

## Modern Transcription of Sanskrit

Sanskrit has an inventory of thirteen vowels and thirty-four consonants: all of which can be unambiguously encoded into Devanāgarī (and various other South Asian scripts). As of recent history, Sanskrit is also often Romanized for simpler typesetting and wider scholarship. The dominant Romanization scheme is the International Alphabet of Sanskrit Transcription (IAST), which evolved out of various earlier Romanization schemes (especially the standard developed in 1894 by the Tenth International Congress of Orientalists). Apart from Romanization schemes, Sanskrit does not have a true alphabet (Brāhmīc scripts are abugidas).

## Devanāgarī

Historically, Sanskrit was encoded in whatever was the native script of the scribe; lately, Devanāgarī (/d̪ɛːʋʌŋɡɑːrɪː/, देवनागरी) has become the dominant standard. Devanāgarī and virtually all other indigenous South Asian scripts descend from the ancient Brāhmī, so they share a complex abugidaic orthography (i.e. consonants contain inherent but modifiable vowels) with many possible ligatures for consonant clusters: the ancient phoneticians took every effort to organize and precisely describe all of Sanskrit's sounds in their scripts.

The following charts display Devanāgarī's most basic glyphs and International Phonetic Alphabet (IPA) equivalents.

अ/अ आइई उऊ ऋॠ ए ऐ ओ औ

/ʌ/, /əʔ/, /ɐʔ/ /aː/ /i/ /iː/ /u/ /uː/ /ɪ/ /ɪː/ /ʊ/ /eː/ /ɛi/ /oː/ /ɔu/

क ख ग घ ङ च छ ज झ ञ

/kʌ/ /kʰʌ/ /gʌ/ /gʱʌ/ /ŋʌ/ /tʃʌ/ /tʃʰʌ/ /dʒʌ/ /dʒʱʌ/ /ɟʌ/

ट ठ ड ढ ण त थ द ध न

/ʈʌ/ /ʈʰʌ/ /ɖʌ/ /ɖʱʌ/ /ɳʌ/ /tʌ/ /tʰʌ/ /dʌ/ /dʱʌ/ /nʌ/

प फ ब भ म य र ल व

/pʌ/ /pʰʌ/ /bʌ/ /bʱʌ/ /mʌ/ /jʌ/ /rʌ/ /lʌ/ /vʌ/

श ष स ह ः \*

/ʃʌ/ /ʂʌ/ /ʂʌ/ /ɦʌ/ /h/ nasals

Consonant-vowel modifications are quite regular.

क्क का किकी कु कू कृ कृ कू के कै को कौ

/k/ /kʌ/ /kaː/ /ki/ /kiː/ /ku/ /kuː/ /kɪ/ /kɪː/ /kʊ/ /keː/ /kɛi/ /koː/ /kɔu/

Ligatures are used to describe consonant clusters. Each preceding consonant has its own specific small/half form which attaches to the succeeding consonant. Many exceptions exist among these ligatures and many of the possible ligatures can be especially far-fetched (Charles Philip Brown once protested, “[every] tutor is apt to teach combinations of letters, such as lkha, sba, vpha, yra, khpa; and when we have acquired them, we find we have taken a fruitless trouble, for these never occur...”). The following are some examples.

क तक्त, क्त

/k/ (half)/<sub>ʌ</sub>/ /k<sub>ʌ</sub>/

Examples क्त क्ष ल्ख स्व ह्व त्य स्म्य

/k:ʌ//dʒʌʌ//kʂʌʌ//lk<sup>h</sup>ʌʌ//ʂbʌʌ//hvʌʌ//ptʃʌʌ//smjʌʌ/

Typing and displaying Devanāgarī on computers is somewhat difficult. The script images were generated by L<sub>A</sub>T<sub>E</sub>X using Charles Wikner's `skt` package.

### Roman transcription and transliteration (Romanization) and IAST

(Due to written Devanāgarī's isomorphic correspondence with spoken Sanskrit, any transliteration of Devanāgarī is automatically a *phonetic* transcription of Sanskrit, *vice versa*. See [Transliteration and Transcription](#).)

Romanization schemes (by both transcription/transliteration) vary significantly depending on their unique purposes. British administrators and surveyors aimed to easily reproduce native words and names, so their scheme became a severely Anglophilic transcription (*pundit, thug, suttee, Himalaya*). But savants intended to translate Sanskrit (and possibly compare it with related languages like Avestan), so their schemes were regular and phonetically lossless.

Compare the following schemes. [\(1\)](#) [\(2\)](#) [\(3\)](#) [\(4\)](#)

Devanāgarī	अ आ इ ई उ ऊ ऋ ॠ ए ऐ ओ औ												
N. Halhed (1776)	ě	ā	ěě	ēē	ōō	rěě	rēē	lěě	âê	ī	ō	ôû	
M. Monier-Williams (1864)	a	á	i	í	u	ú	ṛi	ṛí	ḷri	e	ai	o	au
M. Müller (1866)	a	â	i	î	u	û	ṛi	ṛî	ḷi	e	ai	o	au
W. D. Whitney (1879)	a	ā	i	ī	u	ū	ṛ	ṛī	ḷ	e	āi	o	āu
Congress of Orientalists (1894)	a	ā	i	ī	u	ū	ṛ	ṛī	ḷ	e	ai	o	au
Devanāgarī	N. Halhed	M. Monier-Williams	M. Müller	W. D. Whitney	Orientalists								
क, ख	kě, khě	ka, kha	ka, kha	ka, kha	ka, kha								
ग, घ	gě, ghě	ga, gha	ga, gha	ga, gha	ga, gha								

च, छ	chě, chhě	ća, čha	ch, chha	ca, cha	ca, cha
ज, झ	jě, jhě	ja, jha	ja, jha	ja, jha	ja, jha
ट, ठ	tě, thě	ṭa, ṭha	ṭa, ṭha	ṭa, ṭha	ṭa, ṭha
ड, ढ	dě, dhě	ḍa, ḍha	ḍa, ḍha	ḍa, ḍha	ḍa, ḍha
त, थ	tě, thě	ta, tha	ta, tha	ta, tha	ta, tha
द, ध	dě, dhě	da, dha	da, dha	da, dha	da, dha
प, फ	pě, phě	pa, pha	pa, pha	pa, pha	pa, pha
ब, भ	bě, bhě	ba, bha	ba, bha	ba, bha	ba, bha
ङ, ञ, ण	ñġě, gnyě, ně	n·a, ña, ṇa	ña, ña, ṇa	ña, ña, ṇa	ña, ña, ṇa
न, म	ně, mě	na, ma	na, ma	na, ma	na, ma
य, र, ल, व	yě, rě, lě, wě	ya, ra, la, va	ya, ra, la, va	ya, ra, la, va	ya, ra, la, va
श, ष, स, ह	shě, shě, sě, hě	śa, sha, sa, ha	śa, sha, sa, ha	ṣa, ṣa, sa, ha	śa, śa, sa, ha
ः, +	h, ung	ḥ, ṃ	ḥ, ṃ	ḥ, ṃ	ḥ, ṃ/ṃ

Sējġjēñŷyě hrěēdēyūm nēwēñēētūm  
 Yēdwēduntēě wēěbōōdhā  
 ſtēdēlēēkum.  
 Enyēdāêhē wēēlēšēt pērēētāpāt  
 Sējġjēñō drēwētēě nō nēwēñēētūm.

The good Man's Heart is like Butter,  
 The Poets ſay, but herein they are miſtaken:  
 Upon beholding anothers Life expoſed to  
 Calamities,  
 The good Man melts; &mdash but it is not ſo with  
 Butter.

(PDF) [“Specimen of the Reig Beid,” p. \(xxx\).](#) Halhed, Nathaniel. *A Code of Gentoo Laws* (1776).

In 1894, during the Tenth International Congress of Orientalists, a Commission on Transcription met and decided to amalgamate the two major styles of Sanskrit Romanization (those used by the Royal Asiatic Society<sup>(5)</sup> and the *Deutsche Morgenländische Geſellſchaft*).<sup>(6)</sup> Their *Rapport de la Transcription* deſcribed a ſcheme which later evolved into (what ſome unknown bloke coined as): the International Alphabet of Sanskrit Transliteration (IAST). This ſcheme is ſtill the foremoſt academic ſtandard for Sanskrit Romanization.

“Sur ces baſes, l'alphabet de tranſcription ſerait ainſi conſtitué:

a ā i ī u ū ṛ ṝ ḷ e ai o au

k kh g gh ñ

c ch j jh ñ

ṭ ṭh ḍ ḍh ṇ

t th d dh n

p ph b bh m

y r l v ś ṣ h ḷ ṃ ṃ̄

viſarga ḥ

jihvāmūlīya ḥ

upadhmānīya ḥ

quant aux accents, l'udātta ſerait repréſenté par l'aigu

le svarita par le circonflexe ^

le anudātta par le grave `.”

Excerpt, Emile Senart. "Rapport de la sous-commission pour la transcription des alphabets Sanscrit et Pracrits" (1894).

Instead of including English digraphs such as *ch* and *sh*, the Congress opted for *c* (/tʃ/), *ś* (/ʃ/) and *ṣ* (/ʂ/). In fact, all digraphs define aspiration. The Congress took particular effort to differentiate retroflex (*t*, *ṭh*, *ḍ*, *ḍh*) and dental stops (*t*, *th*, *d*, *dh*) and to identify the nasals (*m̐/m̐*: *ṇ*, *ṅ*, *ṇ*, *n*, *m*). For the *anusvāra* (nasals), the Congress officially standardized *m̐*, but also allowed *m̐*. IAST differs from the Congress's scheme in only three ignorable letters.

### Machine readability and displaying Sanskrit on computers

Despite the versatility of the Congress's scheme, its five diacritics break machine readability (a nuisance during typesetting and word processing). One tremendously popular and strictly ASCII (i.e. using only the English language character set) Sanskrit transcription scheme is the Kyoto-Harvard transcription system (KH) by Hideaki Nakatani and Michael Witzel.<sup>(7)</sup> KH overcomes machine readability by differentiating capital letters.

Relatively recently, several  $\text{L}_A\text{T}_E\text{X}$ -compatible Devanāgarī-rendering packages also came into existence, including: Avinash Chopde's Indian Language Transliteration Package (ITRANS);<sup>(8)</sup> and Frans Velthuis's devnag package (devnag's scheme was later adopted by Charles Wikner for the *skt* package,<sup>(9)</sup> which I use to render Devanāgarī on this website).

It is actually also possible for many modern browsers to render Devanāgarī given a Unicode font; however, I avoided the risk of improper rendering by adhering to Devanāgarī images throughout this website.

The following tables compare these methods of displaying machine/human readable Sanskrit.

Congress of Orientalists	Devanāgarī	Kyoto-Harvard	ITRANS	devnag (F. Velthuis)
a, ā	अ, आ	a, A	a, aa/A	a, aa
i, ī	इ, ई	i, I	i, ii/I	i, ii
u, ū	उ, ऊ	u, U	u, uu/U	u, uu
r, ṛ, ṝ	ऋ, ॠ, ॠ̄	R, RR, IR	RRi/R^i, RRI/R^I, LLi/L^i	.r, .r.r, .l
e, ai	ए, ऐ	e, ai	e, ai	e, ai
o, au	ओ, औ	o, au	o, au	o, au
Congress of Orientalists	Devanāgarī	Kyoto-Harvard	ITRANS	devnag
k, kh, g, gh, ṅ	क, ख, ग, घ, ङ	k, kh, g, gh, G	k, kh, g, gh,	k, kh, g, gh,

	□□		~N	"n
c, ch, j, jh, ñ	□□, □□, □□, □□, □□	c, ch, j, jh, J	ch, Ch, j, Jh, ~n	c, ch, j, jh, ~n
ṭ, ṭh, ḍ, ḍh, ṇ	□□, □□, □□, □□, □□	T, Th, D, Dh, N	T, Th, D, Dh, N	.t, .th, .d, .dh, .n
t, th, d, dh, n	□□, □□, □□, □□, □□	t, th, d, dh, n	t, th, d, dh, n	t, td, d, dh, n
p, ph, b, bh, m	□□, □□, □□, □□, □□	p, ph, b, bh, m	p, ph, b, bh, m	p, ph, b, bh, m
y, r, l, v	□□, □□, □□, □□	y, r, l, v	y, r, l, v/w	y, r, l, v
ś, ṣ, s, h	□□, □□, □□, □□	z, S, s, h	sh, Sh, s, h	"s, .s, s
m̐/m̐, ḥ	□, □	M, H	M, H	.m, .h

Devanāgarī keyboards are rare, and only a few programs vindicate the online use of Devanāgarī; most of these programs require transliterated ASCII input. In this way, modern computing obsoletes Devanāgarī. Sanskrit can still be exceptionally well-conveyed via the aforementioned Roman transcription schemes.

### Notes and sources

As noted above, Devanāgarī and other Brāhmī scripts are abugidas and not alphabets. In the past, I tried to invent an alternate alpehabetical writing system for Sanskrit in order for Sanskrit to technologically mature. It was modeled on IAST, with glyphs developed from Brāhmī, Karoṣṭhī, modern Indic scripts, and pseudo-Yavanānī (my rendition of an ancient Greek alphabet). It is severely incomplete, and an outline can be found at: [New Indic Alphabet](#).

Although I was always interested in Sanskrit scripts, I was particularly obsessed with Sanskrit transcription in Winter 2007-08. The following are just some of the vast number of sources I checked out from the UGLI, Hatcher, JSTOR, Google Books, and ILL.

(1) Halhed, Nathaniel. *A Code of Gentoo Laws* (1776).

This is one of the earliest European researches into Indian culture. *Gentoo Laws* was meant to guide the British mercantilists to better interact and trade with the "Gentoo"/Hindus. In the earlier portions of his text, Halhed outlines Devanāgarī and invents this outlandish scheme to transliterate Sanskrit: download [p. XXIV \(PDF\)](#); download the poetic excerpt a ["Specimen of the Reig Beid." p. \(xxxix\) \(PDF\)](#); public domain source, courtesy of University of Michigan Libraries.

(2) Monier-Williams, Monier. *A Practical Grammar of the Sanskrit Grammar* (1864). Download full text: <http://books.google.com/books?id=qRwJAAAAQAAJ>, courtesy of Google Books.

(3) Müller, F. Max. *A Sanskrit Grammar for Beginners, in Devanāgarī and Roman letters throughout* (1st ed. 1866) (1870).

Download full text: <http://books.google.com/books?id=SKYIAAAAQAAJ>, courtesy of Google Books.

(4) Whitney, William Dwight. *A Sanskrit Grammar, including both the classical language, and the older dialects, of Veda and Brahmana* (1879).

Download full text: <http://books.google.com/books?id=RFAIAAAAQAAJ>, courtesy of Google Books.

(5) Journal of the Royal Asiatic Society of Great Britain and Ireland. v. 26 1894. "Report of the Transliteration Committee".

Not to be confused with the Congress of Orientalists. This committee was one of the driving forces in getting the Congress to address the issue of standardized transcription/transliteration. Download full text: [Report \(PDF\)](#), public domain source, courtesy of University of Michigan Libraries.

(6) X<sup>me</sup> Congrès International des Orientalistes, Session de Genève. *Rapport de la Commission de Transcription* (1894).

Download public domain original in French (missing some pages): [Rapport \(PDF\)](#), courtesy of Ames Library, University of Minnesota. English translation: Journal of the Royal Asiatic Society of Great Britain and Ireland. v. 27 1895. Translation of the "Report of the Transliteration Committee, Tenth International Congress of Orientalists, held at Geneva". Download full text: [translation of Report \(PDF\)](#), public domain source, courtesy of University of Michigan Libraries.

It was somewhat difficult to get a hold of this pamphlet... I finally got an Interlibrary Loan from the Ames Library, University of Minnesota. It describes the transliteration conventions of the age and mitigates the Continental style versus the style of the British Isles.

(7) Through correspondence with Dr. H. Nakatani (Tokyo University of Foreign Studies): KH "was created about twenty years ago at Kyoto by the cooperation of Prof. Witzel and me [Nakatani]." It was officially named "Kyoto-Harvard transcription system" by Nakatani (for his then institution in Kyoto).

(8) Chopde, Avinash. ITRANS: Indian Language Transliteration Package. <http://www.aczoom.com/itrans/>.

ITRANS can be used to render many more Indic scripts than just Devanāgarī. It is widely used by devotional websites to upload Sanskrit documents.

(9) Velthuis, Frans. devnag. <http://devnag.sarovar.org/>.

Wikner, Charles. skt. <http://www.ctan.org/tex-archive/language/sanskrit/>.

Consult sktdoc.pdf at: <http://www.ctan.org/get/fonts/ps-type1/sanskrit/sktdoc.pdf>. I find skt to be an exceptionally well-built for Devanāgarī rendering: it is what I use on this website.

Burgess, James. "The Transliteration of Oriental Alphabets". *Bulletin*. Tenth International Congress of Orientalists at Geveva (1894).

Burgess identifies the many benefits to standardizing Oriental alphabets (for Sanskrit and Arabic). Download full text: [Transliteration](#), public domain source, courtesy of University of Michigan Libraries.

Jones, William. "Orthography of Asiatick Words in Roman Letters". *Asiatick Researches* (1801).

This dissertation is the first part of the vol. 1 of *Asiatick Researches*. Sir William Jones lays out his "System of Indian, Arabian, and Persian Letters"; download [System \(PDF\)](#), public domain source, courtesy of University of Michigan Libraries. Full text of *Asiatick Researches* available at Google Books:  
<http://books.google.com/books?id=u50BAAAAMAAJ>.