Universal Multiple-Octet Coded Character Set
International Organization for Standardization
Organisation Internationale de Normalisation
Международная организация по стандартизации

**Doc Type: Working Group Document**
**Title: Proposal to add Variation Sequences for Latin and Cyrillic letters**
**Source: Karl Pentzlin**
**Status: Expert Contribution**
**Action: For consideration by JTC1/SC2/WG2 and UTC**
**Date: 2011-02-07**
**Replaces: L2/10-280 dated 2010-08-05**

*This revision was done after discovering that L2/10-280 is on the agenda for UTC #126.*
*Compared to the original paper, some of the proposed variants are dropped:*
*— All variation sequences which were related to Jaŋalif, as in the meantime, the author prefers the*
*application of special tailoring to solve the casing issues there.*
*— All variation sequences which were related to the English Phonotypic Alphabet, as the current*
*proposal on that subject (L2/11-040) addresses the casing issues there in a different way.*
*— All variation sequences which were related to 19th century Slovenian alphabets; the issues regarding*
*these characters will be addressed when these characters in fact will be proposed.*

*Besides these droppings, the text is unchanged, to preserve the references in the other papers which*
*are on the UTC #126 agenda referring the same topic.*

# Overview

This proposal consists of two parts:

Part A discusses the use of variation sequences with scripts like Latin and Cyrillic in general, and proposes some principles to be considered when defining such variation sequences in accordance with the rules given in the current Unicode standard 5.2 .

Part B proposes a set of variation sequences for Latin and Cyrillic.

# Part A:
# Benefits of Variation Sequences for scripts like Latin and Cyrillic, and proposals regarding Variation Sequences in general

### 1. The Problem

Unicode encodes characters, not glyphs.
However, there are cases where particular user groups, especially cultural or linguistic groups, recognize only a part of the possible glyphs as adequate representations of that character.

For instance, uppercase forms of newly introduced Latin letters are developed by several user communities separately from their lowercase counterparts, yielding different forms.

Each of these user communities regards "their" forms as adequate, while the other forms may be recognized but perceived as "typographically wrong", "silly", or "cultural ignorant".

As an illustrating example, the letter Ŋ U+014A CAPITAL LETTER ENG has valid glyphs like ŋ. The set of such glyphs constitute the variant which some users in Africa use and expect in their texts (see fig. 6). However, when a user acquainted with the standard form Ŋ encounters this, he will recognize it but will find it silly, like a speaker of Ənglish is expected to regard an Ə-like glyph as an inadequate form of E.

In some cases, this occurs only for text in a specific style (like italic or Roman-serifed). A well-known example are the italic glyphs of some Cyrillic letters, where Serbians use different ones than Russians (see fig. 24, 25).

Requests for getting the wanted form are usually answered by "select a special font for your needs, which provide your preferred glyphs for the affected characters".

However, in these days, "universal fonts" are common which try to serve a part of the user community as large as possible. Such fonts, by their nature, do not give preference to the special needs of any particular groups which may be incompatible with the needs of other parts of the user community.

Especially, such "universal fonts" serve the needs of the Internet community, where the site designer usually cannot rely on the availability of special fonts on the user's side. Instead, the designer relies on the availability of such "universal fonts".
(Technologies to provide fonts on the server side exist, but their use is not widespread. Also, they require technical skills from the site developers, which cannot be assumed in general as internet sites are often created and maintained by amateurs and hobbyists.)

In fact, there are technologies defined which allow a font to contain several glyphs for single characters. But while the general mechanisms for doing so are standardized (i.e. OpenType features), the concrete selection of a specific glyph is not. It is up to the font designer to decide in which way to handle any locale specification provided by a higher level protocol, or which feature called with which parameters are required to select a glyph.

Moreover, it cannot be expected that every font designer is able to incorporate such features in their font. Font designers, especially ones who offer their work for prices affordable to non-commercial users or for free, usually are experts of their subject (i.e. as linguists), and additionally have to be experts of typographic design when the font shall be successful. To be also an expert of a third demanding area— Programming of OpenType features—usually is a burden a little too heavy for a single person or a small team. Professional fonts which are developed by large teams with professionals of all three areas are expensive, and therefore only of a very limited importance for the general public.

Also, the support of such features is restricted today to expensive desktop publication applications like Quark XPress® or Adobe InDesign®. Common text processing applications like older versions of Microsoft Word® do not support OpenType features at all for Latin and Cyrillic (at least, the most recent 2010 version of MS Word starts to support a somewhat limited support of some OpenType features).

Additionally, it must be noted that it cannot be expected that a font incorporates all of the hundreds of different cultural contexts defined as locales, like it cannot expected that all possible cultural contexts are formally defined as locales. Thus, it is a far simpler task to handle some 50 clearly defined variation sequences found in an easily accessible list, than hundreds of locales which require special knowledge to extract the needed information out of them.
In fact, locales now can simply specify the default application of a variation sequence to a given base

character. Thus, a rendering system simply has to request this simple information from the active locale, instead of having the glyph selection rules incorporating repeatedly in every single font.
Thus, the introduction of variation sequences provides an abstraction level which simplifies the glyph selection process in a really significant way.

Also, requesting the user to select an "appropriate" font to get their character to be displayed correctly is a fallback into the era when special fonts were needed to overcome the drawback of limiting the number of characters within a font by 8-bit codes. In that times, you e.g. had to use a special IPA font which maps IPA characters on the position of "ordinary" characters in a proprietary way.

To continue requests like "when you want an 'African f with hook' proper, rather than a slanted 'florin symbol', use a special font where U+0192 uses appropriate glyphs", is something like revitalizing the old-fashioned "when you want an IPA 'ɛ', use a special font where this is mapped onto the code position of the 'e' ".

Also, having to supply an (otherwise) "universal" font with several thousand characters in two different versions "Russian" and "Serbian", which only differ in one character in the Regular style and five ones in the Italic style, should be anachronistic in the age of Unicode.


## 2. The Solution

Fortunately, Unicode already provides a solution which addresses this problem: "Variation Sequences".

According to the current Unicode standard (V5.2, chapter 16.4, p. 511):

- *[They] … provide a mechanism for specifying a restriction on the set of glyphs that are used to represent a particular character. They also provide a mechanism for specifying variants, … that have essentially the same semantics but substantially different ranges of glyphs.*

The current Unicode standard (V5.2) describes variation selectors in chapter 16.4:

- *Variation Sequence.* A variation sequence always consists of a base character followed by a variation selector character. That sequence is referred to as a *variant* of the base character. The variation selector affects *only* the appearance of the base character. The variation selector is *not* used as a general code extension mechanism; only certain sequences are defined …

- … The variation selectors themselves are combining marks of combining class 0 and are default ignorable characters. Thus, if the variation sequence is not supported, the variation selector should be invisible and ignored. …

- … The standardization or support of a particular variation sequence does *not* limit the set of glyphs that can be used to represent the base character alone. If a user *requires* a visual distinction between a character and a particular variant of that character, then fonts must be used to make that distinction. …

Thus, the 16 "variation selectors" defined as Unicode characters (U+FE00 VS1 ... U+FE0F VS16) can be regarded as "invisible accents" which affect the appearance of the base letter itself, instead of adding a visible diacritical mark.

The second cited paragraph states that the insertion of variation selectors cannot harm any existing data in any existing Unicode-compliant environment: If the font or the rendering system does not interpret the

variation selector, it simply has to ignore it.
This rule is not new to Unicode: Even now, if a variation selector is encountered following a Latin or Cyrillic letter, it has to be ignored (as, until now, there are no valid variation sequences starting with such a letter).

The third cited paragraph states that a given variation sequence must be supported by the font in use. This is something trivial: If the font does not support the wanted glyph, it cannot display it; instead another glyph which represents the character (or which denotes the missing availability of the character at all) is displayed as fallback.

## 3. Characters vs. Variants vs. Glyphs,  or:  What Variation Sequences are <u>not</u> for

*Variants* are a means for cases where a part of the glyphs, which are valid representations of the character from a universal point of view, are not considered as appropriate glyphs by specific user groups.
Regarding letters, such user groups usually are groups adhering to a cultural context (defined by language or other criteria).

As a special case, for a character, different uses with different glyph designs may have developed since its encoding, which would justify a disunification today into two (or more) characters. When doing such a disunification is prevented by the quantity of existing data which would be hampered by such a disunification, these uses can be distinguished by introducing new variation sequences, without violating any Unicode stability policy.
(An example are the variation sequences proposed in this paper for U+0192 LATIN SMALL LETTER F WITH HOOK.)

While variants (like character themselves) have a set of glyphs which can represent them (depending of position, style, script variant, etc.), *they are not the glyphs themselves*.
The single difference to a character itself is that for a variant, the set of glyph is usually smaller, and/or that another glyph of the set is the appropriate selection for the representative glyph.

- *This means, a variant is more like a character than like a glyph.*
  As a rule of thumb, a variant is only provided instead of encoding a separate character,
  – either, if the fundamental identity of a single character is given on other grounds (especially when the different variant is tied to the same counterpart by case pairing),
  – or, because the divergent use has emerged after the encoding of the character was done, and a disunification is not appropriate for any reason.

As an example, assume a high-quality font which provides 8 glyphs a1…a8 (contextual glyph variants or whatever), which all are appropriate for "a". Then, selection of VARIANT-1 for "a" may mean (by decision of the font creator) e.g., that the glyphs a3 and a5…a7 and an extra glyph a9 may be appropriate instead (to be selected at last by the context or whatever).
In no case, t"VARIANT-1" is tied a priori to any "glyph no. 1" which may be defined in such a font.

Having said this, variation sequences are a means as inappropriate as character encoding itself for selection of specific glyph variants in specific fonts, like beautiful or cute glyphs, or even specific glyphs in fonts designed for teaching to read and write.

## 4. Application of Variation Sequences

A variant of a character may be selected:

- explicitly, by inserting the code sequence constituting the variation sequence into the text,

- implicitly, by a higher-level protocol, which instructs the rendering system to interpret any occurrence of the base character within its domain, as if it were be followed by the variation selector which constitutes the variation sequence.

### 4.1. Explicit Application

Variation sequences are explicitly applied by inserting their code sequence into the text. E.g. to address the Ŋ variant of the latin capital letter eng explicitly, the character sequence U+014A U+FE01 will be entered into the text.

This explicit application is appropriate for all situations where the selection of the correct variant by a higher-level protocol is not ensured.

Such situations are e.g.:
- Multilingual text.
- Plain text stored in databases, where the environment in which such entries are retrieved is not fixed.
- Text published on the Internet or sent by e-mail, where it is not ensured that the numerous software products which may display that text at last will correctly evaluate any locale or similar information joined with the text.

### 4.2. Implicit Application

Higher-level protocols may specify the application of a specific variation sequence for specified base characters within their area of appliance. Besides other sources, such protocols may use locales, which in turn may contain lists of variation sequences considered appropriate for application within the domain of the locale.

Implicit applicability of variation sequences means:

- There is no need for insertion of the variation selectors into the text itself.

- All existing data (as long as they are subject to a single locale) stay valid and do not need any change, but will be displayed using the correct character variants, as long as the correct locale is in effect.

- *E.g., Serbians do neither need to clutter their texts with VS2s after each б/г/д/п/т/ш, nor to use a specific version of their preferred font (or an expensive one which does the selection inside proprietarily), to get their appropriate italic forms.*

At this time, as no variation sequences for non-ideographic scripts are given yet, probably there is no higher-level protocol in use which advises the appliance of variation sequences. However, once the mechanism of variation sequences is available, the presumption that such higher-level protocols will come in use is plausible.

### 4.3. Overriding Application

An explicit application of a variation selector overrides an implicit one.
Thus, it is possible to access any variant of a given character even in the domain of the implicit appliance of another variant of the same character.

This implies that it is recommended to define not only the "deviating" variant of a character which is applicable for a specific cultural context, but also explicitly give a variation sequence for the "standard" variant, if such exists in contrast to the "deviating" one.

### 5. Benefits of Variation Sequences

- The possible variants of any characters are clearly documented, easily accessible in the Unicode standard.

*Benefits for font designers:*

- A font designer can easy implement them, without needing to understand all of the cultural subtleties of the user groups spread over the whole world, who are served with these variants.

- The font designer only has to use a single special feature (to select a single glyph based on a two-code sequence, rather than based on a single character code).

- A font designer can serve all user groups who need a specific variant **in a standardized way**, without providing drawbacks to the people who do not want this specific variant.

*Benefits for users:*

- Users can access their preferred variants using "universal fonts", without any change on their data, relying on their selected locale specification (when the "implicit application" is in effect as described in the preceding section).

- Users can access any variant using "universal fonts", by explicit selection of a variation sequence. This is valid independent of the presence of any higher-level protocols.

- Users can access any variant when they use plain text.
  This is especially important for storing text in (i.e. linguistic) databases, which are to be correctly displayed when retrieved in another environment.

Ken Lunde (Adobe Systems Inc.) addresses the last point in his paper about ideographic variation sequences L2/10-211: "Adobe-Japan1 IVD Collection: Current Status and Future Directions". In this PowerPoint presentation, on the second sheet is outlined:
- **The power, safety, and reliability of "plain text"**
  - The ability to survive or endure in more environments

### 6. Proposal:  Recommendations for the Proposing of New Variation Sequences

*(In this section, read "variation sequence" as "non-ideographic variation sequence" in all cases.)*

a.  A variation sequence shall be proposed only if the fundamental identity of the base character is obvious, or if an otherwise appropriate disunification of an already encoded

character cannot be done due to existing data.
In any other case, a separate character shall be proposed.

Typical appliances are:

- A lowercase character has developed similar but different uppercase forms in different cultural contexts.

- A character (or a case pair) has developed a specific appearance in a cultural context, but is regarded by the members of that cultural group definitely as the same character, and is accepted by them also when it appears in its original form.

b. It is possible to propose variation sequences when the selection or restriction of glyphs or a set thereof applies only for special positional forms (initial, medial, final, isolated), styles (serifed, sans-serif, italic, …), or script variants (Fraktur/Blackletter, Gaelic, Church Slavonic, …).

c. Whenever the selection or restriction of glyphs obtained by a proposed variation sequence is one usually applied uniformly to all characters in a longer text passage (such as selecting/restricting glyphs according to a locale or a cultural context; rather than to denote the use of different variants within a text passage), the existence of a higher-level protocol shall be assumed which does this selection/restriction implicitly.

In consequence of this, there shall be other variation sequences for the same base character (proposed or existing), by which the implicit appliance of the given variation sequence done by the higher-level protocol can be overridden by explicit appliance. *In doing so, the complete set of variants for a given base character stays available even when a higher-level protocol causes the implicit appliance of a variation sequence.*

d. Whenever more than one variation sequence is given for a base character, and one of them selects the reference glyph of the base character as it is given at the time of the proposal, this variation sequence shall use the variation selector VS1 if applicable.

e. Whenever a variation sequence is proposed for a letter of a bicameral script, the same variation selector (VS1, ...) shall not be used for the counterpart letter of the other case, unless it defines there a variation sequence which usually is used in the same context with the same meaning. (In other words, use the same variation selector for a case pair only if the "results" can be regarded as a case pair again to be used within the intended language, cultural context, etc.).
This shall also apply for known special case pairings (defined by locales etc.) which deviate from the ones given by the character properties, as far as possible.

*This ensures the invariance of variation selectors on simple case conversion: Either the conversion result is the variation sequence which was intended to use anyway, or the result is a sequence which is not defined as a variation sequence, in which case the variation selector is simply to be ignored according to the Unicode rules.*

To accomplish this, gaps in the variant number sequence of a character are allowed.

## 7. Proposal:  Naming of Variation Sequences

It is proposed that each variation sequence, which is not an ideographic variation sequence, is given a unique name, like a character.
This name consists of three consecutive parts, separated by a space:

a.  The name of the character with which the variation sequence starts;
    as an exception, if that character has an Normative Alias at the time the variation sequence is proposed, this Normative Alias is used instead.
    *Remark: It is likely that variation sequences are proposed which start with U+01A2 LATIN CAPITAL LETTER OI (which has the Normative Alias LATIN CAPITAL LETTER GHA), when it comes to encode historical Latin alphabets used in the area of the Former Soviet Union.*

b.  The term VARIANT-n, where n represents the decimal number (1...16) of the used variant selector (U+FE00 VS1 ... U+FE0F VS16), or "M" followed the decimal number (1 … 3) of the used Mongolian free variant selector (U+180B FVS1 … U+180D FVS3).

c.  Something which is associated with the form and/or the purpose of the variant.

As a whole, the name must adhere to all requirements which are valid for a Unicode character name.

Also, each non-ideographic variation sequence automatically constitutes a Named Character Sequence with the same name, denoting the sequence of two Unicode characters which constitute the variation sequence.

In conjunction with the naming rules above, it is proposed that names of characters, and names of Named Character Sequences which are not variation sequences, must not contain parts which begin with "VARIANT-". (This is compatible with the status quo, as no character of Named Characte Sequence currently encoded contains such a part.)

## 8. Proposal:  Stability Policies

It is proposed that adequate stability policies are established which prevent any accepted non-ideographic variation sequence from deleting, changing, or renaming.

## 9. Proposal:  Documentation

At this time, non-ideographic variation sequences are documented in the file StandardizedVariants.txt available in a subdirectory of the Unicode internet site.

To increase the usability for the average user who has access to the text of the standard but is not in comfort with the finding and usage of such files, it is proposed to list the non-ideographic variation sequences also in the Standard text as follows:

a.  In the Unicode code tables, in each block, following the character names list, there is a list of all non-ideographic variation sequences, whose first character is in that block (if such exist). Each such entry consists of:
    1.  the "representative glyph" of the variation sequence, in standard style,
    2.  if the variation sequence concerns a glyph variation which is only relevant in a speci-

fic style (serifed, italic, ...) or script variant (Fraktur, Gaelic, Church Slavonic, ...),
a second glyph in that style or subscript which is representative there.
The style or script variant should be identified in an informative note.
If positional variants are concerned, more than one glyphs may be shown here.

3. the code sequence (consisting of the base character and the variation selector)
4. the name of the variation sequence,
5. if applicable, followed by annotations in separate lines, in the same way as characters in the character names list.

b. Additionally, at the end of each block, there is a list of all Named Character Sequences, whose first character is in that block, and which are not also listed as variation sequences there (if such exist). The format is the same as for the variation sequences list (only that there are no second glyphs).

# Part B:
# Proposal to define 43 Variation Sequences for Latin and Cyrillic

## 1. Introduction

Here, a total of 43 variation sequences for 21 characters (14 Latin, 7 Cyrillic) is proposed. All of them reflect usage in special cultural contexts.

## 2. Encoding Considerations

Regarding the naming, the term ALTERNATIVE in a name denotes that the use of the variant in the addressed cultural context is not thorough.

Regarding the ANGULAR variants of U+01B2 and U+028B, the term "WITH HIGH HOOK" is included in the name as the fonts Doulos SIL and Charis SIL (available at http://www.sil.org ) show a third variant, which is angular but a hook height like the ROUND variants. These are not included here as at this time, no proof of use is at hand, but they may be proposed later. Then, the appropriate naming for these will be "... VARIANT-3 ANGULAR FORM" with no additional term in accordance with the ROUND forms.

## 3. Proposed Variation Sequences

a *a*    0061 FE00    LATIN SMALL LETTER A VARIANT-1 TWO-STOREY FORM
· two-storey form also used when italic (e.g. IPA use)
· in contrast to 0251 latin small letter alpha

g *g*    0067 FE00    LATIN SMALL LETTER G VARIANT-1 LOOP-DESCENDER FORM
· loop-descender form also used when sans-serif style or italic
· in contrast to 0261 latin small letter script g

Ħ    0126 FE00    LATIN CAPITAL LETTER H WITH STROKE VARIANT-1 STANDARD FORM
· showing a horizontal protruding bar above of the central H bar

Ħ    0126 FE01    LATIN CAPITAL LETTER H WITH STROKE VARIANT-2 FORM WITH VERTICAL STROKE
· showing a vertical bar crossing the central H bar in the middle
· used in the Judeo-Tat language

Ŋ 014A FE00 LATIN CAPITAL LETTER ENG VARIANT-1 CAPITAL-N WITH HOOK BELOW FORM

ŋ 014A FE01 LATIN CAPITAL LETTER ENG VARIANT-2 ENLARGED SMALL-LETTER FORM

ŋ 014A FE02 LATIN CAPITAL LETTER ENG VARIANT-3 ENLARGED SMALL-N
　　　　　　 WITH INWARDS HOOK FORM

Ɓ 0181 FE00 LATIN CAPITAL LETTER B WITH HOOK VARIANT-1 STANDARD FORM

Ƃ 0181 FE01 LATIN CAPITAL LETTER B WITH HOOK VARIANT-2 TOPBAR FORM
　　　　　　 · resembles 0182 latin capital letter b with topbar
　　　　　　 · used in the Toma and Dan/Gio languages of Liberia

Ɔ 0186 FE00 LATIN CAPITAL LETTER OPEN-O VARIANT-1 STANDARD FORM
　　　　　　 · in serifed styles, the serif usually is at the bottom

Ɔ 0186 FE01 LATIN CAPITAL LETTER OPEN-O VARIANT-2 TOP-SERIF FORM
　　　　　　 · in serifed styles, the serif is at the top
　　　　　　 · in sans-serif styles, there is no difference to variant-1

ƒ 0192 FE00 LATIN SMALL LETTER F WITH HOOK VARIANT-1 FLORIN SIGN FORM

ƒ 0192 FE01 LATIN SMALL LETTER F WITH HOOK VARIANT-2 LATIN LETTER FORM

Ɲ 019D FE00 LATIN CAPITAL LETTER N WITH LEFT HOOK VARIANT-1 CAPITAL-N WITH HOOK BELOW FORM

Ɲ 019D FE01 LATIN CAPITAL LETTER N WITH LEFT HOOK VARIANT-2 ENLARGED SMALL-LETTER FORM
　　　　　　 · used in some languages of Guinea and the Central African Republic

Ɵ 019F FE00 LATIN CAPITAL LETTER O WITH MIDDLE TILDE VARIANT-1 FORM WITH STRAIGHT BAR INSIDE
　　　　　　 · resembles 04E8 cyrillic capital letter barred o

Ɵ 019F FE01 LATIN CAPITAL LETTER O WITH MIDDLE TILDE VARIANT-2 FORM WITH TILDE INSIDE

ƥ 01A5 FE00 LATIN SMALL LETTER P WITH HOOK VARIANT-1 FORM WITH LEFT HOOK ABOVE

ƥ 01A5 FE01 LATIN SMALL LETTER P WITH HOOK VARIANT-2 FORM WITH RIGHT HOOK ABOVE
　　　　　　 · used in the Serer-Sine language of Senegal

Ʋ 01B2 FE00 LATIN CAPITAL LETTER V WITH HOOK VARIANT-1 ROUND FORM

Ʋ 01B2 FE01 LATIN CAPITAL LETTER V WITH HOOK VARIANT-2 ANGULAR FORM WITH HIGH HOOK
　　　　　　 · used in the Toma language of Guinea

Ʒ 01B7 FE00 LATIN CAPITAL LETTER EZH VARIANT-1 FULL-HEIGHT FORM

Ʒ 01B7 FE01 LATIN CAPITAL LETTER EZH VARIANT-2 CAP-HEIGHT FORM

ɔ 0254 FE00 LATIN SMALL LETTER OPEN-O VARIANT-1 STANDARD FORM
　　　　　　 · in serifed styles, the serif usually is at the bottom

ɔ 0254 FE01 LATIN SMALL LETTER OPEN-O VARIANT-2 TOP-SERIF FORM
　　　　　　 · in serifed styles, the serif is at the top

· in sans-serif styles, there is no difference to variant-1

| | | | |
|---|---|---|---|
| ɵ | 0275 FE00 | LATIN SMALL LETTER BARRED O VARIANT-1 FORM WITH STRAIGHT BAR INSIDE | |
| ɵ | 0275 FE01 | LATIN SMALL LETTER BARRED O VARIANT-2 FORM WITH TILDE INSIDE | |
| ʋ | 028B FE00 | LATIN SMALL LETTER V WITH HOOK VARIANT-1 ROUND FORM | |
| ʋ̂ | 028B FE01 | LATIN SMALL LETTER V WITH HOOK VARIANT-2 ANGULAR FORM WITH HIGH HOOK | |

· used in the Toma language of Guinea

| | | | |
|---|---|---|---|
| Ћ | 040B FE00 | CYRILLIC CAPITAL LETTER TSHE VARIANT-1 STANDARD FORM | |
| Ћ | 040B FE01 | CYRILLIC CAPITAL LETTER TSHE VARIANT-2 FORM WITH STROKE THROUGH ASCENDER | |

· used in some dialects of Juhuri and Tati

| | | | |
|---|---|---|---|
| б *б* | 0431 FE00 | CYRILLIC SMALL LETTER BE VARIANT-1 STANDARD FORM | |
| б *δ* | 0431 FE01 | CYRILLIC SMALL LETTER BE VARIANT-2 SERBIAN FORM | |
| г *г* | 0433 FE00 | CYRILLIC SMALL LETTER GHE VARIANT-1 STANDARD FORM | |
| г *ī* | 0433 FE01 | CYRILLIC SMALL LETTER GHE VARIANT-2 SERBIAN FORM | |

· different from variant-1 only when italic

| | | | |
|---|---|---|---|
| д *д* | 0434 FE00 | CYRILLIC SMALL LETTER DE VARIANT-1 STANDARD FORM | |
| д *g* | 0434 FE01 | CYRILLIC SMALL LETTER DE VARIANT-2 SERBIAN FORM | |

· different from variant-1 only when italic

| | | | |
|---|---|---|---|
| п *n* | 043F FE00 | CYRILLIC SMALL LETTER PE VARIANT-1 STANDARD FORM | |
| п *ū* | 043F FE01 | CYRILLIC SMALL LETTER PE VARIANT-2 SERBIAN FORM | |

· different from variant-1 only when italic

| | | | |
|---|---|---|---|
| т *m* | 0442 FE00 | CYRILLIC SMALL LETTER TE VARIANT-1 STANDARD FORM | |
| т *ш̄* | 0442 FE01 | CYRILLIC SMALL LETTER TE VARIANT-2 SERBIAN FORM | |

· different from variant-1 only when italic

| | | | |
|---|---|---|---|
| ш *ш* | 0448 FE00 | CYRILLIC SMALL LETTER SHA VARIANT-1 STANDARD FORM | |
| ш *ш̱* | 0448 FE01 | CYRILLIC SMALL LETTER SHA VARIANT-2 SERBIAN ALTERNATIVE FORM | |

· different from variant-1 only when italic
· not used thoroughly

## 4. Proposed Names for Existing Non-Ideographic Variation Sequences

As it is proposed in Part A that non-ideographic variation sequences must be named, here is a proposal for naming the existing non-ideographic variation sequences.

The names are derived algorithmically from the descriptions given in:
  http://www.unicode.org/Public/6.0.0/ucd/StandardizedVariants-6.0.0d7.html

```
2229 FE00  INTERSECTION VARIANT-1 WITH SERIFS
222A FE00  UNION VARIANT-1 WITH SERIFS
2268 FE00  LESS-THAN BUT NOT EQUAL TO VARIANT-1 WITH VERTICAL STROKE
2269 FE00  GREATER-THAN BUT NOT EQUAL TO VARIANT-1 WITH VERTICAL STROKE
2272 FE00  LESS-THAN OR EQUIVALENT TO VARIANT-1 FOLLOWING THE SLANT OF THE LOWER LEG
2273 FE00  GREATER-THAN OR EQUIVALENT TO  FOLLOWING THE SLANT OF THE LOWER LEG
228A FE00  SUBSET OF WITH NOT EQUAL TO VARIANT-1 WITH STROKE THROUGH BOTTOM MEMBERS
228B FE00  SUPERSET OF WITH NOT EQUAL TO VARIANT-1 WITH STROKE THROUGH BOTTOM MEMBERS
2293 FE00  SQUARE CAP VARIANT-1 WITH SERIFS
2294 FE00  SQUARE CUP VARIANT-1 WITH SERIFS
2295 FE00  CIRCLED PLUS VARIANT-1 WITH WHITE RIM
2297 FE00  CIRCLED TIMES VARIANT-1 WITH WHITE RIM
229C FE00  CIRCLED EQUALS VARIANT-1 WITH EQUAL SIGN TOUCHING THE CIRCLE
22DA FE00  LESS-THAN EQUAL TO OR GREATER-THAN VARIANT-1 WITH SLANTED EQUAL
22DB FE00  GREATER-THAN EQUAL TO OR LESS-THAN VARIANT-1 WITH SLANTED EQUAL
2A3C FE00  INTERIOR PRODUCT VARIANT-1 TALL VARIANT WITH NARROW FOOT
2A3D FE00  RIGHTHAND INTERIOR PRODUCT VARIANT-1 TALL VARIANT WITH NARROW FOOT
2A9D FE00  SIMILAR OR LESS-THAN VARIANT-1 WITH SIMILAR FOLLOWING THE SLANT OF THE UPPER LEG
2A9E FE00  SIMILAR OR GREATER-THAN VARIANT-1 WITH SIMILAR FOLLOWING THE SLANT OF THE UPPER LEG
2AAC FE00  SMALLER THAN OR EQUAL TO VARIANT-1 WITH SLANTED EQUAL
2AAD FE00  LARGER THAN OR EQUAL TO VARIANT-1 WITH SLANTED EQUAL
2ACB FE00  SUBSET OF ABOVE NOT EQUAL TO VARIANT-1 WITH STROKE THROUGH BOTTOM MEMBERS
2ACC FE00  SUPERSET OF ABOVE NOT EQUAL TO VARIANT-1 WITH STROKE THROUGH BOTTOM MEMBERS

A856 FE00  PHAGS-PA LETTER SMALL A VARIANT-1 REVERSED SHAPING FORM
A85C FE00  PHAGS-PA LETTER HA VARIANT-1 REVERSED SHAPING FORM
A85E FE00  PHAGS-PA LETTER I VARIANT-1 REVERSED SHAPING FORM
A85F FE00  PHAGS-PA LETTER U VARIANT-1 REVERSED SHAPING FORM
A860 FE00  PHAGS-PA LETTER E VARIANT-1 REVERSED SHAPING FORM
A868 FE00  PHAGS-PA SUBJOINED LETTER YA VARIANT-1 REVERSED SHAPING FORM

1820 180B  MONGOLIAN LETTER A VARIANT-M1 SECOND FORM
1820 180C  MONGOLIAN LETTER A VARIANT-M2 THIRD FORM
1821 180B  MONGOLIAN LETTER E VARIANT-M1 SECOND FORM
1822 180B  MONGOLIAN LETTER I VARIANT-M1 SECOND FORM
1823 180B  MONGOLIAN LETTER O VARIANT-M1 SECOND FORM
1824 180B  MONGOLIAN LETTER U VARIANT-M1 SECOND FORM
1825 180B  MONGOLIAN LETTER OE VARIANT-M1 SECOND FORM
1825 180C  MONGOLIAN LETTER OE VARIANT-M2 THIRD FORM
1826 180B  MONGOLIAN LETTER UE VARIANT-M1 SECOND FORM
1826 180C  MONGOLIAN LETTER UE VARIANT-M2 THIRD FORM
1828 180B  MONGOLIAN LETTER NA VARIANT-M1 SECOND FORM
1828 180C  MONGOLIAN LETTER NA VARIANT-M2 THIRD FORM
1828 180D  MONGOLIAN LETTER NA VARIANT-M1 SEPARATE FORM
182A 180B  MONGOLIAN LETTER BA VARIANT-M1 ALTERNATIVE FORM
182C 180B  MONGOLIAN LETTER QA VARIANT-M1 SECOND FORM OR FEMININE SECOND FORM
              ·  second form when initial or medial
              ·  feminine second form form when isolate
182C 180C  MONGOLIAN LETTER QA VARIANT-M2 THIRD FORM
182C 180D  MONGOLIAN LETTER QA VARIANT-M3 FOURTH FORM
```

```
182D 180B   MONGOLIAN LETTER GA VARIANT-M1 SECOND FORM OR FEMININE FORM
                 ·   second form when initial or medial
                 ·   feminine form when final
182D 180C   MONGOLIAN LETTER GA VARIANT-M2 THIRD FORM
182D 180D   MONGOLIAN LETTER GA VARIANT-M3 FEMININE FORM
1830 180B   MONGOLIAN LETTER SA VARIANT-M1 SECOND FORM
1830 180C   MONGOLIAN LETTER SA VARIANT-M2 THIRD FORM
1832 180B   MONGOLIAN LETTER TA VARIANT-M1 SECOND FORM
1833 180B   MONGOLIAN LETTER DA VARIANT-M1 SECOND FORM
1835 180B   MONGOLIAN LETTER JA VARIANT-M1 SECOND FORM
1836 180B   MONGOLIAN LETTER YA VARIANT-M1 SECOND FORM
1836 180C   MONGOLIAN LETTER YA VARIANT-M2 THIRD FORM
1838 180B   MONGOLIAN LETTER WA VARIANT-M1 SECOND FORM
1844 180B   MONGOLIAN LETTER TODO E VARIANT-M1 SECOND FORM
1845 180B   MONGOLIAN LETTER TODO I VARIANT-M1 SECOND FORM
1846 180B   MONGOLIAN LETTER TODO O VARIANT-M1 SECOND FORM
1847 180B   MONGOLIAN LETTER TODO U VARIANT-M1 SECOND FORM
1847 180C   MONGOLIAN LETTER TODO U VARIANT-M2 THIRD FORM
1848 180B   MONGOLIAN LETTER TODO OE VARIANT-M1 SECOND FORM
1849 180B   MONGOLIAN LETTER TODO UE VARIANT-M1 SECOND FORM
184D 180B   MONGOLIAN LETTER TODO QA VARIANT-M1 FEMININE FORM
184E 180B   MONGOLIAN LETTER TODO GA VARIANT-M1 SECOND FORM
185D 180B   MONGOLIAN LETTER SIBE E VARIANT-M1 SECOND FORM
185E 180B   MONGOLIAN LETTER SIBE I VARIANT-M1 SECOND FORM
185E 180C   MONGOLIAN LETTER SIBE I VARIANT-M2 THIRD FORM
1860 180B   MONGOLIAN LETTER SIBE UE VARIANT-M1 SECOND FORM
1863 180B   MONGOLIAN LETTER SIBE KA VARIANT-M1 SECOND FORM
1868 180B   MONGOLIAN LETTER SIBE TA VARIANT-M1 SECOND FORM
1868 180C   MONGOLIAN LETTER SIBE TA VARIANT-M2 THIRD FORM
1869 180B   MONGOLIAN LETTER SIBE DA VARIANT-M1 SECOND FORM
186F 180B   MONGOLIAN LETTER SIBE ZA VARIANT-M1 SECOND FORM
1873 180B   MONGOLIAN LETTER MANCHU I VARIANT-M1 SECOND FORM
1873 180C   MONGOLIAN LETTER MANCHU I VARIANT-M2 THIRD FORM
1873 180D   MONGOLIAN LETTER MANCHU I VARIANT-M3 FOURTH FORM
1874 180B   MONGOLIAN LETTER MANCHU KA VARIANT-M1 SECOND FORM
1874 180B   MONGOLIAN LETTER MANCHU KA VARIANT-M1 FEMININE FIRST FINAL FORM
1874 180C   MONGOLIAN LETTER MANCHU KA VARIANT-M2 FEMININE FIRST MEDIAL FORM
1874 180C   MONGOLIAN LETTER MANCHU KA VARIANT-M2 FEMININE SECOND FINAL FORM
1874 180D   MONGOLIAN LETTER MANCHU KA VARIANT-M3 FEMININE SECOND MEDIAL FORM
1876 180B   MONGOLIAN LETTER MANCHU FA VARIANT-M1 SECOND FORM
1880 180B   MONGOLIAN LETTER ALI GALI ANUSVARA ONE VARIANT-M1 SECOND FORM
1881 180B   MONGOLIAN LETTER ALI GALI VISARGA ONE VARIANT-M1 SECOND FORM
1887 180B   MONGOLIAN LETTER ALI GALI A VARIANT-M1 SECOND FORM
1887 180C   MONGOLIAN LETTER ALI GALI A VARIANT-M2 THIRD FORM
1887 180D   MONGOLIAN LETTER ALI GALI A VARIANT-M3 FOURTH FORM
1888 180B   MONGOLIAN LETTER ALI GALI I VARIANT-M1 SECOND FORM
188A 180B   MONGOLIAN LETTER ALI GALI NGA VARIANT-M1 SECOND FORM
```

## 5. Acknowledgements

## 6. References

*Here, the references are listed which are used in the figures taken from L2/08-034R and submitted by Lorna Priest. This list is copied directly from L2/08-034R.*
*For all other figures, the references are given in the figure legends directly.*

Agarunov, Yakovich Mixhael and Agarunov, Mixhael Yakovich. 1997. *Judeo-Tat – Russian Dictionary*. Moscow: Hebrew University, Moscow.

British and Foreign Bible Society. 1965. *Wofuŋ Jesu Kristo Ere Miti ɗâri Ʒi.* The New Testament in Kâte. London: British and Foreign Bible Society.

___. 1971. Reprinted 2002. *Deʋe Niinɛ* (The New Testament in Loma). Monrovia: The Bible Society in Liberia.

Gurshumov, M. 2004. Миньян (Minyan Journal). No 14.

International Bible Society. 1984. *Nyiɛkpɔɔ Nyu Nunu Fɔŋgɔ* (The New Testament in Nafaanra). Ghana, West Africa.

Lutheran Bible Translators. 1988. *Dùmàì Wòi Dèèléé.* (The New Testament in Belleh (Kuwaa)). Liberia, West Africa. Messengers of Christ, Lutheran Bible Translators, Inc.

Pioneer Bible Translators. 1995. *Yabur Dauda* (Selected Psalms). Republic of Guinea.

___. 1999. *Ɛsita a Fee Kitabuna* (Esther in the Yalunka language). Republic of Guinea.

___. 2004. *Taali e Haala Pular: defⱡere 4* (Pular Fables 4). Republic of Guinea.

___. 2007. *Gɛ́ Ga Undaanɛɛ Ǵona Pagɔ* (2007 calendar in the Loma language). Republic of Guinea.

Trinitarian Bible Society. 1989. *Wodɔwo Dɔu* (The New Testament in the Gio Language of the Republic of Liberia). Trinitarian Bible Society. London, England.

Toumieux, Heesuk and Eric. 2001. *Manuel d'alphabétisation de transition Français - Sérère*. Mbodiène: SIL.

## 6. Examples and Figures

*Note: In the revision which was made within short time to update L2/10-280 for discussion on UTC #126, some figures were dropped. As due to lack of time the figures were not renumbered, the number sequence contains gaps.*

**Fig. 3:**  Example of:
  0067 FE01 LATIN SMALL LETTER G VARIANT-2 LOOP-DESCENDER FORM
*Retrieved 2010-06-02 from:*
*http://en.wikipedia.org/wiki/Obsolete_and_nonstandard_symbols_in_the_International_Phonetic_Alphabet*
This figure shows a table in a sans-serif font, where the in the first column of the second row, a "g" is displayed in the form with loop descender. As such is not available now, a picture of the letter is included instead.



| ɼ | long-leg R | voiced strident apico-alveolar trill (Czech ř) |
| g | looptail g | voiced velar plosive |
| ɤ | ou | close-mid back unrounded vowel or voiced velar fricative |
| я | reversed ʀ or Cyrillic ya | voiced epiglottal trill |

**Fig. 4:** Example of:
0126 FE01 LATIN CAPITAL LETTER H WITH STROKE VARIANT-2 FORM WITH VERTICAL STROKE
*From: Agarunov, 1997, pg 83. – This is Fig. 1 from L2/08-034R.*

83 ҺӘЈ

zɪrɪ подмигивать, делать кому-либо знаки глазами

hyndyr 1. а) высокий; hyndyrɪ mɪrd мужчина высокого роста; рослый мужчина; б) громкий; 2. а) высоко; б) громко; hyndyr birɪ стать высоким; hyndyr varafdɪ повышаться, возвышаться, подниматься

hyndyrɪbuj рослый, высокий, высокого роста

hyndyri вышина, высота, возвышенность; ɪ hyndyri вверх, вверху; навɪрх, наверху; ɪz hyndyri а) с высоты; б) громко

hyndyrirɪpɪjmun высотомер

hyzym дрова, полено, кусок древесины; hyzym vɪnçirɪ рубить дрова

hyzymi древесный; годный на дрова

hyzymvɪnçi лесоруб, дровосек

Һ

hɪb пилюля

hɪcɪ 1. подпорка с раздвоенным концом; 2. развилистый, раздвоенный

hɪcɪjɪzuhun имеющий раздвоенный язык

hɪcɪmɪt 1) медицинские банки, ставящиеся на лоб или затылок с целью отсасывания крови и кровопускания (изготавливаются из рогов домашних животных); 2) кровопускание

hɪcɪtxunɪ (см. тж. zoxut) туалет, уборная

hɪdd 1) границы, предел; 2) мера (степень); ɪz hɪdd vɪdarfdɪ переходить границы (дозволенного), переборщить;

hɪddɪn zijod, ɪz hɪddiş zijod - безмерно

hɪddsyz 1. безграничный, чрезмерный, неограниченный; 2. а) безгранично, беспредельно; б) без меры; множество

hɪddsyz-hisobsyz несметный, бесчисленный, бессчетный

hɪjf 1. жаль, жалко; досадно; hɪjf kɪşirɪ жалеть, сожалеть; 2. месть; hɪjfʃurɪ vɪdɪşɪndɪ мстить, отомстить

hɪjf-hɪjf, hɪjfki межд. жаль! увы!

hɪjfi жалко

hɪjkɪl (hɪkɪl) статуя, изваяние

hɪjkɪlsox, hɪjkɪlşunas ваятель, скульптор

hɪjo стыд, стыдливость; hɪjo kɪşirɪ чувствовать, испытывать сильный стыд; стыдиться, стесняться; hɪjo dorɪ устыдить; выговаривать

hɪjoly стыдливый, скромный, застенчивый

hɪjosyz бесстыдный, беззастенчивый, наглый, нахальный

hɪjosyzi бесстыдство, наглость, нахальство

hɪjot двор

hɪjru очарованный, восхищенный; hɪjru soxdɪ очаровывать, восхищать, пленять; hɪjru birɪ очаровываться, восхищаться, пленяться; mɪ ɪrity hɪjru gɪrdom! (обращение к дорогому человеку) да буду я восхищен тобой!

hɪjvon 1) животное, скотина; vɪhşɪ hɪjvon зверь; дикий зверь; 2) перен. жалкий, несчастный; 3) ругат. скотина, болван; hɪjvon birɪ ругат. озвереть, становиться скотиной

hɪjvondor животновод

hɪjvondori животноводство

hɪjvoni 1. скотство; 2. скотский; звериный; 2) перен. грубость, хам-

**Fig. 5:** Example of:
014A FE00 LATIN CAPITAL LETTER ENG VARIANT-1 CAPITAL-N WITH HOOK BELOW FORM
*From: LBT, 1988. – This is Fig. 2 from L2/08-034R.*

¹⁴Èé Jisè wó wolodí va wɔ bè Kalalii kagbóè nù Ŋgàla Wɔi Nyináà wɔ de kewúì jàdè, wɔ fŭìn gbàvàìdí dè kagbo dɔ̀ɔ̀ ei fandɛ jĕì gbĭn nù. ¹⁵Wɔ̀ bɛdí wɔ dè wòlíî sɛlɛ siya Jusíò Ŋgàla sɔ̀aì bóî nù ŋwɔnɔ̀ wo gbĭn wò sadí wɔ̀ì kaaje tègbèlè gbàà.

**Fig. 6:** Example of:
014A FE01 LATIN CAPITAL LETTER ENG VARIANT-2 ENLARGED SMALL-LETTER FORM
*From: BFBS, 1965, pg 350. – This is Fig. 3 from L2/08-034R.*

32 Paulo e hâmockonec fafahare irec dâŋ mume gâcnezi mu surec embiŋ, â gâcnezi mumbiŋ, «Dâŋ i damɛŋ mocte

33 wâc mutec mananʒepeneŋ.» ı Iŋuc epie Paulozi behec-

34 -joparâ rawec. Ŋ̌ic gâcnezi ine Paulorao qohotahuc Jesure biŋe embiŋ. Jaŋeraonec ʒaʒ sâko moc ʒâŋene Dionisio, erâ ŋokac moc ʒâŋene Damari, erâ gâcne.

**Fig. 7:** Examples of:
014A FE02 LATIN CAPITAL LETTER ENG VARIANT-3 ENLARGED SMALL-N WITH INWARDS HOOK FORM
0186 FE01 LATIN CAPITAL LETTER OPEN-O VARIANT-2 TOP-SERIF FORM
0254 FE01 LATIN SMALL LETTER OPEN-O VARIANT-2 TOP-SERIF FORM
*From: Agarunov, 1997, pg 83. – This is Fig. 4 from L2/08-034R.*

### Waa jaa yaka tɛɛ mbe oyɛ kpɔɔ

²⁷Kena ŋgi ni oo tɛɛ mbe oyɛ kpɔɔ na hin? Yaka n wa Ŋgi na, oo jaa na oyɛ kpɔɔri mbra yuu na o li mbra yirɛ na ree? Ɔɔnhɔn! Mbra na ma Nyiɛkpɔɔ yila mu o fraka la, ŋga o Yesu Kristo shɔ mu li kre na yi. ²⁸Ŋga o chin yi na, shɔnlie na Nyiɛkpɔɔ lewo le na u fraka, ki n pe ŋga we taa Mbose mbra na yi. ²⁹Ndee e jawa na Nyiɛkpɔɔ wre, Judae chiin wo wi ra? Mblɛ pe n pe Judae pre Nyiɛkpɔɔ ma na ree? Pe Nyiɛkpɔɔ wi. U pe mblɛ pe n pe Judae wo na. ³⁰Nyiɛkpɔɔ, uu nunu chom, shɔnlie kolo na we Judae le na pe fraka, a shɔnlie kolo na we mblɛ pe n pe Judae pre le na pe fraka na. ³¹Kena o yire shɔnlie na mu Mbose mbra wa ree? Ɔɔnhɔn! La, mbra yiree wa yi yire haa.

**Fig. 10:** Example of:
0181 FE01 LATIN CAPITAL LETTER B WITH HOOK VARIANT-2 TOPBAR FORM
028B FE00 LATIN SMALL LETTER V WITH HOOK VARIANT-1 ROUND FORM
*From: British and Foreign Bible Society, 1971/2002, pg 255. – This is Fig. 5 from L2/08-034R.*

seelea e vee pote kpaabagini ta te jeany v g
14 Tei Ziize na keelaitiɛ te ɣɛ Yɛlusulɛɛ te mɛnini ga Sameelia
yui yee zeini Gala goi wu, te Pite leveni ta Jɔ̃ te wɔ vɛ.
15 Niitiɛ te vaani te Gala feli te ʋaa ma ga te Gala baa ŋinɛgi
yee maa zɔlɔ wo. 16 Ɓɔvɛlaale ɛ́ lɛ ŋɛni dɛ vaani te ta nɛpɛ
pɔ. Kɛ̀ te bataize ɣɛɛni dɛ nɔ ga tiɛ de maliɣii Ziizɛ na daa-zeigi
zu. 17 Pite ɓɔga Jɔ̃ te yee vɛɛ sa te wũmɔ te Gala baa ŋinɛgi yee
maa zɔlɔ wo.
18 Tei Saimɔ wɛlɛni na va ga Gala baa ŋinɛgi yee maa ɔ́ zɔlɔ
wo Ziizɛ na keelaitiɛ ta yee-vɛɛ-nu-wũmɔi zaala zu, é wuloni ga
naʋolo 19 é ɣɛ te ma, "Nà ɓalaa a gaa-ɓai niɛ́ ʋẹ zèea, kɛɛ ʋaa

---

**Fig. 11:** Example of:
0181 FE01 LATIN CAPITAL LETTER B WITH HOOK VARIANT-2 TOPBAR FORM
A78D FE00 LATIN CAPITAL LETTER TURNED H VARIANT-1 ANGULAR FORM *(first letter in second row)*
*From: Trinitarian Bible Society, 1989, p.380. – This is Fig. 6 from L2/08-034R.*

Ðlanɛ ʋoo, a ʋa ta ka pʋ, wo pʋʋ,
Yi Dɛn Aaɓì, ɥ́ gɥ́ lɔ́ gbée pɔ saalɛɛ l
Ⱨ kəwɔ̌ wò gblɥ́nzɔ̀ lípáɥzɔ̀ ká.
Ɓì, ɥ zɔ̀ɔpúú-mènùa Kín káɛ, ɥ wɔkɔ̀
4 Oo yi Dɛn, dɵ e kâ lɔ́ súó ɥ lɔ̀ɛ?
Dɵ e kâ lɔ́ ɥ tɔ́ ye nɛɛ è kpàozɔ̀ɛ?

---

**Fig. 12:** Example of:
0192 FE01 LATIN SMALL LETTER F WITH HOOK VARIANT-2 LATIN LETTER FORM
0254 FE00 LATIN SMALL LETTER OPEN-O VARIANT-1 STANDARD FORM
*From: http://www.unicode.org/udhr/d/udhr_ewe.pdf (UDHR in Éwé), p.2.*

Esi eme kɔ ƒāa be gomekpɔkpɔ kple Ablɔɖemenɔnɔ gɔmesese tsitotsito nye nu vevitɔ wu, hena takpɔkpɔ
na adzɔgbeɖeɖe sia bliboe ta la,
XEXEME DUKɔ ƑOƑUAWO ƑE TAKPEKPEGÃ LE GBEƑÃ ÐEM:
Amegbetɔwo ƒe Ablɔɖevinyenye ŋu Kpeɖoɖzinya sia be, enye ɖoɖo si dzi mieda asi ɖo na amewo katã le

---

**Fig. 14:** Example of:
019D FE01 LATIN CAPITAL LETTER N WITH LEFT HOOK VARIANT-2 ENLARGED SMALL-LETTER FORM
*From: PBT, 1999, p.5. – This is Fig. 7 from L2/08-034R.*

wayaxan kui. Biiin yi a hhhha aio a waxi a xhh hha hh
naxanyi. Mangan yi a walikɛne yamari fa fala e xa yamaan
nafan fene raba tun.
9 Mangana a ɲaxanla, Baatɛ Fasita fan bata yi ɲaxanle
malan Manga Asuyerusu a banxini dɛɛge ɲaxaɲaxani.
10 Ɲaxa ɲaxana xii soloferede, Manga Asuyerusu to

**Fig. 15:** Example of:
019D FE01 LATIN CAPITAL LETTER N WITH LEFT HOOK VARIANT-2 ENLARGED SMALL-LETTER FORM
*From: PBT, 1999, p.5. – This is Fig. 7 from L2/08-034R.*

**Bono bonaa, ko Jikku mayru on boni**



Ɲande goo Bono fokkiti dennabo ka
esiraaɓe mun. Ndu wiʼi yo jikku mayru
on habbo ndu ka mayru, haa ndu arta,
hita o hersinoy ndu. Jikku on jaabii,

22

**Fig. 16:** Example of:
019D FE01 LATIN CAPITAL LETTER N WITH LEFT HOOK VARIANT-2 ENLARGED SMALL-LETTER FORM
*From: PBT, 1995, p.1. – This is Fig. 9 from L2/08-034R.*

naxan mu yoma Ala ma.
²Ɲɛlɛxinyi na Adama nan bɛ,
Alatala xa sɛriyɛ rafan naxan ma,
naxan a xaxili dɔxɔma Alatala xa sɛriyɛ xɔn ma kɔɛ nun
yanyi.

**Fig. 17:**  Examples of:

019F FE01 LATIN CAPITAL LETTER O WITH MIDDLE TILDE VARIANT-2 FORM WITH TILDE INSIDE

0275 FE01 LATIN SMALL LETTER BARRED O VARIANT-2 FORM WITH TILDE INSIDE

*The left figure shows a page from a primer (Temirxanov M., Alifba. Mahac-Qala, 1935; p.67) of the Kumyk language spoken in Dagestan (Russian Federation), which had a Latin orthography (similar to the Jaŋalif alphabet) for some years about 1930.*

*The right figure shows an enlarged detail of the left one.*

**Fig. 18:** Example of:
01A5 FE01 LATIN SMALL LETTER P WITH HOOK VARIANT-2 FORM WITH RIGHT HOOK ABOVE
*From: Toumieux, 2001. – This is Fig. 10 from L2/08-034R.*

## LEÇON 1

### Présentation générale de l'alphabet sérère

| | | | | | |
|---|---|---|---|---|---|
| A a | B b | Ɓ ɓ | C c | Ć ć | D d |
| Ɗ ɗ | E e | F f | G g | H h | I i |
| J j | K k | L l | M m | Mb mb | N n |
| Nd nd | Ng ng | Nj nj | Nq nq | Ñ ñ | Ŋ ŋ |
| O o | P p | P p̉ | Q q | R r | S s |
| T t | Tf | U u | W w | X x | Y y |
| Y y | ' | | | | |

---

**Fig. 19:** Example of:
01A5 FE01 LATIN SMALL LETTER P WITH HOOK VARIANT-2 FORM WITH RIGHT HOOK ABOVE
*From: Toumieux, 2001. – This is Fig. 11 from L2/08-034R.*

### Exercice 7

Lisez le texte suivant et répondez aux questions :

O tuuduuf sutooxu a duufik. Ye ta sulaa, a seek a leng no ax ke a sama
kam a fat ale. A seek a lakas a yena maa mayeerna mbuuy. A looya. Yaa
njeec naa a sumna, a leep'a, yaam jegee a pay. A seek a lakas a sama
maa jegna kic. A seek a palakaand ale a sama no lanq paax ke.

1) Na naxax alene, a seek ax a podnum njegu ?
2) Xar taxu a seek a fikandeer ale a leep' ?
3) Xar a naxax alene tektu ?

**Fig. 20:** Example of:
01B2 FE01 LATIN CAPITAL LETTER V WITH HOOK VARIANT-2 ANGULAR FORM WITH HIGH HOOK
028B FE01 LATIN SMALL LETTER V WITH HOOK VARIANT-2 ANGULAR FORM WITH HIGH HOOK
*From: PBT, 2007, December. – This is Fig. 12 from L2/08-034R.*

**Fig. 21:** Example of:
01B7 FE01 LATIN CAPITAL LETTER EZH VARIANT-2 CAP-HEIGHT FORM
*From: Girdenis, Alexas: Kalbotyros darbai, (Studies in Linguistics), vol 3: 1988-2000; Vilnius 2001, p.92.*

**4.0.3.** Ir arsyje, ir tezyje, sekant Antikos (ypač Homero, plg. 1.1.2) tradicija, skiemenys kartais dirbtinai ilginami rašybinėmis geminatomis:

Ʒ̂ i n n ô m  jùk, ką Iáuʒei mùms rėplódămĭ riekia,
Kuinai tùr, kàd liepjam' jiems, ʃkraidydămĭ t r u p p į t' [52–53] P416–417

Ʒ̂ i n n ô t'  jùk, kaip Grybas, kàd jį mókĭ pắprówit',
W i r r ă l ŭ s  ir Ăwĭʒu Krŭpàs užgardĭnă ʃʒauney [54–55] W610–611

**Fig. 23:** Example of:
040B FE01 CYRILLIC CAPITAL LETTER TSHE VARIANT-2 FORM WITH STROKE THROUGH ASCENDER
*From: Gurshumov, 2004, p.32. – This is Fig. 13 from L2/08-034R.*

Ћофт сал, ә һәсрәт «Бәбә».
Гузәт бу Кәләбәбә.
Ширинә гофә «Ә-Бә»
Ʒаур, әз ки мишинним?

**Fig. 24:** Examples of:
0431 FE01 CYRILLIC SMALL LETTER BE VARIANT-2 SERBIAN FORM
0433 FE01 CYRILLIC SMALL LETTER GHE VARIANT-2 SERBIAN FORM
0434 FE01 CYRILLIC SMALL LETTER DE VARIANT-2 SERBIAN FORM
043F FE01 CYRILLIC SMALL LETTER PE VARIANT-2 SERBIAN FORM
0442 FE01 CYRILLIC SMALL LETTER TE VARIANT-2 SERBIAN FORM
*From: http://en.wikipedia.org/wiki/File:Special_Cyrillics.png , retrieved 2010-07-29.*
*Accompanying text (2007-10-20):* This image was made by H92.
Description: An overview of special Cyrillic letters used in Macedonian and Serbian, compared with the same letters in standard Cyrillic. The Cyrillic characters are written with Adobe Minion Pro, and Open Type font.

Standard
Regular б г д п т
Italic  б г д n m
Serb./Mac.
Regular δ г д п т
Italic  δ ī g ū ū̄

**Fig. 25:** Examples of:
0431 FE01 CYRILLIC SMALL LETTER BE VARIANT-2 SERBIAN FORM
0433 FE01 CYRILLIC SMALL LETTER GHE VARIANT-2 SERBIAN FORM
0434 FE01 CYRILLIC SMALL LETTER DE VARIANT-2 SERBIAN FORM
043F FE01 CYRILLIC SMALL LETTER PE VARIANT-2 SERBIAN FORM
0442 FE01 CYRILLIC SMALL LETTER TE VARIANT-2 SERBIAN FORM
0448 FE01 CYRILLIC SMALL LETTER SHA VARIANT-2 SERBIAN ALTERNATIVE FORM

*From: http://typography.com/fonts/font_features.php?featureID=14&productLineID=100026*
*retrieved 2010-07-29; Page title: Hoefler & Frere Jones; Whitney features*
 *("Whitney" being a font).*
*The following pictures show excerpts of that page.*

Consulting with H&FJ on the project were two Cyrillists: Maxim Zhukov, former
Typographic Coordinator to the United Nations, and Ilya Ruderman, creator of
the Type & Typography program at the British Higher School of Art and Design
in Moscow. Each brought a culturally enlightened perspective to the critique, and
helped us to evaluate our work in the types of linguistic settings in which
designers would be most likely to use the fonts.

*бδ гī дg nū mū шщ*

SERBIAN/MACEDONIAN LOCALIZATIONS

**Cyrillic Localization**

Three of the most widely-spoken languages that use the Cyrillic alphabet
exhibit regional differences in the shapes of their letters. Not merely stylistic
variations, these are significant morphological alterations without which their
host languages can look foreign.

**Fig. 26:** Examples of:

A78D FE01: LATIN CAPITAL LETTER TURNED H VARIANT-2 CHE FORM
*This figure shows two specimens of Latin alphabets used in Slovenia in the 19th century: the left one is the Dajnko alphabet, the right one the Metelko alphabet. Both alphabets show an adaptation of the Cyrillic che (Чч). While the appearance of the uppercase character (encircled in red) is retained, the lowercase one (enclosed in green) gets the appearance of a turned h in both cases. Therefore, the unification with U+0256 LATIN SMALL LETTER TURNED H is obvious.*
*The other letters of these alphabets which are not encoded yet will be subject of a later proposal, which will discuss these letters in detail.*

**ISO/IEC JTC 1/SC 2/WG 2**

<table>
<tr><td colspan="2">

**PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS**
**FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646**[1]
**Please fill all the sections A, B and C below.**
**Please read Principles and Procedures Document (P & P) from** http://www.dkuug.dk/JTC1/SC2/WG2/docs/principles.html **for guidelines and details before filling this form.**
**Please ensure you are using the latest Form from** http://www.dkuug.dk/JTC1/SC2/WG2/docs/summaryform.html .
**See also** http://www.dkuug.dk/JTC1/SC2/WG2/docs/roadmaps.html **for latest** *Roadmaps.*

</td></tr>
</table>

**A. Administrative**

| | |
|---|---|
| 1. **Title:** | *Proposal to add Variation Sequences for Latin and Cyrillic letters* |
| 2. Requester's name: | *Karl Pentzlin* |
| 3. Requester type (Member body/Liaison/Individual contribution): | *Expert Contribution* |
| 4. Submission date: | *2011-02-07* |
| 5. Requester's reference (if applicable): | |
| 6. Choose one of the following: | |
|     This is a complete proposal: | *Yes* |
|     (or) More information will be provided later: | |

**B. Technical – General**

1. Choose one of the following:
    a. This proposal is for a new script (set of characters):           *No*
        Proposed name of script:
    b. The proposal is for addition of character(s) to an existing block:    *No*
        Name of the existing block:
2. Number of characters in proposal:    *No new characters; 43 Variation Sequences for 21 existing characters*
3. Proposed category (select one from below - see section 2.2 of P&P document):

| A-Contemporary | **X** | B.1-Specialized (small collection) | | B.2-Specialized (large collection) | |
|---|---|---|---|---|---|
| C-Major extinct | | D-Attested extinct | | E-Minor extinct | |
| F-Archaic Hieroglyphic or Ideographic | | | G-Obscure or questionable usage symbols | | |

4. Is a repertoire including character names provided?          *Yes*
    a. If YES, are the names in accordance with the "character naming guidelines"
        in Annex L of P&P document?          *(Yes)*
    b. Are the character shapes attached in a legible form suitable for review?    *Yes*
5. Fonts related:
    a. Who will provide the appropriate computerized font to the Project Editor of 10646 for publishing the standard?
                                      *TBD*
    b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.):

6. References:
    a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?    *Yes*
    b. Are published examples of use (such as samples from newspapers, magazines, or other sources)
    of proposed characters attached?          *Yes*
7. Special encoding issues:
    Does the proposal address other aspects of character data processing (if applicable) such as input,
    presentation, sorting, searching, indexing, transliteration etc. (if yes please enclose information)?   *Yes*
                     *Use of Variation Sequences*

8. Additional Information:
Submitters are invited to provide any additional information about Properties of the proposed Character(s) or Script that will assist in correct understanding of and correct linguistic processing of the proposed character(s) or script. Examples of such properties are: Casing information, Numeric information, Currency information, Display behaviour information such as line breaks, widths etc., Combining behaviour, Spacing behaviour, Directional behaviour, Default Collation behaviour, relevance in Mark Up contexts, Compatibility equivalence and other Unicode normalization related information. See the Unicode standard at http://www.unicode.org for such information on other scripts. Also see http://www.unicode.org/Public/UNIDATA/UCD.html and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.

---

[1] Form number: N3702-F (Original 1994-10-14; Revised 1995-01, 1995-04, 1996-04, 1996-08, 1999-03, 2001-05, 2001-09, 2003-11, 2005-01, 2005-09, 2005-10, 2007-03, 2008-05, 2009-11)

**C. Technical - Justification**

1. Has this proposal for addition of character(s) been submitted before? — *No*
   If YES explain

2. Has contact been made to members of the user community (for example: National Body, user groups of the script or characters, other experts, etc.)? — *No*
   If YES, with whom?
   If YES, available relevant documents:

3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included? — *Yes*
   Reference: — *see text*

4. The context of use for the proposed characters (type of use; common or rare) — *Common*
   Reference: — *see text*

5. Are the proposed characters in current use by the user community? — *Yes*
   If YES, where?  Reference: — *see text*

6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP? — *n/a*
   If YES, is a rationale provided?
   If YES, reference:

7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)? — *n/a*

8. Can any of the proposed characters be considered a presentation form of an existing character or character sequence? — *Yes*
   If YES, is a rationale for its inclusion provided? — *Yes*
   If YES, reference: — *This is what Variation Sequences are for*

9. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters? — *Yes*
   If YES, is a rationale for its inclusion provided? — *Yes*
   If YES, reference: — *Only Variation Sequences consisting of existing characters are proposed*

10. Can any of the proposed character(s) be considered to be similar (in appearance or function) to an existing character? — *Yes*
    If YES, is a rationale for its inclusion provided? — *Yes*
    If YES, reference: — *Variation Sequences are similar to their base characters by definition*

11. Does the proposal include use of combining characters and/or use of composite sequences? — *No*
    If YES, is a rationale for such use provided?
    If YES, reference:
    Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?
    If YES, reference:

12. Does the proposal contain characters with any special properties such as control function or similar semantics? — *No*
    If YES, describe in detail (include attachment if necessary)

13. Does the proposal contain any Ideographic compatibility character(s)? — *No*
    If YES, is the equivalent corresponding unified ideographic character(s) identified?
    If YES, reference: