Script Encoding Working Group

Recommendations to UTC #182 (January 2025) on Script Proposals
Date: 2025-01-16

From: Jan Kučera, Deborah Anderson, Manish Goregaokar, Ken Whistler, Roozbeh Pournader, and Peter Constable
Based on notes by: Deborah Anderson, Manish Goregaokar, Jan Kučera, and Lawrence Wolf-Sonkin

Also participating: Craig Cummings, Lorna Evans, Liang Hai, Ned Holbrook, Denis Moyogo Jacquerye, Kushim Jiang, James Kirby, Robin Leroy, Leona Lin, Steph Toyofuku Maclean, Biswajit Mandal, Kirk Miller, Karl Pentzlin, CheonHyeong Sim, Steve Tinney, Daniel Yacob, Lawrence Wolf-Sonkin, Ben Yang, and Oreen Yousuf

The Script Encoding Working Group met on November 15 and December 13, 2024, and January 3, 2025, in order to review proposals. This document represents feedback on proposals that were available when the group met.

Table of Contents

1. New Characters

1.1 Twelve cuneiform tenû numerals #591

Documents:

      L2/24-270: Twelve cuneiform tenû numerals — Leroy

Recommendation:

1. **[182-C??] Consensus** : Provisionally assign 12 code points U+1246F, U+12475..U+1247F to Cuneiform tenû numerals as described in the file UnicodeData.txt attached to in L2/24-270. [Ref: 1.1 in L2/25-010]

2. **[182-A??] Action Item** for Ken Whistler, RMG: Update the Pipeline to include 12 provisionally assigned code points U+1246F, U+12475..U+1247F for Cuneiform tenû numerals as described in L2/24-270. [Ref: 1.1 in L2/25-010]

3. **[182-A??] Action Item** for Robin Leroy, SAH: Send a font for U+1246F, U+12475..U+1247F (tenû numerals) to Michel Suignard. [Ref: 1.1 in L2/25-010]

Comments:

This document proposes filling the Cuneiform Numbers and Punctuation block with twelve cuneiform numerals used in the third millennium. Three of those are additional numerals in the AŠ (or DIŠ) tenû series, 7–9, while 1-6 are already encoded. The other proposed characters constitute a new series of numerals. Evidence from primary sources was provided for all proposed characters.

We were advised that there are up to 47 other cuneiform numerals yet to be proposed.

1.2 Hebrew Point Sheva Na' #537

Documents:

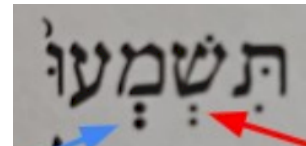L2/24-274: Proposal to encode Hebrew Point Sheva Na — Mosesson

Related:

L2/16-086: Proposal of an additional character to the Hebrew Unicodes: Sheva Na — May

Recommendation:

1. **[182-C??] Consensus** : Provisionally assign one code point U+05C8 HEBREW POINT HEAVY SHEVA based on L2/24-274. [Ref: 1.2 in L2/25-010]

2. **[182-A??] Action Item** for Ken Whistler, RMG: Update the Pipeline to include one provisionally assigned code point U+05C8 HEBREW POINT HEAVY SHEVA based on L2/24-274. [Ref: 1.2 in L2/25-010]

3. **[182-A??] Action Item** for Debbie Anderson, SAH: Work with the proposal author to send a font for U+05C8 HEBREW POINT HEAVY SHEVA to Michel Suignard. [Ref: 1.2 in L2/25-010]

Comments:

The Hebrew vowel sign U+05B0 HEBREW POINT SHEVA can represent either a vocal šewa (naʾ) or a silent šewa (naḥ). Some publishers have been differentiating the two with different glyphs, usually using a bold one or one with bigger dots for the vocal šewa. It is noted that a mark cannot typically be of a different font style from its base character even in rich text, unless a specialized rendering engine is designed to handle this case.

The proposal therefore seeks disunifying the U+05B0 HEBREW POINT SHEVA sign and encoding a separate character for the vocal šewa, which is not as commonly used. This has already been proposed in 2016 by a different author. However, that document included only one unsourced electronic evidence. This proposal contains several printed publications from various publishers and therefore meets the criteria for encoding. The Standards Institution of Israel expressed their support in encoding the new character.

The proposal author suggested the name HEBREW POINT SHEVA NA for the new character, arguing that SHEVA commonly refers to both šewa naʾ and šewa naḥ and when a distinction is needed, it is always šewa naʾ that is spelled fully. This would be consistent with the similar situation of U+05B8 HEBREW POINT QAMATS and U+05C7 HEBREW POINT QAMATS QATAN. The group was not convinced by these arguments and worried about the name SHEVA NA causing confusion about which šewa it represents due to similarity in their names, especially for users with less expertise.

1.3 Devanagari letter Sindhi DDDA #592

Documents:

L2/25-022: Proposal to encode the Devanagari letter Sindhi DDDA in Unicode — Mandal

Related:

L2/05-082: Proposal to add four characters for Sindhi — Everson

Recommendation:

1. **[182-C??] Consensus** : Provisionally assign one code point U+11B0A DEVANAGARI LETTER ALTERNATE DDDA as described in L2/25-022. [Ref: 1.3 in L2/25-010]

2. **[182-A??] Action Item** for Ken Whistler, RMG: Update the Pipeline to include one provisionally assigned code point U+11B0A DEVANAGARI LETTER ALTERNATE DDDA as described in L2/25-022. [Ref: 1.3 in L2/25-010]

3. **[182-A??] Action Item** for Debbie Anderson, SAH: Work with the proposal author to send a font for U+11B0A DEVANAGARI LETTER ALTERNATE DDDA to Michel Suignard. [Ref: 1.3 in L2/25-010]

4. **[182-A??] Action Item** for Ken Whistler, EDC: Annotate U+11B0A DEVANAGARI LETTER ALTERNATE DDDA to include cross reference to U+0926 DEVANAGARI LETTER DA, and place it under separate "Sindhi implosive" section header. [Ref: 1.3 in L2/25-010]

Comments:

The Sindhi language is written in both Arabic and Devanagari scripts. In the Devanagari script, Sindhi implosives tend to use a horizontal line at the bottom, e.g. ग़ U+097B DEVANAGARI LETTER GGA, ॼ U+097C DEVANAGARI LETTER JJA etc. This document proposes encoding another such letter, based on द U+0926 DEVANAGARI LETTER DA.

In Maharashtra, this letter is used for the alveolar implosive in Sindhi. In Madhya Pradesh, the already encoded ॾ U+097E DEVANAGARI LETTER DDDA (based on ड U+0921 DEVANAGARI LETTER DDA) is used instead for the same phoneme. A question therefore came up whether these should not be a glyph variant of each other, but the letters have clearly two different identities, used in two different orthographies.

To avoid having too many Ds in the name, we recommend the name DEVANAGARI LETTER ALTERNATE DDDA.

1.4 Devanagari Sign AAO #593

Documents:

L2/25-021: Proposal to encode the Devanagari Vowel Sign AAO in Unicode — Mandal

Recommendation:

1. **[182-A??] Action Item** for Jan Kučera, EDC: Propose a text addition for the Core Specification pointing out that multiple vowel signs on the same side are allowed in one cluster in the Indic model. [Ref: 1.4 in L2/25-010]

2. **[182-A??] Action Item** for Jan Kučera, SAH: Follow up with the author of L2/25-021 on the issues with representing Devanagari Vowel Sign AAO. [Ref: 1.4 in L2/25-010]

Comments:

This is a request to encode a character in the Devanagari script used to write Avestan. The requested character, a vowel sign AAO, can be currently written using the sequence of U+093E DEVANAGARI VOWEL SIGN AA and U+094B DEVANAGARI VOWEL SIGN O. The proposal author claims that this sequence is not allowed in standard fonts, but that does not seem to be correct. Multiple right-sided vowel signs are explicitly allowed in the Unicode Indic model and members confirmed that the sequence renders correctly at least in the latest versions of Core Text, DirectWrite and Harfbuzz.

This usage in Avestan has been known to Unicode since at least 2007, see L2/07-396 and L2/07-397. It appears the proposed character and the corresponding independent vowel have later been removed from the updated proposals L2/08-042 and L2/08-050, due to an agreement that they are already representable in Unicode. Other characters needed for transcribing Avestan were kept in L2/08-050 and got encoded. We do not see any compelling reason why this decision should be revisited and a confusable character encoded. As such, we recommend no decision to make encoding changes.

1.5 Character additions for Initial Teaching Alphabet #224

Documents:

L2/24-273: Unicode request for Initial Teaching Alphabet — Miller

Related:

L2/23-102R: Unicode request for Initial Teaching Alphabet (revised) — Miller
L2/22-286: Proposal to Encode Latin characters for Initial Teaching Alphabet — Manulov
L2/08-428: Exploratory proposal to encode Germanicist, Nordicist, and other phonetic characters (WG2 N3555) — Everson

Recommendation:

1. **[182-C??] Consensus** : Provisionally assign 13 code points U+1DF5A..U+1DF66 to Initial Teaching Alphabet as described in L2/24-273. [Ref: 1.5 in L2/25-010]

2. **[182-A??] Action Item** for Ken Whistler, RMG: Update the Pipeline to include 13 provisionally assigned code points U+1DF5A..U+1DF66 for Initial Teaching Alphabet as described in L2/24-273. [Ref: 1.5 in L2/25-010]

3. **[182-A??] Action Item** for Kirk Miller, SAH: Send a font for U+1DF5A..U+1DF66 (ITA) to Michel Suignard. [Ref: 1.5 in L2/25-010]

4. **[182-A??] Action Item** for Ken Whistler, EDC: Update annotations for 9 existing Latin characters and for U+1DF5A, U+1DF5E, and U+1DF5F as described in L2/24-273. [Ref: 1.5 in L2/25-010]

Comments:

This is the final version of a proposal to encode additional characters for Initial Teaching Alphabet (ITA), which was used for literacy materials in the UK in the 1960s, whence it spread to Austria and the US. The alphabet is an extension to the Latin alphabet and many of the already encoded Latin characters can be used to represent ITA letters either directly or as a variant in font. This document proposes encoding the remaining few, showing them realized in several fonts and endorsed by the ITA Foundation.

1.6 English Phonotypic Alphabet #456

Documents:

L2/24-277: Fifth Revised Proposal to encode characters for the English Phonotypic Alphabet (EPA) in the UCS — Pentzlin

Recommendation:

1. **[182-C??] Consensus** : Provisionally assign 31 code points U+2E60..U+2E63, U+A7DD and U+1DF68..U+1DF81 to characters for EPA as described in L2/24-277. [Ref: 1.6 in L2/25-010]

2. **[182-A??] Action Item** for Ken Whistler, RMG: Update the Pipeline to include 31 provisionally assigned code points U+2E60..U+2E63, U+A7DD and U+1DF68..U+1DF81 for EPA as described in L2/24-227. [Ref: 1.6 in L2/25-010]

3. **[182-A??] Action Item** for Debbie Anderson, SAH: Work with the proposal author to send a font for U+2E60..U+2E63, U+A7DD and U+1DF68..U+1DF81 (EPA) to Michel Suignard. [Ref: 1.6 in L2/25-010]

Comments:

This is a final proposal to encode 24 letters and 4 punctuation symbols used by the English Phonotypic Alphabet, a precursor of modern IPA, used in the second half of the 19th century. Several iterations of the proposal were submitted by the author to the UTC in 2010 and 2011.

The alphabet has had different case pair mappings through its lifetime and users will not be able to rely on automatic casing no matter how this is approached. We had therefore asked the author to minimize introducing casing pairs to those cases justified by evidence. One upper letter is introduced forming a pair with an existing character, U+0277 LATIN SMALL LETTER CLOSED OMEGA, and hence placed into the same plane. The group is satisfied that all feedback has been addressed and recommends encoding.

1.7 RFE letters #561

Documents:

L2/25-016: Unicode request for letters of the RFE phonetic alphabet — Miller

Recommendation:

1. **[182-C??] Consensus** : Provisionally assign 3 code points U+1DF57..U+1DF59 to RFE letters and 3 code points U+1DFCD..U+1DFCF for RFE modifier letters as described in L2/25-016. [Ref: 1.7 in L2/25-010]

2. **[182-A??] Action Item** for Ken Whistler, RMG: Update the Pipeline to include 6 provisionally assigned code points U+1DF57..U+1DF59 for RFE letters and U+1DFCD..U+1DFCF for RFE modifier letters as described in L2/25-016. [Ref: 1.7 in L2/25-010]

3. **[182-A??] Action Item** for Kirk Miller, SAH: Send a font for U+1DF57..U+1DF59 and U+1DFCD..U+1DFCF (RFE) to Michel Suignard. [Ref: 1.7 in L2/25-010]

4. **[182-A??] Action Item** for Ken Whistler, EDC: Add annotations for U+1DF57 LATIN SMALL LETTER TURNED R WITH MID-HEIGHT LEFT HOOK and U+1DF58 LATIN SMALL LETTER SPLIT O as described in L2/25-016. [Ref: 1.7 in L2/25-010]

Comments:

This phonetic alphabet, Revista de Filología Española (RFE), is used in transcription and linguistic works in Spanish-speaking countries, particularly Spain and Mexico. Most letters in RFE are already in Unicode, but a few letters for lax and reduced sounds are missing.

The proposal contains two pairs of characters with the word SPLIT in their name. An alternative term, BROKEN, was considered, which would match the usage in ⊘ (U+238B BROKEN CIRCLE WITH NORTHWEST ARROW). However, in letters such as Ł (U+A746 LATIN CAPITAL LETTER BROKEN L), the term BROKEN refers to a different feature. We therefore prefer SPLIT as proposed. A couple of already encoded characters in the Sutton SignWriting block contain the word SPLIT in their name (U+1D848, U+1D861..U+1D865), although that refers to a configuration of fingers they represent.

1.8 Phonetic: Unicode request for lezh with curl #586

Documents:

L2/24-272: Unicode request for lezh with curl — Miller

Related:

L2/24-050: Unicode request for letters with palatal hook — Miller

Recommendation:

1. **[182-C??] Consensus** : Provisionally assign one code point U+1DF67 LATIN SMALL LETTER LEZH WITH CURL as described in L2/24-272. [Ref: 1.8 in L2/25-010]

2. **[182-A??] Action Item** for Ken Whistler, RMG: Update the Pipeline to include one provisionally assigned code point U+1DF67 LATIN SMALL LETTER LEZH WITH CURL as described in L2/24-272. [Ref: 1.8 in L2/25-010]

3. **[182-A??] Action Item** for Kirk Miller, SAH: Send a font for U+1DF67 LATIN SMALL LETTER LEZH WITH CURL to Michel Suignard. [Ref: 1.8 in L2/25-010]

Comments:

A previous proposal (L2/24-050) presented a chart with letters with palatal hook, based on the last IPA consonant chart. Several of these letters were unattested and marked as hypothetical. This document presents new evidence showing one of these characters in use and proposes to encode it. The letter name uses CURL rather than PALATAL HOOK as a curl was used on letters with a curved tail.

2. New Symbols

2.1 Music: Miscellaneous musical symbols #560

Documents:

> L2/25-017: Unicode request for miscellaneous musical symbols — Bala and Miller

Recommendation:

1. **[182-C??] Consensus** : Provisionally assign 30 code points U+1D262..U+1D27F to musical symbols as described in L2/25-017. [Ref: 2.1 in L2/25-010]
2. **[182-A??] Action Item** for Ken Whistler, RMG: Update the Pipeline to include 30 provisionally assigned code points U+1D262..U+1D27F for musical symbols as described in L2/25-017. [Ref: 2.1 in L2/25-010]
3. **[182-A??] Action Item** for Kirk Miller, SAH: Send a font for U+1D262..U+1D27F (miscellaneous musical symbols) to Michel Suignard. [Ref: 2.1 in L2/25-010]

Comments:

We reviewed a request for encoding 30 miscellaneous musical symbol characters. All symbols are referred to in evidence and the document cites the PUA code points for the characters in W3C Standard Music Font Layout (SMuFL).

2.2 Music: Additional Baroque ornament #533

Documents:

> L2/25-015: Unicode request for additional Baroque ornament — Bala and Miller

Recommendation:

1. **[182-C??] Consensus** : Provisionally assign one code point U+1D261 MUSICAL SYMBOL SUSPENSION as described in L2/25-015. [Ref: 2.2 in L2/25-010]
2. **[182-A??] Action Item** for Ken Whistler, RMG: Update the Pipeline to include one provisionally assigned code point U+1D261 MUSICAL SYMBOL SUSPENSION as described in L2/25-015. [Ref: 2.2 in L2/25-010]
3. **[182-A??] Action Item** for Kirk Miller, SAH: Send a font for U+1D261 MUSICAL SYMBOL SUSPENSION as described in L2/24-272. to Michel Suignard. [Ref: 2.2 in L2/25-010]

Comments:

We reviewed a request for encoding a Baroque musical ornament denoting suspension. The document includes evidence from two different composers, both in print and in hand. The symbol is also defined in SMuFL.

3. Changes to Characters

3.1 Two Tangut Ideographs #555

Documents:

> L2/25-040: Feedback on L2/24-236 — West

Related:

> L2/24-236: Proposal to encode two Tangut ideographs (WG2 N5286) — Chan et al.

Recommendation:

1. **[182-C??] Consensus** : Update the reference glyph for U+18D1E TANGUT IDEOGRAPH-18D1E and its Radical Stroke value as described in L2/25-040, for Unicode 17.0. [Ref: 3.1 in L2/25-010]
2. **[182-A??] Action Item** for Debbie Anderson, SAH: Work with the proposal author to send an updated font for U+18D1E TANGUT IDEOGRAPH-18D1E to Michel Suignard. [Ref: 3.1 in L2/25-010]
3. **[182-A??] Action Item** for Michel Suignard, Charts: Update the reference glyph for U+18D1E TANGUT IDEOGRAPH-18D1E for Unicode 17.0 as per L2/25-040. [Ref: 3.1 in L2/25-010]
4. **[182-A??] Action Item** for Michel Suignard, PAG: Update the Radical Stroke index of U+18D1E TANGUT IDEOGRAPH-18D1E from 579.14 to 579.15 for Unicode 17.0 as per L2/25-040. [Ref: 3.1 in L2/25-010]

Comments:

This is a request from Andrew West to change the glyph for U+18D1E, one of the two Tangut characters approved for 17.0. With this change, the Radical Stroke value will need to be changed from 579.14 to 579.15. Dr. Sun Bojun, one of the main Tangut experts from China, expressed agreement with the proposal regarding both the glyph change and Radical Stroke value.

3.2 Adlam Names List Modifications #564

Documents:

> L2/24-204: Proposal for Adlam Names List Modifications — Patel

Recommendation:

1. **[182-A??] Action Item** for Ken Whistler, EDC: Consider the proposed annotations for Adlam characters as described in L2/24-204. [Ref: 3.2 in L2/25-010]

Comments:

This proposal argues that the names of characters for Adlam are not commonly used in the community. The author proposes adding short names as an alias to all of the letters. These names were designed for children to be easier to learn and recite the alphabet. It also suggests a clarification of annotation to U+1E94B ADLAM NASALIZATION MARK.

We do not always publish informative aliases reflecting how a community calls letters and the Unicode Standard is not intended as a pedagogical tool. Some precedents exist nevertheless, such as in the name lists for Balinese or Old South Arabian.

The SEW group does not have any objections to the proposed name list aliases on technical grounds and suggests any changes be left to the discretion of the EDC.

3.3 Phonetic: A correction of U+1AB3 (combining downwards 'arrow') #345

Documents:

> L2/25-019: A note on U+1AB3 combining downwards arrow — Miller

Related:

> L2/11-202: Revised proposal to encode "Teuthonista" phonetic characters — Everson et al.

Recommendation:

1. **[182-A??] Action Item** for Ken Whistler, EDC: Add annotations for U+1AB3 COMBINING DOWNWARDS ARROW ABOVE as described in L2/25-019. [Ref: 3.3 in L2/25-010]
2. **[182-A??] Action Item** for Ken Whistler, EDC: Document the use of U+1AB3 in section 7.9.2 (Combining Diacritical Marks Extended) of the Core Specification as suggested in L2/25-019, section "Core Spec". [Ref: 3.3 in L2/25-010]

Comments:

The encoded character U+1AB3 COMBINING DOWNWARDS ARROW is based on a single evidence in L2/11-202, where a downwards arrow mark is used instead of what is supposed to be a compound of a combining caron and acute, as argued in this proposal. The author was seeking a way to remedy this situation and to provide guidance on the correct usage. Noting that documents might have been created where this character is intended to be an

arrow due to its name, the group suggested that for the combined diacritical mark, the sequence of U+030C COMBINING CARON and U+0301 COMBINING ACUTE ACCENT is used, and the standard is updated to provide corresponding guidance. The document proposes changes to this effect.

3.4 More misspellings in Arabic honorifics #484 · #606

Documents:

> L2/25-041: Feedback on the names of Arabic Honorifics — Al-Shahrabi and Kučera

Related:

> L2/24-077: Proposal to encode twenty-five more Arabic honorifics — Pournader et al.
> L2/24-002: Proposal to encode eight more Arabic honorifics — Pournader et al.

Recommendation:

1. **[182-C??] Consensus** : Change the name of provisionally assigned characters U+FBCB from ARABIC LIGATURE RADI ALLAHU TAAALAA ANHUM to ARABIC LIGATURE RADI ALLAAHU TAAALAA ANHUM, U+10ED5 from ARABIC LIGATURE QUDDISA SIRRAHUM to ARABIC LIGATURE QUDDISA SIRRUHUM, U+10ED6 from ARABIC LIGATURE QUDDISA SIRRAHUMAA to ARABIC LIGATURE QUDDISA SIRRUHUMAA, for Unicode 17.0. [Ref: 3.4 in L2/25-010]
2. **[182-A??] Action Item** for Ken Whistler, RMG: Update the Pipeline to reflect the renaming of provisionally assigned characters U+FBCB, U+10ED5 and U+10ED6. [Ref: 3.4 in L2/25-010]

Comments:

We received further feedback on the names of provisionally accepted (179-C49) Arabic honorifics approved for publication in Unicode 17.0 (181-C60) from a native Arabic speaker. The feedback questioned the spelling of SIRRAHUM and SIRRAHUMMA in U+10ED5 resp. U+10ED6, and the spelling of ALLAHU in U+FBCB. A confirmation by four other experts, Mustafa Jibaly, Ayman Aldahleh, Adil Allawi and Khaled Hosny, contacted by Roozbeh Pournader, has resulted in certainty about grammatical and spelling mistakes in the name of these three characters.

The authors also pointed out inconsistency in transcribing kasra and other Arabic marks in the standard. As for the case of kasra in ALAYHI vs. AALIHEE in the name of U+FDCC proposed in L2/24-002, it is noted that this is intentional, reflecting a connective elongation in Quranic Arabic. As for the other inconsistencies, these are a result of Unicode changing the system for transcribing Arabic on several occasions, as well as pushback from WG2.

4. Other

4.1 DoNotEmit for Arabic tashkil #596

Documents:

> L2/25-025: DoNotEmit for Arabic tashkil — Pournader

Recommendation:

1. **[182-A??] Action Item** for Roozbeh Pournader, SAH: Add Arabic Tashkil sequences to DoNotEmit.txt as described in L2/25-025. [Ref: 4.1 in L2/25-010]
2. **[182-A??] Action Item** for Ken Whistler, EDC: Document the discouragement of using pairs of fathas and kasras in the Tashkil Nonspacing Section in chapter 9.2.1 of the Core Specification as suggested in L2/25-025. [Ref: 4.1 in L2/25-010]

Comments:

This proposal seeks to add entries into the DoNotEmit.txt data file and a corresponding explanation in the Core Specification. In Arabic, the character U+064B ARABIC FATHATAN is visually indistinguishable from the sequence of two U+064E ARABIC FATHA characters. Similarly, the character U+064D ARABIC KASRATAN is indistinguishable from the sequence of two U+0650 ARABIC KASRA characters. These characters do not have a decomposition to the respective sequences, which could be considered invalid. Due to the visual confusability in these two cases, we recommend implementations output the fathatan and kasratan characters in favor of the sequences. This does not apply to U+064C ARABIC DAMMATAN, since it visually differs from the sequence of two dammas in most fonts and styles of writing.

5. Public Feedback

5.1 On the naming on some musical symbols #445 · #602

Documents:

> ID20241227092147: Comments on Public Review Issues (July 2, 2024 - Oct 16, 2024)
>
> L2/24-174: Unicode request for Turkish and Arabic accidentals — Bala and Miller

Recommendation:

> No action is requested of the UTC.

Comments:

This feedback did not come in time for it to be reviewed by the proposal authors and will be taken up at a future meeting.

5.2 L2/25-021 (DEVANAGARI VOWEL SIGN AAO)

Documents:

> ID20250107184203: Comments on Public Review Issues (July 2, 2024 - Oct 16, 2024)
>
> L2/25-021: Proposal to encode the Devanagari Vowel Sign AAO in Unicode — Mandal

Recommendation:

1. **[182-A??] Action Item** for Jan Kučera, SAH: Respond to David Corbett regarding their feedback dated Tue Jan 07 18:42:03 CST 2025 with comments from the SEW report. [Ref: 5.2 in L2/25-010]

Comments:

This feedback proposes an alternative solution (i.e. fixing shaping engines) to encoding DEVANAGARI VOWEL SIGN AAO as suggested in L2/25-021. We agree with the feedback author in principle, but since we are not recommending encoding of the character (see above), this feedback does not require any action.

5.3 On the new buzz roll combining character #447 · #601 · #608

Documents:

L2/25-005: Comments on Public Review Issues (Oct 24, 2024 - Jan 2, 2025) — Michelle Perham
ID20241227110909: Comments on Public Review Issues (July 2, 2024 - Oct 16, 2024)

Related:

L2/24-213: Unicode request for additional tremoli — Bala and Miller

Recommendation:

1. **[182-A??] Action Item** for Manish Goragaokar, SAH: Respond to Eduardo Marín Silva regarding their feedback dated Fri Dec 27 11:09:09 CST 2024 with comments from the SEW report. [Ref: 5.3 in L2/25-010]

Comments:

We reviewed this feedback and agreed with the author that U+1D25F should have a different glyph (with the "z" on a stem) and name (MUSICAL SYMBOL COMBINING BUZZ ROLL STEM). A formal request to that effect from Kirk Miller, author of the L2/24-213, will be taken up at a future meeting in a separate document. As with U+1D166 MUSICAL SYMBOL COMBINING SPRECHGESANG STEM, the stem can be suppressed by musical layout software.

6. No Action Required

6.1 Arabic Letter Yeh Barree with Small V #595

Documents:

L2/25-023: Preliminary Proposal for Yeh Barree With Small V — Pournader

Recommendation:

No action is requested of the UTC.

Comments:

We received a preliminary proposal for an Arabic letter Yeh Barree with Small V used in the Shina language in the India-administered region of Jammu and Kashmir. While there are potentially several missing characters, the YEH BARREE WITH SMALL V was present in three independent primers in the language. We are seeking feedback from experts and contacts in the community.

6.2 Mongolian: Proposal to Encode One Manchu Letter #371

Documents:

L2/24-025: Proposal to Encode One Manchu Letter — Sim

Recommendation:

No action is requested of the UTC.

Comments:

We discussed recent comments from Kushim Jiang on U+1879 MONGOLIAN LETTER MANCHU ALTERNATE UE, which was provisionally assigned at the January 2024 UTC meeting (178-C30), but removed from 16.0 after a request from China national body at the 2024 WG2 meeting. In discussions with the Manchu Working Group in China, Kushim mentioned that the character is used for transliteration of traditional Mongolian. This transliteration system is stable and has been studied by scholars. Kushim's recommendation is to change the name and perhaps provide an annotation that the character is used for transliteration. With such a change, there may be no objection from the China national body about encoding the character.

The group recommended the proposal be revised, proposing the name MONGOLIAN LETTER ALTERNATE UE and an annotation about its use in transliteration of traditional Mongolian. We are awaiting an updated document.

6.3 Tulu-Tigalari numerals #599

Documents:

L2/25-020: Proposal to encode Tulu-Tigalari numerals — Mandal

Related:

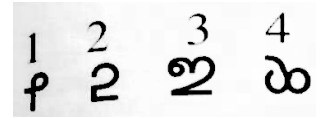L2/23-021: Additional Document for Unicode Tulu Proposal — Jain
L2/22-031: Updated proposal to encode the Tulu-Tigalari script in Unicode (replaces L2/21-210) — Murthy and Rajan

Recommendation:

No action is requested of the UTC.

Comments:

This is a proposal to encode digits for Tulu-Tigalari. The original proposal for Tulu-Tigalari L2/22-031 notes the existence of such digits, but the authors only found a single manuscript with evidence and suggested more research be done. This proposal includes evidence in modern use and assumes unification of the modern Tulu and historical Tulu-Tigalari orthography. The KTSA proposal L2/23-021 also included native digits with evidence from one manuscript, albeit with slightly different shapes. KTSA has requested us to hold off encoding the digits, so we recommend no action at the moment. We encourage as much manuscript evidence as possible to be included in any future document proposing the encoding of digits for Tulu or Tulu-Tigalari.

6.4 False unifications of legacy computing box drawing characters #588

Documents:

L2/25-037: Addressing false unifications of legacy computing box drawing characters — Grochowski

Related:

L2/21-235: Proposal to add further characters from legacy computers and teletext [see also version without attachments, sources spreadsheet, mappings Zip file] (revised) — Bettencourt et al.
L2/19-025: Proposal to add characters from legacy computers and teletext (replaces L2/18-275) [see also version without attachments, mapping files, mapping spreadsheet] — Ewell et al.
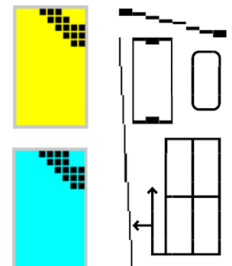
Recommendation:

No action is requested of the UTC.

Comments:

We received a proposal requesting a disunification of some of the legacy computing box drawing characters. The proposal seeks encoding a character that is one pixel different from U+1CE2B LARGE TYPE PIECE DIAGONAL UPPER RIGHT, and encoding 26 characters of box drawing lines for PETSCII and Apple II whose glyphs differ in thickness from the characters already encoded in the Symbols For Legacy Computing and Box Drawing blocks.

We deem the differences demonstrated in the proposal to not constitute differences in plain text. No evidence of a document that would make a distinction between the corresponding characters in the different code pages was provided.

The request appears to come from a misunderstanding of the standard. The document is making glyph distinctions, while the Unicode standard is based on character identities. It is not a goal of the standard to harmonize PETSCII and Apple II code pages. The issues raised in the proposals can be solved by using appropriate fonts and as such no action is recommended to be taken.

6.5 Feedback on L2/24-134 #228 · #603

Documents:

L2/25-024: Feedback on L2/24-134 Proposal to Encode a Pridnestrovian Ruble Sign — West

Related:

L2/25-030: Feedback on L2/24-134 Proposal to Encode a Pridnestrovian Ruble Sign in response to L2/24-271 and L2/25-024 — Pentzlin
L2/24-271: Request for SC2 member comments on proposal to encode Ruble sign with double vertical stem — Constable
L2/24-134R: Proposal to Encode a Pridnestrovian Ruble Sign (revised) — Adrianov

Recommendation:

No action is requested of the UTC.

Comments:
This is a proposal to introduce criteria for encoding currency symbols, namely by requiring an official government agency or a relevant national body to author or support such proposals, and by requiring a currency code to have been assigned to the currency in ISO 4217. The proposal does not contain any rationale for these criteria or explanation how adopting them would improve the Unicode standard and ISO/IEC 10646. However, it notes that the proposed criteria would disqualify the provisionally assigned U+20C1 RUBLE SIGN WITH DOUBLE VERTICAL STEM. However, it would also disqualify U+060B AFGHANI SIGN and others.

This document assumes that currencies are only created by nation states, and overlooks that there are often currencies created by private or administrative entities other than nation states recognized by the United Nations. During World War II, special Allied military currencies were created for use in occupied/liberated countries to provide local economic stability. Within countries that have a national currency, local governments sometimes create a local currency for local economic reasons. Recent examples of local currencies in the UK include the Brixton pound, the Bristol pound, and the Exeter pound; BerkShares is an example of a local currency currently used within the US. In the US, native American communities are independent nations that could create their own currencies but are not nation states recognized by the UN and whose currencies would not be candidates for coding in ISO 4217; Maza is an example of a US tribal currency. Other examples of active local currencies include Eko, Lewes pound, Stroud pound, Chiemgauer, Wirk and Eusko.

The proposal seems to assume that, by encoding, we are endorsing a currency. That is not the case. We encode text elements that are used in human cultures throughout the world in text communication. Currency symbols are text elements that are distinguished from other symbols mainly because they have particular behaviour in text processing, such as bidi layout or line breaking, not because they are controlled by certain kinds of institutions. The group recommends no action.

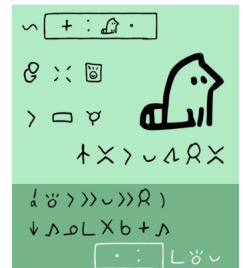## 6.6 Sitelen Pona <small>#466</small>

Documents:
> L2/24-275: Preliminary proposal to encode sitelen pona in the UCS — Sitelen Pona Working Group

Recommendation:
> No action is requested of the UTC.

Comments:
This is a preliminary proposal to encode Sitelen Pona, a script for the toki pona language constructed in 2001. While we appreciate there is a sizable community of speakers of the language, some members were not convinced a need for interchange of text in the script was demonstrated. However, the biggest concern of the group is regarding the stability of the script. While the language was designed to be a limited set of words, the user communities are being inventive and there is no indication that repertoire of the script is stable, or that it is logographic in nature (which would make it constrained to the language). Until the users of the script themselves consider standardizing what the script actually is and can provide an authoritative document or sign list that can be referred to, it is not ready to be included in an international standard.

6.7 Sibe Quotation Marks <small>#369</small>

Documents:
> L2/25-028: Proposal to Add VS3 for Sibe Quotation Marks — Sim

Related:
> L2/25-009: CJK & Unihan Working Group Recommendations for UTC #182 Meeting — Lunde

Recommendation:
> The SEW group has no objections to the recommendations regarding L2/25-028 as proposed by the CJK group in L2/25-009.

Comments:
We reviewed an earlier version of this proposal and provided feedback that has been incorporated.

6.8 KATAKANA DIGRAPH YORI <small>#604</small>

Documents:
> L2/24-279: Proposal to Encode One Kana Ligature — Sim

Related:
> L2/25-009: CJK & Unihan Working Group Recommendations for UTC #182 Meeting — Lunde
> L2/24-150: Proposal for missing three Kana-Ligatures — Kojitani
> L2/23-112: Proposal for missing Kana-Ligatures — Kojitani

Recommendation:

The SEW group has no objections to the recommendations regarding L2/24-279 as proposed by the CJK group in L2/25-009.

7. In Process

Documents:

L2/25-038: Proposal to encode Sirmauri in Unicode — Pandey and Mandal

L2/25-039: Revised Proposal for Encoding the Mwangwego Script in the UCS — Yousuf and Yacob

L2/25-043: On the decision to rename and change the glyphs of Greek phonetic letters — Silva

Appendix

List of documents covered by this proposal:

L2/24-025: Proposal to Encode One Manchu Letter — Sim

L2/24-174: Unicode request for Turkish and Arabic accidentals — Bala and Miller

L2/24-204: Proposal for Adlam Names List Modifications — Patel

L2/24-270: Twelve cuneiform tenû numerals — Leroy

L2/24-272: Unicode request for lezh with curl — Miller

L2/24-273: Unicode request for Initial Teaching Alphabet — Miller

L2/24-274: Proposal to encode Hebrew Point Sheva Na — Mosesson

L2/24-275: Preliminary proposal to encode sitelen pona in the UCS — Sitelen Pona Working Group

L2/24-277: Fifth Revised Proposal to encode characters for the English Phonotypic Alphabet (EPA) in the UCS — Pentzlin

L2/24-279: Proposal to Encode One Kana Ligature — Sim

L2/25-005: Comments on Public Review Issues (Oct 24, 2024 - Jan 2, 2025) — Michelle Perham

L2/25-015: Unicode request for additional Baroque ornament — Bala and Miller

L2/25-016: Unicode request for letters of the RFE phonetic alphabet — Miller

L2/25-017: Unicode request for miscellaneous musical symbols — Bala and Miller

L2/25-019: A note on U+1AB3 combining downwards arrow — Miller

L2/25-020: Proposal to encode Tulu-Tigalari numerals — Mandal

L2/25-021: Proposal to encode the Devanagari Vowel Sign AAO in Unicode — Mandal

L2/25-021: Proposal to encode the Devanagari Vowel Sign AAO in Unicode — Mandal

L2/25-022: Proposal to encode the Devanagari letter Sindhi DDDA in Unicode — Mandal

L2/25-023: Preliminary Proposal for Yeh Barree With Small V — Pournader

L2/25-024: Feedback on L2/24-134 Proposal to Encode a Pridnestrovian Ruble Sign — West

L2/25-025: DoNotEmit for Arabic tashkil — Pournader

L2/25-028: Proposal to Add VS3 for Sibe Quotation Marks — Sim

L2/25-037: Addressing false unifications of legacy computing box drawing characters — Grochowski

L2/25-038: Proposal to encode Sirmauri in Unicode — Pandey and Mandal

L2/25-039: Revised Proposal for Encoding the Mwangwego Script in the UCS — Yousuf and Yacob

L2/25-040: Feedback on L2/24-236 — West

L2/25-041: Feedback on the names of Arabic Honorifics — Al-Shahrabi and Kučera

L2/25-043: On the decision to rename and change the glyphs of Greek phonetic letters — Silva

List of feedback covered by this proposal:

ID20241227092147: Comments on Public Review Issues (July 2, 2024 - Oct 16, 2024)

ID20241227110909: Comments on Public Review Issues (July 2, 2024 - Oct 16, 2024)

ID20250107184203: Comments on Public Review Issues (July 2, 2024 - Oct 16, 2024)