

TO: Unicode Technical Committee
FROM: Ken Whistler
TITLE: Towards a Consensus Encoding of Newa
DATE: 4 December 2014 (revised)

The “Recommendations to the UTC” (L2/14-253) provided input from members of a meeting held in Kathmandu, Nepal, from October 4-7, 2014, on the “Nepaalalipi”/”Newar” script. In an effort to help make progress, the following is a set of revised code charts charts that could serve as the basis for a formal proposal to the UTC and WG2.

This document contains 3 parts:

Part 1: The proposal summary form and code charts with the revised repertoire proposed for encoding in the Unicode Standard. The repertoire includes updated code points and names and other modifications which aim towards a consensus encoding. Note that the Newa deletion mark was considered best encoded as a new, generic combining mark in the Combining Diacritical Marks Supplement block: 1DFB COMBINING DELETION MARK. The character properties are included in this section.

Part 2: A Newa script subset appropriate for the Nepal Bhasa language modern use, omitting characters not needed for modern use, showing translated character names more appropriate to the community, and additional common use punctuation characters from other blocks also used in Nepal Bhasa text. (This information would be appropriate, for example, for establishing the exemplar characters for a Nepal Bhasa localization.)

Part 3: An explanatory chart demonstrating how the encoded Newa characters would be used to represent all required vowel and diphthong sequences for the popular modern orthography of Nepal Bhasa.

**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 the Newa script
2. Requester's name:	<i>Ken Whistler, Dev Dass Manandhar, Samir Karmacharya, Bishnu Chitrakar, Nepal Lipi Guthi, and Anshuman Pandey (Script Encoding Initiative)</i>
3. Requester type (Member body/Liaison/Individual contribution):	<i>Individual and liaison (SEI)</i>
4. Submission date:	<i>7 November 2014</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):	<i>yes</i>	
Proposed name of script:	<i>Newa</i>	
b. The proposal is for addition of character(s) to an existing block:		
Name of the existing block:	<i>(one character is added to Combining Diacritical Marks Supplement)</i>	
2. Number of characters in proposal:	<i>92+1</i>	
3. Proposed category (select one from below - see section 2.2 of P&P document):		
A-Contemporary <input checked="" type="checkbox"/>	B.1-Specialized (small collection) <input type="checkbox"/>	B.2-Specialized (large collection) <input type="checkbox"/>
C-Major extinct <input type="checkbox"/>	D-Attested extinct <input type="checkbox"/>	E-Minor extinct <input type="checkbox"/>
F-Archaic Hieroglyphic or Ideographic <input type="checkbox"/>	G-Obscure or questionable usage symbols <input type="checkbox"/>	
4. Is a repertoire including character names provided?	<i>yes</i>	
a. If YES, are the names in accordance with the "character naming guidelines" in Annex L of P&P document?	<i>yes</i>	
b. Are the character shapes attached in a legible form suitable for review?	<i>yes</i>	
5. Fonts related:		
a. Who will provide the appropriate computerized font to the Project Editor of 10646 for publishing the standard?	<i>Anshuman Pandey</i>	
b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.):	<i>pandey@umich.edu</i>	
6. References:		
a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?	<i>Yes (see docs cited below)</i>	
b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?	<i>See L2-12/03 and L2/14-086</i>	
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)?	<i>See L2-12/03</i>	

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? If YES explain	<i>yes</i> <i>L2/12-003 and L2/14-086 (which cites earlier proposals)</i>
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.)? If YES, with whom? If YES, available relevant documents:	<i>yes</i> <i>Historic and modern script users, see L2/14-253 and L2/12-003</i> <i>See above</i>
3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included?	<i>861,000 modern speakers of Nepal Bhasa/ for historic script users (numbers not available)</i>
Reference:	<i>Ethnologue.com (for number of modern language speakers)</i>
4. The context of use for the proposed characters (type of use; common or rare) Reference:	<i>common</i>
5. Are the proposed characters in current use by the user community? If YES, where? Reference:	<i>yes</i>
6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP? If YES, is a rationale provided? If YES, reference:	<i>no</i>
7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?	<i>yes</i>
8. Can any of the proposed characters be considered a presentation form of an existing character or character sequence? If YES, is a rationale for its inclusion provided? If YES, reference:	<i>no</i>
9. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters? If YES, is a rationale for its inclusion provided? If YES, reference:	<i>See L2/12-003</i> <i>yes</i> <i>See L2/12-003</i>
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? If YES, is a rationale for its inclusion provided? If YES, reference:	<i>yes</i> <i>yes</i> <i>See L2/12-003</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? If YES, reference: Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided? If YES, reference:	<i>yes</i> <i>yes</i> <i>See L2/12-003</i> <i>no</i>
12. Does the proposal contain characters with any special properties such as control function or similar semantics? If YES, describe in detail (include attachment if necessary)	<i>no</i>
13. Does the proposal contain any Ideographic compatibility characters? If YES, are the equivalent corresponding unified ideographic characters identified? If YES, reference:	<i>no</i>

	1DC	1DD	1DE	1DF
0				
1				
2				
3				
4				
5				
6				
7				
8				
9				
A				
B				
C				
D				
E				
F				

Miscellaneous mark

1DFB  COMBINING DELETION MARK
 • used for mhusaa in Newa scribal tradition



	1140	1141	1142	1143	1144	1145	1146	1147
0	ञ 11400	ग 11410	थ 11420	व 11430	ो 11440	० 11450		
1	आ 11401	घ 11411	द 11421	श 11431	ौ 11441	१ 11451		
2	ळ 11402	ढ 11412	ध 11422	ष 11432	् 11442	२ 11452		
3	ळी 11403	ड 11413	न 11423	स 11433	ं 11443	३ 11453		
4	उ 11404	च 11414	ळ 11424	ह 11434	ं 11444	४ 11454		
5	ऊ 11405	छ 11415	प 11425	ा 11435	ं४ 11445	५ 11455		
6	सू 11406	ज 11416	रु 11426	ि 11436	ि 11446	६ 11456		
7	सू 11407	म 11417	व 11427	ी 11437	इ 11447	७ 11457		
8	उ 11408	अ 11418	त 11428	्र 11438	ः 11448	८ 11458		
9	२ 11409	ळ 11419	म 11429	्र 11439	ँ 11449	९ 11459		
A	ए 1140A	ट 1141A	ळ 1142A	्र 1143A	१ 1144A			
B	वे 1140B	० 1141B	य 1142B	्र 1143B	। 1144B	३ 1145B		
C	अ 1140C	उ 1141C	न 1142C	्र 1143C	॥ 1144C			
D	औ 1140D	ट 1141D	ळ 1142D	्र 1143D	॥ 1144D	५ 1145D		
E	क 1140E	घ 1141E	ल 1142E	्र 1143E	॥ 1144E			
F	ख 1140F	ग 1141F	ळ 1142F	्र 1143F	० 1144F			

This script is also known as Nepaalalipi, Nepalakshar, Newah Akhah, Pachumol, Prachalit, and other names.

Independent vowels

Some of the vowels are only used for Sanskrit, and are not needed for Nepal Bhasa.

11400	अ	NEWA LETTER A
11401	आ	NEWA LETTER AA
11402	इ	NEWA LETTER I
11403	ई	NEWA LETTER II
11404	उ	NEWA LETTER U
11405	ऊ	NEWA LETTER UU
11406	ऋ	NEWA LETTER VOCALIC R
11407	ॠ	NEWA LETTER VOCALIC RR
11408	ऌ	NEWA LETTER VOCALIC L
11409	ॡ	NEWA LETTER VOCALIC LL
1140A	ऎ	NEWA LETTER E
1140B	ॐ	NEWA LETTER AI
1140C	॑	NEWA LETTER O
1140D	॒	NEWA LETTER AU

Consonants

6 consonant letters involving ha are encoded for the representation of murmured resonants in Nepal Bhasa, a Tibeto-Burman language. Those letters are not used for the representation of Sanskrit in the Newa script.

1140E	क	NEWA LETTER KA
1140F	ख	NEWA LETTER KHA
11410	ग	NEWA LETTER GA
11411	घ	NEWA LETTER GHA
11412	ङ	NEWA LETTER NGA
11413	झ	NEWA LETTER NGHHA
		• murmured nasal for Nepal Bhasa language
11414	च	NEWA LETTER CA
11415	छ	NEWA LETTER CHA
11416	ज	NEWA LETTER JA
11417	झ	NEWA LETTER JHA
11418	ञ	NEWA LETTER NYA
11419	ट	NEWA LETTER NYHA
		• murmured nasal for Nepal Bhasa language
1141A	ट	NEWA LETTER TTA
1141B	ठ	NEWA LETTER TTHA
1141C	ड	NEWA LETTER DDA
1141D	ढ	NEWA LETTER DDHA
1141E	ण	NEWA LETTER NNA
1141F	त	NEWA LETTER TA
11420	थ	NEWA LETTER THA
11421	द	NEWA LETTER DA
11422	ध	NEWA LETTER DHA
11423	न	NEWA LETTER NA
11424	न्	NEWA LETTER NHA
		• murmured nasal for Nepal Bhasa language
11425	प	NEWA LETTER PA
11426	फ	NEWA LETTER PHA
11427	ब	NEWA LETTER BA
11428	भ	NEWA LETTER BHA
11429	म	NEWA LETTER MA
1142A	म्	NEWA LETTER MHA
		• murmured nasal for Nepal Bhasa language
1142B	य	NEWA LETTER YA
1142C	र	NEWA LETTER RA

1142D	र	NEWA LETTER RHA
		• murmured tap for Nepal Bhasa language
1142E	ल	NEWA LETTER LA
1142F	ल्	NEWA LETTER LHA
		• murmured lateral for Nepal Bhasa language
11430	व	NEWA LETTER WA
11431	श	NEWA LETTER SHA
11432	ष	NEWA LETTER SSA
11433	स	NEWA LETTER SA
11434	ह	NEWA LETTER HA

Dependent vowel signs

Some of the vowel signs are only used for Sanskrit, and are not needed for Nepal Bhasa.

11435	◌ी	NEWA VOWEL SIGN AA
11436	◌ि	NEWA VOWEL SIGN I
11437	◌ि	NEWA VOWEL SIGN II
11438	◌ु	NEWA VOWEL SIGN U
11439	◌ू	NEWA VOWEL SIGN UU
1143A	◌ृ	NEWA VOWEL SIGN VOCALIC R
1143B	◌ॠ	NEWA VOWEL SIGN VOCALIC RR
1143C	◌ॡ	NEWA VOWEL SIGN VOCALIC L
1143D	◌ॢ	NEWA VOWEL SIGN VOCALIC LL
1143E	◌े	NEWA VOWEL SIGN E
1143F	◌ै	NEWA VOWEL SIGN AI
11440	◌ो	NEWA VOWEL SIGN O
11441	◌ौ	NEWA VOWEL SIGN AU

Various signs

11442	◌्	NEWA SIGN VIRAMA = tutisaalaa
11443	◌ं	NEWA SIGN CANDRABINDU = milaaphuti
11444	◌ं	NEWA SIGN ANUSVARA = sinhaphuti
11445	◌ं	NEWA SIGN VISARGA = liphuti
11446	◌्	NEWA SIGN NUKTA
11447	◌ं	NEWA SIGN AVAGRAHA = sulaa
11448	◌्	NEWA SIGN FINAL ANUSVARA = baadipu

Invocation signs

11449	ॐ	NEWA OM
1144A	ॐ	NEWA SIDDHI

Punctuation

1144B		NEWA DANDA = dipu
1144C		NEWA DOUBLE DANDA
1144D	,	NEWA COMMA = jhaasu
1144E	;	NEWA GAP FILLER = thaayjaayekaa
1144F	◌	NEWA ABBREVIATION SIGN

Digits

11450	◌	NEWA DIGIT ZERO = guli
11451	॑	NEWA DIGIT ONE = chi
11452	॒	NEWA DIGIT TWO = nasi

- 11453 𑌓 NEWA DIGIT THREE
= swa
- 11454 𑌔 NEWA DIGIT FOUR
= pi
- 11455 𑌕 NEWA DIGIT FIVE
= njaa
- 11456 𑌖 NEWA DIGIT SIX
= khu
- 11457 𑌗 NEWA DIGIT SEVEN
= nhasa
- 11458 𑌘 NEWA DIGIT EIGHT
= cyaa
- 11459 𑌙 NEWA DIGIT NINE
= gu

Various signs

- 1145B 𑌚 NEWA PLACEHOLDER MARK
= jaayekaa
- 1145C 𑌛 <reserved>
- 1145D 𑌜 NEWA INSERTION SIGN
= tansaa

PROPERTIES

1 General category properties

1DFB;COMBINING DELETION MARK;Mn;230;NSM ;;;;N;;;;;
11400;NEWA LETTER A;Lo;0;L;;;;;N;;;;;
11401;NEWA LETTER AA;Lo;0;L;;;;;N;;;;;
11402;NEWA LETTER I;Lo;0;L;;;;;N;;;;;
11403;NEWA LETTER II;Lo;0;L;;;;;N;;;;;
11404;NEWA LETTER U;Lo;0;L;;;;;N;;;;;
11405;NEWA LETTER UU;Lo;0;L;;;;;N;;;;;
11406;NEWA LETTER VOCALIC R;Lo;0;L;;;;;N;;;;;
11407;NEWA LETTER VOCALIC RR;Lo;0;L;;;;;N;;;;;
11408;NEWA LETTER VOCALIC L;Lo;0;L;;;;;N;;;;;
11409;NEWA LETTER VOCALIC LL;Lo;0;L;;;;;N;;;;;
1140A;NEWA LETTER E;Lo;0;L;;;;;N;;;;;
1140B;NEWA LETTER AI;Lo;0;L;;;;;N;;;;;
1140C;NEWA LETTER O;Lo;0;L;;;;;N;;;;;
1140D;NEWA LETTER AU;Lo;0;L;;;;;N;;;;;
1140E;NEWA LETTER KA;Lo;0;L;;;;;N;;;;;
1140F;NEWA LETTER KHA;Lo;0;L;;;;;N;;;;;
11410;NEWA LETTER GA;Lo;0;L;;;;;N;;;;;
11411;NEWA LETTER GHA;Lo;0;L;;;;;N;;;;;
11412;NEWA LETTER NGA;Lo;0;L;;;;;N;;;;;
11413;NEWA LETTER NGHHA;Lo;0;L;;;;;N;;;;;
11414;NEWA LETTER CA;Lo;0;L;;;;;N;;;;;
11415;NEWA LETTER CHA;Lo;0;L;;;;;N;;;;;
11416;NEWA LETTER JA;Lo;0;L;;;;;N;;;;;
11417;NEWA LETTER JHA;Lo;0;L;;;;;N;;;;;
11418;NEWA LETTER NYA;Lo;0;L;;;;;N;;;;;
11419;NEWA LETTER NYHA;Lo;0;L;;;;;N;;;;;
1141A;NEWA LETTER TTA;Lo;0;L;;;;;N;;;;;
1141B;NEWA LETTER TTHA;Lo;0;L;;;;;N;;;;;
1141C;NEWA LETTER DDA;Lo;0;L;;;;;N;;;;;
1141D;NEWA LETTER DDHA;Lo;0;L;;;;;N;;;;;
1141E;NEWA LETTER NNA;Lo;0;L;;;;;N;;;;;
1141F;NEWA LETTER TA;Lo;0;L;;;;;N;;;;;
11420;NEWA LETTER THA;Lo;0;L;;;;;N;;;;;
11421;NEWA LETTER DA;Lo;0;L;;;;;N;;;;;
11422;NEWA LETTER DHA;Lo;0;L;;;;;N;;;;;
11423;NEWA LETTER NA;Lo;0;L;;;;;N;;;;;
11424;NEWA LETTER NHA;Lo;0;L;;;;;N;;;;;
11425;NEWA LETTER PA;Lo;0;L;;;;;N;;;;;
11426;NEWA LETTER PHA;Lo;0;L;;;;;N;;;;;
11427;NEWA LETTER BA;Lo;0;L;;;;;N;;;;;
11428;NEWA LETTER BHA;Lo;0;L;;;;;N;;;;;
11429;NEWA LETTER MA;Lo;0;L;;;;;N;;;;;
1142A;NEWA LETTER MHA;Lo;0;L;;;;;N;;;;;
1142B;NEWA LETTER YA;Lo;0;L;;;;;N;;;;;
1142C;NEWA LETTER RA;Lo;0;L;;;;;N;;;;;
1142D;NEWA LETTER RHA;Lo;0;L;;;;;N;;;;;
1142E;NEWA LETTER LA;Lo;0;L;;;;;N;;;;;
1142F;NEWA LETTER LHA;Lo;0;L;;;;;N;;;;;
11430;NEWA LETTER WA;Lo;0;L;;;;;N;;;;;
11431;NEWA LETTER SHA;Lo;0;L;;;;;N;;;;;
11432;NEWA LETTER SSA;Lo;0;L;;;;;N;;;;;

11433;NEWA LETTER SA;Lo;0;L;;;;N;;;;;
11434;NEWA LETTER HA;Lo;0;L;;;;N;;;;;
11435;NEWA VOWEL SIGN AA;Mc;0;L;;;;N;;;;;
11436;NEWA VOWEL SIGN I;Mc;0;L;;;;N;;;;;
11437;NEWA VOWEL SIGN II;Mc;0;L;;;;N;;;;;
11438;NEWA VOWEL SIGN U;Mn;0;NSM;;;;N;;;;;
11439;NEWA VOWEL SIGN UU;Mn;0;NSM;;;;N;;;;;
1143A;NEWA VOWEL SIGN VOCALIC R;Mn;0;NSM;;;;N;;;;;
1143B;NEWA VOWEL SIGN VOCALIC RR;Mn;0;NSM;;;;N;;;;;
1143C;NEWA VOWEL SIGN VOCALIC L;Mn;0;NSM;;;;N;;;;;
1143D;NEWA VOWEL SIGN VOCALIC LL;Mn;0;NSM;;;;N;;;;;
1143E;NEWA VOWEL SIGN E;Mn;0;NSM;;;;N;;;;;
1143F;NEWA VOWEL SIGN AI;Mn;0;NSM;;;;N;;;;;
11440;NEWA VOWEL SIGN O;Mc;0;L;;;;N;;;;;
11441;NEWA VOWEL SIGN AU;Mc;0;L;;;;N;;;;;
11442;NEWA SIGN VIRAMA;Mn;9;NSM;;;;N;;;;;
11443;NEWA SIGN CANDRABINDU;Mn;0;NSM;;;;N;;;;;
11444;NEWA SIGN ANUSVARA;Mn;0;NSM;;;;N;;;;;
11445;NEWA SIGN VISARGA;Mc;0;L;;;;N;;;;;
11446;NEWA SIGN NUKTA;Mn;7;NSM;;;;N;;;;;
11447;NEWA SIGN AVAGRAHA;Lo;0;L;;;;N;;;;;
11448;NEWA SIGN FINAL ANUSVARA;Lo;0;L;;;;N;;;;;
11449;NEWA OM;Lo;0;L;;;;N;;;;;
1144A;NEWA SIDDHI;So;0;L;;;;N;;;;;
1144B;NEWA DANDA;Po;0;L;;;;N;;;;;
1144C;NEWA DOUBLE DANDA;Po;0;L;;;;N;;;;;
1144D;NEWA COMMA;Po;0;L;;;;N;;;;;
1144E;NEWA GAP FILLER;Po;0;L;;;;N;;;;;
1144F;NEWA ABBREVIATION SIGN;Po;0;L;;;;N;;;;;
11450;NEWA DIGIT ZERO;Nd;0;L;;0;0;0;N;;;;;
11451;NEWA DIGIT ONE;Nd;0;L;;1;1;1;N;;;;;
11452;NEWA DIGIT TWO;Nd;0;L;;2;2;2;N;;;;;
11453;NEWA DIGIT THREE;Nd;0;L;;3;3;3;N;;;;;
11454;NEWA DIGIT FOUR;Nd;0;L;;4;4;4;N;;;;;
11455;NEWA DIGIT FIVE;Nd;0;L;;5;5;5;N;;;;;
11456;NEWA DIGIT SIX;Nd;0;L;;6;6;6;N;;;;;
11457;NEWA DIGIT SEVEN;Nd;0;L;;7;7;7;N;;;;;
11458;NEWA DIGIT EIGHT;Nd;0;L;;8;8;8;N;;;;;
11459;NEWA DIGIT NINE;Nd;0;L;;9;9;9;N;;;;;
1145B;NEWA PLACEHOLDER MARK;So;0;L;;;;N;;;;;
1145D;NEWA INSERTION SIGN;So;0;L;;;;N;;;;;

2 Linebreaking properties

Linebreaking properties given in the data format of LineBreak.txt:

```

1DFB; CM # COMBINING DELETION MARK
11400..11434; AL # LETTER A .. LETTER HA
11435..11446; CM # SIGN AA .. SIGN NUKTA
11447..1144A; AL # SIGN AVAGRAHA .. SIDDHI
1144B..1144F; BA # DANDA .. ABBREVIATION SIGN
11450..11459; NU # DIGIT ZERO .. DIGIT NINE
1145B..1145D; AL # PLACEHOLDER MARK .. INSERTION SIGN

```

3 Syllabic Categories

Syllabic categories given in the data format of IndicSyllabicCategory.txt:

```
# Indic_Syllabic_Category=Bindu
11443..11444 ; Bindu # Mn [2] SIGN CANDRABINDU .. SIGN ANUSVARA

# Indic_Syllabic_Category=Visarga
11445 ; Visarga # Mc SIGN VISARGA

# Indic_Syllabic_Category=Avagraha
11447 ; Avagraha # Lo SIGN AVAGRAHA

# Indic_Syllabic_Category=Nukta
11446 ; Nukta # Mn SIGN NUKTA

# Indic_Syllabic_Category=Virama
11442 ; Virama # Mn SIGN VIRAMA

# Indic_Syllabic_Category=Vowel_Independent
11400..1140D ; Vowel_Independent # Lo [14] LETTER A .. LETTER AU

# Indic_Syllabic_Category=Vowel_Dependent
11435..11437 ; Vowel_Dependent # Mc [3] VOWEL SIGN AA .. VOWEL SIGN II
11438..1143F ; Vowel_Dependent # Mn [8] VOWEL SIGN U .. VOWEL SIGN AI
11440..11441 ; Vowel_Dependent # Mc [2] VOWEL SIGN O .. VOWEL SIGN AU

# Indic_Syllabic_Category=Consonant
1140E..11434 ; Consonant # Lo [39] LETTER KA .. LETTER HA
```

4 Positional Categories

The following positional categories are given in the data format of IndicPositionalCategory.txt:

```
# Indic_Positional_Category=Right
11435 ; Right # Mc VOWEL SIGN AA
11437 ; Right # Mc VOWEL SIGN II

# Indic_Positional_Category=Left
11436 ; Left # Mc VOWEL SIGN I

# Indic_Positional_Category=Top
1143E..1143F ; Top # Mn [2] VOWEL SIGN E .. VOWEL SIGN AI

# Indic_Positional_Category=Top_And_Right
11440..11441 ; Top_And_Right # Mc [2] VOWEL SIGN O .. VOWEL SIGN AU

# Indic_Positional_Category=Bottom
11438..1143D ; Bottom # Mn [6] VOWEL SIGN U .. VOWEL SIGN VOCALIC LL
11442 ; Bottom # Mn SIGN VIRAMA
```

5 Confusable Categories

Several Newa characters are similar to those for Devanagari, Bengali, and other modern Indic scripts. 'Confusion' data for these similarities are provided below:

```
11400 NEWA LETTER A ; 0985 BENGALI LETTER A
11401 NEWA LETTER AA ; 0986 BENGALI LETTER AA
11404 NEWA LETTER U ; 0909 DEVANAGARI LETTER U
```

11405	NEWA	LETTER	UU	;	090A	DEVANAGARI	LETTER	UU				
1140E	NEWA	LETTER	KA	;	0915	DEVANAGARI	LETTER	KA				
1140F	NEWA	LETTER	KHA	;	0916	DEVANAGARI	LETTER	KHA				
11410	NEWA	LETTER	GA	;	0997	BENGALI	LETTER	GA				
11411	NEWA	LETTER	GHA	;	0918	DEVANAGARI	LETTER	GHA				
11416	NEWA	LETTER	JA	;	099C	BENGALI	LETTER	JA				
11417	NEWA	LETTER	JHA	;	11497	TIRHUTA	LETTER	JHA				
11418	NEWA	LETTER	NYA	;	099E	BENGALI	LETTER	NYA				
1141B	NEWA	LETTER	TTHA	;	1119C	SHARADA	LETTER	TTHA				
1141E	NEWA	LETTER	NNA	;	1149D	TIRHUTA	LETTER	NNA				
11420	NEWA	LETTER	THA	;	0925	DEVANAGARI	LETTER	THA				
11421	NEWA	LETTER	DA	;	0926	DEVANAGARI	LETTER	DA				
11422	NEWA	LETTER	DHA	;	0927	DEVANAGARI	LETTER	DHA				
11423	NEWA	LETTER	NA	;	0928	DEVANAGARI	LETTER	NA				
11425	NEWA	LETTER	PA	;	092A	DEVANAGARI	LETTER	PA				
11427	NEWA	LETTER	BA	;	0935	DEVANAGARI	LETTER	VA				
11429	NEWA	LETTER	MA	;	092E	DEVANAGARI	LETTER	MA				
1142B	NEWA	LETTER	YA	;	092F	DEVANAGARI	LETTER	YA				
1142E	NEWA	LETTER	LA	;	0932	DEVANAGARI	LETTER	LA				
11430	NEWA	LETTER	WA	;	0935	DEVANAGARI	LETTER	VA				
11432	NEWA	LETTER	SSA	;	0937	DEVANAGARI	LETTER	SSA				
11433	NEWA	LETTER	SA	;	0938	DEVANAGARI	LETTER	SA				
11434	NEWA	LETTER	HA	;	0939	DEVANAGARI	LETTER	HA				
11435	NEWA	VOWEL	SIGN	AA	;	093E	DEVANAGARI	VOWEL	SIGN	AA		
11436	NEWA	VOWEL	SIGN	I	;	093F	DEVANAGARI	VOWEL	SIGN	I		
11437	NEWA	VOWEL	SIGN	II	;	0940	DEVANAGARI	VOWEL	SIGN	II		
11438	NEWA	VOWEL	SIGN	U	;	09C1	BENGALI	VOWEL	SIGN	U		
11439	NEWA	VOWEL	SIGN	UU	;	09C2	BENGALI	VOWEL	SIGN	UU		
1143A	NEWA	VOWEL	SIGN	VOCALIC	R	;	0943	DEVANAGARI	VOWEL	SIGN	VOCALIC	R
1143B	NEWA	VOWEL	SIGN	VOCALIC	RR	;	0944	DEVANAGARI	VOWEL	SIGN	VOCALIC	RR
11442	NEWA	SIGN	VIRAMA	;	094D	DEVANAGARI	SIGN	VIRAMA				
11443	NEWA	SIGN	CANDBINDU	;	0F82	TIBETAN	SIGN	NYI	ZLA	NAA	DA	
11444	NEWA	SIGN	ANUSVARA	;	114C0	TIRHUTA	SIGN	ANUSVARA				
11448	NEWA	SIGN	FINAL	ANUSVARA	;	0982	BENGALI	SIGN	ANUSVARA			
1144F	NEWA	ABBREVIATION	SIGN	;	0970	DEVANAGARI	ABBREVIATION					

PART 2: UNICODE NEWA SUBSET FOR NEPAL BHASA LANGUAGE USE

The following is a list of characters needed for Nepal Bhasa language use. The italicized names reflect the preferred names of the characters by the experts attending the script meeting in Nepal. The highlighted letters and words indicate any differences from the code chart names.

Independent Vowels

- 11400 अ *letter a*
- 11401 आ *letter aa*
- 11402 ई *letter i*
- 11403 ऐ *letter ii*
- 11404 उ *letter u*
- 11405 ऊ *letter uu*
- 1140A ए *letter e*
- 1140C ओ *letter o*

Consonants

- 1140E क *letter k*
- 1140F ख *letter kh*
- 11410 ग *letter g*
- 11411 घ *letter gh*
- 11412 ङ *letter ng*
- 11413 ञ *letter ngh*
- 11414 च *letter c*
- 11415 छ *letter ch*
- 11416 ज *letter j*
- 11417 झ *letter jh*
- 11418 ञ *letter nj*
- 11419 ञ *letter njh*
- 1141A ट *letter tt*
- 1141B ठ *letter tth*
- 1141C ड *letter dd*
- 1141D ढ *letter ddh*
- 1141E ण *letter nn*
- 1141F त *letter t*
- 11420 थ *letter th*
- 11421 द *letter d*
- 11422 ध *letter dh*

11423 न letter n
11424 ण letter nh
11425 प letter p
11426 ष letter ph
11427 ब letter b
11428 भ letter bh
11429 म letter m
1142A ऋ letter mh
1142B य letter y
1142C र letter r
1142D ऌ letter rh
1142E ल letter l
1142F ळ letter lh
11430 व letter w
11431 श letter sh
11432 ष letter ss
11433 स letter s
11434 ह letter h

Dependent vowel signs

11435 ा vowel sign aa
11436 ि vowel sign i
11437 ी vowel sign ii
11438 ू vowel sign u
11439 ्र vowel sign uu
1143E े vowel sign e
1143F ै vowel sign ai
11440 ो vowel sign o
11441 ौ vowel sign au

Various signs

11442 ् sign tutisaalaa
11443 ॆ sign milaaphuti
11444 ॆ sign sinhaphuti
11445 ॆ sign liphuti
11447 ॆ sign sulaa
11448 ॆ sign baadipu

Invocation signs

11449 ॐ *om*

1144A ॐ *siddhi*

Punctuation

1144B | *dipu*

1144D \ *jhaasu*

1144E † *thaayjaayekaa*

Digits

11450 ○ *guli*

11451 ॐ *chi*

11452 ॐ *nasi*

11453 ॐ *swa*

11454 ॐ *pi*

11455 ॐ *njaa*

11456 ॐ *khu*

11457 ॐ *nhasa*

11458 ॐ *cyaa*

11459 ॐ *gu*

Various signs

1145B ~ *jaayekaa*

1145D ✕ *tansaa*

Characters from other blocks

0020 SPACE

0021 ! EXCLAMATION MARK

0023 # NUMBER SIGN

0028 (LEFT PARENTHESIS

0029) RIGHT PARENTHESIS

002A * ASTERISK

002C , COMMA

002D - HYPHEN-MINUS

002E . FULL STOP

003F ? QUESTION MARK

2010 - HYPHEN

2018 ' LEFT SINGLE QUOTATION MARK

2019 ' RIGHT SINGLE QUOTATION MARK

2055 ✨ FLOWER PUNCTUATION MARK

= swaan (used with dandas)

1DFB 〰 COMBINING DELETION MARK

Notes:

1. The following characters were deemed controversial amongst experts and are postponed:

☞ *svasti*

· *high spacing dot*

× *abbreviation sign cross*

⚡ *vajra*

☰ *swaapu*

2. Secondary weighting of the retroflex consonants will produce the following order of consonants:

K, KH, G, GH, NG, NGH, C, CH, J, JH, NJ, NJH, T, TT, TH, TTH, D, DD, DH, DDH, N, NN,
NH, P, PH, B, BH, M, MH, Y, R, RH, L, LH, W, S, SH, SS, H

PART 3: UNICODE NEWA SYLLABLE CHARTS FOR NEPAL BHASA LANGUAGE USE

The tables below draw from tables in *Nepālabhāsā: A Monosyllabic Language* by Bishnu Chitrakar (Kathmandu, 2013), but the Devanagari has been converted to Newa (Table 1 below = Table 10 in *Nepālabhāsā* [page 24], Table 2 below =Table 12b [p. 27], Table 3 below=Table 12a [pp. 27]). Note: In the following charts, “ā” indicates /a/, “a” indicates /ɔ/.

Table 1

NON-NASALIZED		NASALIZED	
short	long	short	long
क	कः	कँ	कँः
ka	ka:	kã	kã:
1140E	1140E 11445	1140E	1140E 11443
का	काः	काँ	काँः
kā	kā:	kā̃	kā̃:
1140E	1140E 11435 11445	1140E	1140E 11435 11443
कि	की / किळ	किँ	किँ / किँळ
ki	ki:/kii	kĩ	kĩ:/kĩi
1140E	1140E	1140E	1140E
11436	11437 /1140E 11436 11402	11436	11436 11444 /1140E 11436 11443 11402
कु	कू / कुउ	कुँ	कुँ / कुँउ
ku	ku:/kuu	kũ	kũ:/kũu
1140E	1140E	1140E	1140E
11438	11439 /1140E 11438 11404	11438	11438 11444 /1140E 11438 11443 11404
के	केप्	केँ	केँ / केँप्
ke	ke:/kee	kẽ	kẽ:/kẽe
1140E	1140E	1140E	1140E
1143E	1143E 1142B 11442	1143E	1143E 11444 /1140E 1143E 11443 1142B 11442
को	कोः	कोँ	कोँः
ko	ko:	kõ	kõ:
1140E	1140E	1140E	1140E
11440	11440 11445	11440	11440 11444

TABLE 2

	(a)	(i)	(u)	(e)
क	कऌ = कॡ	कऌ / के	कऌ / को	कॡ = कय्
ka	ka - a = ka:	ka - i / ka - ai	ka - u / ka - au	ka - e
1140E	1140E	1140E	1140E	1140E
	11400	11402	11404	1140A
	=1140E	/1140E	/1140E	=1140E
	11445	1143F	11441	1142B
				11442
का	काऌ = काॡ	काऌ	काऌ	काॡ / काय्
kā	kā - a = kā:	kā - i	kā - u	kā - e
1140E	1140E	1140E	1140E	1140E
11435	11435	11435	11435	11435
	11400	11402	11404	1140A
	=1140E			/1140E
	11435			11435
	11445			1142B
				11442
कि	-	किऌ / की	किऌ / *	किॡ = किऌ
ki	-	ki - i	ki - u	ki - e = ki - i
1140E		1140E	1140E	1140E
11436		11436	11436	11436
		11402	11404	1140A
		/1140E	/* font	=1140E
		11437	missing	11436
			glyphs	11402
कु	-	कुऌ	कुऌ / कु	कुॡ = कुऌ
ku	-	ku - i	ku - u	ku - e = ku - i
1140E		1140E	1140E	1140E
11438		11438	11438	11438
		11402	11404	1140A
			/1140E	=1140E
			11439	11438
				11402
के	-	केऌ	केऌ	केॡ = केय्
ke	-	ke - i	ke - u	ke - e
1140E		1140E	1140E	1140E
1143E		1143E	1143E	1143E
		11402	11404	1140A
				=1140E
				1143E
				1142B
				11442

TABLE 3

	(a)	(i)	(u)	(e)
क़	क़ = क़	क़ / क़	क़ / क़	क़ / क़
kā	kā - a = kā:	kā - i	kā - u	kā - e
1140E	1140E	1140E	1140E	1140E
11443	11443	11443	11443	11443
	11400	11402	11404	1140A
	=1140E	/1140E	/1140E	/1140E
	11444	1143F	11441	11443
		11443	11443	1142B
				11442
का	का = का	का	का	का = का
kā	kā - a = kā:	kā - i	kā - u	kā - e
1140E	1140E	1140E	1140E	1140E
11435	11435	11435	11435	11435
11443	11443	11443	11443	11443
	11400	11402	11404	1140A
	=1140E			=1140E
	11435			11435
	11444			11443
				1142B
				11442
कि	-	कि / कि	कि / **	कि = कि
kī	-	kī - i	kī - u	kī - e = kī-i
1140E		1140E	1140E	1140E
11436		11436	11436	11436
11443		11443	11443	11443
		11402	11404	1140A
		/1140E	/** font	=1140E
		11436	missing	11436
		11444	glyphs	11443
				11402
कु	-	कु	कु / कु	कु = कु
kū	-	kū - i	kū - u	kū - e = kū - i
1140E		1140E	1140E	1140E
11438		11438	11438	11438
11443		11443	11443	11443
		11402	11404	1140A
			/1140E	=1140E
			11438	11438
			11444	11443
				11402
के	-	के	के	के = के
kē	-	kē - i	kē - u	kē - e
1140E		1140E	1140E	1140E
1143E		1143E	1143E	1143E
11443		11443	11443	11443
		11402	11404	1140A
				=1140E
				1143E
				11443
				1142B
				11442