

# **SCRIPT GRAMMAR FOR ODIA LANGUAGE**

**Prepared by**

**Technology Development for Indian Languages (TDIL) Programme**

**Department of Information Technology, Government of India**

**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 Odia.....	18
6.3.2. Consonant Mātrā Combinations. ....	27
6.3.3. The Ligature Set of Odia.....	40
6.3.4 The Collation Order of Odia. ....	50
7. REFERENCES .....	51
8. ANNEXURES .....	52
Annexure 1: Names of experts who have contributed to the script grammar.....	52
Annexure 2: Unicode Table of Odia.....	53

## 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 from the Devanāgarī script 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 Hindi /la/ is different from the Marathi /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 Marathi and Hindi /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 Hindi is different from Hindi although both languages share the same script. Thus in Hindi 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 & 639.3 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 æ 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 ऌ [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 ISCI91 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 /𑌦/ 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 **Eyesh ra**)

**Eyesh ra:** The eyesh ra is used in Konkani, Hindi and Marathi. It is treated as different from the ॠ (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).

**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.

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

---

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



**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 घ्य

**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-supposed 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 ISO 639.1. & 639.3

**Name of the Language: ODIA**

**ISO Mnemonics: *ori***

This refers to a one line description of the language and its mnemonic representation as per the ISO.

### 6.1.2. Identification of the writing system(s) used to inscribe the given language Odia is written using the Odia 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  $\text{କ୍ଷ}$  or alternatively form an entirely new shape  $\text{କ୍ଷ} = \text{ଝ}$ .

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 Odia language

None has been proposed by the experts who have mandated the script grammar.

## 6.2. CONFORMITY TO THE SYLLABLE STRUCTURE

**Odia language complies with the syllable (akshar) structure described above. It can admit up to 3 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 Odia. The full set of Odia is given below

### (C) Consonants

କ	ଖ	ଗ	ଘ	ଙ
ଚ	ଛ	ଜ	ଝ	ଞ
ଟ	ଠ	ଡ	ଢ	ଣ
ତ	ଥ	ଦ	ଧ	ନ
ପ	ଫ	ବ	ଭ	ମ
ୟ	ଋ	ର	ଲ	ଳ
ୱ	୲	ଶ	ଷ	ସ
ହ				

### Flapped Consonants

ଡ଼	ଢ଼
----	----

କ୍ଷ<sup>7</sup>

### (V) Vowels<sup>8</sup>

ଅ	ଆ	ଇ	ଈ	ଉ	ଊ	ଋ	ୠ	ଌ	ୡ	ଏ	ଐ	ଓ	ଔ
---	---	---	---	---	---	---	---	---	---	---	---	---	---

<sup>5</sup> Odia letter va not used in Odia

<sup>6</sup> Odia letter wa

<sup>7</sup> A ligature which is a combination of କ୍ଷ is often treated as a single consonant.

<sup>8</sup> Long rra and lla are rarely used in Odia, but are still listed here

(M) Mātrās or Vowel Modifiers<sup>7</sup>

ଠ	ଠି	ଠି	ଠୁ	ଠୁ	ଠୁ				ଠ	ଠି	ଠି	ଠି
---	----	----	----	----	----	--	--	--	---	----	----	----

(D) Diacritics

ଠ: Anuswara	Anuswara, a nasal, is denoted by a dot above the letter after which it is to be pronounced. This falls under Nasal category.
ଠ: Chandrabindu	Chandrabindu, a nasal, is denoted by a breve with a dot superposed above the letter after which it is to be pronounced. This falls under Nasal category.
ଠ8: Visarga	Visarga, denoted by two dots placed above the other.
ଠ: Avagraha <sup>9</sup>	For extra length with long vowels as seen in the Sanskrit text /उपदेशेऽजनुनासिक इत् /

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

(N)<sup>10</sup> Nukta - used in Odia

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	C	N,M,D,H
C	N	C,M,D,H
-	V	D
C, N	M	D
C, N,V,M	D	-
C, N	H	C

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

N can be preceded by C and followed by any one of the following: N,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.

<sup>9</sup> Avagraha is rarely used in Odia and is used for representing text in Sanskrit here the character is needed..

<sup>10</sup> The nukta is a small dot placed under a character in certain scripts to show that they are flapped or for deriving consonants required for Urdu क,ख,ग,ज़,फ़

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 mātṛā )
- The four other subsets viz. Mātṛās, Vowel Modifiers, Halanta and Nukta cannot constitute a syllable (akshar) by themselves or in combination among themselves.

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

1.1. A pure vowel all by itself: , ଐ/a/ ଐ/ā/ etc.

1.2. A vowel followed by a modifier i.e. either a nasal marker (*anuswara/chandrabindu*) or a visarga: ଐ°/ī̃ /, ଐଃ /āH/

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

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

2.2. A consonant<sup>11</sup> (with or without Nukta) followed by a mātṛā i.e. the inherent vowel being substituted by another vowel: କି /ki:/ କିଃ /de/

2.3. A consonant (with or without Nukta) followed by a modifier: କ° /k̃/, କଃ /haH/<sup>12</sup>/

2.4. A consonant (with or without Nukta) followed by a mātṛā and a modifier: କୁ° /kũ/, କୁଃ /duH/.

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

ତକ /tka/ ତକ୍ /tk̃/, ତକଃ /tkaH/ ତକୁ /tkũ/, ତକୁଃ /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" (secondary derivative from the adjective कर्त्स्न /kṛtsna/ meaning “whole, complete”.)

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).

<sup>11</sup> For purposes of Simplification, C here will automatically be treated as being also consonant+nukta C+ N

<sup>12</sup> This character represents phonetically the weak implicit vowel, termed as schwa and often shown as /a/ also.



Node	Mātrā	Modifier	Mātrā+Modifier
C <sup>13</sup>	CM	CD	CMD
CHC	CHCM	CHCD	CHCMD
CHCHC	CHCHCM	CHCHCD	CHCHCMD
CHCHCHC	CHCHCHCM	CHCHCHCD	CHCHCHCM

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

Since the formal structure script grammar of the syllable (akshar) is common to all Brahmi based scripts, it will not be treated in the sample template, but it will form the basis of an exhaustive description of the characters as well as their ligatural representations.

---

<sup>13</sup> C here will automatically be treated as being also consonant+nukta, C+N to simplify the explanation

## 6.3 SCRIPT GRAMMAR PROPER

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

- 6.3.1. The Character Set of Odia.
- 6.3.2. The Consonant mātrā combinations of Odia.
- 6.3.3. The Ligature Set of Odia.
- 6.3.4. Collocation Order of Odia

### 6.3.1. The Character Set of Odia.

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 Odia.

This comprises the following:

- 6.3.1.1. The Consonant Set
- 6.3.1.2. The Vowel Set<sup>9</sup>
- 6.3.1.3. The Mātrā Set<sup>9</sup>
- 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 Odia comprises the following characters:

A basic Consonant inventory arranged as per their Vargas.

	STOPS				NASAL
	-voiced -aspirated	-voiced +aspirated	+voiced -aspirated	+voiced +aspirated	+voiced -aspirated
Velar	କ	ଖ	ଗ	ଘ	ଙ
Palatal	ଚ	ଛ	ଜ	ଝ	ଞ
Retroflex	ଟ	ଠ	ଡ	ଢ	ଣ
Flaps			ଡ଼	ଢ଼	
Dental	ତ	ଥ	ଦ	ଧ	ନ
B-labial	ପ	ଫ	ବ	ଭ	ମ

### Other consonants

ଘ	ଘ	ର	ଲ	ଳ	
	ଶ	ଷ	ସ	ହ	ଶ

The ligature କ୍+ଷ=କ୍ଷ is not listed since it is a ligatural form

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

	STOPS				NASAL
	-voiced -aspirated	-voiced +aspirated	+voiced -aspirated	+voiced +aspirated	+voiced -aspirated
Velar	କ	ଖ	ଗ	ଘ	ଙ
Palatal	ଚ	ଛ	ଜ	ଝ	ଞ
Retroflex	ଟ	ଠ	ଡ	ଢ	ଣ
Flaps			ଡ଼	ଢ଼	
Dental	ତ	ଥ	ଦ	ଧ	ନ
B-labial	ପ	ଫ	ବ	ଭ	ମ

### Other consonants

ଘ	ଘ	ର	ଲ	ଳ	14
	ଶ	ଷ	ସ	ହ	ଶ

The ligature କ୍+ଷ=କ୍ଷ is not listed since it is a ligatural form

### 6.3.1.2. The Vowel Set

The Vowel set of Odia is as under:

ଅ	ODIA LETTER A
---	---------------

<sup>14</sup> For the explanation of Odia va and wa cf. <http://std.dkuug.dk/jtc1/sc2/WG2/docs/n2525.pdf> for the proposal

ଅ	ODIA LETTER AA
ଇ	ODIA LETTER I
ଈ	ODIA LETTER II
ଉ	ODIA LETTER U
ଊ	ODIA LETTER UU
ଋ	ODIA LETTER VOCALIC R
ୠ	ODIA LETTER VOCALIC RR
ୡ	ODIA LETTER VOCALIC L
ୢ	ODIA LETTER VOCALIC LL
ଏ	ODIA LETTER E
ଐ	ODIA LETTER AI
ଓ	ODIA LETTER O
ଔ	ODIA LETTER AU

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

ଅ	ଅ	ଇ	ଈ	ଉ	ଊ	ଋ	ୠ	ୡ	ୢ	ଏ	ଐ	ଓ	ଔ
---	---	---	---	---	---	---	---	---	---	---	---	---	---

### 6.3.1.3. The Mātrā Set

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

Mātrā Sign	Mātrā Name	Corresponding vowel	Consonant Shapes formed
	ODIA VOWEL SIGN AA	ଅ	କ <sup>୦</sup> + =କ

ᳵ	ODIA VOWEL SIGN I	ଇ	କ <sup>୦</sup> +ᳵ=କି
ᳶ	ODIA VOWEL SIGN II	ଈ	କ <sup>୦</sup> +ᳶ=କୀ
᳷	ODIA VOWEL SIGN U	ଉ	କ <sup>୦</sup> +᳷=କୁ
᳸	ODIA VOWEL SIGN UU	ଊ	କ <sup>୦</sup> +᳸=କୁ
᳹	ODIA VOWEL SIGN VOCALIC R	ଋ	କ <sup>୦</sup> +᳹=କୃ
	ODIA VOWEL SIGN VOCALIC RR	ୠ	କ <sup>୦</sup> +᳹=କୃ
	ODIA VOWEL SIGN VOCALIC L	ୡ	କ <sup>୦</sup> +᳹=କୃ
	ODIA VOWEL SIGN VOCALIC LL	ୢ	କ <sup>୦</sup> +᳹=କୃ
ᳺ	ODIA VOWEL SIGN E	ଏ	କ <sup>୦</sup> +ᳺ=କେ
᳻	ODIA VOWEL SIGN AI	ଐ	କ <sup>୦</sup> +᳻=କୈ
᳼	ODIA VOWEL SIGN O	ଓ	କ <sup>୦</sup> +᳼=କୋ
᳽	ODIA VOWEL SIGN AU	ଔ	କ <sup>୦</sup> +᳽=କୌ

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

।	ᳵ	ᳶ	᳷	᳸	᳹				ᳺ	᳻	᳼	᳽
---	---	---	---	---	---	--	--	--	---	---	---	---

#### 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). Odia has no such displaced catenator. However as in the case of certain Dravidian languages, Odia admits two part vowel dependent signs

CATENATOR	POSITION	EXAMPLE
᳼	᳼ to left and । to right	କ᳼।

୧୩	୧୩ to left and ୧୩ to right	୧୩୧୩
----	----------------------------	------

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

The ra consonant 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. In Odia when it is conjoined after a consonant with the help of a

halanta, it appends to the consonant in the shape of a ୧୩, placed below the consonant

Odia has the following combinations of ra:

#### RAFARS ୧୩

Top rafar ୧୩୩ ୧୩୩ as in

କ୧୩୩, ଆ୧୩୩ ଦିଆସ୍ତାଏ
---------------------

#### RAKARS ୧୩

Bottom rakar ୧୩୩ ୧୩୩ ୧୩୩

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

Bottom rakar	ରାଷ୍ଟ୍ରରାଜ୍ୟତ୍ୱ, ଦ୍ରାବିଡ଼
--------------	---------------------------

### 6.3.1.6. Diacritics

These are as under in the case of Odia:

- ୠ: Anunasika e.g. ଏକଂ ନଂ
- ୡ: Chandrabindu e.g. ଘାଙ୍ଗ, ଝାଙ୍ଗ
- ୢ Visarga e.g. ଅଃ, ଅଃପର
- ୣ: Avagraha For extra length with long vowels. In Odia this sign is used in the formation of a limited number of compound words formed according to the rules of Sanskrit grammar.

### 6.3.1.7. Numerals

The International number set (Latino-Arabic set: 0,1,2,3,4,5,6,7,8,9) is used in official documents in Odia.

Following are the numbers used in Odia language.

୦ ୧ ୨ ୩ ୪ ୫ ୬ ୭ ୮ ୯

Numeral	Explanation
---------	-------------

Shapes	
୦	Odia Digit Zero
୧	Odia Digit One
୨	Odia Digit Two
୩	Odia Digit Three
୪	Odia Digit Four
୫	Odia Digit Five
୬	Odia Digit Six
୭	Odia Digit Seven
୮	Odia Digit Eight
୯	Odia Digit Nine

**In addition Odia admits the following Fraction Signs<sup>15</sup>:**

୧	ODIA FRACTION ONE QUARTER
୨	ODIA FRACTION ONE HALF
୩	ODIA FRACTION THREE QUARTERS
୧/୧୬	ODIA FRACTION ONE SIXTEENTH
୧/୮	ODIA FRACTION ONE EIGHTH
୩/୧୬	ODIA FRACTION THREE SIXTEENTHS

---

<sup>15</sup> These are rarely used in everyday use



### 6.3.1.8. Punctuation Markers

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

Purna and Deerga Virama (full-stop/danda) Devanagari code block: U+0964, U+0965

।,॥. । is used to mark the full stop and । and ॥ are used for representing poetry as in the sample shlokas below:

अनिरोधम् अनुत्पादम् अनुच्छेदम् अशाश्वतम् । अनेकार्थम् अनानार्थम् अनागमम् अनिर्गमम् ॥

यः प्रतीत्यसमुत्पादं प्रपञ्चोपशमं शिवम् । देशयामास संबुद्धस्तं वन्दे वदतां वरम् ॥

ଅନିରୋଧମ୍ ଅନୁତ୍ପାଦମ୍ ଅନୁଚ୍ଛେଦମ୍ ଅଶାଶ୍ଵତମ୍ । ଅନେକାର୍ଥମ୍ ଅନାନାର୍ଥମ୍ ଅନାଗମମ୍ ଅନିର୍ଗମମ୍ ॥

ଯଃ ପ୍ରତୀତ୍ୟସମୁତ୍ପାଦଂ ପ୍ରପଞ୍ଚୋପଶମଂ ଶିବମ୍ । ଦେଶୟାମାସ ସଂବୁଦ୍ଧସ୍ତଂ ବନ୍ଦେ ବଦତାଂ ବରମ୍ ॥

A list of punctuations is provided below:

Sr. No.	Name of the marker	Marker Shape
1.	Full Stop or Period	.
2.	Question Mark	?
3.	Comma	,
4.	Exclamation Mark	!
5.	Apostrophe	'
6.	Semi Colon	;
7.	Colon	:
8.	Hyphen	-
9.	Dash	--
10.	Ellipsis mark	...
11.	Oblique	/
12.	Double quotation mark	" "
13.	Single quotation mark	‘ ’
14.	Cross	XXX
15.	As Above	-- " --
16.	Round Brackets	( )
17.	Square Brackets	[ ]
18.	Curly Brackets	{ }
19.	Odia Danda	।
20.	Odia Double Danda	॥

### 6.3.1.9 Other Symbols

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

✓ Odia Isshar

₹: Rupee Sign as mandated by Government of India.

Fraction signs are listed in 6.3.1.7 above



### 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 for reasons of clarity, for each class a series of 3 tables are provided.

Table 1:	କ	ଖ	ଗ	ଘ	ଙ	ଚ	ଛ	ଜ	ଝ	ଞ
Table 2:	ଟ	ଠ	ଡ	ଢ	ଣ	ତ	ଥ	ଦ	ଧ	ନ
Table 3:	ପ	ଫ	ବ	ଭ	ମ	ଯ	ର	ଲ	ଳ	
Table 4:	ଶ	ଷ	ସ	ହ	ଓ	ଊ	କ୍ଷ <sup>17</sup>			

16

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 ଓ + 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: ଓୂ

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.

These are with Anuswara and Chandrabindu. 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>16</sup> Not used in Odia. Retained for expert's advice

<sup>17</sup> କ୍ଷ = କ୍ଷ is ଓ has been retained although a ligature since in Odia, it is treated as an atomic character

### 6.3.2.1 Consonant and Mātrā combinations.

This set refers to a simple concatenation of Consonant and Mātrā<sup>18</sup>.

#### Consonant and Mātrā combinations Set 1

	କ	ଖ	ଗ	ଘ	ଙ	ଚ	ଛ	ଜ	ଝ	ଞ <sup>19</sup>
।	କ।	ଖ।	ଗ।	ଘ।	x	ଚ।	ଛ।	ଜ।	ଝ।	x
ୂ	କୂ	ଖୂ	ଗୂ	ଘୂ	x	ଚୂ	ଛୂ	ଜୂ	ଝୂ	x
ୃ	କୃ	ଖୃ	ଗୃ	ଘୃ	x	ଚୃ	ଛୃ	ଜୃ	ଝୃ	x
ୌ	କୌ	ଖୌ	ଗୌ	ଘୌ	x	ଚୌ	ଛୌ	ଜୌ	ଝୌ	x
େ	କେ	ଖେ	ଗେ	ଘେ	x	ଚେ	ଛେ	ଜେ	ଝେ	x
ୌ	କୌ	ଖୌ	ଗୌ	ଘୌ	x	ଚୌ	ଛୌ	ଜୌ	ଝୌ	x
େ।	କେ।	ଖେ।	ଗେ।	ଘେ।	x	ଚେ।	ଛେ।	ଜେ।	ଝେ।	x
ୌ।	କୌ।	ଖୌ।	ଗୌ।	ଘୌ।	x	ଚୌ।	ଛୌ।	ଜୌ।	ଝୌ।	x

Remark 1- ଙ and ଞ are rarely used only as the first members of clusters and mostly as alternatives of Anuswara

<sup>18</sup> In the case of matras, vocalic RR, vocalic L and vocalic LL have been left out since they occur infrequently in Odia.

<sup>19</sup> କଞ = କଞ is ଞ has been retained although a ligature since in Odia, it is treated as an atomic character

## Consonant and Mātrā combinations Set 2

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

	ଟ	ଠ	ଡ	ଢ	ଣ	ତ	ଥ	ଦ	ଧ	ନ
।	ଟ।	ଠ।	ଡ।	ଢ।	ଣ।	ତ।	ଥ।	ଦ।	ଧ।	ନ।
ୂ	ଟୂ	ଠୂ	ଡୂ	ଢୂ	ଣୂ	ତୂ	ଥୂ	ଦୂ	ଧୂ	ନୂ
ୱ	ଟୱ	ଠୱ	ଡୱ	ଢୱ	ଣୱ	ତୱ	ଥୱ	ଦୱ	ଧୱ	ନୱ
୲	ଟ୲	ଠ୲	ଡ୲	ଢ୲	ଣ୲	ତ୲	ଥ୲	ଦ୲	ଧ୲	ନ୲
୳	ଟ୳	ଠ୳	ଡ୳	ଢ୳	ଣ୳	ତ୳	ଥ୳	ଦ୳	ଧ୳	ନ୳
୴	ଟ୴	ଠ୴	ଡ୴	ଢ୴	ଣ୴	ତ୴	ଥ୴	ଦ୴	ଧ୴	ନ୴
୵	ଟ୵	ଠ୵	ଡ୵	ଢ୵	ଣ୵	ତ୵	ଥ୵	ଦ୵	ଧ୵	ନ୵
୶	ଟ୶	ଠ୶	ଡ୶	ଢ୶	ଣ୶	ତ୶	ଥ୶	ଦ୶	ଧ୶	ନ୶
୷	ଟ୷	ଠ୷	ଡ୷	ଢ୷	ଣ୷	ତ୷	ଥ୷	ଦ୷	ଧ୷	ନ୷
୸	ଟ୸	ଠ୸	ଡ୸	ଢ୸	ଣ୸	ତ୸	ଥ୸	ଦ୸	ଧ୸	ନ୸
୹	ଟ୹	ଠ୹	ଡ୹	ଢ୹	ଣ୹	ତ୹	ଥ୹	ଦ୹	ଧ୹	ନ୹
୺	ଟ୺	ଠ୺	ଡ୺	ଢ୺	ଣ୺	ତ୺	ଥ୺	ଦ୺	ଧ୺	ନ୺
୻	ଟ୻	ଠ୻	ଡ୻	ଢ୻	ଣ୻	ତ୻	ଥ୻	ଦ୻	ଧ୻	ନ୻
୼	ଟ୼	ଠ୼	ଡ୼	ଢ୼	ଣ୼	ତ୼	ଥ୼	ଦ୼	ଧ୼	ନ୼
୽	ଟ୽	ଠ୽	ଡ୽	ଢ୽	ଣ୽	ତ୽	ଥ୽	ଦ୽	ଧ୽	ନ୽
୾	ଟ୾	ଠ୾	ଡ୾	ଢ୾	ଣ୾	ତ୾	ଥ୾	ଦ୾	ଧ୾	ନ୾
୿	ଟ୿	ଠ୿	ଡ୿	ଢ୿	ଣ୿	ତ୿	ଥ୿	ଦ୿	ଧ୿	ନ୿

### Consonant and Mātrā combinations Set 3

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

	ପ	ଫ	ବ	ଭ	ମ	ଯ	ସ୍ୱ	ର	ଲ	ଳ	20	
।	ପ।	ଫ।	ବ।	ଭ।	ମ।	ଯ।	ସ୍ୱ।	ର।	ଲ।	ଳ।	।	।
ୂ	ପୂ	ଫୂ	ବୂ	ଭୂ	ମୂ	ଯୂ	ସ୍ୱୂ	ରୂ	ଲୂ	ଳୂ	ୂ	ୂ
ୃ	ପୃ	ଫୃ	ବୃ	ଭୃ	ମୃ	ଯୃ	ସ୍ୱୃ	ରୃ	ଲୃ	ଳୃ	ୃ	ୃ
ଌ	ପଌ	ଫଌ	ବଌ	ଭଌ	ମଌ	ଯଌ	ସ୍ୱଌ	ରଌ	ଲଌ	ଳଌ	ଌ	ଌ
ୠ	ପୠ	ଫୠ	ବୠ	ଭୠ	ମୠ	ଯୠ	ସ୍ୱୠ	ରୠ	ଲୠ	ଳୠ	ୠ	ୠ
ଐ	ପଐ	ଫଐ	ବଐ	ଭଐ	ମଐ	ଯଐ	ସ୍ୱଐ	ରଐ	ଲଐ	ଳଐ	ଐ	ଐ
ଓ	ପଓ	ଫଓ	ବଓ	ଭଓ	ମଓ	ଯଓ	ସ୍ୱଓ	ରଓ	ଲଓ	ଳଓ	ଓ	ଓ
ଌ	ପଌ	ଫଌ	ବଌ	ଭଌ	ମଌ	ଯଌ	ସ୍ୱଌ	ରଌ	ଲଌ	ଳଌ	ଌ	ଌ
ୠ	ପୠ	ଫୠ	ବୠ	ଭୠ	ମୠ	ଯୠ	ସ୍ୱୠ	ରୠ	ଲୠ	ଳୠ	ୠ	ୠ

<sup>20</sup> Not used in Odia. Replaced by Retained for expert's advice

### Consonant and Mātrā combinations Set 4

This set is in continuation of set 3 which shows consonant and Matra combinations.

	ଗ	ଞ	ସ	ହ	ଡ଼	ଢ଼	କ୍ଷ
।	ଗ।	ଞ।	ସ।	ହ।	ଡ଼।	ଢ଼।	କ୍ଷ।
ˆ	ଗି	ଞି	ସି	ହି	ଡ଼ି	ଢ଼ି	କ୍ଷି
୧	ଗ୧	ଞ୧	ସ୧	ହ୧	ଡ଼୧	ଢ଼୧	କ୍ଷ୧
ୃ	ଗୃ	ଞୃ	ସୃ	ହୃ	ଡ଼ୃ	ଢ଼ୃ	କ୍ଷୃ
ୂ	ଗୁ	ଞୁ	ସୁ	ହୁ	ଡ଼ୁ	ଢ଼ୁ	କ୍ଷୁ
ୃ̇	ଗୃ̇	ଞୃ̇	ସୃ̇	ହୃ̇	ଡ଼ୃ̇	ଢ଼ୃ̇	କ୍ଷୃ̇
ୃ̣	ଗୃ̣	ଞୃ̣	ସୃ̣	ହୃ̣	ଡ଼ୃ̣	ଢ଼ୃ̣	କ୍ଷୃ̣
୧̣	ଗ୧̣	ଞ୧̣	ସ୧̣	ହ୧̣	ଡ଼୧̣	ଢ଼୧̣	କ୍ଷ୧̣
୧̣̇	ଗ୧̣̇	ଞ୧̣̇	ସ୧̣̇	ହ୧̣̇	ଡ଼୧̣̇	ଢ଼୧̣̇	କ୍ଷ୧̣̇
୧̣̣̇	ଗ୧̣̣̇	ଞ୧̣̣̇	ସ୧̣̣̇	ହ୧̣̣̇	ଡ଼୧̣̣̇	ଢ଼୧̣̣̇	କ୍ଷ୧̣̣̇
୧̣̣̣̇	ଗ୧̣̣̣̇	ଞ୧̣̣̣̇	ସ୧̣̣̣̇	ହ୧̣̣̣̇	ଡ଼୧̣̣̣̇	ଢ଼୧̣̣̣̇	କ୍ଷ୧̣̣̣̇

### 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

	କ	ଖ	ଗ	ଘ	ଙ	ଚ	ଛ	ଜ	ଝ	ଞ
ଂ	କଂ	ଖଂ	ଗଂ	ଘଂ	x	ଚଂ	ଛଂ	ଜଂ	ଝଂ	x
।ଂ	କ।ଂ	ଖ।ଂ	ଗ।ଂ	ଘ।ଂ	x	ଚ।ଂ	ଛ।ଂ	ଜ।ଂ	ଝ।ଂ	x
ିଂ	କିଂ	ଖିଂ	ଗିଂ	ଘିଂ	x	ଚିଂ	ଛିଂ	ଜିଂ	ଝିଂ	x
ୀଂ	କୀଂ	ଖୀଂ	ଗୀଂ	ଘୀଂ	x	ଚୀଂ	ଛିୀଂ	ଜୀଂ	ଝୀଂ	x
ୂଂ	କୂଂ	ଖୂଂ	ଗୂଂ	ଘୂଂ	x	ଚୂଂ	ଛିୂଂ	ଜୂଂ	ଝୂଂ	x
ଂ	x	x	x	x	x	x	x	x	x	x
ୃଂ	କୃଂ	ଖୃଂ	ଗୃଂ	ଘୃଂ	x	ଚୃଂ	ଛିୃଂ	ଜୃଂ	ଝୃଂ	x
ୂଂ	କୂଂ	ଖୂଂ	ଗୂଂ	ଘୂଂ	x	ଚୂଂ	ଛିୂଂ	ଜୂଂ	ଝୂଂ	x
େଂ	କେଂ	ଖେଂ	ଗେଂ	ଘେଂ	x	ଚେଂ	ଛିେଂ	ଜେଂ	ଝେଂ	x
ୌ	କ	ଖ	ଗ	ଘ	x	ଚ	ଛ	ଜ	ଝ	x
େ।ଂ	କ	ଖ	ଗ	ଘ	x	ଚ	ଛ	ଜ	ଝ	x
େିଂ	କ	ଖ	ଗ	ଘ	x	ଚ	ଛ	ଜ	ଝ	x



### 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

	ଟ	ଠ	ଡ	ଢ	ଣ	ତ	ଥ	ଦ	ଧ	ନ
°	ଟ°	ଠ°	ଡ°	ଢ°	ଣ°	ତ°	ଥ°	ଦ°	ଧ°	ନ°
।°	ଟ।°	ଠ।°	ଡ।°	ଢ।°	ଣ।°	ତ।°	ଥ।°	ଦ।°	ଧ।°	ନ।°
ି°	ଟି°	ଠି°	ଡି°	ଢି°	ଣି°	ତି°	ଥି°	ଦି°	ଧି°	ନି°
ୀ°	ଟୀ°	ଠୀ°	ଡୀ°	ଢୀ°	ଣୀ°	ତୀ°	ଥୀ°	ଦୀ°	ଧୀ°	ନୀ°
ୁ°	ଟୁ°	ଠୁ°	ଡୁ°	ଢୁ°	ଣୁ°	ତୁ°	ଥୁ°	ଦୁ°	ଧୁ°	ନୁ°
°	x	x	x	x	x	x	x	x	x	x
ୂ°	ଟୁୂ°	ଠୁୂ°	ଡୁୂ°	ଢୁୂ°	ଣୁୂ°	ତୁୂ°	ଥୁୂ°	ଦୁୂ°	ଧୁୂ°	ନୁୂ°
ୃ°	ଟୃ°	ଠୃ°	ଡୃ°	ଢୃ°	ଣୃ°	ତୃ°	ଥୃ°	ଦୃ°	ଧୃ°	ନୃ°
େ°	ଟେ°	ଠେ°	ଡେ°	ଢେ°	ଣେ°	ତେ°	ଥେ°	ଦେ°	ଧେ°	ନେ°
ୌ°	ଟୌ°	ଠୌ°	ଡୌ°	ଢୌ°	ଣୌ°	ତୌ°	ଥୌ°	ଦୌ°	ଧୌ°	ନୌ°
େ.।°	ଟେ.। °	ଠେ.। °	ଡେ.। °	ଢେ.। °	ଣେ.। °	ତେ.। °	ଥେ.। °	ଦେ.। °	ଧେ.। °	ନେ.। °
ୌ.ି°	ଟୌ.ି °	ଠୌ.ି °	ଡୌ.ି °	ଢୌ.ି °	ଣୌ.ି °	ତୌ.ି °	ଥୌ.ି °	ଦୌ.ି °	ଧୌ.ି °	ନୌ.ି °

### 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

	ପ	ଫ	ବ	ଭ	ମ	ଯ	ସ	ର	ଲ	ଳ	21	
°	ପ°	ଫ°	ବ°	ଭ°	ମ°	ଯ°	ସ°	ର°	ଲ°	ଳ°	°	।°
।°	ପ।°	ଫ।°	ବ।°	ଭ।°	ମ।°	ଯ।°	ସ।°	ର।°	ଲ।°	ଳ।°	।°	°
୦°	ପ୦°	ଫ୦°	ବ୦°	ଭ୦°	ମ୦°	ଯ୦°	ସ୦°	ର୦°	ଲ୦°	ଳ୦°	୦°	୧°
୧°	ପ୧°	ଫ୧°	ବ୧°	ଭ୧°	ମ୧°	ଯ୧°	ସ୧°	ର୧°	ଲ୧°	ଳ୧°	୧°	୨°
୨°	ପ୨°	ଫ୨°	ବ୨°	ଭ୨°	ମ୨°	ଯ୨°	ସ୨°	ର୨°	ଲ୨°	ଳ୨°	୨°	x
୩°	x	x	x	x	x	x	x	x	x	x	x	୩°
୪°	ପ୪°	ଫ୪°	ବ୪°	ଭ୪°	ମ୪°	ଯ୪°	ସ୪°	ର୪°	ଲ୪°	ଳ୪°	୪°	୪°
୫°	ପ୫°	ଫ୫°	ବ୫°	ଭ୫°	ମ୫°	ଯ୫°	ସ୫°	ର୫°	ଲ୫°	ଳ୫°	୫°	୫°
୬°	ପ୬°	ଫ୬°	ବ୬°	ଭ୬°	ମ୬°	ଯ୬°	ସ୬°	ର୬°	ଲ୬°	ଳ୬°	୬°	୬°
୭°	ପ୭°	ଫ୭°	ବ୭°	ଭ୭°	ମ୭°	ଯ୭°	ସ୭°	ର୭°	ଲ୭°	ଳ୭°	୭°	୭°
୮°	ପ୮°	ଫ୮°	ବ୮°	ଭ୮°	ମ୮°	ଯ୮°	ସ୮°	ର୮°	ଲ୮°	ଳ୮°	୮°	୮°
୯।°	ପ୯।°	ଫ୯।°	ବ୯।°	ଭ୯।°	ମ୯।°	ଯ୯।°	ସ୯।°	ର୯।°	ଲ୯।°	ଳ୯।°	୯।°	୯।°
୯୧°	ପ୯୧°	ଫ୯୧°	ବ୯୧°	ଭ୯୧°	ମ୯୧°	ଯ୯୧°	ସ୯୧°	ର୯୧°	ଲ୯୧°	ଳ୯୧°	୯୧°	୧°

<sup>21</sup> Not used in Odia. Replaced by . Retained for expert's advice

### Consonant and Mātrā +Nasal combinations - Set 4

This set is in continuation of set 3 above which shows combinations of Consonant and Mātrā + Nasal marker

	ଗ	ଞ	ସ	ହ	ଡ଼	ଢ଼	କ୍ଷ
ଂ	ଗଂ	ଞଂ	ସଂ	ହଂ	ଡ଼ଂ	ଢ଼ଂ	କ୍ଷଂ
।ଂ	ଗ।ଂ	ଞ।ଂ	ସ।ଂ	ହ।ଂ	ଡ଼।ଂ	ଢ଼।ଂ	କ୍ଷ।ଂ
ିଂ	ଗିଂ	ଞିଂ	ସିଂ	ହିଂ	ଡ଼ିଂ	ଢ଼ିଂ	କ୍ଷିଂ
ୀଂ	ଗୀଂ	ଞୀଂ	ସୀଂ	ହୀଂ	ଡ଼ୀଂ	ଢ଼ୀଂ	କ୍ଷୀଂ
ୁଂ	ଗୁଂ	ଞୁଂ	ସୁଂ	ହୁଂ	ଡ଼ୁଂ	ଢ଼ୁଂ	କ୍ଷୁଂ
ୂଂ	ଗୁଂ	ଞୁଂ	ସୁଂ	ହୁଂ	ଡ଼ୁଂ	ଢ଼ୁଂ	କ୍ଷୁଂ
ୃଂ	ଗୁଂ	ଞୁଂ	ସୁଂ	ହୁଂ	ଡ଼ୁଂ	ଢ଼ୁଂ	କ୍ଷୁଂ
େଂ	ଗେଂ	ଞେଂ	ସେଂ	ହେଂ	ଡ଼େଂ	ଢ଼େଂ	କ୍ଷେଂ
ୌଂ	ଗୌଂ	ଞୌଂ	ସୌଂ	ହୌଂ	ଡ଼ୌଂ	ଢ଼ୌଂ	କ୍ଷୌଂ
େ।ଂ	ଗେ।ଂ	ଞେ।ଂ	ସେ।ଂ	ହେ।ଂ	ଡ଼େ।ଂ	ଢ଼େ।ଂ	କ୍ଷେ।ଂ
ୌ।ଂ	ଗୌ।ଂ	ଞୌ।ଂ	ସୌ।ଂ	ହୌ।ଂ	ଡ଼ୌ।ଂ	ଢ଼ୌ।ଂ	କ୍ଷୌ।ଂ

### Consonant and Mātrā + Nasal combinations: With Chandrabindu - Set 1

	କ	ଖ	ଗ	ଘ	ଙ	ଚ	ଛ	ଜ	ଝ	ଞ
ଂ	କଂ	ଖଂ	ଗଂ	ଘଂ	x	ଚଂ	ଛଂ	ଜଂ	ଝଂ	x
ଂ	କାଁ	ଖାଁ	ଗାଁ	ଘାଁ	x	ଚାଁ	ଛାଁ	ଜାଁ	ଝାଁ	x
ଂ	କିଁ	ଖିଁ	ଗିଁ	ଘିଁ	x	ଚିଁ	ଛିଁ	ଜିଁ	ଝିଁ	x
ଂ	କୌଁ	ଖୌଁ	ଗୌଁ	ଘୌଁ	x	ଚୌଁ	ଛୌଁ	ଜୌଁ	ଝୌଁ	x
ଂ	x	x	x	x	x	x	x	x	x	x
ଂ	କୃଁ	ଖୃଁ	ଗୃଁ	ଘୃଁ	x	ଚୃଁ	ଛୃଁ	ଜୃଁ	ଝୃଁ	x
ଂ	କ୍ଳୌଁ	ଖ୍ଳୌଁ	ଗ୍ଳୌଁ	ଘ୍ଳୌଁ	x	ଚ୍ଳୌଁ	ଛ୍ଳୌଁ	ଜ୍ଳୌଁ	ଝ୍ଳୌଁ	x
ୌ	କୌ	ଖୌ	ଗୌ	ଘୌ	x	ଚୌ	ଛୌ	ଜୌ	ଝୌ	x
ୌ	କୌଁ	ଖୌଁ	ଗୌଁ	ଘୌଁ	x	ଚୌଁ	ଛୌଁ	ଜୌଁ	ଝୌଁ	x
ୌ	କୌ	ଖୌ	ଗୌ	ଘୌ	x	ଚୌ	ଛୌ	ଜୌ	ଝୌ	x

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

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

	ଚ	ଠ	ଡ	ଢ	ଣ	ତ	ଥ	ଦ	ଧ	ନ
ଂ	ଚଂ	ଠଂ	ଡଂ	ଢଂ	ଣଂ	ତଂ	ଥଂ	ଦଂ	ଧଂ	ନଂ
ି	ଚି	ଠି	ଡି	ଢି	ଣି	ତି	ଥି	ଦି	ଧି	ନି
ୂ	ଚୁ	ଠୁ	ଡୁ	ଢୁ	ଣୁ	ତୁ	ଥୁ	ଦୁ	ଧୁ	ନୁ
ୱ	ଚୱ	ଠୱ	ଡୱ	ଢୱ	ଣୱ	ତୱ	ଥୱ	ଦୱ	ଧୱ	ନୱ
ୃ	ଚୃ	ଠୃ	ଡୃ	ଢୃ	ଣୃ	ତୃ	ଥୃ	ଦୃ	ଧୃ	ନୃ
ୌ	ଚୌ	ଠୌ	ଡୌ	ଢୌ	ଣୌ	ତୌ	ଥୌ	ଦୌ	ଧୌ	ନୌ
ଂ	x	x	x	x	x	x	x	x	x	x
ୂ	ଚୁଂ	ଠୁଂ	ଡୁଂ	ଢୁଂ	ଣୁଂ	ତୁଂ	ଥୁଂ	ଦୁଂ	ଧୁଂ	ନୁଂ
ୱ	ଚୱଂ	ଠୱଂ	ଡୱଂ	ଢୱଂ	ଣୱଂ	ତୱଂ	ଥୱଂ	ଦୱଂ	ଧୱଂ	ନୱଂ
ୃ	ଚୃଂ	ଠୃଂ	ଡୃଂ	ଢୃଂ	ଣୃଂ	ତୃଂ	ଥୃଂ	ଦୃଂ	ଧୃଂ	ନୃଂ
ୌ	ଚୌଂ	ଠୌଂ	ଡୌଂ	ଢୌଂ	ଣୌଂ	ତୌଂ	ଥୌଂ	ଦୌଂ	ଧୌଂ	ନୌଂ
ୌ	ଚୌଂ	ଠୌଂ	ଡୌଂ	ଢୌଂ	ଣୌଂ	ତୌଂ	ଥୌଂ	ଦୌଂ	ଧୌଂ	ନୌଂ
ୌ	ଚୌଂ	ଠୌଂ	ଡୌଂ	ଢୌଂ	ଣୌଂ	ତୌଂ	ଥୌଂ	ଦୌଂ	ଧୌଂ	ନୌଂ
ୌ	ଚୌଂ	ଠୌଂ	ଡୌଂ	ଢୌଂ	ଣୌଂ	ତୌଂ	ଥୌଂ	ଦୌଂ	ଧୌଂ	ନୌଂ
ୌ	ଚୌଂ	ଠୌଂ	ଡୌଂ	ଢୌଂ	ଣୌଂ	ତୌଂ	ଥୌଂ	ଦୌଂ	ଧୌଂ	ନୌଂ
ୌ	ଚୌଂ	ଠୌଂ	ଡୌଂ	ଢୌଂ	ଣୌଂ	ତୌଂ	ଥୌଂ	ଦୌଂ	ଧୌଂ	ନୌଂ
ୌ	ଚୌଂ	ଠୌଂ	ଡୌଂ	ଢୌଂ	ଣୌଂ	ତୌଂ	ଥୌଂ	ଦୌଂ	ଧୌଂ	ନୌଂ
ୌ	ଚୌଂ	ଠୌଂ	ଡୌଂ	ଢୌଂ	ଣୌଂ	ତୌଂ	ଥୌଂ	ଦୌଂ	ଧୌଂ	ନୌଂ

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

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

	ପ	ଫ	ବ	ଭ	ମ	ଯ	ସ	ର	ଲ	ଳ		
ଂ	ପଁ	ଫଁ	ବଁ	ଭଁ	ମଁ	ଯଁ	ସଁ	ରଁ	ଲଁ	ଳଁ	ଂ	ଂ
ା	ପାଁ	ଫାଁ	ବାଁ	ଭାଁ	ମାଁ	ଯାଁ	ସାଁ	ରାଁ	ଲାଁ	ଳାଁ	ା	ା
ଃ	ପୃଁ	ଫୃଁ	ବୃଁ	ଭୃଁ	ମୃଁ	ଯୃଁ	ସୃଁ	ରୃଁ	ଲୃଁ	ଳୃଁ	ଃ	ଃ
ୀ	ପାଁ	ଫାଁ	ବାଁ	ଭାଁ	ମାଁ	ଯାଁ	ସାଁ	ରାଁ	ଲାଁ	ଳାଁ	ୀ	ୀ
ୈ	ପୈଁ	ଫୈଁ	ବୈଁ	ଭୈଁ	ମୈଁ	ଯୈଁ	ସୈଁ	ରୈଁ	ଲୈଁ	ଳୈଁ	ୈ	ୈ
ଂ	x	x	x	x	x	x	x	x	x	x	x	x
ୂ	ପୂଁ	ଫୂଁ	ବୂଁ	ଭୂଁ	ମୂଁ	ଯୂଁ	ସୂଁ	ରୂଁ	ଲୂଁ	ଳୂଁ	ୂ	ୂ
ୃ	ପୃଁ	ଫୃଁ	ବୃଁ	ଭୃଁ	ମୃଁ	ଯୃଁ	ସୃଁ	ରୃଁ	ଲୃଁ	ଳୃଁ	ୃ	ୃ
ୌ	ପୌଁ	ଫୌଁ	ବୌଁ	ଭୌଁ	ମୌଁ	ଯୌଁ	ସୌଁ	ରୌଁ	ଲୌଁ	ଳୌଁ	ୌ	ୌ
ୌ	ପୌଁ	ଫୌଁ	ବୌଁ	ଭୌଁ	ମୌଁ	ଯୌଁ	ସୌଁ	ରୌଁ	ଲୌଁ	ଳୌଁ	ୌ	ୌ
ୌ	ପୌଁ	ଫୌଁ	ବୌଁ	ଭୌଁ	ମୌଁ	ଯୌଁ	ସୌଁ	ରୌଁ	ଲୌଁ	ଳୌଁ	ୌ	ୌ
ୌ	ପୌଁ	ଫୌଁ	ବୌଁ	ଭୌଁ	ମୌଁ	ଯୌଁ	ସୌଁ	ରୌଁ	ଲୌଁ	ଳୌଁ	ୌ	ୌ

### Consonant and Mātrā +Nasal combinations With Chandrabindu - Set 4

This set is in continuation of set 3 above which shows combinations of Consonant and Mātrā + Chandrabindu

	ଗ	ଝ	ଞ	ଟ	ଡ	ଢ	କ୍ଷ
ଂ	ଗଁ	ଝଁ	ଞଁ	ଟଁ	ଡଁ	ଢଁ	କ୍ଷଁ
ଐ	ଗାଁ	ଝାଁ	ଞାଁ	ଟାଁ	ଡାଁ	ଢାଁ	କ୍ଷାଁ
ଋ	ଗୌ	ଝୌ	ଞୌ	ଟୌ	ଡୌ	ଢୌ	କ୍ଷୌ
ୠ	ଗୌ	ଝୌ	ଞୌ	ଟୌ	ଡୌ	ଢୌ	କ୍ଷୌ
ଌ	ଗ୍ଳିଁ	ଝ୍ଳିଁ	ଞ୍ଳିଁ	ଟ୍ଳିଁ	ଡ୍ଳିଁ	ଢ୍ଳିଁ	କ୍ଷ୍ଳିଁ
ୡ	x	x	x	x	x	x	x
ୣ	ଗ୍ଲିଁ	ଝ୍ଲିଁ	ଞ୍ଲିଁ	ଟ୍ଲିଁ	ଡ୍ଲିଁ	ଢ୍ଲିଁ	କ୍ଷ୍ଲିଁ
୤	ଗ୍ଲିଁ	ଝ୍ଲିଁ	ଞ୍ଲିଁ	ଟ୍ଲିଁ	ଡ୍ଲିଁ	ଢ୍ଲିଁ	କ୍ଷ୍ଲିଁ
୥	ଗର୍	ଝର୍	ଞର୍	ଟର୍	ଡର୍	ଢର୍	କ୍ଷର୍
୦	ଗର୍	ଝର୍	ଞର୍	ଟର୍	ଡର୍	ଢର୍	କ୍ଷର୍
୧	ଗର୍	ଝର୍	ଞର୍	ଟର୍	ଡର୍	ଢର୍	କ୍ଷର୍
୨	ଗର୍	ଝର୍	ଞର୍	ଟର୍	ଡର୍	ଢର୍	କ୍ଷର୍

### 6.3.3. The Ligature Set of Odia.

Odia 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 Consonants)

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 ligature is not possible in the language but is theoretically possible. 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 e.g. Combination of consonant “କ” and “ ଗ ” is ligature “କ୍ଗ”.

The Odia letter ୍ is not retained in the matrix of characters with Halanta

CHC( combination of two consonants) - Set 1										
	କ	ଖ	ଗ	ଘ	ଙ	ଚ	ଛ	ଜ	ଝ	ଞ
କ୍	କ୍କ	କ୍ଖ	କ୍ଗ	କ୍ଘ	x	କ୍ଚ	କ୍ଛ	କ୍ଜ	କ୍ଝ	x
ଖ୍	ଖ୍କ	ଖ୍ଖ	ଖ୍ଗ	ଖ୍ଘ	x	ଖ୍ଚ	ଖ୍ଛ	ଖ୍ଜ	ଖ୍ଝ	x
ଗ୍	ଗ୍କ	ଗ୍ଖ	ଗ୍ଗ	ଗ୍ଘ	x	ଗ୍ଚ	ଗ୍ଛ	ଗ୍ଜ	ଗ୍ଝ	x
ଘ୍	ଘ୍କ	ଘ୍ଖ	ଘ୍ଗ	ଘ୍ଘ	x	ଘ୍ଚ	ଘ୍ଛ	ଘ୍ଜ	ଘ୍ଝ	x
ଙ୍	ଙ୍କ	ଙ୍ଖ	ଙ୍ଗ	ଙ୍ଘ	x	ଙ୍ଚ	ଙ୍ଛ	ଙ୍ଜ	ଙ୍ଝ	x
ଚ୍	ଚ୍କ	ଚ୍ଖ	ଚ୍ଗ	ଚ୍ଘ	x	ଚ୍ଚ	ଚ୍ଛ	ଚ୍ଜ	ଚ୍ଝ	x
ଛ୍	ଛ୍କ	ଛ୍ଖ	ଛ୍ଗ	ଛ୍ଘ	x	ଛ୍ଚ	ଛ୍ଛ	ଛ୍ଜ	ଛ୍ଝ	x
ଜ୍	ଜ୍କ	ଜ୍ଖ	ଜ୍ଗ	ଜ୍ଘ	x	ଜ୍ଚ	ଜ୍ଛ	ଜ୍ଜ	ଜ୍ଝ	ଜ୍ଞ
ଝ୍	ଝ୍କ	ଝ୍ଖ	ଝ୍ଗ	ଝ୍ଘ	x	ଝ୍ଚ	ଝ୍ଛ	ଝ୍ଜ	ଝ୍ଝ	x
ଞ୍	ଞ୍କ	ଞ୍ଖ	ଞ୍ଗ	ଞ୍ଘ	x	ଞ୍ଚ	ଞ୍ଛ	ଞ୍ଜ	ଞ୍ଝ	x
ଟ୍	ଟ୍କ	ଟ୍ଖ	ଟ୍ଗ	ଟ୍ଘ	x	ଟ୍ଚ	ଟ୍ଛ	ଟ୍ଜ	ଟ୍ଝ	x
ଠ୍	ଠ୍କ	ଠ୍ଖ	ଠ୍ଗ	ଠ୍ଘ	x	ଠ୍ଚ	ଠ୍ଛ	ଠ୍ଜ	ଠ୍ଝ	x
ଡ୍	ଡ୍କ	ଡ୍ଖ	ଡ୍ଗ	ଡ୍ଘ	x	ଡ୍ଚ	ଡ୍ଛ	ଡ୍ଜ	ଡ୍ଝ	x



ଢ	ଢକ	ଢଖ	ଢଗ	ଢଘ	x	ଢଚ	ଢଛ	ଢଜ	ଢଝ	x
ଣ	ଣକ	ଣଖ	ଣଗ	ଣଘ	x	ଣଚ	ଣଛ	ଣଜ	ଣଝ	x
ତ	ତକ	ତଖ	ତଗ	ତଘ	x	ତଚ	ତଛ	ତଜ	ତଝ	x
ଥ	ଥକ	ଥଖ	ଥଗ	ଥଘ	x	ଥଚ	ଥଛ	ଥଜ	ଥଝ	x
ଦ	ଦକ	ଦଖ	ଦଗ	ଦଘ	x	ଦଚ	ଦଛ	ଦଜ	ଦଝ	x
ଧ	ଧକ	ଧଖ	ଧଗ	ଧଘ	x	ଧଚ	ଧଛ	ଧଜ	ଧଝ	x
ନ	ନକ	ନଖ	ନଗ	ନଘ	x	ନଚ	ନଛ	ନଜ	ନଝ	x
ପ	ପକ	ପଖ	ପଗ	ପଘ	x	ପଚ	ପଛ	ପଜ	ପଝ	x
ଫ	ଫକ	ଫଖ	ଫଗ	ଫଘ	x	ଫଚ	ଫଛ	ଫଜ	ଫଝ	x
ବ	ବକ	ବଖ	ବଗ	ବଘ	x	ବଚ	ବଛ	ବଜ	ବଝ	x
ଭ	ଭକ	ଭଖ	ଭଗ	ଭଘ	x	ଭଚ	ଭଛ	ଭଜ	ଭଝ	x
ମ	ମକ	ମଖ	ମଗ	ମଘ	x	ମଚ	ମଛ	ମଜ	ମଝ	x
ଯ	ଯକ	ଯଖ	ଯଗ	ଯଘ	x	ଯଚ	ଯଛ	ଯଜ	ଯଝ	x
ସ	ସକ	ସଖ	ସଗ	ସଘ	x	ସଚ	ସଛ	ସଜ	ସଝ	x
ର	ରକ	ରଖ	ରଗ	ରଘ	x	ରଚ	ରଛ	ରଜ	ରଝ	x
ଲ	ଲକ	ଲଖ	ଲଗ	ଲଘ	x	ଲଚ	ଲଛ	ଲଜ	ଲଝ	x
ଳ	ଳକ	ଳଖ	ଳଗ	ଳଘ	x	ଳଚ	ଳଛ	ଳଜ	ଳଝ	x
ୠ	ୠକ	ୠଖ	ୠଗ	ୠଘ	x	ୠଚ	ୠଛ	ୠଜ	ୠଝ	x
ଶ	ଶକ	ଶଖ	ଶଗ	ଶଘ	x	ଶଚ	ଶଛ	ଶଜ	ଶଝ	x
ଷ	ଷକ	ଷଖ	ଷଗ	ଷଘ	x	ଷଚ	ଷଛ	ଷଜ	ଷଝ	x
ସା	ସାକ	ସାଖ	ସାଗ	ସାଘ	x	ସାଚ	ସାଛ	ସାଜ	ସାଝ	x
ହ	ହକ	ହଖ	ହଗ	ହଘ	x	ହଚ	ହଛ	ହଜ	ହଝ	x
ଝ	x	x	x	x	x	x	x	x	x	x
ଞ	x	x	x	x	x	x	x	x	x	x

## CHC Set 2:

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 e.g. Combination of consonant “କ” and “ଟ” is ligature “କଟ”.

The Odia letter ୟ is not retained in the matrix of characters with Halanta

CHC( combination of two consonants) - Set 2										
	ଚ	ଠ	ଡ	ଢ	ଣ	ତ	ଥ	ଦ	ଧ	ନ
କ	କଚ	କଠ	କଡ	କଢ	କଣ	କତ	କଥ	କଦ	କଧ	କନ
ଖ	ଖଚ	ଖଠ	ଖଡ	ଖଢ	ଖଣ	ଖତ	ଖଥ	ଖଦ	ଖଧ	ଖନ
ଗ	ଗଚ	ଗଠ	ଗଡ	ଗଢ	ଗଣ	ଗତ	ଗଥ	ଗଦ	ଗଧ	ଗନ
ଘ	ଘଚ	ଘଠ	ଘଡ	ଘଢ	ଘଣ	ଘତ	ଘଥ	ଘଦ	ଘଧ	ଘନ
ଙ	ଙଚ	ଙଠ	ଙଡ	ଙଢ	ଙଣ	ଙତ	ଙଥ	ଙଦ	ଙଧ	ଙନ
ଚ	ଚଚ	ଚଠ	ଚଡ	ଚଢ	ଚଣ	ଚତ	ଚଥ	ଚଦ	ଚଧ	ଚନ
ଛ	ଛଚ	ଛଠ	ଛଡ	ଛଢ	ଛଣ	ଛତ	ଛଥ	ଛଦ	ଛଧ	ଛନ
ଜ	ଜଚ	ଜଠ	ଜଡ	ଜଢ	ଜଣ	ଜତ	ଜଥ	ଜଦ	ଜଧ	ଜନ
ଝ	ଝଚ	ଝଠ	ଝଡ	ଝଢ	ଝଣ	ଝତ	ଝଥ	ଝଦ	ଝଧ	ଝନ
ଞ	ଞଚ	ଞଠ	ଞଡ	ଞଢ	ଞଣ	ଞତ	ଞଥ	ଞଦ	ଞଧ	ଞନ
ଟ	ଟଚ	ଟଠ	ଟଡ	ଟଢ	ଟଣ	ଟତ	ଟଥ/ X	ଟଦ/ X	ଟଧ/ X	ଟନ/ X
ଢ	ଢଚ	ଢଠ	ଢଡ	ଢଢ	ଢଣ	ଢତ	ଢଥ	ଢଦ	ଢଧ	ଢନ
ଣ	ଣଚ	ଣଠ	ଣଡ	ଣଢ	ଣଣ	ଣତ	ଣଥ	ଣଦ	ଣଧ	ଣନ
ତ	ତଚ	ତଠ	ତଡ	ତଢ	ତଣ	ତତ	ତଥ	ତଦ	ତଧ	ତନ
ଥ	ଥଚ	ଥଠ	ଥଡ	ଥଢ	ଥଣ	ଥତ	ଥଥ	ଥଦ	ଥଧ	ଥନ
ଦ	ଦଚ	ଦଠ	ଦଡ	ଦଢ	ଦଣ	ଦତ	ଦଥ	ଦଦ	ଦଧ	ଦନ
ଧ	ଧଚ	ଧଠ	ଧଡ	ଧଢ	ଧଣ	ଧତ	ଧଥ	ଧଦ	ଧଧ	ଧନ
ନ	ନଚ	ନଠ	ନଡ	ନଢ	ନଣ	ନତ	ନଥ	ନଦ	ନଧ	ନନ
ଂ	ଂଚ	ଂଠ	ଂଡ	ଂଢ	ଂଣ	ଂତ	ଂଥ	ଂଦ	ଂଧ	ଂନ

ଫ	ଫା	ଫଠ	ଫଡ଼	ଫୁ	ଫଣ	ଫଟ	ଫଥ	ଫଦ	ଫଧ	ଫନ
ବ	ବା	ବଠ	ବଡ଼	ବୁ	ବଣ	ବଟ	ବଥ	ବଦ	ବଧ	ବନ
ଭ	ଭା	ଭଠ	ଭଡ଼	ଭୁ	ଭଣ	ଭଟ	ଭଥ	ଭଦ	ଭଧ	ଭନ
ମ	ମା	ମଠ	ମଡ଼	ମୁ	ମଣ	ମଟ	ମଥ	ମଦ	ମଧ	ମନ
ଯ	ଯା	ଯଠ	ଯଡ଼	ୟ	ଯଣ	ଯଟ	ଯଥ	ଯଦ	ଯଧ	ଯନ
ସ୍	ସ୍ତା	ସ୍ତଠ	ସ୍ତଡ଼	ସ୍ତୁ	ସ୍ତଣ	ସ୍ତଟ	ସ୍ତଥ	ସ୍ତଦ	ସ୍ତଧ	ସ୍ତନ
ର	ରା	ରଠ	ରଡ଼	ରୁ	ରଣ	ରଟ	ରଥ	ରଦ	ରଧ	ରନ
ଲ	ଲା	ଲଠ	ଲଡ଼	ଲୁ	ଲଣ	ଲଟ	ଲଥ	ଲଦ	ଲଧ	ଲନ
ଳ	ଳା	ଳଠ	ଳଡ଼	ଳୁ	ଳଣ	ଳଟ	ଳଥ	ଳଦ	ଳଧ	ଳନ
	ଟା	ଠ	ଡ଼	ଢୁ	ଣ	ଟ	ଥ	ଦ	ଧ	ନ
ଶ	ଶା	ଶଠ	ଶଡ଼	ଶୁ	ଶଣ	ଶଟ	ଶଥ	ଶଦ	ଶଧ	ଶନ
ଷ	ଷା	ଷଠ	ଷଡ଼	ଷୁ	ଷଣ	ଷଟ	ଷଥ	ଷଦ	ଷଧ	ଷନ
ସ	ସା	ସଠ	ସଡ଼	ସୁ	ସଣ	ସଟ	ସଥ	ସଦ	ସଧ	ସନ
ଝ	ଝା	ଝଠ	ଝଡ଼	ଝୁ	ଝଣ	ଝଟ	ଝଥ	ଝଦ	ଝଧ	ଝନ
ଡ	x	x	x	x	x	x	x	x	x	x
ଢ	x	x	x	x	x	x	x	x	x	x

### CHC SET 3:

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

The Odia letter is not retained in the matrix of characters with Halanta

CHC( combination of two consonants) - Set 3														
	ପ	ଫ	ବ	ଭ	ମ	ଯ	ର	ଲ			ଶ	ଷ	ସ	ହ
କ	କ୍ପ	କ୍ଫ	କ୍ବ	କ୍ଭ	କ୍ମ	କ୍ଯ	କ୍ର	କ୍ଲ	କ୍	କ୍	କ୍ଶ	କ୍ଷ	କ୍ସ	କ୍ହ
ଖ	ଖ୍ପ	ଖ୍ଫ	ଖ୍ବ	ଖ୍ଭ	ଖ୍ମ	ଖ୍ଯ	ଖ୍ର	ଖ୍ଲ	ଖ୍	ଖ୍	ଖ୍ଶ	ଖ୍ଷ	ଖ୍ସ	ଖ୍ହ
ଗ	ଗ୍ପ	ଗ୍ଫ	ଗ୍ବ	ଗ୍ଭ	ଗ୍ମ	ଗ୍ଯ	ଗ୍ର	ଗ୍ଲ	ଗ୍	ଗ୍	ଗ୍ଶ	ଗ୍ଷ	ଗ୍ସ	ଗ୍ହ
ଘ	ଘ୍ପ	ଘ୍ଫ	ଘ୍ବ	ଘ୍ଭ	ଘ୍ମ	ଘ୍ଯ	ଘ୍ର	ଘ୍ଲ	ଘ୍	ଘ୍	ଘ୍ଶ	ଘ୍ଷ	ଘ୍ସ	ଘ୍ହ
ଙ	ଙ୍ପ	ଙ୍ଫ	ଙ୍ବ	ଙ୍ଭ	ଙ୍ମ	ଙ୍ଯ	ଙ୍ର	ଙ୍ଲ	ଙ୍	ଙ୍	ଙ୍ଶ	ଙ୍ଷ	ଙ୍ସ	ଙ୍ହ
ଚ	ଚ୍ପ	ଚ୍ଫ	ଚ୍ବ	ଚ୍ଭ	ଚ୍ମ	ଚ୍ଯ	ଚ୍ର	ଚ୍ଲ	ଚ୍	ଚ୍	ଚ୍ଶ	ଚ୍ଷ	ଚ୍ସ	ଚ୍ହ
ଛ	ଛ୍ପ	ଛ୍ଫ	ଛ୍ବ	ଛ୍ଭ	ଛ୍ମ	ଛ୍ଯ	ଛ୍ର	ଛ୍ଲ	ଛ୍	ଛ୍	ଛ୍ଶ	ଛ୍ଷ	ଛ୍ସ	ଛ୍ହ
ଜ	ଜ୍ପ	ଜ୍ଫ	ଜ୍ବ	ଜ୍ଭ	ଜ୍ମ	ଜ୍ଯ	ଜ୍ର	ଜ୍ଲ	ଜ୍	ଜ୍	ଜ୍ଶ	ଜ୍ଷ	ଜ୍ସ	ଜ୍ହ
ଝ	ଝ୍ପ	ଝ୍ଫ	ଝ୍ବ	ଝ୍ଭ	ଝ୍ମ	ଝ୍ଯ	ଝ୍ର	ଝ୍ଲ	ଝ୍	ଝ୍	ଝ୍ଶ	ଝ୍ଷ	ଝ୍ସ	ଝ୍ହ
ଞ	ଞ୍ପ	ଞ୍ଫ	ଞ୍ବ	ଞ୍ଭ	ଞ୍ମ	ଞ୍ଯ	ଞ୍ର	ଞ୍ଲ	ଞ୍	ଞ୍	ଞ୍ଶ	ଞ୍ଷ	ଞ୍ସ	ଞ୍ହ
ଟ	ଟ୍ପ	ଟ୍ଫ	ଟ୍ବ	ଟ୍ଭ	ଟ୍ମ	ଟ୍ଯ	ଟ୍ର	ଟ୍ଲ	ଟ୍	ଟ୍	ଟ୍ଶ	ଟ୍ଷ	ଟ୍ସ	ଟ୍ହ
ଠ	ଠ୍ପ	ଠ୍ଫ	ଠ୍ବ	ଠ୍ଭ	ଠ୍ମ	ଠ୍ଯ	ଠ୍ର	ଠ୍ଲ	ଠ୍	ଠ୍	ଠ୍ଶ	ଠ୍ଷ	ଠ୍ସ	ଠ୍ହ
ଡ	ଡ୍ପ	ଡ୍ଫ	ଡ୍ବ	ଡ୍ଭ	ଡ୍ମ	ଡ୍ଯ	ଡ୍ର	ଡ୍ଲ	ଡ୍	ଡ୍	ଡ୍ଶ	ଡ୍ଷ	ଡ୍ସ	ଡ୍ହ
ଣ	ଣ୍ପ	ଣ୍ଫ	ଣ୍ବ	ଣ୍ଭ	ଣ୍ମ	ଣ୍ଯ	ଣ୍ର	ଣ୍ଲ	ଣ୍	ଣ୍	ଣ୍ଶ	ଣ୍ଷ	ଣ୍ସ	ଣ୍ହ
ତ	ତ୍ପ	ତ୍ଫ	ତ୍ବ	ତ୍ଭ	ତ୍ମ	ତ୍ଯ	ତ୍ର	ତ୍ଲ	ତ୍	ତ୍	ତ୍ଶ	ତ୍ଷ	ତ୍ସ	ତ୍ହ
ଥ	ଥ୍ପ	ଥ୍ଫ	ଥ୍ବ	ଥ୍ଭ	ଥ୍ମ	ଥ୍ଯ	ଥ୍ର	ଥ୍ଲ	ଥ୍	ଥ୍	ଥ୍ଶ	ଥ୍ଷ	ଥ୍ସ	ଥ୍ହ
ଦ	ଦ୍ପ	ଦ୍ଫ	ଦ୍ବ	ଦ୍ଭ	ଦ୍ମ	ଦ୍ଯ	ଦ୍ର	ଦ୍ଲ	ଦ୍	ଦ୍	ଦ୍ଶ	ଦ୍ଷ	ଦ୍ସ	ଦ୍ହ
ଧ	ଧ୍ପ	ଧ୍ଫ	ଧ୍ବ	ଧ୍ଭ	ଧ୍ମ	ଧ୍ଯ	ଧ୍ର	ଧ୍ଲ	ଧ୍	ଧ୍	ଧ୍ଶ	ଧ୍ଷ	ଧ୍ସ	ଧ୍ହ
ନ	ନ୍ପ	ନ୍ଫ	ନ୍ବ	ନ୍ଭ	ନ୍ମ	ନ୍ଯ	ନ୍ର	ନ୍ଲ	ନ୍	ନ୍	ନ୍ଶ	ନ୍ଷ	ନ୍ସ	ନ୍ହ
ପ୍	ପ୍ପ	ପ୍ଫ	ପ୍ବ	ପ୍ଭ	ପ୍ମ	ପ୍ଯ	ପ୍ର	ପ୍ଲ	ପ୍	ପ୍	ପ୍ଶ	ପ୍ଷ	ପ୍ସ	ପ୍ହ
ଫ୍	ଫ୍ପ	ଫ୍ଫ	ଫ୍ବ	ଫ୍ଭ	ଫ୍ମ	ଫ୍ଯ	ଫ୍ର	ଫ୍ଲ	ଫ୍	ଫ୍	ଫ୍ଶ	ଫ୍ଷ	ଫ୍ସ	ଫ୍ହ
ବ୍	ବ୍ପ	ବ୍ଫ	ବ୍ବ	ବ୍ଭ	ବ୍ମ	ବ୍ଯ	ବ୍ର	ବ୍ଲ	ବ୍	ବ୍	ବ୍ଶ	ବ୍ଷ	ବ୍ସ	ବ୍ହ
ଭ୍	ଭ୍ପ	ଭ୍ଫ	ଭ୍ବ	ଭ୍ଭ	ଭ୍ମ	ଭ୍ଯ	ଭ୍ର	ଭ୍ଲ	ଭ୍	ଭ୍	ଭ୍ଶ	ଭ୍ଷ	ଭ୍ସ	ଭ୍ହ
ମ୍	ମ୍ପ	ମ୍ଫ	ମ୍ବ	ମ୍ଭ	ମ୍ମ	ମ୍ଯ	ମ୍ର	ମ୍ଲ	ମ୍	ମ୍	ମ୍ଶ	ମ୍ଷ	ମ୍ସ	ମ୍ହ

ଯ	ଯପ	ଯଫ	ଯବ	ଯଭ	ଯମ	ଯଯ	ଯର	ଯଲ	ଯ	ଯ	ଯଶ	ଯଷ	ଯସ	ଯହ
ସ୍ଯ	ସ୍ଯପ	ସ୍ଯଫ	ସ୍ଯବ	ସ୍ଯଭ	ସ୍ଯମ	ସ୍ଯଯ	ସ୍ଯର	ସ୍ଯଲ	ସ୍ଯ	ସ୍ଯ	ସ୍ଯଶ	ସ୍ଯଷ	ସ୍ଯସ	ସ୍ଯହ
ର	ରପ	ରଫ	ରବ	ରଭ	ରମ	ରଯ	ରର	ରଲ	ର	ର	ରଶ	ରଷ	ରସ	ରହ
ଲ	ଲପ	ଲଫ	ଲବ	ଲଭ	ଲମ	ଲଯ	ଲର	ଲଲ	ଲ	ଲ	ଲଶ	ଲଷ	ଲସ	ଲହ
ଳ	ଳପ	ଳଫ	ଳବ	ଳଭ	ଳମ	ଳଯ	ଳର	ଳଲ	ଳ	ଳ	ଳଶ	ଳଷ	ଳସ	ଳହ
	ପ	ଫ	ବ	ଭ	ମ	ଯ	ର	ଲ			ଶ	ଷ	ସ	ହ
ଶ	ଶପ	ଶଫ	ଶବ	ଶଭ	ଶମ	ଶଯ	ଶର	ଶଲ	ଶ	ଶ	ଶଶ	ଶଷ	ଶସ	ଶହ
ଷ	ଷପ	ଷଫ	ଷବ	ଷଭ	ଷମ	ଷଯ	ଷର	ଷଲ	ଷ	ଷ	ଷଶ	ଷଷ	ଷସ	ଷହ
ସ	ସପ	ସଫ	ସବ	ସଭ	ସମ	ସଯ	ସର	ସଲ	ସ	ସ	ସଶ	ସଷ	ସସ	ସହ
ହ	ହପ	ହଫ	ହବ	ହଭ	ହମ	ହଯ	ହର	ହଲ	ହ	ହ	ହଶ	ହଷ	ହସ	ହହ
ଓ	x	x	x	x	x	x	x	x	x	x	x	x	x	x
ଔ	x	x	x	x	x	x	x	x	x	x	x	x	x	x

#### CHC SET 4:

The following set shows a combination of two consonants of which the second form is the nukta . To know how particular combinations forms, select one consonant from the first column and second from first row. For e.g. Combination of consonant “କ୍” and “ଡ଼” is a theoretical ligature “କ୍ଡ଼”. The experts have deemed that such combinations do not occur in the language.

CHC Set 4		
	ଡ଼	ଡ଼
କ୍	x	x
ଖ୍	x	x
ଗ୍	x	x
ଘ୍	x	x
ଙ୍	x	x
ଚ୍	x	x
ଛ୍	x	x
ଜ୍	x	x
ଝ୍	x	x
ଞ୍	x	x
ଟ୍	x	x
ଠ୍	x	x
ଡ଼	x	x
ଣ୍	x	x
ତ୍	x	x
ଥ୍	x	x
ଦ୍	x	x
ଧ୍	x	x
ନ୍	x	x
ପ୍	x	x
ଫ୍	x	x
ବ୍	x	x
ଭ୍	x	x
ସ୍	x	x
ଷ୍	x	x
ହ୍	x	x
ମ୍	x	x
ୟ୍	x	x
ର୍	x	x
ଲ୍	x	x
ୱ୍	x	x

୨	X	X
୩	X	X
୪	X	X
୫	X	X
୬	X	X
୭	X	X
୮	X	X

### 6.3.3.2 CHCHC ( combination of three Consonants)

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

With a few exceptions these are mainly linear in nature.

କ+ତ+ର = କ୍ତର

କ+ଷ+ଶ = କ୍ଷଶ

କ+ଷ+ମ = କ୍ଷମ

କ+ଷ+ସ = କ୍ଷସ

ଢ+କ+ଷ = ଢକ୍ଷ

ଜ+ଜ+ = ଜ୍ଜ

ଶ+ତ+ର = ଶ୍ତର

ତ+ତ+ = ତ୍ତ

ତ+ତ+ = ତ୍ତ

ତ+ତ+ର = ତ୍ତର

ତ+ମ+ସ = ତ୍ମସ

ତ+ସ+ନ = ତ୍ସନ

ତ+ସ+ସ = ତ୍ସସ

ନ+ତ+ସ = ନ୍ତସ

ନ+ତ+ = ନ୍ତ

ନ+ତ+ = ନ୍ତ

ନ+ଦ+ର = ନ୍ଦର

ନ+ଦ+ = ନ୍ଦ

ନ+ଦ+ = ନ୍ଦ

ନ+ଧ+ର = ନ୍ଧର

ନ+ଧ+ସ = ନ୍ଧସ

ନ+ନ+ସ = ନ୍ନସ

ମ+ପ+ର = ମ୍ପର

ମ+ଭ+ର = ମ୍ଭର

ର+ଶ+ଶ = ର୍ଶଶ

ର+ତ+ସ = ର୍ତସ

ର+ଦ+ର = ର୍ଦର

ର+ଧ+ସ = ର୍ଧସ

ର+ଶ+ = ର୍ଶ

ର+ଶ+ = ର୍ଶ

ଷ+କ+ର = ଷ୍କର

ଷ+ଟ+ର = ଷ୍ଟର

ଷ+ପ+ର = ଷ୍ପର

ସ+ଧ+ସ = ସ୍ଧସ



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

This is rare and is found only in the following:

ନ+ତ+ର+ୟ = ନ୍ତରୟ

ର+ତ+ତ+ୟ = ର୍ତତୟ

ର+ତ+ତ+ର = ର୍ତତର

ର+ଦ+ଧ+ = ର୍ଦଧ୍

ର+ଦ+ଧ+ = ର୍ଦଧ୍

Collation is one of the most important features of a script grammar. It determines the order in which a given culture indexes its characters. This is best seen in a dictionary sort where for easy search words are sorted and arranged in a specific order. Within a given script, each allo-script may have a different sort-order. Thus in Odia the conjunct glyph ଝଁ is sorted as in Marathi or Nepali at the end. The same is not the case with Hindi which admits a different sort order.

In the case of Odia the following is the traditional sort order as determined by the experts. 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.

The dictionary sort order as prescribed by lexicographers and used for dictionaries for lexicographic sorting

## The Computational collation order as used by computers

[illegible]

<sup>24</sup> A ligature which is a combination of 𐌵 and 𐌹 and is often treated as a single consonant.

## 7. REFERENCES

- Diringer, David. 1996 (1947). The alphabet: a key to the history of mankind. New Delhi: Munshiram Manoharlal. ISBN 81-215-0780-0
- Faulmann, Carl. 1990 (1880). Das Buch der Schrift. Frankfurt am Main: Eichborn. ISBN 3-8218-1720-8
- Gebrauch mit besonderer Berücksichtigung des Buchgewerbes. Unter Mitwirkung von Fachgelehrten zusammengestellt in der Reichsdruckerei. Berlin: Reichsdruckerei.
- Haarmann, Harald. 1990. Die Universalgeschichte der Schrift. Frankfurt: Campus. ISBN 3-593-34346-0
- ISCII'91
- Mahapatra, B. P. 1996. "Odia Writing", in Peter T. Daniels and William Bright, eds. The world's writing systems. New York; Oxford: Oxford University Press. ISBN 0-19-507993-0
- Praharaj, G. C. 1931. Purnacandra Odia Bhasa kosa. Cuttack.
- Reichsdruckerei. 1924. Alphabete und Schriftzeichen des Morgen- und Abendlandes, zum allgemeinen
- Taylor, Isaac. 1883. The alphabet: an account of the origin and development of letters. Vol. 1: Semitic alphabets; Vol. 2: Aryan alphabets. London: Kegan Paul.
- Unicode 5.2 and upwards. <http://www.unicode.org>

## 8. ANNEXURES

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

## Annexure 2: Unicode Table of Odia<sup>25</sup>

---

<sup>25</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