L2/24-166

TO: **UTC**
FROM: **Deborah Anderson, Manish Goregaokar, Jan Kučera, Ken Whistler, Roozbeh Pournader, and Peter Constable** [1]
SUBJECT: **Recommendations to UTC #180 (July 2024) on Script Proposals**
DATE: **2024-07-18**

The Script Encoding Working Group met on May 17, June 4, and July 5, 2024, in order to review proposals. The following represents feedback on proposals that were available when the group met.

Table of Contents

[1] Also participating were Basit Ali, Nat Baca, Craig Cummings, Quinn Dombrowski, Lorna Evans, Karljurgen Feuerherm, Audrey Gao, Andrew Glass, Liang Hai, Ned Holbrook, Anushah Hossain, John Hudson, Robin Leroy, Norbert Lindenberg, Kamal Mansour, Kirk Miller, Denis Moyogo Jacquerye, Cheon Hyeong Sim, Michel Suignard, Harald Tveiten, Lawrence Wolf-Sonkin, and Ben Yang. The text for the comments and recommendations was based on notes primarily by Manish Goregaokar with assistance from Debbie Anderson, Jan Kučera, Robin Leroy, and Quinn Dombrowski.

A. PROPOSALS REQUIRING UTC ACTION

I. AFRICAN SCRIPTS

1. Egyptian Hieroglyphs

1a Rotations [#378]

Document: L2/24-177 Additional Variation Selectors as Substitutes for Legacy v5.2–v15 Hieroglyphs -- Werning et al.
Background doc: L2/24-175 Rationale for commenting out the variation sequence U+1333B U+FE00 for Ancient Egyptian -- Dils et al.

Recommendation: We recommend the UTC make the following dispositions:

1. Consensus: Rescind the Egyptian Hieroglyph standardized variation sequence 1333B FE00 as described in document L2/24-177 for Unicode version 16.0. [Reference: Section 1a of L2/24-166]
2. Action Item for Ken Whistler, PAG: Comment out the Egyptian Hieroglyph standardized variation sequence 1333B FE00 in StandardizedVariants.txt described in document L2/24-177, and add a reference to the UTC consensus, for Unicode version 16.0. [Reference: Section 1a of L2/24-166]
3. Consensus: Accept the ten Egyptian Hieroglyph standardized variation sequences as described in document L2/24-177 for Unicode version 16.0. [Reference: Section 1a of L2/24-166]
4. Action Item for Ken Whistler, UTC: Update the Pipeline page with the ten Egyptian Hieroglyph standardized variation sequences, as described in L2/24-177. [Reference: Section 1a of L2/24-166]
5. Action Item for Ken Whistler, PAG: Add the ten Egyptian Hieroglyph standardized variation sequences, as described in L2/24-177, for Unicode version 16.0. [Reference: Section 1a of L2/24-166]

Comments:
This document requests the addition of ten new standardized variant sequences for Egyptian Hieroglyphs and requests one sequence that was added in Unicode Version 15.0, be commented out.

Draft UAX #57 Unicode Egyptian Hieroglyph Database (Unikemet), dated July 1 2024, identifies three provisional properties for characters, C(ore), L(egacy), and N(one). For the 1,071 Egyptian Hieroglyph characters published in Unicode 5.2, the only property options are C(ore) or L(egacy). “Legacy” denotes characters that were used in the past, but should no longer be used and are not expected to be supported in (new) fonts. For the 3,995 new characters that will appear in 16.0, the only property options are C(ore) or N(one). “None” denotes characters that are not necessarily expected to be supported by Core fonts.

In Unicode version 5.2, ten rotated variants of hieroglyphs were separately encoded, though the rotated versions may have the same function as the unrotated, encoded characters. These rotated characters are being given the status “legacy” by way of the provisional property assignment kEH_Core=L in the Unikemet database (UAX #57).

The approach for handling non-meaningful rotations that was agreed upon with Egyptologists and standardized in Unicode 15.0 is to use variation sequences. This is expected to greatly improve the behavior of untailored search and indexing of hieroglyphic corpora, allowing, e.g., general-purpose web search and string search to be useful to the user community.

Unicode version 16.0 is seen by Egyptologists as the first one usable for large-scale encoding of

hieroglyphic corpora. In order to represent the ten commonly used “legacy” characters and to not encourage users to use the encoded rotations in this version, ten variation sequences are needed, and should be included in 16.0.

In addition to the ten new sequences, the sequence 1333B FE00, which was approved in Unicode version 15.0, should be commented out. The rotation corresponds to U+13339, which is a legacy character. However, the base of that sequence, U+1333B, is also a legacy character, best represented as 13338 FE06, one of the newly proposed sequences. U+13339 should instead be represented as 13338 FE03, another one of the newly proposed sequences, so that all of them share the base U+13338. (For discussion and evidence, see L2/24-175.) This request is reflected in the 16.0 Beta Feedback comment by Anderson (ReportID: ID20240606172234.)

We recommend the UTC accept the ten new standardized variation sequences and rescind the standardized variation sequence 1333B FE00.

1b Unapproved change in Core Spec [#492]

Document: L2/24-161 Comments on Public Review Issues (PRI 502 Beta Feedback on Unicode 16.0; ReportID: ID20240523093428)

Recommendation: We recommend the UTC make the following disposition:

1. **Action Item** for Debbie Anderson, Andrew Glass, Robin Leroy, EDC: Update the Core Spec text based on this discussion to clarify that the choices of either sequences or precomposed characters are somewhat "soft". Potentially include an illustration of encoding principles [Reference: Section 1b of L2/24-166]
2. **Action Item** for Michel Suignard, EDC: Incorporate the encoding principles of Egyptian hieroglyphs in UAX #57 , either with new text or by incorporating existing text in L2/23-181R2. [Reference: Section 1b of L2/24-166]

Comments: This topic arose in response to Unicode 16.0 Beta feedback ReportID: ID20240523093428 from Charlotte Buff.

There are many cases of ambiguity in Egyptian Hieroglyphs where a certain character can be found as an atomic encoding but can also be built as a sequence of components.

The proposed new DoNotEmit (L2/24-021) file included a `Precomposed_Hieroglyph` category to recommend avoiding the use of such precomposed sequences. This matches Unicode's guidance in 15.0, however this is relatively new guidance: in Unicode 14.0 the recommendation was systematically in favor of compound signs.

This guidance is still under flux and is changing again in Unicode 16.0 to a more nuanced approach. As it is yet in flux, we should avoid making DoNotEmit decisions based on it, and rather put effort into

providing clarity on the status quo where possible.

The `Precomposed_Hieroglyph` category and associated codepoints have already been removed from the draft DoNotEmit file for 16.0.

In the long run as we obtain a better understanding of usage, and as the Unikemet database grows, we may be able to produce useful DoNotEmit data for Egyptian.

1c 3-tier property system for Ancient Egyptian hieroglyphs [#496]

Document: L2/24-176 3-tier system for Ancient Egyptian hieroglyphs -- Nederhof et al.

Recommendation: We recommend the UTC make the following disposition:

1. **Consensus**: Change the status for the proposed kEH_Core property to Provisional, with three enumerated values. [Reference: Section 1c of L2/24-166]
2. **Action Item for** Michel Suignard, PAG: In Draft UAX # 57, change the status of kEH_Core from Normative to Provisional. [Reference: L2/24-176 and Section 1c of L2/24-166]
3. **Action Item for** Michel Suignard, PAG: In Draft UAX # 57, change the Syntax of kEH_Core to C|L|N. [Reference: L2/24-176 and Section 1c of L2/24-166]

Comments:
The Dec 26 2023 version of the Unicode Egyptian Hieroglyph Database (Unikemet) included the normative property kEH_Core, denoting characters that should be included in fonts. The property had "Yes" / "No" values. However, Egyptologists would like 3-way values, C(ore), L(egacy), and N(one). "Legacy" denotes characters that were used in the past, but should no longer be used and are not expected to be supported in fonts. For the 1,071 Egyptian Hieroglyph characters published in Unicode 5.2, the only property options would be C(ore) or L(egacy). For the 3,995 new characters that will appear in 16.0, the only property options would be C(ore) or N(one). "None" denotes characters that are not necessarily expected to be supported by Core fonts.

The Script Encoding Working Group agreed with the three-tiered system, and recommended the property values be changed to: C[ore], L[egacy], or N[one], and the property be changed from normative to provisional.

---

2. Sheek Bakrii Saphaloo (Oromo) [#389]

Document: L2/24-109 Proposal for Encoding the Sheek Bakrii Saphaloo Script in the UCS -- Oreen Yousuf and Daniel Yacob

Recommendation: We recommend the UTC make the following disposition:

1. **Action Item** for V.S. Umamaheswaran, RMP: Add Oromo (Sheek Bakrii Saphaloo) to the Roadmap at the following location: U+1C800..U+1CB2F [Reference: L2/24-109 and Section 2 of L2/24-166]

Comments: The Sheek Bakrii Saphaloo script was introduced for the Oromo language in Ethiopia in 1956.

The script is an abugida, with 33 base consonants, 33 geminated consonants, 10 vowel modifications (and 1 pure consonant version of each consonant). Each consonant also has an "unmodified" base version distinct from the pure consonant, used in pedagogical contexts. The proposal also includes digits and "word space" and "full stop" punctuation marks, coming to a total of 804 characters.

The proposal uses the Ethiopic model for encoding abugidas, with atomically encoded syllables.

Alternate spellings for the script name are "Sheikh Bakri Sapalo" (prevalently used in English) and in the past "Shaykh Bakri Saṗalō" or "Sheek Bakrii Saphaalloo". The chosen name "Sheek Bakrii Saphaloo" is based on the Qubee Latin orthography for Oromo. It should probably be replaced by an English spelling of the name, which is the norm among other script names in Unicode.

Furthermore, there is a "geminated" vowel and glottal stop series analogous to the other consonant series, but we're not sure of the actual usage of this series.

We are working with the authors to address issues in the proposal and currently request it be listed in the Roadmap.

II. MIDDLE EASTERN SCRIPTS

3. Arabic

3a Crown Letters [#353]

Document: L2/24-131 Proposal to encode Arabic Crown Letters (حروف التاج) -- Goregaokar, Hosny, Yang, Hasan

Recommendation: We recommend the UTC make the following disposition:

1. **Consensus**: Provisionally assign 22 Arabic crown letter code points and one crown letter combining mark U+10ED9..U+10EEE, U+10EF9 in the Arabic Extended-C block as described in L2/24-131 [Reference: Section 3a of L2/24-166]
2. **Action Item** for Ken Whistler, UTC: Update the Pipeline to include 22 Arabic crown letter code points and one crown letter combining mark U+10ED9..U+10EEE, U+10EF9 in the Arabic Extended-C block as described in L2/24-131 [Reference: Section 3a of L2/24-166]

3. **Action Item** for Roozbeh Pournader, EDC: Prepare additions/revisions for the Core Spec (in Table 9-3 and the subsequent text) to mention the joining behavior of the crown letters as described in L2/24-131 [Reference: Section 3a of L2/24-166]
4. **Action Item** for Manish Goregaokar, Roozbeh Pournader, EDC: Prepare a new table for Core Spec Chapter 9 (similar to Table 9-8) to list the Left-joining joining groups of the crown letters as described in L2/24-131 [Reference: Section 3a of L2/24-166]
5. **Action Item** for Manish Goregaokar, Roozbeh Pournader, EDC: Prepare text for the Arabic block description alerting Arabic implementers that left-joining Arabic characters may occur. [Reference: Section 3a of L2/24-166]
6. **Action Item** for Manish Goregaokar, SEW: Send a font to Michel Suignard for 22 Arabic crown letters and one combining mark. [Reference: Section 3a of L2/24-166]

Comments: This document proposes the addition of "crown letters", a form of capital letters designed for Arabic in the early 1900s, which found use in (mostly official) Egyptian documents for some time.

The crown is a looped flourish that sits atop the letter. For most dual-joining letters (except LAM) it joins to them, goes over and around, and then joins to the rest of the word. For non-left-joining letters (i.e., right-joining, non-joining), as well as LAM, it simply sits above them with no additional joining behavior. Outside of artistic contexts, it is not attested in non-word-initial position, as it was designed as a form of titlecase.

The proposal suggests atomically encoding 22 code points for each left-joining letter attested with a crown, and one combining mark for use with non-left-joining letters.

Despite being a form of titlecase, this proposal does not ask for any changes to casing data: this would break stability, and these characters are archaic.

A point worth explicitly highlighting for implementations is that this proposal introduces Arabic letters that have `Joining_Type = Left_Joining`, which have not been found in the Arabic blocks until these.

3b Fixing KASHMIRI YEH [#365]

Document: L2/24-152 Fixing the Kashmiri Yeh – Manish Goregaokar (मनीष गोरेगांवकर) and Roozbeh Pournader (روزبه پورنادر)

Recommendation: We recommend the UTC make the following disposition:

1. **Consensus**: Change the representative glyph for U+0620 ARABIC LETTER KASHMIRI YEH as detailed in L2/24-152 for Unicode 16.0 [Reference: Section 3b of L2/24-166]
2. **Consensus**: In ArabicShaping.txt, change the `Joining_Group` for U+0620 from `YEH` to a new joining group `KASHMIRI YEH`, and change its schematic name from `DOTLESS YEH WITH SEPARATE RING BELOW` to `KASHMIRI YEH` for Unicode 16.0. [Reference: Section 3b of L2/24-166]

3. **Action Item** for Lorna Evans, SEW: Send an updated font for U+0620 to Michel Suignard for Unicode 16.0. [Reference: Section 3b of L2/24-166]
4. **Action Item** for Roozbeh Pournader, PAG: Update ArabicShaping.txt for U+0620 per L2/24-152 for Unicode 16.0. [Reference: Section 3b of L2/24-166]
5. **Action Item** for Lorna Evans, EDC: Change the Core Specification text in Chapter 9 to provide an updated set of Kashmiri Yeh glyphs and correct guidance for its positional forms as detailed in L2/24-152 for Unicode 16.0. [Reference: Section 3b of L2/24-166]

Comments: Kashmiri has a letter used for palatalization, represented as U+0620 ARABIC LETTER KASHMIRI YEH. It takes on two rather distinct forms: one with a ring below ("gol"), and one with a "half" tail. The half-tail form is only found in isolated and final positions. According to the Core Spec (and many fonts), the ring form is found everywhere, however real-world Kashmiri users seem to disagree with this, and as a result have come up with alternate representations for the tail form at the end of words, including the Pashto letter U+06CD ARABIC LETTER YEH WITH TAIL as well as occasionally U+06C1 ARABIC LETTER HEH GOAL.

Previously, L2/23-273R documented research done to fully understand the situation, coming up with the conclusion that the Core Spec text is a mistake and should be remedied. This document contains concrete recommendations for how to do so.

3c Misspellings in some honorific character names [#484]

Recommendation: We recommend the UTC make the following disposition:

1. **Consensus**: Change the name of the following provisionally assigned characters [Reference: Section 3c of L2/24-166]:
    - U+FBD1: from ARABIC LIGATURE **AJJIL** ALLAAHU FARAJAHU ASH-SHAREEF to ARABIC LIGATURE **AJJAL** ALLAAHU FARAJAHU ASH-SHAREEF
    - U+FDCD: from ARABIC LIGATURE **AJJIL** ALLAAHU TAAALAA FARAJAHU ASH-SHAREEF to ARABIC LIGATURE **AJJAL** ALLAAHU TAAALAA FARAJAHU ASH-SHAREEF
    - U+10ED4: from ARABIC LIGATURE **QUDDISA** ALLAAHU SIRRAH to ARABIC LIGATURE **QADDASA** ALLAAHU SIRRAH.
2. **Action Item** for Ken Whistler, UTC: Update the pipeline with the new character names. [Reference: Section 3c of L2/24-166]

Comments: After getting some feedback on Twitter/X, Roozbeh Pournader noticed that the vowels (corresponding to the tashkeel) in some of the honorifics may be wrong. After further research and contacting some native Arabic speakers, he proposed these changes.

The proposed fixes appear reasonable and have also been verified by some SEW members with access to Arabic expertise. The SEW recommends the fixes.

3d Sixteen Quranic Arabic Characters [#497]

Document: L2/24-178 PROPOSAL TO ENCODE SIXTEEN QURANIC ARABIC CHARACTERS -- Rikza F. Sh.

Recommendation: We recommend the UTC make the following disposition:

1. **Consensus**: Provisionally assign 16 Arabic code points U+10EC9..U+10ECF, U+10EF0..U+10EF8 in the Arabic Extended-C block as described in L2/24-178 [Reference: Section 3d of L2/24-166]
2. **Action Item** for Ken Whistler, UTC: Update the Pipeline to include 16 Arabic code points U+10EC9..U+10ECF, U+10EF0..U+10EF8 in the Arabic Extended-C block as described in L2/24-178[Reference: Section 3d of L2/24-166]
3. **Action Item** for Lorna Evans, SAH: Send a font for the sixteen Quranic Arabic characters U+10EC9..U+10ECF, U+10EF0..U+10EF8 to Michel Suignard. [Reference: Section 3d of L2/24-166]
4. **Action Item** for Lorna Evans and Roozbeh Pournader, SAH: Prepare a Working Draft Update for UAX53 to mention the proposed MCM characters in the text of the annex. [Reference: Section 3d of L2/24-166]

Comments: The proposed characters are used as a technical notation in some Quranic *mushaf*'s that are marked up to distinguish various minority readings of the Quran. The proposed names are consistent with existing Unicode character names and the proposed code points are also appropriate. It should be noted that two of the proposed characters are modifier combining marks (as described in UAX #53).

---

4. Sidetic [#213]

Document: L2/23-019 Revised proposal to encode Sidetic in Unicode -- Pandey

Recommendation: We recommend the UTC make the following disposition:

1. **Consensus**: Change the character names for provisionally assigned code points U+10940..U+10948 SIDETIC LETTER N1..SIDETIC LETTER N9 to SIDETIC LETTER N01..SIDETIC LETTER N09 [Reference: Section 4 of L2/24-166]
2. **Action Item** for Ken Whistler, EDC: Change the character names for provisionally assigned code points U+10940..U+10948 SIDETIC LETTER N1..SIDETIC LETTER N9 to SIDETIC LETTER N01..SIDETIC LETTER N09 [Reference: Section 4 of L2/24-166]
3. **Action Item** for Robin Leroy and Anshuman Pandey, EDC: Provide annotations to Ken Whistler for the Sidetic characters U+10940..U+1095C that provide more information on their phonetic values where known. [Reference: Section 4 of L2/24-166]

Comments: Feedback was received at the June 2024 WG2 meeting that the provisionally assigned Sidetic characters SIDETIC LETTER N1..N9 should use some zero-padding in their names. Also, there was a recommendation to add the known phonetic values as annotations.

III. EAST ASIAN SCRIPTS

5. Ideographic Symbols and Punctuation [#469]

Document: L2/24-071R3 Proposal to encode three stable extended Suzhou Numeral-like letters for Cantonese Music -- Eiso Chan

Recommendation: We recommend the UTC make the following disposition:

1. **Consensus**: Rename 3 previously provisionally assigned Yangqin notational code points from U+16FF4 YANGQIN SIGN SLOW TWO, U+16FF5 YANGQIN SIGN SLOW THREE, and U+16FF6 YANGQIN SIGN SLOW FOUR (captured in UTC-179-C44) to U+16FF4 YANGQIN SIGN SLOW ONE BEAT, U+16FF5 YANGQIN SIGN SLOW THREE HALF BEATS, U+16FF6 YANGQIN SIGN SLOW TWO BEATS as described in L2/24-071R3 [Reference: Section 5 of L2/24-166]
2. **Action Item** for Ken Whistler, UTC: Update the Pipeline to reflect the new names U+16FF4 YANGQIN SIGN SLOW ONE BEAT, U+16FF5 YANGQIN SIGN SLOW THREE HALF BEATS, U+16FF6 YANGQIN SIGN SLOW TWO BEATS in Ideographic Symbols and Punctuation block as described in L2/24-071R3 [Reference: Section 5 of L2/24-166]

Comments: UTC-179-C44 previously accepted these three code points with the names U+16FF4 YANGQIN SIGN SLOW TWO, U+16FF5 YANGQIN SIGN SLOW THREE, and U+16FF6 YANGQIN SIGN SLOW FOUR. After some discussion, SEW and the authors feel that the names should include the word BEAT.

---

6. Jurchen [#256]

Document: L2/24-139 Proposal to Encode the Jurchen Script and L2/24-140 Proposal to Encode Radicals for the Jurchen Script -- Andrew West, Sun Bojun, Zhōnghuá Zìkù, Michael Everson

Recommendation: We recommend the UTC make the following disposition:

1. **Consensus**: Provisionally assign 914 Jurchen code points U+18E00..U+19191 in a new Jurchen block (U+18E00..U+1919F) as described in L2/24-139 [Reference: Section 6 of L2/24-166]
2. **Consensus**: Provisionally assign 51 Jurchen radicals code points U+191A0..U+191D2 in a new Jurchen Radicals block (U+191A0..U+191DF) as described in L2/24-140 [Reference: Section 6 of L2/24-166]
3. **Action Item** for Ken Whistler, UTC: Update the Pipeline to include 914 Jurchen code points U+18E00..U+19191 in a new Jurchen block (U+18E00..U+1919F) as described in L2/24-139 [Reference: Section 6 of L2/24-166]

4. **Action Item** for Ken Whistler, UTC: Update the Pipeline to include 51 Jurchen radicals code points U+191A0..U+191D2 in a new Jurchen Radicals block (U+191A0..U+191DF) as described in L2/24-140 [Reference: Section 6 of L2/24-166]
5. **Action Item** for Andrew West, SAH: Send a Jurchen and Jurchen Radicals font to Michel Suignard. [Reference: Section 6 of L2/24-166]

Comments: Jurchen is a historical ideographic script used to write the Jurchen language primarily during the Jin dynasty (1115 - 1234).

L2/24-139 proposes encoding a repertoire of Jurchen ideographs based on a list of sources it documents. Furthermore, L2/24-140 proposes encoding standalone forms of various Jurchen radicals; these are artificial radicals used for ordering Jurchen ideographs in the 1984 dictionary of Jīn Qǐzōng. Many of these are similar to existing Han or Tangut radicals, but cannot be unified because of their `Script` property.

L2/24-140 highlights that a lot of the Jurchen radicals are usable for Khitan Large Script as well, and there may be reason to have their `Script` property be `Common`. SEW feels that the script property for the Jurchen radicals could be changed once Khitan Large Script is proposed, but that the radicals should have a Jurchen script property for now.

---

7. Miao UTN [#419]

Document: UTN 56 Representing Miao -- Lorna Evans

Recommendation: We recommend the UTC make the following disposition:

1. **Action Item** for Ken Whistler, EDC: Include a reference to the new Miao UTN 56 in the Core Spec [Reference: Section 7 of L2/24-166]

Comments: This document provides an introduction and overview on how to represent Miao/Pollard script text. It also gives information on the languages using the script and resources which are available.

SEW endorses this as a good model to follow for future UTNs for scripts.

IV. SOUTH AND CENTRAL ASIAN SCRIPTS

8. Bengali [#476]

Document: L2/24-153 Proposal to encode Bengali Sign Combining Anusvara Above -- Jan Kučera

Recommendation: We recommend the UTC make the following disposition:

1. **Consensus**: Provisionally assign one Bengali code point U+0984 BENGALI SIGN COMBINING ANUSVARA ABOVE as described in L2/24-153 [Reference: Section 8 of L2/24-166]
2. **Action Item** for Ken Whistler, UTC: Update the Pipeline to include one Bengali code point U+0984 BENGALI SIGN COMBINING ANUSVARA ABOVE as described in L2/24-153
3. **Action Item** for Jan Kučera, SAH: Send a font for U+0984 BENGALI SIGN COMBINING ANUSVARA ABOVE to Michel Suignard. [Reference: Section 8 of L2/24-166]

Comments: The proposal author Jan Kučera was approached by an expert, Charles Li, who works with Bengali manuscripts. Dr. Li asked how to represent combining anusvara in Bengali, since a combining anusvara is found in other scripts, such as Newa U+11444 NEWA SIGN ANUSVARA.

While there are no modern typeset examples, the proposal contains evidence from manuscripts and an example showing contrast between the proposed character and U+0982 BENGALI SIGN ANUSVARA. A question was raised whether the name should contain “combining”, but “combining” is found in names for comparable anusvaras in Kannada, Telugu, and Grantha, for example. As a result, leaving “combining” in the name was deemed acceptable.

We recommend a code point for BENGALI SIGN COMBINING ANUSVARA ABOVE be provisionally assigned.

---

9. Mongolian [#507]

Document: L2/24-180 Proposal to refer to UTN 57 for implementing the Mongolian script -- Liang Hai

Recommendation: We recommend the UTC make the following disposition:

1. **Consensus**: Section 13.5 of the Unicode Standard will reference UTN 57 for compatibility with Chinese standard GB/T 25914-2023 [Reference: Section 9 of L2/24-166]
2. **Action Item** for Liang Hai, EDC: Reference UTN 57 from the Core Spec. [Reference: Section 9 of L2/24-166]

Comments: A new UTN on Mongolian by Kushim Jiang will supplement the content in the Mongolian section of the Core Spec (section 13.5). The Mongolian section in Unicode 16.0 will now point to the new UTN, UTN 57, which contains guidance for implementers. Revisions to the Mongolian chapter will remove conflicting information.

Publication of the new UTN and revisions to section 13.5 are important, because the new information will now be compatible with the Chinese national standard GB/T 25914-2023, which went into effect June 1, 2024.

V. PHONETIC CHARACTERS

10. Modifier click letters [#416]

Document: L2/24-052R Unicode request for modifier pre-Kiel click letters (revised) -- Miller

Recommendation: This is an FYI to the UTC, but this item should be mentioned during review of minutes from meeting UTC 179 that the reference to consensus 179-C60 should be to L2/24-052R (not L2/24-052).

Comments: UTC-179-C60, and section 16 of L2/24-068, provisionally assigned 5 code points U+107BB..U+107BF based on L2/24-052, without listing the characters explicitly. However, L2/24-052 did not contain a consistent mapping of code points (the list of characters on the top of page 2 varied from the properties list and the code chart). The document has since been updated in L2/24-052R with consistent listing of characters across the proposal. This is to clarify the previous report that SEW considers the mappings in L2/24-052R to be acceptable for these characters.

---

11. Phonetic characters: Modifier Voiceless Implosive Letters [#442]

Document: L2/24-130R Unicode request for modifier voiceless implosive letters -- Miller

Recommendation: We recommend the UTC make the following disposition:

1. **Consensus**: Provisionally assign six modifier letter code points U+1DFFA MODIFIER LETTER SMALL C WITH HOOK, U+1DFFB MODIFIER LETTER SMALL K WITH HOOK, U+1DFFC MODIFIER LETTER SMALL P WITH HOOK, U+1DFFD MODIFIER LETTER SMALL Q WITH HOOK, U+1DFFE MODIFIER LETTER SMALL T WITH HOOK, U+1DFFF MODIFIER LETTER SMALL T WITH HOOK AND RETROFLEX HOOK in the Latin Extended-G block as described in L2/24-130R [Reference: Section 11 of L2/24-166]
2. **Action Item** for Ken Whistler, UTC: Update the Pipeline to include six modifier letter code points U+1DFFA MODIFIER LETTER SMALL C WITH HOOK, U+1DFFB MODIFIER LETTER SMALL K WITH HOOK, U+1DFFC MODIFIER LETTER SMALL P WITH HOOK, U+1DFFD MODIFIER LETTER SMALL Q WITH HOOK, U+1DFFE MODIFIER LETTER SMALL T WITH HOOK, U+1DFFF MODIFIER LETTER SMALL T WITH HOOK AND RETROFLEX HOOK in the Latin Extended-G block as described in L2/24-130R [Reference: Section 11 of L2/24-166]
3. **Action Item** for Kirk Miller, SAH: Send a font for six modifier letter code points U+1DFFA..U+1DFFF to Michel Suignard. [Reference: Section 11 of L2/24-166]

Comments: This proposes six modifier letter versions of existing encoded phonetic letters for voiceless implosives. These letters are not a part of modern recommended IPA orthography, but are still used by

some linguists.

This proposal does not provide attestation for any direct uses of these as modifier letters, but makes a case for the various reasons IPA letters are useful to have as modifiers, with specific examples of phonetic notation this encoding would enable.

This proposal has received support from the IPA in L2/24-049.

---

12. Phonetic characters: Modifier Letters with Palatal Hook [#443]

Document: L2/24-144 Unicode request for modifier letters with palatal hook -- Miller

Recommendation: We recommend the UTC make the following disposition:

1. **Consensus**: Provisionally assign 5 phonetic modifier letter code points U+1DFF5..U+1DFF9 in Latin Extended-G block as described in L2/24-144 [Reference: Section 12 of L2/24-166]
2. **Action Item** for Ken Whistler, UTC: Update the Pipeline to include 5 phonetic modifier letter code points U+1DFF5..U+1DFF9 in Latin Extended-G block as described in L2/24-144 [Reference: Section 12 of L2/24-166]
3. **Action Item** for Kirk Miller, SAH: Send a font for 5 phonetic modifier letter code points U+1DFF5..U+1DFF9 to Michel Suignard. [Reference: Section 12 of L2/24-166]

Comments: This document proposes 5 modifier letter versions of existing encoded phonetic letters that have a palatal hook. All but U+1DFF7 MODIFIER LETTER SMALL N WITH PALATAL HOOK are attested in the document.

The IPA has expressed support for that character in L2/24-049.

---

13. Phonetic characters: Modifier psi and omega [#485]

Document: L2/24-145R Unicode request for modifier psi and omega -- Miller

Recommendation: We recommend the UTC make the following disposition:

1. **Consensus**: Provisionally assign two modifier letter code points U+1DFF3 MODIFIER LETTER SMALL GREEK PSI, U+1DFF4 MODIFIER LETTER SMALL GREEK OMEGA in the Latin Extended-G block as described in L2/24-145R [Reference: Section 13 of L2/24-166]
2. **Action Item** for Ken Whistler, UTC: Update the Pipeline to include two modifier letter code points U+1DFF3 MODIFIER LETTER SMALL GREEK PSI, U+1DFF4 MODIFIER LETTER SMALL GREEK OMEGA in the Latin Extended-G block as described in L2/24-145R [Reference: Section 13 of L2/24-166]
3. **Action Item** for Kirk Miller, SAH: Send a font for U+1DFF3 MODIFIER LETTER SMALL

GREEK PSI, U+1DFF4 MODIFIER LETTER SMALL GREEK OMEGA to Michel Suignard. [Reference: Section 13 of L2/24-166]

Comments: This proposes modifier versions of the lowercase Greek letters psi and omega for use in phonetic notation. Modifier Omega is used in Sinological notation as a vowel, and in Americanist notation for labialization.

There may appear to be a choice in what to base the modifier letter small omega on, Greek or Latin, considering there also exists a lowercase Latin omega at U+A7B7. However, the Latin lowercase omega looks different from the Greek lowercase omega, as can be seen in the representative glyph in the code charts. The modifier letter being proposed here looks like the Greek omega, not the Latin omega.

It should be noted that there is existing precedent for encoding Greek script characters in Latin (or predominantly-Latin) blocks—e.g., U+AB65 GREEK LETTER SMALL CAPITAL OMEGA in the Latin Extended-E block.

The encoding of these characters are supported by the IPA in L2/24-049.

---

14. Phonetic characters: Greek letters with palatal hook [#486]

Document: L2/24-146 Unicode request for Greek letters with palatal hook -- Miller

Recommendation: We recommend the UTC make the following disposition:

1. **Consensus**: Provisionally assign three phonetic code points U+1DF3B GREEK SMALL LETTER BETA WITH PALATAL HOOK, U+1DF3C GREEK SMALL LETTER THETA WITH PALATAL HOOK, U+1DF3D GREEK SMALL LETTER CHI WITH PALATAL HOOK in the Latin Extended-G block as described in L2/24-146 [Reference: Section 14 of L2/24-166]
2. **Action Item** for Ken Whistler, UTC: Update the Pipeline to include three phonetic Latin code points U+1DF3B GREEK SMALL LETTER BETA WITH PALATAL HOOK, U+1DF3C GREEK SMALL LETTER THETA WITH PALATAL HOOK, U+1DF3D GREEK SMALL LETTER CHI WITH PALATAL HOOK in the Latin Extended-G block as described in L2/24-146 [Reference: Section 14 of L2/24-166]
3. **Action Item** for Kirk Miller, SAH: Send a font for U+1DF3B GREEK SMALL LETTER BETA WITH PALATAL HOOK, U+1DF3C GREEK SMALL LETTER THETA WITH PALATAL HOOK, U+1DF3D GREEK SMALL LETTER CHI WITH PALATAL HOOK to Michel Suignard. [Reference: Section 14 of L2/24-166]

Comments: This proposes three attested Greek-derived phonetic symbols with palatal hooks, first introduced by L2/24-050, but split out into its own proposal to resolve questions about the script and character names.

The IPA has expressed support for the encoding of these letters in L2/24-049.

15. IPA: Modifier Sinological extensions to the IPA [#493]

Document: L2/24-147 Modifier Sinological extensions to the IPA -- Miller

Recommendation: We recommend the UTC make the following disposition:

1. **Consensus**: Provisionally assign ten modifier letter code points U+1DFE9..U+1DFF2 in the Latin Extended-G block as described in L2/24-147 [Reference: Section 15 of L2/24-166]
2. **Action Item** for Ken Whistler, UTC: Update the Pipeline to include ten modifier letter code points U+1DFE9..U+1DFF2 in the Latin Extended-G block as described in L2/24-147 [Reference: Section 15 of L2/24-166]
3. **Action Item** for Kirk Miller, SAH: Send a font with the ten Sinological modifier letters to Michel Suignard. [Reference: Section 15 of L2/24-166]
4. **Action Item** for Ken Whistler, EDC: Consider adding annotations to U+027F, U+0285, U+02AE, U+02AF, U+1DFEB, U+1DFEC, U+1DFED, U+1DFEE as described in L2/24-147 [Reference: Section 15 of L2/24-166]

Comments: This document proposes ten modifier letter code points for phonetic symbols found in Sinological and Americanist phonetic conventions.

The document highlights a potential discrepancy between the symbols that are used by Y.R. Chao and Bernhard Karlgren, and that there are concerns that encoding modifier versions of these letters will prevent potential future disunification. SEW does not feel that encoding these modifiers will limit the possibility of disunification in the future.

The IPA has expressed support for this proposal in L2/24-049.

---

16. IPA: Miscellaneous historical and para-IPA modifier letters [#494]

Document: L2/24-171 Miscellaneous historical and para-IPA modifier letters - Miller

Recommendation: We recommend the UTC make the following disposition:

1. **Consensus**: Provisionally assign two barred letters (U+1DF3E, U+1DF3F) and 17 modifier letters (U+1DFD8..U+1DFE7) in the Latin Extended-G block as described in L2/24-171 [Reference: Section 16 of L2/24-166]
2. **Action Item** for Ken Whistler, UTC: Update the Pipeline to include two barred letters (U+1DF3E, U+1DF3F) and 17 modifier letters (U+1DFD8..U+1DFE7) in the Latin Extended-G block as described in L2/24-171 [Reference: Section 16 of L2/24-166]
3. **Action Item** for Kirk Miller, SAH: Send a font for two barred letters (U+1DF3E, U+1DF3F) and 17 modifier letters (U+1DFD8..U+1DFE7) to Michel Suignard. [Reference: Section 16 of L2/24-166]

Comments: This proposes 19 phonetic letters, largely modifier variants of historical and para-IPA phonetic letters.

The IPA has expressed support for the encoding of these letters in L2/24-049.

VI. SYMBOLS

17. Numbers with slashes used in figured bass (musical symbols) [#346]

Document: L2/23-277 Unicode request for digits with slashes used in figured bass -- Gavin Jared Bala, Kirk Miller

Recommendation: We recommend the UTC make the following disposition:

1. **Consensus**: Change the name of the provisionally assigned characters U+1D1F7..U+1D1FE to replace the word "SLASH" with "STROKE" [Reference: Section 17 of L2/24-166]
2. **Action Item** for Ken Whistler, SAH: Update the pipeline for U+1D1F7..U+1D1FE to replace the word "SLASH" with "STROKE" [Reference: Section 17 of L2/24-166]

Comments: This reflects feedback that was received orally during the 2024 WG2 in Prague, Czech Republic. Unicode tends to use "with stroke" as a generic naming for "characters with some line drawn through them", when there's no specificity of the type of line. "With slash" would mean specifically a solidus overstrike.

---

18. Music Symbols: Musical notes and rests [#392]

Document: L2/24-172 Unicode request for 256th, 512th, and 1024th notes and rests --Gavin Jared Bala, Kirk Miller

Recommendation: We recommend the UTC make the following disposition:

1. **Consensus**: Provisionally assign 6 musical symbol note and rest code points U+1D250..U+1D255 in a new Musical Symbols Supplement block U+1D250..U+1D28F as described in L2/24-172 [Reference: Section 18 of L2/24-166]
2. **Action Item** for Ken Whistler, UTC: Update the Pipeline to include 6 musical symbol note and rest code points U+1D250..U+1D255 in a new Musical Symbols Supplement block U+1D250..U+1D28F as described in L2/24-172 [Reference: Section 18 of L2/24-166]
3. **Action Item** for Kirk Miller, SAH: Send a font for musical symbol note and rest code points U+1D250..U+1D255 to Michel Suignard. [Reference: Section 18 of L2/24-166]

Comments: This is a proposal for three combining characters, used as flags (or beams) to write very short-duration notes, and three corresponding short-duration rest characters. Currently, the shortest note Unicode supports is the 128th note, with five flags (or beams). Adding flags to a flagless note can create a 256th note (with 6 flags), 512th note (with 7 flags) and 1024th note (with 8 flags). A similar pattern can be used for rests. Notes and rests with these durations appear in works of Mozart, Beethoven, Vivaldi, and others. These notes and rests are supported in SMuFL and MusicXML, which is a format used for exchanging music score files between applications.

---

19. Music Symbols: Turkish and Arabic accidentals [#445]

Document: L2/24-174 Unicode request for Turkish and Arabic accidentals -- Gavin Jared Bala, Kirk Miller

Recommendation: We recommend the UTC make the following disposition:

1. **Consensus**: Provisionally assign 5 accidental musical symbol code points U+1D256..U+1D25A in a new Musical Symbols Supplement block U+1D250..U+1D28F as described in L2/24-174 [Reference: Section 19 of L2/24-166]
2. **Action Item** for Ken Whistler, UTC: Update the Pipeline to include 5 accidental musical symbol code points U+1D256..U+1D25A in a new Musical Symbols Supplement block U+1D250..U+1D28F as described in L2/24-174 [Reference: Section 19 of L2/24-166]
3. **Action Item** for Kirk Miller, SAH: Send a font for musical symbol accidental code points U+1D256..U+1D25A to Michel Suignard. [Reference: Section 19 of L2/24-166]

Comments: This proposes some accidentals used in the notation of Turkish and Arabic music, primarily versions of sharp/flat symbols with a stroke for marking quarter tones.

The `Bidi_Class` property of these characters should be `L`.

---

20. Middle Asterisk [#503]

Document: L2/24-173 Proposal to encode a Middle Asterisk as referred to in the German standard DIN 2137 -- Pentzlin

Recommendation: We recommend the UTC make the following disposition:

1. **Action Item** for Ken Whistler, EDC: Consider adding an annotation to U+2217 that it may be used to represent the telephony asterisk. [Reference: L2/24-173 and Section 20 of L2/24-166]

Comments: This proposes a "middle asterisk" symbol for representing the telephony asterisk found on phone keyboards, for use in the new German keyboard standard DIN 2137. The ASCII asterisk is not necessarily centered, so the DIN standard currently uses the mathematical asterisk sign (U+2217

ASTERISK OPERATOR). However U+2217 ASTERISK OPERATOR is treated as the start of a mathematical expression by Word. A separate code point explicitly for representing the telephony asterisk is requested.

SEW feels that the existing set of code points (U+002A U+204E U+2217 U+26B9 U+2731 U+1F7B6) covers all use cases of asterisk-like characters. SEW recommends usage of U+2217 ASTERISK OPERATOR.

VII. FEEDBACK

21. PRI 500 Draft UAX #57, Unicode Egyptian Hieroglyph Database (Unikemet) [#471]

Document: L2/24-161 Comments on Public Review Issues (PRI 500 Draft UAX #57, Unicode Egyptian Hieroglyph Database (Unikemet)

Recommendation: We recommend the UTC make the following disposition:

1. **Consensus**: Approve Draft UAX #57 (draft 3, revision 2, as amended in discussion) as a UAX component of Unicode version 16.0.

Comments: The feedback items listed below (from PRI # 500) have been addressed in the latest version of Unikemet, the relevant code charts, and the text of draft UAX # 57:
ReportID: ID20240229132224
ReportID: ID20240509174621
ReportID: ID20240508101047
ReportID: ID20240607142556
ReportID: ID20240614105616
ReportID: ID20240618154946
ReportID: ID20240619172020

The following feedback was received before the April 2024 UTC and, if deemed actionable, has already been handled in Unikemet:
ReportID: ID20240211061152
ReportID: ID20240225095538
ReportID: ID20240422064429

In addition, the PRI 502 16.0 beta feedback on the Egyptological arrows (ReportID: ID20240508072349) has also been addressed; the glyphs have been fixed.

PRI 502 16.0 beta feedback item from Deborah Anderson regarding commenting out a sequence (ReportID: ID20240606172234) is handled separately in this report (see section 1a: "Rotations" above,

and doc L2/24-177).

---

22. Beta Feedback: Errors in Bamum Supplement names [#481]

Document: L2/24-161 Comments on Public Review Issues (PRI 502 Beta Feedback on Unicode 16.0; ReportID: ID20240504174305)

Recommendation: This is an FYI to the UTC. SEW is working on this and will propose fixes at a later UTC.

Comments: A formal name alias has already been approved for U+1680B "MAEMBGBIEE" for MAEMGBIEE (consensus 178-C7). For the rest, SEW is still investigating the issue.

---

23. Beta Feedback: Inconsistent representative glyphs for consonant-preceding rephas [#489]

Document: L2/24-161 Comments on Public Review Issues (PRI 502 Beta Feedback on Unicode 16.0; ReportID: ID20240624144320)

Recommendation: We recommend the UTC make the following disposition:

1. **Action Item** for Michel Suignard, Charts WG: Update reference glyphs for U+113D1 TULU-TIGALARI REPHA to not use a dotted circle for Unicode 16.0 [Reference: Section 23 of L2/24-166]

Comments: Unicode 16 includes five characters (U+0D4E, U+113D1, U+11941, U+11D46, U+11F02) that have `InSC=Consonant_Preceding_Repha`. These characters are nonspacing marks encoded in phonetic order, in other words in text stream order they *precede* the consonant they combine with.

In draft charts for Unicode 16.0, all five characters are represented with a dashed box, except U+113D1 TULU-TIGALARI REPHA, which has a dotted circle inside the dashed box.

The dashed box is used for a variety of purposes that tend to mean that the reference glyph is not actually what the character looks like, while the dotted circle typically references combining on a preceding character.

In this case because these characters do not combine onto the character preceding them, they should not have a dotted circle. Hence, we recommend the dotted circle be removed from the glyph for U+113D1 TULU-TIGALARI REPHA.

---

24. Beta Feedback: Runic punctuation [#513]

Document: L2/24-161 Comments on Public Review Issues (PRI 502 Beta Feedback on Unicode 16.0; ReportID: ID20240629152410)

Recommendation: We recommend the UTC make the following disposition:

1. **Consensus**: Add `Runr` to the `Script_Extensions` of U+16EB RUNIC SINGLE PUNCTUATION, U+16EC RUNIC MULTIPLE PUNCTUATION, U+16ED RUNIC CROSS PUNCTUATION for Unicode 16.0 [Reference: Section 24 of L2/24-166]
2. **Action Item** for Roozbeh Pournader, PAG: Add `Runr` to the `Script_extensions` of U+16EB RUNIC SINGLE PUNCTUATION, U+16EC RUNIC MULTIPLE PUNCTUATION, U+16ED RUNIC CROSS PUNCTUATION for Unicode 16.0 [Reference: Section 24 of L2/24-166]

Comments: SEW agrees with the feedback and feels that these characters, not attested outside of Runic contexts, should have the appropriate `Script_Extensions` property.

B. DOCUMENTS NOT REQUIRING UTC ACTION

25. Currency Symbol: Transnistrian (/Pridnestrovian) Ruble [#228]

Document: L2/24-134 Proposal to encode a Pridnestrovian ruble sign - Ivan Adrianov

Recommendation: This is FYI to the UTC.

Comments: We reviewed this proposal for the "Pridnestrovian ruble." A proposal for the same symbol was proposed in 2023 (L2/23-022) from a different author (with SAH comments on page 18 of L2/23-012).

The proposal L2/24-134 needs to have the user community involved, such as a national bank authority. (Cf. the Kyrgyz SOM currency symbol proposal from The National Bank of Kyrgyz Republic L2/20-261 and Lari currency symbol from Georgia, where the National Bank of Georgia submitted the proposal, L2/14-161.)

In addition, a more diverse set of examples should be provided.

---

26. Symbols: Historic scientific characters used by Leibniz [#289]

Documents: L2/24-141 Proposal to add historic scientific characters (WG2 N5277) -- Uwe Mayer, et al
L2/24-142 Leibniz project - Supplemental documents -- Stötzner et al.

Recommendation: We recommend the UTC take no action.

Comments: The proposal document L2/24-141 proposes 228 characters used in the works of Leibniz. Responses from the proposal authors to earlier SEW feedback is contained in L2/24-142, which also includes information on the ambiguity signs and slashed digits. Feedback sent to the authors also noted that some proposed characters are mathematical notation and are not plain text.

The documents are posted in the Unicode document register in order to elicit input from other experts and the public at large.

Note: The Leibniz proposal L2/24-141 was presented at the June 2024 WG2 meeting. The proposers were encouraged to break this proposal into smaller, separate proposals, starting with characters that are likely not controversial.

---

27. Mongolian: Proposal to Encode One Manchu Letter [#371]

Documents:

- L2/24-025 Proposal to Encode One Manchu Letter -- CheonHyeong Sim, et al.
- L2/24-156 Feedback on the Proposal to Revise and Add Manchu Character U+1879 (N5270) - China NB
- L2/24-137 Response to WG2N5265 [L2/24-156] -- CheonHyeong Sim

Recommendation: This is an FYI to the UTC

Comments: The character MONGOLIAN LETTER MANCHU ALTERNATE UE was provisionally assigned a code point at the January 2024 UTC. At the June 2024 WG2 meeting, China submitted feedback (WG2 N5270=L2/24-156), to which the proposal author responded (L2/24-137). At the WG2 meeting, China requested a year to review the topic, but did not object when the character was identified as eligible for a future version of the standard (see M71.11 of WG2 recommendations, N5254), presumably to appear in the 7th edition of ISO/IEC 10646, which will likely cover Unicode versions 17.0 and 18.0.

---

28. Punctuation: Old Turkic one and two-dot punctuation [#406]

Document: L2/24-132 Proposal to Include Punctuation of Old Turkic - Yazga

Recommendation: This is FYI to the UTC.

Comments: This document proposes an Old Turkic colon and interpunct-shaped punctuation which are

used as word/phrase/sentence separation in Old Turkic.

SEW has corresponded with the author and suggested the user community select some of the existing codepoints for use for this, and develop better fonts that include the correct stylistic form for these code points. We welcome a UTN or other form of documentation to identify and describe these characters as intended for use with Old Turkic.

---

29. ExtIPA cartouche [#431]

Document: L2/24-182 Unicode request for ExtIPA cartouche -- Miller and Ball

Recommendation: We recommend the UTC take no action.

Comments: This proposes an encircling "cartouche" as a start/end pair for use with IPA characters to mark transcriptions of unidentifiable or indeterminate sounds. It is a convention used by Extensions to the IPA for Disordered Speech (extIPA)

The proposed characters are a LEFT CARTOUCHE END CAP and a RIGHT CARTOUCHE END CAP that could, for IPA usage, enclose Latin, Greek, and possibly Cyrillic phonetic notation. A font should ideally automatically join the end caps, without requiring overlines or underlines to support the top and bottom lines of a cartouche. Diacritics would be enclosed by the structure.

There was a question whether the mechanism proposed was actually plain text, and whether markup, other styling, or higher level protocol might be more appropriate. Use of U+20DD COMBINING ENCLOSING CIRCLE was mentioned as one mechanism, though this is tricky to implement.

One of the authors of the proposal is checking with the ICPLA on the possibility of using existing parentheses or brackets to represent the "cartouche" in print.

---

30. Latin: English Phonotypic Alphabet (4th revised version) [#456]

Document: L2/24-136 Draft Fifth Revised Proposal to encode characters for the English Phonotypic Alphabet (EPA) in the UCS -- Pentzlin

Recommendation: This is an FYI to the UTC.

Comments: This is a proposal for letters in the English Phonotypic Alphabet (EPA), which was originally devised in the 19c as an English spelling reform. EPA contains elements later incorporated into IPA. Several earlier proposals were submitted by the author to the UTC in 2010 and 2011.

A number of members of the Script Encoding Working Group have corresponded with the proposal author and provided feedback and comments. Subsequent discussion in the Script Encoding Working Group recommended the proposal to not inject casing where it is not called for by the text. The author

agreed to revise his proposal.

---

31. Kawi: Preliminary Proposal to Add Kawi jihwāmūlīya and upadhmānīya [#459]

Document: L2/24-128 Preliminary Proposal to Add Kawi jihwāmūlīya and upadhmānīya -- Aditya Bayu Perdana

Recommendation: This is an FYI to the UTC.

Comments: This proposal suggests adding jihvāmūlīya and upadhmānīya code points for the Kawi script.

The proposal has only one attestation for each of these characters used on a recently discovered Watu Genuk slab inscription.

Some SEW members suggested using the corresponding Brahmi code points (U+11003 and U+11004) for now until further attestations are found.

---

32. Runic: Article for discussion [#487]

Document: L2/24-129 (R)Unicode: Encoding and Sustainability Issues in Runology -- Elisabeth Maria Magin, Marcus Smith

Recommendation: This is an FYI to the UTC.

Comments: This article by Marcus Smith and Elisabeth Maria Magin was originally published in 2023 and was sent to Deborah Anderson in May 2024. The authors identify problems in the currently encoded Runic block and suggest some approaches on how to handle the problems.

The authors are invited to document issues outlined in the article with examples, for further discussion in the Script Encoding Working Group.

---

33. VSes: 128 User-Defined Variation Selectors [#495]

Document: L2/24-148 Proposal to Encode a Set of 128 User-Defined Variation Selectors (WG2 N5266) - West and Everson

Recommendation: This is an FYI to the UTC.

Comments: The group looked at this proposal, and largely opposed the specific mechanism proposed by

this document, noting that this opens up the issue of having private use characters with specific properties, which the UTC has consistently been avoiding, and the interoperability issues of completely unchecked nonstandard variation sequences.

At the same time, some members of the group recognize the need for interchanging such variants in historic scripts, and the fact that standardizing them as standardized variation sequences is not tractable. One member of the group offers to work with the proposal authors on a proposal for a registration mechanism similar to UTS37, but aimed at non-ideographic historic scripts. The narrower scope should make it possible to avoid duplicate registrations (that is, distinct variation sequences should map to distinct glyphic subsets, whereas this is not guaranteed in UTS37).

---

34. Ethiopic UTN: Application of the Zero Width Joiner Mark for Selective Ethiopic Ligation [#500]

Document: UTN 55 Application of the Zero Width Joiner Mark for Selective Ethiopic Ligation -- Daniel Yacob

Recommendation: This is an FYI to the UTC

Comments: We support this UTN being posted.

C. IN PROCESS

- L2/24-133 On the future of MANDAIC LETTER KAD - Ben Joeng (Yang) [#425]
- L2/24-151 Proposal for two geometric shapes for Japanese traditional calendars -- Gen Kojitani [#515]
- L2/24-125 Proposal to add a new CJK Abbreviations block -- Gen Kojitani [#516]
    - L2/24-126 Comments on L2/24-125 -- CheonHyeong Sim
    - L2/24-135 Response to L2/24-126 (CheonHyeong Sim's Comments on L2/24-125) -- Gen Kojitani