L2/21-130

TO: **UTC**

FROM: **Deborah Anderson, Ken Whistler, Roozbeh Pournader, and Liang Hai**[1]

SUBJECT: **Recommendations to UTC #168 July 2021 on Script Proposals**

DATE: **July 26, 2021**

The Script Ad Hoc group met on May 21, June 11, and July 16, 2021, in order to review proposals. The following represents feedback on proposals that were available when the group met.

Table of Contents

[1] Also participating were Fred Brennan, Peter Constable, Craig Cornelius, Craig Cummings, Lorna Evans, Manish Goregaokar, Ned Holbrook, Cibu Johny, Frank van de Kasteelen, Jan Kučera, Norbert Lindenberg, Kamal Mansour, Lisa Moore, Lawrence Wolf-Sonkin, and Ben Yang. The text for the comments and recommendations were based on notes taken by Fred Brennan and Debbie Anderson.

I. EUROPE

1 Cyrillic

1a. Cyrillic Phonetic Letters

Document: L2/21-107 Unicode request for Cyrillic phonetic letters – Miller

Comments: We reviewed this request for 59 superscript and subscript Cyrillic phonetic letters. These characters are used in Cyrillic-based phonetic transcriptions, in a manner similar to IPA. They can be found in academic works as well as in standard dictionaries.

Character names have been changed based on feedback from the May Script Ad Hoc and appear acceptable. Evidence for the various characters is provided.

Recommendation: We recommend that the UTC make the following disposition:
***SAH-UTC168-R1*:** Accepts the 59 characters from U+1E030..U+1E06A in the new Cyrillic Extended-D block (U+1E030..U+1E08F) for encoding in a future version of the standard, with glyphs and properties as documented in L2/21-107 (Reference: Section 1a of L2/21-130).

We also recommend the UTC make the following dispositions:
Action Item for Ken Whistler: Update the Pipeline (Reference: L2/21-107 and Section 1a of L2/21-130)
Action Item for Debbie Anderson and Kirk Miller: Provide a font to Michel Suignard (Reference: L2/21-107 and Section 1a of L2/21-130)
Action Item for Debbie Anderson: Confirm the Roadmap reflects the new Cyrillic Extended-D block (U+1E030..U+1E08F). This will require Chola and Chalukya be moved. (Reference: L2/21-107 and Section 1a of L2/21-130)

1b. Addendum to L2/21-107 Cyrillic modifier letters

Document: L2/21-142 Addendum to L2/21-107, Cyrillic modifier letters – Miller

Comments: We reviewed this proposal to add two characters to the set listed in L2/21-107, discussed above.

The first modifier character, MODIFIER LETTER CYRILLIC SMALL ES WITH DESCENDER, was mentioned on page 5 of L2/21-107 as being one of a set that could be handled as a base letter and diacritic. However, as noted by the Public Feedback from David Corbett (Section X, below), the character U+04AB CYRILLIC SMALL LETTER ES WITH DESCENDER is not decomposed, so the modifier should also be handled as a single atomic character.

The second character, MODIFIER LETTER CYRILLIC SMALL YERU WITH BACK YER, was a request from Sebastian Kempgen from the Commission for Computer Supported Processing of Medieval Slavonic Manuscripts and Early Printed Books. The letter is an old-style YERU treated as a distinct letter, especially in older documents.

Evidence is provided and the code points are acceptable. The proposed location is immediately after the set of 59 characters in L2/21-107, in a new Cyrillic Extended-D block (U+1E030..U+1E08F).

Recommendation: We recommend that the UTC make the following disposition:
***SAH-UTC168-R2*:** Accepts U+1E06B MODIFIER LETTER CYRILLIC SMALL ES WITH DESCENDER and U+1E06C MODIFIER LETTER CYRILLIC SMALL YERU WITH BACK YER for encoding in a future version of the standard, with glyphs and properties as documented in L2/21-142 (Reference: Section 1b of L2/21-130).

We also recommend the UTC make the following dispositions:
Action Item for Ken Whistler: Update the Pipeline (Reference: L2/21-142 and Section 1b of L2/21-130)
Action Item for Debbie Anderson and Kirk Miller: Provide a font to Michel Suignard (Reference: L2/21-142 and Section 1b of L2/21-130)

2 Old Hungarian

Document: L2/21-115 Proposal for Modifying the Old Hungarian Block – FEHÉR and KATONA

Comments: We reviewed this proposal requesting modifications to Old Hungarian.

This document proposes a change to the name of the block and to character names. Due to the Unicode character stability policy, the name of the script and character names cannot be changed, once they have been encoded. (Old Hungarian was published in Unicode 8.0, June 2015.)

Suggested glyph changes (with adequate rationale) can be submitted for consideration. Similarly, new characters can be proposed, with evidence. We recommend this proposal be forwarded to the original Old Hungarian script proposal authors, Michael Everson and Andre Szabolcs Szelp, for their review.

It was noted that the comment, "Furthermore, there are manufacturers (e.g. Apple) which do not support ligatures, thus the use of ZWJ (Zero Width Joiner) should be avoided," is not accurate.

Recommendation: We recommend that the UTC make the following disposition:
Action Item for Debbie Anderson: Forward the proposal L2/21-115 to Michael Everson and Andre Szabolcs Szelp. (Reference: Section 2 of L2/21-130)
Action Item for Rick McGowan: Forward the comments in Section 2 of L2/21-130 to the proposal authors.

---

3 Sidetic

Document: L2/21-111 Preliminary Proposal to encode the Sidetic script – Pandey

Comments: We reviewed this preliminary proposal for the Sidetic script, a right-to-left alphabetic historic script used in Anatolia. The script was listed in UTR #3 under the heading "Scripts Not Considered for Encoding" because information was lacking. The script was added to the Roadmap based

on Pandey's L2/19-106 "Introducing the Sidetic Script". While not fully deciphered, the script is used in scholarly publications. There is scholarly consensus on the values of 26 of the 29 letters.

The following comments are noted for further Script Ad Hoc consideration or as feedback for the author:

- Issues that need scholarly input are contained in section 4 (on page 4).
- Some glyph variants are significant (i.e., #12) and could be candidates for separate encoding.
- The Wikipedia article on Sidetic is incorrect in stating that word borders are marked with a vertical stroke.
- Bidi value was confirmed as "R."
- The two ligatures (with possible names SIDETIC LETTER LIGATURE-1 and SIDETIC LETTER LIGATURE-2) are still uncertain. The marks appeared on coins, but the components are not clear (i.e., a ligature of which two elements?). Since ligature formation is not productive, atomic encoding appears warranted.
- The code chart shows three columns but the Roadmap only has two. Explain why three columns are needed.

(The comments have been forwarded to the proposal author.)

Recommendations: We recommend that the UTC make the following disposition:
Notes this document but takes no further action.

II. AMERICAS

4 Unified Canadian Aboriginal Syllabics

Document: L2/21-141 Proposed changes to the representative glyphs of the Unified Canadian Aboriginal Syllabics code charts – Kevin King

Comments: We reviewed this proposal, which has been reviewed by the Script Ad Hoc several times.

This version has incorporated changes (listed on page 1) based on feedback from the Script Ad Hoc. The author modified earlier wording about the impact of the changes on other Dene orthographies and made changes to proposed wording for the names list, identifying the language whose orthography uses the letter(s).

Colors are used in the charts to show different kinds of glyph change:

• Aqua: changes to vertical positioning of the glyph
• Pink: changes to glyph shape or proportions
• Yellow: changes to glyph orientation (rotation)

The third set (yellow—U+1628, U+163B, U+18DB) should be called out in an erratum notice.

Section 1.1 includes text with a figure showing the appropriate vertical positioning of the Carrier final characters. We recommend this text be included in the Core Spec for section 20.2 of the Core Spec.

The document includes extensive examples and a letter of support from a member of the Carrier user community. In addition, it traces some of the history of the encoding of these characters in Unicode, which is useful.

Recommendations: The Script Ad Hoc recommends to the UTC the following disposition:
SAH-UTC168-R3*:* Accepts 186 changes for glyphs in the Unified Canadian Aboriginal Syllabics and Unified Canadian Aboriginal Syllabics Extended blocks as shown in L2/21-141 for correction in a future version of the standard. (Reference: Section 4 of L2/21-130)

We also recommend the following dispositions:
Action Item for the Editorial Committee: Post an erratum for U+1628, U+163B and U+18DB, which show significant character orientation changes from the current glyphs, for posting when the glyph changes are published. (Reference: L2/21-141 and Section 4 of L2/21-130)

Action Item for the Editorial Committee: Include the figure and accompanying text from Section 1.1 of L2/21-141, showing the appropriate vertical positioning of the Carrier final characters for section 20.2 of the Core Spec, when the glyph changes are published. (Reference: L2/21-141 and Section 4 of L2/21-130)

III. AFRICA

5 Egyptian Hieroglyphs

Document: L2/21-108 Consideration for the encoding of an extended Egyptian Hieroglyphs repertoire – Suignard

Comments: We briefly reviewed this FYI document, which is a snapshot of the ongoing work by Michel Suignard on extended Egyptian Hieroglyphs as of June 14, 2021.

Two files are appended to the PDF:

(1) an exploratory set of 3,090 glyphs deemed to have the highest priority, based on number of attestations, described on page 8
(2) a spreadsheet with 10,350 glyphs, from which characters can be selected for encoding.

A proposal for Egyptian hieroglyph format control characters is needed in order to decide on the repertoire; the format controls proposal is expected at the October UTC meeting (UTC #169).

Recommendation: We recommend the UTC make the following disposition:
Notes this document but takes no further action.

IV. MIDDLE EAST

6 Arabic

6a. Quranic Marks used in Turkey

Document: L2/21-133 Proposal to encode Quranic marks used in Turkey – Lateef Sagar Shaikh

Comments: We reviewed this proposal for three symbols that appear in a Quran published in Turkey. The proposal provides evidence from a single source, a Quran formally published by the government of Turkey. In the past, UTC has accepted such evidence for Quranic Arabic, see for example characters proposed in L2/06-358R based on a Quran published by the Iranian government, or various characters in L2/19-306 based on Qurans published by the government of Tunisia, Libya, Morocco, or Saudi Arabia. We consider attested usage of the proposed characters in governmental publications of religious materials provides sufficient evidence of need for representation in public interchange.

The names appear to be acceptable. (Members of the group checked with two Arabic speakers and confirmed the transcription of the second name, QASR, is correct.)

Recommendation: We recommend that the UTC make the following disposition:
***SAH-UTC168-R4*:** Accepts
U+0895 ARABIC SMALL LOW WORD SAKTA
U+0896 ARABIC SMALL LOW WORD QASR
U+0897 ARABIC SMALL LOW WORD MADDA
for encoding in a future version of the standard, with glyphs and properties as documented in L2/21-133 (Reference: L2/21-133 and Section 6a of L2/21-130)

We also recommend the UTC make the following dispositions:
Action Item for Ken Whistler: Update the Pipeline (Reference: L2/21-133 and Section 6a of L2/21-130)
Action Item for Debbie Anderson: Provide a font to Michel Suignard (Reference: L2/21-133 and Section 6a of L2/21-130)

V. SOUTH AND CENTRAL ASIA

7 Devanagari

7a. Avagraha followed by Anusvara or Visarga

Document: L2/21-113 On the sequence of Avagraha followed by Anusvara or Visarga – Srinidhi and Sridatta

Comments: We reviewed this document which presents evidence of *avagraha* followed by *anusvara* in Sanskrit texts, with examples in Kannada (pp. 1-2) and Devanagari (p. 3). The document also suggests the sequence of *avagraha* and *visarga* could occur in grammatically valid Sanskrit sentences, but no examples from printed books or texts are provided. The author requests the sequences be considered valid in the Core Spec and by rendering engines. (The general category value of Devanagari and Kannada *avagraha* is Lo, Devanagari *visarga* is Mc, Devanagari *anusvara* is Mn and Kannada *visarga* and *anusvara* are Mc.)

The Script Ad Hoc had previously requested (L2/20-105, p. 28) evidence as to whether Devanagari *avagraha* can be followed by either *anusvara* or *candrabindu*. Srinidhi and Sridatta provided evidence for *avagraha* followed by *anusvara* but not for *avagraha* followed by candrabindu.

The following discussion points were raised:

- Andrew Glass reported that *avagraha* + *anusvara* is a Microsoft shaping engine restriction; Harfbuzz renders it correctly. A bug has been filed for MS shaping engine. The author of L2/21-113 has been notified about the bug.
- Orthographies have different traditions, and it is not possible to know all the possible combinations of consonant-like base characters that take marks. For some implementers, it might be prudent to allow combinations and only impose restrictions until the issue is well understood.
- Is there any text in the Core Spec that contradicts *avagraha* followed by *anusvara* or *visarga*?
- Devanagari cluster validation was discussed by Norbert Lindenberg in L2/21-112.

Recommendation: We recommend the UTC make the following disposition:
Action Item for Debbie Anderson: Relay comments in Section 7a of L2/21-130 to the authors of L2/21-113.

7b. Devanagari 'bhale mīṇḍu'

Document: L2/21-102 Proposal for representing the Devanagari 'bhale mīṇḍu' in Unicode

Comments: We reviewed this revised proposal for a class of auspicious signs.

- Earlier comments from the Script Ad Hoc have been accommodated, and we now recommend approval of the characters.
- The characters are located in the Devanagari Extended-A block (U+11B00..U+11B4F). However, Anshuman Pandey reports that additional Devanagari characters will be requested in the near future and the block needs to be extended by at least one additional column. As a result, we recommend the Roadmap Committee extend the Devanagari Extended-A block to end at U+11B5F, and move Sharada Extensions over one column, from U+11B50..U+11B6F to U+11B60..U+11B7F.
- It was agreed that one character name should be changed, because "MIRRORED" is not used in character names: U+11B08 DEVANAGARI SIGN MIRRORED NINE-LIKE BHALE should be U+11B08 DEVANAGARI SIGN REVERSED NINE-LIKE BHALE.

Recommendation: We recommend that the UTC make the following disposition:
***SAH-UTC168-R5*:** Accepts 10 Devanagari 'bhale mīṇḍu' characters from U+11B00..U+11B09 for encoding in a future version of the standard, with glyphs and properties as documented in L2/21-102, with one name change: U+11B08 DEVANAGARI SIGN MIRRORED NINE-LIKE BHALE should be U+11B08 DEVANAGARI SIGN REVERSED NINE-LIKE BHALE. (Reference: Section 7b of L2/21-130)

We also recommend the UTC make the following dispositions:
Action Item for Ken Whistler: Update the Pipeline (Reference: L2/21-102 and Section 7b of L2/21-130)
Action Item for Debbie Anderson and Anshuman Pandey: Provide a font to Michel Suignard (Reference: L2/21-102 and Section 7b of L2/21-130)

Action Item for Debbie Anderson: Confirm changes are made to the Roadmap (extend the Devanagari Extended-A block to end at U+11B5F, and move Sharada Extensions over one column, from U+11B50..U+11B6F to U+11B60..U+11B7F). (Reference: Section 7b of L2/21-130)

---

8 Gurung Khema

Document: L2/21-145 Proposal to Encode the Gurung Khema Script in the Universal Character Set – Biwajit Mandal

Note: The Script Ad Hoc saw an earlier version of the posted proposal and has not yet reviewed this version. The comments below were made in response to the version seen by the group.

Comments*:* We reviewed this proposal, which is one of two scripts used to write the Gurung language of Nepal and Sikkim, India. In 2000, six characters were introduced to the script (four medials, a "virama" and an anusvara) and in 2019, four new vowel letters and four new vowel signs were introduced to represent long vowels, and many characters have modified glyphs.

The following were noted during discussion:

- Concerns were raised about the official status of the script and whether it is stable enough for encoding in the Unicode Standard.
- The name "virama" may be misleading. Its combining class value 9 needs to be reconsidered. (In the standard, "virama" is associated with special behavior, specifically conjoining behavior, but the proposed virama doesn't conjoin. Note that in the Indic_Syllabic_Category file, "Virama" is a contextually conjoining graphic mark.)
- 3.9 Basis of Character Shapes: A question was raised about the old forms of the letters: should they be separately encoded or handled by a font? Cf. the circle below the letter in the last *akshara* in the word for "excuse" on page 49. How much material is written with the old forms?
- 3.9: The proposal mentions that Sikkim Herald Gurung Edition uses its own font, but is not readable to most users (page 3). Is the newspaper still publishing in this font? It was noted that the new vowel letters could cause confusion for some, since the new letter for VOWEL II is the same as the letter VOWEL U in the Sikkim Herald.
- Syllabic Structure: Can the proposal author provide any examples where the graphical and phonetic order contradict one another (i.e., left-side medial YA or VA is pronounced after a bottom-side sign virama)?

The above comments have been forwarded to the proposal author.

Recommendation: We recommend the UTC make the following disposition:
Notes this document but takes no further action.

---

9 Kannada

Document: L2/21-114 Proposal to encode KANNADA SIGN COMBINING ANUSVARA ABOVE RIGHT – Shashank Shenoy Basty

Comments: We reviewed this document which proposes one character for Kannada. The author made the changes based on feedback from the October 2020 Script Ad Hoc recommendations (p. 11 of SAH Recs L2/20-250).

The proposal is still missing Indic positional and syllabic categories, but the group agreed they should be: Indic_Positional_Category=Right and Indic_Syllabic_Category=Bindu. Also, the Bidi property should be L, not NSM. Hence the UnicodeData.txt properties should be:
0CF3; KANNADA SIGN COMBINING ANUSVARA ABOVE RIGHT;Mc;0;L;;;;;N;;;;

Recommendation: We recommend that the UTC make the following disposition:
***SAH-UTC168-R6*:** Accepts U+0CF3 KANNADA SIGN COMBINING ANUSVARA ABOVE RIGHT for encoding in a future version of the standard, with glyph as in L2/21-114 and properties in Section 9 of L2/21-130.

We also recommend the UTC make the following disposition:
Action Item for Ken Whistler: Update the Pipeline (Reference: L2/21-114 and Section 9 of L2/21-130).

---

10 Khojki

10a. Vocalic R

Document: L2/21-110 Proposal to encode the Khojki vowel sign VOCALIC R in Unicode – Pandey

Comments: The character KHOJKI VOWEL SIGN R was described in the original Khojki proposal (L2/11-021) and was separately proposed by Srinidhi and Sridatta in L2/17-307. The SAH discussed L2/17-307 in its April-May 2018 recommendations (p. 24 of L2/18-168), asking (a) whether the evidence was based on one source and (b) how the independent vowel is represented. Srinidhi and Sridatta replied that more research was needed.

Comments from discussion:

- The character may not have been part of the original repertoire, or may have been re-introduced later. Inclusion of the character now will support ongoing digitization work at the Institute of Ismaili Studies in London.
- The author clarified that there no independent letter for vocalic r (or at least not yet found).
- A general rule-of-thumb is for proposal authors to provide examples from two different publishers or authors. If that is not possible, demonstrate the importance of the character. The author mentioned that the source document is a liturgical text, at the level of a Catholic encyclical or hymn. In an earlier version of the proposal examples only came from one source. However, in this version, the author has added examples from different sources and different publishers.

Based on the information provided in the proposal, we recommend the character be approved.

Recommendation: We recommend that the UTC make the following disposition:
***SAH-UTC168-R7*:** Accepts U+11241 KHOJKI VOWEL SIGN VOCALIC R for encoding in a future version of the standard, with glyph and properties as documented in L2/21-110. (Reference: Section 10a of L2/21-130)

We also recommend the UTC make the following dispositions:
Action Item for Ken Whistler: Update the Pipeline (Reference: L2/21-110 and Section 10a of L2/21-130)
Action Item for Debbie Anderson and Anshuman Pandey: Provide a font to Michel Suignard (Reference: L2/21-110 and Section 10a of L2/21-130)

10b. QA

Document: L2/21-103 Proposal to encode the Khojki letter Qa in Unicode—Pandey

Comments*:* We reviewed this proposal to encode one Khojki letter, letter QA. This character was mentioned in the Khojki script proposal (L2/11-021), but was not included in the repertoire, due to lack of sufficient evidence.

This character is used to represent Arabic /q/ in Khojki. (The Arabic sound /q/ can also be represented by the use of KA or KA with NUKTA.) Contrastive usage of the character is shown in section 3.

The proposal incorporates changes requested by the Script Ad Hoc, and hence we recommend U+1123F KHOJKI LETTER QA for encoding.

Recommendation: We recommend that the UTC make the following disposition:
***SAH-UTC168-R8*:** Accepts U+1123F KHOJKI LETTER QA for encoding in a future version of the standard, with glyph and properties as documented in L2/21-103. (Reference: Section 10b of L2/21-130)

We also recommend the UTC make the following dispositions:
Action Item for Ken Whistler: Update the Pipeline (Reference: L2/21-103 and Section 10b of L2/21-130)
Action Item for Debbie Anderson and Anshuman Pandey: Provide a font to Michel Suignard (Reference: L2/21-103 and Section 10b of L2/21-130)

10c. Short I

Document: L2/21-104 Proposal to encode the Khojki letter SHORT I in Unicode – Pandey

Comments: We reviewed this proposal to encode one Khojki letter for SHORT I.

The repertoire for the Khojki script does not have distinctive letters for independent vowels ī and i, although it does have distinctive signs for long and short dependent i. In order to distinguish long and short independent i, printers used U+11202 𑈂 KHOJKI LETTER I for long i, and 𑈀 for short i. This

document proposes the new character. Because the name KHOJKI LETTER I is already used, the proposed new character is named “KHOJKI LETTER SHORT I”.

The evidence demonstrates the character is well-justified. The proposal incorporates changes requested by the Script Ad Hoc, so we recommend approval of U+11240 KHOJKI LETTER SHORT I.

Recommendation: We recommend that the UTC make the following disposition:
***SAH-UTC168-R9*:** Accepts U+ 11240 KHOJKI LETTER SHORT I for encoding in a future version of the standard, with glyph and properties as documented in L2/21-104. (Reference: Section 10c of L2/21-130)

We also recommend the UTC make the following dispositions:
Action Item for Ken Whistler: Update the Pipeline (Reference: L2/21-104 and Section 10c of L2/21-130)
Action Item for Debbie Anderson and Anshuman Pandey: Provide a font to Michel Suignard (Reference: L2/21-104 and Section 10c of L2/21-130)

---

11 “Kirat Rai”

Document: L2/21-132 Advancing the Kirat Rai script – Evans and Smith

Comments: We reviewed this request to progress work on “Kirat Rai,” a script that was first proposed in 2011 (L2/11-145) by Anshuman Pandey. A second proposal for Khambu Rai was also written by Pandey (L2/11-105). Khambu Rai appears only to have one attestation, and seems to be the same as Kirat Rai or a version of it.

The authors will work with Biswajit Mandal on a proposal.

The authors described the use of the terms (Rai, Kirat, Khambu) and suggest the most appropriate name would be Rai, with a second choice being Khambu Rai and the last choice would be Kirat Rai. During discussion, Anshuman Pandey recommended against Rai, an ethnonym in South Asia and Nepal, and instead recommended a combination of two names (such as the script creator/community and “Rai”) Discussion of the name should be included in the proposal.

We agreed that work on a proposal for “Kirat Rai” could progress. The current “Kirat Rai” location on the Roadmap (U+16D40..U+16D7F) is acceptable. Based on discussion, we agreed “Khambu Rai” should be removed from the Roadmap from U+11BC0..U+11BFF, and to move Nag Mundari there.

Recommendation: We recommend the UTC make the following disposition:
Action Item for Debbie Anderson: Confirm changes are made to the Roadmap (remove “Khambu Rai” U+11BC0..U+11BFF and move Nag Mundari to that location). (Reference: Section 11 of L2/21-130)

---

12 Syloti Nagri

12a. Candrabindu

Document: L2/21-138 Proposal to Encode Syloti Nagri Sign Candrabindu -- Biwajit Mandal

Comments: We reviewed this document to encode one Syloti Nagri character.

The following summarizes the comments:

- The only evidence is contained in figures 1-3. Evidence showing more widespread usage outside of primers is needed, particularly from examples showing a longer section of text. (The author may wish to note the comments in L2/02-388, which mentions a purported candrabindu in primers, but the authors of L2/02-388 had not seen any examples.)
- Provide Indic syllabic and positional categories.

Recommendations: We recommend the UTC make the following disposition:
Action Item for Debbie Anderson: Relay comments in Section 12a of L2/21-130 to the author of L2/21-138.

12b. Numerals

Document: L2/21-140 Proposal to Encode the Sylheti Nagri Numerals in the UCS -- Harris Mowbray

Comments: We reviewed this proposal to encode ten Syloti Nagri numbers.

The following comments were made:

- Evidence is based largely on social media. Provide additional examples (with sourced citations), preferably from printed materials.
- Section II.1.6 of L2/02-388, mentions use of a mix of Bengali, Arabic-Indic, Arabic [Latin] digits. Are the proposed characters similarly an ad hoc collection of digits or is there evidence of an emerging conventional set of digits?
- Some of the examples do not match the proposed glyphs, such as loop for '3'.
- Number the pages, tables, and images.
- Provide bibliographic information on the book example. Clarify where the example on page 10 came from.
- Provide a better font, preferably one matching the glyphs in the current Syloti Nagri chart.

Recommendations: We recommend the UTC make the following disposition:
Action Item for Debbie Anderson: Relay comments in Section 12b of L2/21-130 to the author of L2/21-140.

13 Tulu/Tigalari

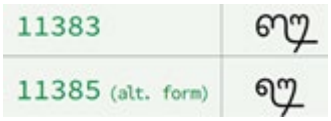
13a. Tulu-Tigalari Proposal

Document: L2/21-146 Updated proposal to encode Tulu-Tigalari script in Unicode – Vaishnavi Murthy and Vinodh Rajan

Comments: We reviewed this revised proposal. A list of changes made from the earlier version is on page 3.

The following summarizes points raised during discussion:

- The proposal is a lengthy document. To aid reviewers, include a Table of Contents, and, if possible, use PDF bookmarks.
- Character names in the properties list should be all in uppercase (i.e, CONJOINER).
- Most of the discussion focused on the encoding model for handling the vowel letters that could be conceived of as being two parts.
    - The authors generally prefer an atomic approach to encoding these vowel letters. However, reliance on a Do Not Use table (p. 10) has been problematic in the past: the list is not strictly enforced and users tend to type combinations to get the correct looking letter. Note that since this is a classical script, the decision may need to consider more heavily the impact on users, who may prefer the atomic encoding of vowel letters in other closely related scripts (Malayalam, Kannada, …) that also conform to the Sanskrit analysis.
    - If the atomic encoding of vowels is adopted with a Do Not Use table, the entry for U+1138E AI should be changed, so in the Do Not Use sequence the base character VOWEL LETTER EE precedes the combining sign U+113C2 VOWEL SIGN EE.
    - In section 10.2 Character properties, U+11391 LETTER AU has incorrect code points in the decomposition (if the decision is to provide decompositions): U+113C3 is an unassigned code point. The code points should be: 11390 113C9. Also check carefully the code points for other decompositions: there are errors also in the decompositions for U+113C7 and U+113C8.
    - Are the first two rows (for U+11383 LETTER II and U+11385 [alternate] LETTER UU) in the Do Not Use chart decomposable or not? Segmenting these two graphically is difficult.



    - Provide more information on the alternate glyph shapes (p. 33). This can inform the choice of atomic character (with decomposition) or a sequence. If decomposition or a sequence is used, the graphical elements need to be separately encoded.
    - If the model for handling independent vowels is to avoid atomic encoding, and instead rely on use of encoded graphical pieces, then we may end up with the situation as in Lithuanian, in which users are searching for the combination in the chart. Named sequences may not necessarily be the answer. Another approach could be atomic encoding with decompositions, thereby avoided the reliance on a Do Not Use table.

    - It was noted that the most common form of LETTER UU uses a spiral overlapping the tail of LETTER U; the spiral component does not occur by itself, so the atomic encoding of this independent vowel could be warranted.
    - In sum, there was no consensus on the best approach amongst group members. There was support for canonical decompositions of the two-part dependent vowels, but the group was divided on how to handle independent vowels, in part because more information was needed, especially on the alternate glyphs.
- One member discouraged discussion of stylistic sets, as it could be misleading for implementers.

We recommend the authors revise their proposal, and expand the section 6.1 Alternate glyph shapes (page 33), specifying which alternates are common and which are rare, where each is used, and are they the only shapes seen, etc.

The comments above have already been sent to the authors.

Recommendation: We recommend the UTC make the following disposition:
Notes this document but takes no further action.

13b. Tulu-Tigalari replies to Script Ad Hoc etc.

Document: L2/21-147 Replies to Recommendations to UTC #167 April 2021 on Script Proposals (L2/21-073) & Norbert Lindenberg's email with comments 10 April 2021 – Vaishnavi Murthy and Vinodh Rajan

Comments: We reviewed this document that responded to earlier comments from recommendations from the Script Ad Hoc (L2/21-073) and email comments from Norbert Lindenberg.

While most of the comments from Norbert Lindenberg were accommodated, the suggestion of following the Kawi model, that is, to avoid atomic encoding if the letters can be represented by character sequences, was not adopted by the proposal authors, because the users are used to the Indic model. As Lindenberg notes, use of sequences avoids Do Not Use tables and canonical decomposition. (See further discussion above.)

Recommendation: We recommend the UTC make the following disposition:
Notes this document but takes no further action.

---

14 Vatteluttu

Documents: L2/21-139 Virama as a Solid Dot and Atomic Code Points for Short E and Short O in Vatteluttu – Santhalingam and Ganesan
Related documents:
L2/16-068 Preliminary proposal to encode Vatteluttu in Unicode – Pandey

L2/21-052 Request to Change Pulli Representation in the Proposed Vatteluttu Encoding – Cibu Johny
L2/21-073 Recommendations to UTC #167 April 2021 on Script Proposals (especially Section 10)

Comments: We reviewed this document, which commented on L2/21-052 "Request to Change Pulli Representation in the Proposed Vatteluttu Encoding."

Background: A preliminary proposed for Vatteluttu was submitted by Anshuman Pandey in 2016 (L2/16-068). Later that year, TVA/ T. Udhayachandran (L2/16-118) submitted a response, requesting the proposal be put on hold while TVA prepares its own proposal "in about six months". However, no follow-up proposal has yet been submitted. The document L2/21-052 by Cibu requested a change from the encoding model in Pandey's proposal, which had proposed atomically encoding the vowels with the virama (*pulli*), but used a separate virama as a vowel-killer. Cibu's proposal advocated a model where the virama is a separate combining mark for use with a vowel or consonant. The Script Recommendations felt the changes in Cibu's document made sense, but requested seeing the changes within the entire proposal.

The following highlights the comments made during discussion:

- The document from Santhalingam and Ganesan argues in section 1 for the solid dot as the most common glyph for the virama, providing examples. Cibu was ok with the glyph change, but questioned some of the examples.
- The document from Santhalingam and Ganesan further recommends the atomic encoding of the independent short vowel letters e and o and the corresponding vowel signs, retention of the naming convention for vowels as in Pandey's proposal, and a code chart pattern for Vatteluttu that follows other southern Indic scripts.
- In order to make a decision, a new proposal is needed to provide the full context. A joint proposal from the various parties would be welcome.
- Cibu reports he will prepare a document asking for guidelines on how to define the boundary between Tamil-Brahmi and Vatteluttu, with examples showing the issues.

Recommendation: We recommend the UTC make the following disposition:
Notes this document but takes no further action.

---

15 Vedic

Document: L2/21-054 Interaction of Vedic svara markers with post-base spacing marks – Sharma

Comments: We reviewed this document that identified a problem with the current encoding of Vedic text, specifically nonspacing svara marks (tone marks) and post-base markers (primarily visarga and anusvara in Bengali and South Indian scripts).

The problem is that some users expect the sequence to be syllable + nonspacing svara mark(s) + the spacing mark. However, current text shaping engines mark this sequence as illegal and a dotted circle appears before the spacing mark; the permitted order is syllable + spacing mark + svara.

The author states that Vedic support is still in its infancy and requests *TUS* recommend svara marks be allowed before post-base visarga and anusvara.

The following points were raised:

- The shaping engines agree on the behavior of svara markers and post-base spacing marks (i.e., syllable + spacing mark + svara) and they all follow the Core Spec (13.0 p.460: R10) and follow the documentation of the OpenType Devanagari shaping engine.
- Modern Input methods, such as Keyman, can re-order what users input to correct logical order.
- It is not new for Brahmic scripts to have the logical order differ from the visual order: left-side vowels are encoded post-base for most Brahmic scripts.
- Making the change as proposed by Shriramana would mean the strings encoded in the old and new orders would not be canonically equivalent.
- The Indic syllable structure needs to be investigated and clarified. The current documentation in the Unicode Standard and in OpenType shaping engine documentation is imprecise, inconsistent, and, for Vedic characters, incompatible with normalization. Members are invited to read N. Lindenberg's Devanagari cluster validation document L2/21-112.
- If a proposal were to be considered, what would users expect if one of the svaras co-occurs with vowel signs or other signs above?

In sum, there was no consensus to make a change.

The comments above have been sent to the document author.

Recommendation: We recommend the UTC make the following disposition:
Action Item for Liang Hai: Work with Norbert Lindenberg to investigate the Indic syllable structure and how dotted circles are working. (Reference: Section 15 of L2/21-130)

VI. SOUTHEAST ASIA, INDONESIA, AND OCEANIA

16 Lao

16a. Proposal for Lao Yamakkan

Document: L2/21-093 Proposal to Encode Lao Yamakkan – Rajan and Bhikkhu

Comments: The Script Ad Hoc had recommended approval of Lao Yamakkan in its April 2021 set of recommendations (page 12-13 of L2/21-073). However, during the April UTC a question was raised about the character's Indic_Syllabic_Category, so the proposal was returned to the Script Ad Hoc.

The proposed Indic_Syllabic_Category for the character is Syllable_Modifier. After discussion, the group agreed Syllable_Modifier was probably the best fit, but Roozbeh Pournader will check the Indic_Syllabic_Category before it appears in the release (15.0 or later). It was noted that the assignment of Indic_Syllabic_Category is not immutable; the value can be changed even after the character has been published.

Recommendation: We recommend the UTC make the following disposition:
***SAH-UTC168-R10*:** Accepts U+0ECE LAO YAMAKKAN for encoding in a future version of the standard, with glyphs and properties as documented in L2/21-093. (Reference: Section 16a of L2/21-130)

We also recommend that the UTC make the following dispositions:
Action Item for Ken Whistler: Update the Pipeline. (Reference: L2/21-093 and Section 16a of L2/21-130)
Action Item for Debbie Anderson and Vinodh Rajan: Provide Michel Suignard with a font. (Reference: L2/21-093 and Section 16a of L2/21-130)

16b. Comments on Lao Yamakkan

Document*:* L2/21-143 Reply to Comments on Lao Yamakkan – Vinodh Rajan

Comments: We reviewed this document, which provides comments made over the Unicore email list by Richard Wordingham on the disunification of LAO YAMAKKAN and LAO CANCELLATION MARK and the later email exchange between Wordingham and Vinodh Rajan, who proposed LAO YAMAKKAN (L2/21-093).

The example on page two shows the two marks appearing contrastively in the same document. Based on this example, we agree there is no reason to consider unifying LAO YAMAKKAN with LAO CANCELLATION MARK.

Recommendation: We recommend the UTC make the following disposition:
Notes this document but takes no further action.

---

17 Surat Ulu (Rejang)

Document: L2/21-116 Towards an Encoding for Surat Ulu – M. Mahali Syarifuddin

Comments: We reviewed this document. According to the author, Surat Ulu is the name of the Brahmi-based script that is found in manuscripts of southwestern Sumatra. It was used to write Malay or its dialects, including Rejang, Lembak, Serawai, and Pasemah. Because Unicode has encoded the Rejang script separately, the question arises how to handle other variants of Surat Ulu.

The following summarizes the discussion points:

- As noted by the author in section 3 and 4, additional research is needed. More information will make it possible for decisions to be made on which characters to encode and how best to handle the variants.

  For example, two KA glyphs are shown for Serawai on page 4, both of which vary from the Rejang character KA. Are these two in free variation or does each have specific meaning? Or are they used by groups in different regions? Is the loop just a cursive form of the other shape? If so, it may not need to be encoded, since not every shape in a document needs to be separately encoded.

- Page 6 mentions that the authors of the 2014 source publication asked Serawai informants about the current use of the script in their area. How active is the current use in the different regional areas and how does modern use vary from the attestations in manuscripts? How many people are using the script today? Is it being taught and used in publications today? Also, how do the separate groups see themselves in relation to other groups?
- When revising the proposal, include some text introducing the diacritics topic in 2.2. before figure 3.
- In footnote 4, the author notes the glyph for U+A939 REJANG LETTER CA in *TUS* is angular instead of circular. The author is invited to request the glyph be changed, providing evidence.
- Under the first bullet point on page 10, the author states that "Rejang" is not a suitable name, since it is a subset of Sural Ulu. Due to the Unicode character stability policy, once a script is encoded, its name cannot be changed. One approach would be to fill the remaining open spots in the Rejang block with new characters and add an additional "Rejang" block, but include notes in the names list and in the Core Spec documentation about the relation of the new characters to Rejang. This approach would be better than having separate blocks for Lembak etc., in our view, but a final decision on the best approach ultimately requires having more information on the letters.
- Are all the characters in Jaspan's book *Redjang Ka-Ga-Nga Texts* included in the Rejang block?
- It may be worth considering a more visual encoding order, and not necessarily following the model prescribed (but rarely implemented) for Batak.
- It would be useful to get expert input on the proposal, including Anshuman Pandey for his comments, since he had written a proposal for Lampung (L2/16-073), a script that should be investigated for possible unification with Surat Ulu.

Recommendation: We recommend the UTC make the following dispositions:
Action Item for Norbert Lindenberg: Invite Indonesian script experts to review the Surat Ulu proposal (Reference: Section 17 of L2/21-130 and L2/21-116).
Action Item for Debbie Anderson: Invite Anshu Pandey to review the Surat Ulu proposal (Reference: Section 17 of L2/21-130 and L2/21-116).

---

18 Tagalog and Hanunoo

Document: L2/21-117 Pamudpod properties – Pournader

Comments: We reviewed this short document about the character properties for the Tagalog and Hanunoo *pamudpod* characters. Review of the properties by Roozbeh Pournader uncovered discrepancies between the two.

After discussion, the group agreed that the general category for both characters should be Mc, the bidi property L, and the Indic_Positional_Category should be changed to Right. (Note: The document recommends HANUNOO SIGN PAMUDPOD retain its current bidi class, NSM, but the group agreed that L was appropriate.)

The following summarizes the corrected UnicodeData and Indic_Positional_Category entries:

UnicodeData.txt:
1715;TAGALOG SIGN PAMUDPOD;Mc;9;L;;;;;N;;;;;
1734;HANUNOO SIGN PAMUDPOD;Mc;9;L;;;;;N;;;;;

Indic_Positional_Category.txt:
1715 ; Right
1734 ; Right

Recommendation: We recommend that the UTC make the following disposition:

***SAH-UTC168-R11*:** Make the following adjustments to the properties for the following two characters for Unicode 14.0:
For U+1734 HANUNOO SIGN PAMUDPOD:
Change the gc from Mn to Mc
Change the bidi class from NSM to L
Change the Indic_Positional_Category from Bottom to Right.

For U+1715 TAGALOG SIGN PAMUDPOD:
Change the Indic_Positional_Category Bottom_and_Right to Right.

Note: The above changes have already been made in the data files, so no action by the UTC is required.

VII. EAST ASIA

19 Jurchen

Document: L2/21-049 A Supplementary Proposal to Encode the Jurchen Characters in UCS – China NB

Comments: This is FYI to the UTC. The document requests adding 15 “glyphs” to the earlier proposal on Jurchen, based on an inscription discovered in 2015. (Note: The reference to the latest proposal should have been to N3788 [L2/10-101] “Revised Proposal to Encode the Jurchen Characters in UCS” from China NB. L2/21-049 refers to N3628 [L2/09-351] as the last proposal, in error.)

Recommendation: We recommend the UTC make the following disposition:
Notes this document but takes no further action.

VIII. OTHER SCRIPTS

20 Toki Pona

Document: L2/21-137 Proposal to encode Toki Pona – Gabriel Tellez

Comments: We reviewed this document, which is not a full proposal, for a conlang. In our view, it is out of scope and not appropriate for encoding at this time.

Recommendation: We recommend that the UTC make the following disposition
Action Item for Debbie Anderson: Relay comments in Section 20 of L2/21-130 to the author of L2/21-137.

IX. SYMBOLS, PUNCTUATION, AND NOTATIONAL SYSTEMS

21 Nine Pointed Star

Document: L2/21-144 Proposal for New Character - Nine Pointed Star – Adib Behjat, Joop Kiefte

Comments: We reviewed this proposal, which has accommodated comments from an earlier Script Ad Hoc meeting. Earlier proposals seen by the UTC or ESC were L2/17-237 and L2/20-095.

The following captures the comments made during discussion:

- While not all the evidence provided was strong, the preponderance of evidence was enough for the group to recommend encoding a nine-pointed star. Some of the examples were "textish" (such as figure 5, which resemble bullets), but figure 17 was deemed acceptable as evidence.
- The name we recommend is NINE POINTED WHITE STAR to be located at U+1F7D9, with a glyph like that shown on the gravestone in figure 14.
- An annotation should be added, "Baha'i faith"

Recommendation: We recommend that the UTC make the following disposition:
***SAH-UTC168-R12*:** Accepts U+1F7D9 NINE POINTED WHITE STAR for encoding in a future version of the standard, with a glyph based on the gravestone marker in figure 14 and properties as in L2/21-144. (Reference: L2/21-144 and Section 21 of L2/21-130)

We also recommend the UTC make the following dispositions:
Action Item for Ken Whistler: Update the Pipeline (Reference: L2/21-144 and Section 21 of L2/21-130)
Action Item for Fred Brennan: Provide a font to Michel Suignard (Reference: L2/21-144 and Section 21 of L2/21-130)

---

22 Persian Siyaq

Document: L2/21-105 Proposal to encode Persian Siyaq Numbers – Pandey

Comments: The Persian Siyaq proposal has been seen several times by the Script Ad Hoc. The last posted version of the proposal was L2/20-249 (with comments from SAH on page 16 of L2/20-250).

In this proposal, the author focused on plain text representation of the numbers in the system, and considers rendering – including vertical stacking and nesting of the characters – to be outside the scope of the proposal.

The following comments were raised during discussion:

- Provide more information on the placeholder character. Further examples could justify it as a non-combining character.
- Invite others to review the proposal, checking the examples and the sources on pages 27-82. Be sure the proposal is not missing anything.

Recommendation: We recommend that the UTC make the following disposition:
Notes this document but takes no further action.

X. PUBLIC REVIEW FEEDBACK

Document: L2/21-125 Public Review Feedback

Book Pahlavi

Date/Time: Sun Jun 13 19:07:29 CDT 2021
Name: Eduardo Marín Silva
Report Type: Feedback on an Encoding Proposal
Opt Subject: On the proposed Book Pahlavi encoding model

[Three suggestions were made in this feedback for L2/21-090 Advancing the encoding model for Book Pahlavi]

We briefly reviewed these comments. The comments have been forwarded to the Book Pahlavi proposal author, so no UTC action is required.

Cyrillic

Date/Time: Tue Jun 15 13:45:10 CDT 2021
Contact:corbett.dav@northeastern.edu
Name: David Corbett
Report Type: Feedback on an Encoding Proposal Opt Subject: Comment on L2/21-107

L2/21-107 proposes “that spacing superscript й, ў, ҫ, ҙ etc. [...] be typeset with diacritics”. Because U+04AB CYRILLIC SMALL LETTER ES WITH DESCENDER and U+0499 CYRILLIC SMALL LETTER ZE WITH DESCENDER are encoded without decompositions, if modifier letter versions of them are attested, shouldn’t the modifier letter versions be encoded without decompositions too?

The author of L2/21-107 agreed, as do we. MODIFIER LETTER CYRILLIC SMALL ES WITH DESCENDER is now proposed, described above under Section 1b. (Kirk Miller, author of L2/21-107, reports <ҙ> is found in Bashkir, but he has not found any examples, so he did not propose it.)
No action is required by the UTC.

XI. 14.0 ALPHA REVIEW FEEDBACK

Note: Alpha feedback from Patrik Sjöwall on Arabic (dated April 11) was received too late for review at the April Script Ad Hoc meeting, but was taken up at the May meeting and is discussed below. His comment on renaming OLD POLISH O, which was also discussed by Michael Everson in feedback dated April 16, is mentioned in the April Editorial Committee report. L2/21-070. (Michael Everson had also disagreed with Sjöwall on the proposed change to O ROGATE.)

Similarly, feedback from Michael Everson on COMBINING OVERCURL and Latin Extended-D was received after the Script Ad Hoc had met. Recommendations on the feedback are listed below.

Document*:* https://www.unicode.org/review/pri428/ PRI #428 Alpha Feedback

Arabic

Date/Time: Sun Apr 11 02:28:55 CDT 2021
Name: Patrik Sjöwall
Report Type: Public Review Issue
Opt Subject: Unicode 14.0 Alpha review

[1] I found a few issues with some characters for Unicode 14.0 that seem to have gone unnoticed:

0874 ARABIC LETTER ALEF WITH ATTACHED KASRA
0875 ARABIC LETTER ALEF WITH ATTACHED BOTTOM RIGHT KASRA
0879 ARABIC LETTER ALEF WITH ATTACHED ROUNDDOT BELOW
087C ARABIC LETTER ALEF WITH RIGHT MIDDLE STROKE AND DOT ABOVE
087D ARABIC LETTER ALEF WITH ATTACHED BOTTOM RIGHT KASRA AND DOT ABOVE
0880 ARABIC LETTER ALEF WITH ATTACHED BOTTOM RIGHT KASRA AND LEFT RING

These letters reqiure more shaping information. It is not clear how the attached fatha or dot will behave in an obligatory LAM-ALEF ligature.

Under [1], the author requested more shaping information for the six new 14.0 Arabic characters (U+0874, U+0875, U+0879, U+087C, U+087D, and U+0880, proposed in L2/19-306), and specifically asked about how the attached fatha or dot would behave in an obligatory LAM-ALEF ligature.

It is not known if the characters mentioned appear in lam-alef ligatures or not. The proposal (L2/19-306 Arabic additions for Quranic orthographies) lists its references, which could be investigated to see if such combinations exist. If they don't exist, implementations would not be required to support them.

Kamal Mansour raised his concerns about the proposed set of characters in 14.0 (U+0870..U+0884), and questioned the statement in L2/19-306 that when elements touch or intersect, they are encoded as a single precomposed character. Kamal advocated the characters be encoded as a series of independent combining marks, instead of new types of alefs. While it is good to get feedback, the window for requesting changes to the 14.0 repertoire now was deemed too tight and ultimately, unfortunately, too late.

[2] 088E ARABIC VERTICAL TAIL

This character is missing in ArabicShaping-14.0.0.txt, but it always joins with the preceding letter. It should be included in that file, either as Right_Joining or be given a new joining type (since it does not change its shape, only causes the character to its right to join), and with either a joining group of its own or No_Joining_Group.

VERTICAL TAIL has been included in ArabicShaping.txt as Right_Joining in the Unicode version 14.0 data file.

[3] 08FB ARABIC DOUBLE RIGHT ARROWHEAD ABOVE
08FC ARABIC DOUBLE RIGHT ARROWHEAD ABOVE WITH DOT

The comment "also used in Quranic text in African and other orthographies to represent dammatan" should come after 08FB, not 08FC. The "right arrowhead" is an angular-shaped damma, and the "dammatan" is a double damma (not a double damma with dot).

The correction has been made in the data file and is noted in the Editorial Committee report L2/21-070.

Recommended **Action Item** for Rick McGowan: Send comments from Section XI. (Arabic) to Patrik Sjöwall.

Latin

Date/Time: Sun Apr 11 02:28:55 CDT 2021
Name: Patrik Sjöwall
Report Type: Public Review Issue
Opt Subject: Unicode 14.0 Alpha review

The author commented that A7D3 LATIN SMALL LETTER DOUBLE THORN and A7D5 LATIN SMALL LETTER DOUBLE WYNN were added without the corresponding capitals, which he considers inconsistent with past practice. He cites instances of capital letters that were encoded without being needed outside of all-caps.

New character additions for the next version of Unicode should be approved at or before the January UTC meeting. The UTC did not have consensus to encode the uppercase forms of DOUBLE THORN and DOUBLE WYNN at the January 2021 meeting.

On the topic of missing uppercase forms, one way to address the issue is suggested in the Script Ad Hoc's comment on page 13 of L2/21-016:

> Because the argument for case pairs regularly arises, it might be helpful to have a document discussing the formal criteria for the addition of case pairs, which can guide making decisions when no evidence is presented. It was also noted that newly created uppercase forms for letters only attested with lowercases in historical texts can be problematic regarding what exact glyph shape they should actually take.

Recommended **Action Item** for Rick McGowan: Send comments from Section XI (Latin) to Patrik Sjöwall.

Latin

Date/Time: Fri Apr 16 17:24:19 CDT 2021
Name: Michael Everson

Report Type: Public Review Issue
Opt Subject: Latin Extended-D

The author requested the empty spaces at U+A7D2 and U+A7D4 be filled with LATIN CAPITAL LETTER DOUBLE THORN and LATIN CAPITAL LETTER DOUBLE WYNN.

We agree it would be helpful to have the uppercase letters beside their lowercase counterparts. However, the UTC was not convinced at the January 2021 meeting that the evidence was strong enough to encode them at this time.

The above comment has been sent to Michael Everson, so no UTC action is required.

Combining Diacritical Marks

Date/Time: Mon Apr 12 17:58:47 CDT 2021
Name: Michael Everson
Report Type: Public Review Issue
Opt Subject: Encode COMBINING OVERCURL at 1ACF

This is a lengthy request to encode U+1ACF COMBINING OVERCURL.

New characters for the next version should be approved at the January UTC meeting (or before). An updated document for COMBINING OVERCURL is invited.

The above comment has been sent to Michael Everson, so no UTC action is required.

XII. 14.0 BETA REVIEW FEEDBACK

Document: https://www.unicode.org/review/pri433/ PRI #433 Beta Feedback

Kana

Date/Time: Fri Jun 18 01:50:26 CDT 2021
Name: Lim Hian-tong
Report Type: Public Review Issue
Opt Subject: Issues related to Kana Extended-B (Public Review Issue #433)

[Detailed comments supporting a request to remove two annotations “also used for tone six” from Kana Extended-B names list.]

We agree with the request. The Names List editor has already removed the annotations from the Names List for Unicode version 14.0.

Ethiopic

Date: Mon, 5 Jul 2021 12:19:48 -0400
Name: Daniel Yacob
Subject: 3 Name Defects in Ethiopic Extended-B Tables

[Three character name errors in the beta chart for Ethiopic Extended-B were identified:
In beta chart:
1E7E9 ETHIOPIC SYLLABLE HWI
1E7EA ETHIOPIC SYLLABLE HWEE
1E7EB ETHIOPIC SYLLABLE HWE

The corrected names should be:
1E7E9 ETHIOPIC SYLLABLE HHWI
1E7EA ETHIOPIC SYLLABLE HHWEE
1E7EB ETHIOPIC SYLLABLE HHWE]

We agree. The names have been corrected in the data files for the Unicode version 14.0 release.