). On the basis of pragmatism ज्ञ (jna) is proposed rather than ज्ञ (jnj'). Moreover combination of ज् (j) and न (na) is grammatically invalid.]

Nukta Consonants

क (k.a)	ख (kh.a)	ग (g.a)	ज (j.a)
ड (d'.a)	ढ (d'h.a)	फ (ph.a)	य (y.a)
ळ (l'.a)	र (r.a)	न (n.a)	

Sindhi Sounds

ग (g_a)	ज (j_a)	ड (d'_a)	ब (b_a)
---------	---------	----------	---------

Special Character

अँ (a'um')	ऽ (~)	र (r_') [in Marathi]
------------	-------	----------------------



Non-join \

पक्व	= pakva
पक्व	= pak\va
सत्य	= sat\ya
पंक	= pam'ka
पङ्क	= pang'ka
पङ्क	= pang\ka
अर्थात्	= arthaat

Intra Syllabic Vowels Forms combined with consonant character

क (ka)	का (kaa)	कि (ki)	की (kii)
कु (ku)	कू (kuu)	कृ (kr'i)	के (ke)
कै (kai)	को (ko)	कौ (kau)	कॉ (ko')
कं (kam')	कँ (kam_')	कः (ka:)	क्र (kra)

Rule for decoding from Romanized text.

- 'a' is integral part of consonant. Omission of 'a' results into 'pure' consonants. e.g. ka (क), k (क्)
- Vowel combined with 'pure' consonant transforms into matra. e.g. khi (खि), ghuu (घू), ko' (कॉ)
- h is added to consonants/vowels to denote aspirated or closer sounds, eg. kha (ख), c'ha (छ), jha (झ), ...
- m' denotes anusvar/ bindu and combines with previous vowel as in am' (अं), kam' (कं), kim' (किं), svim'ga (स्विंग)
- h'a is used following a consonant to render explicit ह rather an aspirated character. To illustrate, ख (kha), कह (kh'a), घ (gha), गह (gh'a), छ (c'ha), चह (c'h'a), झ (jha), जह (jh'a), ठ (t'ha), टह (t'h'a), ढ (d'ha), डह (d'h'a), थ (tha), तह (th'a), ध (dha), दह (dh'a), फ (pha), फह (ph'a), भ (bha), बह (bh'a), श (sha), सह (sh'a). Such requirements will be rare.
- ◌ and ◌: are proposed for Nukta and Visarga. When ◌ and ◌: are used as punctuation marks,

minimum a space must follow. However ambiguity resolution during decoding in context is simple.

Example of word/phrase level transliteration

कट (kat'a), खाट (khaat'a), घृत (ghr'ita), चिप (chipa), झील (jhiila), बाण (baan'a), तुक (tuka), टूट (t'uut'a), थूक (thuuka), फोड़ (phod'.a), मैल (maila), षट (s'hat'a), साथ (saatha), आशा (aashaa), ईख (iikha), ऋषि (r'is'hi), और (aura), बूढ़ा (buud'h.aa), मक्खन (makkhana), झाड़वर (d'raaivara), अण्डा (an'd'aa), गद्दी (gaddii), सत्ताईस (sattaaiisa), पृथ्वी (pr'ithvii), प्रथम (prathama), कुँवर (kum_'vara), ख़्वाब (kh.baaba), श्वान (shvaana), प्राण (praan'a), पुण्य (pun'ya), गङ्गा (gang'gaa), ज्ञानज्योति (jnaanajyoti), नूपुर (nuupura), ध्वनि (dhvani), कॉलेज (ko'leja), हंस (ham'sa), हँस, (ham_'sa), आँखें (aam_'khem'), विपत्ति (vipatti), परिस्थिति (paristhiti), पॅन (pa'na), बॉल (bo'la), बाँके बिहारी (baam_'ke bih'aarii), गज्जा (ganj'\gaa), क्वाथ (k\vaatha), क्वा (kvaa), पद्य (padya), पद्य (pad'ya), निष्क्रिय (nis'hkriya), कृषि (kr'is'hi), जमदग्नि (jamadagni), डमरू (d'amaruu), रत्न (ratna), तन (tana), टन (t'ana), दाल (daala), डाल (d'aala), कृपा (kr'ipaa), क्रिया (kriyaa) ज्योति (jyoti), खजाना (kh.ajaanaa), चण्डीगढ़ (chan'd'iigad'h.a), दुःख (du:kha)

न हि ज्ञानेन सदृशं पवित्रमिह विद्यते.

(na hi jnaanena sad'r'isham' pavitrmiha vidyate)

Feedback on the basis of usage is solicited.

(Courtesy : TDIL Programme
Department of Information Technology
Ministry of Communications &
Information Technology
(Government of India)
6, CGO Complex, New Delhi-110003
Tel. : 011-4363076
E-mail : tdilinfo@mit.gov.in)