L2/13-148

ISO/IEC JTC1/SC2/WG2 N 4453
Date: 2013-06-13

ISO/IEC JTC1/SC2/WG2
Coded Character Set
Secretariat: Japan (JISC)

Doc. Type: **Disposition of comments**

Title: **Disposition of comments on DAM2 text to ISO/IEC 10646 3rd edition**

Source: **Michel Suignard (project editor)**
Project: JTC1 02.10646.00.02.00.03
Status: For review by WG2
Date: 2013-06-13
Distribution: WG2
Reference: SC2 N4275 N4278, WG2 N4391 N4396 N4398 N4422
Medium: Paper, PDF file

Comments were received from Germany, Ireland, Japan, South Korea (ROK), and USA. The following document is the disposition of those comments. The disposition is organized per country.

> Note – With some minor exceptions, the full content of the ballot comments have been included in this document to facilitate the reading. The dispositions are inserted in between these comments and are marked in **Underlined Bold Serif text,** *with explanatory text in italicized serif***.**

As a result of these dispositions, all countries have now Positive votes.

Germany: Positive with comments

Germany votes “Yes with comments”.

Technical comment

T1. U+A7AF LATIN LETTER SMALL CAPITAL OMEGA

The DAM2 (subject of this comment) contains the character: U+A7AF LATIN LETTER SMALL CAPITAL OMEGA.
Another document "Draft additional repertoire for ISO/IEC 10646:2014" (Document SC2/WG2 N4383, accepted as such according to SC2/WG2 N4398) contains the following two characters:
A7B6 LATIN CAPITAL LETTER OMEGA
A7B7 LATIN SMALL LETTER OMEGA
Until now, it is common practice that Latin letters named after Greek letters resemble the small form of the Greek letter for both of the capital and small form. This is the case for the letters proposed in N4383.
Also, until now it is common practice that the shape of "small capital" letters resembles the one of the (pure) capital form of the letter named the same beyond the script and case designation.
Thus, a "Latin letter small capital omega" has to resemble the "Latin capital letter omega", which in turn resembles the "Greek small letter omega".
However, A7AF resembles a Greek capital letter omega.
Thus, "Latin letter small capital omega" is a misnomer.

Proposed change by Germany
Germany requests a name change for A7AF.
Germany suggests the character to be named LATIN LETTER SMALL CAPITAL GREEK OMEGA.

Accepted in principle
The new proposed name and location are:
AB65 GREEK LETTER SMALL CAPITAL OMEGA
This letter is in fact similar to a set of characters in the Phonetic Extensions block (1D00..1D7F) encoded in the range U+1D26..U+1D2A named GREEK LETTER SMALL CAPITAL GAMMA to PSI. Short of creating a new extension to the Phonetic Extensions block (itself full), the character may be added in the Latin Extended-E block with a note added in the block mentioning that the block contains ‘Greek’ letters used for phonetic purpose.
The location move is a result of disposition of Irish comment T2 of the CD ballot on 10646 the edition (WG2 N4454).

Ireland: Negative

Ireland disapproves the draft with the technical and editorial comments given below.
Acceptance of these comments and appropriate changes to the text will change our vote to approval.

Technical comments

T1. Page 52: Row 10C8: Hungarian.

With reference to ISO/IEC JTC1/SC2/WG2 N4374R "Old Hungarian/Szekely-Hungarian Rovas Ad-hoc Report", and L2/13-049 [aka ISO/IEC JTC1/SC2/WG2 N4422] "Declaration for declining the 'Hungarian' block of the DAM", Ireland requests the change of the name of the script from HUNGARIAN to either OLD HUNGARIAN or to SZEKLER-HUNGARIAN in both the block names and the character names. Ireland notes the following from N4374R:

> In N4197 "Remarks on Old Hungarian and other scripts with regard to N4183", it is noted that "the preferred term in current Hungarian scientific literature is *'székely írás'* i.e. 'Szekler script'." Other terms for the script which have been used are "Hungarian Runic", "Hungarian script", and "Szekler-Hungarian script" (the last of which is similar to "Székely-Hungarian Rovas" promoted by "the Rovas side").

Discussions regarding the encoding of the Old Hungarian/Szekler-Hungarian script have not progressed efficiently, in part due to arguments similar to those given in L2/13-049. Those arguments have not been shared by our own Hungarian colleagues who have supported the encoding of this script as presented in ISO/IEC JTC1/SC2/WG2 N4268R "Consolidated proposal for encoding the Old Hungarian script in the UCS" and previous documents by that same document's authors. Following the text given in N4197, we will describe the authors of L2/13-049 as "the Rovas side" in the discussion below.

With regard to the name of the script, it has been demonstrated that the common name of the script in the English language is "Old Hungarian", and has been for many years. Members of "the Rovas side", however, have attempted (for several years) to get the neologism "Rovás" into the standard, with doubtful argument about the meaning of the word "Old" and with obvious attempts (via editing of various Wikipedia articles etc.) to establish the term "Rovash" in various spellings in other European languages, thus leading to possible argument that evidence exists that the name in the English language "must" be "Székely-Hungarian Rovás". This name, however, is not an acceptable name for the script in the English language, because the Hungarian word *rovás* is a general category word which means 'incised script' (Germanic runes, Ogham, and Old Turkic are also "incised scripts"). In point of fact, many attempts to devise a suitable compromise name have been attempted:

1) The first proposal was made using the preferred English term "Old Hungarian", on the basis of an agreement made following on from a meeting in 2008-07 in Budapest, which some of the present objectors attended; they did not object to the term at the time.
2) "The Rovas side" objected, insisting on "Szekely-Hungarian Rovas".
3) The committee chose the compromise "Hungarian Runic" (WG2 Dublin 2009-04).
4) "The Rovas side" objected, insisting on "Szekely-Hungarian Rovas".
5) The committee switched back to "Old Hungarian" (WG2 Helsinki 2011-06).
6) "The Rovas side" objected, insisting on "Szekely-Hungarian Rovas".
7) The committee chose the compromise "Hungarian" (WG2 Chiang Mai 2012-10).
8) "The Rovas side" has objected *yet again*, rejecting "Hungarian".

All that the objections of "the Rovas side" do is prevent genuine users of this script from making use of it on computers in a standardized way. It is nearly five years since the meeting in Budapest—at the end of which there was general agreement on the proposal, which was subsequently drawn up as ISO/IEC JTC1/SC2/WG2 N3483 "Preliminary proposal for encoding the Old Hungarian script in the UCS". Delays occasioned by repeated considerations by "the Rovas side" do not serve the majority of users of this script.

Document L2/13-049 makes the following unsupported claims:

1) The name of the script is erroneous and contradictory.

Ireland agrees that “Hungarian” on its own is problematic, since the usual script used for Hungarian is Latin. It is not clear to us what is “contradicted” by it, but we support shifting to “Old Hungarian” or “Szekler-Hungarian” as specified above. The name “Hungarian” on its own for this script is simply not found in the literature, and the name “magyar írás” seems to refer, in Hungarian, to the Latin alphabet as used for the Hungarian language. We note that “Szekler” does not require an accent where “Székely” ought to have one. Moreover, the name “Szekler” is more widely-used than “Székely” in English-language materials (Encyclopaedia Britannica, OED, etc.). Indeed the Szekler National Council http://www.sznt.ro/en/ uses this as its name in English.

We also point out that the ad-hoc in Chiang Mai acted pro-actively in changing the name from “Old Hungarian” to “Hungarian” in the hope that “the Rovas side” would find itself in consensus with that term. *Since they reject it, the pro-active change can be said to have failed. It would be better to revert to the previously-used common English-language name for the script, namely “Old Hungarian”.* Failing that, the term “Szekler-Hungarian” could be used, as it has some arguments for it. We have sought clarification from the user community as regards which of these two would be preferable. Our first preference has always been “Old Hungarian”, and the preference of our contacts in the user community is also for “Old Hungarian”.

2) Significant part of the character names are erroneous.

This comment refers to the preference of “the Rovas side” to Latin-alphabet letternames (A, AA, B, C, CS, D etc). The character names under ballot are based on the native names of the characters (A, AA, EB, EC, ECS, ED, etc) attested in primary documents (the Bologna MS, the Rudimenta MS, the Nikolsburg MS). The use of these names is not “erroneous”: it is well-justified and has been supported in working-group ad-hoc reports:

- ISO/IEC JTC1/SC2/WG2 N4110R “Hungarian Runic/Szekely-Hungarian Rovas Adhoc Report” 2011-06-08
- ISO/IEC JTC1/SC2/WG2 N4374R “Old Hungarian/Szekely-Hungarian Rovas Ad-hoc Report” 2012-11-12

These affirm the preference of the standardization committee for the names as presented in the primary source materials for the script. Those names are not “erroneous”.

3) Significant part of the glyph shapes of the characters are erroneous.

No explicit discussion of the glyph shapes of the characters has ever been offered by “the Rovas side”, and in fact previous ballotted charts used a font distributed by them.

- ISO/IEC JTC1/SC2/WG2 N4196 “Code chart fonts for Old Hungarian”)

Examination of a wide variety of non-UCS fonts available shows that the glyphs in the code table are well within the range of acceptable glyphs for this script. They are not “erroneous”.

4) The order of the characters is erroneous.

The order of characters is based on that found in the primary source materials, as in the Nikolsburg MS (dated 1483). In general this follows the order of the Latin alphabet as used for Hungarian, while inserting a few ligatures and homorganic nasals into the sequence near their non-nasalized base characters (A, AA, EB, AMB, EC, ENC, ECZ, ED, AND. […] EZ, EZS, […]) This order is well-justified and, in fact, useful for finding text in ordered word-lists. It is not “erroneous”.

“The Rovas side” has never given a justification or descripion of its preferred order, but it is significant to note that the order presented in N4367 “Revised proposal for encoding the Rovas in the UCS” 2012-10-14 (A, AA, B, C, CS, D, […] Z, ZS, AMB, AND, ANT, EMP, […]) is completely different from the order presented in N4183 “Revised proposal for encoding the Szekely-Hungarian Rovas, Carpathian Basin Rovas and Khazarian Rovas scripts into the Rovas block in the SMP of the UCS” 2012-01-11 (A, AMB, AND, ANT, AA, B, C, CS, D, […] Z, ZS).

5) The deficiency of the code set is substantially restricts the comprehensive use.

A number of mapping tables have been made demonstrating that the overwhelming majority of characters requested in the competing proposals is the same.

- ISO/IEC JTC1/SC2/WG2 N4042 "Mapping between Hungarian Runic proposals in N3697 and N4007" 2011-05-08
- ISO/IEC JTC1/SC2/WG2 N4064 "Comparison of Hungarian Runic and Szekely‐Hungarian Rovas proposals" 2011-05-07
- ISO/IEC JTC1/SC2/WG2 N3532 "Mapping between Old Hungarian proposals in N3531, N3527, and N3526" 2008-11-02)

In general, "the Rovas side" has never responded to the specific technical arguments given in these mapping documents, nor to the decisions made and presented in the adhoc meeting reports. In particular, characters proposed by them but not accepted for encoding were not accepted due to a lack of convincing data and argument regarding their existence.

"Comprehensive use" of this script is well-served by the set of characters under ballot.

Accepted in principle
The name change from 'Hungarian' to 'Old Hungarian' is accepted, but the block will be moved to the 4th edition of 10646 to allow further technical review.
Otherwise, the answers by the Irish NB to the various claims expressed in WG2 4422 are noted.

T2. Page 87, Row 1E80: Mende.

With reference to ISO/IEC JTC1/SC2/WG2 N4396 "Rationale for script name change from Mende to Kikakui", and after consultation with script expert Konrad Tuchscherer cited in that document, Ireland requests the change of the name of the script from MENDE to MENDE KIKAKUI in both the block names and the character names. Ireland would not, however, support the change of the name of the script to KIKAKUI alone..

Accepted
See also comment TE2 from US.
The block name and character names are changed from 'MENDE' to 'MENDE KIKAKUI'. Note that this also requires a change to the Amendment title.

Editorial comments

E1. Page 49, Row 1035: Old Permic.

Ireland requests a change to the glyph of COMBINING OLD PERMIC LETTER ZATA so that it centres better over the dotted circle.
Accepted

E2. Page 86, Row 16B0: Pahawh Hmong.

Ireland requests that the following informative notes be used for two of the Pahawh Hmong characters:
𖭰 16B5E PAHAWH HMONG NUMBER MILLIONS
= roob
𖭱 16B5F PAHAWH HMONG NUMBER HUNDRED MILLIONS
= neev
Accepted

As a result of these dispositions Ireland changed its vote to Yes.

Japan: Positive with comments

General comment

Note that Japanese comments on this sheet, except the two, one on (a part of) the title and another on Siddham which is written as a comment against Clause 31, are all contained in the comment to CD 10646 (4th Ed.)

Technical and Editorial comments (noted as T or E)

E1. Title

The title of this amendment contains a phrase "Old Hungarian", although the corresponding names, such as character names, the block name, and the collection name, in this amendment are HUNGARIAN (no "Old".)

Proposed change by Japan
Remove the word "Old" from the title.

Accepted

E2. Top of first page

There is a title of the standard with an extra dash at its end as follows:
Information technology — Universal
Coded Character Set (UCS) —

Proposed change by Japan
Remove the extra dash.
Accepted

E3. Page 6, Sub-clause 16.5 – Variation selector sequences

The second sentences of the third list item (for Phags-pa variation sequences) and the fifth list item (for CJK Unified Ideographs variation sequences) include a phrase "variation selector sequences". It should be "variation sequences" (without "selector").

Proposed change by Japan
Replace "variation selector sequences" with "variation sequences" (removing "selector".)
Accepted

T4. Page 6, Sub-clause 16.5 – CJK Compatibility Ideographs variation sequences

The current draft says that the newly introduced standardized variation sequences for CJK Unified Ideographs are equivalent to CJK Compatibility Ideographs and that they are preferred representation (over CJK Compatibility Ideographs), but such statements are misleading.

The intention of this list item appears that "the visual appearances specified by these variation sequences are that of CJK compatibility ideographs" and that "if an application needs to normalize the text data, and it needs to distinguish compatibility ideographs and corresponding unified ideographs after the normalization, then use of the standardized variation sequences for CJK Unified Ideographs may help."

It is better to say the point simply. Note that the second sentence is just a hint to the users and not a requirement, and it appears better to be written as a part of the NOTE.

Proposed change by Japan
Replace the list item with the following:

- CJK Unified Ideographs. Each of these variation sequences corresponds to a CJK compatibility ideograph. Its specified appearance is that of the corresponding CJK compatibility ideograph.

Replace the NOTE 7 to the list item with the following:

> NOTE 7 – If an application normalizes text data containing CJK compatibility ideographs, the CJK compatibility ideographs are replaced with the corresponding CJK unified ideographs, and the distinction between the two is lost. It makes lossless two-way code conversion impossible. On the other hand, variation sequences are unchanged by normalization process. If an application needs normalization, and it needs to distinguish appearances of CJK compatibility ideographs and corresponding CJK unified ideographs, use of the standardized variation sequences for CJK Unified Ideographs in place of CJK compatibility ideographs may be a solution. No equivalence between these variation sequences and the corresponding compatibility ideographs are defined. Conversion considerations are out of scope of this International Standard.

Partially accepted

The list item replacement is accepted as it is. However the proposed note needs to be altered to show that the use of normalization is more prevalent than suggested by Japan and is often beyond the control of applications. The new note reads as follows:

> NOTE 7 – All normalization forms replace CJK compatibility ideographs with the corresponding CJK unified ideographs, but leave the variation sequences unchanged (see 21). In contexts where normalization forms are used and the distinction between the CJK compatibility ideographs and CJK unified ideographs is desired, the usage of variation sequences is a mechanism to maintain that distinction. No equivalence between these variation sequences and the corresponding compatibility ideographs are defined. Conversion considerations are out of scope of this International Standard.

T5. Page 6, Clause 18 Compatibility characters – Note 3

Normalization and compatibility ideographs are, in a sense, incompatible in both ways. Stating this fact from one side will mislead users.

Also, the current sentence uses a vague phrase "the distinct identity of compatibility characters". Variation sequences are neither compatibility characters nor compatibility ideographs. As the standard says, variation sequences only specify appearance.

There are some other problems in the current sentences: the NOTE 3 uses a phrase "compatibility characters" although the message strictly aims to users of compatibility ideographs as opposed to general compatibility characters.

Proposed change by Japan

Replace the NOTE 3 with the following:

> NOTE 3 - Because compatibility ideographs are not preserved through any normalization forms, use of standardized variation sequences for CJK Unified Ideographs (See 16.5) may be better if the application needs to perform normalization and the distinction between CJK compatibility ideographs and the corresponding CJK Unified ideographs needs to be preserved. Another alternative is to avoid normalization at all.

Partially accepted

While normalization forms and compatibility ideographs are in a sense incompatible as stated by Japan it is not true that it is only stated from one side. Both the compatibility clause (18) and the normalization form clause (21) mention that situation. If there is bias toward normalization, it is because it is now prevalent in many contexts. And it is also why many experts are reluctant to encode more compatibility ideographs. Furthermore, variation sequences with definitions such as '7DF4 FE00; CJK COMPATIBILITY IDEOGRAPH-F996' might not be compatibility ideographs but they are clearly specified to preserve the concept of compatibility ideographs through context where normalizations forms are used. While variation sequences are clearly intended to specify appearance there is nothing that prevents them to be used to create a distinction between a regular character and

its compatibility 'equivalent'. Variation sequences may not be the perfect vehicle to preserve the compatibility concept (including round-tripping where normalization forms are prevalent) but it was felt that using variations sequences avoided the introduction of a whole new mechanism to preserve the separate identity of compatibility ideographs. A new Note 3 is proposed as follows:

> NOTE 3 - Because compatibility ideographs are not preserved through any normalization forms, use of standardized variation sequences for CJK Unified Ideographs (see 16.5) may be preferred in contexts where normalization forms are used and the distinction between CJK compatibility ideographs and the corresponding CJK Unified ideographs needs to be preserved. In context where compatibility ideographs should be preserved normalization forms cannot be used.

T6. Page 7, Clause 21 – Normalization forms, Note 4

The NOTE begins with "Because normalization forms preserve the variation selectors", assuming the reader knows it and the reader also understand normalization replaces some compatibility characters, specifically CJK compatibility ideographs, with the corresponding characters, although it is not always the case. 10646 doesn't explain normalization procedure and does refer to the Unicode Standard, so this NOTE is better to explain more on the point.

Also, this NOTE tells the user only one side of the issue. Doing so is misleading.

Proposed change by Japan
Replace the NOTE 4 with the following:

> NOTE 4 - In all of the four normalization forms, CJK Compatibility Ideographs are replaced with the corresponding CJK Unified Ideographs. Normalization, however, doesn't alter variation selectors, and variation sequences are preserved. Because of this, it may be better to use standardized variation sequences for CJK Unified Ideographs than to use CJK Compatibility Ideographs, in the context of normalization (See 16.5). In other words, if an application needs to use CJK Compatibility ideographs and the distinction between the corresponding CJK Unified Ideographs need to be preserved, use of normalization should be avoided.

Partially accepted

Explaining in better terms the situation between normalization forms and variations selectors/sequences is a good thing. However presenting this is a one-sided presentation is in itself misleading. Stating that an option is that normalization should be avoided is unrealistic. In many contexts the benefit of normalization forms are such that they are prevalent and applications have no control on the data set they are served.

Furthermore the proposed sentence: <<In other words, if an application needs to use CJK Compatibility ideographs and the distinction between the corresponding CJK Unified Ideographs need to be preserved, use of normalization should be avoided. *>> is not accurate. The whole idea of the new CJK unified ideographs variation sequences is to allow maintaining the distinction between CJK compatibility ideograph and CJK unified ideograph without using CJK compatibility code points.*

A new Note 4 is proposed as follows:

> NOTE 4 - In all of four normalization forms, CJK Compatibility Ideographs are replaced with the corresponding CJK Unified Ideographs. Normalization, however, doesn't alter variation selectors, and variation sequences are preserved. Because of this, the use of standardized variation sequences for CJK Unified Ideographs over the CJK Compatibility Ideographs is preferred in the context of normalization (see 16.5).

T7. Page 63, Clause 31 – Siddham block

The new Siddham block may be incomplete. See WG2N4407 for details. Japan wants WG2 to consider the proposal in N4407.

Proposed change by Japan
(No concrete change proposal is given at this time.)

Noted

See also comment TE1 from US.

E8. Page 8, Sub-clause 31.2 – Characters name list

The text says a TILDE precedes a variation sequence in the name list. However, in the actual name list, a SWUNG DASH does. The definition text and the actual name list should use the same character.

Proposed change by Japan

Change the TILDE sign that appears in "Variation sequences preceded by '~'," to a SWUNG DASH sign.

Accepted

E9. Page 8, Sub-clause 31.2 – EXAMPLE

The amendment introduces into the character name list a new sign (whichever TILDE or SWUNG DASH it is). The EXAMPLE for 31.2 should also be amended to show the usage of the new sign.

Proposed change by Japan

Add a new amendment text for 31.2 to add an appropriate part from the name list to show the use of "~" signs, e.g., a name list entry for 1820 (MONGOLIAN LETTER A), into the EXAMPLE.

Accepted

T10. Page 10, Annex I.1 – Syntax of an ideographic description sequence

The draft updates the definition of IDS by allowing private use characters as its DCs. Although Japan understands a requirement to allow something unencoded in UCS as a DC, it is afraid of opening up an unrestricted distribution of data containing private use characters.

Yes, IRG did use some private use characters as DCs in its own use of IDC-look-alikes, it already caused some problems even in IRG works; many IRG editors misunderstood what shapes those particular private use characters were meant, because their PC showed a different private use characters in place. In practice, it is not easy to detect a given text data contained any private use characters.

Japan considers it was a mistake that we used private use characters in IRG works. Japan worries about the issues IRG experienced may confuse world-wide UCS users.

As an alternative to private use characters, Japan would like to propose use of REPLACEMENT CHARACTER to represent a DC that is not encoded in UCS. REPLACEMENT CHARACTER is better than private use characters in the following ways: REPLACEMENT CHARACTER is expected to appear as its own glyph, that is very unlikely to be mistakenly recognized as an intended component of an ideograph by a receiving person. On the other hand, a private use code point may, by accident, have some ideograph-like character assigned by the receiver-side PC, and the receiving person may not be aware of the use of private character in the IDC, while he/she sees totally different shape than the sender's.

Proposed change by Japan

Replace the following list item to be inserted

"a private use character (as long as the interchanging parties have agreed that the particular private use character represents a particular ideograph or component of an ideograph)"

with the following:

"FFFD REPLACEMENT CHARACTER"

Partially accepted

See also comment K1 by South Korea (ROK) and its disposition.

The concern about IRG editors not being able to communicate effectively the information using private use characters is valid. However, the purpose of the new formulation is to make the use of Private Use characters conformant in that context, not to encourage their usage. This is not different from usages of Private Use characters for other purposes. It is up to the IRG group to determine its own policy concerning the use or not of Private Use characters for their own context.
At the same time, it is useful to indicate an un-encoded DC (Description Component) by a special character as suggested by Japan. However the character U+FF1F ？ FULLWIDTH QUESTION MARK is preferred. To that effect the following item will be added in sub-clause I.1in the first list (after 'a coded radical…'):

- the character FF1F FULLWIDTH QUESTION MARK to represent an otherwise un-described Description Component.

South Korea (ROK): Negative

Technical comment:

T1. Page 10, Annex I.1 – Syntax of an ideographic description sequence – New item in list
[The propose change is to add a new Description Component (DC) as "a private use character (as long as the interchanging parties have agreed that the particular private use character represents a particular ideograph or component of an ideograph)]

With this change, ISO/IEC allows users to use PUA chars fairly widely. It does not seem desirable.

Proposed change by ROK:
Do not insert "a private use character\ (as ... ideograph).

OR we could encode those chars in UCS.

Furthermore, it is suggested that the list of those chars (and a font) be supplied somewhere publicly so that people can reference (and use).
Not accepted
See also comment T10 from Japan and its disposition.
Encoding these private use characters used for Description Components is possible and can be done independently of the change in Annex I

As a result of this disposition and discussion during WG2 meeting #61, South Korea changed its vote to YES.

USA: Positive with comments

Technical comments:

TE.1. Latin Extended-D

These characters represent a subset of the Siddham section marks proposed in WG2 N4336 and have been identified as required for representing the text in the Jogon and Annen traditions. Although no names have been identified, the usage of these characters and their attestation has been provided in WG2 N4391.

Proposed change by US:

The US requests the addition of the following 7 Siddham section marks to the Siddham block, currently under ballot, with glyphs and properties as shown on pages 9 and 10 of N4336:

U+115CB SIDDHAM SECTION MARK-2
U+115CC SIDDHAM SECTION MARK-3
U+115CE SIDDHAM SECTION MARK-5
U+115CF SIDDHAM SECTION MARK-6
U+115D0 SIDDHAM SECTION MARK-7
U+115D1 SIDDHAM SECTION MARK-8
U+115D4 SIDDHAM SECTION MARK-11.

Accepted in principle

See also comment T7 from Japan that concerns the same block.
This request was superseded by WG2 N4457with more characters with new names. The repertoire will be added in the 4th edition, not Amendment 2.

TE.2. Mende

The rationale is provided in WG2 N4396.

Proposed change by US:

The US requests the block name for the "Mende" script be changed to "Kikakui" since the Mende language is most commonly written in a Latin-based orthography, and the script name "Kikakui" would be less ambiguous. The character names should also be changed accordingly.

Accepted in principle

See also comment T2 from Ireland.
The block name and character names are changed from 'MENDE' to 'MENDE KIKAKUI'. Note that this also requires a change to the Amendment title.

TE.3. Miscellaneous Symbols and Pictographs

The original CLAPPER BOARD character derived from Emoji, and the design with lines was based on what is generally recognized as a clapper board in Japan. The reverted glyph more closely reflects the original shape of the Emoji character.

Proposed change by US:

The US requests the UCS glyph for U+1F3AC CLAPPER BOARD be reverted back to the glyph with the lines.

Accepted

For reference the two glyphs are shown above (first, glyph prior to this amendment and second, glyph changed by this amendment)

Editorial comments:

E.1. Page 1 – Title

"Hungarian" is the block name of the script under ballot.

Proposed change by US:

The title of the ballot (page 1) currently reads "Old Hungarian". This should be corrected to "Hungarian."

Accepted

See also comment E1 from Japan.

E.2. Sub-clause 16.5

The current definition of a variation selector states that it only follows a [non] decomposable base character. However, there are bases that contain decomposition mappings (cf. http://unicode.org/Public/6.2.0/ucd/StandardizedVariants.txt). The correction of the above text to "canonical decomposable base character" will correct the error.

Proposed change by US:

The wording in section 16.5 "Variation selectors and variation sequences" of the 3rd edition should be adjusted, adding "canonical" before "decomposable base character" (3 instances) in the following text:

> Variation selectors are a specific class of combining characters immediately following a non decomposable base character and which indicate a specific variant form of graphic symbol for that character. A decomposable character is a character for which there exists an equivalent composite sequence. The character sequence consisting of a non decomposable base character followed by a variation selector is called a variation sequence.

Accepted

The reference in this ballot should have been UCSVariants.txt, not the hyperlink referenced above although their content is identical. Furthermore it may be useful to provide at least an example:

24C2 FE0E; text style; # CIRCLED LATIN CAPITAL LETTER M

In this example, U+24C2 has a decomposition mapping to a compatibility equivalent (U+004D but not to a canonical equivalent. No entry in that file has a decomposition mapping to a canonical equivalent.