Technical Manual of Itranslator 99

written by Ulrich Stiehl, 5-Oct-2002
U.Stiehl@t-online.de
http://www.sanskritweb.net/itrans

Revised 24-Apr-2003, 15-Dec-2003, and 29-May-2005

This update reflects changes in the transliteration font implemented in update of Itranslator 99 build 1.3.0.72

Please read Chapter 2 first, as it may serve as a tutorial

This documentation refers to Itranslator 99, available since October 2002. Itranslator 99 has been developed by Omkarananda Ashram Himalayas.

Any suggestions, complaints, etc. may kindly be addressed to:

> Swami Satchidananda
> Omkarananda Ashram Himalayas
> Swami Omkarananda Road
> Muni-ki-reti (Rishikesh)
> Distt. Tehri Garhwal
> P.O. Shivanandanagar - 249 192
> Uttaranchal, INDIA
>
> or by email to:
> itranslator@omkarananda-ashram.org
>
> Please visit our website at:
> http://www.omkarananda-ashram.org

1. Structure and Character Set of ITX File

The **ITX file** is a special ANSI file consisting of portions treated as ITX encoding, which will be converted to Devanagari or to Transliteration, and portions treated as English etc. insertions (comments, translations, etc.), which will remain as is. Example:

candraH	##moon##	>— Itranslator conversion —>	चन्द्रः	moon
ITX	English		Sanskrit	English

1.1. Character Set of ITX Encoding

1.1.1. Control Characters (0-31)

All characters in the range 0–31 are forbidden with the exception of CR + LF (= 13 + 10), treated as the end of a paragraph, whereby a paragraph can be of any length and hence is not restricted to a line of text as seen on the screen when word-wrapping is turned on.

CR + LF (Carriage Return + Linefeed) should always be used as a pair in this sequence. Therefore CR alone without LF or LF + CR (wrong sequence) will not work.

Usage of any other control characters except CR + LF provokes unpredictable results.

1.1.2. ASCII Characters (32–127)

Of the characters in the range 32–127, only those marked **pink and green** on the following chart will be evaluated for the conversion to Devanagari or Transliteration:

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
032		!	"	#	$	%	&	'	(	)	*	+	,	-	.	/
048	0	1	2	3	4	5	6	7	8	9	:	;	<	=	>	?
064	@	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
080	P	Q	R	S	T	U	V	W	X	Y	Z	[	\	]	^	_
096	`	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o
112	p	q	r	s	t	u	v	w	x	y	z	{			}	~	DEL

Pink = letters and figures, **Green** = delimiters and modifiers, **Yellow** = unused, forbidden

The usage of the 31 **yellow** characters !"#$%&()*,/:;<=>?@BEFPQVWXY[]`DEL is forbidden and has unpredictable results. In addition, **all** characters in the 8-bit ANSI range 128-255 are strictly forbidden in ITX encoded portions, which are clean 7-bit ASCII texts.

1.2. Character Set of English etc. Insertions

There are two types of insertions, and hence there are two different character sets.

1.2.1. Insertion of Single Punctuation Marks

In this case, the escape character "\" is prefixed, and **one each** of these marks may follow:

`\! \" \$ \& \( \) \* \+ \, \- \. \/ \: \; \< \= \> \? \@ \[ \] \^ \` \| \~`

which will be converted to the respective punctuation marks with "\" being stripped off:

`! " $ & ( ) * + , - . / : ; < = > ? @ [ ] ^ ` | ~`

The following escape character sequences are **forbidden**: `\' \" \_ \# \% \\ \{ \}`.
Note: `\'` (1 apostrophe), `\"` (2 apostrophes) – **But:** `\"` (quotation mark). See section 2.2.6.

1.2.2. Insertion of English etc. Passages

In this case, the English etc. passage has to be encapsulated by a pair of `## ... ##`, whereby ## serves as a simple toggle switch, so you should take care to use ## ... ## always in pairs, because the insertion mode is not terminated by CR + LF and thus may be of any length.

The character set of an insertion may comprise – besides the two control codes CR + LF – **all** characters in the range 32-255 **except** these six: `#  %  \  DEL-Code 127  {  }`. These forbidden codes have been marked by lowercase underlined "x" in following chart. In addition the codes marked by uppercase underlined "X" should best be avoided too, since these codes are undefined in the 1252 codepage of TrueType fonts:

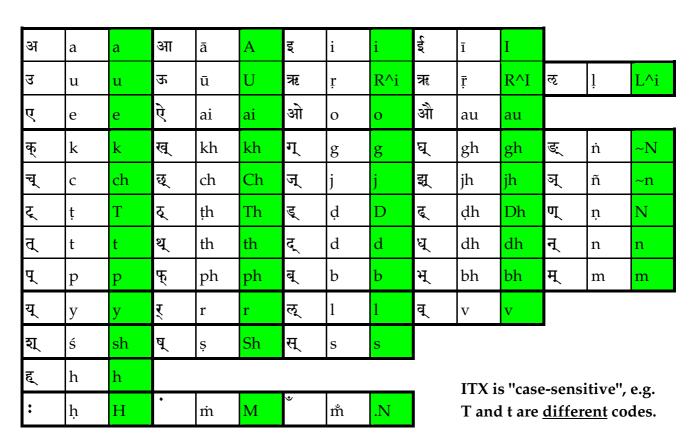| ITX file with Insertions	Itranslator Result

`saMskR^itam`
`##`
` !"x$x&'()*+,-./`
`0123456789:;<=>?`
`@ABCDEFGHIJKLMNO`
`PQRSTUVWXYZ[x]^_`
`` `abcdefghijklmno ``
`pqrstuvwxyzx|x~x`
`€X‚ƒ„…†‡ˆ‰Š‹ŒXŽX`
`X''""•–—˜™š›œXžŸ`
` ¡¢£¤¥¦§¨©ª«¬-®¯`
`°±²³´µ¶·¸¹º»¼½¾¿`
`ÀÁÂÃÄÅÆÇÈÉÊËÌÍÎÏ`
`ÐÑÒÓÔÕÖ×ØÙÚÛÜÝÞß`
`àáâãäåæçèéêëìíîï`
`ðñòóôõö÷øùúûüýþÿ`
`##`
`saMskR^itam`
	संस्कृतम्
	
	` !"x$x&'()*+,-./`
	`0123456789:;<=>?`
	`@ABCDEFGHIJKLMNO`
	`PQRSTUVWXYZ[x]^_`
	`` `abcdefghijklmno ``
	`pqrstuvwxyzx|x~x`
	`€X‚ƒ„…†‡ˆ‰Š‹ŒXŽX`
	`X''""•–—˜™š›œXžŸ`
	` ¡¢£¤¥¦§¨©ª«¬-®¯`
	`°±²³´µ¶·¸¹º»¼½¾¿`
	`ÀÁÂÃÄÅÆÇÈÉÊËÌÍÎÏ`
	`ÐÑÒÓÔÕÖ×ØÙÚÛÜÝÞß`
	`àáâãäåæçèéêëìíîï`
	`ðñòóôõö÷øùúûüýþÿ`
	
	संस्कृतम्

2. Fundamentals of ITX Encoding Scheme

When you start Itranslator by clicking on the program icon, it will open up with the ITRANS Editor window (labeled "New File 1"). You are encouraged to enter some words into this window to acquaint yourself with the encoding scheme. For instance, if you enter **devanAgarI** into the Editor and hit on the function key F2 or click on the Convert button, देवनागरी is shown in the output window. And when you hit on key F6 (= Transliteration) and then on F2, the output window will show **devanāgarī** in its diacritical transliteration.

2.1. Simplified ITX Encoding Scheme for Classical Sanskrit

Itranslator uses the 7-bit ASCII encoding scheme ITX (acronym for Indian Text Exchange). The following chart designed for Classical Sanskrit shows Devanagari (Sanskrit 99 font), Transliteration (URW Palladio IT font) and – in the the green columns – ITX encoding.

अ	a	a	आ	ā	A	इ	i	i	ई	ī	I			
उ	u	u	ऊ	ū	U	ऋ	ṛ	R^i	ॠ	ṝ	R^I	ऌ	ḷ	L^i
ए	e	e	ऐ	ai	ai	ओ	o	o	औ	au	au			
क्	k	k	ख्	kh	kh	ग्	g	g	घ्	gh	gh	ङ्	ṅ	~N
च्	c	ch	छ्	ch	Ch	ज्	j	j	झ्	jh	jh	ञ्	ñ	~n
ट्	ṭ	T	ठ्	ṭh	Th	ड्	ḍ	D	ढ्	ḍh	Dh	ण्	ṇ	N
त्	t	t	थ्	th	th	द्	d	d	ध्	dh	dh	न्	n	n
प्	p	p	फ्	ph	ph	ब्	b	b	भ्	bh	bh	म्	m	m
य्	y	y	र्	r	r	ल्	l	l	व्	v	v			
श्	ś	sh	ष्	ṣ	Sh	स्	s	s						
ह्	h	h												
ः	ḥ	H	ं	ṁ	M	ँ	m̐	.N						

**ITX is "case-sensitive", e.g.
T and t are underlined different codes.**

The green boxes show those ITX codes which are compatible with ITRANS 5.1 since 1997. To avoid confusion for beginners, only **one** ITX code is shown for each single character. For those who want to encode Classical Sanskrit only, no further codes need to be learnt.

Itranslator supports several alternative codes, e.g. **aa** instead of **A** for encoding आ. And it can be used not only for Classical Sanskrit, but for encoding accented Vedic Sanskrit too. With some restrictions, Itranslator can also be used for encoding those modern Indian languages which are based on the Devanagari alphabet (e.g. Hindi, Marathi, Nepali).

2.2. Complete Description of the ITX Codes with Examples

(For Pronunciation of the Sanskrit Alphabet see http://www.sanskritweb.net/deutsch/ipa_sans.pdf)

2.2.1. Sanskrit Punctuation Marks *(for English Punctuation Marks see 1.2.1.)*

ITX	D.	T.	ITX – Devanagari –Transliteration – English Translation
-	−	-	kuru-kShetram – कुरु-क्षेत्रम् – kuru-kṣetram – *Kuru-country*
| or .	।	|	pR^icChAmi | ko .aham || – पृच्छामि । को ऽहम् ॥
|| or ..	॥	||[1]	pṛcchāmi | ko 'ham || – *I ask: Who am I?*
.a	ऽ	'	Avagraha (apostrophe, i.e. elision of a), as above in को ऽहम्

2.2.2. Sanskrit Vowels

ITX	D.	T.	ITX – Devanagari –Transliteration – English Translation
a	अ	a	arthaH – अर्थः – arthaḥ – *money*
A, aa	आ	ā	AnandaH or aanandaH – आनन्दः – ānandaḥ – *bliss*
i	इ	i	iti – इति – iti – *thus*
I, ii	ई	ī	IkShate or iikShate – ईक्षते – īkṣate – *he/she watches (present tense)*
u	उ	u	uDupaH – उडुपः – uḍupaḥ – *boat*
U, uu	ऊ	ū	UcuH or uucuH – ऊचुः – ūcuḥ – *they said (perfect tense)*
R^i, RRi	ऋ	ṛ	R^itvik or RRitvik – ऋत्विक् – ṛtvik – *priest*
R^I, RRI	ॠ	r̥̄ [1]	svasR^INAm or svasRRINAm – स्वसॄणाम् – svasṝṇām – *of sisters*
L^i, LLi	ऌ	ḷ [2]	kL^iptam or kLLiptam – कॢप्तम् – kḷptam – *prepared*
L^I, LLI	ॡ	ḹ [3]	Not used in Sanskrit words
e	ए	e	ekadA – एकदा – ekadā – *once upon a time*
ai	ऐ	ai	aikShata – ऐक्षत – aikṣata – *he/she watched (imperfect tense)*
o	ओ	o	odanaH – ओदनः – odanaḥ – *boiled rice*
au	औ	au	auShadham – औषधम् – auṣadham – *herbal medicine*
OM, AUM	ॐ	om	This special ligature ॐ is used instead of om ओम् or aum औम्

[1] ṝ never occurs at the beginning of a word. Therefore the full vowel ॠ is never used.

[2] ḷ only occurs in some forms of the root klp and hence the full vowel ऌ is never used. Itranslator transliterates this extremely rare ऌ to ḷ since the new build of April 2003.

[3] ḹ (long version of ḷ) has been invented by ancient Sanskrit grammarians for reasons of symmetry, since the other simple vowels exist as pairs of short and long vowels.

2.2.3. Sanskrit Consonants

The Devanagari consonants in column "D." are depicted without the Virama (্), as the Itranslator conversion routine evaluates by itself, where to add the Virama and where not.

ITX	D.	T.	ITX – Devanagari –Transliteration – English Translation
k	क	k	kusumam – कुसुमम् – kusumam – *flower*
kh	ख	kh	khalu – खलु – khalu – *indeed*
g	ग	g	gadyam – गद्यम् – gadyam – *prose*
gh	घ	gh	ghoShaH – घोषः – ghoṣaḥ – *noise*
~N, N^	ङ	ṅ	a~NgulI or aN^gulI – अङ्गुली – aṅgulī – *finger*
c, ch	च	c [1]	cAraH or chAraH – चारः – cāraḥ – *spy*
Ch, chh	छ	ch [1]	ChAttraH or chhAttraH – छात्रः – chāttraḥ – *pupil*
j	ज	j	jihvA – जिह्वा – jihvā – *tongue*
jh	झ	jh [2]	jhaTiti – झटिति – jhaṭiti – *suddenly*
~n, JN	ञ	ñ	ra~njayati or raJNjayati – रञ्जयति – rañjayati – *he/she dyes in red*
T	ट	ṭ	kaTaH – कटः – kaṭaḥ – *straw mat*
Th	ठ	ṭh	kaNThaH – कण्ठः – kaṇṭhaḥ – *neck*
D	ड	ḍ	daNDaH – दण्डः – daṇḍaḥ – *stick*
Dh	ढ	ḍh	leDhum – लेढुम् – leḍhum – *to lick (infinitive)*
N	ण	ṇ	varNaH – वर्णः – varṇaḥ – *color*
t	त	t	talam – तलम् – talam – *floor*
th	थ	th	sthAnam – स्थानम् – sthānam – *place*
d	द	d	sarvadA – सर्वदा – sarvadā – *always*
dh	ध	dh	dhanam – धनम् – dhanam – *wealth*
n	न	n	nArI – नारी – nārī – *woman*
p	प	p	pashyati – पश्यति – paśyati – *he/she sees*
ph	फ	ph	phalam – फलम् – phalam – *fruit*
b	ब	b	bAlaH – बालः – bālaḥ – *boy*
bh	भ	bh	bhadraM te – भद्रं ते – bhadraṁ te – *good luck!*
m	म	m	mitram – मित्रम् – mitram – *friend*

[1] To achieve compatibility with ITRANS (see 5.3.), encode **ch** for च and **Ch** (or **chh**) for छ.

[2] jh, regularly used in Marathi, occurs in Sanskrit only in a few onomatopoetic words.

Sanskrit Consonants Continued

ITX	D.	T.	ITX – Devanagari –Transliteration – English Translation
y	य	y	yogaH – योगः – yogaḥ – *Yoga*
r	र	r	rAjA – राजा – rājā – *king*
l	ल	l	likhati – लिखति – likhati – *he/she writes*
v, w	व	v	vanam or wanam – वनम् – vanam – *forest*
sh	श	ś	shukaH – शुकः – śukaḥ – *parrot*
S, Sh, shh, Z	ष	ṣ [1]	SaT or ShaT or shhaT or ZaT – षट् – ṣaṭ – *six*
s	स	s	svapnaH – स्वप्नः – svapnaḥ – *sleep*
h	ह	h	hantA – हन्ता – hantā – *killer*

[1] To achieve compatibility with ITRANS (see 5.3.) encode **Sh** (or **shh**) for ष.

2.2.4. Additional Characters for Classical and Vedic Sanskrit

ITX	D.	T.	ITX – Devanagari –Transliteration – English Translation
H	ः	ḥ [1]	ashvaH – अश्वः – aśvaḥ – *horse*
M, .n, .m	ं	ṁ [2]	saMshayaH or sa.nsh... or sa.msh... – संशयः – saṁśayaḥ – *doubt*
.N	ँ	m̐l [3]	tA.NllokAn – ताँल्लोकान् – tām̐llokān – *these people (accusative pl.)*
{\m+}	ꣳ	m̐ [4]	ida{\m+} or ida.N – इदꣳ or इदँ – idam̐ (for idaṁ) – *this (Vedic)*
L	ळ	ḻ [5]	agnimILe – अग्निमीळे – agnimīḻe – *I praise (īḻe) the fire god (agnim)*
kS, kSh, x	क्ष	kṣ [6]	akSi or akShi or axi – अक्षि – akṣi – *eye.* **Do not encode "ksh" !!!**
j~n, GY, dny	ज्ञ	jñ [6]	Aj~nA or AGYA or AdnyA – आज्ञा – ājñā – *order*

[1] Visarga: Replacement for s or r at the end of a word or syllable as per sandhi rules.

[2] Anusvara: Shortcut character for m and other nasals at the end of a word or syllable.

[3] Sanskrit Anunasika: m̐l. In Sanskrit texts, it is only used as sandhi, if "l" follows "n".

[4] Vedic Anunasika: m̐: ँ or ꣳ. In Vedic texts often used before sibilants, semi-vowels and h. Note: ꣳ is preferred, if ँ cannot be fitted on top for lack of space. Compare सर्वाँ with सर्वꣳ.

[5] Vedic Sanskrit: If ḍ is preceded **and** followed by a **vowel**, ḻ may be used instead of ḍ.

[6] These are the only two ligatures, that – for historial reasons – have **additional** ITX codes (**x** in addition to usual encoding kS or kSh, **GY** or **dny** in addition to usual encoding j~n). All other ligatures are encoded by combining the usual ITX codes for simple characters, which are converted by Itranslator to the respective Devanagari ligatures automatically.

2.2.5. Additional Characters for Modern Indian Languages

ITX	D.	T.	ITX – Devanagari –Transliteration – English Translation
q	क़	q [1]	qImata – क़ीमत – qīmata – *price*
K	ख़	k͟h [1]	KatarA – ख़तरा – k͟hatarā – *danger*
G	ग़	ġ [1]	Galata – ग़लत – ġalata – *wrong*
z, J	ज़	z [1]	zamIna or JamIna – ज़मीन – zamīna – *earth*
f	फ़	f [1]	sirfa – सिर्फ़ – sirfa – *only*
.D	ड़	ṛ [2]	ba.DA – बड़ा – baṛā – *big* (Same diacritic as for vowel ṛ = ऋ !!!)
.Dh	ढ़	ṛh [2]	ba.DhAnA – बढ़ाना – baṛhānā – *to increase*
.c	ॅ	ô [3]	jA.cb– जॉब – jôba – *job*; kA.cleja – कॉलेज – kôleja – *college*
L	ळ	ḻ [4]	kALa – काळ – kāḻa – *time, Death (fig.)* **Do not encode "ld" !!!**
R	ऱ	ř [5]	taRhA – तऱ्हा – tařhā – *variety*

[1] Nukta (underdot): Used in Hindi for words of Persian and Arabic origin as well as for other foreign words to modify the original pronunciation of Devanagari characters.

[2] Nukta is here used in Hindi to build diacritical Devanagari characters for retroflex flap (transliteration ṛ) and for aspirated retroflex flap (transliteration ṛh).

[3] Ardhacandra: Anglizising of pronunciation of vowels of English words used in Hindi.

[4] Retroflex l (transliteration ḻ) in Marathi. Compare the letter ळ (ḻ) in Vedic Sanskrit. Note: If you encode "ld", you will get ल्द, which is used e.g. in word galda (see Rig-Veda 8-1-20).

[5] This special ř-character ऱ is used in Marathi in combination with the letters y and h.

2.2.6. Vedic Accents

Vedic accents are encoded by \' \" _ – single/double Udatta and Anudatta. Example:

ITX	puru\'Sha e_vedaM sarvaM_ yadbhU_taM yacca_ bhavya\'m
D.	पुरुष एवेदं सर्वं यद्भूतं यच्च भव्यम्
T.	puruṣa evedaṁ sarvaṁ yadbhūtaṁ yacca bhavyam
E.	*The primordial being is verily all this that has been and that will be (Rig-Veda, 10-90-2)*

Note: \' = \ + apostrophe (ASCII 39), \" = \ + **two** apostrophes, _ = \ + underscore

2.2.7. Figures

The Arabic figures 0123456789 are converted to the respective Indian figures ०१२३४५६७८९ . Decimal figures are encoded with "\," and "\.", e.g. 1\,500\.99 results in १,५००.९९ (1,500.99).

3. Structure of Devanagari Font Sanskrit 99

3.1. Character Codes of Sanskrit 99 Font versus Latin 1 Fonts (ANSI 1252)

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
032		॒	घ	इ	ई	उ	ॢ	ङ	च	फ	द्य	॑	।	–	॥	
032		!	"	#	$	%	&	'	(	)	*	+	,	-	.	/
048	०	१	२	३	४	५	६	७	८	९	९	ष	॰	ऽ	ः	
048	0	1	2	3	4	5	6	7	8	9	:	;	<	=	>	?
064	ए	अ	ह	च	छ	ॎ		ढ	र	झ	ी	ज	व	ल	भ	ख
064	@	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
080	ट	ठ	°	स	त	ॢ	ड	थ	द	ट	इ	ण	ऋ	क्ष	ऊ	भ
080	P	Q	R	S	T	U	V	W	X	Y	Z	[	\	]	^	_
096	ॐ	ा	ब	च	द्	ॎ	ड	ग	ह	ि	ज	क	ल	म	न	ख
096	`	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o
112	प	ट	र	स	त	ॣ	व	थ	घ	य	श	ण	ज्ञ	ज्ञ	ॕ	■
112	p	q	r	s	t	u	v	w	x	y	z	{	|	}	~	
128	■	■	ॖ	ॗ	कृ	ॖ	ह	ँ	.	॒	क्क	क्क	ह्न	■	■	■
128	€		‚	ƒ	„	…	†	‡	^	‰	Š	‹	Œ		Ž	
144	■	ळ	ह्ल	ङ्क	ट्ट	≈	.	·	द्द	द्य		॓	॒	■	■	ह्ल
144		'	'	"	"	•	–	—	~	™	š	›	œ		ž	ÿ
160		ॏ	ग्र	ॅ	ल	लृ	ळ	ऋ	६	६	७	८	क्रू	–	त्त	८
160		¡	¢	£	¤	¥	¦	§	¨	©	ª	«	¬	-	®	¯
176	६	रू	रू	रू	रू	ख	म्र	·	६	घ्र	घ्र	ङ्क	ङ्ग	च्च	ज्ज	ञ्ञ
176	°	±	²	³	´	µ	¶	·	¸	¹	º	»	¼	½	¾	¿
192	छ	ङ	ऋ	ऋ	त्त	र	ह्न	त्र	३	भ	दु	दू	द्र	द्र	द्र	द्ध
192	À	Á	Â	Ã	Ä	Å	Æ	Ç	È	É	Ê	Ë	Ì	Í	Î	Ï
208	द्र	द्ध	ध्र	द्ध	द्र	द्य	द्र	भ्र	भ्र	त्र	त्र	६	त्र	त्त	६	भ
208	Ð	Ñ	Ò	Ó	Ô	Õ	Ö	×	Ø	Ù	Ú	Û	Ü	Ý	Þ	ß
224	प्र	फ	ञ	न्न	ब	भ्र	भ्र	म्र	म्र	रु	रू	न्न	व	श्र	श्र	श्र
224	à	á	â	ã	ä	å	æ	ç	è	é	ê	ë	ì	í	î	ï
240	श्ल	श्र	ष्ट	ष्ठ	स्त	स्र	स्र	हु	हृ	ह्र	४	ह्न	ह्न	ह्य	ह्न	ह
240	ð	ñ	ò	ó	ô	õ	ö	÷	ø	ù	ú	û	ü	ý	þ	ÿ

White rows: Sanskrit 99 font, **Green rows:** Revised ANSI 1252 (incl. € , Ž , ž): Courier New
Black: unused, reserved

3.2. Improvements of New Sanskrit 99 over Old Sanskrit 98

In order to achieve compatibility with all word processing programs and printers, the font has been made compatible with the **original** specifications for TTF and PS Type 1 fonts:

1. The codes 127, 128-129, 141-144, 157-158 (not used in original TTF and PS fonts) have been left empty, because characters in these slots are often skipped by printer drivers.

2. Printer's Specials: Two spaces: hard space (code 32) and no-break space (code 160). Two hyphens: hard hyphen (code 45) and soft hyphen (code 173). Middot (code 183).

3.3. Classification of Glyphs of Sanskrit 99

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
032		◌̑	घ	इ	ई	उ	ॽ	ङ	च	फ	द्य	◌̂	।	–	॥	
048	०	१	२	३	४	५	६	७	८	९	६	ष	ॱ	ऽ	ः	
064	ए	अ	ड	च	छ	◌ॕ	ढ	ग	झ	◌ी	ज	व	ल	फ	र	रु
080	द	ठ	◌ॕ	स	त	◌ॖ	ड	थ	ध	ट	इ	ण	ऋ	क्ष	ऊ	∦
096	ॐ	ा	ब	च	द्	◌ॅ	ड	ग	ह	ि	ज	क	ल	म	न	ख
112	प	ट	र	स	त	◌ॗ	व	थ	ध	य	श	ण	ज	ज्ञ	◌ॕ	
128			◌ॖ	◌ॕ	क्रू	◌ॖ	दृ	◌ॆ	·	◌̣	क्र	क्र	ह्र				
144		ल्ल	ह्ल	ङ्क	इृ	≈	◌.	◌̇	द्द	द्य	◌ॖ	◌ॕ	◌̑			ह्ल
160		◌॓	ग्र	◌̆	ळ	ॡ	ळ	ॠ	◌॒	◌॒	◌॒	◌॒	क्रू	–	त्त	◌॒
176	◌॒	क्र	क्र	क्र	क्त	ख	म्र	·	द	घ्र	घ्र	ङ्क	ङ्न	च	ज	ज्र
192	छ	ञ	श्च	ञ्ज	त्त	र	ल	त्र	ड	द	दू	द्र	द्र	द्र	द्र	द्र
208	द्र	द्र	ध्र	द्ध	द्र	ध्य	द्र	ध्र	ध्र	न्त्र	न्त्र	ड	न्त	स	६	म
224	प्र	फ	ज	ब्र	ब्र	भ्र	भ्र	म्र	म्र	रु	रू	न्र	व्र	श्व	श्र	श्र
240	श्ल	श्च	ष्ट	ष्ठ	स्त्र	स्त्र	स्त्र	हु	हू	हृ	ढ़	ह्र	ह्म	ह्य	ह्र	ह

Pink: Complete glyphs (complete simple characters and complete precomposed ligatures)

Green: Half-glyphs (left halves of simple letters and ligatures – *Exception:* च = right half)

Yellow: Zero-width glyphs (matras, accents, virama) inserted **after** the respective glyph, e.g. क + ◌ु = कु – *Exception:* Nukta is inserted **before** the respective glyph, e.g.: ◌़ + क = क़

White: Printer's Specials (spaces, hyphens and centered period = middle dot).

Black: unused, reserved

3.4. Selective List of Ligatures of Sanskrit 99 in Large Print for Beginners

This list of ligatures, drawn from my book "Sanskrit-Kompendium", has been updated to reflect changes made in the font Sanskrit 99 as opposed to Sanskrit 98: Due to lack of 1252 code page space, a few ligatures, e.g. **ṭhṭh** and **phra** used in Hindi and **khna** and **thna** very rarely used in Sanskrit had to be removed. In this way it was possible to re-standardize the code page to achieve compatibility with all word processors and printer drivers.

Not all ligatures listed occur in **Sanskrit**. For detailed information on ligatures that really occur in Vedic and Classical Sanskrit, consult my book "Conjunct Consonants in Sanskrit".

क्क	kka	क्ष्व	kṣva	ङ्क	ṅka[+]	–	cra
क्ख	kkha	क्स	ksa	ङ्क्त	ṅkta	च्व	cva
क्च	kca	ख्न	khna	ङ्क्त्य	ṅktya	छ्म	chma
क्ण	kṇa	ख्य	khya[++]	ङ्क्य	ṅkya	छ्य	chya
क्त	kta[++]	ख्र	khra	ङ्क्ष	ṅkṣa	छ्र	chra
क्त्य	ktya	ख्व	khva	ङ्क्ष्व	ṅkṣva	छ्व	chva
क्त्र	ktra	ग्द	gda	ङ्ख	ṅkha	ज्ज	jja[+]
क्त्र्य	ktrya	ग्ध	gdha	ङ्ख्य	ṅkhya	ज्ज्ञ	jjña
क्त्व	ktva[+]	ग्ध्व	gdhva	ङ्ग	ṅga[++]	ज्ज्व	jjva
क्थ	ktha	ग्न	gna[++]	ङ्ग्य	ṅgya	ज्झ	jjha
क्थ्न	kthna	ग्न्य	gnya	ङ्ग्र	ṅgra	ज्ञ	jña[++]
क्न	kna	ग्भ	gbha	ङ्घ	ṅgha	ज्ञ्य	jñya
क्न्य	knya	ग्भ्य	gbhya	ङ्घ्य	ṅghya	ज्म	jma
क्म	kma	ग्म	gma	ङ्घ्र	ṅghra	ज्य	jya[++]
क्य	kya[++]	ग्य	gya	ङ्ङ	ṅṅa	ज्र	jra
क्र	kra[+++]	ग्र	gra[++]	ङ्न	ṅna	ज्व	jva
क्र्य	krya	ग्र्य	grya	ङ्म	ṅma	ञ्च	ñca[++]
क्ल	kla	ग्ल	gla	ङ्य	ṅya	ञ्च्म	ñcma
क्व	kva	ग्व	gva	च्च	cca[++]	ञ्च्य	ñcya
क्व्य	kvya	घ्न	ghna[+]	च्छ	ccha[++]	ञ्छ	ñcha
क्ष	kṣa[+++]	घ्न्य	ghnya	च्छ्र	cchra[+]	ञ्ज	ñja[++]
क्ष्ण	kṣṇa	घ्म	ghma	च्छ्व	cchva	ञ्ज्य	ñjya
क्ष्म	kṣma	घ्य	ghya	च्ञ	cña	ञ्ञ	ñña
क्ष्म्य	kṣmya	घ्र	ghra[+]	च्म	cma	ञ्श	ñśa[+]
क्ष्य	kṣya[++]	घ्ल	ghla	च्य	cya[++]	ट्क	ṭka

ट्ट	ṭṭa	ण्ढ	ṇḍha	थ्व	thva	ध्र्य	dhrya
ट्ठ	ṭṭha	ण्ण	ṇṇa	द्ग	dga	ध्व	dhva+
ट्प	ṭpa	प्न	ṇna	द्ग्य	dgya	न्क	nka+
ट्म	ṭma	ण्म	ṇma	द्ग्र	dgra	न्त	nta+++
ट्य	ṭya	ण्य	ṇya++	द्घ	dgha	न्त्य	ntya+
ट्र	ṭra	ण्व	ṇva	द्घ्र	dghra	न्त्र	ntra+
ट्व	ṭva	त्क	tka++	द्द	dda+	न्त्र्य	ntrya
ट्ष	ṭsa	त्क्र	tkra	द्द्य	ddya	न्त्व	ntva
ट्स	ṭsa	त्त	tta+++	द्द्र	ddra	न्त्स	ntsa
ठ्ठ	ṭhṭha	त्त्य	ttya	द्द्व	ddva	न्थ	ntha
ठ्म	ṭhma	त्त्र	ttra	द्ध	ddha+++	न्द	nda++
ठ्य	ṭhya	त्त्व	ttva++	द्ध्न	ddhna	न्द्ध	nddha
ठ्र	ṭhra	त्थ	ttha+	द्ध्य	ddhya	न्द्र	ndra++
ड्ग	ḍga	त्न	tna+	द्ध्व	ddhva	न्ध	ndha++
ड्ग्य	ḍgya	त्न्य	tnya	द्न	dna	न्ध्य	ndhya
ड्घ	ḍgha	त्प	tpa++	द्ब	dba	न्ध्र	ndhra
ड्घ्र	ḍghra	त्र	tpra+	द्ब्र	dbra	न्न	nna+++
ड्ड	ḍḍa	त्फ	tpha	द्भ	dbha++	न्य	nnya
ड्ढ	ḍḍha	त्म	tma++	द्भ्र	dbhra	न्प	npa++
ड्भ	ḍbha	त्म्य	tmya	द्भ्य	dbhya	न्प्र	npra+
ड्म	ḍma	त्य	tya+++	द्म	dma	न्फ	npha
ड्य	ḍya	त्र	tra+++	द्य	dya+++	न्म	nma++
ड्र	ḍra	त्र्य	trya	द्र	dra+++	न्य	nya+++
ढ्ण	ḍhṇa	त्व	tva+++	द्र्य	drya	न्र	nra+
ढ्य	ḍhya	त्स	tsa++	द्व	dva+++	न्व	nva++
ढ्र	ḍhra	त्स्न	tsna	द्व्य	dvya	न्स	nsa++
ढ्व	ḍhva	त्स्न्य	tsnya	द्व्र	dvra	न्ह	nha
ण्ट	ṇta	त्स्म्य	tsmya	ध्न	dhna	प्त	pta++
ण्ठ	ṇtha	त्स्य	tsya+	ध्न्य	dhnya	प्त्य	ptya
ण्ड	ṇḍa++	त्स्व	tsva	ध्म	dhma	प्त्र	ptra
ण्ड्य	ṇḍya	थ्न	thna	ध्य	dhya++	प्त्र्य	ptrya
ण्ड्र	ṇḍra	थ्य	thya	ध्र	dhra	प्न	pna+

प्प	ppa	भ्व	bhva	व्व	vva	स्क	ska[+]
प्फ	ppha	म्न	mna[+]	श्क	śka	स्क्र	skra
प्म	pma	म्प	mpa	श्च	śca[+++]	स्ख	skha
प्य	pya[++]	म्प्र	mpra	श्च्य	ścya	स्ट	sṭa
प्र	pra[+++]	म्ब	mba[+]	श्न	śna	स्त	sta[+++]
प्ल	pla	म्ब्य	mbya	श्म	śma	स्त्य	stya
प्व	pva	म्भ	mbha	श्य	śya[++]	स्त्र	stra[++]
प्व्य	pvya	म्म	mma	श्र	śra[+++]	स्त्व	stva[++]
प्स	psa[+]	म्य	mya[++]	श्र्य	śrya	स्थ	stha[++]
प्स्व	psva	म्र	mra	श्ल	śla	स्थ्य	sthya
फ्म	phma	म्ल	mla	श्व	śva[++]	स्न	sna
–	phra	म्व	mva	श्व्य	śvya	स्न्य	snya
फ्य	phya	म्स	msa	श्श	śśa	स्प	spa[++]
ब्घ	bgha	म्ह	mha	ष्क	ṣka[+]	स्फ	spha
ब्ज	bja	य्य	yya	ष्क्र	ṣkra	स्म	sma[++]
ब्ट	bṭa	य्व	yva	ष्ट	ṣṭa[++]	स्म्य	smya
ब्त	bta	ल्क	lka	ष्ट्य	ṣṭya[++]	स्य	sya[+++]
ब्द	bda	ल्ग	lga	ष्ट्र	ṣṭra[++]	स्र	sra[++]
ब्ध	bdha	ल्द	lda	ष्ट्र्य	ṣṭrya	स्ल	sla
ब्ध्व	bdhva	ल्न	lna	ष्ट्व	ṣṭva[++]	स्व	sva[++]
ब्न	bna	ल्प	lpa[+]	ष्ठ	ṣṭha[++]	स्स	ssa
ब्ब	bba	ल्म	lma	ष्ठ्य	ṣṭhya	ह्ण	hṇa
ब्भ	bbha	ल्य	lya[++]	ष्ण	ṣṇa[++]	ह्न	hna
ब्भ्य	bbhya	ल्ल	lla[+]	ष्ण्य	ṣṇya	ह्म	hma[++]
ब्य	bya	ल्व	lva	ष्प	ṣpa	ह्य	hya[++]
ब्र	bra[++]	ल्ह	lha	ष्प्र	ṣpra	ह्र	hra
ब्व	bva	व्न	vna	ष्म	ṣma[++]	ह्र्य	hrya
भ्न	bhna	व्य	vya[+++]	ष्य	ṣya[++]	ह्ल	hla
भ्य	bhya[++]	व्र	vra[++]	ष्व	ṣva[++]	ह्व	hva
भ्र	bhra[++]	व्ल	vla	ष्ष	ṣṣa	ह्व्य	hvya

[+++] p > 5 %; [++] p > 1 %; [+] p > 0.5 %; others p < 0.5 %. Probability p = 1 % means that the ligature occurs **once** in every 100 lines (=16-syllable half-verses) of the Mahabharata, e.g. ligature ग्म occurs in 577 lines out of 158,484 lines of the Mahabharata, hence p = 0.364%.

4. Conversion of Transliterated Texts to ITX

The font **URW Palladio IT** (= Itranslator Transliteration) has the following encoding:

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
032		!	"	#	$	%	&	'	(	)	*	+	,	-	.	/
048	0	1	2	3	4	5	6	7	8	9	:	;	<	=	>	?
064	@	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
080	P	Q	R	S	T	U	V	W	X	Y	Z	[	\	]	^	_
096	`	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o
112	p	q	r	s	t	u	v	w	x	y	z	{			}	~	`
128	`	`	,	ƒ	„	…	†	‡	^	‰	Š	‹	Œ	`		`
144	`	'	'	"	"	•	–	—	~	™	š	›	œ	`		Ÿ
160		¡	¢	£	¤	¥	¦	§	¨	©	ª	«	¬	-	®	¯
176	°	±	²	³	´	µ	¶	·	¸	¹	º	»	¼	½	¾	¿
192	Ṁ	Á	Â	Ã	Ā	Ṛ	Æ	Ś	Ṝ	Ī	Ê	Ṇ	Ṅ	Í	Î	Ñ
208	Ṃ	Ṣ	Ḍ	Ó	Ô	Ř	Ṭ	'	Ḷ	Ḥ	Ġ	Kh	Ū	Ḹ	_	Ḻ
224	ṁ	á	â	ã	ā	ṛ	æ	ś	ṝ	ī	ê	ṇ	ṅ	í	î	ñ
240	ṃ	ṣ	ḍ	ó	ô	ř	ṭ	"	ḷ	ḥ	ġ	kh	ū	ḹ	KH	ḻ

Green = Sanskrit; **Pink** = Vedic accents; **Yellow** = *additional* diacritics for Hindi/Marathi

This font was reworked in April 2003 to include *additional* diacritics for Hindi/Marathi. Note: Itranslator does not use uppercase diacritics. It converts to lowercase only.

This font is used for **output**, if you select "Transliteration" in Itranslator. However, most Western indologists do not use 7-bit ITX encoding for **input**. They enter Sanskrit texts with any of various 8-bit transliteration fonts such as URW Palladio IT. In this case you will need the Code Converter **Koko** converting 8-bit encoded transliterated "clean textfiles" (without English etc. insertions) into ITX encoded textfiles. These ITX files can thereafter be converted by the Itranslator program to Sanskrit 99 files.

Input	Program	Output
Transliterated file	>——— Koko ———>	a) ITX encoded file, **or** b) Transliterated file *(with other encoding)*
ITX encoded file	>—— Itranslator ——>	a) Devanagari file *(Sanskrit 99)*, **or** b) Transliterated file *(URW Palladio IT)*

"Koko" is available to programmers only. See http://www.sanskritweb.net/koko

5. Miscellaneous Details on Various Topics

This chapter deals with a variety of different additional topics.

5.1. The Ligature Problem

Itranslator allows the user to build (a) ligatures, which really occur in the ancient texts, and (b) fancy ligatures, which may have never been used in old Sanskrit manuscripts. Example: If you encode "**mbha**", Itranslator generates म्भ, which is a correct conversion of a ligature that *does occur* in Sanskrit texts. However, if you encode "**bmha**", Itranslator produces ब्म्ह, a typographically pleasant ligature that is phonetically impossible and hence obviously *does not occur* in extant Sanskrit texts. Itranslator, therefore, can generate many ligatures looking nice from the typographical point of view, but which are phonetically impossible and hence are unlikely to be ever discovered. On the other hand, Itranslator has to generate many typographically unpleasant pseudo-ligatures for very rare ligatures. Example: "ṅkhy" (~Nkhya) is built as the pseudo-ligature ङ्ख्य (i.e. ङ् with virama + ख्य), but you should bear in mind, that this ligature occurs *not even once* in the Mahabharata, though – theoretically – it *could* occur, e.g. in प्राङ्ख्याति (*"the one looking east is revealing"*).

Readers who are interested in ligature research may consult my comprehensive book "Conjunct consonants in Sanskrit" documenting almost all conjunct consonants ever used in Vedic and Classical Sanskrit by quotations from a large variety of original texts.

5.2. Ligature Inhibitor "{}" and Virama Enforcer ".h"

Occasionally, it may be necessary to prevent Itranslator from using build-in ligatures by defining alternative glyph combinations. Likewise, it may be necessary to enforce virama.

5.2.1. Ligature Inhibitor "{}"

The ITX code {} is used as *"ligature inhibitor"*, *"half-glyph enforcer"* and *"diphthong-divider"*

"Ligature inhibitor" means that Itranslator is guided not to resort to the ordinary ligature building routine, but rather to build the ligature in the way defined by the inserted code {}.

"Half-glyph enforcer" means that Itranslator will generate a half-glyph for the combination "consonant" + "code {}", *provided* a half-glyph (e.g क्) is available for the consonant (e.g. क). If no half-glyph is available in the Sanskrit 99 font (e.g. छ has no half-glyph), Itranslator uses consonant + virama, so that you need not know, whether a half-glyph exists or not.

As there are so many extremely rare ligatures, the ITX encoding is ambiguous in certain circumstances. A good example is the ligature "kṣ" (क्ष). As this ligature belongs to the ten most often used ligatures, there are in use two *contradicting* ITX codes "ksha" (क्शा) and "kSha" (क्ष), which were converted by Itranslator both to the ligature क्ष, unless the other क्शा was enforced by inserting the half-glyph-enforcer: "k{}sha".

When ITX encoding was devised about ten years ago, is was not known, that there exist several rare ligatures beginning with "kś". Therefore for the new build 1.3.0.58 (April 2003) of Itranslator it was decided to no longer permit the ambiguous ITX code "ksh" for "kṣ" (क्ष) to allow for encoding of ligatures beginning with "kś" (क्शा) by encoding "ksh" <u>without</u> {}, e.g. "kśm", "kśr", "kśl", and "kśv". So now simply encode "kshm" for क्श्म , "kshl" for क्श्ल , etc.

The ITX code {} is also used as *"diphthong-divider"* for the very rare occurrence of two consecutive vowels in *Sanskrit* (whereas in *Hindi*, two consecutive vowels are required frequently). Here the {} can be used to prevent automatic generation of diphthongs. Example: The ITX codes ai and au always denote the two respective Sanskrit diphthongs, and Itranslator automatically converts "au" to औ., e.g. devau = देवौ (Dual). However, if you need two consecutive vowels for writing one of those extremely rare Sanskrit words such as "titäu" (*sieve*, ä = a-dieresis, a-trema), i.e. "tita"+ "u", you may either encode "tita{u}" or "tita{}u", both of which result in तितउ, transliterated by Itranslator as "titau" (without "ä").

5.2.2. Halant = Virama Enforcer ".h"

Itranslator automatically generates Virama (्), if a consonant as final character of a word is followed by one of these delimiter codes:

Delimiter	Examples
CR + LF	खादति फलम् consonant terminates a line of text – T: "*He eats the fruit.*"
Danda	खादति फलम् । consonant is followed by single/double danda
Space	अखादत् फलम् consonant is followed by space– T: "*He ate the fruit.*"
\	किं खादति फलम्? consonant is followed by "\" + punctuation mark, e.g. \?
-	ऋग्-वेदः ṛg-vedaḥ hyphen in compound as a help for Sanskrit learners

Occasionally it may be necessary to enforce the virama by inserting the code **".h"**, but Itranslator itself does not require the Halant code ".h", whereas in mixed Hindi-Sanskrit encoding environments, e.g. ITRANS (see 5.3.), it is usually necessary to enforce virama.

Note: In Hindi, "virām" denotes the period or full-stop, while "halant" denotes the sign "्". In Sanskrit, "virāmaḥ" means sign "्", while "halantaḥ" is a consonant at the end of word.

5.3. ITRANS Textfiles

ITRANS textfiles encoded for use by the converter ITRANS (current release 5.3, 2001), developed by Avinash Chopde, contain – beside the straightforward ITX encoded texts – a lot of additional **meta-commands** not used by Itranslator. Luckily, these commands are usually only found in the file header preceding the respective Sanskrit text. Example:

```
%@@1                                                                Header
% File name        : akshi.itx
%------------------------------------------
% Text title        : Akshi Upanishad
% Author            : Vedic tradition
% Language          : Sanskrit
% Subject           : philosophy/hinduism
% Description/comments  : # 72 / 108; Krishna Yajurveda - Samanya upanishad
% Latest update      : Mar. 24, 2000
% Send corrections to  : sanskrit@cheerful.com
% Special Instructions:
% i1h.hdr,ijag.inc,itrans.sty,multicol.sty,iarticle.sty
% Transliteration scheme: ITRANS 5.2
% Site access  :
% http://sanskrit.gde.to
%----------------------------------------------------
% The text is to be used for personal studies and research only.
% Any use for commercial purpose is prohibited as a 'gentleman's'
% agreement.
% @@2
% Please ignore the following commands upto #indian and others with
% \ mark,
% which are needed for devanaagarii output and formatting.
%-------------------------------------------------------
\documentstyle[11pt,multicol,itrans]{article}
#include=ijag.inc
#endwordvowel=.h
\portraitwide
\parindent=50pt
\let\usedvng=\Largedvng      % for 1 column
\pagenumbering{itrans}
\def\engtitle#1{\hrule\medskip\centerline{\large #1}}
\def\itxtitle#1{\medskip\centerline{\LARGEdvng #1}\medskip\hrule}
\def\endtitles{\medskip\obeyspaceslines}
%%
\begin{document}
\engtitle{.. Akshi Upanishad ..}##
\itxtitle{.. akshyupanishhat.h ..}##\endtitles##
```

```
yatsaptabhuumikaavidyaavedyaanandakalevaram.h .          Sanskrit Text
vikalevarakaivalya.n raamachandrapadaM bhaje ..
etc. etc.
```

ITRANS textfiles can be best converted by Itranslator, **if you delete the complete header**. For further details on ITRANS textfiles visit the website http://www.aczoom.com/itrans

5.4. Flush-Right Hyphenated Typesetting of Prose Sanskrit Texts

Word processors (MS-Word etc.) require conditional hyphens for flush-right composition. For this reason, the **soft hyphen** (code 173) has been implemented into the Sanskrit 99 font to enable hyphenated typesetting of prose Sanskrit texts. After importing an Itranslator-generated RTF file in your word processor, you may refine typesetting by insertion of soft hyphens, if possible also by a word processor "macro". A **hard hyphen** (code 45) is always visible, while a soft hyphen remains invisible, unless it is needed as hyphen at the end of a line. Therefore you will never be annoyed by visible hyphens in the middle of lines of text, when you reformat the text in another point size or make corrections, deletions, insertions.

Prose Text without Soft Hyphens
अस्ति भागीरथीतीरे गृध्रकूटनाम्नि पर्वते महान्पर्कटीवृक्षः। तस्य कोटरे देवदुर्विपाकाद्गलितनखनयनो जरद्‌ध्वो नाम गृध्रः प्रतिवसति। तज्जीवनाय तद्‌वृक्षवासिनः पक्षिणः स्वाहारात्किञ्चित्किञ्चिद्‌ददति। तेनासौ जीवति। शावकानां रक्षणं करोति। अथ कदाचिद्दीर्घकर्णो नाम मार्जारः पक्षिशावकान्भक्षयितुमागतः। तत्तमायान्तं दृष्ट्वा शावकैर्भयाकुलैर्महान्कोलाहलः कृतः। तच्छ्रुत्वा जरद्‌ध्वेनोक्तम्। कोऽयमायाति। दीर्घकर्णो गृध्रमवलोक्य सभयं हा हतोऽस्मि। अधुनातिसंनिधाने पलायितुमक्षमः। तद्यथा प्राप्तव्यं तद्‌भवतु। एतत्समीपमुपगच्छामि। इत्यालोच्योपसृत्याब्रवीत्। आर्य त्वामभिवन्दे। गृध्रो ब्रूते। कस्त्वम्। सोऽब्रवीत्। मार्जारोऽहम्। तदपसर। नो चेद्धन्तव्योऽसि मया। मार्जारोऽवदत्। श्रूयतां तावन्मम वचनम्। ततो यद्यहं वध्यस्तदा हन्तव्यः।
Prose Text with Soft Hyphens
अस्ति भागीरथीतीरे गृध्रकूटनाम्नि पर्वते महान्पर्कटीवृक्षः। तस्य कोटरे देवदुर्विपाकाद्गलितनख– नयनो जरद्‌ध्वो नाम गृध्रः प्रतिवसति। तज्जीवनाय तद्‌वृक्षवासिनः पक्षिणः स्वाहारात्किञ्चित्किञ्चिद्‌– ददति। तेनासौ जीवति। शावकानां रक्षणं करोति। अथ कदाचिद्दीर्घकर्णो नाम मार्जारः पक्षिशाव– कान्भक्षयितुमागतः। तत्तमायान्तं दृष्ट्वा शावकैर्भयाकुलैर्महान्कोलाहलः कृतः। तच्छ्रुत्वा जरद्‌– ध्वेनोक्तम्। कोऽयमायाति। दीर्घकर्णो गृध्रमवलोक्य सभयं हा हतोऽस्मि। अधुनातिसंनिधाने पलायितुमक्षमः। तद्यथा प्राप्तव्यं तद्‌भवतु। एतत्समीपमुपगच्छामि। इत्यालोच्योपसृत्याब्रवीत्। आर्य त्वामभिवन्दे। गृध्रो ब्रूते। कस्त्वम्। सोऽब्रवीत्। मार्जारोऽहम्। तदपसर। नो चेद्धन्त– व्योऽसि मया। मार्जारोऽवदत्। श्रूयतां तावन्मम वचनम्। ततो यद्यहं वध्यस्तदा हन्तव्यः।

The text has been drawn from Hitopadeśa, Part I, Fable 3: "Vulture, Cat and Little Birds".

5.5. Variants of Floating Accents or Matras

Occasionally you may wish to refine a Devanagari word according to your personal taste, e.g. Itranslator converts "vartate" to वर्तते with very elegantly overlapping floating accents. But Sanskrit beginners might be confused, so that वर्तते might be better readable for them. For making such refinements, use following chart of floating accents (zero-width glyphs), where न is used as reference glyph to make visible positional differences among variants:

ॖ	नॖ (033)	नॖ (137)	नॖ (156)
�	नं (151)	नं (060)	
॰	नॕ (082)	नँ (155)	
ॢ	नृ (038)	नृ (175)	न्न (171)
ॣ	नॣ (168)	नॣ (176)	नॄ (169)
ॗ	नु (117)	नु (130)	न्नु (133)
ॖ	नू (085)	नू (131)	न्नू (170)
॰	नॢ (150)[1]	न्न (136)[1]	
ॷ	नँ (126)[2]	नॅ (163)[2]	
॰	नॕ (161)[2]		
ॖ	ने (101)[2]	नै (069)[2]	

[1] Nukta is inserted **before** the respective letter ("nukta" is of Persian origin denoting "dot").
[2] These matras do not have variants. A matra is a (non-a) vowel sign put above or below a consonant. In a broader sense, the term "matra" includes all other floating accents too.

The floating accents are inserted by pressing the ALT key and by entering the respective ANSI character code, prefixed by 0, on the numerical key pad, for example ALT 0137.

5.6. Sanskrit 99 ps (PostScript Type 1 Font)

In addition to the TrueType font "**Sanskrit 99**" (file name: Sanskrit99.ttf), you may also use the PostScript Type 1 font "**Sanskrit 99 ps**" (file names sans99ps.pfb, sans99ps.pfm), which is recommended for those who own a PostScript printer **or** who want to make PDF files with Acrobat Distiller **or** who have to create imagesetter files for book printing plants. Professsional imagesetters and computer-to-plate systems usually prefer PostScript fonts.

5.7. Extracting Sanskrit 99 ps Texts from PDF Files

It is easy to extract the entire Devanagari text from a PDF file created with Sanskrit 99 ps, provided that the text is a "clean" Devanagari text (without English etc. insertions) typeset flush left and with one column per page. However, owing to internal rounding errors relating to zero-width glyphs (in fact not "zero", but with a width of 1 unit per em quad), the clipped file may contain a couple of unwanted spaces, depending on printer driver.

Experiment: 1) In Acrobat Reader: Clip and Copy the following text. 2) In Word Processor: Paste the text and mark it up with Sanskrit 99 ps or with Sanskrit 99 (both is possible).

धृतराष्ट्र उवाच ।

धर्मक्षेत्रे कुरुक्षेत्रे समवेता युयुत्सवः ।

मामकाः पाण्डवाश्चैव किमकुर्वत संजय ॥ १-१ ॥

संजय उवाच ।

दृष्ट्वा तु पाण्डवानीकं व्यूढं दुर्योधनस्तदा ।

आचार्यमुपसंगम्य राजा वचनमब्रवीत्॥ १-२ ॥

पश्यैतां पाण्डुपुत्राणामाचार्य महतीं चमूम् ।

व्यूढां द्रुपदपुत्रेण तव शिष्येण धीमता ॥ १-३ ॥

अत्र शूरा महेष्वासा भीमार्जुनसमा युधि ।

युयुधानो विराटश्च द्रुपदश्च महारथः ॥ १-४ ॥

धृष्टकेतुश्चेकितानः काशिराजश्च वीर्यवान् ।

पुरुजित्कुन्तिभोजश्च शैब्यश्च नरपुंगवः ॥ १-५ ॥

युधामन्युश्च विक्रान्त उत्तमौजाश्च वीर्यवान् ।

सौभद्रो द्रौपदेयाश्च सर्व एव महारथाः ॥ १-६ ॥

अस्माकं तु विशिष्टा ये तान्निबोध द्विजोत्तम ।

नायका मम सैन्यस्य संज्ञार्थं तान्ब्रवीमि ते ॥ १-७ ॥

भवान्भीष्मश्च कर्णश्च कृपश्च समितिंजयः ।

अश्वत्थामा विकर्णश्च सौमदत्तिस्तथैव च ॥ १-८ ॥

अन्ये च बहवः शूरा मदर्थे त्यक्तजीविताः ।

नानाशस्त्रप्रहरणाः सर्वे युद्धविशारदाः ॥ १-९ ॥

Pages 1-20 of this manual were typeset in the PostScript Type 1 typefaces URW Palladio IT and Sanskrit 99 ps

Appendix: Sanskrit 99 vs. Sanskrit 2003

The 360 most often used ligatures of Classical Sanskrit (frequency of occurrence > 0.010%) are typeset below in "Sanskrit 2003" (column S03) and in "Sanskrit 99 ps" (column S99):

S03	S99	
क्क	क्क	kk (0,070)
क्च	क्च	kc (0,064)
क्त	क्त	kt (4,850)
क्त्य	क्त्य	kty (0,162)
क्त्र	क्त्र	ktr (0,118)
क्त्व	क्त्व	ktv (0,687)
क्थ	क्थ	kth (0,013)
क्न	क्न	kn (0,102)
क्प	क्प	kp (0,088)
क्प्र	क्प्र	kpr (0,037)
क्म	क्म	km (0,140)
क्य	क्य	ky (1,308)
क्र	क्र	kr (5,207)
क्ल	क्ल	kl (0,391)
क्व	क्व	kv (0,300)
क्श	क्श	kś (0,043)
क्ष	क्ष	kṣ (9,976)
क्ष्ण	क्ष्ण	kṣṇ (0,305)
क्ष्म	क्ष्म	kṣm (0,247)
क्ष्म्य	क्ष्म्य	kṣmy (0,026)
क्ष्य	क्ष्य	kṣy (1,476)
क्ष्व	क्ष्व	kṣv (0,176)
क्स	क्स	ks (0,069)
ख्य	ख्य	khy (1,307)
ग्ग	ग्ग	gg (0,025)
ग्ज	ग्ज	gj (0,025)
ग्ज्य	ग्ज्य	gjy (0,030)
ग्ण	ग्ण	gṇ (0,020)

ग्द	ग्द	gd (0,034)
ग्ध	ग्ध	gdh (0,249)
ग्ध्व	ग्ध्व	gdhv (0,018)
ग्न	ग्न	gn (1,290)
ग्न्य	ग्न्य	gny (0,081)
ग्ब	ग्ब	gb (0,023)
ग्भ	ग्भ	gbh (0,101)
ग्भ्य	ग्भ्य	gbhy (0,018)
ग्म	ग्म	gm (0,364)
ग्य	ग्य	gy (0,191)
ग्र	ग्र	gr (2,453)
ग्र्य	ग्र्य	gry (0,150)
ग्ल	ग्ल	gl (0,052)
ग्व	ग्व	gv (0,126)
घ्न	घ्न	ghn (0,545)
घ्र	घ्र	ghr (0,743)
घ्व	घ्व	ghv (0,016)
ङ्क	ङ्क	ṅk (0,505)
ङ्क्त	ङ्क्त	ṅkt (0,091)
ङ्क्ष	ङ्क्ष	ṅkṣ (0,195)
ङ्क्ष्व	ङ्क्ष्व	ṅkṣv (0,025)
ङ्ख	ङ्ख	ṅkh (0,330)
ङ्ख्य	ङ्ख्य	ṅkhy (0,052)
ङ्ग	ङ्ग	ṅg (2,091)
ङ्ग्य	ङ्ग्य	ṅgy (0,018)
ङ्ग्र	ङ्ग्र	ṅgr (0,019)
ङ्घ	ङ्घ	ṅgh (0,038)
ङ्घ्र	ङ्घ्र	ṅghr (0,109)
ङ्न	ङ्न	ṅn (0,030)

ङ्म	ङ्म	ṅm (0,137)
च्च	च्च	cc (1,716)
च्च्य	च्च्य	ccy (0,022)
च्छ	च्छ	cch (4,297)
च्छ्र	च्छ्र	cchr (0,535)
च्छ्व	च्छ्व	cchv (0,033)
च्य	च्य	cy (1,111)
ज्ज	ज्ज	jj (0,584)
ज्ज्ञ	ज्ज्ञ	jjñ (0,061)
ज्ज्य	ज्ज्य	jjy (0,017)
ज्ज्व	ज्ज्व	jjv (0,033)
ज्ञ	ज्ञ	jñ (4,640)
ज्म	ज्म	jm (0,061)
ज्य	ज्य	jy (1,937)
ज्र	ज्र	jr (0,329)
ज्व	ज्व	jv (0,435)
ञ्च	ञ्च	ñc (1,779)
ञ्छ	ञ्छ	ñch (0,044)
ञ्ज	ञ्ज	ñj (1,057)
ञ्ज्ञ	ञ्ज्ञ	ñjñ (0,021)
ञ्श	ञ्श	ñś (0,528)
ञ्श्र	ञ्श्र	ñśr (0,073)
ञ्श्व	ञ्श्व	ñśv (0,018)
ट्क	ट्क	ṭk (0,030)
ट्ट	ट्ट	ṭṭ (0,091)
ट्प	ट्प	ṭp (0,015)
ट्य	ट्य	ṭy (0,037)
ट्श	ट्श	ṭś (0,014)
ट्स	ट्स	ṭs (0,032)

ठ्य	ठ्य	ṭhy (0,013)	त्र	त्र	tr (14,060)	द्र	द्र	dr (5,763)
ड्ग	ड्ग	ḍg (0,167)	त्र्य	त्र्य	try (0,158)	द्र्य	द्र्य	dry (0,016)
ड्भ	ड्भ	ḍbh (0,049)	त्व	त्व	tv (8,125)	द्व	द्व	dv (5,162)
ड्य	ड्य	ḍy (0,097)	त्व्य	त्व्य	tvy (0,015)	द्व्य	द्व्य	dvy (0,115)
ड्र	ड्र	ḍr (0,018)	त्स	त्स	ts (3,350)	द्व्र	द्व्र	dvr (0,020)
ड्व	ड्व	ḍv (0,037)	त्स्त	त्स्त	tst (0,017)	ध्न	ध्न	dhn (0,033)
ढ्य	ढ्य	ḍhy (0,043)	त्स्त्र	त्स्त्र	tstr (0,016)	ध्म	ध्म	dhm (0,066)
ण्ट	ण्ट	ṇṭ (0,068)	त्स्थ	त्स्थ	tsth (0,072)	ध्य	ध्य	dhy (2,680)
ण्ठ	ण्ठ	ṇṭh (0,095)	त्स्न	त्स्न	tsn (0,217)	ध्र	ध्र	dhr (0,335)
ण्ड	ण्ड	ṇḍ (4,570)	त्स्म	त्स्म	tsm (0,033)	ध्व	ध्व	dhv (0,897)
ण्ड्य	ण्ड्य	ṇḍy (0,045)	त्स्य	त्स्य	tsy (0,529)	न्क	न्क	nk (0,806)
ण्ड्र	ण्ड्र	ṇḍr (0,018)	त्स्र	त्स्र	tsr (0,016)	न्क्र	न्क्र	nkr (0,066)
ण्ण	ण्ण	ṇṇ (0,077)	त्स्व	त्स्व	tsv (0,191)	न्क्ल	न्क्ल	nkl (0,013)
ण्म	ण्म	ṇm (0,050)	थ्य	थ्य	thy (0,363)	न्क्ष	न्क्ष	nkṣ (0,112)
ण्य	ण्य	ṇy (1,523)	थ्व	थ्व	thv (0,018)	न्ख	न्ख	nkh (0,025)
ण्व	ण्व	ṇv (0,137)	द्ग	द्ग	dg (0,483)	न्ग	न्ग	ng (0,304)
त्क	त्क	tk (1,883)	द्ग्र	द्ग्र	dgr (0,018)	न्घ	न्घ	ngh (0,043)
त्क्र	त्क्र	tkr (0,135)	द्घ	द्घ	dgh (0,061)	न्त	न्त	nt (12,919)
त्क्व	त्क्व	tkv (0,018)	द्द	द्द	dd (0,868)	न्त्य	न्त्य	nty (0,820)
त्क्ष	त्क्ष	tkṣ (0,166)	द्द्य	द्द्य	ddy (0,025)	न्त्र	न्त्र	ntr (0,536)
त्ख	त्ख	tkh (0,025)	द्द्र	द्द्र	ddr (0,109)	न्त्र्य	न्त्र्य	ntry (0,088)
त्त	त्त	tt (7,464)	द्द्व	द्द्व	ddv (0,119)	न्त्व	न्त्व	ntv (0,144)
त्त्य	त्त्य	tty (0,081)	द्ध	द्ध	ddh (6,155)	न्त्स्य	न्त्स्य	ntsy (0,023)
त्त्र	त्त्र	ttr (0,099)	द्ध्य	द्ध्य	ddhy (0,217)	न्थ	न्थ	nth (0,133)
त्त्व	त्त्व	ttv (1,339)	द्ध्र	द्ध्र	ddhr (0,017)	न्द	न्द	nd (2,632)
त्थ	त्थ	tth (0,536)	द्ध्व	द्ध्व	ddhv (0,215)	न्द्य	न्द्य	ndy (0,095)
त्न	त्न	tn (0,625)	द्न	द्न	dn (0,037)	न्द्र	न्द्र	ndr (2,427)
त्न्य	त्न्य	tny (0,038)	द्ब	द्ब	db (0,418)	न्द्व	न्द्व	ndv (0,085)
त्प	त्प	tp (1,918)	द्ब्र	द्ब्र	dbr (0,277)	न्ध	न्ध	ndh (2,396)
त्प्र	त्प्र	tpr (0,910)	द्भ	द्भ	dbh (1,687)	न्ध्य	न्ध्य	ndhy (0,067)
त्फ	त्फ	tph (0,071)	द्भ्य	द्भ्य	dbhy (0,071)	न्ध्र	न्ध्र	ndhr (0,046)
त्म	त्म	tm (4,163)	द्भ्र	द्भ्र	dbhr (0,057)	न्न	न्न	nn (5,086)
त्म्य	त्म्य	tmy (0,052)	द्म	द्म	dm (0,263)	न्न्य	न्न्य	nny (0,027)
त्य	त्य	ty (9,763)	द्य	द्य	dy (5,591)	न्व	न्व	nnv (0,025)

न्प	न्प	np (1,195)
न्प्र	न्प्र	npr (0,714)
न्फ	न्फ	nph (0,014)
न्ब	न्ब	nb (0,319)
न्ब्र	न्ब्र	nbr (0,120)
न्भ	न्भ	nbh (0,493)
न्भ्र	न्भ्र	nbhr (0,054)
न्म	न्म	nm (2,803)
न्य	न्य	ny (7,223)
न्र	न्र	nr (0,539)
न्व	न्व	nv (2,650)
न्व्य	न्व्य	nvy (0,089)
न्स	न्स	ns (1,767)
न्स्थ	न्स्थ	nsth (0,038)
न्स्म	न्स्म	nsm (0,017)
न्स्व	न्स्व	nsv (0,103)
न्ह	न्ह	nh (0,345)
प्त	प्त	pt (2,664)
प्त्व	प्त्व	ptv (0,035)
प्न	प्न	pn (0,608)
प्म	प्म	pm (0,055)
प्य	प्य	py (2,285)
प्र	प्र	pr (21,172)
प्ल	प्ल	pl (0,300)
प्स	प्स	ps (0,523)
प्स्य	प्स्य	psy (0,235)
ब्ज	ब्ज	bj (0,014)
ब्द	ब्द	bd (0,460)
ब्ध	ब्ध	bdh (0,464)
ब्ध्व	ब्ध्व	bdhv (0,095)
ब्य	ब्य	by (0,030)
ब्र	ब्र	br (4,583)
ब्व	ब्व	bv (0,031)
भ्ण	भ्ण	bhṇ (0,017)

भ्य	भ्य	bhy (3,240)
भ्र	भ्र	bhr (1,242)
भ्व	भ्व	bhv (0,055)
म्ण	म्ण	mṇ (0,021)
म्न	म्न	mn (0,542)
म्प	म्प	mp (0,245)
म्ब	म्ब	mb (0,755)
म्ब्य	म्ब्य	mby (0,014)
म्भ	म्भ	mbh (0,483)
म्य	म्य	my (2,118)
म्र	म्र	mr (0,129)
म्ल	म्ल	ml (0,059)
य्य	य्य	yy (0,200)
य्व	य्व	yv (0,017)
र्क	र्क	rk (0,321)
र्क्ष्य	र्क्ष्य	rkṣy (0,024)
र्ख	र्ख	rkh (0,020)
र्ग	र्ग	rg (1,698)
र्ग्य	र्ग्य	rgy (0,023)
र्ग्र	र्ग्र	rgr (0,030)
र्घ	र्घ	rgh (0,379)
र्घ्य	र्घ्य	rghy (0,035)
र्ङ	र्ङ	rṅg (0,030)
र्च	र्च	rc (0,365)
र्च्छ	र्च्छ	rcch (0,026)
र्च्य	र्च्य	rcy (0,038)
र्छ	र्छ	rch (0,081)
र्ज	र्ज	rj (2,240)
र्ज्ञ	र्ज्ञ	rjñ (0,028)
र्ज्य	र्ज्य	rjy (0,050)
र्ज्व	र्ज्व	rjv (0,020)
र्ण	र्ण	rṇ (3,357)
र्ण्य	र्ण्य	rṇy (0,047)
र्त	र्त	rt (4,314)

र्त्त	र्त्त	rtt (0,033)
र्त्म	र्त्म	rtm (0,040)
र्त्य	र्त्य	rty (0,182)
र्त्र	र्त्र	rtr (0,021)
र्त्व	र्त्व	rtv (0,023)
र्त्स	र्त्स	rts (0,017)
र्त्स्न्य	र्त्स्न्य	rtsny (0,032)
र्थ	र्थ	rth (4,883)
र्थ्य	र्थ्य	rthy (0,032)
र्द	र्द	rd (1,873)
र्द्ध	र्द्ध	rddh (0,017)
र्द्य	र्द्य	rdy (0,035)
र्द्र	र्द्र	rdr (0,132)
र्द्व	र्द्व	rdv (0,142)
र्ध	र्ध	rdh (1,434)
र्ध्न	र्ध्न	rdhn (0,062)
र्ध्म	र्ध्म	rdhm (0,015)
र्ध्य	र्ध्य	rdhy (0,021)
र्ध्र	र्ध्र	rdhr (0,031)
र्ध्व	र्ध्व	rdhv (0,185)
र्न	र्न	rn (1,118)
र्न्य	र्न्य	rny (0,023)
र्प	र्प	rp (0,532)
र्ब	र्ब	rb (0,731)
र्ब्र	र्ब्र	rbr (0,112)
र्भ	र्भ	rbh (1,285)
र्भ्य	र्भ्य	rbhy (0,033)
र्भ्र	र्भ्र	rbhr (0,042)
र्म	र्म	rm (9,114)
र्म्य	र्म्य	rmy (0,080)
र्य	र्य	ry (6,633)
र्ल	र्ल	rl (0,225)
र्व	र्व	rv (11,898)
र्व्य	र्व्य	rvy (0,146)

र्श	र्श	rś (1,443)	श्प	श्प	śp (0,026)	स्त्य	स्त्य	sty (0,332)
र्श्व	र्श्व	rśv (0,103)	श्म	श्म	śm (0,380)	स्त्र	स्त्र	str (2,764)
र्ष	र्ष	rṣ (4,117)	श्य	श्य	śy (3,276)	स्त्र्य	स्त्र्य	stry (0,020)
र्ष्ट	र्ष्ट	rṣṭ (0,025)	श्र	श्र	śr (5,604)	स्त्व	स्त्व	stv (1,339)
र्ष्ण	र्ष्ण	rṣṇ (0,174)	श्ल	श्ल	śl (0,204)	स्थ	स्थ	sth (3,492)
र्ष्य	र्ष्य	rṣy (0,026)	श्व	श्व	śv (3,454)	स्थ्य	स्थ्य	sthy (0,030)
र्ह	र्ह	rh (1,688)	श्व्य	श्व्य	śvy (0,032)	स्न	स्न	sn (0,432)
र्ह्य	र्ह्य	rhy (0,018)	ष्क	ष्क	ṣk (0,576)	स्प	स्प	sp (1,003)
र्ह्र	र्ह्र	rhr (0,019)	ष्क्र	ष्क्र	ṣkr (0,060)	स्फ	स्फ	sph (0,145)
ल्क	ल्क	lk (0,117)	ष्ट	ष्ट	ṣṭ (4,855)	स्म	स्म	sm (4,964)
ल्क्य	ल्क्य	lky (0,097)	ष्ट्य	ष्ट्य	ṣṭy (0,237)	स्म्य	स्म्य	smy (0,055)
ल्ग	ल्ग	lg (0,232)	ष्ट्र	ष्ट्र	ṣṭr (1,252)	स्य	स्य	sy (13,483)
ल्प	ल्प	lp (0,754)	ष्ट्व	ष्ट्व	ṣṭv (1,405)	स्र	स्र	sr (1,481)
ल्ब	ल्ब	lb (0,095)	ष्ठ	ष्ठ	ṣṭh (4,521)	स्व	स्व	sv (4,801)
ल्म	ल्म	lm (0,145)	ष्ठ्य	ष्ठ्य	ṣṭhy (0,015)	स्स्व	स्स्व	ssv (0,016)
ल्य	ल्य	ly (1,178)	ष्ण	ष्ण	ṣṇ (1,793)	ह्ण	ह्ण	hṇ (0,273)
ल्ल	ल्ल	ll (0,563)	ष्ण्य	ष्ण्य	ṣṇy (0,047)	ह्न	ह्न	hn (0,201)
ल्व	ल्व	lv (0,145)	ष्प	ष्प	ṣp (0,451)	ह्म	ह्म	hm (2,980)
व्य	व्य	vy (6,095)	ष्प्र	ष्प्र	ṣpr (0,073)	ह्य	ह्य	hy (2,118)
व्र	व्र	vr (1,172)	ष्म	ष्म	ṣm (1,504)	ह्र	ह्र	hr (0,341)
श्च	श्च	śc (12,999)	ष्य	ष्य	ṣy (3,440)	ह्ल	ह्ल	hl (0,134)
श्च्य	श्च्य	ścy (0,044)	ष्व	ष्व	ṣv (1,145)	ह्व	ह्व	hv (0,420)
श्छ	श्छ	śch (0,187)	स्क	स्क	sk (0,713)			
श्न	श्न	śn (0,297)	स्त	स्त	st (13,762)			