

# **SCRIPT GRAMMAR FOR KONKANI LANGUAGE**

**Prepared by**

**Technology Development for Indian Languages (TDIL) Programme  
of DIT, GoI in association with**

**Centre for Development of Advanced Computing (C-DAC)**

## Table of Contents

0. INTRODUCTION .....	3
1. OBJECTIVES OF SCRIPT GRAMMAR .....	4
2. END USERS FOR SCRIPT GRAMMAR .....	5
3. SCOPE .....	6
4. TERMINOLOGY .....	7
5. PHILOSOPHY AND UNDERLYING PRINCIPLES.....	11
6. SCRIPT GRAMMAR STRUCTURE.....	12
6.1. PERIPHERAL ELEMENTS OF THE SCRIPT GRAMMAR.....	13
6.2. CONFORMITY TO THE SYLLABLE STRUCTURE .....	14
6.3 SCRIPT GRAMMAR PROPER.....	18
6.3.1. The Character Set of Konkani.....	18
6.3.2. Consonant Mātrā Combinations. ....	24
6.3.3. The Ligature Set of Konkani.....	29
6.3.4 The Collation Order of Konkani. ....	34
7. REFERENCES .....	35
8. ANNEXURES .....	36
Annexure 1: Names of experts who have contributed to the script grammar.....	36
Annexure 2: Unicode Table of Devanāgarī .....	37

## 0. INTRODUCTION

The term **script grammar** refers to the behaviour pattern of the writing system of a given language. Languages which have written representations do not use a haphazard manner of storing the information within the system, but use a coherent pattern which is similar to the linguistic grammar of a given language. With the help of specialists (not necessarily linguist) who work in the area of the written representation of the language, the manner in which the shapes of the characters of the language and the representation of the conjunct forms is provided. In other words the Script Grammar deals with the surface structure of the language and tries to provide the best possible “fit” for shapes and their representation. Since this is a highly subjective issue, the shapes provided here are recommendations at the best and conform to the perception of the mandating body/evaluators who consensually arrive at the “best possible fit” which is acceptable to a majority of users. An example will make the above clear. Although Marathi and Hindi share the same script Devanāgarī, not only do they not share the same character inventory but in addition the representation of certain characters is different. Thus the Marathi /la/ is different from the Hindi /la/ in so far as the placement of the stem is concerned Hindi ल Marathi ल. This ensures that the Script Grammar conforms to the language in question and provides the character shapes acceptable to a given user community. It should be noted that this does not mean monotony. The Hindi and Marathi /la/ can have a variety of forms once the intrinsic structure of the character is determined.

Script Grammar is the term used to define:

- the writing system used to inscribe a given language
- the history of the script and language (wherever available)
- the syllabic structure of the writing system of the language
- the rule ordering of the characters within the syllable (akshar)
- description of the syllabic clusters
- collation order of the characters: lexical / dictionary sorting order

## 1. OBJECTIVES OF SCRIPT GRAMMAR

The Objectives of the script grammar for each language can be divided into two major parts:

### **Societal:**

- Provide a visual representation of shapes that are deemed to be in conformity with the perception of a given community
- Ensure thereby that this perception is safe-guarded
- Through wide-spread dissemination and creation of appropriate tools ensure that within the given linguistic community, all media tries to adopt the given shape.

### **Technical:**

- Classify the language in terms of its ISO and also whether it belongs to the Abjad, Akshar (Alphasyllabary) class.
- Provide an inventory of the characters pertinent to the language and classify the same in terms of their taxonomy.
- As a corollary determine whether the inventory is in conformity to the Syllable formalism as stipulated in ISCII'91 and subsequently adopted by Unicode.
- Since Brahmi is written from left to right, and since certain characters do not follow the linear L to R order, provide an inventory of displaced concatenators i.e. characters such as Mātrās that concatenate to the Consonant
- Propose the best shape representation of the individual characters as well as of the ligatures used within a given script. As a corollary request the expert(s) to identify the largest possible strings of such ligatures.
- Finally provide the collation order pertinent to that script/language, which would be of great utility to high-end NLP as well as to CLDR's in the pertinent language. The collation order for Marathi is different from Hindi although both languages share the same script. Thus in Marathi झ, ञ are placed at the end of the consonant inventory i.e. after ह in the sort order. In Hindi झ is sorted along with क and ञ with ज

## 2. END USERS FOR SCRIPT GRAMMAR

The script-grammar specific to a given language can be used by a large number of users.

- Most importantly it can be used by font developers desirous of developing a font which is compliant with the perception of the characters and ligatures of a language by its user community.
- Certain features of the script grammar such as the shapes can also be used for testing OCR and OHWR. Similarly information regarding Ligatures as well as collation order can help in high-end NLP work such as detecting invalid combinations, correct implementation of syllable structure, prediction routines to name a few. Information regarding collation and character sets can be also used for CLDR.
- They allow the font designer to design a font which is in compliance with the norms and standards of that particular script. A major problem which will be dealt with in the template is one of ligatures. The final list of ligatures defined by the script grammar allows the font designer to write specific rules for such glyphs.
- It permits the software developer to design and implement the keyboard and the input mechanism which will meet the requirement of the particular linguistic community.
- The collation or sort order as described in a Script Grammar permits the software developer to write software functions/ routines for sorting data in all applications.
- Script Grammars are equally important for keyboard design, especially when supplemented by frequency data from a corpus.

As can be seen the script grammar has a wide range of use and can be of utility to font developers, Indian language developers and linguists in the area of computation.

### 3. SCOPE

This script grammar document contains following information about the language and the script used for writing the language.

1. Name of the language and its representation in the 3 letter mnemonic as per ISO 639.1 & 639.2 standard.
2. Script used to inscribe the given language
3. The structure of the script used for writing the language
  - Rule ordering of the characters within the syllable formation is a language
  - Description of the syllabic clusters of the script
  - Collation order of the characters: lexical / dictionary sorting order
  - Compliance of the script with Unicode.

These will be treated within the relevant sections of the script grammar

## 4. TERMINOLOGY<sup>1</sup>

**Abjad:** A writing system in which each symbol always or usually stands for a consonant. The long vowels are indicated. However the short vowels are rarely marked and the reader needs to supply these. Example: Urdu written in Perso-Arabic Script is an example of this writing system.

**Abugida:** also called an alphasyllabary, is a segmental writing system in which consonant–vowel sequences are written as a unit: each unit is based on a consonant letter, and vowel notation is obligatory but secondary<sup>2</sup>

**Akshar:** see **Abugida**

**Allographs:** Variants of the representation of a character. Thus ae and æ [U+00E6] in Latin alphabet are allographs.

**Allo-Script:** The term relates to languages which share a common script. Thus Devanāgarī is used to write 9 official languages. However these languages do not use the same set of characters. Thus Marathi uses the retroflex lla ऴ [U+ 0933] which Hindi does not use. Flaps used in Hindi ढ [U+095C] ढ [U+095D] are not used in Konkani. These sub-sets of scripts based on a single “matricial” script are termed as allo-scripts.

**Alphabet:** A set of letters used in writing a language. Example: The English Alphabet.

**Aspirated consonant:** A consonant which is pronounced with an extra puff of air coming out at the time of release of the oral obstruction. This has a sound of an extra "h".

**Basic alphabet:** The minimal set of letters which can be used for uniquely encoding every word of a language. The basic alphabet for English consists of only the upper-case letters A-Z

**Catenators:** Also termed as Concatenators are characters which are concatenated to another character. In the Brahmi script these are the Mātrās or Vowel modifiers which are adjoined to the consonant and add a vocalic value to the consonant.

**Conjunct:** The Indic scripts are noted for a large number of consonant conjunct forms that serve as orthographic abbreviations (ligatures) of two or more adjacent letterforms. This abbreviation takes place only in the context of a consonant cluster...Under normal circumstances, a consonant cluster is depicted with a conjunct glyph if such a glyph is available in the current font. In the absence of a conjunct glyph, the one or more dead

---

<sup>1</sup> As in the case of the BIS Document, in order to make the terminology accessible for all readers, examples have been chosen from English/Latin scripts, wherever possible. Some definitions have been excerpted from the BIS ISCII91 document and suitably modified where necessary.

<sup>2</sup> Wikipedia definition

consonants that form part of the cluster are depicted using half-form glyphs. In the absence of half-form glyphs, the dead consonants are depicted using the nominal consonant forms combined with visible virama signs.<sup>3</sup>

**Consonant:** A letter representing a speech sound in which the breath is at least partly obstructed,

**Diacritic:** A mark added to a letter which distinguishes it from the same letter without a mark, usually having a different phonetic value or stress.

**Displaced Catenator:** (see Catenator) Within the Brahmi script, the writing system is linear and moves from left to right. However in the case of some catenators this rule is not observed and the catenator (wholly or partially) is placed to the right of the consonant to which it relates. The short vowel I in Devanāgarī is an example of a displaced catenator.

**Display composing:** The process of organizing the basic shapes available in a font in order to display (or print) a word.

**Display rendition:** The process by which a string of characters is displayed (or printed). In this process several consecutive characters may combine with each other on the screen. The sequence of display of the characters may become different.

**Eyebrow repha:** (See Eyelash ra)

**Eyelash ra:** The eyelash ra is used in Konkani, Nepali and Marathi. It is treated as different from the र् (repha) by certain linguists. While the former is treated as a flap, the latter is a continuant trill (*cf.*, Kalyan Kale and Anjali Soman. 1986). There are cases in Marathi of minimal pairs such as: आचार्यास “to the teacher” vs. आचाऱ्यास “to the cook” or दर्या /darya/ “ocean” vs. दऱ्या /darya/ “valleys”.

**Font:** A set of symbols used for display or printing of a script in a particular style.

**International numerals:** The conventional 0 to 9 digits used in English for denoting numbers. These are also known as Indo-Arabic numerals (to differentiate them from the Roman numerals like IX for 9).

**Latin alphabet:** The alphabet used for writing the language of ancient Rome. Also known as the Roman alphabet. The alphabet is used today for writing English and European languages.

**Letter:** A character representing one or more of the simple or compound sounds used in speech. It can be any of the alphabetic symbols.

---

<sup>3</sup> Unicode 6.0 Chapter 9.0 pp 6-7



**Ligature:** (see **Conjunct**)

**Nasal consonant:** A consonant pronounced with the breath passing through the nose.  
Example *m n* in English.

**Nasalized vowel:** A vowel pronounced with the breath passing both through the nose and the mouth. In Indian scripts this is denoted by a Chandrabindu and gives the vowel/vowel modifier over which it placed a nasal value. Example: जाँच

**Phonetic alphabet:** An alphabet which has direct correspondence between letters and sounds Example: The International Phonetic Alphabet..

**Pure consonant:** A consonant which does not have any vowel implicitly associated with it.

**Rafar:** A special case of a ligature constituted by the adjunction of ra followed by a halanta to consonant. The resultant combination places the ra on top of the consonant to which it is adjoined. In case the consonant itself is adjoined to another consonant, the rafar is placed above the consonant e.g. र्+क कर्क , र्+घ्+य घ्यर्

**Rakar:** A special case of a ligature constituted by the adjunction of a consonant followed by a halanta to ra. In a large number of Brahmi scripts the ra is adjoined to the stem of consonant to which it relates. In the case of consonants which have no stem such as the dental retroflexes in Devanāgarī, the rakar is placed below the consonant to which it relates.

**Repha:** (see **Rafar**)

**Roman script:** The script based on the ancient Roman alphabet, with the letters A-Z and additional diacritic marks. Used for writing a language which is not usually written in the Roman alphabet.

**Script:** A distinctive and complete set of characters used for the written form of one or more languages.

**Script numerals:** The 0 to 9 digits in a script, which have shapes distinct from their international counterparts.

**Syllable:** A unit of pronunciation uttered without interruption, forming whole or part of a word, and usually having one vowel or diphthong sound optionally surrounded by one or more consonants

**Transliteration:** Representation of words with the closest corresponding letters in an alphabet of a different language.

**Vowel:** A letter representing a speech sound made with the vibration of the vocal cords, but without audible obstruction

**Vowel sign:** A graphic character associated with a letter, to indicate a vowel to be associated with that character (Mātrā in Hindi).

## 5. PHILOSOPHY AND UNDERLYING PRINCIPLES

The script grammar is based on the following principles:

1. The Grammar aims to depict the surface grammar of the written language: the manner in which characters as well as conjuncts are depicted
2. Where a given script admits many languages, it is pre-suppose that such languages will prescribe different representations for a given shape or conjunct according to the perception of the native users of that language
3. Corollary to the above the result is a script and allo-scripts i.e. a given script shared by many languages is not uniformly deployed across all the languages but is subject to variations and modulations.
4. The term Grammar is used here in a non-normative sense: what is prescribed is in the form of recommendations provided by experts who visualize the shape of the given script in their mother tongue in a specific manner. Subjective variations may occur<sup>4</sup>
5. The Grammar is limited to its synchronic use i.e. the manner in which a given language as of today admits a character set within the script used to write it. It is not diachronic or historical in nature and does not study the evolution of the given script across centuries.

---

<sup>4</sup> It is recommended that such variations be culled by placing the Grammars of different scripts in public review.

## 6. SCRIPT GRAMMAR STRUCTURE

The script grammar provided below has the following parts.

Part 6.1. deals with peripheral elements such as the ISO of the language, the writing system used: (Alphasyllabic) Abugida or Abjad.

Part 6.2. treats of the syllabic structure. It verifies whether the character set of the language complies with the ISCII syllabic structure and if not which cases are not compliant.

Part 6.3 is the script grammar proper and describes the character set as well as the conjunct shapes of the given script along with the collation order

## 6.1. PERIPHERAL ELEMENTS OF THE SCRIPT GRAMMAR

These constitute the elements that are peripheral to the Script Grammar. The main parameters considered are the mnemonic and name of the language (needed for CLDR and also for language tags), the writing system used to inscribe the language and wherever possible a short history of the language.

### 6.1.1. Name of the language and its representation in the 3 letter mnemonic as per 639.3

**Name of the Language: KONKANI**  
**ISO Mnemonics: *kok***

This refers to a one line description of the language and its mnemonic representation as per the ISO. In the case of Marathi, the above information is pertinent.

### 6.1.2. Identification of the writing system(s) used to inscribe the given language **Konkani is written using the Devanāgarī script. It is an alphasyllabary with the akshar as its core.**

This is a one line description of the script used to write the language. However in case the language uses more than one script, all the scripts in question are specified, provided these constitute the official language of the given state.

All scripts derived from Brahmi are Abugidas i.e. syllabary driven systems. The main features of Abugidas are as under:

- The consonant has an implicit vowel built-in which is normally the schwa.
- The inherent vowel can be modified by the addition of other vowels or muted by a diacritic termed as a Virama or Halanta
- Vowels can be handled as full vowels with a vocalic value
- When two or more consonants join together they form ligatures which can be recognized by their shape क्ख or alternatively form an entirely new shape क्ख = क्ष.

Abugidas/Alphasyllabaries because of their syllabic structure require a special description which is the subject of the discussion in 6.2. below.

### 6.1.3. Amendments needed in Unicode for Konkani language

**The experts of the Goa Konkani Akademi saw no need for any amendments.**

## 6.2. CONFORMITY TO THE SYLLABLE STRUCTURE

**Konkani language complies with the syllable (akshar) structure described above. It can admit up to h consonant clusters.**

Alphasyllabaries are determined by the notion of the syllable or the Akshar. The compositional grammar of the syllable determines its well-formedness. This is through a series of formal constraints based on a Backus-Naur Formalism which is given below. The syllable (akshar), first defined in the ISCII document (1991), identifies the following character ‘sub-sets’ for the purposes of identifying the syllable (akshar). In what follows the syllable analysis will be restricted to Konkani.

### (C) Consonants

क	ख	ग	घ	ङ
च	छ	ज	झ	ञ
ट	ठ	ड	ढ	ण
त	थ	द	ध	न
प	फ	ब	भ	म
य	र	ऌ <sup>5</sup>	ल	व
श	ष	स	ह	ळ

### (V) Vowels<sup>6</sup>

अ	आ	इ	ई	उ	ऊ	ऋ	ॠ	ए	ऐ
ऐ	ओ	औ	औ						

### (M) Mātrās or Vowel Modifiers

ा	ि	ी	ु	ू	ृ	ॠ	े	ै
ो	ौ	ँ	ॉ					

### (D) Diacritics

ं: Anuswar	Anuswara, a nasal, is denoted by a dot above the letter after which it is to be pronounced. This falls under Nasal category.
ँ: Chandrabindu <sup>7</sup>	Chandrabindu, a nasal, is denoted by a breve with a dot

<sup>5</sup> There are differences of expert opinion regarding this character to generate eyelash ra

<sup>6</sup> As per experts an addendum needs to be done to the existing Sscript grammar and the half Chandra vowels and vowel modifiers need to be incorporated, which is reflected in this script grammar document.

	superposed above the letter after which it is to be pronounced. This falls under Nasal category.
◌ः Visarga	Visarga, denoted by two dots placed above the other.

(H):Halanta ◌ - Halant used in most writing systems of the Indian subcontinent to signify the lack of an inherent vowel.

(N)<sup>8</sup> Nukta ◌ - is Not used in Konkani

Each of these sub-types has its restrictions in terms of what can precede or follow it within a syllable (akshar), as shown in the table below:

PRECEDED BY	SUBTYPE	FOLLOWED BY
-, H	<b>C</b>	M,D,H
-	<b>V</b>	D
C	<b>M</b>	D
C,V,M	<b>D</b>	-
C	<b>H</b>	C

C can be preceded by H or no subtype and followed by any one of the following: M,D,H

V can be preceded by no subtype and followed by D but not by another sub-type.

M can be preceded by C and followed by D.

D can be preceded by C, V, M and followed by no other subtype. It closes the syllable (akshar).

H can be preceded by C alone and followed only by C and no other sub-set.

### 6.2.1. Syllable (akshar) Types

The formalism defines the syllable (akshar) in terms of both what can constitute a syllable (akshar) and what cannot. A valid syllable (akshar) as per this definition can be of only two types:

1. A vowel syllable (akshar): a full vowel.
2. A consonant syllable (akshar): a full consonant (having a weak vowel or a mātrā )

---

<sup>7</sup> Chandrabindu is not part of the Konkani character set and is used only to denote a अँ + ◌ and आँ + ◌ Anuswar as in बँक

<sup>8</sup> The nukta is a small dot placed under a character in Northern scripts to show that they are flapped or for deriving 5 other consonants in the Devanāgarī and Punjabi scripts, required for Urdu क़,ख़,ग़,ज़,फ़

The three other subsets viz. Mātrās, Diacritics, Halanta cannot constitute a syllable (akshar) by themselves or in combination among themselves.

6.2.1.1. The Vowel syllable (akshar) is of the following types:

6.2.1.1.1. A pure vowel all by itself: अ, /a/ आ /ā/ etc.

6.2.1.1.2. A vowel followed by a modifier i.e. either a nasal marker or a visarga or an avagraha: ई /ī /, आ: /āH/ आऽ /ā:/

6.2.1.2. The Consonant syllable (akshar) can be of the following types:

6.2.1.2.1. A full consonant (with or without Nukta) i.e. with the inherent vowel : क : /ka/

6.2.1.2.2. A consonant followed by a mātrā i.e. the inherent vowel being substituted by another vowel: की /ki:/

6.2.1.2.3. A consonant followed by a modifier: कं /k̃/, हः /haH<sup>9</sup>/

6.2.1.2.4. A consonant followed by a mātrā and a modifier: कुं /kū/, दुः /duH/.

6.2.1.2.5. A consonant cluster i.e. a dead or half consonant (Consonant+Halanta) followed by a full consonant followed optionally by a mātrā, a modifier or a combination of both. These result in a ligature or what is often termed as *yuktakshara*.

त्क /tka/ त्कं /tk̃/, त्कः /tkaH/ त्कुं /tkū/, त्दु /tdu/.

The above permutations and combinations result in 7 major syllable (akshar) types. Of these the last type introduces the problem of the number of consonant clusters. ISCII (91, p.23) provides for up to three consonant clusters as the “worst case” i.e. the largest possible string. This is functional for modern Prakrits where the largest consonantal cluster rarely exceeds three consonant. Sanskrit is an exception where in a single word, four consonants can come together: कार्त्स्न्य /kārtsnya/ "wholeness", "entirety".

This means that theoretically the following forms can be postulated:

1. Vowel Set: With the Vowel as the node.  
V VD
2. Consonant set: With the Consonant as the node (an implicit or modified vowel is pre-implied).

Node	Mātrā	Modifier	Mātrā+Modifier
C	CM	CD	CMD
CHC	CHCM	CHCD	CHCMD
CHCHC	CHCHCM	CHCHCD	CHCHCMD
CHCHCHC	CHCHCHCM	CHCHCHCD	CHCHCHCMD

A total number of 16 theoretical syllables is therefore possible. It will be seen that the written syllable (akshar) is not very different in structure from the phonetic syllable and

<sup>9</sup> This character represents phonetically the weak implicit vowel, termed as schwa and often shown as /a/ also.



that the movement from the written to the spoken levels is made feasible by application of certain rules.

This formal structure of the syllable (akshar) explained above is common to all Brahmi based scripts (with a few variations). It will form the basis of an exhaustive description of the characters as well as their ligatural representations.

## 6.3 SCRIPT GRAMMAR PROPER

This section lays down in detail the different parameters of the Script Grammar for Konkani. These are:

- 6.3.1. The Character Set of Konkani.
- 6.3.2. The Consonant mātṛā combinations of Konkani.
- 6.3.3. The Ligature Set of Konkani.
- 6.3.4. Collocation Order of Konkani
- 6.3.5. Cardinal Numbers used in Konkani.

### 6.3.1. The Character Set of Konkani.

This section provides detailed information about the characters in the language and the list of the same and also more importantly shows the manner in which the character is to be written. Each subsection comprises therefore two parts: the basic character set and the shape each character should have, as mandated by the experts who have designed the script grammar of Konkani.

This comprises the following:

- 6.3.1.1. The Consonant Set
- 6.3.1.2. The Vowel Set
- 6.3.1.3. The Mātṛā Set
- 6.3.1.4. Displaced Catenators
- 6.3.1.5. Shape of the combination of ra (rakar, repha)
- 6.3.1.6. The Set of Diacritics
- 6.3.1.7. Numerals
- 6.3.1.8. Punctuation marks
- 6.3.1.9. Other symbols

Each of these will be analysed in detail:

#### 6.3.1.1. The Consonant Set

The Consonant set of Konkani comprises the following characters:

Basic Consonant inventory arranged as per their Vargas.

	-voiced -aspirated	-voiced +aspirated	+voiced -aspirated	+voiced +aspirated	Nasal
Velar	क	ख	ग	घ	ङ
Palatal	च	छ	ज	झ	ञ
Retroflex	ट	ठ	ड	ढ	ण
Dental	त	थ	द	ध	न
B-labial	प	फ	ब	भ	म

Other consonants

य	र	र <sup>10</sup>	ल	ळ
व	श	ष	स	ह

**Note:** Ligatures क्षत्रज्ञ are not listed in the consonants list.

The exact shapes as desired by the experts are provided in the table below:

	-voiced -aspirated	-voiced +aspirated	+voiced -aspirated	+voiced +aspirated	Nasal
Velar	क	ख	ग	घ	ङ
Palatal	च	छ	ज	झ	ञ
Retroflex	ट	ठ	ड	ढ	ण
Dental	त	थ	द	ध	न
B-labial	प	फ	ब	भ	म

Other consonants

य	र	र <sup>11</sup>	ल	ळ
व	श	ष	स	ह

### 6.3.1.2. The Vowel Set

The Vowel set of Konkani is as under:

अ	DEVANĀGARĪ LETTER A
आ	DEVANĀGARĪ LETTER AA
इ	DEVANĀGARĪ LETTER I
ई	DEVANĀGARĪ LETTER II
उ	DEVANĀGARĪ LETTER U

<sup>10</sup> There are differences of opinion re. this character used for generating the eyelash ra

<sup>11</sup> There are differences of opinion re. this character used for generating the eyelash ra

ऊ	DEVANĀGARĪ LETTER UU
ऋ	DEVANĀGARĪ LETTER VOCALIC R
ए	DEVANĀGARĪ LETTER E
अँ	DEVANĀGARĪ LETTER CANDRA A
ऐ	DEVANĀGARĪ LETTER AI
ओ	DEVANĀGARĪ LETTER O
औ	DEVANĀGARĪ LETTER CANDRA O
औ	DEVANĀGARĪ LETTER AU

As per expert recommendations the character set should be written as under:

अ आ इ ई उ ऊ ऋ ए अँ ऐ ओ औ

#### 6.3.1.3. The Mātrā Set

The Mātrā (Vowel Modifier Set) of Marathi is as under:

Mātrās Names	Mātrās Sign	Where is it used ?	Consonant Shapes formed
1. Devanāgarī sign AA	ा	आ	क् + आ = का
2. Devanāgarī sign I ( stands to the left of the consonat)	ि	इ	क् + इ = कि
3. Devanāgarī sign II	ी	ई	क् + ई = की
4. Devanāgarī sign U	ु	उ	क् + उ = कु
5. Devanāgarī sign UU	ू	ऊ	क् + ऊ = कू
6. Devanāgarī sign vocalic R	ृ	ऋ	क् + ऋ = कृ
7. Devanāgarī sign E	े	ए (There can not be any matra on this)	क् + ए = के
8. Devanāgarī sign AI	ै	ऐ (There can not be any matra on this)	क् + ऐ = कै

9.Devanāgarī sign O	ो	ओ	क् + ओ = को
10. Devanāgarī sign AU	ौ	औ	क् + औ = कौ
11. Devanāgarī sign Candra E	ँ	अँ	क् + अँ = कै
12. Devanāgarī sign Candra O	ॉ	आँ	क् + आँ = काँ

As per expert recommendations the character set should be written as under:

ा ि िी ु ू ृ े ँ ै ो ँ ौ

#### 6.3.1.4. Displaced Catenators

Under normal circumstances Vowel Modifiers also known as catenators (since they concatenate to the preceding consonant) in Brahmi based scripts are written from left to right in linear order (with the exception of Consonant stacks). However certain modifiers are displaced and are placed to the left of the consonant to which they concatenate. As a general rule in all Devanāgarī script driven languages there is only one displaced catenator: ि

CATENATOR	POSITION	EXAMPLE
ि	To left of character	कि, रि, पि

#### 6.3.1.5. Shape of the combination of ra (rakar, repha)

The र takes a variety of shapes known as rakar and repha (rafar) depending on its position. When conjoined before a consonant by means of the halanta, it changes shape and is placed on top of the consonant or consonant clusters to which it relates. This is called a repha or rafar. Konkani admits a special repha known as eyelash ra. When it is conjoined after a consonant with the help of a halanta, it appends to the consonant in the shape of a slanting stroke attached to the stem (side rakar) or in the case of consonants which have no stem such as ट, it is appended in the shape of a ^ to the bottom of the character (bottom rakar). Konkani has the following combinations of ra:

#### RAFARS

Top rafar: कँ तँ example

स्वर्ग	पूर्व
--------	-------

#### RAKARS

1. Bottom rakar द्र ड्र

2. Side rakar      क्र      च
3. Eyelash RA      ँ      ह्य

The eyelash ra is appended only to two consonants: य and ह

Examples of words where Rakars are used in Konkani language are given below:

Bottom rakar	महाराष्ट्र, ड्रम
Side rakar	क्रमांक, प्रशांत
Eyelash RA	ह्रस्व

### 6.3.1.6. Diacritics

These are as under in the case of Konkani:

ं: - Anuswar

्: - Halant

ः: - Visarga

ऌ: - For extra length with long vowels e.g. / उपदेशेऽ जनुनासिक /

Chandrabindu is not part of the Konkani character set and is used only to denote a अँ + ं and औँ + ं as in बैँक

### 6.3.1.7. Numerals

Following are the numbers used in Konkani language. Use of English numerals occurs in handwritten text as well as some of the official documents. The English (Latino-Arabic set: 0,1,2,3,4,5,6,7,8,9) is used only in books on mathematics where complicated equations and computations have to be shown.

Numeral Shapes	Explanation
०	Devanāgarī Digit Zero
१	Devanāgarī Digit One
२	Devanāgarī Digit Two
३	Devanāgarī Digit Three
४	Devanāgarī Digit Four
५	Devanāgarī Digit Five

६	Devanāgarī Digit Six
७	Devanāgarī Digit Seven
८	Devanāgarī Digit Eight
९	Devanāgarī Digit Nine

The shapes of some of these numbers vary

depending upon the user.

Variant shapes are provided below:

### 6.3.1.8. Punctuation Markers

Konkani uses punctuation markers from the Latin set. such as . , ; : “ ‘ ( ) [ ] etc.

Purna and Deerga Virama (full-stop/danda) Devanāgarī code block: 0964, 0965 ।,॥ are used in poetry alone.

ॐ नमो जी आद्या| वेद प्रतिपाद्या| जय जय स्वसंवेद्या| आत्मरूपा ||१||

देवा तूंचि गणेशु| सकलार्थमतिप्रकाशु| म्हणे निवृत्तिदासु| अवधारिजो जी ||२||

ज्ञानेश्वरी भावार्थदीपिका अध्याय १.

A list of punctuations is provided below:

Sr. No.	Name of the marker	Marker Shape
01	Full Stop or Period	.
02	Question Mark	?
03	Exclamation Mark	!
04	Apostrophe	,
05	Semi Colon	;
06	Colon	:
07	Hyphen	-
08	Dash	--
09	Ellipsis mark	...
10	Oblique	/
11	Double quotation mark	" "
12	Single quotation mark	‘ ‘
13	Cross	XXX
14	As Above	-- " --
15	Round Brackets	( )
16	Square Brackets	[ ]
17	Curly Brackets	{ }
18	Abbreviation Sign	(°)
19	Devanāgarī Danda	

### 6.3.1.9 Other Symbols

These are religious, currency markers etc. included in Unicode:

ॐ: Om (as written in Konkani)

₹: Rupee Sign as mandated by Government of India.

### 6.3.2. Consonant Mātrā Combinations.

These refer to the shapes generated when a Mātrā is adjoined to the Consonant. The layout of these is in the shape of a matrix where the first horizontal row refers to the active consonant and the first vertical column refers to the vowel-modifier.

Due to constraints of space and also clarity, for each class a series of 3 tables are provided.

Table 1: क ख ग घ ङ च छ ज झ ञ

Table 2: ट ठ ड ढ ण त थ द ध न

Table 3: प फ ब भ म य र<sup>12</sup> ल व श  
ष स ह ळ

Wherever there is an X it implies that the combination does not exist. For the font developer this is an indication that for this particular combination which is not possible in the language but needs to be accommodated in the font table, a simple linear combination be provided.

e.g. Although the combination of ङ + Mātrā is theoretically not possible it needs to be handled at the font level in the anticipation that a user could type this combination. The font would show the following: ङा

The classes are as under:

6.3.2.1. refers to a simple concatenation of Consonant and Mātrā combinations.

6.3.2.2. refers to a concatenation of Consonant and Mātrā + Nasal marker combinations. Other diacritics such as avagraha and visarga have been avoided, since these are linear in nature, are adjoined to the combination and do not in any way modify the structure of the shapes.

---

<sup>12</sup> र is not included in the list since it can occur only in combinations with य & ह and cannot constitute a stand-alone consonant



### 6.3.2.1 Consonant and Mātrā combinations.

This set refers to a simple concatenation of Consonant and Mātrā.

#### Consonant and Mātrā combinations Set 1

	क	ख	ग	घ	ङ	च	छ	ज	झ	ञ
ा	का	खा	गा	घा	ङा	चा	छा	जा	झा	जा
ि	कि	खि	गि	घि	ङि	चि	छि	जि	झि	ञि
ी	की	खी	गी	घी	ङी	ची	छी	जी	झी	जी
ु	कु	खु	गु	घु	ङु	चु	छु	जु	झु	जु
ू	कू	खू	गू	घू	ङू	चू	छू	जू	झू	जू
ृ	कृ	खृ	गृ	घृ	ङृ	चृ	छृ	जृ	झृ	जृ
े	के	खे	गे	घे	ङे	चे	छे	जे	झे	जे
ै	कै	खै	गै	घै	ङै	चै	छै	जै	झै	जै
ो	को	खो	गो	घो	ङो	चो	छो	जो	झो	जो
ौ	कौ	खौ	गौ	घौ	ङौ	चौ	छौ	जौ	झौ	जौ
ँ	कँ	खँ	गँ	घँ	ङँ	चँ	छँ	जँ	झँ	जँ
ॉ	काँ	खाँ	गाँ	घाँ	ङाँ	चाँ	छाँ	जाँ	झाँ	जाँ

Remark 1- ङ and ञ are rarely used only as the first members of clusters and mostly as परसवर्ण or alternatives of अनुस्वारः

#### Consonant and Mātrā combinations Set 2

This set is in continuation of set 1 which shows consonant and Matra combinations.

	ट	ठ	ड	ढ	ण	त	थ	द	ध	न
ा	टा	ठा	डा	ढा	णा	ता	था	दा	धा	ना
ि	टि	ठि	डि	ढि	णि	ति	थि	दि	धि	नि
ी	टी	ठी	डी	ढी	णी	ती	थी	दी	धी	नी
ु	टु	ठु	डु	ढु	णु	तु	थु	दु	धु	नु
ू	टू	ठू	डू	ढू	णू	तू	थू	दू	धू	नू
ृ	टृ	ठृ	डृ	ढृ	णृ	तृ	थृ	दृ	धृ	नृ
े	टे	ठे	डे	ढे	णे	ते	थे	दे	धे	ने
ै	टै	ठै	डै	ढै	णै	तै	थै	दै	धै	नै
ो	टो	ठो	डो	ढो	णो	तो	थो	दो	धो	नो

ौ	टै	ठै	डै	ढै	णै	तै	थै	दै	धै	नै
ँ	टँ	ठँ	डँ	ढँ	णँ	तँ	थँ	दँ	धँ	नँ
ॉ	टॉ	ठॉ	डॉ	ढॉ	णॉ	तॉ	थॉ	दॉ	धॉ	नॉ

### Consonant and Mātrā combinations Set 3

This set is in continuation of set 2 which shows consonant and Matra combinations.

	प	फ	ब	भ	म	य	र	ल	व	श	ष	स	ह	ळ
ा	पा	फा	बा	भा	मा	या	रा	ला	वा	शा	षा	सा	हा	ळा
ि	पि	फि	बि	भि	मि	यि	रि	लि	वि	शि	षि	सि	हि	ळि
ी	पी	फी	बी	भी	मी	यी	री	ली	वी	शी	षी	सी	ही	ळी
ु	पु	फु	बु	भु	मु	यु	रु	लु	वु	शु	षु	सु	हु	ळु
ू	पू	फू	बू	भू	मू	यू	रू	लू	वू	शू	षू	सू	हू	ळू
ृ	पृ	फृ	बृ	भृ	मृ	यृ	ऋ	लृ	वृ	शृ	षृ	सृ	हृ	ळृ
े	पे	फे	बे	भे	मे	ये	रे	ले	वे	शे	षे	से	हे	ळे
ै	पै	फै	बै	भै	मै	यै	रै	लै	वै	शै	षै	सै	है	ळै
ो	पो	फो	बो	भो	मो	यो	रो	लो	वो	शो	षो	सो	हो	ळो
ौ	पौ	फौ	बौ	भौ	मौ	यौ	रौ	लौ	वौ	शौ	षौ	सौ	हौ	ळौ
ँ	पँ	फँ	बँ	भँ	मँ	यँ	रँ	लँ	वँ	शँ	षँ	सँ	हँ	ळँ
ॉ	पॉ	फॉ	बॉ	भॉ	मॉ	यॉ	रॉ	लॉ	वॉ	शॉ	षॉ	सॉ	हॉ	ळॉ

### 6.3.2.2 Consonant and Mātrā +Nasal combinations.

This set refers to a Consonant and Mātrā + Nasal marker combinations.

#### Consonant and Mātrā + Nasal combinations - Set 1

	क	ख	ग	घ	ङ	च	छ	ज	झ	ञ
ं	कं	खं	गं	घं	ङं	चं	छं	जं	झं	ञं
ां	कां	खां	गां	घां	ङां	चां	छां	जां	झां	जां
िं	किं	खिं	गिं	घिं	ङिं	चिं	छिं	जिं	झिं	ञिं
ीं	कीं	खीं	गीं	घीं	ङीं	चीं	छीं	जीं	झीं	जीं
ृं	कृं	खृं	गृं	घृं	ङृं	चृं	छृं	जृं	झृं	ञृं
ुं	कुं	खुं	गुं	घुं	ङुं	चुं	छुं	जुं	झुं	नुं
ूं	कूं	खूं	गूं	घूं	ङूं	चूं	छूं	जूं	झूं	जूं
ें	कें	खें	गें	घें	ङें	चें	छें	जें	झें	जें
ैं	कैं	खैं	गैं	घैं	ङैं	चैं	छैं	जैं	झैं	जैं
ों	कों	खों	गों	घों	ङों	चों	छों	जों	झों	जों
ौं	कौं	खौं	गौं	घौं	ङौं	चौं	छौं	जौं	झौं	जौं
ँ	कँ	खँ	गँ	घँ	ङँ	चँ	छँ	जँ	झँ	जँ
ॉ	काँ	खाँ	गाँ	घाँ	ङाँ	चाँ	छाँ	जाँ	झाँ	जाँ

#### Consonant and Mātrā +Nasal combinations - Set 2

This set is in continuation of set 1 above which shows combinations of Consonant and Mātrā + Nasal marker

	ट	ठ	ड	ढ	ण	त	थ	द	ध	न
ं	टं	ठं	डं	ढं	णं	तं	थं	दं	धं	नं
ां	टां	ठां	डां	ढां	णां	तां	थां	दां	धां	नां
िं	टिं	ठिं	डिं	ढिं	णिं	तिं	थिं	दिं	धिं	निं
ीं	टीं	ठीं	डीं	ढीं	णीं	तीं	थीं	दीं	धीं	नीं

ृं	टं	ठं	डं	ढं	णं	तं	थं	दं	धं	नं
ुं	टुं	ठुं	डुं	ढुं	णुं	तुं	थुं	दुं	धुं	नुं
ूं	टू	ठू	डू	ढू	णू	तू	थू	दू	धू	नू
ें	टे	ठे	डे	ढे	णे	ते	थे	दे	धे	ने
ैं	टै	ठै	डै	ढै	णै	तै	थै	दै	धै	नै
ों	टो	ठो	डो	ढो	णो	तो	थो	दो	धो	नो
ौं	टौ	ठौ	डौ	ढौ	णौ	तौ	थौ	दौ	धौ	नौ
ँ	टँ	ठँ	डँ	ढँ	णँ	तँ	थँ	दँ	धँ	नँ
ॉं	टाँ	ठाँ	डाँ	ढाँ	णाँ	ताँ	थाँ	दाँ	धाँ	नाँ

### Consonant and Mātrā +Nasal combinations - Set 3

This set is in continuation of set 2 above which shows combinations of Consonant and Mātrā + Nasal marker

	प	फ	ब	भ	म	य	र	ल	व	श	ष	स	ह	ळ
ं	पं	फं	बं	भं	मं	यं	रं	लं	वं	शं	षं	सं	हं	ळं
ां	पां	फां	बां	भां	मां	यां	रां	लां	वां	शां	षां	सां	हां	ळां
िं	पिं	फिं	बिं	भिं	मिं	यिं	रिं	लिं	विं	शिं	षिं	सिं	हिं	ळिं
ीं	पीं	फीं	बीं	भीं	मीं	यीं	रीं	लीं	वीं	शीं	षीं	सीं	हीं	ळीं
ृं	पृं	फृं	बृं	भृं	मृं	यृं	रृं	लृं	वृं	शृं	षृं	सृं	हृं	ळृं
ुं	पुं	फुं	बुं	भुं	मुं	युं	रुं	लुं	वुं	शुं	षुं	सुं	हुं	ळुं
ूं	पू	फू	बू	भू	मू	यू	रू	लू	वू	शू	षू	सू	हू	ळू
ें	पे	फे	बे	भे	मे	ये	रे	ले	वे	शे	षे	से	हे	ळे
ैं	पै	फै	बै	भै	मै	यै	रै	लै	वै	शै	षै	सै	है	ळै
ों	पो	फो	बो	भो	मो	यो	रो	लो	वो	शो	षो	सो	हो	ळो
ौं	पौ	फौ	बौ	भौ	मौ	यौ	रौ	लौ	वौ	शौ	षौ	सौ	हौ	ळौ
ँ	पँ	फँ	बँ	भँ	मँ	यँ	रँ	लँ	वँ	शँ	षँ	सँ	हँ	ळँ
ॉं	पाँ	फाँ	बाँ	भाँ	माँ	याँ	राँ	लाँ	वाँ	शाँ	षाँ	साँ	हाँ	ळाँ

### 6.3.3. The Ligature Set of Konkani.

Konkani has a large set of ligatural forms. These are combinations of Consonant+Halanta+Consonant (CHC) or CHCHC or even rarer CHCHCHC. The CHC combinations which are the most frequent are arranged in the shape of a matrix: the abscissa or horizontal axis refers to the Consonant which constitutes the ligature and the ordinate or vertical axis shows the consonant which forms the ligature and which is followed by a halanta.

As in 6.3.2. the ligature sets are divided into the following

6.3.3.1 CHC (in a matrix)

6.3.3.2 CHCHC

6.3.3.3. CHCHCHC

#### 6.3.3.1. CHC ( combination of two consonanats)

These ligatures are presented as in the earlier case of Consonant+Mātrā combinations in three sets. A lot of “slots” have an X marked, showing that the experts have deemed that such a ligature is not possible in the language. However in these cases, the font developer is to assume that the ligature is linear in nature.

The following set shows a combination of two consonants. To know how particular combinations forms, select one consonant from the first column and second from first row. For eg. Combination of consonant “क” and “क” is ligature “क्क”.

CHC( combination of two consonanats) - Set 1										
	क	ख	ग	घ	ङ	च	छ	ज	झ	ञ
क	क्क	क्ख	क्ग	क्घ	क्ङ	क्च	क्छ	क्ज	क्झ	क्ञ
ख	क्क	क्ख	क्ग	क्घ	क्ङ	क्च	क्छ	क्ज	क्झ	क्ञ
ग	गक	गख	गग	गघ	गङ	गच	गछ	गज	गझ	गञ
घ	घक	घख	घग	घघ	घङ	घच	घछ	घज	घझ	घञ
ङ	ङक	ङख	ङग	ङघ	ङङ	ङच	ङछ	ङज	ङझ	ङञ
च	चक	चख	चग	चघ	चङ	चच	चछ	चज	चझ	चञ
छ	छक	छख	छग	छघ	छङ	छच	छछ	छज	छझ	छञ
ज	जक	जख	जग	जघ	जङ	जच	जछ	जज	जझ	जञ
झ	झक	झख	झग	झघ	झङ	झच	झछ	झज	झझ	झञ
ञ	ञक	ञख	ञग	ञघ	ञङ	ञच	ञछ	ञज	ञझ	ञञ
ट	टक	टख	टग	टघ	टङ	टच	टछ	टज	टझ	टञ
ठ	ठक	ठख	ठग	ठघ	ठङ	ठच	ठछ	ठज	ठझ	ठञ
ड	डक	डख	डग	डघ	डङ	डच	डछ	डज	डझ	डञ
ढ	ढक	ढख	ढग	ढघ	ढङ	ढच	ढछ	ढज	ढझ	ढञ
ण	णक	णख	णग	णघ	णङ	णच	णछ	णज	णझ	णञ
त	तक	तख	तग	तघ	तङ	तच	तछ	तज	तझ	तञ

१	थक	थ्र	था	थघ	थङ	थच	थछ	थज	थझ	थञ
द्	दक	द्र	दा	दघ	दङ	दच	दछ	दज	दझ	दञ
६	धक	ध्र	धा	धघ	धङ	धच	धछ	धज	धझ	धञ
न्	नक	न्र	ना	नघ	नङ	नच	नछ	नज	नझ	नञ
८	पक	प्र	पा	पघ	पङ	पच	पछ	पज	पझ	पञ
फ	फक	फ्र	फा	फघ	फङ	फच	फछ	फज	फझ	फञ
ब	बक	ब्र	बा	बघ	बङ	बच	बछ	बज	बझ	बञ
भ	भक	भ्र	भा	भघ	भङ	भच	भछ	भज	भझ	भञ
म्	मक	म्र	मा	मघ	मङ	मच	मछ	मज	मझ	मञ
८	यक	य्र	या	यघ	यङ	यच	यछ	यज	यझ	यञ
ल्	लक	ल्र	ला	लघ	लङ	लच	लछ	लज	लझ	लञ
८	वक	व्र	वा	वघ	वङ	वच	वछ	वज	वझ	वञ
१	शक	श्र	शा	शघ	शङ	शच	शछ	शज	शझ	शञ
८	षक	ष्र	षा	षघ	षङ	षच	षछ	षज	षझ	षञ
८	स्क	स्त्र	सा	सघ	सङ	सच	सछ	सज	सझ	सञ
ह	हक	ह्र	हा	हघ	हङ	हच	हछ	हज	हझ	हञ
ळ	ळक	ळ्र	ळा	ळघ	ळङ	ळच	ळछ	ळज	ळझ	ळञ

The following set shows a combination of two consonants. To know how particular combinations forms, select one consonant from the first column and second from first row. For eg. “ट” is ligature “कट”.

CHC( combination of two consonanats) – Set 2										
	ट	ठ	ड	ढ	ण	त	थ	द	ध	न
क	कट	कठ	कड	कढ	कण	क्त	कथ	कद	कध	कन
ख	खट	खठ	खड	खढ	खण	ख्त	खथ	खद	खध	खन
ग	गट	गठ	गड	गढ	गण	गत	गथ	गद	गध	गन
घ	घट	घठ	घड	घढ	घण	घ्त	घथ	घद	घध	घन
ङ	ङट	ङठ	ङड	ङढ	ङण	ङ्त	ङथ	ङद	ङध	ङन
च	चट	चठ	चड	चढ	चण	च्त	चथ	चद	चध	चन
छ	छट	छठ	छड	छढ	छण	छ्त	छथ	छद	छध	छन
ज	जट	जठ	जड	जढ	जण	ज्त	जथ	जद	जध	जन
झ	झट	झठ	झड	झढ	झण	झ्त	झथ	झद	झध	झन
ञ	ञट	ञठ	ञड	ञढ	ञण	ञ्त	ञथ	ञद	ञध	ञन

ट	ट्ट	टर	टड	टढ	टण	टत	टथ	टद	टध	टन
ठ	ठ्ट	ठर	ठड	ठढ	ठण	ठत	ठथ	ठद	ठध	ठन
ड	ड्ट	डर	डड	डढ	डण	डत	डथ	डद	डध	डन
ढ	ढ्ट	ढर	ढड	ढढ	ढण	ढत	ढथ	ढद	ढध	ढन
ण	ण्ट	णर	णड	णढ	णण	णत	णथ	णद	णध	णन
त	त्त	तर	तड	तढ	तण	तत	तथ	तद	तध	तन
थ	थ्ट	थर	थड	थढ	थण	थत	थथ	थद	थध	थन
द	द्ट	दर	दड	दढ	दण	दत	दथ	दद	दध	दन
ध	ध्ट	धर	धड	धढ	धण	धत	धथ	धद	धध	धन
न	न्त	नर	नड	नढ	नण	नत	नथ	नद	नध	नन
प	प्प	पर	पड	पढ	पण	पत	पथ	पद	पध	पन
फ	फ्ट	फर	फड	फढ	फण	फत	फथ	फद	फध	फन
ब	ब्ट	बर	बड	बढ	बण	बत	बथ	बद	बध	बन
भ	भ्ट	भर	भड	भढ	भण	भत	भथ	भद	भध	भन
म	म्ट	मर	मड	मढ	मण	मत	मथ	मद	मध	मन
य	य्ट	यर	यड	यढ	यण	यत	यथ	यद	यध	यन
र	रत्त	रर	रड	रढ	रण	रत	रथ	रद	रध	रन
ल	ल्ट	लर	लड	लढ	लण	लत	लथ	लद	लध	लन
ळ	ळ्ट	ळर	ळड	ळढ	ळण	ळत	ळथ	ळद	ळध	ळन
व	वत्त	वर	वड	वढ	वण	वत	वथ	वद	वध	वन
श	श्ट	शर	शड	शढ	शण	शत	शथ	शद	शध	शन
ष	ष्ट	षर	षड	षढ	षण	षत	षथ	षद	षध	षन
स	स्ट	सर	सड	सढ	सण	सत	सथ	सद	सध	सन
ह	ह्ट	हर	हड	हढ	हण	हत	हथ	हद	हध	हन

The following set shows a combination of two consonants. To know how particular combinations forms, select one consonant from the first column and second from first row. For eg. Combination of consonant “क” and “प” is कप

CHC( combination of two consonants) – Set 3														
	प	फ	ब	भ	म	य	र	ल	व	श	ष	स	ह	ळ
क	कप	कफ	कब	कभ	कम	कय	क्र	कल	कव	कश	कष	कस	कह	कळ
ख	खप	खफ	खब	खभ	खम	खय	ख्र	खल	खव	खश	खष	खस	खह	खळ
ग	गप	गफ	गब	गभ	गम	गय	ग्र	गल	गव	गश	गष	गस	गह	गळ
घ	घप	घफ	घब	घभ	घम	घय	घ्र	घल	घव	घश	घष	घस	घह	घळ

ઙ	ઙપ	ઙફ	ઙબ	ઙભ	ઙમ	ઙય	ઙ્ર	ઙલ	ઙવ	ઙશ	ઙષ	ઙસ	ઙહ	ઙઠ
ચ	ચપ	ચફ	ચબ	ચભ	ચમ	ચય	ચ્ર	ચ્લ	ચ્વ	ચ્શ	ચ્ષ	ચ્સ	ચ્હ	ચ્ઠ
છ	છપ	છફ	છબ	છભ	છમ	છય	છ્ર	છલ	છવ	છશ	છષ	છસ	છહ	છઠ
જ	જપ	જફ	જબ	જભ	જમ	જય	જ્ર	જ્લ	જ્વ	જ્શ	જ્ષ	જ્સ	જ્હ	જ્ઠ
ઙ	ઙપ	ઙફ	ઙબ	ઙભ	ઙમ	ઙય	ઙ્ર	ઙલ	ઙવ	ઙશ	ઙષ	ઙસ	ઙહ	ઙઠ
ઙ	ઙપ	ઙફ	ઙબ	ઙભ	ઙમ	ઙય	ઙ્ર	ઙલ	ઙવ	ઙશ	ઙષ	ઙસ	ઙહ	ઙઠ
ટ	ટપ	ટફ	ટબ	ટભ	ટમ	ટય	ટ્ર	ટ્લ	ટ્વ	ટ્શ	ટ્ષ	ટ્સ	ટ્હ	ટ્ઠ
ટ	ટપ	ટફ	ટબ	ટભ	ટમ	ટય	ટ્ર	ટ્લ	ટ્વ	ટ્શ	ટ્ષ	ટ્સ	ટ્હ	ટ્ઠ
ડ	ડપ	ડફ	ડબ	ડભ	ડમ	ડય	ડ્ર	ડ્લ	ડ્વ	ડ્શ	ડ્ષ	ડ્સ	ડ્હ	ડ્ઠ
ડ	ડપ	ડફ	ડબ	ડભ	ડમ	ડય	ડ્ર	ડ્લ	ડ્વ	ડ્શ	ડ્ષ	ડ્સ	ડ્હ	ડ્ઠ
ણ	ણપ	ણફ	ણબ	ણભ	ણમ	ણય	ણ્ર	ણ્લ	ણ્વ	ણ્શ	ણ્ષ	ણ્સ	ણ્હ	ણ્ઠ
ટ	ટપ	ટફ	ટબ	ટભ	ટમ	ટય	ટ્ર	ટ્લ	ટ્વ	ટ્શ	ટ્ષ	ટ્સ	ટ્હ	ટ્ઠ
થ	થપ	થફ	થબ	થભ	થમ	થય	થ્ર	થ્લ	થ્વ	થ્શ	થ્ષ	થ્સ	થ્હ	થ્ઠ
દ	દપ	દફ	દબ	દભ	દમ	દય	દ્ર	દ્લ	દ્વ	દ્શ	દ્ષ	દ્સ	દ્હ	દ્ઠ
ધ	ધપ	ધફ	ધબ	ધભ	ધમ	ધય	ધ્ર	ધ્લ	ધ્વ	ધ્શ	ધ્ષ	ધ્સ	ધ્હ	ધ્ઠ
ન	નપ	નફ	નબ	નભ	નમ	નય	ન્ર	ન્લ	ન્વ	ન્શ	ન્ષ	ન્સ	ન્હ	ન્ઠ
પ	પપ	પફ	પબ	પભ	પમ	પય	પ્ર	પ્લ	પ્વ	પ્શ	પ્ષ	પ્સ	પ્હ	પ્ઠ
પ	પપ	પફ	પબ	પભ	પમ	પય	પ્ર	પ્લ	પ્વ	પ્શ	પ્ષ	પ્સ	પ્હ	પ્ઠ
ભ	ભપ	ભફ	ભબ	ભભ	ભમ	ભય	ભ્ર	ભ્લ	ભ્વ	ભ્શ	ભ્ષ	ભ્સ	ભ્હ	ભ્ઠ
મ	મપ	મફ	મબ	મભ	મમ	મય	મ્ર	મ્લ	મ્વ	મ્શ	મ્ષ	મ્સ	મ્હ	મ્ઠ
ય	યપ	યફ	યબ	યભ	યમ	યય	ય્ર	ય્લ	ય્વ	ય્શ	ય્ષ	ય્સ	ય્હ	ય્ઠ
લ	લપ	લફ	લબ	લભ	લમ	લય	લ્ર	લ્લ	લ્વ	લ્શ	લ્ષ	લ્સ	લ્હ	લ્ઠ
વ	વપ	વફ	વબ	વભ	વમ	વય	વ્ર	વ્લ	વ્વ	વ્શ	વ્ષ	વ્સ	વ્હ	વ્ઠ
શ	શપ	શફ	શબ	શભ	શમ	શય	શ્ર	શ્લ	શ્વ	શ્શ	શ્ષ	શ્સ	શ્હ	શ્ઠ
ષ	ષપ	ષફ	ષબ	ષભ	ષમ	ષય	ષ્ર	ષ્લ	ષ્વ	ષ્શ	ષ્ષ	ષ્સ	ષ્હ	ષ્ઠ
સ	સપ	સફ	સબ	સભ	સમ	સય	સ્ર	સ્લ	સ્વ	સ્શ	સ્ષ	સ્સ	સ્હ	સ્ઠ
હ	હપ	હફ	હબ	હભ	હમ	હય	હ્ર	હ્લ	હ્વ	હ્શ	હ્ષ	હ્સ	હ્હ	હ્ઠ
ઝ	ઝપ	ઝફ	ઝબ	ઝભ	ઝમ	ઝય	ઝ્ર	ઝ્લ	ઝ્વ	ઝ્શ	ઝ્ષ	ઝ્સ	ઝ્હ	ઝ્ઠ



### 6.3.3.2 CHCHC ( combination of three consonants)

These are not as frequent as the CHC combinations. Only the major are listed below.

क्+त्+य = क्त्य	ष्+ट्+य=ष्ट्य
क्+त्+र =क्त्र	त्+त्+व=त्त्व
च+ त्+र =र्त्र	र+ध्+य=र्ध्य
ङ् + क् + त = ङ्क्त	ष्+ट्+व=ष्ट्व
र+च्+छ=र्च्छ	र+थ्+य=र्थ्य
र+ष्+य=र्ष्य	क्+ष्+य=क्ष्य
ल्+ल्+य= ल्ल्य	ष्+ट्+र=ष्ट्र
च्+छ+र=च्छ्र	द्+च+य=द्र्य
न्+द्+र=न्द्र	न्+द्+य=न्द्य
र+द्+र=र्द्र	उ+द्+न= उद्न
स्+त्+र=स्त्र	न्+द्+व=न्द्र
क्+ष्+म=क्ष्म	त्+क्+र=त्क्र
क्+ष्+ण=क्ष्ण	च+ म्+य=र्म्य
त्+त्+य= त्त्य	ट्+ट्+य= ट्ट्य
त्+स्+न=त्स्न	द्+ध्+य= द्ध्य
द्+भ्+य= द्भ्य	च+ ल्+य=र्ल्य
च्+छ+व=च्छ्व	द्+ब्+य= द्ब्य
ज्+ज्+व=ज्ज्व	च+ क्+य=र्क्य
ह्+म्+य= ह्म्य	द् + ध् + य = द्धय
न्+ध्+य=न्द्य	च+ ध् + व = र्ध्व
च+ ह्+य= र्ह्य	त्+च+य=र्च्य

### 6.3.3.3.CHCHCHC ( Combination of four Consonanats)

This is rare and is found in the following

र्+ ष+ट्+य=ष्ट्य

And the Sanskrit word: कात्स्य

र्+त्+स्+न्+य=त्स्य

### 6.3.4 The Collation Order of Konkani.

The collation order refers to the order in which the characters in a given language are sorted. In the case of Konkani the following is the traditional sort order as determined by the experts. Unlike Hindi, Konkani treats क्ष, ज्ञ as individual characters and these come at the end of the list. The order as given below is pertinent to sorting by a computer program and is compliant with CLDR as laid down by Unicode and W3C.

ं	ः	अ	आ	इ	ई	उ	ऊ	ऋ	ए
अँ	ऐ	ओ	ऑ	औ	क	ख	ग	घ	ङ
च	छ	ज	झ	ञ	ट	ठ	ड	ढ	ण
त	थ	द	ध	न	प	फ	ब	भ	म
य	र	ल	व	श	ष	स	ह	ळ	क्ष
ज्ञ	ा	ि	ी	ु	ू	ृ	्र	े	ँ
ै	ो	ॉ	ौ	उ					

## 7. REFERENCES

1. <http://www.unicode.org>
2. ISCII'91

## **8. ANNEXURES**

### **Annexure 1: Names of experts who have contributed to the script grammar**

## Annexure 2: Unicode Table of Devanāgarī <sup>13</sup>

0900

Devanagari

097F

	090	091	092	093	094	095	096	097
0	ऐ	ठ	र	ी	ॐ	ऋ	०	
1	ँ	ऑ	ड	र	ु	ं	ॠ	ॠ
2	ं	ओ	ढ	ल	ॡ	ॡ	ॡ	अँ
3	ः	ओ	ण	ळ	ॢ	ॢ	ॢ	
4	ऐ	औ	त	ळ	ॣ	ॣ	ॣ	
5	अ	क	थ	व	॥	॥	॥	
6	आ	ख	द	श	॥	॥	॥	
7	इ	ग	घ	ष	॥	॥	॥	
8	ई	घ	न	स	॥	॥	॥	
9	उ	ड	न	ह	॥	॥	॥	
A	ऊ	च	प	॥	॥	॥	॥	
B	ऋ	छ	फ	॥	॥	॥	॥	
C	ॠ	ज	ब	॥	॥	॥	॥	
D	ॡ	झ	भ	॥	॥	॥	॥	
E	ॢ	ञ	म	॥	॥	॥	॥	
F	ॣ	ट	य	॥	॥	॥	॥	

The Unicode Standard 5.1, Copyright © 1991-2008 Unicode, Inc. All rights reserved.

65

<sup>13</sup> The Unicode chart provided is for version 5.1 since the Script Grammar was prepared at that time. No considerable change in the script grammar can be seen in the updated versions of Unicode, with the possible addition of the Rupee Sign U+02B9

## Based on ISCII 1988

## Various signs

0901	◌ं	DEVANAGARI SIGN CANDRABINDU = anusaka → 0310 ◌ं combining candrabindu
0902	◌ः	DEVANAGARI SIGN ANUSVARA = bindu
0903	◌ः	DEVANAGARI SIGN VISARGA

## Independent vowels

0904	अ	DEVANAGARI LETTER SHORT A
0905	आ	DEVANAGARI LETTER A
0906	इ	DEVANAGARI LETTER AA
0907	ई	DEVANAGARI LETTER I
0908	उ	DEVANAGARI LETTER II
0909	ऊ	DEVANAGARI LETTER U
090A	ऋ	DEVANAGARI LETTER UU
090B	ॠ	DEVANAGARI LETTER VOCALIC R
090C	ॡ	DEVANAGARI LETTER VOCALIC L
090D	ए	DEVANAGARI LETTER CANDRA E
090E	ऐ	DEVANAGARI LETTER SHORT E • for transcribing Dravidian short e
090F	ऑ	DEVANAGARI LETTER E
0910	ऋ	DEVANAGARI LETTER AI
0911	ॠ	DEVANAGARI LETTER CANDRA O
0912	ॡ	DEVANAGARI LETTER SHORT O • for transcribing Dravidian short o
0913	ओ	DEVANAGARI LETTER O
0914	औ	DEVANAGARI LETTER AU

## Consonants

0915	क	DEVANAGARI LETTER KA
0916	ख	DEVANAGARI LETTER KHA
0917	ग	DEVANAGARI LETTER GA
0918	घ	DEVANAGARI LETTER GHA
0919	ङ	DEVANAGARI LETTER NGA
091A	च	DEVANAGARI LETTER CA
091B	छ	DEVANAGARI LETTER CHA
091C	ज	DEVANAGARI LETTER JA
091D	झ	DEVANAGARI LETTER JHA
091E	ञ	DEVANAGARI LETTER NYA
091F	ट	DEVANAGARI LETTER TTA
0920	ठ	DEVANAGARI LETTER TTHA
0921	ड	DEVANAGARI LETTER DDA
0922	ढ	DEVANAGARI LETTER DDHA
0923	ण	DEVANAGARI LETTER NNA
0924	त	DEVANAGARI LETTER TA
0925	थ	DEVANAGARI LETTER THA
0926	द	DEVANAGARI LETTER DA
0927	ध	DEVANAGARI LETTER DHA
0928	न	DEVANAGARI LETTER NA
0929	न्	DEVANAGARI LETTER NNA • for transcribing Dravidian alveolar n = 0928 ण 093C ◌
092A	प	DEVANAGARI LETTER PA
092B	फ	DEVANAGARI LETTER PHA
092C	ब	DEVANAGARI LETTER BA
092D	भ	DEVANAGARI LETTER BHA
092E	म	DEVANAGARI LETTER MA
092F	य	DEVANAGARI LETTER YA
0930	र	DEVANAGARI LETTER RA

0931	ऌ	DEVANAGARI LETTER RRA • for transcribing Dravidian alveolar r • half form is represented as "Eyelash RA" = 0930 र 093C ◌
0932	ल	DEVANAGARI LETTER LA
0933	ळ	DEVANAGARI LETTER LLA
0934	ॢ	DEVANAGARI LETTER LLLA • for transcribing Dravidian l = 0933 ळ 093C ◌
0935	व	DEVANAGARI LETTER VA
0936	श	DEVANAGARI LETTER SHA
0937	ष	DEVANAGARI LETTER SSA
0938	स	DEVANAGARI LETTER SA
0939	ह	DEVANAGARI LETTER HA

## Various signs

093C	◌्	DEVANAGARI SIGN NUKTA • for extending the alphabet to new letters
093D	◌ः	DEVANAGARI SIGN AVAGRAHA

## Dependent vowel signs

093E	◌ः	DEVANAGARI VOWEL SIGN AA
093F	◌ः	DEVANAGARI VOWEL SIGN I • stands to the left of the consonant
0940	◌ः	DEVANAGARI VOWEL SIGN II
0941	◌ः	DEVANAGARI VOWEL SIGN U
0942	◌ः	DEVANAGARI VOWEL SIGN UU
0943	◌ः	DEVANAGARI VOWEL SIGN VOCALIC R
0944	◌ः	DEVANAGARI VOWEL SIGN VOCALIC RR
0945	◌ः	DEVANAGARI VOWEL SIGN CANDRA E = candra
0946	◌ः	DEVANAGARI VOWEL SIGN SHORT E • for transcribing Dravidian vowels
0947	◌ः	DEVANAGARI VOWEL SIGN E
0948	◌ः	DEVANAGARI VOWEL SIGN AI
0949	◌ः	DEVANAGARI VOWEL SIGN CANDRA O
094A	◌ः	DEVANAGARI VOWEL SIGN SHORT O • for transcribing Dravidian vowels
094B	◌ः	DEVANAGARI VOWEL SIGN O
094C	◌ः	DEVANAGARI VOWEL SIGN AU

## Various signs

094D	◌्	DEVANAGARI SIGN VIRAMA = halant (the preferred Hindi name) • suppresses inherent vowel
094E	◌्	<reserved>
094F	◌्	<reserved>
0950	◌्	DEVANAGARI OM
0951	◌्	DEVANAGARI STRESS SIGN UDATTA = Vedic tone svarita • mostly used for Rigvedic svarita, with rare use for Yajurvedic udatta • used also in Vedic texts written in other scripts
0952	◌्	DEVANAGARI STRESS SIGN ANUDATTA = Vedic tone anudatta • used also in Vedic texts written in other scripts
0953	◌्	DEVANAGARI GRAVE ACCENT
0954	◌्	DEVANAGARI ACUTE ACCENT

## Additional consonants

0958	◌्	DEVANAGARI LETTER QA = 0915 क 093C ◌
0959	◌्	DEVANAGARI LETTER KHHA = 0916 ख 093C ◌

095A	ग	DEVANAGARI LETTER GHHA = 0917 ग 093C ◌
095B	झ	DEVANAGARI LETTER ZA = 091C ञ 093C ◌
095C	ढ	DEVANAGARI LETTER DDDHA = 0921 ढ 093C ◌
095D	ण	DEVANAGARI LETTER RHA = 0922 ण 093C ◌
095E	फ	DEVANAGARI LETTER FA = 092B फ 093C ◌
095F	य	DEVANAGARI LETTER YYA = 092F य 093C ◌

**Additional vowels for Sanskrit**

0960	ॠ	DEVANAGARI LETTER VOCALIC RR
0961	ॡ	DEVANAGARI LETTER VOCALIC LL
0962	◌ॢ	DEVANAGARI VOWEL SIGN VOCALIC L
0963	◌ॣ	DEVANAGARI VOWEL SIGN VOCALIC LL

**Generic punctuation for scripts of India**

*These punctuation marks are for common use for the scripts of India despite being named "DEVANAGARI".*

0964	।	DEVANAGARI DANDA = purna viram • phrase separator
0965	॥	DEVANAGARI DOUBLE DANDA = deergh viram

**Digits**

0966	०	DEVANAGARI DIGIT ZERO
0967	१	DEVANAGARI DIGIT ONE
0968	२	DEVANAGARI DIGIT TWO
0969	३	DEVANAGARI DIGIT THREE
096A	४	DEVANAGARI DIGIT FOUR
096B	५	DEVANAGARI DIGIT FIVE
096C	६	DEVANAGARI DIGIT SIX
096D	७	DEVANAGARI DIGIT SEVEN
096E	८	DEVANAGARI DIGIT EIGHT
096F	९	DEVANAGARI DIGIT NINE

**Devanagari-specific additions**

0970	॥	DEVANAGARI ABBREVIATION SIGN
0971	˙	DEVANAGARI SIGN HIGH SPACING DOT

**Additional vowel for Marathi**

0972	ॡ	DEVANAGARI LETTER CANDRA A • Marathi
------	---	---

**Sindhi implosives**

097B	ɠ	DEVANAGARI LETTER GGA
097C	ɟ	DEVANAGARI LETTER JJA

**Glottal stop**

097D	ʔ	DEVANAGARI LETTER GLOTTAL STOP • used for writing Limbu in Devanagari • a glyph variant has the connecting top bar
------	---	--

**Sindhi implosives**

097E	ɗ	DEVANAGARI LETTER DDDA
097F	ɗ	DEVANAGARI LETTER BBBA

