

# Brahmic TECKit mappings

David M. Jones

Version 0.01 beta, 2015/06/30

## Abstract

This file defines a set of TECKit mappings based on ISO 15919 *Transliteration of Devanagari and related Indic scripts into Latin characters* for reverse transliteration of Latin-based characters into various Indic scripts in the Brahmic family.

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
1.1	Acknowledgements . . . . .	2
1.2	Caveats . . . . .	3
1.3	Supported scripts . . . . .	3
<b>2</b>	<b>Transcription table</b>	<b>3</b>
2.1	Special conventions . . . . .	5
2.2	Special extensions for Sanskrit . . . . .	5
<b>3</b>	<b>TECKit header</b>	<b>6</b>
<b>4</b>	<b>Overview</b>	<b>6</b>
<b>5</b>	<b>Normalized input characters</b>	<b>7</b>
5.1	Vowels and diphthongs . . . . .	7
5.2	Combining accents . . . . .	8
<b>6</b>	<b>Output Characters</b>	<b>8</b>
6.1	Script-independent characters . . . . .	8
6.2	Devanagari . . . . .	8
6.3	Bengali . . . . .	13
6.4	Gujarati . . . . .	16
6.5	Gurmukhi . . . . .	18
6.6	Kannada . . . . .	21
6.7	Malayalam . . . . .	24
6.8	Oriya . . . . .	27
6.9	Sinhala . . . . .	30

6.10	Tamil . . . . .	32
6.11	Telugu . . . . .	35
7	<b>Phase 1: Normalization</b>	<b>38</b>
7.1	Normalization Pass 1: Decompose compound characters. . . . .	38
7.2	Normalization Pass 2: Downcase base characters. . . . .	38
7.3	Normalization Pass 3: Special cases and alternate transcriptions. . . . .	39
8	<b>Sanskrit External Vowel Sandhi</b>	<b>40</b>
8.1	Irregular sandhi . . . . .	41
9	<b>Phase 2: Transliteration</b>	<b>41</b>
9.1	Transliteration Pass 1: Transliterate blindly. . . . .	41
9.2	Transliteration Pass 1.5 . . . . .	45
9.3	Transliteration Pass 2: Attach vowels to consonants. . . . .	46
9.4	Transliteration Pass 3: Form consonant conjuncts. . . . .	46
9.5	Transliteration Pass 4: Add viramas. . . . .	47
9.6	Transliteration Pass 5: Remove inherent a. . . . .	47
9.7	Transliteration Pass 6: Resolve PUA characters. . . . .	47

## 1 Introduction

This grew out of my experiments with typesetting Sanskrit text using Xe<sub>La</sub>TeX, which in turn stemmed from my attempts to satisfy my curiosity about how the Devanagari script works. With these mappings and suitable macros, you can write<sup>1</sup>

```
A simple \textbf{liquid vowel} followed by any \emph{other} vowel or
by a diphthong is changed into its semivowel;—e.g. \N{dadhi atra} =
\NS{dadhy atra}; \N{karṭṛ uta} = \NS{karṭṛ uta}; \N{madhu iva} = \NS{madv
iva}; \N{nadī artham} = \N{nadyartham}.
```

and get

A simple **liquid vowel** followed by any *other* vowel or by a diphthong is changed into its semivowel;—e.g. दधि अत्र **dadhi atra** = दध्यत्र **dadhy atra**; कर्तृ उत **karṭṛ uta** = कर्तृत **karṭṛ uta**; मधु इव **madhu iva** = मद्विव **madv iva**; नदी अर्थम् **nadī artham** = नद्यर्थम् **nadyartham**.

Since the same input is used to generate both the Roman and the Devanagari text, this eliminates the possibility that a typo in one or the other will cause a mismatch—assuming the mapping is implemented correctly, of course!

### 1.1 Acknowledgements

The first draft map was based on version 1.2 of `velthuis-sanskrit.map`. I’ve mauled the code almost beyond recognition, but its influence is still apparent throughout Phase 2.

<sup>1</sup>Quoted from section 20 of *A Sanskrit Grammar for Students*[3]

## 1.2 Caveats

I'm fairly confident that the Devanagari transliteration works,<sup>2</sup> but I'm much less sanguine about the other scripts, of which I lack even the modest knowledge I have about Devanagari. I'd be grateful for any feedback from more knowledgeable users.

## 1.3 Supported scripts

deva	Devanagari
beng	Bengali
gujr	Gujarati
guru	Gurmukhi
knda	Kannada
mlym	Malayalam
orya	Oriya
sinh	Sinhala
taml	Tamil
telu	Telugu

## 2 Transcription table

This table is adapted from [1]. The main (intentional) departures are in the rows for e, ē, ê, o, ô, and ɱ. Strictly speaking, ISO 15919 only allows them in selected scripts, but for convenience, I've allowed them to be used in all scripts as synonyms for something similar.

For purposes of illustration, I tried to find fonts that contain all of the characters covered, but there are still some missing glyphs in Devanagari, Kannada, Malayalam, Oriya, Telugu. I'd be grateful for suggestions of fonts containing those characters.

Latin	Dev.	San.	Gur.	Guj.	Ben.	Ori.	Tam.	Mal.	Kan.	Tel.	Sin.
a	अ	অ	ਅ	અ	অ	ଅ	அ	അ	ಅ	అ	අ
ā	आ	আ	ਆ	આ	আ	ଆ	ஆ	ആ	ಆ	ఆ	ආ
æ											ඇ
æ̃											ඈ
i	इ	ই	ਇ	ઈ	ই	ଇ	இ	ഇ	ಇ	ఇ	ඉ
ī	ई	ঈ	ਈ	ઈ	ঈ	ଈ	ஈ	ഈ	ಈ	ఈ	ඊ
u	उ	উ	ਉ	ઉ	উ	ଉ	உ	ഉ	ಉ	ఉ	උ
ū	ऊ	ঊ	ਊ	ઊ	ঊ	ଊ	ஊ	ഊ	ಊ	ఊ	ඌ
ũ								ഘ			
r	ऋ	ঋ		ઋ	ঋ	ଋ		ഠ	ಠ	ఠ	ඹ
ṛ	ॠ	ৠ		ॠ	ৠ	ॠ		ഡ	ॡ	డ	ඹ
l	ल	ল			ল	ଳ		ണ	ల	ల	ల
ḷ	ळ	৳			৳	ḷ		ణ	ల	ల	ల
e	ऐ	এ	ਏ	એ	এ	ଏ	ஏ	എ	ఎ	ఎ	එ
ē	ए	এ	ਏ	એ	এ	ଏ	ஏ	എ	ఎ	ఎ	එ

<sup>2</sup>At least, it did before I started adding re-arranging things to add other scripts.

Latin	Dev.	San.	Gur.	Guj.	Ben.	Ori.	Tam.	Mal.	Kan.	Tel.	Sin.
ê	ऐँ	ए	ਏ	ઐ	ଏ	ଏ	ஏ	ഏ	ಏ	ఏ	ඒ
ai	ऐ	ऐ	ਐ	ઐ	ਐ	ਐ	ஐ	ഐ	ಐ	ఐ	ඒ
o	ओ	ओ	ਓ	ઓ	ਓ	ਓ	ஓ	ഓ	ಓ	ఓ	ඔ
ô	ओ	ओ	ਓ	ઓ	ਓ	ਓ	ஓ	ഓ	ಓ	ఓ	ඔ
au	औ	औ	ਔ	ઔ	ਔ	ਔ	ஔ	ഔ	ಔ	ఔ	ඔ
m̐	.	.	.	.	.	.		.	.	.	◻
m̐	.	.	.	.	.	.		.	.	.	◻
m̐	.	.	.	.	.	.		.	.	.	◻
n̐	.	.	.	.	.	.		.	.	.	◻
ḥ	:	:	:	:	:	:		:	:	:	◻
ḥ	◻	◻	◻	◻	◻	◻		◻	◻	◻	◻
ḥ	◻	◻	◻	◻	◻	◻		◻	◻	◻	◻
k							ஃ				
ka	क	क	ਕ	ક	ক	କ	க	ക	ಕ	క	ක
kha	ख	ख	ਖ	ખ	খ	ଖ		ഖ	ಖ	ఖ	ක
ga	ग	ग	ਗ	ગ	গ	ଗ		ഗ	ಗ	గ	ක
gha	घ	घ	ਘ	ઘ	ঘ	ଘ	஝	ഘ	ಘ	ఘ	ක
ṅa	ङ	ङ	ਙ	ઙ	ঙ	ଙ	ஞ	ങ	ಙ	ఙ	ක
ṅga											
ca	च	च	ਚ	ચ	চ	ଚ	ச	ച	ಚ	చ	ච
ĉa										◻	
cha	छ	छ	ਛ	છ	ছ	ଛ		ചറ	ಚ	చ	ච
ja	ज	ज	ਜ	જ	জ	ଜ	ஐ	ജ	ಜ	జ	ච
jha	झ	झ	ਝ	ઝ	ঝ	ଝ		ఙ	ఙ	ఙ	ක
ṅa	ञ	ञ	ਞ	ઞ	ঞ	ଞ	ஐ	ഞ	ಞ	ఞ	ක
ṅja											
ṭa	ट	ट	ਟ	ટ	ট	ଟ	ட	ട	ಟ	ట	ට
ṭha	ठ	ठ	ਠ	ઠ	ঠ	ଠ		ഠ	ഠ	ఠ	ට
ḍa	ड	ड	ਡ	ડ	ড	ଡ		ഡ	ഡ	డ	ඨ
ṛa	ड़		ੜ		ড়	ଢ					
ḍha	ढ	ढ	ਢ	ढ	ঢ়	ଢ		ഢ	ഢ	ढ	ඨ
ṛha	ढ़				ড়়	ঢ়়					
ṇa	ण	ण	ਣ	ṇ	ণ	ṇ	ண	ണ	ಣ	ణ	ඬ
ṇḍa											
ta	त	त	ਤ	ત	ত	ତ	த	ത	ತ	త	ට
tha	थ	थ	ਥ	થ	থ	ଥ		ഥ	ഥ	थ	ට
da	द	द	ਦ	દ	দ	ଦ		ദ	ദ	द	ට
dha	ध	ध	ਧ	ધ	ধ	ଧ		ധ	ധ	ध	ට
na	न	न	ਨ	ન	ন	ନ	ந	ന	ನ	న	ට
ṇḍa											

Latin	Dev.	San.	Gur.	Guj.	Ben.	Ori.	Tam.	Mal.	Kan.	Tel.	Sin.
pa	प	प	ਪ	પ	প	ପ	॥	പ	ಪ	ప	ප
pha	फ	फ	ਫ	ફ	ফ	ଫ		ഫ	ಫ	ఫ	ආ
ba	ब	ब	ਬ	બ	ব	ବ		ബ	ಬ	బ	බ
bha	भ	भ	ਭ	ભ	ভ	ଭ		ഭ	ಬ	భ	භ
ma	म	म	ਮ	મ	ম	ମ	॥	മ	ಮ	మ	ම
m̐ba									𑌕		𑌕
ra	र						॥	റ	ര	ర	ර
ta								ത			
na	न						॥	ന			න
la	ल						॥	ല	ല	ల	ල
ya	य	य	ਯ	ય	য	ଯ	॥	യ	ಯ	య	ය
ya	य़				য়	ୟ					
ra	र	र	ਰ	ર	র	ର	॥	റ	ര	ర	ර
ṛ	ठ	ठ									
la	ल	ल	ਲ	લ	ল	ଲ	॥	ല	ല	ల	ල
ḷa	ळ	ळ	ਲ਼	ળ		ଲ਼	॥	ള	ള	ఱ	ඌ
va	व	व	ਵ	વ	ব	ବ	॥	വ	വ	వ	ඹ
śa	श	श	ਸ਼	શ	শ	ଶ	॥	ശ	ശ	ష	ශ
ṣa	ष	ष		ષ	ষ	ଷ	॥	ഷ	ഷ	ఱ	ඳ
sa	स	स	ਸ	સ	স	ସ	॥	സ	സ	స	ස
ha	ह	ह	ਹ	હ	হ	ହ	॥	ഹ	ഹ	హ	හ
'	ऽ	ऽ		ઽ	ং	ଽ		ഃ	ഃ	ః	
qa	क		ਕ	ક	ক						
kha	ख		ਖ	ખ	খ						
ḡa	ग़		ਗ਼	ગ	গ						
za	ज़		ਜ਼	જ	জ				𑌕		
fa	फ़		ਫ਼	ફ	ফ				𑌕		ඳ
wa					ব	𑌕					

## 2.1 Special conventions

- A colon (:) can be used to break a ligature or diphthong: ब्ह b:ha vs. भ bha, प्रउ prau vs. प्रौ prau.
- Use U+007C VERTICAL LINE for dandas: ।, ॥.
- Use U+00B0 DEGREE SIGN for U+0970 DEVANAGARI ABBREVIATION SIGN

## 2.2 Special extensions for Sanskrit

- ṛ can be used instead of ॠ for vocalic r.
- ḷ can be used immediately following a k for the vocalic l, thus the root क्लृप् can be typed “klp” rather than “kḷp”.
- ḹ is always recognized as a long vocalic l.
- A diaeresis can also be to break a diphthong: प्रउ prau̐ vs. प्रौ prau.
- ṁ is a synonym for ँ.

- ç is a synonym for ś.
- āi and āu are synonyms for ai and au, respectively.

## References

- [1] Anthony P. Stone. "Transliteration of Indic Scripts: How to use ISO 15919." <http://homepage.ntlworld.com/stone-catend/trind.htm>
- [2] Arthur A. Macdonell, *A Sanskrit-English Dictionary*. Longmans, Green, and Co., London, 1893.
- [3] Arthur A. MacDonell, *A Sanskrit Grammar for Students*, third edition. Oxford University Press, 1927.
- [4] The Unicode Consortium. *The Unicode Standard: Version 7.0—Core Specification*. <http://www.unicode.org/versions/Unicode7.0.0/UnicodeStandard-7.0.pdf> (modification date 09/26/2014 03:27:41 PM; md5sum aa6c89d24d7075f93259069fae273251).
- [5] William Dwight Whitney, *Sanskrit Grammar: Including both the Classical Language, and the older Dialects, of Veda and Brahmana*. Harvard University Press, fourteenth printing, 1975.

## 3 TECKit header

```

1 LHSName "Unicode/ISO 15919"
2 <deva&!san>RHSName "Unicode/Devanagari"
3 <deva & san>RHSName "Unicode/Sanskrit"
4 <beng>RHSName "Unicode/Bengali"
5 <gujr>RHSName "Unicode/Gujarati"
6 <guru>RHSName "Unicode/Gurmukhi"
7 <knda>RHSName "Unicode/Kannada"
8 <mlym>RHSName "Unicode/Malayalam"
9 <orya>RHSName "Unicode/Oriya"
10 <sinh>RHSName "Unicode/Sinhala"
11 <taml>RHSName "Unicode/Tamil"
12 <telu>RHSName "Unicode/Telugu"
13 Version "0.01 beta"

```

## 4 Overview

Conceptually, the mapping process is divided into two main phases, the first of which translates user input into an internal normalized form and the second of which translates that normalized form into the desired output alphabet. So, we are actually dealing with three different sets of characters:

1. User input: The specific Latin-based Unicode characters supplied by the user. This can be a mix of upper and lower case and precomposed or fully decomposed characters.

2. Normalized input: During Phase 1, the user input is converted into a normalized form, based on, but not identical with, Unicode Normalization Form D. In broad terms, all characters are downcased, and combining marks representing long vowels or tone marks are separated from their base characters, while combining marks designating distinct consonants are combined with their base characters. In addition, alternative representations are merged and in some cases context-sensitive disambiguation is performed.
3. Output: During Phase 2, the normalized input is converted into the desired output script.

## 5 Normalized input characters

### 5.1 Vowels and diphthongs

```

14 Define a      U+0061
15 Define ae     U+00E6
16 Define b      U+0062
17 Define c      U+0063
18 Define d      U+0064
19 Define e      U+0065
20 Define f      U+0066
21 Define g      U+0067
22 Define h      U+0068
23 Define i      U+0069
24 Define j      U+006A
25 Define k      U+006B
26 Define l      U+006C
27 Define m      U+006D
28 Define n      U+006E
29 Define o      U+006F
30 Define p      U+0070
31 Define q      U+0071
32 Define r      U+0072
33 Define s      U+0073
34 Define t      U+0074
35 Define u      U+0075
36 Define v      U+0076
37 Define w      U+0077
38 ; Define x    U+0078
39 Define y      U+0079
40 Define z      U+007A

```

It will be convenient to have single-character representations for vocalic r and vocalic l, but Unicode doesn't have precomposed characters for l or r with a ring below. So, for convenience, we will map them to the PUA in our internal representation.

```

41 Define ri     U+E000 ; PUA
42 Define lry    U+E001 ; PUA

```

## Symbols

```
43Define apos      U+0027
44; Define colon    U+003A ; predefined
45Define degree     U+00B0
46Define bar        U+007C
47
48Define lbrack     U+005B
49Define rbrack     U+005D
```

## 5.2 Combining accents

Define short names for the various combining accents we need.<sup>3</sup>

```
50Define grave      U+0300
51Define acute      U+0301
52Define circumflex U+0302
53Define tildeabove U+0303
54Define macronabove U+0304
55Define breveabove U+0306
56Define dotabove   U+0307
57Define dieresis   U+0308
58Define vert       U+030D
59Define CANDRABINDU U+0310
60Define dotbelow   U+0323
61Define ringbelow  U+0325
62Define ccedilla   U+0327
63Define hookbelow  U+0328
64Define brevebelow U+032E
65Define macronbelow U+0331
```

## 6 Output Characters

Now define the output space.

### 6.1 Script-independent characters

```
66Define danda      U+0964
67Define dbldanda   U+0965
68
69Define zwnj        U+200C ; ZERO WIDTH NON-JOINER
70Define zwj         U+200D ; ZERO WIDTH JOINER
71Define hiatus      U+0020 ; SPACE
72Define nbsp        U+00A0 ; NO-BREAK SPACE
73
74Define NOTDEF      U+FFFD
```

### 6.2 Devanagari

---

<sup>3</sup>The names tilde, macron, breve, dot\_above, and diaeresis are predefined the compiler to refer to the spacing variants of those accents, so I've had to choose alternate names.



```

75 ⟨*deva⟩
76 ; Define altcandrabindu U+0900
77 Define candrabindu U+0901
78 Define anusvara U+0902
79 Define visarga U+0903
80
81 ; Define ashort U+0904
82 Define asvara U+0905
83 Define Asvara U+0906
84
85 Define isvara U+0907
86 Define Isvara U+0908
87
88 Define usvara U+0909
89 Define Usvara U+090A
90
91 Define rsvara U+090B
92 Define lsvara U+090C
93
94 ⟨*san⟩
95 Define esvara U+090F
96 Define Esvara U+090F
97 Define ecandra U+090F
98 ⟨/san⟩
99 ⟨*!san⟩
100 Define esvara U+090E
101 Define Esvara U+090F
102 Define ecandra U+090D
103 ⟨!/san⟩
104
105 Define aisvara U+0910
106
107 ⟨*san⟩
108 Define osvara U+0913
109 Define Osvara U+0913
110 Define ocandra U+0913
111 ⟨/san⟩
112 ⟨*!san⟩
113 Define osvara U+0912
114 Define Osvara U+0913
115 Define ocandra U+0911
116 ⟨!/san⟩
117
118 Define ausvara U+0914
119
120 Define ka U+0915
121 Define kha U+0916
122 Define ga U+0917
123 Define gha U+0918
124 Define nga U+0919

```

125		
126	Define ca	U+091A
127	Define cha	U+091B
128	Define ja	U+091C
129	Define jha	U+091D
130	Define nya	U+091E
131		
132	Define tta	U+091F
133	Define ttha	U+0920
134	Define dda	U+0921
135	Define ddha	U+0922
136	Define nna	U+0923
137		
138	Define ta	U+0924
139	Define tha	U+0925
140	Define da	U+0926
141	Define dha	U+0927
142	Define na	U+0928
143		
144	Define nnna	U+0929
145		
146	Define pa	U+092A
147	Define pha	U+092B
148	Define ba	U+092C
149	Define bha	U+092D
150	Define ma	U+092E
151		
152	Define ya	U+092F
153	Define ra	U+0930
154	Define rra	U+0931
155	Define la	U+0932
156	Define lla	U+0933
157	Define llla	U+0934
158	Define va	U+0935
159		
160	Define sha	U+0936
161	Define ssa	U+0937
162	Define sa	U+0938
163	Define ha	U+0939
164		
165	; Define oematra	U+093A
166	; Define OEmatra	U+093B
167		
168	Define nukta	U+093C
169	Define avagraha	U+093D
170		
171	Define Amatra	U+093E
172		
173	Define imatra	U+093F
174	Define Imatra	U+0940

175  
176 Define umatra U+0941  
177 Define Umatra U+0942  
178  
179 Define rmatra U+0943  
180 Define Rmatra U+0944  
181  
182 ⟨\*san⟩  
183 Define ematra U+0947  
184 Define Ematra U+0947  
185 Define ecandramatra U+0947  
186 ⟨/san⟩  
187 ⟨\*!san⟩  
188 Define ematra U+0946  
189 Define Ematra U+0947  
190 Define ecandramatra U+0945  
191 ⟨/!san⟩  
192  
193 Define aimatra U+0948  
194  
195 ⟨\*san⟩  
196 Define omatra U+094B  
197 Define Omatra U+094B  
198 Define ocandramatra U+094B  
199 ⟨/san⟩  
200 ⟨\*!san⟩  
201 Define omatra U+094A  
202 Define Omatra U+094B  
203 Define ocandramatra U+0949  
204 ⟨/!san⟩  
205  
206 Define aumatra U+094C  
207  
208 Define virama U+094D  
209 ; Define prishthamatra U+094E  
210 ; Define awmatra U+094F  
211 ; Define om U+0950  
212  
213 Define udatta U+0951  
214 Define anudatta U+0952  
215  
216 ; Define devgrave U+0953  
217 ; Define devacute U+0954  
218  
219 ; Define Ecandra U+0955  
220  
221 ; Define uematra U+0956  
222 ; Define UEmatra U+0957  
223  
224 ⟨\*nukta⟩

225	Define qa	U+0958
226	Define khha	U+0959
227	Define ghha	U+095A
228	Define za	U+095B
229	Define dddha	U+095C
230	Define rha	U+095D
231	Define fa	U+095E
232	Define yya	U+095F
233	</nukta>	
234		
235	<!*nukta>	
236	Define qa	NOTDEF
237	Define khha	NOTDEF
238	Define ghha	NOTDEF
239	Define za	NOTDEF
240	Define dddha	NOTDEF
241	Define rha	NOTDEF
242	Define fa	NOTDEF
243	Define yya	NOTDEF
244	</!nukta>	
245		
246	Define Rsvara	U+0960
247	Define Lsvara	U+0961
248	Define lmatra	U+0962
249	Define Lmatra	U+0963
250		
251	Define zero	U+0966
252	Define one	U+0967
253	Define two	U+0968
254	Define three	U+0969
255	Define four	U+096A
256	Define five	U+096B
257	Define six	U+096C
258	Define seven	U+096D
259	Define eight	U+096E
260	Define nine	U+096F
261		
262	Define abbrev	U+0970
263		
264	; Define devdot	U+0971
265	; Define acandra	U+0972
266	; Define oesvara	U+0973
267	; Define OEsvara	U+0974
268	; Define awsvara	U+0975
269	; Define uesvara	U+0976
270	; Define UEsvara	U+0977
271	; Define ddamarwari	U+0978
272	; Define zha	U+0979
273	; Define jjya	U+097a
274	; Define gga	U+097b

```

275; Define jja          U+097c
276; Define devglottal U+097d
277; Define ddda        U+097e
278; Define bba         U+097f
279
280⟨san⟩Define svarita_k_i U+1CD7
281
282Define jihvamuliya U+1CF5
283Define upadhmaniya U+1CF6
284
285Define extravyanjana dddha fa ghha khha lla llla qa rha yya za
286⟨/deva⟩

```

### 6.3 Bengali

```

287⟨*beng⟩
288; Define anji        U+0980
289Define candrabindu U+0981
290Define anusvara      U+0982
291Define visarga       U+0983
292
293Define asvara        U+0985
294Define Asvara       U+0986
295
296Define isvara        U+0987
297Define Isvara       U+0988
298
299Define usvara        U+0989
300Define Usvara       U+098A
301
302Define rsvara        U+098B
303Define lsvara        U+098C
304
305Define esvara        U+098F
306Define Esvara       U+098F
307Define ecandra       U+098F
308Define aisvara       U+0990
309
310Define osvara        U+0993
311Define Osvara       U+0993
312Define ocandra       U+0993
313Define ausvara       U+0994
314
315Define ka            U+0995
316Define kha          U+0996
317Define ga           U+0997
318Define gha          U+0998
319Define nga          U+0999
320
321Define ca            U+099A

```

322	Define cha	U+099B
323	Define ja	U+099C
324	Define jha	U+099D
325	Define nya	U+099E
326		
327	Define tta	U+099F
328	Define ttha	U+09A0
329	Define dda	U+09A1
330	Define ddha	U+09A2
331	Define nna	U+09A3
332		
333	Define ta	U+09A4
334	Define tha	U+09A5
335	Define da	U+09A6
336	Define dha	U+09A7
337	Define na	U+09A8
338		
339	Define pa	U+09AA
340	Define pha	U+09AB
341	Define ba	U+09AC
342	Define bha	U+09AD
343	Define ma	U+09AE
344		
345	Define ya	U+09AF
346	Define ra	U+09B0
347	Define la	U+09B2
348	Define lla	NOTDEF
349	Define va	U+09F1
350		
351	Define sha	U+09B6
352	Define ssa	U+09B7
353	Define sa	U+09B8
354	Define ha	U+09B9
355		
356	Define nukta	U+09BC
357	Define avagraha	U+09BD
358		
359	Define Amatra	U+09BE
360		
361	Define imatra	U+09BF
362	Define Imatra	U+09C0
363		
364	Define umatra	U+09C1
365	Define Umatra	U+09C2
366		
367	Define rmatra	U+09C3
368	Define Rmatra	U+09C4
369		
370	Define ematra	U+09C7
371	Define Ematra	U+09C7

```

372 Define ecandramatra U+09C7
373 Define aimatra      U+09C8
374
375 Define omatra       U+09CB
376 Define Omatra       U+09CB
377 Define ocandramatra U+09CB
378 Define aumatra      U+09CC
379
380 Define virama       U+09CD
381
382 Define khandata     U+09CE
383
384 Define udatta       NOTDEF
385 Define anudatta     NOTDEF
386
387 Define aulength     U+09D7
388
389 Define dddha        U+09DC
390
391 Define qa           U+E002 ; PUA
392 Define khha        U+E005 ; PUA
393 Define ghha        U+E007 ; PUA
394 Define za          U+E008 ; PUA
395 Define rra         NOTDEF
396 Define rha         U+09DD
397 Define fa          U+E009 ; PUA
398 Define yya         U+09DF
399
400 Define Rsvara      U+09E0
401 Define Lsvara      U+09E1
402 Define lmatra      U+09E2
403 Define Lmatra      U+09E3
404
405 Define zero        U+09E6
406 Define one         U+09E7
407 Define two         U+09E8
408 Define three       U+09E9
409 Define four        U+09EA
410 Define five        U+09EB
411 Define six         U+09EC
412 Define seven       U+09ED
413 Define eight       U+09EE
414 Define nine        U+09EF
415
416 Define wa          U+E010 ; PUA
417
418 ; Define ramd U+09F0
419 ; Define rald U+09F1
420 ; Define rupeemark U+09F2
421 ; Define rupeesign U+09F3

```

422; Define curnumone U+09F4  
 423; Define curnumtwo U+09F5  
 424; Define curnumthree U+09F6  
 425; Define curnumfour U+09F7  
 426; Define curnumoneless U+09F8  
 427; Define curdenosixteen U+09F9  
 428; Define issnar U+09FA  
 429; Define ganda mark U+09FB  
 430  
 431Define abbrev U+0970  
 432  
 433Define extravyanjana dddha fa ghha khha qa rra va wa za  
 434</beng>

## 6.4 Gujarati

435<\*gujr>  
 436Define candrabindu U+0A81  
 437Define anusvara U+0A82  
 438Define visarga U+0A83  
 439  
 440Define asvara U+0A85  
 441Define Asvara U+0A86  
 442  
 443Define isvara U+0A87  
 444Define Isvara U+0A88  
 445  
 446Define usvara U+0A89  
 447Define Usvara U+0A8A  
 448  
 449Define rsvara U+0A8B  
 450Define lsvara U+0A8C  
 451  
 452Define esvara U+0A8F  
 453Define Esvara U+0A8F  
 454Define ecandra U+0A8D  
 455Define aisvara U+0A90  
 456  
 457Define osvara U+0A93  
 458Define Osvara U+0A93  
 459Define ocandra U+0A91  
 460Define ausvara U+0A94  
 461  
 462Define ka U+0A95  
 463Define kha U+0A96  
 464Define ga U+0A97  
 465Define gha U+0A98  
 466Define nga U+0A99  
 467  
 468Define ca U+0A9A



469	Define cha	U+0A9B
470	Define ja	U+0A9C
471	Define jha	U+0A9D
472	Define nya	U+0A9E
473		
474	Define tta	U+0A9F
475	Define ttha	U+0AA0
476	Define dda	U+0AA1
477	Define ddha	U+0AA2
478	Define nna	U+0AA3
479		
480	Define ta	U+0AA4
481	Define tha	U+0AA5
482	Define da	U+0AA6
483	Define dha	U+0AA7
484	Define na	U+0AA8
485		
486	Define pa	U+0AAA
487	Define pha	U+0AAB
488	Define ba	U+0AAC
489	Define bha	U+0AAD
490	Define ma	U+0AAE
491		
492	Define ya	U+0AAF
493	Define ra	U+0AB0
494	Define la	U+0AB2
495	Define lla	U+0AB3
496	Define va	U+0AB5
497		
498	Define sha	U+0AB6
499	Define ssa	U+0AB7
500	Define sa	U+0AB8
501	Define ha	U+0AB9
502		
503	Define nukta	U+0ABC
504	Define avagraha	U+0ABD
505		
506	Define Amatra	U+0ABE
507		
508	Define imatra	U+0ABF
509	Define Imatra	U+0AC0
510		
511	Define umatra	U+0AC1
512	Define Umatra	U+0AC2
513		
514	Define rmatra	U+0AC3
515	Define Rmatra	U+0AC4
516		
517	Define ematra	U+0AC7
518	Define Ematra	U+0AC7

```

519 Define ecandramatra U+0AC5
520 Define aimatra      U+0AC8
521
522 Define omatra       U+0ACB
523 Define Omatra       U+0ACB
524 Define ocandramatra U+0AC9
525 Define aumatra      U+0ACC
526
527 Define virama       U+0ACD
528 ; Define om        U+0ADO
529
530 Define udatta       NOTDEF
531 Define anudatta     NOTDEF
532
533 Define qa           U+E003 ; PUA
534 Define khha        U+E006 ; PUA
535 Define ghha        U+E007 ; PUA
536 Define za          U+E008 ; PUA
537 Define rra         NOTDEF
538 Define rha         NOTDEF
539 Define fa           U+E009 ; PUA
540 Define yya         NOTDEF
541
542 Define Rsvara       U+0AE0
543 Define Lsvara       U+0AE1
544 Define lmatra       U+0AE2
545 Define Lmatra       U+0AE3
546
547 Define zero         U+0AE6
548 Define one          U+0AE7
549 Define two          U+0AE8
550 Define three        U+0AE9
551 Define four         U+0AEA
552 Define five         U+0AEB
553 Define six          U+0AEC
554 Define seven        U+0AED
555 Define eight        U+0AEE
556 Define nine         U+0AEF
557
558 Define abbrev       U+0AF0
559
560 ; Define rupeemark NOTDEF
561 ; Define rupeesign U+0AF1
562
563 Define extravyanjana fa ghha khha qa za
564 </gujr>

```

## 6.5 Gurmukhi

```

565 <*guru>

```

566	Define candrabindu	U+0A01
567	Define anusvara	U+0A02
568	Define visarga	U+0A03
569		
570	Define asvara	U+0A05
571	Define Asvara	U+0A06
572		
573	Define isvara	U+0A07
574	Define Isvara	U+0A08
575		
576	Define usvara	U+0A09
577	Define Usvara	U+0A0A
578		
579	Define rsvara	NOTDEF
580	Define Rsvara	NOTDEF
581		
582	Define lsvara	NOTDEF
583	Define Lsvara	NOTDEF
584		
585	Define esvara	U+0A0F
586	Define Esvara	U+0A0F
587	Define ecandra	U+0A0F
588	Define aisvara	U+0A10
589		
590	Define osvara	U+0A13
591	Define Osvara	U+0A13
592	Define ocandra	U+0A13
593	Define ausvara	U+0A14
594		
595	Define ka	U+0A15
596	Define kha	U+0A16
597	Define ga	U+0A17
598	Define gha	U+0A18
599	Define nga	U+0A19
600		
601	Define ca	U+0A1A
602	Define cha	U+0A1B
603	Define ja	U+0A1C
604	Define jha	U+0A1D
605	Define nya	U+0A1E
606		
607	Define tta	U+0A1F
608	Define ttha	U+0A20
609	Define dda	U+0A21
610	Define ddha	U+0A22
611	Define nna	U+0A23
612		
613	Define ta	U+0A24
614	Define tha	U+0A25
615	Define da	U+0A26

616	Define dha	U+0A27
617	Define na	U+0A28
618		
619	Define pa	U+0A2A
620	Define pha	U+0A2B
621	Define ba	U+0A2C
622	Define bha	U+0A2D
623	Define ma	U+0A2E
624		
625	Define ya	U+0A2F
626	Define ra	U+0A30
627	Define la	U+0A32
628	Define lla	U+0A33
629	Define va	U+0A35
630		
631	Define sha	U+0A36
632	Define ssa	NOTDEF
633	Define sa	U+0A38
634	Define ha	U+0A39
635		
636	Define nukta	U+0A3C
637	Define avagraha	NOTDEF
638		
639	Define Amatra	U+0A3E
640		
641	Define imatra	U+0A3F
642	Define Imatra	U+0A40
643		
644	Define umatra	U+0A41
645	Define Umatra	U+0A42
646		
647	Define rmatra	NOTDEF
648	Define Rmatra	NOTDEF
649		
650	Define ematra	U+0A47
651	Define Ematra	U+0A47
652	Define ecandramatra	U+0A47
653	Define aimatra	U+0A48
654		
655	Define omatra	U+0A4B
656	Define Omatra	U+0A4B
657	Define ocandramatra	U+0A4B
658	Define aumatra	U+0A4C
659		
660	Define virama	U+0A4D
661		
662	Define udatta	U+0A51
663	Define anudatta	NOTDEF
664		
665	Define qa	U+E004 ; PUA

666	Define khha	U+0A59
667	Define ghha	U+0A5A
668	Define za	U+0A5B
669	Define rra	U+0A5C
670	Define rha	NOTDEF
671	Define fa	U+0A5E
672	Define yya	NOTDEF
673		
674	Define lmatra	NOTDEF
675	Define Lmatra	NOTDEF
676		
677	Define zero	U+0A66
678	Define one	U+0A67
679	Define two	U+0A68
680	Define three	U+0A69
681	Define four	U+0A6A
682	Define five	U+0A6B
683	Define six	U+0A6C
684	Define seven	U+0A6D
685	Define eight	U+0A6E
686	Define nine	U+0A6F
687		
688	Define abbrev	NOTDEF
689		
690	Define tippi	U+0A70
691		
692	Define addak	U+0A71
693	Define iri	U+0A72
694	Define ura	U+0A73
695	Define ek_onkar	U+0A74
696	Define yakash	U+0A75
697		
698	Define extravyanjana fa ghha khha qa rra za	
699	</guru>	

## 6.6 Kannada

700	<*knda>	
701	Define candrabindu	U+0C81
702	Define anusvara	U+0C82
703	Define visarga	U+0C83
704		
705	Define asvara	U+0C85
706	Define Asvara	U+0C86
707		
708	Define isvara	U+0C87
709	Define Isvara	U+0C88
710		
711	Define usvara	U+0C89
712	Define Usvara	U+0C8A

713		
714	Define rsvara	U+0C8B
715		
716	Define lsvara	U+0C8C
717		
718	Define esvara	U+0C8E
719	Define Esvara	U+0C8F
720	Define ecandra	U+0C8F
721	Define aisvara	U+0C90
722		
723	Define osvara	U+0C92
724	Define Osvara	U+0C93
725	Define ocandra	U+0C93
726	Define ausvara	U+0C94
727		
728	Define ka	U+0C95
729	Define kha	U+0C96
730	Define ga	U+0C97
731	Define gha	U+0C98
732	Define nga	U+0C99
733		
734	Define ca	U+0C9A
735	Define cha	U+0C9B
736	Define ja	U+0C9C
737	Define jha	U+0C9D
738	Define nya	U+0C9E
739		
740	Define tta	U+0C9F
741	Define ttha	U+0CA0
742	Define dda	U+0CA1
743	Define ddha	U+0CA2
744	Define nna	U+0CA3
745		
746	Define ta	U+0CA4
747	Define tha	U+0CA5
748	Define da	U+0CA6
749	Define dha	U+0CA7
750	Define na	U+0CA8
751		
752	Define pa	U+0CAA
753	Define pha	U+0CAB
754	Define ba	U+0CAC
755	Define bha	U+0CAD
756	Define ma	U+0CAE
757		
758	Define ya	U+0CAF
759	Define ra	U+0CB0
760	Define rra	U+0CB1
761	Define la	U+0CB2
762	Define lla	U+0CB3

```

763 Define va          U+0CB5
764
765 Define llla        U+0CDE
766
767 Define sha         U+0CB6
768 Define ssa        U+0CB7
769 Define sa         U+0CB8
770 Define ha         U+0CB9
771
772 Define nukta       U+0CBC
773 Define avagraha    U+0CBD
774
775 Define Amatra      U+0CBE
776
777 Define imatra      U+0CBF
778 Define Imatra      U+0CC0
779
780 Define umatra      U+0CC1
781 Define Umatra      U+0CC2
782
783 Define rmatra      U+0CC3
784 Define Rmatra      U+0CC4
785
786 Define ematra      U+0CC6
787 Define Ematra      U+0CC7
788 Define ecandramatra U+0CC7
789 Define aimatra     U+0CC8
790
791 Define omatra      U+0CCA
792 Define Omatra      U+0CCB
793 Define ocandramatra U+0CCB
794 Define aumatra     U+0CCC
795
796 Define udatta      NOTDEF
797 Define anudatta    NOTDEF
798
799 Define virama       U+0CCD
800 ; Define length_mark U+0CD5
801 ; Define ai_length_mark U+0CD6
802
803 Define za          U+E008 ; PUA
804 Define fa          U+E014 ; PUA
805
806 Define Rsvara      U+0CE0
807 Define Lsvara      U+0CE1
808 Define lmatra      U+0CE2
809 Define Lmatra      U+0CE3
810
811 Define zero        U+0CE6
812 Define one         U+0CE7

```

813 Define two U+0CE8  
 814 Define three U+0CE9  
 815 Define four U+0CEA  
 816 Define five U+0CEB  
 817 Define six U+0CEC  
 818 Define seven U+0CED  
 819 Define eight U+0CEE  
 820 Define nine U+0CEF  
 821  
 822 Define abbrev NOTDEF  
 823  
 824 Define jihvamuliya U+0CF1  
 825 Define upadhmaniya U+0CF2  
 826  
 827 Define extravyanjana fa llla za  
 828  $\langle /knda \rangle$

## 6.7 Malayalam

829  $\langle *mlym \rangle$   
 830 Define candrabindu U+0D01  
 831 Define anusvara U+0D02  
 832 Define visarga U+0D03  
 833  
 834 Define asvara U+0D05  
 835 Define Asvara U+0D06  
 836  
 837 Define isvara U+0D07  
 838 Define Isvara U+0D08  
 839  
 840 Define usvara U+0D09  
 841 Define Usvara U+0D0A  
 842  
 843 Define rsvara U+0D0B  
 844 Define lsvara U+0D0C  
 845  
 846 Define esvara U+0D0E  
 847 Define Esvara U+0D0F  
 848 Define ecandra U+0D0F  
 849 Define aisvara U+0D10  
 850  
 851 Define osvara U+0D12  
 852 Define Osvara U+0D13  
 853 Define ocandra U+0D13  
 854 Define ausvara U+0D14  
 855  
 856 Define ka U+0D15  
 857 Define kha U+0D16  
 858 Define ga U+0D17  
 859 Define gha U+0D18



860	Define nga	U+0D19
861		
862	Define ca	U+0D1A
863	Define cha	U+0D1B
864	Define ja	U+0D1C
865	Define jha	U+0D1D
866	Define nya	U+0D1E
867		
868	Define tta	U+0D1F
869	Define ttha	U+0D20
870	Define dda	U+0D21
871	Define ddha	U+0D22
872	Define nna	U+0D23
873		
874	Define ta	U+0D24
875	Define tha	U+0D25
876	Define da	U+0D26
877	Define dha	U+0D27
878	Define na	U+0D28
879		
880	Define nnna	U+0D29
881		
882	Define pa	U+0D2A
883	Define pha	U+0D2B
884	Define ba	U+0D2C
885	Define bha	U+0D2D
886	Define ma	U+0D2E
887		
888	Define ya	U+0D2F
889	Define ra	U+0D30
890	Define rra	U+0D31
891	Define la	U+0D32
892	Define lla	U+0D33
893	Define llla	U+0D34
894	Define va	U+0D35
895		
896	Define sha	U+0D36
897	Define ssa	U+0D37
898	Define sa	U+0D38
899	Define ha	U+0D39
900		
901	Define ttta	U+0D3A
902		
903	Define avagraha	U+0D3D
904		
905	Define Amatra	U+0D3E
906		
907	Define imatra	U+0D3F
908	Define Imatra	U+0D40
909		

```

910 Define umatra      U+0D41
911 Define Umatra      U+0D42
912
913 Define rmatra      U+0D43
914 Define Rmatra      U+0D44
915
916 Define ematra      U+0D46
917 Define Ematra      U+0D47
918 Define ecandramatra U+0D47
919 Define aimatra     U+0D48
920
921 Define omatra      U+0D4A
922 Define Omatra      U+0D4B
923 Define ocandramatra U+0D4B
924 Define aumatra     U+0D4C
925
926 Define virama      U+0D4D
927
928 Define dot_reph     U+0D4E
929
930 Define udatta      U+0D51
931 Define anudatta     U+0D52
932
933 ; Define au_length_mark U+0D57
934
935 Define fa           NOTDEF
936
937 Define Rsvara      U+0D60
938 Define Lsvara      U+0D61
939 Define lmatra      U+0D62
940 Define Lmatra      U+0D63
941
942 Define zero        U+0D66
943 Define one         U+0D67
944 Define two         U+0D68
945 Define three       U+0D69
946 Define four        U+0D6A
947 Define five        U+0D6B
948 Define six         U+0D6C
949 Define seven       U+0D6D
950 Define eight       U+0D6E
951 Define nine        U+0D6F
952
953 ; Define ten        U+0D70
954 ; Define hundred    U+0D71
955 ; Define thousand   U+0D72
956
957 ; Define quarter    U+0D73
958 ; Define half       U+0D74
959 ; Define three_quarters U+0D75

```

```

960
961; Define date_mark    U+0D79
962
963Define abbrev         NOTDEF
964
965Define chillu_nn U+0D7A
966Define chillu_n  U+0D7B
967Define chillu_rr U+0D7C
968Define chillu_l  U+0D7D
969Define chillu_ll U+0D7E
970Define chillu_k  U+0D7F
971
972Define extravyanjana llla ttta
973</mlym>

```

## 6.8 Oriya

```

974<*orya>
975Define candrabindu U+0B01
976Define anusvara    U+0B02
977Define visarga      U+0B03
978
979Define asvara       U+0B05
980Define Asvara       U+0B06
981
982Define isvara       U+0B07
983Define Isvara       U+0B08
984
985Define usvara       U+0B09
986Define Usvara       U+0B0A
987
988Define rsvara       U+0B0B
989Define lsvara       U+0B0C
990
991Define esvara       U+0B0F
992Define Esvara       U+0B0F
993Define ecandra      U+0B0F
994Define aisvara      U+0B10
995
996Define osvara       U+0B13
997Define Osvara       U+0B13
998Define ocandra      U+0B13
999Define ausvara      U+0B14
1000
1001Define ka          U+0B15
1002Define kha         U+0B16
1003Define ga          U+0B17
1004Define gha         U+0B18
1005Define nga         U+0B19
1006

```

1007	Define	ca	U+0B1A
1008	Define	cha	U+0B1B
1009	Define	ja	U+0B1C
1010	Define	jha	U+0B1D
1011	Define	nya	U+0B1E
1012			
1013	Define	tta	U+0B1F
1014	Define	ttha	U+0B20
1015	Define	dda	U+0B21
1016	Define	ddha	U+0B22
1017	Define	nna	U+0B23
1018			
1019	Define	ta	U+0B24
1020	Define	tha	U+0B25
1021	Define	da	U+0B26
1022	Define	dha	U+0B27
1023	Define	na	U+0B28
1024			
1025	Define	pa	U+0B2A
1026	Define	pha	U+0B2B
1027	Define	ba	U+0B2C
1028	Define	bha	U+0B2D
1029	Define	ma	U+0B2E
1030			
1031	Define	ya	U+0B2F
1032	Define	ra	U+0B30
1033	Define	la	U+0B32
1034	Define	lla	U+0B33
1035	Define	va	U+0B35
1036			
1037	Define	sha	U+0B36
1038	Define	ssa	U+0B37
1039	Define	sa	U+0B38
1040	Define	ha	U+0B39
1041			
1042	Define	nukta	U+0B3C
1043	Define	avagraha	U+0B3D
1044			
1045	Define	Amatra	U+0B3E
1046			
1047	Define	imatra	U+0B3F
1048	Define	Imatra	U+0B40
1049			
1050	Define	umatra	U+0B41
1051	Define	Umatra	U+0B42
1052			
1053	Define	rmatra	U+0B43
1054	Define	Rmatra	U+0B44
1055			
1056	Define	ematra	U+0B47

```

1057 Define Ematra      U+0B47
1058 Define ecandramatra U+0B47
1059 Define aimatra      U+0B48
1060
1061 Define omatra        U+0B4B
1062 Define Omatra        U+0B4B
1063 Define ocandramatra U+0B4B
1064 Define aumatra      U+0B4C
1065
1066 Define virama        U+0B4D
1067
1068 Define udatta        NOTDEF
1069 Define anudatta      NOTDEF
1070
1071 ; Define ai_length_mark U+0B56
1072 ; Define au_length_mark U+0B57
1073
1074 Define rra           U+0B5C
1075 Define rha           U+0B5D
1076 Define yya          U+0B5F
1077
1078 Define fa            NOTDEF
1079
1080 Define Rsvara        U+0B60
1081 Define Lsvara        U+0B61
1082 Define lmatra       U+0B62
1083 Define Lmatra       U+0B63
1084
1085 Define zero          U+0B66
1086 Define one           U+0B67
1087 Define two           U+0B68
1088 Define three         U+0B69
1089 Define four         U+0B6A
1090 Define five          U+0B6B
1091 Define six           U+0B6C
1092 Define seven         U+0B6D
1093 Define eight         U+0B6E
1094 Define nine         U+0B6F
1095
1096 Define abbrev        NOTDEF
1097
1098 Define isshar        U+0B70
1099 Define wa            U+0B71
1100 Define quarter       U+0B72
1101 Define half          U+0B73
1102 Define three_quarters U+0B74
1103 Define sixteenth     U+0B75
1104 Define eighth        U+0B76
1105 Define three_sixteenths U+0B77
1106

```

1107 Define extravyanjana rha rra wa yya  
 1108 ⟨/orya⟩

## 6.9 Sinhala

1109 ⟨\*sinh⟩  
 1110 Define candrabindu NOTDEF  
 1111 Define anusvara U+0D82  
 1112 Define visarga U+0D83  
 1113  
 1114 Define asvara U+0D85  
 1115 Define Asvara U+0D86  
 1116  
 1117 Define aesvara U+0D87  
 1118 Define AEsvara U+0D88  
 1119  
 1120 Define isvara U+0D89  
 1121 Define Isvara U+0D8A  
 1122  
 1123 Define usvara U+0D8B  
 1124 Define Usvara U+0D8C  
 1125  
 1126 Define rsvara U+0D8D  
 1127 Define Rsvara U+0D8E  
 1128  
 1129 Define lsvara U+0D8F  
 1130 Define Lsvara U+0D90  
 1131  
 1132 Define esvara U+0D91  
 1133 Define Esvara U+0D92  
 1134 Define ecandra U+0D92  
 1135 Define aisvara U+0D93  
 1136  
 1137 Define osvara U+0D94  
 1138 Define Osvara U+0D95  
 1139 Define ocandra U+0D95  
 1140 Define ausvara U+0D96  
 1141  
 1142 Define ka U+0D9A  
 1143 Define kha U+0D9B  
 1144 Define ga U+0D9C  
 1145 Define gha U+0D9D  
 1146 Define nga U+0D9E  
 1147 Define nnga U+0D9F  
 1148  
 1149 Define nnna U+E011 ; PUA  
 1150  
 1151 Define ca U+0DA0  
 1152 Define cha U+0DA1  
 1153 Define ja U+0DA2

1154	Define	jha	U+0DA3
1155	Define	nya	U+0DA4
1156	Define	jnya	U+0DA5
1157	Define	nyja	U+0DA6
1158			
1159	Define	tta	U+0DA7
1160	Define	ttha	U+0DA8
1161	Define	dda	U+0DA9
1162	Define	ddha	U+0DAA
1163	Define	nna	U+0DAB
1164	Define	nndda	U+0DAC
1165			
1166	Define	ta	U+0DAD
1167	Define	tha	U+0DAE
1168	Define	da	U+0DAF
1169	Define	dha	U+0DB0
1170	Define	na	U+0DB1
1171			
1172	Define	nda	U+0DB3
1173			
1174	Define	pa	U+0DB4
1175	Define	pha	U+0DB5
1176	Define	ba	U+0DB6
1177	Define	bha	U+0DB7
1178	Define	ma	U+0DB8
1179	Define	mba	U+0DB9
1180			
1181	Define	ya	U+0DBA
1182	Define	ra	U+0DBB
1183	Define	rra	U+E012 ; PUA
1184	Define	la	U+0DBD
1185	Define	va	U+0DC0
1186			
1187	Define	sha	U+0DC1
1188	Define	ssa	U+0DC2
1189	Define	sa	U+0DC3
1190	Define	ha	U+0DC4
1191	Define	lla	U+0DC5
1192	Define	fa	U+0DC6
1193			
1194	Define	llla	U+E013 ; PUA
1195			
1196	Define	avagraha	NOTDEF
1197			
1198	Define	virama	U+0DCA
1199			
1200	Define	Amatra	U+0DCF
1201			
1202	Define	aematra	U+0DD0
1203	Define	AEmatra	U+0DD1

1204  
 1205 Define imatra U+0DD2  
 1206 Define Imatra U+0DD3  
 1207  
 1208 Define umatra U+0DD4  
 1209 Define Umatra U+0DD6  
 1210  
 1211 Define rmatra U+0DD8  
 1212  
 1213 Define ematra U+0DD9  
 1214 Define Ematra U+ODDA  
 1215 Define ecandramatra U+ODDA  
 1216 Define aimatra U+ODDB  
 1217  
 1218 Define omatra U+ODDC  
 1219 Define Omatra U+ODDD  
 1220 Define ocandramatra U+ODDD  
 1221 Define aumatra U+ODDE  
 1222  
 1223 Define lmatra U+ODDF  
 1224  
 1225 Define udatta NOTDEF  
 1226 Define anudatta NOTDEF  
 1227  
 1228 Define zero U+ODE6  
 1229 Define one U+ODE7  
 1230 Define two U+ODE8  
 1231 Define three U+ODE9  
 1232 Define four U+ODEA  
 1233 Define five U+ODEB  
 1234 Define six U+ODEC  
 1235 Define seven U+ODED  
 1236 Define eight U+ODEE  
 1237 Define nine U+ODEF  
 1238  
 1239 Define Rmatra U+ODF2  
 1240 Define Lmatra U+ODF3  
 1241  
 1242 Define kunddaliya U+ODF4  
 1243  
 1244 Define abbrev NOTDEF  
 1245  
 1246 Define extravyanjana fa lla llla mba nda nnda nga nna nyja rra  
 1247 </sinh>

## 6.10 Tamil

1248 <\*taml>  
 1249 Define candrabindu NOTDEF  
 1250 Define anusvara U+0B82



1251	Define visarga	U+0B83
1252		
1253	Define asvara	U+0B85
1254	Define Asvara	U+0B86
1255		
1256	Define isvara	U+0B87
1257	Define Isvara	U+0B88
1258		
1259	Define usvara	U+0B89
1260	Define Usvara	U+0B8A
1261		
1262	Define rsvara	NOTDEF
1263	Define lsvara	NOTDEF
1264		
1265	Define esvara	U+0B8E
1266	Define Esvara	U+0B8F
1267	Define ecandra	U+0B8F
1268	Define aisvara	U+0B90
1269		
1270	Define osvara	U+0B92
1271	Define Osvara	U+0B93
1272	Define ocandra	U+0B93
1273	Define ausvara	U+0B94
1274		
1275	Define ka	U+0B95
1276	Define kha	NOTDEF
1277	Define ga	NOTDEF
1278	Define gha	NOTDEF
1279	Define nga	U+0B99
1280		
1281	Define ca	U+0B9A
1282	Define cha	NOTDEF
1283	Define ja	U+0B9C
1284	Define jha	NOTDEF
1285	Define nya	U+0B9E
1286		
1287	Define tta	U+0B9F
1288	Define ttha	NOTDEF
1289	Define dda	NOTDEF
1290	Define ddha	NOTDEF
1291	Define nna	U+0BA3
1292		
1293	Define ta	U+0BA4
1294	Define tha	NOTDEF
1295	Define da	NOTDEF
1296	Define dha	NOTDEF
1297	Define na	U+0BA8
1298		
1299	Define nnna	U+0BA9
1300		

1301	Define pa	U+OBAA
1302	Define pha	NOTDEF
1303	Define ba	NOTDEF
1304	Define bha	NOTDEF
1305	Define ma	U+OBAE
1306		
1307	Define ya	U+OBAF
1308	Define ra	U+OBB0
1309	Define rra	U+OBB1
1310	Define la	U+OBB2
1311	Define lla	U+OBB3
1312	Define llla	U+OBB4
1313	Define va	U+OBB5
1314		
1315	Define sha	U+OBB6
1316	Define ssa	U+OBB7
1317	Define sa	U+OBB8
1318	Define ha	U+OBB9
1319		
1320	Define avagraha	NOTDEF
1321		
1322	Define Amatra	U+OBBE
1323		
1324	Define imatra	U+OBBF
1325	Define Imatra	U+OBC0
1326		
1327	Define umatra	U+OBC1
1328	Define Umatra	U+OBC2
1329		
1330	Define rmatra	NOTDEF
1331	Define Rmatra	NOTDEF
1332		
1333	Define ematra	U+OBC6
1334	Define Ematra	U+OBC7
1335	Define ecandramatra	U+OBC7
1336	Define aimatra	U+OBC8
1337		
1338	Define omatra	U+OBCA
1339	Define Omatra	U+OBCB
1340	Define ocandramatra	U+OBCB
1341	Define aumatra	U+OBCC
1342		
1343	Define virama	U+OBCD
1344		
1345	Define om	U+OBDO
1346		
1347	Define udatta	NOTDEF
1348	Define anudatta	NOTDEF
1349		
1350	; Define au_length_mark U+OBD7	

1351  
1352 Define fa NOTDEF  
1353  
1354 Define Rsvara NOTDEF  
1355 Define Lsvara NOTDEF  
1356 Define lmatra NOTDEF  
1357 Define Lmatra NOTDEF  
1358  
1359 Define zero U+0BE6  
1360 Define one U+0BE7  
1361 Define two U+0BE8  
1362 Define three U+0BE9  
1363 Define four U+0BEA  
1364 Define five U+0BEB  
1365 Define six U+0BEC  
1366 Define seven U+0BED  
1367 Define eight U+0BEE  
1368 Define nine U+0BEF  
1369  
1370 Define abbrev NOTDEF  
1371  
1372 ; Define ten U+0BF0  
1373 ; Define hundred U+0BF1  
1374 ; Define thousand U+0BF2  
1375  
1376 ; Define daysign U+0BF3  
1377 ; Define monthsign U+0BF4  
1378 ; Define yearsign U+0BF5  
1379 ; Define debitsign U+0BF6  
1380 ; Define creditsign U+0BF7  
1381 ; Define asabovesign U+0BF8  
1382 ; Define rupeesign U+0BF9  
1383 ; Define numbersign U+0BFA  
1384  
1385 Define extravyanjana llla  
1386 </taml>

## 6.11 Telugu

1387 <\*telu>  
1388 Define candrabindu U+0C00  
1389 Define altcandrabindu U+0C01  
1390 Define anusvara U+0C02  
1391 Define visarga U+0C03  
1392  
1393 Define asvara U+0C05  
1394 Define Asvara U+0C06  
1395  
1396 Define isvara U+0C07  
1397 Define Isvara U+0C08

1398		
1399	Define usvara	U+0C09
1400	Define Usvara	U+0C0A
1401		
1402	Define rsvara	U+0C0B
1403	Define lsvara	U+0C0C
1404		
1405	Define esvara	U+0C0E
1406	Define Esvara	U+0C0F
1407	Define ecandra	U+0C0F
1408	Define aisvara	U+0C10
1409		
1410	Define osvara	U+0C12
1411	Define Osvara	U+0C13
1412	Define ocandra	U+0C13
1413	Define ausvara	U+0C14
1414		
1415	Define ka	U+0C15
1416	Define kha	U+0C16
1417	Define ga	U+0C17
1418	Define gha	U+0C18
1419	Define nga	U+0C19
1420		
1421	Define ca	U+0C1A
1422	Define cha	U+0C1B
1423	Define ja	U+0C1C
1424	Define jha	U+0C1D
1425	Define nya	U+0C1E
1426		
1427	Define tta	U+0C1F
1428	Define ttha	U+0C20
1429	Define dda	U+0C21
1430	Define ddha	U+0C22
1431	Define nna	U+0C23
1432		
1433	Define ta	U+0C24
1434	Define tha	U+0C25
1435	Define da	U+0C26
1436	Define dha	U+0C27
1437	Define na	U+0C28
1438		
1439	Define pa	U+0C2A
1440	Define pha	U+0C2B
1441	Define ba	U+0C2C
1442	Define bha	U+0C2D
1443	Define ma	U+0C2E
1444		
1445	Define ya	U+0C2F
1446	Define ra	U+0C30
1447	Define rra	U+0C31

1448	Define la	U+0C32
1449	Define lla	U+0C33
1450	Define llla	U+0C34
1451	Define va	U+0C35
1452		
1453	Define sha	U+0C36
1454	Define ssa	U+0C37
1455	Define sa	U+0C38
1456	Define ha	U+0C39
1457		
1458	Define avagraha	U+0C3D
1459		
1460	Define Amatra	U+0C3E
1461		
1462	Define imatra	U+0C3F
1463	Define Imatra	U+0C40
1464		
1465	Define umatra	U+0C41
1466	Define Umatra	U+0C42
1467		
1468	Define rmatra	U+0C43
1469	Define Rmatra	U+0C44
1470		
1471	Define ematra	U+0C46
1472	Define Ematra	U+0C47
1473	Define ecandramatra	U+0C47
1474	Define aimatra	U+0C48
1475		
1476	Define omatra	U+0C4A
1477	Define Omatra	U+0C4B
1478	Define ocandramatra	U+0C4B
1479	Define aumatra	U+0C4C
1480		
1481	Define virama	U+0C4D
1482		
1483	Define udatta	NOTDEF
1484	Define anudatta	NOTDEF
1485		
1486	; Define length_mark	U+0C55
1487	; Define ai_length_mark	U+0C56
1488		
1489	Define tsa	U+0C58
1490	Define dza	U+0C59
1491		
1492	Define fa	NOTDEF
1493		
1494	Define Rsvara	U+0C60
1495	Define Lsvara	U+0C61
1496	Define lmatra	U+0C62
1497	Define Lmatra	U+0C63

```

1498
1499 Define zero      U+0C66
1500 Define one      U+0C67
1501 Define two      U+0C68
1502 Define three    U+0C69
1503 Define four     U+0C6A
1504 Define five     U+0C6B
1505 Define six      U+0C6C
1506 Define seven    U+0C6D
1507 Define eight    U+0C6E
1508 Define nine     U+0C6F
1509
1510 Define abbrev    NOTDEF
1511
1512 Define zero1     U+0C78
1513 Define one1      U+0C79
1514 Define two1      U+0C7A
1515 Define three1    U+0C7B
1516
1517 Define one2      U+0C7C
1518 Define two2      U+0C7D
1519 Define three2    U+0C7E
1520
1521 Define tuumu      U+0C7F
1522
1523 Define extravanja dza llla tsa
1524 </telu>

```

## 7 Phase 1: Normalization

By first normalizing the input, we can simplify the transliteration rules without sacrificing flexibility on the input side.

### 7.1 Normalization Pass 1: Decompose compound characters.

Separate diacritics from their base letters and sort them into a standard order. This ensures that, for example, all long vowels are encoded as base letter + macron and not by a precomposed character.

```

1525 pass(NFD)

```

### 7.2 Normalization Pass 2: Downcase base characters.

Convert all uppercase letters to lowercase.

```

1526 pass(Unicode)
1527
1528 UniClass[upper] = (U+0041 .. U+005A) ; A..Z
1529 UniClass[lower] = (U+0061 .. U+007A) ; a..z
1530
1531 [upper] > [lower]

```

This is also a good time to delete optional characters.

```
1532 lbrack .* rbrack >
```

### 7.3 Normalization Pass 3: Special cases and alternate transcriptions.

ᳵ and ᳶ always represent vocalic l and vocalic r.

```
1533 pass(Unicode)
```

```
1534
```

```
1535 l ringbelow > lry
```

```
1536 r ringbelow > ri
```

For compatibility with IAST and common practice, allow ᳶ as a synonym for vocalic r in Sanskrit.

```
1537 <san>r dotbelow > ri
```

Similarly, accept ᳵ as a synonym for vocalic l in an appropriate context. Specifically, as per section 26 of Whitney's grammar, vocalic l is only found in derivatives of the rare root **klp** and (as far as I can tell) it will always immediately following a **k**. The following rule allows ᳵ to be used in those contexts. (See "Transliteration pass 1.5" below for another way to disambiguate dual uses of ᳵ.)

```
1538 <san>l dotbelow / k _ > lry
```

ᳶ is always a vowel.

```
1539 <san>l dotbelow / _ macronabove > lry
```

Allow āi and āu as synonyms for ai and au.

```
1540 <san>a macronabove / _ (i | u) > a
```

Recognize some older transcription conventions used in, for example, works by Macdonell or Whitney.

```
1541 <*san>
```

```
1542 n macronabove <> n dotabove
```

```
1543 c ccedilla <> s acute
```

```
1544 circumflex <> macronabove
```

```
1545 </san>
```

```
1546 <*san>
```

```
1547 a macronbelow u macronbelow > a u macronbelow
```

```
1548 a macronbelow i macronbelow > a i macronbelow
```

```
1549 </san>
```

Diereses are used to keep "i" or "u" from forming diphthongs with a preceding "a".

```
1550 i dieresis > zwnj i
```

```
1551 u dieresis > zwnj u
```

In NFD, marks under the base character are encoded before marks above the character, but we need the long vowel mark and the candrabindu to come first.

```
1552 macronbelow macronabove > macronabove macronbelow
```

```
1553 macronbelow CANDRABINDU > CANDRABINDU macronbelow
```

## 8 Sanskrit External Vowel Sandhi

This section was inspired by Macdonell's practice of writing, for example,

अध्यूढ *adhi\_ūḍha*, *pp.* (√ūh) placed upon (*lc.*).

in his Sanskrit-English Dictionary to illustrate the action of vowel sandhi in compound words. With these rules, you can use U+203F UNDERTIE to request vowel coalescence in the Devanagari text:

\N{adhi\_ūḍha}, ...

I've attempted to implement all of sections 19–22 of *A Sanskrit Grammar for Students* [3]. The rules below are quoted from there, with minor adaptations and elisions. The order and division into passes is chosen to ensure that, for example, the digraphs *ai* and *au* are processed before any of the patterns beginning with *i* or *u*.

```
1554 (*san)
1555 pass(Unicode)
1556
1557 Define coalesce U+203F ; UNDERTIE
```

For simplicity (and because Macdonell doesn't display them in the Devanagari in his dictionary), drop all accents.

```
1558 UniClass[vaccent] = (acute grave macronbelow vert)
1559
1560 [vaccent] / _ coalesce >
1561 [vaccent] / coalesce _ >
1562 [vaccent] / coalesce . macronabove? _ >
```

SGS §21a: The Guṇa vowels *e* and *o* remain unchanged before *a*, which is elided:

```
1563 coalesce a / (e | o) _ > apos
```

NB: *e* or *o* followed *ā* is handled by the next rule because more specific (i.e., longer) rules take precedence.

SGS §21b: The Guṇa vowels *e* and *o* become *a* (through *ay* and *av*, which drop the semivowel) before every other vowel or diphthong:

```
1564 UniClass[othervowel] = (i u ri lry e o)
1565
1566 (e | o) coalesce / _ a macronabove > a hiatus
1567 (e | o) coalesce / _ [othervowel] > a hiatus
```

SGS §22: The Vṛddhi vowels *ai* and *au* respectively become *ā* and *āv* before every vowel (or diphthong):

```
1568 a i coalesce / _ (a | [othervowel]) > a macronabove hiatus
1569 a u coalesce / _ (a | [othervowel]) > a macronabove v
```

SGS §22a: The (secondary) hiatus occasioned by the dropping of *y* and *v* in the above three cases (21b and 22) remains.

```
1570 pass(Unicode)
```



SGS §19a: a and ā coalesce with a following simple liquid vowel to Guṇa:

```
1571 a macronabove? coalesce i macronabove? > e
1572 a macronabove? coalesce u macronabove? > o
1573 a macronabove? coalesce ri macronabove? > a r
1574 a macronabove? coalesce lry macronabove? > a l
```

SGS §19b: a and ā coalesce with Guṇa vowels to Vṛddhi

SGS §19c: a and ā are absorbed by Vṛddhi vowels

```
1575 a macronabove? coalesce (e | a i) > a i
1576 a macronabove? coalesce (o | a u) > a u
```

Other cases of §19b,c are digraphs handled by §18 below.

SGS §20: A simple liquid vowel followed by any other vowel or by a diphthong is changed to its semivowel:

```
1577 UniClass[anyvowel] = (a i u ri lry e o)
1578 UniClass[liquid]    = ( i u ri lry)
1579 UniClass[sv]        = ( y v r l)
1580
1581 [liquid] macronabove? coalesce / _ [anyvowel] > [sv]
```

SGS §18: If the same simple vowel (short or long) occurs at the end and beginning of words, the result is a long vowel:

```
1582 a macronabove? coalesce a macronabove? > a macronabove
1583 i macronabove? coalesce i macronabove? > i macronabove
1584 u macronabove? coalesce u macronabove? > u macronabove
1585 ri macronabove? coalesce ri macronabove? > ri macronabove
1586 lry macronabove? coalesce lry macronabove? > lry macronabove
```

## 8.1 Irregular sandhi

U+2054 INVERTED UNDERTIE can be used to invoke an alternate set of rules, for example,

उपोषण  $\text{upa}_\text{̣oṣaṇa}$ , *n.* fasting.

where अ a and ओ o have coalesced to form औ o rather than the regular औ au.

```
1587 Define altcoalesce U+2054 ; INVERTED UNDERTIE
1588
1589 a altcoalesce o > o
1590 </san>
```

(This is the only exception I've come across so far; if there are others, I'll add them.)

## 9 Phase 2: Transliteration

### 9.1 Transliteration Pass 1: Transliterate blindly.

After this step, all characters will be replaced by their approximate Devanagari equivalents, but the result will not be well-formed.

```
1591 pass(Unicode)
```

## Numerals

```
1592 U+0030 > zero
1593 U+0031 > one
1594 U+0032 > two
1595 U+0033 > three
1596 U+0034 > four
1597 U+0035 > five
1598 U+0036 > six
1599 U+0037 > seven
1600 U+0038 > eight
1601 U+0039 > nine
```

## Vowels

```
1602 UniClass[short_in] = (a      i      u      ri      lry)
1603 UniClass[short_out] = (asvara isvara usvara rsvara lsvara)
1604 UniClass[long_out]  = (Asvara Isvara Usvara Rsvara Lsvara)
1605
1606 [short_in] macronabove > [long_out]
1607
1608 [short_in] > [short_out]
1609
1610 ⟨mlym⟩u breveabove > virama ; ???
1611 ⟨!mlym⟩u breveabove > usvara
1612
1613 e                > esvara
1614 e macronabove   > Esvara
1615 a i             > aisvara
1616
1617 e circumflex    > ecandra
1618 o circumflex    > ocandra
1619
1620 o                > osvara
1621 o macronabove    > Osvara
1622 a u             > ausvara
1623
1624 ⟨sinh⟩ae > aesvara
1625 ⟨sinh⟩ae macronabove > AEsvara
1626
1627 ⟨telu⟩n breveabove > altcandrabindu ; ???
1628
1629 ⟨deva | knda⟩h macronbelow > jihvamuliya
1630 ⟨deva | knda⟩h brevebelow  > upadhmaniya
1631
1632 ⟨taml⟩k macronbelow > visarga
1633
1634 apos > avagraha
```

## Tone marks

```
1635 vert          > udatta
```

1636 macronbelow > anudatta  
1637 <san>hookbelow > svarita\_k\_i

Delete non-Vedic accents.

1638 grave >  
1639 acute >  
1640 tildeabove >

#### Bindus and visargas

1641 m dotabove > anusvara  
1642 h dotbelow > visarga

The International Alphabet of Sanskrit Transliteration uses **m̐** instead of **m̐̐** for the anusvara. Since **m̐** is only otherwise used for the tippi in Gurmukhi, allow it as a synonym for the anusvara everywhere else.

1643 <guru>m dotbelow > tippi  
1644 <!guru>m dotbelow > anusvara  
1645 m CANDRABINDU > candrabindu  
1646 n CANDRABINDU > candrabindu

#### Consonants (no conjuncts)

##### Nuktas

1647 <\*nukta>  
1648 q <> qa  
1649 k macronbelow h macronbelow > khha  
1650 g dotabove > ghha  
1651 <!telu>z > za  
1652 f > fa  
1653 </nukta>  
1654 <knda>z > za  
1655 <knda | sinh>f > fa  
1656 <telu>z > dza  
1657 <beng | orya>w > wa

##### Velar

1658 k <> ka  
1659 k h <> kha  
1660 g <> ga  
1661 g h <> gha  
1662 n dotabove <> nga  
1663 <sinh>n breveabove g > nnga

##### Palatal

1664 c <> ca  
1665 <telu>c circumflex <> tsa  
1666 c h <> cha  
1667 j <> ja  
1668 j h <> jha  
1669 n tildeabove <> nya  
1670 <sinh>n breveabove j <> nyja

### Retroflex

1671 t dotbelow <> tta  
1672 t dotbelow h <> ttha  
1673 d dotbelow <> dda  
1674 d dotbelow h <> ddha  
1675 n dotbelow <> nna  
1676  
1677 ⟨sinh⟩n breveabove d <> nda  
1678 ⟨sinh⟩n breveabove d dotbelow <> nndda  
1679  
1680 ⟨deva&!san⟩r dotbelow > dddha  
1681 ⟨beng⟩r dotbelow > dddha  
1682 ⟨guru | orya⟩r dotbelow > rra  
1683  
1684 ⟨deva | beng | orya⟩r dotbelow h > rha

### Dental

1685 t <> ta  
1686 t h <> tha  
1687 d <> da  
1688 d h <> dha  
1689 n <> na

### Labial

1690 p <> pa  
1691 p h <> pha  
1692 b <> ba  
1693 b h <> bha  
1694 m <> ma  
1695 ⟨sinh⟩m breveabove b <> mba

### Semi-vowels

???

1696 r macronbelow <> rra  
1697 ⟨mlym⟩t macronbelow <> tttta  
1698 ⟨deva | tam | mlym⟩n macronbelow <> nnna  
1699 ⟨sinh⟩n macronbelow <> nnna  
1700 ⟨deva | tam | mlym | knda | telu | sinh⟩l macronbelow <> lllla

### Semi-vowels

1701 y <> ya  
1702 ⟨deva | beng | orya⟩y dotabove > yya  
1703 r breveabove > rra zwj  
1704 r <> ra  
1705 l <> la  
1706 v <> va

### Sibilants

```

1707 s acute      <> sha
1708 s dotbelow  <> ssa
1709 s           <> sa
1710 h           <> ha

```

### Vedic lingual l

```

1711 l dotbelow   <> lla
1712 ⟨*san⟩
1713 l dotbelow h <> lla ha
1714 ⟨/san⟩

```

### Punctuation and symbols

```

1715 degree  <> abbrev
1716 bar     <> danda
1717 bar bar <> dbldanda

```

### Ligature breaking

Per note 12 of Stone's *vade mecum*, allow a colon to break ligatures.

```

1718 UniClass[vyanjana] = (ka .. ha extravyanjana)
1719
1720 ; colon / [vyanjana] _ [vyanjana] <> virama zwj
1721
1722 colon <> zwj

```

## 9.2 Transliteration Pass 1.5

Adding an additional pass here would enable automatic disambiguation between **l̥** representing lingual l and **l̥** representing vocalic l to be made based on the phonetic environment (see section 54 of Whitney's grammar), but only at the cost of making it more complicated to type an isolated lingual l. It hardly seems worth it.

```

1723 ; pass(Unicode)
1724 ;
1725 ; UniClass[vyanjana] = (ka .. ha)
1726 ;
1727 ; lla / [vyanjana] _ > lsvara
1728 ; lla / # _ > lsvara
1729 ; lla / _ # > lsvara
1730 pass(unicode)

```

Per rule R10 in section 12.1 of Unicode 7.0 [4], vedic accents (“svaras”) should be encoded following bindus. This is necessary to get correct rendering in, e.g., XeTeX (luaTeX is more forgiving).

```

1731 UniClass[svaras] = (udatta anudatta)
1732 UniClass[bindus] = (anusvara candrabindu)
1733
1734 [svaras]=v1 [bindus]=b1 > @b1 @v1

```

### 9.3 Transliteration Pass 2: Attach vowels to consonants.

At this stage all vowels are still represented as independent vowels. This step will convert vowels following consonants into dependent vowel signs.

```
1735 pass(Unicode)
1736
1737 UniClass[vyanjana] = (ka .. ha extravyanjana)
1738
1739 UniClass[svara] = (      Asvara \
1740                      isvara  Isvara \
1741                      usvara  Usvara \
1742                      rsvara  Rsvara \
1743                      lsvara  Lsvara \
1744                      ecandra esvara Esvara \
1745                      aisvara \
1746                      ocandra osvara Osvara \
1747                      ausvara)
1748
1749 UniClass[matra] = (      Amatra \
1750                      imatra Imatra \
1751                      umatra Umatra \
1752                      rmatra Rmatra \
1753                      lmatra Lmatra \
1754                      ecandramatra ematra Ematra \
1755                      aimatra \
1756                      ocandramatra omatra Omatra \
1757                      aumatra)
1758
1759 [svara] / [vyanjana] _ > [matra]
```

Attach isolated vowel signs and other combining marks to a non-breaking space (cf. The Unicode Standard 7.0, page 60, “Exhibiting Nonspacing Marks in Isolation”. HarfBuzz currently requires this for proper rendering.

NB: visarga is a *spacing* combiner.

```
1760 UniClass[combiners] = ([matra] \
1761                        candrabindu anusvara visarga \
1762                        udatta anudatta \
1763 <san>                  svarita_k_i \
1764 )
1765
1766 [combiners]=c1 / # _ > nbsp @c1
```

### 9.4 Transliteration Pass 3: Form consonant conjuncts.

```
1767 pass(Unicode)
1768
1769 UniClass[vyanjana] = (ka .. ha extravyanjana)
1770
1771 [vyanjana] / _ [vyanjana] > [vyanjana] virama
```

## 9.5 Transliteration Pass 4: Add viramas.

```
1772 pass(Unicode)
1773
1774 UniClass[vyanjana] = (ka .. ha extravyanjana)
1775
1776 UniClass[novirama] = (asvara Amatra \
1777                       imatra Imatra \
1778                       umatra Umatra \
1779                       rmatra Rmatra \
1780                       lmatra Lmatra \
1781                       ematra Ematra ecandramatra aimatra \
1782                       omatra Omatra ocandramatra aumatra \
1783                       virama)
1784
1785 [vyanjana] / _ ^[novirama] <> [vyanjana] virama
```

## 9.6 Transliteration Pass 5: Remove inherent a.

Finally, remove the inherent short a.

```
1786 pass(Unicode)
1787
1788 UniClass[vyanjana] = (ka .. ha extravyanjana)
1789
1790 [vyanjana] asvara <> [vyanjana]
```

## 9.7 Transliteration Pass 6: Resolve PUA characters.

```
1791 <beng | gujr | guru | knda | sinh>pass(Unicode)
1792
1793 <guru | gujr | beng>qa > ka nukta
1794 <beng | gujr>khha > kha nukta
1795 <beng | gujr>ghha > ga nukta
1796 <beng | gujr>fa > pha nukta
1797 <beng | gujr | knda>za > ja nukta
1798 <beng>wa > ba nukta
1799 <knda>fa > pha nukta
1800 <sinh>nnna > na U+002E
1801 <sinh>rra > ra U+002E
1802 <sinh>llla > lla U+002E
```