



7. Transliteration Tables

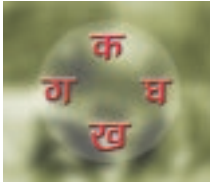
Punjabi to Romanized English Punjabi to Urdu Transliteration

This document presents the mapping from Punjabi letters to equivalent Urdu letters in romanized form, for transliterating from Punjabi text to Urdu Text, (that is from ISCII to USCII).

Following table has first entry for Punjabi letters, second Entry for the corresponding romanized English mapping for that character, and the third entry is the phonetic representation of that character in Urdu.

Punjabi	Roman	Urdu
ਕ	k	kaf ک
ਖ	kh	kha کہ
ਗ	g	gaf گ
ਘ	gh	Gha گھ
ਙ	Mg	نگ
ਚ	ch	cheem چ
ਛ	chh	chha چھ
ਜ	j	jeem ج
ਝ	J	zay ژ
ਞ	jh	jhha جه
ਟ	Mj	نج
ਠ	T	Te ٹ
ਡ	Th	Tha ٹھ
ਢ	D	Daal ڈ
ਣ	Dh	Dhaal ڈھ
ਨ	n	noon ن
ਤ	t	te ت
ਤ	t1	toe ط
ਥ	th	they تھ
ਦ	d	daal د

ਧ	dh	dhaal دھ
ਨ	n	noon ن
ਪ	p	pe پ
ਫ	ph	phey پھ
ਬ	b	be ب
ਭ	bh	bhey بھ
ਮ	m	meem م
ਯ	y	bari-ye ے
ਰ	r	re ر
ਲ	l	lam ل
ਵ	w	vao و
ਸ਼	sh	sheen ش
ਸ	s	seen س
ਸ	S	Se ث
ਸ	s1	swad ص
ਹ	H	bari - he ح
ਹ	x	chhoti-he ہ
ਹ	h	dochashmi he ه
ਕ	q	qaf ق
ਖ	Kh	khe خ
ਗ	G	ghain غ
ਫ	f	fe ف
ਜ	Z	zal ذ
ਜ	z	ze ز
ਜ	z1	zwad ض
ਜ	z2	zoe ظ
ੜ	R	Re ڑ
ੜ	Rh	Rhe ڑھ
ੌ (Anuswar)	M	noon-ghunna ن
ਅ]	ain ع
ਅ	O	hamza ء



Mappings For Vowel (Not At Beginning)

इ	When in between or end	OI
ਈ	When in between or end	OI
ਉ	When in between or end	OU
ਊ	When in between or end	uU
ਏ	When in between or end	OI
ਐ	When in between or end	O/I
ਓ	When in between or end	OU
ਔ	When in between or end	O/U

Below are the mappings that are used in production of vowel combinations

Vowel Letter	Matra	Mapping for vowel letter (Roman)	Mapping for matra (Roman)	Representation for vowel letter Urdu	Representation for matra Urdu
ਅ	halant	a	NILL	alif ا	
ਆ	ਾ	A	a	alif -mad آ	aliph ا
ਇ	ی	aY	Y	alif Zer اِ	Zer ِ
ਈ	ੀ	aYI	YI	alif zer ye اِی	zer ye یِ
ਉ	ੁ	au	u	alif pesh اُ	pesh ُ
ਊ	ੂ	auU	uU	alif vao pesh اُو	pesh vao ُو
ਏ	ੇ	e	e	baRi ye اے	baRi ye ے
ਐ	ੈ	ay	I	alif choti -ye اِی	aliph choti -ye اِی
ਓ	ੋ	aU	U	vao او	و
ਔ	ੌ	a/U	/U	vao اُو	وُ

˘ (To Represent W
repetition of
same character)
used when
character's
sound is halved

ˆ jazm

ˆ tashdid ˆ

Following are the mappings for numerals and punctuation's

۰	0	zero ۰
۱	1	one ۱
۲	2	two ۲
۳	3	three ۳
۴	4	four ۴
۵	5	five ۵
۶	6	six ۶
۷	7	seven ۷
۸	8	eight ۸
۹	9	Nine ۹
-	-	minus sign -
*	*	asterisk *
?	~	question mark ؟
({	open parenthesis (
)	}	close parenthesis)
/	#	forward slash /
+	\$	plus sign +
=	@	equal sign =
:	%	colon :
;	&	Semicolon ؛
,	,	Comma ،
“	“	Open double quote ”
”	”	Close double quote “
‘	‘	Open single quote ’
’	’	Close single quote ‘
-	-	sentence dash -
..	-	sentence dash ..



Hindi to Romanized English Hindi to Urdu Transliteration

This document presents the mapping from Hindi letters to equivalent Urdu letters in romanized form, for transliterating from Punjabi text to Urdu Text, (that is from ISCII to USCII).

Following table has first entry for Hindi letters, second Entry for the corresponding romanized English mapping for that character, and the third entry is the phonetic representation of that character in Urdu.

Hindi	Roman	Urdu
क	k	kaf ک
ख	kh	kha کھ
ग	g	gaf گ
घ	gh	gha گھ
ङ	Mg	نگ
च	ch	chim چ
छ	chh	chha चھ
ज	j	jim ج
झ	J	zay ज़
झ	jh	zay
झ	jh	jha جه
ञ	Mj	نج
ट	T	Te ٹ
ठ	Th	They ٹھ
ड	D	Daal ڈ
ढ	Dh	Dhaal ڈھ
ण	n	noon ن

त	t	te ت
त	t1	toe ط
त	t2	te-marbuta
थ	th	they تھ
द	d	daal د
ध	dh	dhaal دھ
न	n	noon ن
प	p	pe پ
फ	ph	phey फھ
ब	b	be ب
भ	bh	bhey भह
म	m	meem م
य	y	bari-ye ے
र	r	re ر
ल	l	lam ل
व	w	vao و
श	sh	sheen ش
ष	sh	sheen ش
स	s	seen س
स	S	Se سٹ
स	s1	swad ص
ह	H	bari - he ح
ह	h	dochashmi-he ه
ह	x	chhoti-he ه



क	q	qaf ق
ख	Kh	khe خ
ग	G	ghain غ
फ	f	fe ف
ज़	Z	zal ذ
ज़	z	ze ز
ज़	z1	zwad ض
ज़	z2	zoe ظ
र	ri	re ر
र	ri	re ر

Mappings For Vowel (Not At Beginning)

ि	When in between or end	OI
ी	When in between or end	OI
ु	When in between or end	OU
ू	When in between or end	uU
े	When in between or end	OI
ै	When in between or end	O/I
ो	When in between or end	OU
ौ	When in between or end	O/U

Below are the mappings that are used in production of vowel combinations

Vowel Letter	Matra	Mapping for vowel letter	Mapping for matra	Representation for vowel letter Urdu	Representation for matra Urdu
अ	halant	a	NILL	alif ا	NILL
आ	।	A	a	alif-mad ا	aliph ا

इ	ि	aY	Y	aliph Zer ا	Zer َ
ई	ी	aYI	YI	alif zer ye اِي	zer ye ِي
उ	ु	au	u	aliph pesh اُ	pesh ُ
ऊ	ू	auU	uU	aliph vao pesh اُو	pesh vao ُو
ए	े	e	e	baRi -ye اے	baRi -ye ے
ऐ	ै	ay	ay	aliph choti -ye اِي	aliph choti -ye ِي
ओ	ो	aU	U	vao او	و
औ	ौ	a/U	/U	vao او	و
		(To Represent repetition of same character)	W	tashdid	ّ
		used when character's sound is halved	^	jazm	
अं		//		tanveen (do-zabar)	
आ		^		khaRa-alif	

Following are the mappings for numerals and punctuation's

०	0	zero ۰
१	1	one ۱
२	2	two ۲
३	3	three ۳
४	4	four ۴
५	5	five ۵



६	6	six ८
७	7	seven ७
८	8	eight ८
९	9	Nine ९
—	-	minus sign -
*	*	asterisk *
?	~	question mark १
({	open parenthesis)
)	}	close parenthesis (
/	#	forward slash /
+	\$	plus sign +
=	@	equal sign =
:	%	colon :
;	&	Semicolon ;
,	,	Comma ,
“	“	Open double quote ”
”	”	Close double quote “
‘	‘	Open single quote ’
’	’	Close single quote ‘
	-	sentence dash .
	-	sentence dash .

(Courtesy : Mr. Munawwar Ali,
Resource Centre project, C-DAC, GIST.
Tel. : 020-5694000 extn. 243
e-mail : munawwar@cadacindia.com)

Revision of The Unicode Standard Version 3.0

The Unicode Standard is a character encoding standard which represents the international character code for information processing that includes all major scripts of the world and is the foundation for development of software for world wide use. This standard also defines character properties and algorithms for use and implementations. For Indian languages, Unicode consortium had taken basic inputs from ISCII-1988 document, whereas, at the national level, ISCII-1988 was subsequently revised as ISCII-1991.

The specific problems of sorting/ searching for Indian languages are not fully addressed in the current representation in Unicode, however, within the guidelines for revision of Unicode as defined by the Unicode, DIT had undertaken the task of reviewing the standard through consensus of experts and State Governments. The Ministry has collected inputs for each of the Indian scripts in order to make a single Unified presentation of Indian scripts to the Unicode consortia. Draft code charts for each of the scripts have been prepared along with code details. In the previous issues of the **Vishwabharat@tdil**, the draft for proposed changes in the Unicode standards for Devanagari, Gujarati, Gurmukhi, Oriya, Malayalam, Telugu and Bengali scripts have been published. The draft for proposed changes in the Unicode Standards for Tamil, Kannada and Arabic (Urdu, Sindhi & Kashmiri) are published in this issue of the newsletter.

The Readers are requested to provide their feedback on the proposed modifications as per revised charts. Sanskrit language is not represented so far, therefore, vedic code set has been evolved after discussion with Sanskrit experts.