

**TO:** UTC

**FROM:** Deborah Anderson, Ken Whistler, Roozbeh Pournader, Lisa Moore, and Liang Hai<sup>1</sup>

**SUBJECT:** Recommendations to UTC #167 April 2021 on Script Proposals

**DATE:** April 23, 2021

The Script Ad Hoc group met on February 12, March 5, and April 9, 2021, in order to review proposals. The following represents feedback on proposals that were available when the group met.

## Table of Contents

|   |   |
|---|---|
| I. EUROPE .....   | 2 |
| 1 Latin.....  | 2 |
| 1a. Latin Letter Tone Six.....  | 2 |
| 1b. Modifier capital letters .....  | 3 |
| II. AMERICAS .....  | 3 |
| 2 Unified Canadian Aboriginal Syllabics.....  | 3 |
| III. AFRICA.....  | 4 |
| 3 Egyptian Hieroglyphs .....  | 4 |
| IV. MIDDLE EAST .....   | 4 |
| 4 Arabic .....  | 4 |
| 4a. Damma over Damma .....  | 4 |
| 4b. China comments on WG2 N5155 “Request for glyph changes and annotations for Kazakh, Kyrgyz, and Uyghur”..... | 5 |
| 4c. Behdini Kurdish HEH.....  | 5 |
| 4d. Urdu Digits .....   | 6 |
| 5 Book Pahlavi.....   | 6 |
| V. SOUTH AND CENTRAL ASIA.....  | 7 |
| 6 Malayalam.....  | 7 |
| 7 Nag Mundari .....   | 7 |
| 8 Old Turkic .....  | 8 |

---

<sup>1</sup> Also participating were Fred Brennan, Peter Constable, Craig Cornelius, Craig Cummings, Lorna Evans, Andrew Glass, Manish Goregaokar, Ned Holbrook, John Hudson, Marek Jeziorek, Cibu Johny, Jan Kučera, Norbert Lindenberg, Kamal Mansour, Michel Suignard, Lawrence Wolf-Sonkin, and Ben Yang.

|  |    |
|--|----|
| 9 Tulu/Tigalari .....  | 8  |
| 9a. Postpone New Tulu script from Karnataka Tulu Sahitya Academy proposal..... | 8  |
| 9b. Tulu/Tigalari proposal .....   | 9  |
| 10 Vatteluttu .....  | 11 |
| VI. SOUTHEAST ASIA, INDONESIA, AND OCEANIA .....                               | 12 |
| 11 Lao .....   | 12 |
| VII. EAST ASIA .....   | 13 |
| 12 Kanbun .....  | 13 |
| 13 Nushu .....   | 14 |
| VIII. SYMBOLS, PUNCTUATION, AND NOTATIONAL SYSTEMS .....                       | 14 |
| 14 Arrows .....  | 14 |
| 15 Blissymbols .....   | 14 |
| 16 Kaktovik Numerals .....   | 15 |
| 17 Znamenny Musical Notation .....   | 15 |
| 17a. Changes to Znamenny Musical Notation .....                                | 15 |
| 17b. Canonical Combining Classes of Znamenny Notation Characters .....         | 16 |
| IX. PUBLIC REVIEW FEEDBACK .....   | 17 |
| X. 14.0 ALPHA REVIEW FEEDBACK .....  | 20 |
| XI. COMMENTS AND DOCUMENTS NOT YET DISCUSSED .....                             | 25 |

## I. EUROPE

### 1 Latin

#### 1a. Latin Letter Tone Six

**Document:** [L2/21-083](#) Proposal to provide the glyphs of Latin letter tone six -- Kushim Jiang

**Comments:** In 2019, the Script Ad Hoc ([L2/19-286](#)) reviewed [L2/19-201](#) which requested a glyph change for the two characters U+0184 LATIN CAPITAL LETTER TONE SIX and U+0185 LATIN SMALL LETTER TONE SIX. The Script Ad Hoc deemed the request justified. This document completes the request by providing a font with the glyphs (attached to the document).

We recommend the glyph be changed.

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

Assigns an Action Item to Michel Suignard to change the glyphs for U+0184 LATIN CAPITAL LETTER TONE SIX and U+0185 LATIN SMALL LETTER TONE SIX for Unicode 14.0. (Reference: [L2/21-083](#))

---

## 1b. Modifier capital letters

**Document:** [L2/21-089](#) Unicode request for modifier capital letters – Miller

**Comments:** We reviewed this document, which requested two characters, modifier capital letter S and X.

The examples do not show the two modifier capital letters being used for phonetic transcription, but instead for morphological analysis apparatus, tone reconstruction and historical analysis, for which arbitrary letters could be used. In our view, rich text (or font-based stylistic sets) would be appropriate, but not plain text.

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

---

---

## II. AMERICAS

### 2 Unified Canadian Aboriginal Syllabics

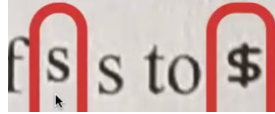
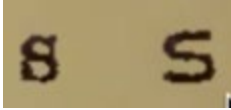
**Document:** [L2/21-088](#) Proposed revisions to the representative code chart characters of the Unified Canadian Aboriginal Syllabics -- Typotheque / Kevin King

**Comments:** We reviewed this set of 120 proposed revisions to the glyphs in the UCAS main and extended blocks for Carrier, Sayisi, and Ojibway. The document provides extensive examples, code charts showing the current shapes versus the proposed shapes, and figures from some of the original character proposals. A letter of support for the proposed Carrier changes is also included.

For Carrier, the characters should have similar height and width (known as the “square” form), which differs from the graphic appearance of other syllabic orthographies in UCAS (the “round” form). For Sayisi Dene, eight shapes should be changed to a “square” form, not the “round” form as in the code charts presently. For Ojibway, the orientation of one character needs to be modified.

The following comments were made:

- Use different colors in the code charts to indicate the following: (a) proposed characters with proportion change, (b) those characters with change in their vertical position (as in figure 1), and (c) those characters whose structure has significantly changed in addition to proportion. Include a legend describing what the different colors indicate. Use a lighter color to indicate those changes that are updated as part of a systematic font update (minor changes), versus cases with significant changes, shown in darker color.
- Mention that the vertical positioning of Carrier characters should be noted in the Core Spec.
- Inquire about Carrier serif forms (see below, left), and whether users have a preference? Also, ask whether the shape of “s” should also be raised (as in the right-hand example, below)?



- The author is welcome to recommend subheadings for groupings in the UCAS names list, as well as comments to indicate design of Carrier glyphs
- The author to provide a revised document for review at the next Script Ad Hoc.

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

---

---

## III. AFRICA

### 3 Egyptian Hieroglyphs

**Document:** [L2/21-096](#) Egyptian Hieroglyphs: Summary of Script Ad Hoc Discussion, April 2021 -- Anderson

**Comments:** This document is a summary of the Script Ad Hoc discussion on Egyptian Hieroglyphs with Michel Suignard, Andrew Glass and the Script Ad Hoc about ongoing work being done in conjunction with Egyptologists from Europe.

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

---

---

## IV. MIDDLE EAST

### 4 Arabic

#### 4a. Damma over Damma

**Document:** [L2/21-085](#) Damma over Damma comments in response to SAH recs -- Lateef Sagar

**Comments:** We reviewed this reply from Lateef Sagar, who was responding to comments from the January 2021 Script Ad Hoc recommendations in [L2/21-016r](#).

Sagar agrees the *damma over damma* shape, which he proposed in [L2/20-292](#), is a glyph variant of the normal *dammatan*. He wants to type *damma over damma* as it appears in the Tunisian Quran without requiring a custom font. He voiced concern the glyph would be lost whenever a font is changed to a standard Arabic font. He also mentions that the alternative *dammatan* can't be used if one relies on the country or locale in the font, because the normal and alternative *dammatan* are used for the same language (Arabic). As a result, he requests a new character.

In our view, in order to make a case in favor of the proposed *damma* over *damma* character a thorough analysis of *dammatan* is needed, including discussion of the Arabic model and examples of contrastive usage.

Note that documentation of the different styles of *dammatan* is a 14.0 task assigned to Ben Yang, et al., and it will include vertical positioning of *dammatan*. This documentation can be used by font designers to support the different shapes. Once the new text (with figures) on *dammatan* has been incorporated into the Core Spec, Sagar can see if it suffices or if he has additional comments.

**Recommendation:** We recommend the UTC make the following disposition:  
Assigns an Action Item to Debbie Anderson to relay feedback to Lateef Sagar, including comments from the April 2021 Script Ad Hoc Recommendations L2/21-073.

---

#### 4b. China comments on WG2 N5155 “Request for glyph changes and annotations for Kazakh, Kyrgyz, and Uyghur”

**Document:** [L2/21-050](#) Chinese comments on WG2 N5155 (= [L2/20-289](#))

**Comments:** We reviewed the Chinese comments on WG2 N5155 ([L2/20-289](#)) “Request for glyph changes and annotations for Kazakh, Kyrgyz, and Uyghur” by Lorna Evans. (Note that UTC #166, which took place in January 2021, already has taken action to make some changes proposed in L2/20-289.)

In our view, some explanation and clarifications to the comments from China are warranted. We recommend the Unicode liaison to WG2, Peter Constable, draft a response and the UTC review it before it is sent to China (and posted in the WG2 document register).

**Recommendation:** We recommend the UTC make the following disposition:  
Assigns an Action Item to Peter Constable to draft a response to [L2/21-050](#) for review by the UTC.  
(Reference: Section 4b of L2/21-073)

---

#### 4c. Behdini Kurdish HEH

**Document:** [L2/21-087](#) Regarding the Behdini Kurdish Heh -- Evans

**Comments:** We reviewed this document which has incorporated feedback from the Script Ad Hoc. It also includes recent comments from Jonathan Kew.

The document provides several options on how to handle a glyph variant for *heh* found in Behdini Kurdish, namely:

- Consider it a glyph variant of U+0647 ARABIC LETTER HEH
- Consider it a glyph variant of U+06BE ARABIC LETTER HEH DOACHASHMEE
- Encode a new character.

The document considers the outcomes of each option. Samples are provided, and the author reported she is in touch with Kurdish experts.

In our view, a meeting of a group of Unicode experts on Arabic is needed in order to decide which option is preferable. We also recommend asking Jonathan Kew if he would like to participate, as well as any other Unicode members interested in the topic.

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

Assigns an Action Item to Debbie Anderson to set up a meeting with Unicode experts on Arabic to discuss selected Arabic topics, including Behdini Kurdish HEH. (Reference: Section 4c of L2/21-073)

---

#### 4d. Urdu Digits

**Document:** [L2/21-080](#) Missing Urdu digits 4, 6 and 7 -- Aleem Qureshi

**Comments:** We discussed this request to encode three digits to represent 4, 6, and 7 used by Urdu.

We recommend the author be directed to Table 9-2 in the Core Spec, and he can respond if he has further questions. A FAQ on this topic could be written.

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

Assigns an Action Item to Debbie Anderson to relay feedback to the author of [L2/21-080](#) from the April 2021 Script Ad Hoc Recommendations L2/21-073.

---

## 5 Book Pahlavi

**Document:** L2/21-090 Advancing the encoding model for Book Pahlavi -- Anshuman Pandey

**Note:** *The Script Ad Hoc reviewed an earlier version of the proposal than the document posted in the document register. The comments below are based on the earlier version.*

**Comments:** We reviewed this document, which builds off Roozbeh Pournader’s document, [L2/20-135](#) “Next Steps on Book Pahlavi” and discusses the encoding model. (For further background on Book Pahlavi, see the topical document list for Book Pahlavi, <https://unicode.org/L2/topical/bookpahlavi/>.)

Discussion focused on the fact that some of the examples on page 11, for example, are based on poor fonts. As a result, some distinctions shown in the examples may not be meaningful, and hence should not be encoded. We recommend more examples be included as well as an explanation of the use of colors in the figures (i.e., page 10).

We recommend that after Pandey has posted the document in the document register, experts be invited to review the document and make comments.

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

Assigns an Action Item to Debbie Anderson and Anshuman Pandey to post the revised document and invite experts to review it. (Reference: Section 5 of L2/21-073)

---

---

## V. SOUTH AND CENTRAL ASIA

### 6 Malayalam

**Document:** [L2/20-294](#) Clarification Request on Malayalam */ṅṭa/* Conjunct Specification Proposed in L2/19-345r2 -- Cibu Johny

**Comments:** We reviewed this request for improved wording in the Malayalam block introduction on the legacy representation of the conjunct */ṅṭa/*. The problematic wording is the following underlined text: “It is recommended that implementations be prepared to treat *<na, virama, rra>* as an equivalent sequence of the recommended representation,” since it does not explicitly state whether newly created text can use *<na, virama, rra>* or not. As a result, users may use legacy sequences instead of the recommended sequences.

The proposed text change specifies that *<na, virama, rra>* in existing text is an equivalent sequence to the recommended representation of *<0D7B chillu-n, 0D4D virama, 0D31 rra>* in newly generated text.

In our opinion, the recommended change is a clear improvement, and the text should be sent to the Editorial Committee for incorporation into the Core Spec.

The SAH also discussed the definition of “equivalent” and “equivalent sequences” in Core Spec documentation. Since this more general topic falls outside L2/20-294 proper, members who wish to take up the topic are invited to write a document and forward it to the Properties and Algorithms group for consideration.

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

Assigns an Action Item to Liang Hai and the Editorial Committee to provide text for the Core Spec on the legacy representation of the conjunct */ṅṭa/*, based on the proposed wording on page 2 of [L2/20-294](#).

---

---

### 7 Nag Mundari

**Document:** [L2/21-031R](#) Proposal to Encode the Nag Mundari Script -- Wolf-Sonkin, Mandal

**Note:** *The Script Ad Hoc reviewed an earlier version of the proposal than the latest document posted in the document register.*

**Comments:** We discussed this script proposal, which had not yet been recommended for approval, due to discussion about the script’s name (originally “Mundari Bani”). Since “Bani” means “language”, another name should be selected. The proposed name “Nag Mundari” was deemed acceptable to one

group that has been involved. An explanation of the name is also needed. We recommend the author update his proposal and post it in the document register.

**Recommendation:** Pending review of the revised proposal, we recommend that the UTC approve the following:

**SAH-UTC167-R1:** The UTC accepts 42 Nag Mundari characters in a new Nag Mundari block (U+1E4D0..U+1E4FF) for encoding in a future version of the standard, with glyphs and properties as documented in L2/21-031R.

We also recommend that the UTC make the following dispositions:

Assigns an AI to Ken Whistler to update the Pipeline. (Reference: L2/21-031R)

Assigns an AI to Debbie Anderson to confirm the SMP Roadmap reflects the new script name.

Assigns an AI to Lawrence Wolf-Sonkin to provide Michel Suignard with a font.

---

---

## 8 Old Turkic

**Document:** [L2/21-081](#) Application for Adding Letters to Old Turkish Alphabet -- Gökbeý Uluç

**Comments:** This document proposes two Old Turkish ligatures. Encoding ligatures in Old Turkic can be a slippery slope. Note that an earlier proposal by Anshuman Pandey ([L2/19-069](#)) had proposed the Old Turkic ligature ORKHON CI, but the UTC decided to not accept it as a separate character, instead recommending it be handled as a ligature of two letters with ZWJ.

Regarding this proposal, Anshuman Pandey remarked, “I think SAH should table the OT ligatures. It would be useful to understand how many such ligatures exist in the extant records and the nature and structure of ligatures. It would be useful to know the frequency of their usage within a given text, etc. If there is a limited set, then it would be useful to encode them all.”

We recommend other experts also review this proposal. In addition, the author should explain the orientation of “ICH” in example 5.

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

Assigns an Action Item to Debbie Anderson to seek feedback from other Old Turkic experts and to ask the author of [L2/21-081](#) to explain the orientation of “ICH” in example 5. (Reference: Section 8 of L2/21-073).

---

---

## 9 Tulu/Tigalari

### 9a. Postpone New Tulu script from Karnataka Tulu Sahitya Academy proposal

**Document:** [L2/21-092](#) Proposal to postpone encoding of the new Tulu script from the Karnataka Tulu Sahitya Academy -- Rajan, et al.

Background document: [L2/21-019](#) Proposal to encode Tulu -- Pavanaja



**Comments:** We reviewed this document which basically proposes the Tulu-Tigalari script proposal proceed through the review and approval process, without waiting for a full proposal from the Tulu Sahitya Academy (“Academy”) for the reformed orthography.

The document points out specific issues with the Academy proposal.

- The Tulu-Tigalari script is used to write Sanskrit and closely represents forms attested in manuscripts. The Academy proposal, on the other hand, is intended to write the modern Tulu language. Instead of extending the original, historical script and orthography, the Academy proposal made changes to the original script.

For example, the Academy proposal borrowed some characters from Kannada (such as Vowel Sign Vocalic RR), instead of using forms already existing in Tulu-Tigalari. Other characters proposed by the Academy differ from the original value of characters in Tulu-Tigalari (i.e., Academy NGA is NYA in Tulu-Tigalari), which will cause confusion to users. The behavior of ligatures and conjuncts differs between the Academy version and Tulu-Tigalari, which has a major impact on the encoding model decision.

- The authors of L2/21-092 note that the Academy proposal for encoding is not yet stable. Recent books, for example, vary in their representation of the vowels and vowel signs E, EE, and OO. Some primers published by the Academy contain ligature forms not supported in the Academy proposal. In contrast, the Tulu-Tigalari script proposal, which covers the traditional Tulu-Tigalari script, is well-attested and is stable.

We agree that the Tulu-Tigalari script can proceed separately from the Tulu proposal from the Academy. In our view, if at a later point the reformed orthography has matured and is stable, a decision can then be made whether to unify the reformed orthography with Tulu-Tigalari (adding new characters if needed) or whether a separate encoding is warranted.

It was noted that a few technical issues still remain in the Tulu-Tigalari proposal (see 9b. below). Norbert Lindenberg suggested changing the name of TULU-TIGALARI JOINER to TULU-TIGALARI CONJOINER. The Script Ad Hoc agreed. (The group also agreed the name CONJOINER would be a better name for Kawi, replacing KAWI SUBJOINER.) Norbert Lindenberg will send his other comments to the proposal authors

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

Assigns an Action Item to Debbie Anderson to send [L2/21-092](#) to the Tulu Sahitya Academy contacts.

---

## 9b. Tulu/Tigalari proposal

**Document:** [L2/21-086](#) Updated proposal to encode Tulu-Tigalari script in Unicode -- V. Murthy and V. Rajan

**Comments:** This revised proposal notes recent updates to the proposal on page 2. The proposal states the repertoire supports use of the script for the Tulu, Kannada, and Sanskrit languages “for archival use,” as well as contemporary usage “without distorting the traditional orthography.” The proposal does

not include some characters proposed by the Karnataka Tulu Sahitya Academy (such as the numbers and long and short /e/, /o/ and /ae/).

V. Murthy reported she was not able to meet with members of the Karnataka Tulu Sahitya Academy. (For a comparison between an earlier version of this proposal and the [proposal from Karnataka Tulu Sahitya Academy](#), see [L2/20-279](#).)

### *Repha*

In the proposal (page 34ff.), the ligated form of *repha* is invoked by *repha* with the conjoiner (proposed character name: TULU-TIGALARI JOINER), a novel approach that may be problematic. A conjoiner in Indic encoding is meant to connect two normal bases, while trying to conjoin this special base (*repha*) that essentially has a built-in conjoiner creates an irregularity in the encoding logic, and does not work with the Universal Shaping Engine (USE). If the approach described in the proposal is adopted, USE would need to be extended. (Note: During discussion of [L2/21-092](#) [9a., above], Norbert Lindenberg suggested changing the name of TULU-TIGALARI JOINER to TULU-TIGALARI CONJOINER. The Script Ad Hoc agreed. The group also agreed the name CONJOINER would be a better name for Kawi, replacing KAWI SUBJOINER.)

Since the use of the ligated *repha* is stylistic, does it need to be handled at the encoding level – could it be handled at the font-level? Or should ZWJ or ZWNJ be used in this case?

In our view, the authors should consider use of general joiners (ZWJ and ZWNJ) or a font-level approach for ligating *repha*. In order to make a decision, the authors may try comparing this ligation behavior with other unpredictable ligatures in the script and other scripts. The cost of introducing a novel encoding logic is high, and a largely stylistic variation may not be worth it. Also they should consider that, in practice, fonts will only be able to provide a selective set of ligatures and will fall back to the non-ligated form. Such a fallback situation will lead to an unpredictable behavior of the conjoiner in a *repha*-ed akshara, similar to how virama’s behavior is not fully predictable in the Devanagari encoding model.

### *Looped Virama*

The name (previously “ligating virama”) appears to be acceptable.

### *Conjuncts*

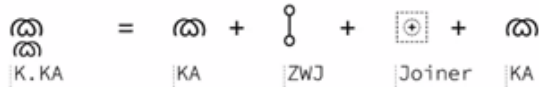
One difference between the “traditional” (manuscript) use of the script versus the modern “reformed” version is conjoining behavior, with conjuncts (known as “horizontal conjuncts” in the document) vs. subscripts (i.e., “vertical conjuncts”):

|  | Traditional | Reformed |
|--|-------------|----------|
| <div style="border: 1px solid black; padding: 5px; display: inline-block;">kku<br/>ktu</div> | ಕು<br>ಕು    | ಕು<br>ಕು |

(Appendix 1. p. 49)

The proposal recommends handling the different conjuncts as described below:

Forced **Vertical** Conjunct : C1 + ZWJ + Joiner + C2 = <C1><C2>



Forced **Horizontal** Conjunct : C1 + Joiner + ZWJ + C2 = <C1><C2>



The use of the vertical versus the horizontal conjunct appears to be stylistic in the traditional orthography. Should they be handled at the encoding level?

Note: Liang Hai also noted that the proposed specification of using ZWJ to “force” (horizontal) conjuncts also breaks the general assumption of Indic encoding models.

Tulu-Tigalari could handle subscript conjuncts like the encoded traditional Malayalam orthography, i.e., subscript is the default, and use font-level opt-in for horizontal conjuncts. However, this would make the orthography that only uses subscripts (vertical conjuncts) a secondary citizen. Would this be acceptable to modern users of the reformed orthography?

This proposal has already been sent to the Tulu Sahitya Academy contacts. Those who wish to make specific comments on the proposal are encouraged to contact the proposal authors.

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

---

## 10 Vatteluttu

**Documents:** [L2/21-052](#) Request to Change Pulli Representation in the Proposed Vatteluttu Encoding -- Cibu

Background doc: [L2/16-068](#) Preliminary proposal to encode Vatteluttu in Unicode -- Pandey

**Comments:** We reviewed this document, which requested a change in the representation of the *pulli* as described in the preliminary Vatteluttu proposal from 2016 by Anshuman Pandey ([L2/16-068](#)). (Note: The proposal generated a response from the Tamil Virtual Academy [[L2/16-118](#)], which stated TVA would be proposing the script separate, though no document has been received to date.) The Vatteluttu script is an historical script used in Tamil Nadu and the Kerala region.

Vatteluttu has a virama (“*pulli*”) based model. The *pulli* appears in older inscriptions from Tamil Nadu, but has not been found in inscriptions in Kerala that are dated to later centuries. The *pulli* acts as a vowel-killer and a vowel reducer (to either zero or schwa), though the script does not have conjuncts or ligatures. The *pulli* is relatively rare and appears as a combining dot above or the top-right of the base character, which can be certain independent vowels, consonants, or consonant-vowel sign clusters.

The following comments were made:

- The overall argument in this document is that the *pulli* should be separately encoded for use with a vowel or consonant, and not as in the preliminary proposal, which atomically encodes the vowels with the *pulli* but uses a separate *pulli* as a consonant vowel-killer.
- This document recommends the independent vowel letters *e* and *o* and dependent vowel signs *e* and *o* be represented by sequences of the vowel followed by the *pulli*.
- The preliminary proposal contained canonical decompositions for the vowel signs *e*, *o*, and *oo*. Because they have canonical decompositions a “Do Not Use” table is not required (as suggested on pages 3-4).
- In theory the request to change the *pulli* representation makes sense in our opinion, but the change needs to be seen within the context of the entire proposal.
- The document should specify that the *pulli* always comes last in a sequence when representing the dependent and independent vowels.

The author has already noted the above comments. A revised Vatteluttu proposal will at some point be submitted.

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

---



---

## VI. SOUTHEAST ASIA, INDONESIA, AND OCEANIA

### 11 Lao

**Document:** [L2/21-093](#) Proposal to Encode Lao Sign Yamakkan -- Rajan and Bhikkhu

**Note:** *The Script Ad Hoc reviewed an earlier version of the proposal than the document posted in the document register. The comments below are based on the earlier version.*

**Comments:** We reviewed this proposal to add one new character, Lao Sign Yamakkan, which is needed to support Lao Pali. The proposed character is cognate with U+0E4E THAI CHARACTER YAMAKKAN.

The following comments were made:

- In general, it is helpful to add information on the interaction of tone marks which have fixed position classes. However, Pali doesn’t use tone marks, so no examples can be provided.
- Norbert Lindenberg suggested that for new Brahmic characters, CCC must be 0 unless it is a virama.
- Recommended changes include:
  - Add the Indic\_Syllabic\_Category and CCC values of U+0E4E THAI CHARACTER YAMAKKAN. If the Thai and Lao character values differ, explain why.
  - Change the character’s name to LAO YAMAKKAN
  - Make CCC=0 and specify LAO YAMAKKAN goes at the end of the syllable.

With the above changes, we recommend the UTC accept the character U+0ECE LAO YAMAKKAN.

**Recommendation:** We recommend that the UTC approve the following:

**SAH-UTC167-R2:** The UTC 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: [L2/21-093](#))

We also recommend that the UTC make the following dispositions:

Assigns an AI to Ken Whistler to update the Pipeline. (Reference: L2/21-093)

Assigns an AI to Debbie Anderson and Vinodh Rajan to provide Michel Suignard with a font.

---

## VII. EAST ASIA

### 12 Kanbun

**Document:** [L2/21-082](#) Proposal to encode 20 Additional Kanbun Marks -- Wang Yifan

**Comments:** This proposal requests 20 Kanbun marks, and recommends space be reserved for 12 as yet unattested characters. The green coloring in the list is used for the different columns.

The proposal builds on [L2/20-232](#), which proposed 9 characters. The September 2020 [Unihan Ad Hoc](#) group had recommended encoding the 9 characters in a block of two columns and the SAH concurred. The Oct. 2020 UTC, however, did not accept them. Ken Lunde corresponded with the author after the UTC, recommending he add compatibility decompositions with the superscript value and that the glyphs be shown in reduced size (as in the current Kanbun block). This document is a follow-up to that recommendation.

Specific comments raised during discussion:

- We recommend the Roadmap Committee allocate the three columns U+16FB0..U+16FDF for KANBUN EXTENDED-A. The location in SMP is preferred (in alignment with [Jan 2020 SAH recommendations](#), which responds to [L2/20-276](#)).
- In the next revision provide a table showing how the names of the existing Kanbun characters are related to the proposed characters (i.e., FIVE vs. ...FIFTH, etc.). The proposal should only include those characters with attested usage, removing “reserved” code points.
- Fred Brennan offered to read the proposal and get in touch with the author.
- Ken Lunde recommended the proposal author resubmit the proposal after evidence for the additional characters have been found. The Kanbun block has been unchanged since the very beginning of the Unicode Standard, so there is obviously no particular rush to get additional Kanbun marks encoded.

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

Assigns an Action Item to Debbie Anderson to forward the relevant comments in Section 12 of L2/21-073 to the proposal author and also ensure the Roadmap request is forwarded to the Roadmap Committee.

---

## 13 Nushu

**Document:** [L2/21-024](#) Updates for 3 values for 3 Nushu characters in UCD -- Eiso Chan and Kushim Jiang

**Comments:** We reviewed this proposal to correct Nushu kReading values for U+1B1FE, U+1B2EB and U+1B2EC in NushuSources.txt. The request provides evidence. The proposed updates were reviewed by Chen Zhuang of the China National Body in December 2020 (and his experts), and were deemed acceptable.

**Recommendation:** We recommend the UTC make the following disposition:  
Assigns an Action Item to Michel Suignard to update the three readings in NushuSources.txt for Unicode 14.0. (Reference: [L2/21-024](#))

---

---

## VIII. SYMBOLS, PUNCTUATION, AND NOTATIONAL SYSTEMS

### 14 Arrows

**Document:** [L2/21-084](#) Proposal to include four-arrow symbols in Unicode -- Diego Guella

**Comments:** We reviewed this request for additional arrow symbols found in cursor icons on desktop operating systems.

The author refers to use of the symbols in software documentation at his company. Is the software public, so the documentation would be in the public sphere?

In our view, accepting the proposed symbols would equally apply to any user-interface icon; such icons are not automatically candidates for encoding. Stronger evidence and rationale are needed.

It was noted that the proposal is similar to the Powerline symbols proposal ([L2/19-068](#)). The Powerline symbols are currently in a set of additions to Legacy Computing symbols (though the Script Ad Hoc only felt a few were good candidates for encoding, in its [Oct. 2019 recommendations](#)).

**Recommendation:** We recommend the UTC make the following disposition:  
Assigns an Action Item to Debbie Anderson to forward the comments in Section 14 of L2/21-073 to the proposal author.

---

---

### 15 Blissymbols

**Document:** [L2/20-271](#) On the keyboard inputting of Blissymbols (WG2 N5149) – Everson

**Comments:** This document is informational. (A full Blissymbols proposal will be forthcoming.)

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

---

---

## 16 Kaktovik Numerals

**Document:** [L2/21-058](#) Unicode request for Kaktovik numerals -- Marin Silva, Miller

**Comments:** We reviewed this proposal for a set of twenty Kaktovik numeral characters. The numerals are part of a penta-vigesimal system and are in active use for Iñupiaq Eskimo (ISO 639: ipk) in Alaska.

The name “Kaktovik” is acceptable to the various groups. Rationale for the atomic encoding is provided. We agree that the atomic encoding is appropriate.

We recommend the 20 characters be approved. The SMP Roadmap already reflects the new block.

**Recommendation:** We recommend that the UTC approve the following:

**SAH-UTC167-R3:** The UTC accepts 20 Kaktovik numeral characters in a new Kaktovik Numerals block (U+1D2C0..U+1D2DF) for encoding in a future version of the standard, with glyphs and properties as documented in [L2/21-058](#). (Reference: Section 16 of L2/21-073)

We recommend that the UTC make the following dispositions:

Assigns an Action Item to Kirk Miller to forward a font to Michel Suignard.

Assigns an Action Item to Ken Whistler to update the pipeline.

---

---

## 17 Znamenny Musical Notation

### 17a. Changes to Znamenny Musical Notation

**Document:** [L2/21-060](#) Changes to Znamenny Musical Notation—Andreev

**Comments:** We reviewed this document which requested a limited number of specific changes to the Znamenny Musical Notation characters, currently in the alpha release for 14.0 (PRI #428).

We agree with the proposed changes (items #1-#4, listed under the Recommendations below).

We also agree with the General Category property change proposed by Buff (#5), i.e., that the general category for the Znamenny priznak modifiers should be changed from Cf to Mn. Format control characters are ignored by many processes unlike Mn, which participate in normal rendering. For the Znamenny priznak modifiers, Mn is appropriate. (See [UAX 29](#), rules GB4 and 5, which show grapheme clusters break before and after controls, whereas in GB 9 grapheme clusters do not break before extending characters or ZWJ [and most combining marks are extending characters].)

This change will require a change in the glyphs so they are combining marks. (Note: Michel Suignard has already received a font with the glyphs for U+1CF42...U+1CF46 as dotted circles.)

Regarding the question of the canonical combining marks, changing the current CCC=0 values would very likely result in postponing the script to Unicode 15.0. Note that once assigned, the CCC values are immutable, so it is prudent to make the decision carefully. If changed and the non-zero CCC values are

incorrect, the user community would need to create a workaround similar to the Arabic Mark Transient Reordering Algorithm in UTR #53. The consensus after discussion was to keep CCC=0.

**Recommendation:** We recommend that the UTC approve the following:

**SAH-UTC167-R4:** The UTC accepts the following changes:

1. Removes U+1CFC0 ZNAMENNY NEUME STRELA ROZHAYA
2. Renames U+1CFBF ZNAMENNY NEUME ROZHEK to ZNAMENNY NEUME KUFISMA
3. Modifies the name of U+1CF42 ZNAMENNY PRIZNAK MODIFIER LEVEL 2 to U+1CF42 ZNAMENNY PRIZNAK MODIFIER LEVEL-2
4. Modifies the name of U+1CF43 ZNAMENNY PRIZNAK MODIFIER LEVEL 3 to U+1CF43 ZNAMENNY PRIZNAK MODIFIER LEVEL-3
5. Moves U+1CFC4 ZNAMENNY NEUME OBLAKO to U+1CFC0.
6. Changes the gc property of the Znamenny priznak modifiers (U+1CF42..U+1CF46) from gc=Cf to gc=Mn.

(Reference: Section 17a of L2/21-073)

We also recommend that the UTC make the following dispositions:

Assigns an Action Item to Ken Whistler to update the pipeline.

Assigns an Action Item to Ken Whistler to update UnicodeData.txt (and other data files) for 14.0.

---

## 17b. Canonical Combining Classes of Znamenny Notation Characters

**Document:** [L2/21-033r](#) Canonical Combining Classes of Znamenny Notation Characters -- Buff

**Comments:** This document was revised, based on comments from the Script Ad Hoc. Changes in this revised version are listed on page 11.

Comments:

- Earlier concerns of the Script Ad Hoc were raised again, namely, that if mistakes are made in the initial assignments, they cannot be corrected in the future, due to stability guarantees. If the CCC remains zero, a Unicode Technical Note could be created to guide users on how to encode the text, documenting ordering.
- Because the Znamenny Notation is likely to be used by a very small circle of users and will not likely appear in a system font, erring on the side of caution (i.e., keeping CCC=0) is probably the most prudent approach.

The outcome of the discussion was to keep the CCC values as zero.

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

Notes this document but takes no further action.



---

---

## IX. PUBLIC REVIEW FEEDBACK

**Note:** The following comments are drawn from [L2/21-068](#) Comments on Public Review Issues (unless otherwise noted)

### *Arabic*

**Date/Time:** Fri Mar 5 20:39:31 CST 2021

**Name:** David Corbett

**Opt Subject:** Representing hamza in lam–alef ligature

**Comments:** The question from David Corbett was how a font should handle the sequence of hamza in a lam-alef ligature.

The Unicode Standard is not encyclopedic in providing examples of all possible combinations and even if examples are provided, fonts may or may not support the sequence. There are, however, documents in the document register from Roozbeh Pournader and Tom Milo that talk about this (e.g., [L2/13-226](#) and [L2/14-109](#)).

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

Assigns an Action Item for Lorna Evans, Roozbeh Pournader, and Debbie Anderson to review the text in Chapter 9, “Arabic”, subsection “Quranic Texts” and see if the wording can be improved.

### *Kannada*

**Date/Time:** Tue Feb 23 20:12:22 CST 2021

**Name:** David Corbett

**Opt Subject:** Kannada

**Comments:** The author asks about a statement in section 5.21 of the Core Spec, stating format characters may have no visible effect on display. The feedback submitter specifically refers to ZWJ and ZWNJ and asks how the Kannada sequence <ra, ZWNJ, virama, consonant> (Chapter 12) should be rendered.

It was noted that the behavior of ZWJ and ZWNJ in Kannada does not differ from Indic generally, and the question of the behavior of the format characters applies equally to other ISCII-inspired Indic encoding models in the Unicode Standard.

Section 23.2 was mentioned during discussion, but the section on joiners and non-joiners in that section is very generic, and shouldn’t be relied on for specifics. (The text in 23.2 was purposely not included in the chapter on conformance [chapter 3 of Core Spec], because the behavior of ZWJ and ZWNJ is so complicated.)

The group agreed that it is not practical for Unicode to offer an opinion on any arbitrary undefined behavior of format controls, but the definition should be more systematically specified and thus are tightened up to provide reasonable fallbacks for situations that are not directly defined.

**Recommendation:** No specific action is required at this time.

*Kawi*

**Date/Time:** Tue Jan 19 19:39:21 CST 2021

**Name:** Eduardo Marín Silva

**Opt Subject:** On the Kawi space filler and the names of punctuation characters

**Related documents:**

[L2/21-048](#) Reply to 'On the Kawi Space Filler and the Names of Punctuation Characters -- Aditya Bayu Perdana et al.

**Comments:** We reviewed the public review feedback from Eduardo Marín Silva, and responses to the public review from Aditya Bayu Perdana et al. In sum, no changes were recommended by the SAH.

The following are comments raised during discussion:

- *Public Review comment:* Remove SPACE FILLER and annotate DIGIT FOUR, noting its use as a space filler.  
*SAH comments:* The examples in Bayu’s response demonstrated that “4” and the space filler are distinct. The SAH agreed that it might be useful to pick another representative glyph for the SPACE FILLER which is graphically different from the “4” representative glyph, but Bayu reports selecting which glyph to use is difficult, given that the script spanned 800 years and there are many variants.
- *Public Review comment:* Change the name of KAWI PUNCTUATION ALTERNATE SECTION MARK to SECTION MARK WITH REPHA  
*SAH comments:* Bayu noted that the shape above or beside the section mark can vary, so a name with “*repha*” is not useful (and could cause confusion, if one expected it to be a sequence with a *repha*). The Core Spec or an annotation in the names list could document that the section mark can have different “flourish” shapes.
- *Public Review comment:* Change the name of KAWI PUNCTUATION CIRCLE to CIRCLE WITH DOT.  
*SAH comments:* As noted in the response document, the circle can have different shapes, so in our view, the current name should be retained.
- *Public Review comment:* Change name KAWI PUNCTUATION CLOSING SPIRAL to KAWI PUNCTUATION SPIRAL WITH WAVY TAIL  
*SAH comments:* The SAH saw no reason to make the change.
- *Public Review comment:* Annotate SECTION MARKER as Siddham  
*SAH comments:* “Siddham” could apply to several characters, and Bayu reports identifying the two SECTION MARKERS with a “Siddham” annotation is not clear at this point. No change is needed in our view.

**Recommendation:** No change is needed.

## Malayalam

**Date/Time:** Tue Feb 23 20:40:37 CST 2021

**Name:** David Corbett

**Report Type:** Other Question, Problem, or Feedback

**Opt Subject:** Edge case for ZWJ and ZWNJ in Malayalam

**Comments:** The author asks about the rendering of unexpected sequences with joiners, specifically Malayalam <C, ZWJ, ZWNJ, virama, C> and <C, ZWNJ, ZWJ, virama, C>, since the general rule is to ignore them when they appear unexpectedly.

The situation is complicated. To answer the feedback, a theory paper on how to deal with unexpected sequences is needed, proposing how much more specific the Core Spec should be. See also the Kannada section for the same general problem.

The group agreed more work on Indic is needed, which defines Indic shaping in a systematic way (such as, defining the expected behavior of all the controls).

**Recommendation:** No specific action required at this time.

## Mandaic

(from [L2/21-011](#))

**Date/Time:** Sun Dec 20 09:57:01 CST 2020

**Name:** David Corbett

**Report Type:** Other Question, Problem, or Feedback

**Opt Subject:** Mandaic kad

**Comments:** The feedback asks whether the two ways to represent *kad* in Mandaic (U+0857 MANDAIC LETTER KAD or the sequence <U+084A MANDAIC LETTER AK, U+0856 MANDAIC LETTER DUSHENNA>) mean the same thing and should be rendered identically (and if identical, which should be used).

**Recommendation:** No additional action is required. Ben Yang is already working on a document discussing the options.

## Myanmar

**Date/Time:** Thu Feb 25 00:46:04 CST 2021

**Name:** Norbert Lindenberg

**Report Type:** Error Report

**Opt Subject:** Incorrect Indic Syllabic Category for Myanmar Sign Asat

**Comments:** The feedback author considers the current IndicSyllabicCategory for U+103A MYANMAR SIGN ASAT to be incorrect, recommending Virama instead of Pure\_Killer.

We discussed this topic and decided to take no action. Norbert Lindenberg will do further research.

**Recommendation:** No action is required.

## Old Hungarian

(from [L2/21-011](#))

**Date/Time:** Thu Nov 26 10:22:07 CST 2020

**Name:** David Corbett

**Report Type:** Other Question, Problem, or Feedback

**Opt Subject:** Diacritics in Old Hungarian

**Comments:** The feedback refers to L2/11-242R, an ad hoc report for Old Hungarian that was trying to reach consensus between competing script proposals. At the time, the Script Ad Hoc had recommended U+1DC4 COMBINING MACRON-ACUTE for the “duplicating mark” proposed in WG2 [N4007](#), but this recommendation need not be binding today.

In our view, the Core Spec is not meant to spell out all the details for each script. In this case, U+0304, the generic character, should be used.

**Recommendation:** No action is required.

---

---

## X. 14.0 ALPHA REVIEW FEEDBACK

**Document:** <https://www.unicode.org/review/pri428/>

**Date/Time:** Sun Feb 14 09:01:03 CST 2021

**Name:** Charlotte Buff

**Opt Subject:** PRI #428: Incorrect CCC of U+10F83

Proposed Character U+10F83 OLD UYGHUR COMBINING DOT BELOW currently has canonical combining class 230 (Above), but the correct value would be 220 (Below).

**Comment:** We agree. This has been corrected in the datafile and is so noted on [PRI #428 feedback page](#).

**Date/Time:** Mon Feb 15 19:56:28 CST 2021

**Name:** Eduardo Marín Silva

**Opt Subject:** Suggestions on the alpha code chart of Diacritical Marks Extended

The request (#3 in the feedback) is to move U+1ACD COMBINING DOUBLE PLUS ABOVE and U+1ACE COMBINING DOUBLE PLUS BELOW so they are located immediately after U+1AC8 COMBINING PLUS SIGN ABOVE.

Note: This same request was submitted by Michael Everson in Alpha feedback (Mon Apr 12 16:47:19 CDT 2021)

**Comment:** In our view, moving the characters has no real benefit, but incurs costs. The Script Ad Hoc does not recommend this change.

**Date/Time:** Sun Feb 14 08:59:40 CST 2021

**Name:** Charlotte Buff

**Opt Subject:** PRI #428: Incorrect decomposition mapping of U+107A9

Proposed character U+107A9 MODIFIER LETTER SMALL R WITH FISHHOOK currently decomposes to U+207E SUPERSCRIPT RIGHT PARENTHESIS, but the correct mapping would be to U+027E LATIN SMALL LETTER R WITH FISHHOOK.

**Comment:** This has been corrected in the datafile and is so noted on PRI #428 feedback page

**Date/Time:** Sun Feb 14 09:29:15 CST 2021

**Name:** Charlotte Buff

**Opt Subject:** PRI #428: U+1CF42 and U+1CF43 have nonconformant names

The names of proposed characters U+1CF42 (ZNAMENNY PRIZNAK MODIFIER LEVEL 2) and U+1CF43 (ZNAMENNY PRIZNAK MODIFIER LEVEL 3) currently do not conform to section 4.8 of the Unicode Standard. A hyphen-minus needs to be inserted before the final digit in both names because a digit must not immediately follow a space.

**Comment:** We agree. See the Script Ad Hoc recommendations above.

**Date/Time:** Sun Feb 14 09:58:07 CST 2021

**Name:** Charlotte Buff

**Opt Subject:** PRI #428: Names of U+1FAF1 and U+1FAF2

The names of proposed characters U+1FAF1 RIGHTWARD BACKHAND and U+1FAF2 LEFTWARD HAND could potentially be changed to RIGHTWARDS BACKHAND and LEFTWARDS HAND respectively. The words “rightward” and “leftward” do not occur in any other Unicode character names; instead the spellings “rightwards” and “leftwards” are used every single time.

**Comment:** We agree. This change will make the names more consistent.

**SAH-UTC167-R5:** We recommend the UTC accept the following name changes:

U+1FAF1 RIGHTWARD BACKHAND > RIGHTWARDS BACKHAND

U+1FAF2 LEFTWARD HAND > LEFTWARDS HAND

(Reference: PRI #428, Buff feedback Feb. 14 2021)

We also recommend that the UTC make the following dispositions:

Assigns an Action Item to Ken Whistler to update the pipeline.

Assigns an Action Item to Ken Whistler to update UnicodeData.txt (and other data files) for 14.0.

**Note:** After the Script Ad Hoc had met, later feedback from Michael Everson was received (April 16). He recommended the following, but the Script Ad Hoc did not discuss this option:

U+1FAF1 RIGHT-POINTING BACK OF HAND

U+1FAF2 LEFT-POINTING FRONT OF HAND

**Date/Time:** Sun Feb 14 10:01:09 CST 2021

**Name:** Charlotte Buff

**Opt Subject:** PRI #428: Defective glyph for U+1FAE2

The code chart glyph for proposed character U+1FAE2 FACE WITH OPEN EYES AND HAND OVER MOUTH is inverted, showing a solidly filled face instead of an outline drawing like the other faces.

**Comment:** We agree there is an error in glyph. The glyph has been fixed.

**Date/Time:** Sun Feb 14 10:25:15 CST 2021

**Name:** Charlotte Buff

**Opt Subject:** PRI #428: Names of dezH and tesh digraphs with hooks

The names of the following proposed characters should be adjusted to include the word “digraph” for consistency with their respective hook-less counterparts (U+02A4 LATIN SMALL LETTER DEZH DIGRAPH and U+02A7 LATIN SMALL LETTER TESH DIGRAPH):

U+1DF12: LATIN SMALL LETTER DEZH WITH PALATAL HOOK → LATIN SMALL LETTER DEZH DIGRAPH WITH PALATAL HOOK

U+1DF17: LATIN SMALL LETTER TESH WITH PALATAL HOOK → LATIN SMALL LETTER TESH DIGRAPH WITH PALATAL HOOK

U+1DF19: LATIN SMALL LETTER DEZH WITH RETROFLEX HOOK → LATIN SMALL LETTER DEZH DIGRAPH WITH RETROFLEX HOOK

U+1DF1C: LATIN SMALL LETTER TESH WITH RETROFLEX HOOK → LATIN SMALL LETTER TESH DIGRAPH WITH RETROFLEX HOOK

**Comment:** We agree; the proposed changes will make the names more consistent. Kirk Miller, who proposed the characters, is okay with the new names.

**SAH-UTC167-R6:** We recommend the UTC accept the following name changes:

U+1DF12 LATIN SMALL LETTER DEZH WITH PALATAL HOOK > LATIN SMALL LETTER DEZH DIGRAPH WITH PALATAL HOOK

U+1DF17 LATIN SMALL LETTER TESH WITH PALATAL HOOK > LATIN SMALL LETTER TESH DIGRAPH WITH PALATAL HOOK

U+1DF19 LATIN SMALL LETTER DEZH WITH RETROFLEX HOOK > LATIN SMALL LETTER DEZH DIGRAPH WITH RETROFLEX HOOK

U+1DF1C LATIN SMALL LETTER TESH WITH RETROFLEX HOOK > LATIN SMALL LETTER TESH DIGRAPH WITH RETROFLEX HOOK

(Reference: PRI #428, Buff feedback Feb. 14 2021)

We also recommend that the UTC make the following dispositions:

Assigns an Action Item to Ken Whistler to update the pipeline.

Assigns an Action Item to Ken Whistler to update UnicodeData.txt (and other data files) for 14.0.

**Date/Time:** Sun Feb 14 10:57:51 CST 2021

**Name:** Charlotte Buff

**Opt Subject:** PRI #428: General category of Znamenny priznak modifiers

The Znamenny priznak modifiers (U+1CF42..U+1CF46) were given the general category Cf (Format). A more appropriate value would be Mn (Nonspacing\_Mark) because they apply directly to the preceding character, comparable to variation selectors for instance. Other properties like bidi class and grapheme cluster break would need to be adjusted accordingly as well.

**Comment:** We agree. See Script Ad Hoc recommendations above (Section 17a).

**Date/Time:** Fri Feb 26 15:42:43 CST 2021

**Name:** Vinodh Rajan

**Opt Subject:** Sharada Code Chart

In the character list on Page 3, SHARADA VOWEL SIGN VOCALIC LL and SHARADA VOWEL SIGN E are overlapping. This needs to be fixed.

**Comment:** Based on this feedback, the Sharada font has been modified to prevent overlapping in the names list.

**Date/Time:** Fri Feb 26 15:56:05 CST 2021

**Name:** Vinodh Rajan

**Opt Subject:** Telugu Nukta Glyph in the Code Chart

As per L2/20-085, Telugu Nukta should have the combining circle below as its representative glyph to avoid confusion with the aspirate marker.

(If the current shape will be retained) The annotation “can also appear as a large dot” is moot. The glyph is already a dot.

**Comment:** The Script Ad Hoc recommended a circle glyph as shown in L2/20-085, and the character was approved by the UTC. The glyph has been fixed.

**Date/Time:** Sat Feb 27 22:19:11 CST 2021

**Name:** Norbert Lindenberg

**Opt Subject:** UAX 44: Indic data for Toto

The proposed update for UAX #44, Unicode character database, has notes on IndicPositionalCategory and IndicSyllabicCategory for the Toto script. The two data files this refers to are not available for review yet, but these notes assume that the Toto script has at least some of the characteristics of a Brahmic script that make Indic properties necessary.

According to the proposal L2/19-330, that is not the case: It states that “This Toto writing system is not syllable-based and doesn’t have an inherent vowel.” In addition, the combining class 230 for the U+1E2AE TOTO LETTER RISING TONE would be inappropriate if the script were Brahmic, as combining classes ≠ 0 are in general incompatible with the phonetic character order used for Brahmic scripts.

Comment: In general, we recommend comments on UAX #44 be made during the beta period. However, a note on the PRI #428 feedback page reports that this comment has been taken care of in the UAX #44 draft.

**Date/Time:** Mon Mar 1 15:54:54 CST 2021

**Name:** Lorna Evans

**Report Type:** Error Report

**Opt Subject:** Arabic U+089D..U+089F, U+08D0..U+08D2 have wrong property

These characters have “ON” in UnicodeData:

089D;ARABIC SUPERSCRIPT ALEF MOKHASSAS;Mn;230;ON;;;;;N;;;;;

089E;ARABIC DOUBLED MADDA;Mn;230;ON;;;;;N;;;;;

089F;ARABIC HALF MADDA OVER MADDA;Mn;230;ON;;;;;N;;;;;

and

08D0;ARABIC SUKUN BELOW;Mn;220;ON;;;;;N;;;;;

08D1;ARABIC LARGE CIRCLE BELOW;Mn;220;ON;;;;;N;;;;;

08D2;ARABIC LARGE ROUND DOT INSIDE CIRCLE BELOW;Mn;220;ON;;;;;N;;;;;

They should be “NSM”.

See Unicode proposal: <https://www.unicode.org/L2/L2019/19306-quranic-additions.pdf>

Comment: We agree. This has been fixed in the datafile and is so noted on PRI #428 feedback page.

**Date/Time:** Mon Mar 1 16:47:35 CST 2021

**Name:** Erik Carvalhal Miller

**Report Type:** Public Review Issue

**Opt Subject:** PRI #428: Comment for U+02B9

The request is to remove “primary” from the annotation (“primary stress, emphasis”). Background to the request is contained in [L2/20-286](#).

Comment: We agreed to remove “primary” from the annotation for U+02B9. (For the Script Ad Hoc’s response to the original proposal [L2/20-286](#), see the Script Ad Hoc comments in [L2/21-016r](#).)

The latest master annotation file has been fixed and will appear in the beta version of NamesList.txt. No action is required.

**Date/Time:** Tue Mar 2 13:52:43 CST 2021

**Name:** Marc Lodewijck

**Opt Subject:** PRI #428: decomposition of 107A9

107A9 MODIFIER LETTER SMALL R WITH FISHHOOK

# <sup> 207E



Decomposition of 107A9 must read <sup> 027E (instead of <sup> 207E).

**Comment:** This has been corrected in the datafile and is so noted on PRI #428 feedback page.

**Date/Time:** Sat Apr 10 18:49:00 CDT 2021

**Name:** Mikoto Ohtsuki

**Report Type:** Public Review Issue

**Opt Subject:** 1B11F-1B122 in Unicode 14.0 Alpha (PRI #428: Unicode 14.0 Alpha Review)

[Summary] This is a request to postpone encoding the four Kana characters located at U+1B11F..U+1B122 because the author considers the rationale in L2/19-381 to be insufficient. The feedback submitter considers the books that were referenced in the proposal to be written without scholarly knowledge. The submitter also recommends "ARCHAIC" in the names be reconsidered.

**Comment:** We disagree with the feedback submitter. Examples of the characters are provided in L2/19-381, with multiple sources using these characters as text. The kana could also be used for linguistic transcriptions of Japanese of foreign words. In addition, the Japanese National Body queried experts and in an email dated 24 December 2020, the Japanese NB confirmed the characters were suitable for encoding. No action is required. We did not discuss the name change.

---

---

## XI. COMMENTS AND DOCUMENTS NOT YET DISCUSSED

### Alpha Review Comments

Alpha comments from Patrik Sjöwall on Arabic (April 11) were received too late for review, and will be discussed at the next Script Ad Hoc meeting. His comment on renaming OLD POLISH O, which was also discussed by Michael Everson in feedback dated April 16, was not discussed by the Script Ad Hoc. (Michael Everson disagreed with Sjöwall on the proposed change to O ROGATE.)

Similarly, feedback from Michael Everson on COMBINING OVERCURL at 1ACF was received April 12 and a comment on Latin Extended-D was received April 16, after the Script Ad Hoc had met. These will be taken up at the next Script Ad Hoc meeting.

### Documents

The following documents have not yet been reviewed by the Script Ad Hoc, but will be discussed at a future meeting:

[L2/21-079](#) Improvements to the code chart annotations of "C0 Controls and Basic Latin" and ... Latin 1 Supplement -- Marin Silva

[N5131](#) A Supplementary Proposal to Encode the Jurchen Characters in UCS -- China NB