

HUNGARIAN NATIVE WRITING DRAFT PROPOSAL

BAKONYI Gábor. (Hungary, Budapest, Csillaghegy)

September 30, 2008

1 The Origin of the writing

The hungarian native writing no doubt has the same origin as the old turkish writing, which was for example found at the river Orhun and at the river Yenisey, and the writing used by the Huns. (In the hungarian language the word "betű" means letter near same as the old-turkish word "bitig". **𐰃𐰆𐰇𐰈**.) After the Huns, the Avar Empire ruled the Carpathian basin. At the end of the VIII. century the Avar Empire collapsed. Little Avar groups stayed in Transylvania, but several groups went to the east. In 896 the Magyars returned to the Carpathian Basin, and the chronicles say, that the groups which had stayed, were using this kind of writing. Two chronicles say the opposite about the origin of the writing. The "Illustrated chronicle" says, that the Székely stayers had learnt this writing from the bulgar-turkish "blak" groups during the IX. century, despite of the fact, that any example of the Bulgar-Turkish writing has not been found. (Today we can not read the remaining short texts written by the Avars.) In conflict with the "Illustrated chronicle", the "Thuróczy chronicle" says, that the Székelys were still keeping the old, Hun style writing. The returned magyars in all probability used an old-turkish style writing system too, but the count of the remaining texts are very low.

Similarities of Old Turkish writing and Hungarian Native Writing

(Bakonyi Gábor.)








(The Székelys was appointed for border-guarding of Hungary. so they were a little bit separated from the other parts of the country. The Székely dialect of the Hungarian language is used in eastern Transylvania.)

bellow. All the lines are goes from the right to the left. We have an example for an old text written on a four sided wood lath, but it has only remained as a paper copy, so we don't know exactly, what the writing direction was in general on a wood stick. During the latest centuries, the paper become a usual surface for writing. In general, the writing direction is the same as on wood tables: from the right to the left. I don't know any boustrophedon example written on paper, stone or wooden board. Only the Marsigli stick is supposed to have been written this style, but we have only a copy about it on paper, and on the paper, there is no trace of the original positions. I know only three examble from the early ages, which was written from the left to the right. Two of them used latin language with native letters, and both of them was cryptographic using: short texts mixed among latin letters, without mirrored letters. One of them was the letter of Szamosközi István from 1604. In this letter, the title of the poem was written from tr rightt to the left, but the following lines was written from the left to the right. The other unmirrored left to the right text was written by Zakariás János in 1756, using the latin language, the text was mixed into a latin text written by latin characters. The only one example which was witten with mirrored letters, is the text from Constantinople, written by the mission in Istanbul in 1515, using hungarian language. (Only a paper copy exist from the text originally written to a stone.) Leonardo Da Vici wrote mirrored letters from the right to the left, but this cryptographic tradition is not part of the Unicode.

4 Capital Letters and Small Letters

Despite of the mentioned letter of Szamosközi István from 1604, in which in the tittle of the poem, greater size letters was used at the starts of the words, in the Hungarian Native Writing, there are no distinct lower-case and upper-case letters, contrast with the latin, cyrillic and greek writings.

5 The Letters of the Latent Alphabetical Writing



The hungarian writing has several written abc. The most important of them is the so called Nicolsburgian abc from 1483. The abc was written down several times, and ligatures was enclosed to most of to these abc's. In the hungarian writing, it is difficult to find the base letters, because the ligatures are present in a very great number, and this number continuously increasing. In some case, it is impossible to determine the original components of the ligatures. For example the "eMP" , "uNK" , "TPRUS" , "uS"  signs are this kind of signs. The consonant "K" has two letter sign: "aK"  and "eK" . It is not exactly clear, why two different signs exists. Today, none of the hungarian dialects can distinguish two kind of "K". It is supposed, that the "aK"  was used in the case of deep concomitant vowels, or it was used at the end of the words. Its use has not been clarified yet.

The traditional style prefered skipping some vowels, especially the 'E'. The traditional style also prefered the ligatures. The alphabetical abc:

SIGN	DEFFINITION
	LETTER A

SIGN	DEFINITION
X	LETTER B
↑	LETTER C
H	LETTER Cs
+	LETTER D
1	LETTER J
3	LETTER E
⊗	LETTER F
^	LETTER G
≠	LETTER Gy
X	LETTER Ty
8	LETTER H
†	LETTER I
1	LETTER aK
^	LETTER L
⊙	LETTER Ly
4	LETTER M
)	LETTER N
D	LETTER Ny
∩	LETTER O
≡	LETTER P
◇	LETTER K
H	LETTER R
Y	LETTER Zs
^	LETTER S
Y	LETTER T
≡	LETTER U
M	LETTER V
h	LETTER Ü
8	LETTER Ö
1	LETTER Sz
H	LETTER Z

6 Ligatures and "Bug"-Characters

There is not a sharp border between the ligatures and the "bug"-characters. In the Hungarian Native Writing system in the past the ligatures was preferred. The ligatures are occasional, optional, and sometimes ambiguous. The "bug"-characters are optional too, maybe in the far past they were ligatures or the remaining characters of an old Turkish syllabism. We don't know the exact components of these signs, but sometimes we can have slight ideas. The "bug"-characters are optional parts of the abc's. From some old abc sets they are missing, in others, they are present. The 'MB'  and 'NT'  are the most frequently used bug-characters. The meaning and using of some "bug"-character is not

clear: some of them have not found in text examples, only in abc's. The "bug"-characters as it is in the Nicolsburgian-abc:

SIGN	DEFFINITION
↕	Letter eNC
✖	Letter aND
✘	Letter aTyaTy
✙	Letter uNK
⊗	Letter MB
⋈	Letter eNT
⊖	Letter uS
⦿	Letter TPRUS

Some sample for ligatures:

SIGN	DEFFINITION
↗	Ligature CK
↘	Ligature CK
↙	Ligature NC
↘↗	Ligature NCs
↘↙	Ligature NK
↘↖	Ligature ND
↘↗↖	Ligature NGy
↘↗↙	Ligature NyJ
↘↗↙↖	Ligature ZT
↘↗↙↖↗	Ligature ST
↘↗↙↖↗↙	Ligature OE=Ö

7 The main problem: Long and short vowels, long and short consonants

In the traditional use, the long vowels and long consonants were not designated, in the writing there were not any difference between the long and short vowels. It is the main deficiency of the hungarian writing. (In the turkish language long vowels not exists.) In the hungarian language the long "É" is not exactly the long pair of the short 'E', but it is a close "E" vowel, so a variant of the letter "É" was created to solve the problem. (In paralel with it, the long 'Á' is not exactly the long pair of the short 'A', but in some dialects the difference is not important at all.) The letter "É" ↗ was the first try to sign the long vowels. For example in the abc of Harsányi from 1678-1680 this two variant of the original "E" letter had been separated. For example the text found in the book of the gyulafehérvárian school, from 1655, had made distinction between the 'E' and 'É' letters. The letter 'É' is old enough to accept it as a traditional one.

SIGN	DEFFINITION
Œ	Letter É

8 Numeral signs

We have some numeral signs too: "one", "five", "ten", "fifty", "hundred", and "thousand".

SIGN	DEFFINITION
I	Number 1
V	Number 5
X	Number 10
∨	Number 50
* (with a dot)	Number 100
* (with a dot)	Number 1000

9 Other Signs: Puncuatation Marks

Due to the crowded style, sometimes the words are separated by one point, or two points, or four points arranged along a vertical line, but sometimes a blank space is used as a separator. In the text from Csikszentmárton village, from 1501, four dots are the separators between the words. In the letter of Szamosközi István from 1604 one dot separates the words, etc. We can found additional signs in the so called "Rudimenta" from 1598: A single point separates the sentences, a double point separates the subordinate sentences, in the same role, as the comma in the latin abc. In the Rudimenta, two near vertical paralel lines signs the word break. It means, we use it, when we want to continue the word in the next line. In the text sample of Bél Mátyás, from 1718, also a single point closes the sentences, and two paralel lines signs the word break. In the register of Furta village from 1786 a slanted line from the direction of the right top to direction of left bottom, means the same, as the colon in the latin abc. It means: something important follows it.

Unfortunetly the question mark is allways missing from the old texts. (In the turkish language the question mark is needn't.)

SIGN	DEFFINITION
·	Singele dot separator
∴	Double dot separator
⋮	Qvad dot separator
=	Break
/	Colon

SIGN	DEFINITION
˘	Long-Mark

As an other solution, using the letter 'É' as a model, some researcher try to create new letters from the glyph variants of the old ones. The great number of the newly created characters caused a total chaos. Furthermore different reformer groups used different variants of the letters.

Unfortunately due to an orthographical mistake, on the Marsigli-stick, the letter Ʒ was used as 'Ö' sound and the letter Ɔ was used as 'Ü' sound; the two letters was used transposed. (Total 8+2 times.) The researcher Sebestyén Gyula said, that it was only a mistake, but this mistake was a good chance for the reformers, to increase a count of the characters, so the reformer Forrai Sándor said, that they are independent new letters. He also said, that the same transposed letters are used in the Nicorsburgian abc, despite of its latin transcription is ambiguous. (The transposed letters and was not uncommon in the early latin transcriptions, for example in the Telegdi-Kájoni abc, the signs of the 'cz' and 'cs' were transposed too.) Based on this mistake, the long and short 'Ö' and the long and short 'Ü' letters created newly by the reformer groups are the mix of the glyph variants of the original 'Ö' and 'Ü' letters and the 'OE' ligatures.

We have no traditional ABC or text where two different type of letter 'Ö' or 'Ü' was present, therefore we either have to say, that that is an orthographical error in the Marsigli text, or we have to say, that the Marsigli 'Ö' and 'Ü' letters are very distorted, but the reformers selected the worst solution.

Furthermore the different reformer groups use different mixes.

	a	á	e	é	ë	i	í	o	ó	ö	ő	u	ú	ü	ű
<i>Petrovay János</i> 1903	4	4̄	ʒ	ʒ̄		†	†̄	∩	∩̄	K	K̄	⊗	⊗̄	z	z̄
<i>Magyar Adorján</i>	4	4	ʒ	ʒ		†	∩	∩	∩	⊗	⊗	⊗	⊗	z	z
<i>Szakonyi István</i> 1963	4	4	ʒ	ʒ		†	∩	∩	∩	⊗	⊗	⊗	⊗	z	z
<i>Forrai</i> 1996	4	4	ʒ	ʒ		†	†	∩	∩	K	⊗	⊗	⊗	z	z
<i>Forrai</i> 1994	4	4	ʒ	ʒ		†	†	∩	∩	K	⊗	⊗	⊗	z	z
<i>Libsch</i> 1998	4	4	ʒ	ʒ	ʒ	†	†	∩	∩	z	K	⊗	⊗	z	z
<i>Nyíri Attila</i> <2008	4	4	ʒ	ʒ		†	†	∩	∩	∩	∩	⊗	⊗	z	z
Badinyi	4	4	ʒ	ʒ		†	∩	∩	∩	⊗	⊗	⊗	⊗	z	z
	P	P	ʒ	ʒ		†	∩	∩	∩	⊗	⊗	⊗	⊗	z	z
Kiszely 2001	4	4	ʒ	ʒ		†	∩	∩	∩	⊗	⊗	⊗	⊗	z	z
	4	4	ʒ	ʒ		†	∩	∩	∩	⊗	⊗	⊗	⊗	z	z

Different reformers created a great number of different long vowels. The long and short vowels are created from different styles of the abc, therefore the long and short vowels became style dependent. The claim for the using of different styles never let standstill the solutions, if we use this method. Every person, who try a new style, shall solve again the problem of the long vowels. for example Nyíri Attila's 'U' and 'Ú' letters are the opposit of the letters of Forrai. The new letters, created by the reformers, has a lot of problems. I think, that the best solution is to separate them.

The Hungarian language uses an open, short 'Ě' sound too, together with the open, long 'É' and the closed, short 'E'. This sound form is near to the 'Ö' in the south-dialect, but in the other dialects, it is sounds like the 'E'. Neither in the latin writing system, nor in the native hungarian writing system, it never was distinguished, none of the old character sets contents the letter 'Ě'. Despite of these, the reformers are clinging atherence to create a new letter for this. In the Rudimenta of Telegdi, a glyph variant of the letter 'Ö' X was used, which is a little bit similar to the letter 'H' X. (The Rudimenta is exists obly in copies. The hamburgian copy, and the giessenian copy.) Therefore the closed 'Ě' sounded like 'Ö' and the letter 'Ö' X was similar to the letter 'H' X. A distorted form of the letter 'E' ʒ is also similar to the letter 'H' X. The reformers likes to increase the count of the characters, so a new letter was created from a glyph variant of the letter 'H' for the sound 'Ě'. I think, that it is very stange, that the theories of the reformers are founded on copy errors, or orthographical errors like this.

The letter 'R' has two glyph variant, but they have the same meaning, same sound-form, but the shorter form is for lazy writers.

Some of the reformers try to change the writing directions too, because it was easier to write on a computer from the left to the right. In the early ages the stick was in the left hand, and the knife in the right hand. It would be very difficult to write from the left to the right!

The reformers doesn't bother with traditional punctuation marks, they use the mirrored punctuation marks of the latin character set.

In our age, some creator tries to distinguish small letters and capital letters, but it is against the tradition. In the earlier times, the letters was wrote to wood sticks, so all the letters had to be the same height!

As a result of the reforms, the reformed writings are sometimes illegible. The newly created letters are only glyph variants of the original letters, so the difference is small. It makes impossible the using of the different character styles too.

The reformers doesn't bother with the problem of the long consonants. They want to sign them, like in the latin abc: by writing the letter two times. In the the traditional writing, both in the case of the long consonants, both in the case of the short consonants, the letter is written only one times, like in the case of the vowels. It seems, that the reformers try to reform the traditions using the latin abc as a model.

Some of the reformers is using arabic numeral signs, none of them is using indian numeral signs, and some of the reformers try to create new numeral signs too.

The traditional style prefered the ligatures, but in our modern age, this signs causes only problems. Despite of this, the reformers are clinging atherence to use them.

None of the reformers studied the problem of the long consonants and the missing question mark.

12 Conclusion

The Hungarian native writing had only an informal use during the last eleven centuries. Due to the only informal use and the traditional jam-packed writing style, its has a great number of ligatures and variants, so not only very difficult to find the best solutions, but the base letters too. Its cryptographic use caused additional problems. We have to satisfy the new requirements in this chaos. Looking for new solutions increased this chaos. First of all, we had to the earliest style to create peace among the great number goups using different styles and variants. As a second step had to simplify everything, and on the base of a simplified writing system, must to solve the problems of the long vowels and long consonants, and the ligatures. Difficult to value the count of the users, probably some hundred thousand, but the count of the potential users are some millions.

13 My suggestions

The most inportant is the 32+1 base character. The standard have to contain them.

The sorting order of the characters in the traditional ABC's was allways near in the order of the latin ABC, so I suggest, that the best sorting order would be an order based on the the order of the latin transcription. The long vowels and long consonants have to follow immediatly their short pairs. If we use combining dot or combining macron, as proposed by Michael Everson, we have to solve it too. It is possible, if the combining mark follows

the letter.

The "bug"-characters are might be part of the standard, at least the two most important one: "MB" 𐞀 and "NT" 𐞁 . This two bug-characters are present is some text examples, but the others are not. We could live without the others...

We could solve the problem of the ligatures using the proposal for the Zero-Width-Joiner U+200D from the document n1758.pdf created by Michael Everson in 1998.

Michael Everson says, that the problem of the numeral signs should be solved later.

The punctuation marks are important too. The space-character is either an U+0020 space, or vertical four dots. Michael Everson says, that in this second case we could use the U+205E. (Of course, if the word separator is not the space, in the ttf files, have to set it in the usBreakChar field of the OS/2 table.) Michael Everson says, that as a single point we could use U+2E31 (???), and as a double dot: U+205A (???). For the "break" and "colon", he have not said anything yet.

To implement the Petrovay style long-mark, Michael Everson says, that a combining macron U+0304 would be a good solution. This way, we could solve both the problem of the long vowels, both the problem of the long consonants. In the case of the letter 'É', maybe a combining dot above a letter 'É' 𐞂 would solve the problem: 'Ë' 𐞃 . It would be strange to put a combining macron over a letter 'É' 𐞄 because only this original character is long.

SIGN	DEFFINITION
·	Short-Mark

A question mark is needed! I am suggest to use a question mark as three vertically arranged points, because the modern use requires it.

SIGN	DEFFINITION
⋮	Tripledote separator

All of our problems are already solved, but with respect to the work of the XX. century reformers, I suggest 14 freely useable character code, where the reformers could place their conflicting characters.

I also suggest, that it would be wonderful, if our writing would be on the BMP.

I suggest the following sorting order.

SIGN	NAME	CHARACTER PROPERTIES
𐞀	Native Hungarian Letter A	;Lo; 0;R; ;; ;N;;;;;
𐞁	Native Hungarian Letter eB	;Lo; 0;R; ;; ;N;;;;;
𐞂	Native Hungarian Letter eC	;Lo; 0;R; ;; ;N;;;;;
𐞃	Native Hungarian Letter eCs	;Lo; 0;R; ;; ;N;;;;;
𐞄	Native Hungarian Letter eD	;Lo; 0;R; ;; ;N;;;;;
𐞅	Native Hungarian Letter E	;Lo; 0;R; ;; ;N;;;;;
𐞆	Native Hungarian Letter EI	;Lo; 0;R; ;; ;N;;;;;
𐞇	Native Hungarian Letter eF	;Lo; 0;R; ;; ;N;;;;;
𐞈	Native Hungarian Letter eG	;Lo; 0;R; ;; ;N;;;;;
𐞉	Native Hungarian Letter eGy	;Lo; 0;R; ;; ;N;;;;;

SIGN	NAME	CHARACTER PROPERTIES
☒	Native Hungarian Letter eH	;Lo; 0;R; ;;; ;N;;;;;
†	Native Hungarian Letter I	;Lo; 0;R; ;;; ;N;;;;;
1	Native Hungarian Letter eJ	;Lo; 0;R; ;;; ;N;;;;;
◊	Native Hungarian Letter eK	;Lo; 0;R; ;;; ;N;;;;;
ℓ	Native Hungarian Letter aK	;Lo; 0;R; ;;; ;N;;;;;
♠	Native Hungarian Letter eL	;Lo; 0;R; ;;; ;N;;;;;
⊙	Native Hungarian Letter eLy	;Lo; 0;R; ;;; ;N;;;;;
♣	Native Hungarian Letter eM	;Lo; 0;R; ;;; ;N;;;;;
♠	Native Hungarian Letter eMB	;Lo; 0;R; ;;; ;N;;;;;
)	Native Hungarian Letter eN	;Lo; 0;R; ;;; ;N;;;;;
↕	Native Hungarian Letter eNC	;Lo; 0;R; ;;; ;N;;;;;
✕	Native Hungarian Letter aND	;Lo; 0;R; ;;; ;N;;;;;
☒	Native Hungarian Letter uNK	;Lo; 0;R; ;;; ;N;;;;;
†	Native Hungarian Letter eNT	;Lo; 0;R; ;;; ;N;;;;;
D	Native Hungarian Letter eNy	;Lo; 0;R; ;;; ;N;;;;;
⊙	Native Hungarian Letter O	;Lo; 0;R; ;;; ;N;;;;;
⊗	Native Hungarian Letter OE	;Lo; 0;R; ;;; ;N;;;;;
☒	Native Hungarian Letter eP	;Lo; 0;R; ;;; ;N;;;;;
☒	Native Hungarian Letter TPRUS	;Lo; 0;R; ;;; ;N;;;;;
H	Native Hungarian Letter eR	;Lo; 0;R; ;;; ;N;;;;;
^	Native Hungarian Letter eS	;Lo; 0;R; ;;; ;N;;;;;
⊙	Native Hungarian Letter uS	;Lo; 0;R; ;;; ;N;;;;;
1	Native Hungarian Letter eSz	;Lo; 0;R; ;;; ;N;;;;;
†	Native Hungarian Letter eT	;Lo; 0;R; ;;; ;N;;;;;
✕	Native Hungarian Letter eTy	;Lo; 0;R; ;;; ;N;;;;;
✕	Native Hungarian Letter aTyaTy	;Lo; 0;R; ;;; ;N;;;;;
⊗	Native Hungarian Letter U	;Lo; 0;R; ;;; ;N;;;;;
†	Native Hungarian Letter UE	;Lo; 0;R; ;;; ;N;;;;;
M	Native Hungarian Letter eV	;Lo; 0;R; ;;; ;N;;;;;
☒	Native Hungarian Letter eZ	;Lo; 0;R; ;;; ;N;;;;;
†	Native Hungarian Letter eZs	;Nl; 0;R; ;;; ;N;;;;;
1	Native Hungarian Numeral One	;Nl; 0;R; ;;; 1;N;;;;;
V	Native Hungarian Numeral Five	;Nl; 0;R; ;;; 5;N;;;;;
X	Native Hungarian Numeral Ten	;Nl; 0;R; ;;; 10;N;;;;;
V	Native Hungarian Numeral Fifty	;Nl; 0;R; ;;; 50;N;;;;;
*	Native Hungarian Numeral Hundred	;Nl; 0;R; ;;; 100;N;;;;;
*	Native Hungarian Numeral Thousand	;Nl; 0;R; ;;; 1000;N;;;;;
.	Native Hungarian Mark Singele Dot	;Po; 0;R; ;;; ;N;;;;;
:	Native Hungarian Mark Double Dot	;Po; 0;R; ;;; ;N;;;;;
:	Native Hungarian Question-Mark	;Po; 0;R; ;;; ;N;;;;;
:	Native Hungarian Mark Qvad Dot	;Zs; 0;R; ;;; ;N;;;;;
=	Native Hungarian Mark Break	;Pd; 0;R; <compat> 0020; ;;; ;N;;;;;

SIGN	NAME	CHARACTER PROPERTIES
/	Native Hungarian Mark Colon	;Po; 0;R; ;;; ;N;;;;;
.	Native Hungarian Short-Mark	;Mn; 230;R; ;;; ;N;;;;;
-	Native Hungarian Long-Mark	;Mn; 230;R; ;;; ;N;;;;;

14 Input

The Hungarian Native writing had a DOS coding standard and keyboard, created by Hosszú Gábor László in the years after 1990, but this standard neither doesn't fit to the UNICODE philosophy at all, nor to this proposal. I propose, that the 1 byte decoding of the letters should be in the range 128-191 continuously. The latin standard 7-bit codes should be in the 32-127 range.

KEY	CHAR	UNICODE (HEXA)	KEY	CHAR	UNICODE (HEXA)
193	'Á'	00C1	225	'á'	00E1
194	'Â'	01CD	226	'â'	01CE
199	'Ç'	00C7	231	'ç'	00E7
201	'É'	00C9	233	'é'	00E9
203	'Ë'	00CB	235	'ë'	00EB
205	'Í'	00CD	237	'í'	00ED
206	'Î'	01CF	238	'î'	01D0
208	'Ĝ'	011E	240	'ğ'	011F
211	'Ó'	00D3	243	'ó'	00F3
213	'Õ'	0150	245	'õ'	0151
214	'Ö'	00D6	246	'ö'	00F6
217	'Û'	01D3	249	'û'	01D4
218	'Ú'	00DA	250	'ú'	00FA
219	'Ű'	0170	251	'ű'	0171
220	'Ü'	00DC	252	'ü'	00FC
221	'Ï'	0130	253	'ï'	0131
222	'Ş'	015E	254	'ş'	015F
255	ZWJ	200D			

The Zero-Width-Joiner character had to be present in this 1 byte code too.

In this proposal I wrote only the most necessary requirements. I hope, it will be accepted. I think, that in the future, other, additional proposals will complete this fundament. For example the ligatures are missing from my proposal. The count of the possible ligatures are too great for me. The characters created newly in the recent times by the reformer groups are missing from my proposal too, because the different groups has different solutions, and I didn't want to stir up a hornet's nest.

15 PROPOSAL SUMMARY FORM

A. Administrative

1. Title:

Proposal for encoding the Native Hungarian script in the UCS.

2. Requester's name:

Bakonyi Gábor.

3. Requester type (Member body/Liaison/Individual Contribution)

Individual contribution

4. Submission date:

5. Requester's reference (if applicable)

6. Choose one of the following:

6a. This is a complete proposal

Yes.

6b. More Information will be provided later.

No.

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)?

No.

B. Technical - General

1a. Choose one of the following:

This proposal is for a new script (set of characters)

Yes.

1b. Proposed name of script:

Native Hungarian.

1c. The proposal is for addition of character(s) to an existing block.

No.

1d. Name of the existing block

Number of characters in proposal

55

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

Category B.1

4a. Is a repertoire including character names provided?

Yes.

4b. If YES, are the names in accordance with the "character naming guidelines" in Annex L of P & P document?

Yes.

4c. Are the character shapes attached in a legible form suitable for review?

Yes.

5a. Who will provide the appropriate computerized font (ordered preference: True Type, or PostScript format) for publishing the standard?

Bakonyi Gábor.

5b. If Available now, identify source(s) for the font (include address, e-mail, ftp-site, etc.) and indicate the tools used:

Bakonyi Gábor, Type Designer 3.1 + own tool created by Borland C++ Builder

6a. Are references (to other character sets, dictionaries, descriptive texts, etc.) provided?

Yes.

6b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?

Yes.

7. 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.

8. Submitters are invited to provide any additional information about Properties 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.

C. Technical - Justification

1. Has this proposal for addition of character(s) been submitted before?

If YES, explain

YES, N1758 (1998-05-02) and N1686 (1998-01-20) discussed as " Old Hungarian" previously.

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?

Reference:

Historical and contemporary cultural use by native Hungarians.

4. The context of use for the proposed characters (type of use, common or rare)

Reference:

Rare, but pervasive. Only in informal contacts.

5. Are the proposed characters in current use by the user community?

If YES, where? Reference:

In Transylvania, Hungary, in the all territory of the Carpathian basin, and everywhere else among the native Hungarians, maybe in Canada and in the USA too.

6. After giving due considerations to the principles, in the P&P document must the

proposed characters be entirely in the BMP?

If YES, is a rationale provided?

If YES, reference:

Yes. Due to the contemporary use, the BMP encoding would be comfortable.

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

Yes.

8. Can any of the proposed characters be considered a presentation form of an existing character or character sequence?

If YES, is a rationale for its inclusion provided?

If YES, reference:

No.

9. Can any of the proposed character(s) be encoded using a composed character sequence of either existing characters or other proposed characters?

If YES, is a rationale for its inclusion provided?

If YES, reference:

No.

10. Can any of the proposed character(s) be considered to be similar (in appearance or function) to an existing character?

If YES, is a rationale for its inclusion provided?

If YES, reference:

No.

11. Does the proposal include use of combining characters and/or use of composite sequences ?

Yes.

If YES, is a rationale for its inclusion provided?

If YES, reference:

A very great number of ligatures would be created with the use of the Zero-Width-Joiner, but it will be the theme of an other proposal in the future, because in this proposal I wrote only about the most necessary things.

The Long-mark sign and Short-mark sign above the characters modifies the meaning of it.

Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?

If YES, reference:

No.

12. Does the proposal contain characters with any special properties such as control function or similar semantics?

If YES, describe the detail (include attachment if necessary.)

No.

13. Does the proposal contain any Ideographic compatibility character(s)?

If YES, is the equivalent corresponding unified ideographic character(s) identified?

No.