

Proposal to encode CHAKMA LETTER LHAA, DEPENDENT VOWEL SIGNS AA & EI for Chakma

Bivuti Chakma and Andrew Glass
(bsereye at hotmail dot com | anglass at outlook dot com)

November 2, 2016

Chakma script was added to Unicode in version 6.1 (2012) on the basis of Michael Everson’s proposal documents submitted in 2008 and 2009 (L2/08-133 & L2/09-187R). The present encoding for Chakma lacks three characters that are used by the Baarah Maatraa orthography for Chakma language. The present document addresses the gap in the Chakma encoding by proposing these signs as additions to the existing Chakma block (U+11100–U+1114F). The three characters are:

Glyph	Code	Character name
𑌒	11144	CHAKMA LETTER LHAA = fetfada lhaa
𑌓	11145	CHAKMA VOWEL SIGN AA = deil-dya aa
𑌔	11146	CHAKMA VOWEL SIGN EI = dvi adi gadei-dya ei

Michael Everson's original proposal was aware of the development of orthographies for Tanchangya language and anticipated the need for encoding additional characters. These signs are not being proposed in connection with Tachangya language. This document is not a statement regarding the completeness of the Chakma encoding for any other language or orthography other than Baarah Maatraa orthography for Chakma language. The characters proposed here, Chakma LETTER LHAA, VOWEL SIGNS AA & EI have also been used traditionally in regional Chakma orthographies and are not simply modern innovations.

The LETTER LHAA is used for the aspirated counterpart to LETTER LAA (figures 1, 2, 3 & 4). This sign is a CV aksara that may combine with dependent vowel signs, for example:

[illegible]

It can also occur in conjunct consonant clusters, for example:

ལྷ (lhyaa)

8 (11haa)

The VOWEL SIGN AA ၣ is a dependent vowel sign used to mark an explicit long A vowel (see figure 4, 5, 6, 7, 8, 12, 13, 14 & 15) in some orthographies. It is related to the Myanmar sign TALL AA (U+102B). There is no corresponding independent form of this vowel.

The VOWEL SIGN EI is a dependent vowel sign used to mark an explicit long E vowel mix with I vowel (see figure 9, 10, 11, 12, 13, 14 & 15). It is derived from the CHAKMA LETTER E (U+11106) & I (U+11104). There is no corresponding independent form of this vowel.

Character Data

Character properties

In the format of UnicodeData.txt:

```
11144;CHAKMA LETTER LHAA;Lo;0;L;;;;;N;;;;;
11145;CHAKMA VOWEL SIGN AA; Mc;0;L;;;;;N;;;;;
11146;CHAKMA VOWEL SIGN EI; Mc;0;L;;;;;N;;;;;
```

Linebreaking

In the format of LineBreak.txt:

```
11144;AL # Lo [1] CHAKMA LETTER LHAA
11145..11146;CM # Mc [2] CHAKMA VOWEL SIGN AA..CHAKMA VOWEL SIGN EI
```

Syllabic categories

In the format of IndicSyllabicCategory.txt:

```
# Indic_Syllabic_Category=Consonant

11144; Consonant # Lo [1] CHAKMA LETTER LHAA

# Indic_Syllabic_Category=Vowel_Dependent

11145..11146 ; Vowel_Dependent # Mc [2] CHAKMA VOWEL SIGN AA..CHAKMA VOWEL SIGN EI
```

Positional categories

In the format of IndicPositionalCategory.txt:

```
# Indic_Positional_Category=Right

11145; Right # Mc CHAKMA VOWEL SIGN AA

# Indic_Positional_Category=Right

11146; Right # Mc CHAKMA VOWEL SIGN EI
```

Sorting

Sort values should be assigned to these characters so that the following sort order is achieved:

𑌛 (U+11126) < 𑌛 (U+11144) < 𑌛 (U+11103)

𑌛 (U+11134) < 𑌛 (U+11145) < 𑌛 (U+11146)

𑌛 (U+11145) < 𑌛 (U+11146) < 𑌛 (U+11100)

References

Everson, Michael, 2008 “Preliminary proposal for encoding the Chakma script in the UCS”. N3428 L2/08-133.
<http://www.unicode.org/L2/L2008/08133-n3428-chakma.pdf>

Everson, Michael, 2009. “Proposal for encoding the Chakma script in the UCS”. N3645R L2/09-187R.
<http://www.unicode.org/L2/L2009/09187r-n3645r-chakma.pdf>

00

Chakma

1114F

	1110	1111	1112	1113	1114
0	◌̇ 11100	ꣳ 11110	ꣴ 11120	ꣵ 11130	ꣶ 11140
1	◌̈ 11101	ꣷ 11111	꣸ 11121	꣹ 11131	꣺ 11141
2	◌̉ 11102	ꣻ 11112	꣼ 11122	ꣽ 11132	ꣾ 11142
3	ꣿ 11103	꤀ 11113	꤁ 11123	꤂ 11133	꤃ 11143
4	꤄ 11104	꤅ 11114	꤆ 11124	꤇ 11134	꤈ 11144
5	꤉ 11105	ꤊ 11115	ꤋ 11125	ꤌ 11135	ꤍ 11145
6	ꤎ 11106	ꤏ 11116	ꤐ 11126	ꤑ 11136	ꤒ 11146
7	ꤓ 11107	ꤔ 11117	ꤕ 11127	ꤖ 11137	ꤗ 11147
8	ꤘ 11108	ꤙ 11118	ꤚ 11128	ꤛ 11138	ꤜ 11148
9	ꤝ 11109	ꤞ 11119	ꤟ 11129	ꤠ 11139	ꤡ 11149
A	ꤢ 1110A	ꤣ 1111A	ꤤ 1112A	ꤥ 1113A	ꤦ 1114A
B	ꤧ 1110B	ꤨ 1111B	ꤩ 1112B	ꤪ 1113B	꤫ 1114B
C	꤬ 1110C	꤭ 1111C	꤮ 1112C	꤯ 1113C	ꤰ 1114C
D	ꤲ 1110D	ꤳ 1111D	ꤴ 1112D	ꤵ 1113D	ꤶ 1114D
E	ꤷ 1110E	ꤸ 1111E	ꤹ 1112E	ꤺ 1113E	ꤻ 1114E
F	ꤼ 1110F	ꤽ 1111F	ꤾ 1112F	ꤿ 1113F	ꥀ 1114F

Add the three additional characters:

1. Chakma Consonant letter LHAA
= fetfada lhaa
= U+11144
2. Dependent Vowel Sign AA
= deil-dya aa
= U+11145
3. Dependent Vowel Sign EI
= dvi adi gadei-dya ei
= U+11146

The Unicode Standard 9.0, Copyright © 1991-2016 Unicode, Inc. All rights reserved.

11100

Various signs

- 11100 𑄀 CHAKMA SIGN CANDRABINDU
= caanaphupudaa
11101 𑄁 CHAKMA SIGN ANUSVARA
= ekaphudaa
11102 𑄂 CHAKMA SIGN VISARGA
= dviphudaa

Independent vowels

- 11103 𑄃 CHAKMA LETTER AA
= pichapujhaa aa
11104 𑄄 CHAKMA LETTER I
= delabhaangagaa i
11105 𑄅 CHAKMA LETTER U
= bacacu u
11106 𑄆 CHAKMA LETTER E
= lejaubaa e

Consonants

- 11107 𑄇 CHAKMA LETTER KAA
= cucyaangyaa kaa
11108 𑄈 CHAKMA LETTER KHAA
= grajaangyaa khaa
11109 𑄉 CHAKMA LETTER GAA
= caandyaa gaa
1110A 𑄊 CHAKMA LETTER GHAA
= tinaddaalyaa ghaa
1110B 𑄋 CHAKMA LETTER NGAA
= cilaama ngaa
1110C 𑄌 CHAKMA LETTER CAA
= dvibhalyaa caa
1110D 𑄍 CHAKMA LETTER CHAA
= majaraa chaa
1110E 𑄎 CHAKMA LETTER JAA
= dvipadala jaa
1110F 𑄏 CHAKMA LETTER JHAA
= uraauraa jhaa
11110 𑄐 CHAKMA LETTER NYAA
= silaacyaa nyaa
11111 𑄑 CHAKMA LETTER TAA
= dviyaadaat taa
11112 𑄒 CHAKMA LETTER THAA
= phudaadviyaat thaa
11113 𑄓 CHAKMA LETTER DDAA
= aadudaangaat ddaa
11114 𑄔 CHAKMA LETTER DDHAA
= lejabharaat ddhaa
11115 𑄕 CHAKMA LETTER NNAA
= petttuyaa nnaa
11116 𑄖 CHAKMA LETTER TAA
= ghangadaat taa
11117 𑄗 CHAKMA LETTER THAA
= jagadaat thaa
11118 𑄘 CHAKMA LETTER DAA
= dolaniit daa
11119 𑄙 CHAKMA LETTER DHAA
= talamuyaat dhaa
1111A 𑄚 CHAKMA LETTER NAA
= phaaraabaanyaa naa
1111B 𑄛 CHAKMA LETTER PAA
= paalyaa paa
1111C 𑄜 CHAKMA LETTER PHAA
= ubaraphudaa phaa
1111D 𑄝 CHAKMA LETTER BAA
= ubaramuyaa baa

Chakma

1113F

- 1111E 𑄞 CHAKMA LETTER BHAA
= ciraddaalyaa bhaa
1111F 𑄟 CHAKMA LETTER MAA
= bugatpadalaa maa
11120 𑄠 CHAKMA LETTER YYAA
= cimayyaa yyaa
11121 𑄡 CHAKMA LETTER YAA
= jilyaa yaa
11122 𑄢 CHAKMA LETTER RAA
= dvidaayyaa raa
11123 𑄣 CHAKMA LETTER LAA
= talamuyaa laa
11124 𑄤 CHAKMA LETTER WAA
= bajhonyaa waa
11125 𑄥 CHAKMA LETTER SAA
= bhudibukyaa saa
11126 𑄦 CHAKMA LETTER HAA
= ubaramuyaa haa

Dependent vowel signs

- 11127 𑄧 CHAKMA VOWEL SIGN A
= ubaratulyaa a
11128 𑄨 CHAKMA VOWEL SIGN I
= bahryaa i
11129 𑄩 CHAKMA VOWEL SIGN II
= baaniiphadaa ii
1112A 𑄪 CHAKMA VOWEL SIGN U
= ekattaana u
1112B 𑄫 CHAKMA VOWEL SIGN UU
= dvittaana uu
1112C 𑄬 CHAKMA VOWEL SIGN E
= ekaara e
1112D 𑄭 CHAKMA VOWEL SIGN AI
= delabhaanga ai
1112E 𑄮 CHAKMA VOWEL SIGN O
= okaara o
1112F 𑄯 CHAKMA VOWEL SIGN AU
= aukaara au
11130 𑄰 CHAKMA VOWEL SIGN OI
= oikaara oi
11131 𑄱 CHAKMA O MARK
11132 𑄲 CHAKMA AU MARK

Various signs

- 11133 𑄳 CHAKMA VIRAMA
• used to form conjuncts
→ 1039 𑄳 myanmar sign virama
11134 𑄴 CHAKMA MAAYYAA
• killer
→ 103A 𑄴 myanmar sign asat

Digits

- 11136 𑄵 CHAKMA DIGIT ZERO
11137 𑄶 CHAKMA DIGIT ONE
11138 𑄷 CHAKMA DIGIT TWO
11139 𑄸 CHAKMA DIGIT THREE
1113A 𑄹 CHAKMA DIGIT FOUR
1113B 𑄺 CHAKMA DIGIT FIVE
1113C 𑄻 CHAKMA DIGIT SIX
1113D 𑄼 CHAKMA DIGIT SEVEN
1113E 𑄽 CHAKMA DIGIT EIGHT
1113F 𑄾 CHAKMA DIGIT NINE

- 11144 𑄿 CHAKMA LETTER LHAA
= fetfada lhaa

- 11145 𑅀 CHAKMA VOWEL SIGN AA
= deil-dya aa

- 11146 𑅁 CHAKMA VOWEL SIGN EI
= dvi adi gadei-dya ei

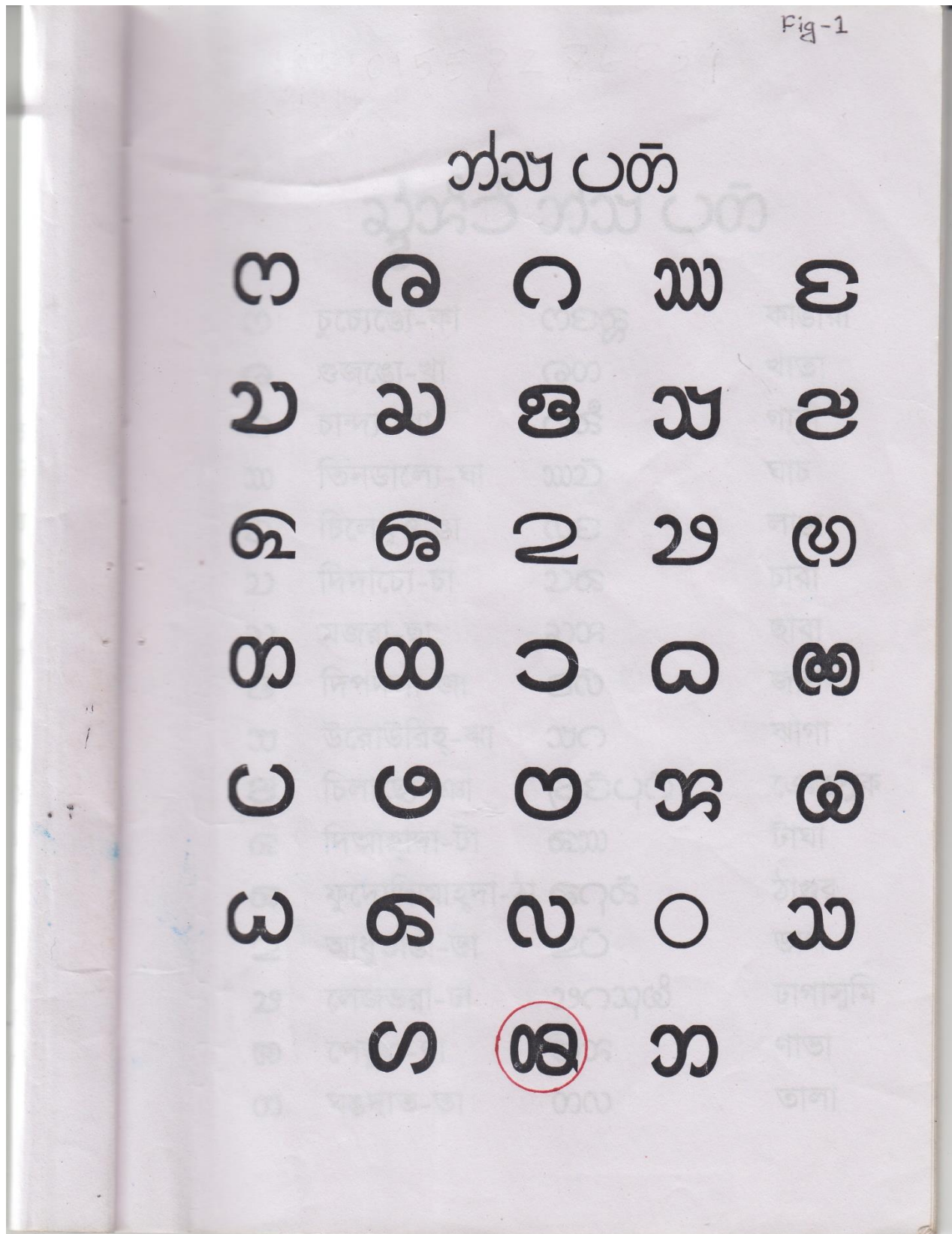


Figure 1: Usage of the proposed character consonant letter **LHAA** of Chakma [from *Chakma Legha Shigheni Bozar 2011 (Chakma Script learning year-2011)*. Source: Online Published from Tripura, India]

Fig-3.

। ৪৭ লেখা ।

এবার তলর লেখানি র' গরি মাদেনেই লেখি চা'

৐৐=মালাহ্ ২৩৬=চাবারা
৩৬=আধা ৬৬৬=পাগালা

সালে আমি জানিলঙ-

৐=মা, ৐=লাহ্, ২=চা, ৩=বা, ৬=রা, ৩=আ,
৬=ধা, ৬=পা, ৬=গা, আ ৬=লা।

রিনি চা' -

চাঙমাদি দিভে লা আঘন - ইক্কো লা (৬) আইক্কো লাহ্
৐।

ন চেয়ে গরি লিখি চা' -

মা = ...৐...	রা = ...৬.
লাহ্ = ...৐.	আ = ...৩.
চা = ...২.	ধা = ...৬.
বা = ...৩.	পা = ...৬.
গা = ...৬.	লা = ...৐.

Figure 3: Usage of the proposed character consonant letter **LHAA** of Chakma [from *30 Minide Chakma Legha Shigho (Learn Chakma Script within 30 minutes)*. Souce: Online Published, Tripura, India]

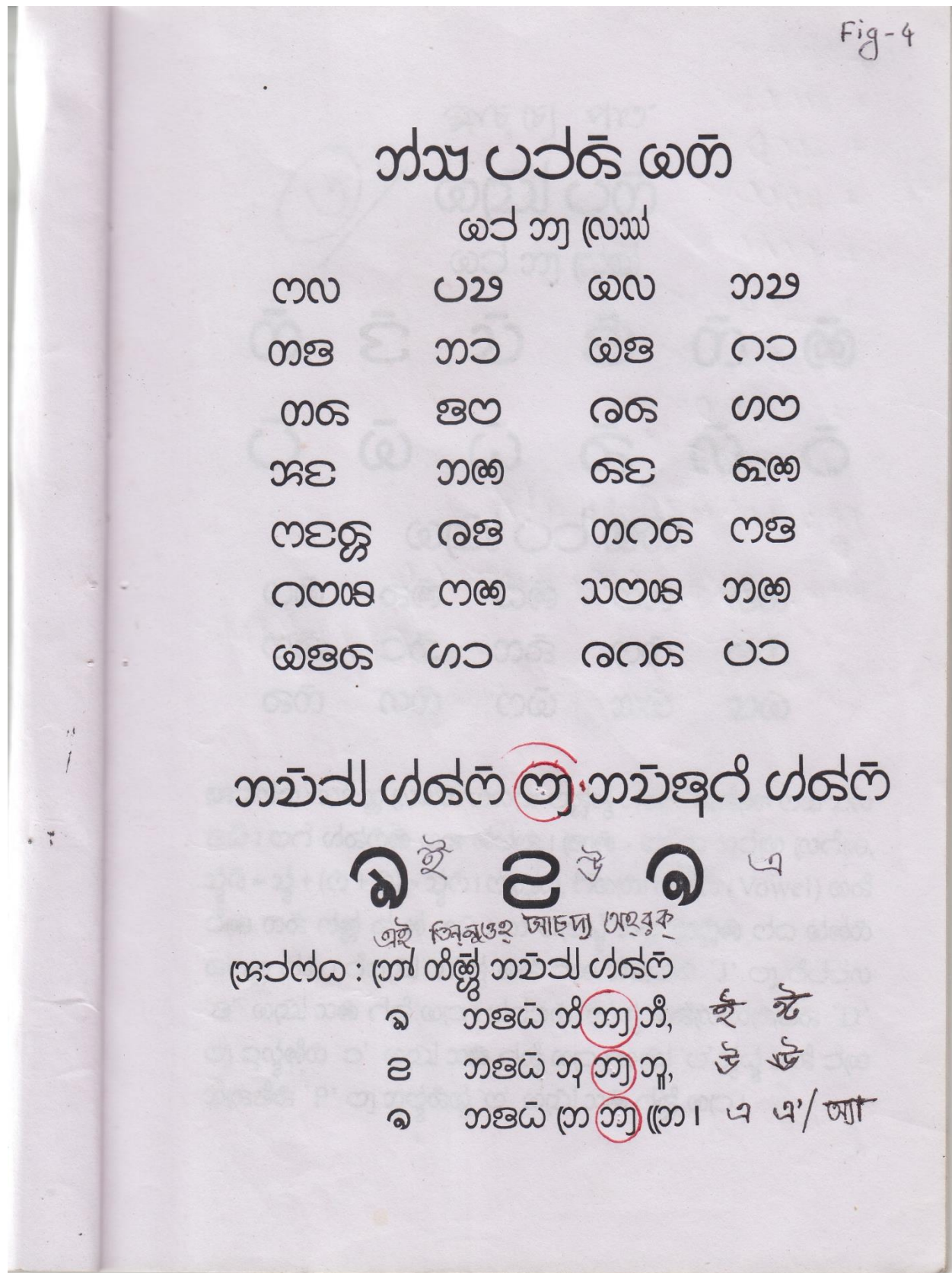


Figure 5: Usage of the proposed character Chakma Dependent Vowel Sign **AA** of Chakma [from *Chakma Legha Shigheni Bozar 2011 (Chakma Script learning year-2011)*. Source: Online Published from Tripura, India]

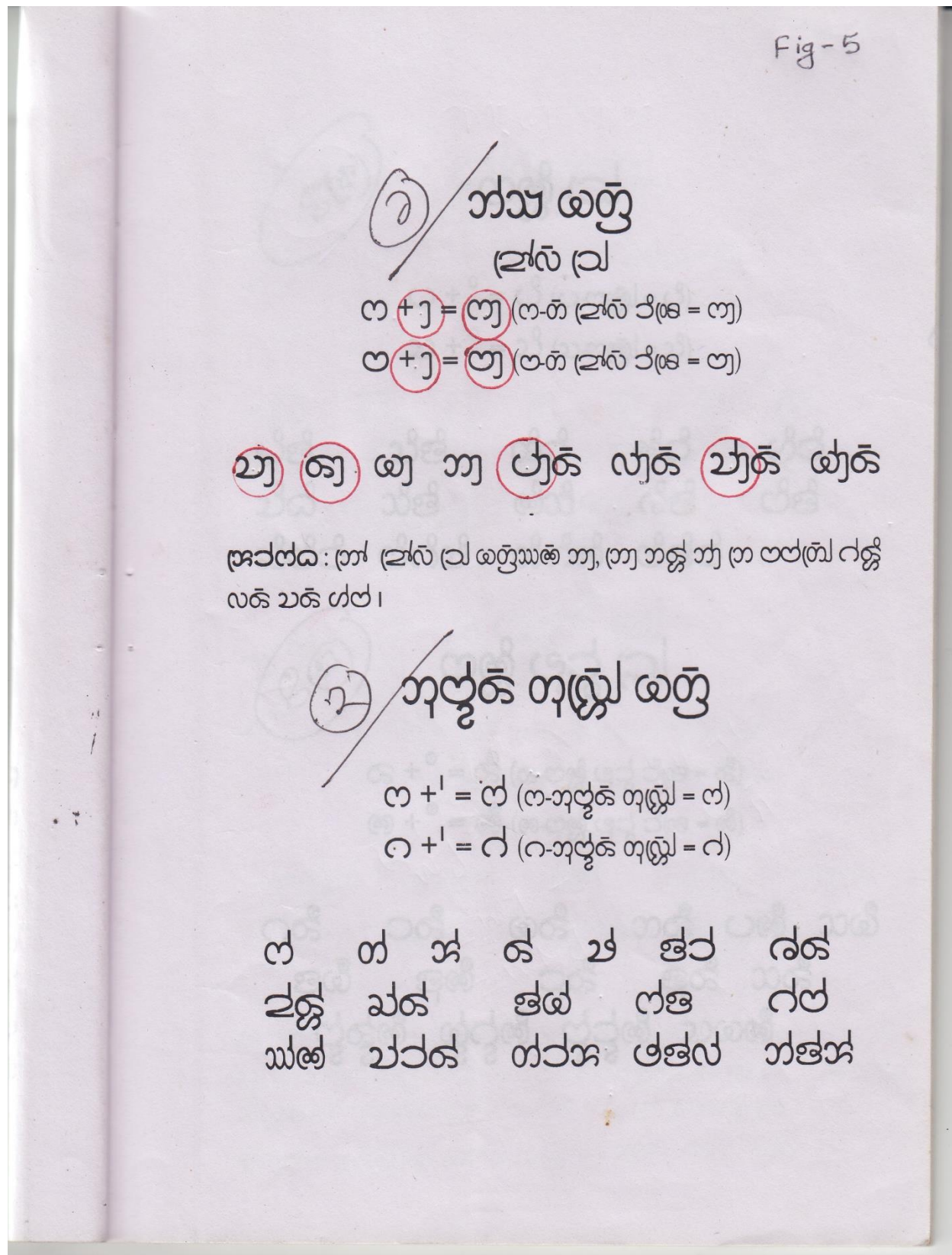


Figure 6: Usage of the proposed character Chakma Dependent Vowel Sign AA of Chakma [from *Chakma Legha Shigheni Bozar 2011 (Chakma Script learning year-2011)*. Source: Online Published from Tripura, India]

Fig - 6

■ ലെ യിരൂശലൈം ■

ইয়েনিহ্ মাদেই চ'-

ଓଡ଼ିଆ=ମାଜା

৐৐৐=মা'দালা

৩৫৫=বানা

৩০০=বানাহ

(৭) ইয়েনরে হলে ‘আ-কার’ আ (৯) ইয়েনরে হলে ‘আহ্ বাজেই দেনা’। পথমান অহরগর বাজগরে লাম্বা গরে আ জেররান গো’র গরে। তলর র’ঘুন মাদেই চ’ -

৮৭=পা'

২৭=চা'

୦୦୭୭=ଗାଦାନାହ

৐=মানাহ

ଘୋଷଣା=ମା'ଜନ

৬৯ = প'র

৩৩৫ = বদনাহ

ଓଠାଠୁ=ଜଦନାହ

ଐଠି=ନାରାହ

৳৳৳=মা'রাঙা

২৭৯ = চ'র

১০৮ = ম'র

ଓଢ଼ିଆ=ଫାଗାରା' ଲଢ଼ିଆ=କାନାହ

၁၇၇

ଓଢ଼ୁ=ବାରାହ

ଚାଓ ସାରା=୧୫ ଯକ ବର ଚା'ଓ=୧୫ ୧୫

ল'র দেনা=৩০ টাকা আহুদত ল'র=১০০ টাকা

মাগানা = ১০০০

মাগানাহ্=৫০০

ଜାର କାଳ=ପ୍ରତି ମନ୍ତ୍ର

ଜା'ର ଗରାନାହ୍=ଆଉଁ ଠକଞ୍ଜୁ

বাজানা=৩৩৩৩

ବା' ଜାନା= ୩ ଗୁଣ

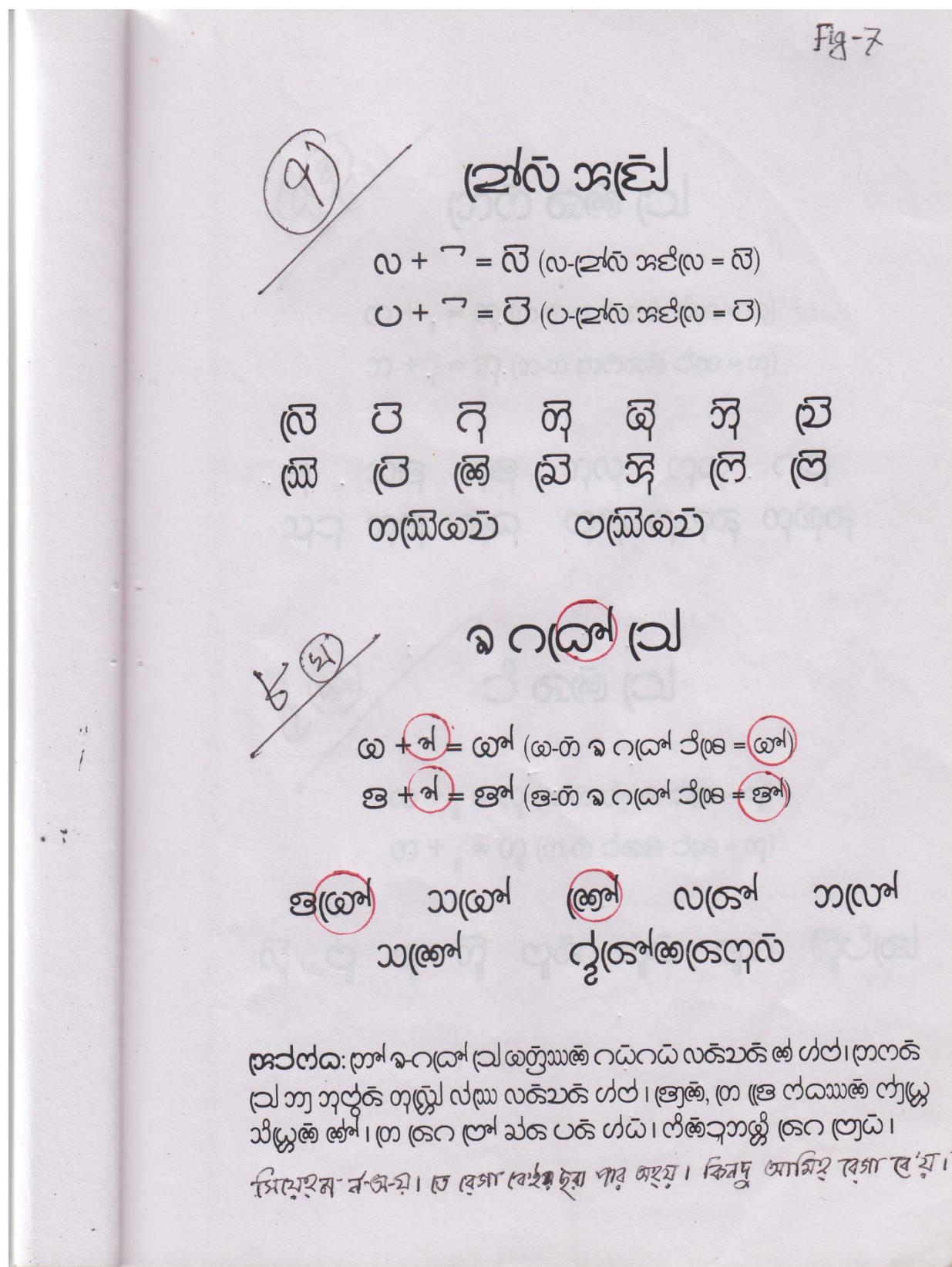


Figure 9: Usage of the proposed character Chakma Dependent Vowel Sign EI of Chakma [from *Chakma Legha Shigheni Bozar 2011 (Chakma Script learning year-2011)*. Source: Online Published from Tripura, India]

Fig - 8

। ৪৪ ঝিঝিট ।

গমে রিনি চ'-

৮ ৮=গায় গায়	৩৩ ৩৩=তাঘেই মাচ
৮ ৮=ঐখ গাঙ	৩৩ (৩৩=ভাত নেই
(৩৩)=এহুইত	৩৩=নয়
(৩৩)=নেই	৩৩(৩৩=সামেই

(৩) ইয়েনরে 'ধেল ভাঙে', (৮) ইয়েনরে 'এ-গাধেই দ্যে', আ (৩) ইয়েনরে 'ই গাধেই দ্যে' হুও অহুয়। তলর র'ঘুন মাদেই চ' -

(৩৩)=পেইক	৩৩=অয়
৮=গুই	৩৩ ৩৩=বাঘেই মাচ
৩৩(৩৩)=হাবিনেই	(৩৩)=ন'য়
(৩৩)=বেইন	৩৩ ৩৩=ভাত পৈ
(৩৩) ৩৩ ৩৩ ৩৩ ৩৩=রেগা বেই সরা পার অয়।	
হেইম=(৩৩)	অহুয়=(৩৩)
সানেই=৩৩(৩৩)	বাঘেইসুরি=৩৩৩৩
লুহুঙিনেই=৩৩(৩৩)	সদর ভেই=৩৩৩৩
মারি কেইম ফারি ন জেচ	=৩৩ (৩৩ ৩৩ ৩৩)।

৪৪/১৫

Figure 10: Usage of the proposed character Chakma Dependent Vowel Sign **EI** of Chakma [from 30 Minide Chakma Legha Shigho (Learn Chakma Script within 30 minutes)]

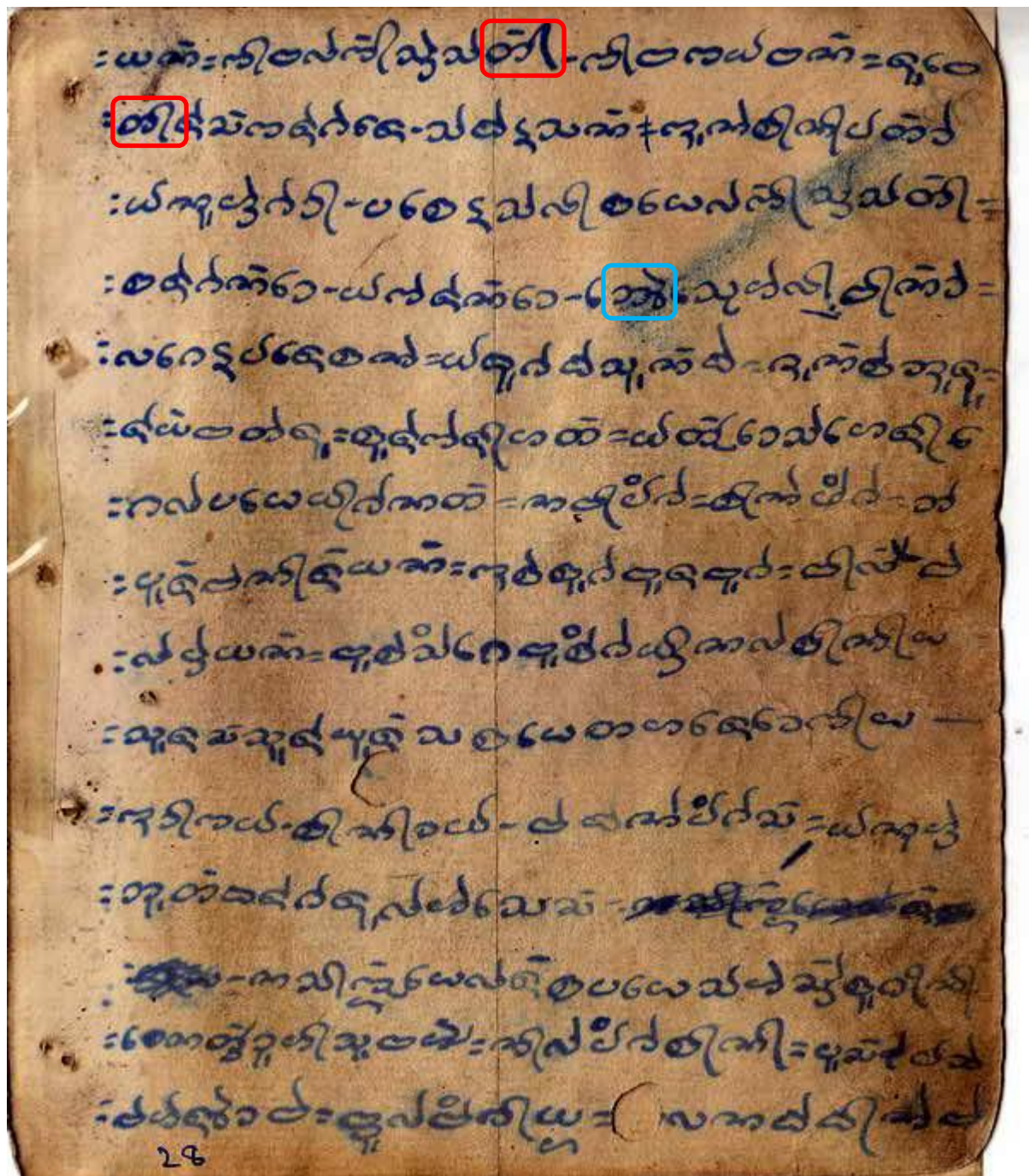


Figure 13: Usage of the proposed character Chakma Dependent Vowel Sign **AA** & **EI** of Chakma [from Chandobi Baaro Mach manuscript, circa 100 years old].

ཀའ་ཀ་ ལྟལ་ རྟལ་ ལའི་མའི་

ཀའི་ རྟལ་ ཀ་ ལྟལ་ (Follow and write)

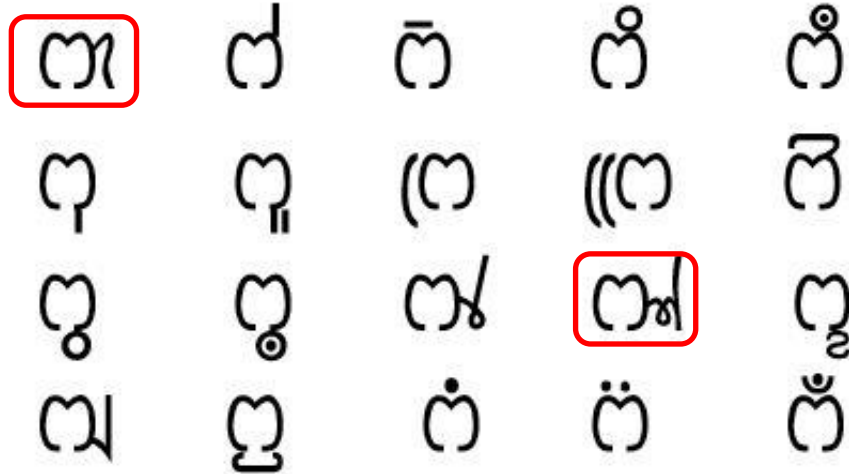


Figure 15: Usage of the proposed character Chakma Dependent Vowel Sign AA & EI of Chakma [from Prescribed as a Text Book by the Secretary District School Education Board, CADC, India. vide Notification No. 26/TMSTB/2008- 2009/84/200 Dt. 12/1/09.

ISO/IEC JTC 1/SC 2/WG 2

**PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS
FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646¹**

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.

A. Administrative

1. Title:	Proposal to encode CHAKMA LETTER LHAA, DEPENDENT VOWEL SIGNS AA & EI for Chakma		
2. Requester's name:	Bivuti Chakma, Andrew Glass		
3. Requester type (Member body/Liaison/Individual contribution):	Individual contribution		
4. Submission date:			
5. Requester's reference (if applicable):			
6. Choose one of the following:			
This is a complete proposal:	Complete		
(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:	11100–1114F		
Name of the existing block:	Chakma		
2. Number of characters in proposal:	3		
3. Proposed category (select one from below - see section 2.2 of P&P document):			
A-Contemporary	A	B.1-Specialized (small collection)	
C-Major extinct		D-Attested 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” in Annex L of P&P document?	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?	Bivuti Chakma		
b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.):	bsereye at hotmail.com		
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 issues:			
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)?	Yes		
	Shaping		

8. Additional Information:

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 Unicode Character Database (<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.

¹ 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)

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.)?	Yes
If YES, with whom? Bivuti Chakma	
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?	No
Reference:	
4. The context of use for the proposed characters (type of use; common or rare)	Common
Reference:	
5. Are the proposed characters in current use by the user community?	Yes
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)?	Yes
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?	No
If YES, is a rationale for its inclusion provided?	
If YES, reference:	
11. Does the proposal include use of combining characters and/or use of composite sequences?	No
If YES, is a rationale for such use provided?	
If YES, reference:	
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?	No
If YES, are the equivalent corresponding unified ideographic characters identified?	
If YES, reference:	