

Proposal to encode Phonetic Correction Mark in Devanagari

Srinidhi A
srinidhi.pinkpetals24@gmail.com

Sridatta A
sridatta.jamadagni@gmail.com

May 30, 2026

1 Introduction

This proposal requests to encode a new character in Devanagari. A phonetic correction mark used in Sanskrit and Prakrit manuscripts to indicate the intended phoneme when a nearby or phonetically similar letter is written under the influence of regional language pronunciation. Its encoding is necessary to ensure accuracy in the editing and representation of manuscripts.

Typically used to indicate and correct interchange between following consonants

- y / j
- kh / ʃ
- n / ṇ
- ś / ʃ / s etc

Example: मणि maṇi is grammatically correct word in Sanskrit. If it is written as मनि mani in manuscript then to indicate appropriate and correct Sanskrit/Prakrit reading the proposed character is used. The correct forms are identified based on the context. These are mostly used in Literary manuscripts (Kāvya) where reading correctly is very essential.

In response to the grapheme/phoneme issue raised by SEWG: A grapheme is the smallest unit of writing and is broadly comparable to an akshara in Indic scripts, while a phoneme is the smallest unit of sound. The purpose here is to indicate correct pronunciation while reading based on the sound values represented, not to modify the written graphemes. Hence, the name “phonetic correction mark” is appropriate.

2 Character Data

- U+11B0E  DEVANAGARI PHONETIC CORRECTION MARK
- Character properties:
11B0E; DEVANAGARI PHONETIC CORRECTION MARK ;Mn;0;NSM;;;;;N;;;;;
- Indic_Positional_Category=Top

3 References

1. सामान्यपांडुलिपि विज्ञान by Dr. महावीरप्रसाद शर्मा. Sāmānyapāṇḍulipi Vijnān by Dr. Mahāvīrprasād Śarmā. Apabhramsha Sahitya Academy , Rajasthan. 2002
2. Raghuvamśa of Kālidāsa - manuscript digitised by eGangotri
<https://archive.org/details/RaghuvanshKalidas780GhaAlm4Shlf3DevanagariKavya/page/n3/mode/1up>
3. Raghuvamśa of Kālidāsa - manuscript digitised by eGangotri
<https://archive.org/details/raghu-vamsha-sanskrit-manuscript-district-library-varanasi/>
4. Nitiśataka of Bhaṭṭhari- manuscript digitised by eGangotri
<https://archive.org/details/NitiShatakaBrihatHari229GhaAlm2Shlf2DevanagariKavya/page/n3/mode/1up>
5. Raghuvamśa of Kālidāsa - manuscript digitised by National Mission for Manuscripts
<https://www.pandulipipatala.nic.in/manus/pdf-view?content=MjkyNTYyL3BkZi9WQkkwMDAwMTEyODRfUmFnaHV2YeG5g8WbYV8xMzQ3X1BERkEucGRm>
6. Prabodha-candrodaya-nāṭaka
<https://gyanbharatam.com/manuscript/view?id=b7369a20-34a8-4120-98f2-77b4a470f507>

4 Attestations

The language of Figure 2 manuscript is Prakrit and rest of the manuscripts are in Sanskrit language.

Abbreviation:

GCW : Grammatically correct words - represent the correct words/spellings as per grammar.

GIW : Grammatically incorrect words- It reproduces and transliterates grammatically incorrect words/spellings exactly as they appear in the manuscripts graphically.

The corresponding letters are highlighted in green and red respectively.

27. कतिपय हस्तलिखित प्रतियों में संशोधकों के द्वारा किये हुए चिह्न भी मिलते हैं :-

(i) 'तेन' ^जजानन्ति' इस विधि से पद के अंतिम अक्षर के ऊपर खड़ी रेखा मिलती है। यह खड़ी रेखा (।) इस ओर संकेत करती है कि 'तेन' और 'जानन्ति' - ये दोनों पद अलग हैं।

 (ii) 'ख', 'क', ष या 'ज', 'य' अथवा 'श', 'ष', 'स' - इन अक्षरों में से किसी अक्षर के ऊपर यदि '∪' इस तरह का चिह्न किया गया हो तो वह उचित अक्षर पढ़ने का संकेत करता है।

Figure 1. Description of proposed sign in a Hindi language book on Manuscriptology. From Ref 1.

Translation :

In the letters 'kha', 'ka', 'ṣa' or 'ja', 'ya' or 'śa', 'ṣa' and 'sa' if **u** sign is placed above a letter, it indicates that the correct letter should be read from the context.

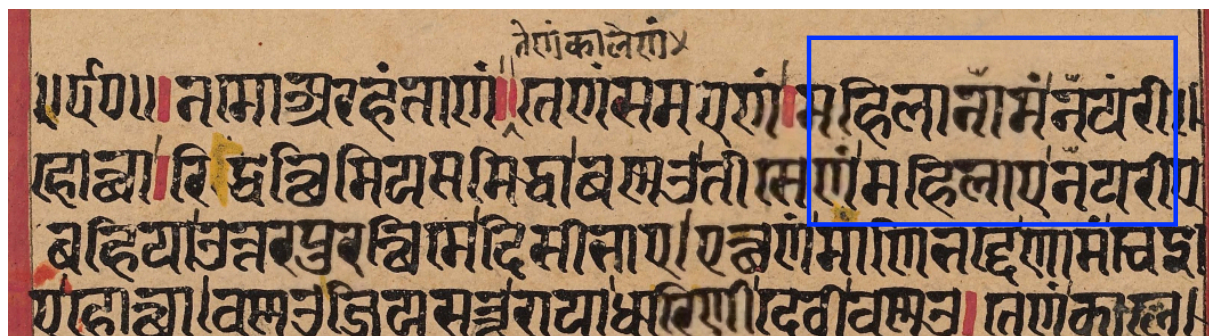


Figure 2. GCW: णामं , णयरी ṇāmaṁ , ṇayarī, GIW : नामं , नयरी nāmaṁ , nayarī

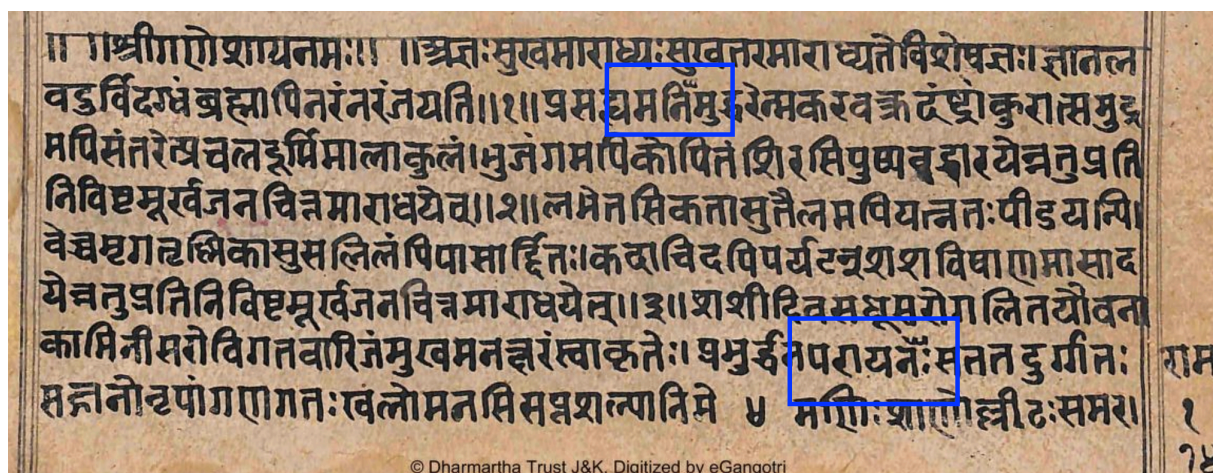


Figure 3. GCW: मणि, परायणः maṇi, parāyaṇaḥ, GIW : मनि, परायनः maṇi, parāyaṇaḥ. From Ref 4.

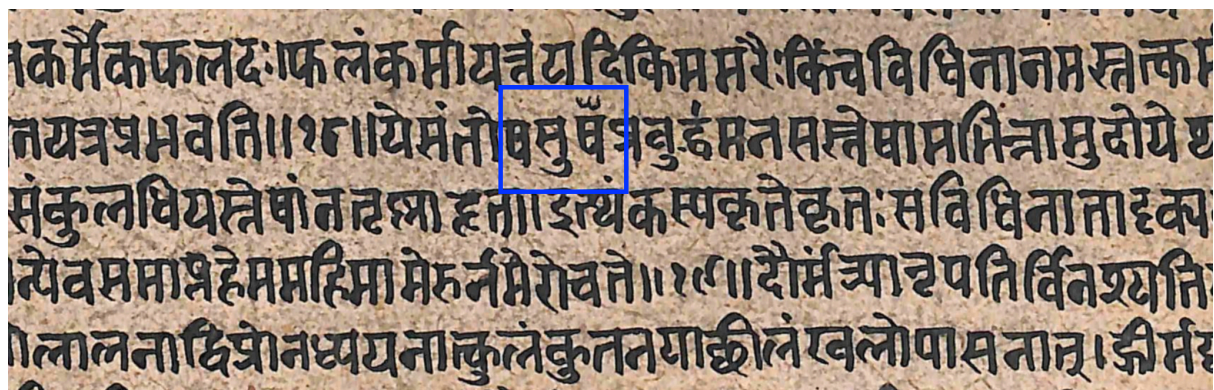


Figure 4. GCW: सुख sukha, GIW: सुष suṣa. From Ref 4.

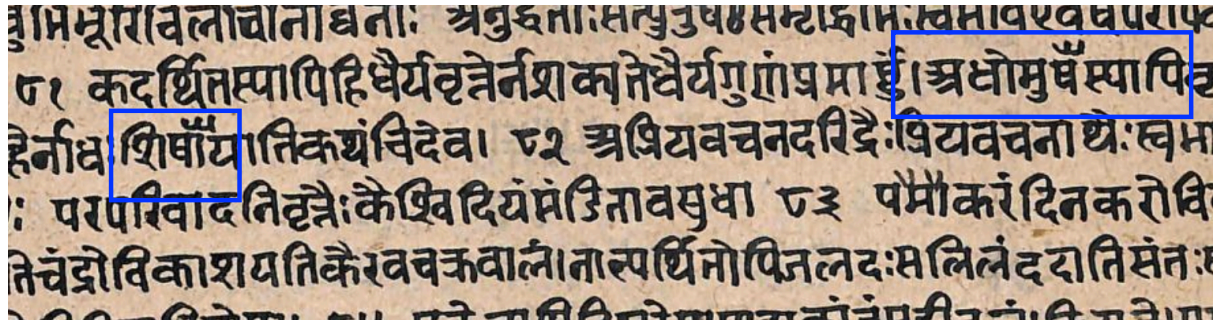


Figure 5. GCW: अधोमुखस्यापि, शिखा adhomukhasyāpi, śikhā, GIW: अधोमुषस्यापि, शिषा adhomuṣasyāpi, śiṣā. From Ref 4.

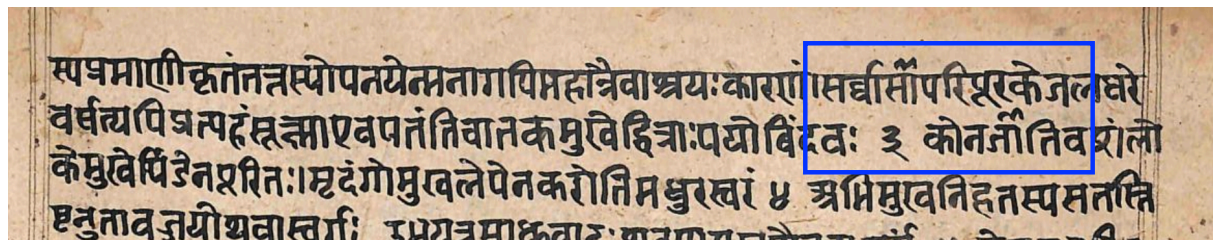


Figure 6. A variant with upturned form . This should be handled at font level as this form is rare. From Ref 4.

GCW: सर्वाशापरिपूरके sarvāśāparipūrake, याति yāti,

GIW: सर्वासापरिपूरके sarvāsāparipūrake, जाति jāti

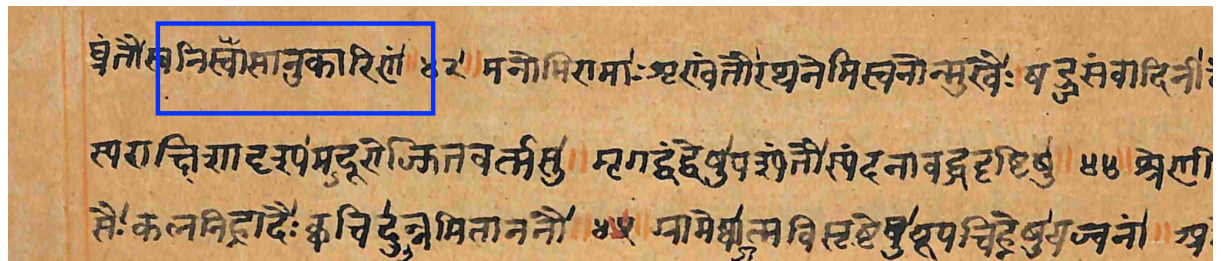


Figure 7. GCW: निश्वासा.. niśvāsā.., GIW: निस्वासा.. niṣvāsā.. From Ref 2.

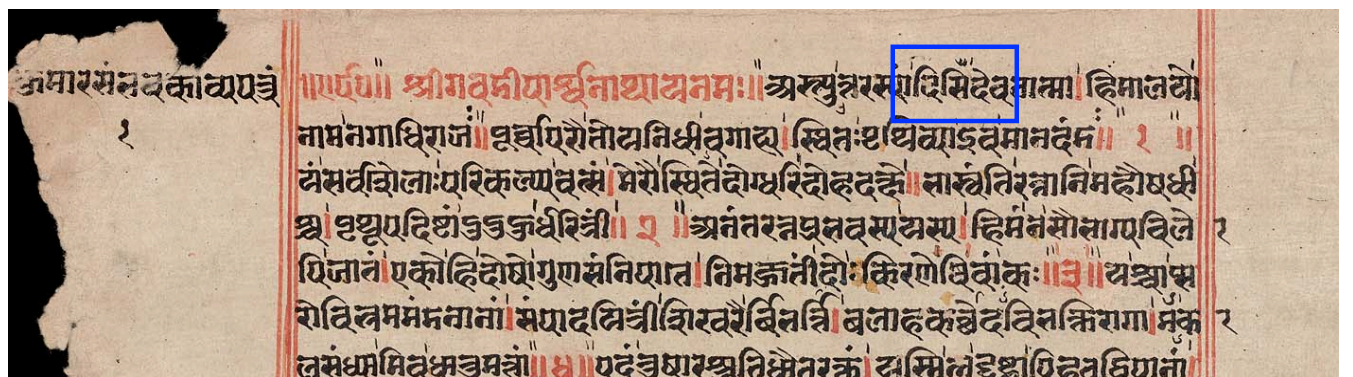


Figure 8. From the manuscript of Kumārasambhava by Kālidāsa.

GCW: दिसि di*si*, GIW: दिसि di*si*

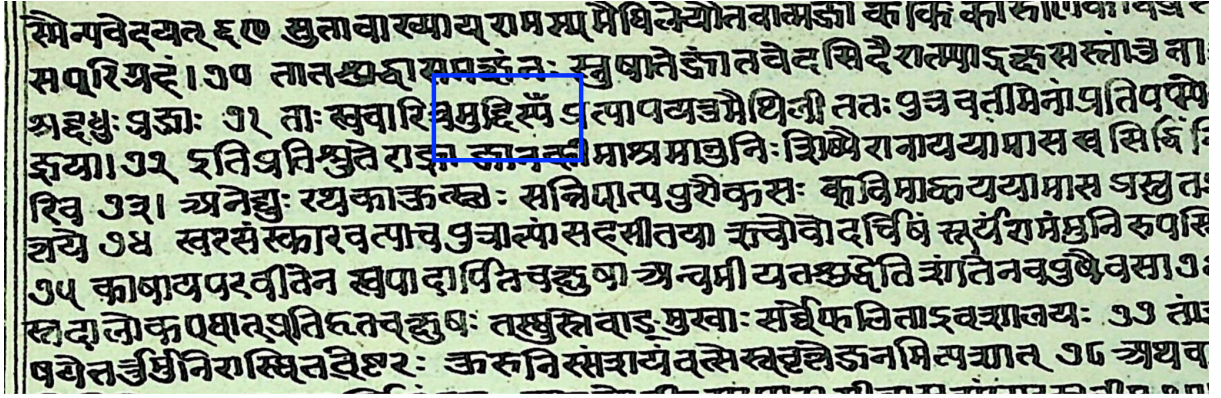


Figure 9. GCW: स्वचारित्रमुद्दिश्य svacāritramuddiśya, GIW: स्वचारित्रमुद्दिश्य svacāritramuddiśya. From Ref 3.

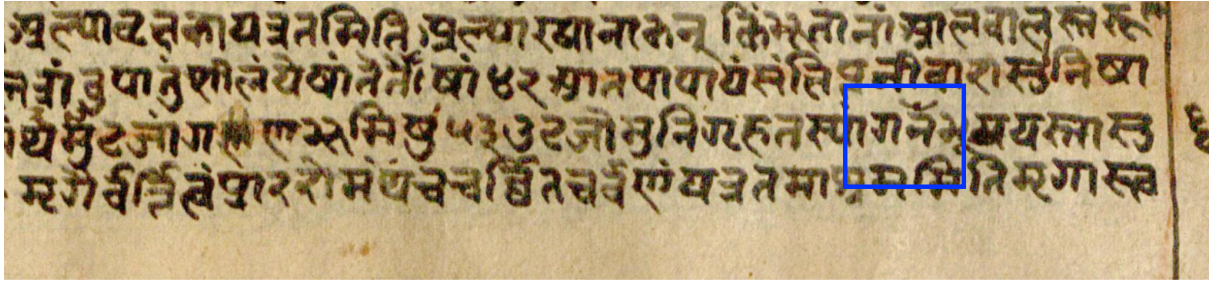


Figure 10. GCW: गण ga*ṇa*, GIW: गन ga*ṇa*. From Ref 5.

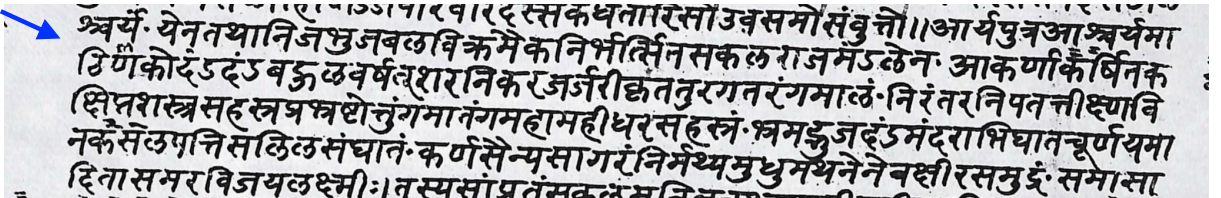


Figure 11. This uses the variant with upturned form. GCW: कठिन kaṭhi*ṇa*. GIW: कठिण kaṭhi*ṇa*. From Ref 6.

ISO/IEC JTC 1/SC 2/WG 2

PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646

Please fill all the sections A, B and C below.

Please read Principles and Procedures Document (P & P) from

<http://std.dkuug.dk/JTC1/SC2/WG2/docs/principles.html> for guidelines and details
before filling this form.

Please ensure you are using the latest Form from

<http://std.dkuug.dk/JTC1/SC2/WG2/docs/summaryform.html>.

See also <http://std.dkuug.dk/JTC1/SC2/WG2/docs/roadmaps.html> for latest *Roadmaps*.

Form number: N4502-F (Original 1994-10-14; Revised 1995-01, 1995-04, 1996-04, 1996-08, 1999-03, 2001-05, 2001-09, 2003-11, 2005-01, 2005-09, 2005-10, 2007-03, 2008-05, 2009-11, 2011-03, 2012-01)

A. Administrative

1. Title:	<i>Proposal to encode Phonetic correction mark in Devanagari</i>	
2. Requester's name:	<i>Srinidhi A and Sridatta A</i>	
3. Requester type (Member body/Liaison/Individual contribution):	<i>Individual Contribution</i>	
4. Submission date:	<i>May 30, 2026</i>	
5. Requester's reference (if applicable):		
6. Choose one of the following:		
This is a complete proposal:	<i>Yes</i>	
(or) More information will be provided later:		

B. Technical - General

1. Choose one of the following:

a. This proposal is for a new script (set of characters):

Proposed name of script:

b. The proposal is for addition of character(s) to an existing block:

Name of the existing block:

Devanagari

2. Number of characters in proposal:

1

3. Proposed category (select one from below - see section 2.2 of P&P document):

A-Contemporary

B.1-Specialized (small collection)

B.2-Specialized (large collection)

X

C-Major extinct

D-Attested extinct

E-Minor extinct

F-Archaic Hieroglyphic or Ideographic

G-Obscure or questionable usage symbols

4. Is a repertoire including character names provided?

Yes

a. If YES, are the names in accordance with the "character naming guidelines"

Yes

b. Are the character shapes attached in a legible form suitable for review?

Yes

5. Fonts related:

a. Who will provide the appropriate computerized font to the Project Editor of 10646 for publishing the standard?

Font glyph can be borrowed from Noto Sans Khojki. Glyph is identical to KHOJKI SIGN SHADDA

b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.):

Noto Sans Khojki is an Opensource font.

6. References:

a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?

Yes

b. Are published examples of use (such as samples from newspapers, magazines, or other sources)

of proposed characters attached?

Yes

7. Special encoding issue

Does the proposal address other aspects of character data processing (if applicable) such as input,

presentation, sorting, searching, indexing, transliteration etc. (if yes please enclose information)?

N/A

8. Submitters are invited to provide any additional information about Properties of the proposed Character(s) or Script that will assist in correct understanding of and correct linguistic processing of the proposed character(s) or script. Examples of such properties are: Casing information, Numeric information, Currency information, Display behaviour information such as line breaks, widths etc., Combining behaviour, Spacing behaviour, Directional behaviour, Default Collation behaviour, relevance in Mark Up contexts, Compatibility equivalence and other Unicode normalization related information. See the Unicode standard at <http://www.unicode.org> for such information on other scripts. Also see UAX#44: <http://www.unicode.org/reports/tr44/> and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.

C. Technical - Justification

1. Has this proposal for addition of character(s) been submitted before?	No
If YES explain	
2. Has contact been made to members of the user community (for example: National Body, user groups of the script or characters, other experts, etc.)?	No
If YES, available relevant documents:	
3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included?	Yes
Reference:	<i>Used by manuscript scholars of Devanagari</i>
4. The context of use for the proposed characters type of use; common or rare)	Common
Reference:	<i>Common in manuscripts, rare in modern use.</i>
5. Are the proposed characters in current use by the user community?	No
If YES, where? Reference:	
6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP?	No
If YES, is a rationale provided?	
If Yes, reference:	
7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?	N/A
8. Can any of the proposed characters be considered a presentation form of an existing character or character sequence?	No
If YES, is a rationale for its inclusion provided?	
If Yes, reference:	
9. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters?	No
If YES, is a rationale for its inclusion provided?	

If Yes, reference:	
10. Can any of the proposed character(s) be considered to be similar (in appearance or function)	
to, or could be confused with, an existing character?	Yes
If YES, is a rationale for its inclusion provided?	Yes
If Yes, reference:	<i>Glyph shape resembles Khojki Sign Shaddha</i>
11. Does the proposal include use of combining characters and/or use of composite sequences?	
If YES, is a rationale for such use provided?	Yes
If Yes, reference:	<i>Proposed character is combining mark</i>
Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?	No
If Yes, reference:	
12. Does the proposal contain characters with any special properties such as	
control function or similar semantics?	No
If YES, describe in detail (include attachment if necessary)	
13. Does the proposal contain any Ideographic compatibility characters?	
If YES, are the equivalent corresponding unified ideographic characters identified?	No
If Yes, reference:	

UnicodeData.txt

1	11B0E; DEVANAGARI PHONETIC CORRECTION MARK ;Mn;0;NSM;;;;;N;;;;;
---	---

IndicPositionalCategory.txt

11B0E	; Top
-------	-------