ISO/IEC JTC1/SC2/WG2 N3357R2

L2/08-071 2008-01-25

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 for encoding the Old Turkic script in the SMP of the UCS Source: China, Ireland, the UK, and UC Berkeley Script Encoding Initiative

(Universal Scripts Project)

Status: National Body and Liaison Contribution

Action: For consideration by JTC1/SC2/WG2 and UTC

Date: 2008-01-25

Introduction. The Old Turkic script has its roots in the Aramaic script, probably via the scripts which developed for Iranian languages. It was used in a wide area stretching across the grasslands of Mongolia, in particular concentrated near the Orkhon River in Mongolia and the Yenisei River in Siberia. Old Turkic script appears to have been put to widespread and popular use for grave stelae, border signs or graffiti of all sorts in hundreds of inscriptions in South Siberia. When, around 700 CE, the Türk founded their empire in Mongolia (which went on to conquer a vast area between China and present day Turkmenistan, all the way down to Afghanistan in the South and the Urals in the North), they needed a national script for their national monuments. The script has a number of names, which can be rendered in English as Old Turkic, Turkic Runiform, Turkic Runes, Orkhon, Yenisei. In Turkish it is sometimes called *Kök Turki* 'Old Turkic'. Resemblances to Germanic Runes are entirely superficial and accidental.

The texts found in the Orkhon River valley in the 19th century are the earliest-known examples of writing in any Turkic language. The Orkhon inscriptions date from the early 8th century CE. The inscriptions on the bank of the Orkhon river were the first to be deciphered because they are bilingual with Chinese. There are no more than 6 or 7 of them. About 50 years later the Uighur empire replaced the Türk empire, and its rulers also put up some stelae, even fewer in number, this time along the (nearby) river Selenge (also in Mongolia). Later inscriptions dating from the later 8th century occur in a variant of the script alphabet, known as Yenisei or Siberian runes, have also been found around Yenisei and other parts of Siberia. Old Turkic is used to write Iranian text in a few manuscripts.

Structure. Old Turkic writing is more-or-less alphabetic, and is inscribed from right to left in rows running from bottom to top. Modern editors use right-to-left directionality but usually set their text in rows running from top to bottom. (Boustrophedon writing is not unknown, though it is rather rare. When the writing changes direction, the letter shapes are mirrored. This happens regardless of the character's base shape; \$\mathcal{1}\$ U+10C00 ORKHON A appears as \$\mathcal{1}\$ when written in boustrophedon, and \$\mathcal{1}\$ U+10C01 YENISEI A appears as \$\mathcal{1}\$.)

Describing Old Turkic as "more-or-less alphabetic" is more-or-less accurate. Consonants have a sort of implicit vowel associated with them; the vowel precedes as in the character names here. It also happens that a vowel *after* a letter is also implicit. But that only happens because it is implicit with reference to the *next* consonant letter. If there is no *next* consonant letter, that is, if the word ends in a vowel, it *always* is explicit. This fact is supported also by characters which appear to be ideograms: 'arrow' is /ok/ in Old Turkic, and the characters \$\dagge\$ U+10C36 ORKHON OQ and \$\dagge\$ U+10C37 YENISEI OQ have arrow shapes. /ay/ is 'moon', and \$\dagge\$ U+10C16 ORKHON AY looks like one; /\dagge\$b/ means 'home, tent' and \$\dagge\$ U+10C0B ORKHON AEB

and ★ U+10C0C YENISEI AEB have a tent shape (evidently very similar to the actual tents used in the area).

All of this affects reading and orthography, not processing, however. Initial and medial vowels are often not explicitly written, though in final syllables they often are; compare LYSS ÇĞBT ← reversed as ATABAĞEÇ otboğaç 'a ferocious bull, a fire-breathing dragon' and LYH IRNT ← reversed as ATANĀRI tengri 'deity'.

Character set and naming convention. The set of characters proposed here includes both Orkhon and Yenisei "variants". Old Turkic orthographies vary and there is overlap between the letters used in different areas, just as there are letters shared between Orkhon and Yenisei. In the character names, ORKHON is applied to letters typically found in the Orkhon area, and YENISEI to letters typically found in the Yenisei area, and those names do not imply more than that. Non-significant glyph variants have been unified. The possibility of treating Orkhon and Yenisei characters as font-dependent (which would reduce the number of characters from 71 to 46) has been explored but rejected, as the distribution of the forms cannot simply be predicted by the provenance of an inscription. The proposed encoding is favoured by experts both in China and in Europe and permits a comprehensive representation of texts of all kinds.

While some consonants are neutral, others are associated with either front or back vowels. In von Gabain's transcription, L¹ is written for back-vowel consonants, L² for front-vowel consonants, and L⁰ for neutral consonants (the example $\lambda Y \delta$ and $\Gamma Y h$ above could be transliterated as $T^1B^1\check{G}^1 \varsigma^0$ and $T^1N^1R^2I$). The naming convention here writes A in front of L¹ consonants, AE (for \ddot{a}) in front of L² consonants, and E in front of neutral consonants. There are also I consonants, and O is used to represent some consonants which are either o or u; two vowel letters have the name OE which similarly represents \ddot{o} or \ddot{u} . All of this is indicative of the strength of the feature of vowel harmony in Turkic languages.

Ordering. The order of the characters in the code charts is indicative of their alphabetical order. However, since the character model for Old Turkic supports a superset of characters used in different geographic areas, about half the characters are sorted together at the first level and only differentiated for tie-breaking. This ensures that a word like $\lambda Y \delta \delta$, which could also be written $\lambda Y \delta \delta$ (using all of the Yenisei letters), can be found in one place in a word-list. Superscript Y in the list below identifies the Yenisei letters.

Punctuation. The character U+205A TWO DOT PUNCTUATION is used for a variety of purposes. The actual appearance of this varies in different manuscripts and editions. Many instances of this use a much tighter form of the punctuation, with the dots closer together, as in Figures 3 and 12. The character U+2E30 RING POINT is sometimes used as a separator just as U+205A TWO DOT PUNCTUATION is, as in Figure 9.

Unicode Character Properties

```
10C00;OLD TURKIC LETTER ORKHON A;Lo;0;R;;;;N;;;;
10C01;OLD TURKIC LETTER YENISEI A;Lo;0;R;;;;N;;;;
10C02;OLD TURKIC LETTER YENISEI AE;Lo;0;R;;;;N;;;;
10C03;OLD TURKIC LETTER ORKHON I;Lo;0;R;;;;N;;;;
10C04;OLD TURKIC LETTER YENISEI I;Lo;0;R;;;;N;;;;
10C05;OLD TURKIC LETTER YENISEI E;Lo;0;R;;;;N;;;;
10C06;OLD TURKIC LETTER ORKHON O;Lo;0;R;;;;N;;;;
10C07;OLD TURKIC LETTER ORKHON OE;Lo;0;R;;;;N;;;;
```

```
10C08;OLD TURKIC LETTER YENISEI OE;Lo;0;R;;;;;N;;;;;
10C09;OLD TURKIC LETTER ORKHON AB;Lo;0;R;;;;;N;;;;;
10C0A;OLD TURKIC LETTER YENISEI AB;Lo;0;R;;;;N;;;;
10C0B; OLD TURKIC LETTER ORKHON AEB; Lo; 0; R;;;;; N;;;;;
10C0C;OLD TURKIC LETTER YENISEI AEB;Lo;0;R;;;;N;;;;;
10C0D; OLD TURKIC LETTER ORKHON AG; Lo; 0; R;;;;; N;;;;;
10C0E; OLD TURKIC LETTER YENISEI AG; Lo; 0; R;;;;; N;;;;;
10C0F; OLD TURKIC LETTER ORKHON AEG; Lo; 0; R;;;;; N;;;;;
10C10;OLD TURKIC LETTER YENISEI AEG;Lo;0;R;;;;;N;;;;;
10C11;OLD TURKIC LETTER ORKHON AD;Lo;0;R;;;;;N;;;;;
10C12; OLD TURKIC LETTER YENISEI AD; Lo; 0; R;;;;; N;;;;;
10C13;OLD TURKIC LETTER ORKHON AED;Lo;0;R;;;;;N;;;;;
10C14;OLD TURKIC LETTER ORKHON EZ;Lo;0;R;;;;;N;;;;;
10C15;OLD TURKIC LETTER YENISEI EZ;Lo;0;R;;;;;N;;;;;
10C16;OLD TURKIC LETTER ORKHON AY;Lo;0;R;;;;;N;;;;;
10C17;OLD TURKIC LETTER YENISEI AY;Lo;0;R;;;;;N;;;;;
10C18;OLD TURKIC LETTER ORKHON AEY;Lo;0;R;;;;;N;;;;;
10C19; OLD TURKIC LETTER YENISEI AEY; Lo; 0; R;;;;; N;;;;;
10C1A;OLD TURKIC LETTER ORKHON AEK;Lo;0;R;;;;;N;;;;;
10C1B;OLD TURKIC LETTER YENISEI AEK;Lo;0;R;;;;N;;;;
10C1C; OLD TURKIC LETTER ORKHON OK; Lo; 0; R;;;;; N;;;;
10C1D;OLD TURKIC LETTER ORKHON AL;Lo;0;R;;;;;N;;;;;
10C1E; OLD TURKIC LETTER YENISEI AL; Lo; 0; R;;;;; N;;;;;
10C1F;OLD TURKIC LETTER ORKHON AEL;Lo;0;R;;;;;N;;;;;
10C20;OLD TURKIC LETTER ORKHON ELT;Lo;0;R;;;;;N;;;;;
10C21;OLD TURKIC LETTER ORKHON EM;Lo;0;R;;;;;N;;;;;
10C22;OLD TURKIC LETTER YENISEI AN;Lo;0;R;;;;;N;;;;;
10C23;OLD TURKIC LETTER ORKHON AEN;Lo;0;R;;;;;N;;;;;
10C24;OLD TURKIC LETTER YENISEI AEN;Lo;0;R;;;;;N;;;;;
10C25; OLD TURKIC LETTER ORKHON ENT; Lo; 0; R;;;;; N;;;;;
10C26;OLD TURKIC LETTER ORKHON ENC;Lo;0;R;;;;;N;;;;;
10C27;OLD TURKIC LETTER YENISEI ENC;Lo;0;R;;;;;N;;;;;
10C28;OLD TURKIC LETTER ORKHON ENY;Lo;0;R;;;;;N;;;;;
10C29;OLD TURKIC LETTER YENISEI ENY;Lo;0;R;;;;;N;;;;;
10C2A;OLD TURKIC LETTER YENISEI ANG;Lo;0;R;;;;;N;;;;;
10C2B;OLD TURKIC LETTER ORKHON ENG;Lo;0;R;;;;;N;;;;;
10C2C;OLD TURKIC LETTER YENISEI ENG;Lo;0;R;;;;;N;;;;;
10C2D; OLD TURKIC LETTER ORKHON EP; Lo; 0; R;;;;; N;;;;
10C2E; OLD TURKIC LETTER ORKHON OC; Lo; 0; R;;;;; N;;;;;
10C2F;OLD TURKIC LETTER ORKHON IC;Lo;0;R;;;;;N;;;;;
10C30;OLD TURKIC LETTER ORKHON EC;Lo;0;R;;;;;N;;;;;
10C31;OLD TURKIC LETTER YENISEI EC;Lo;0;R;;;;;N;;;;;
10C32;OLD TURKIC LETTER ORKHON AQ;Lo;0;R;;;;;N;;;;;
10C33;OLD TURKIC LETTER YENISEI AQ;Lo;0;R;;;;;N;;;;;
10C34;OLD TURKIC LETTER ORKHON IQ;Lo;0;R;;;;;N;;;;
10C35;OLD TURKIC LETTER YENISEI IQ;Lo;0;R;;;;;N;;;;
10C36;OLD TURKIC LETTER ORKHON OQ;Lo;0;R;;;;;N;;;;
10C37;OLD TURKIC LETTER YENISEI OQ;Lo;0;R;;;;;N;;;;;
10C38;OLD TURKIC LETTER ORKHON AR;Lo;0;R;;;;;N;;;;;
10C39;OLD TURKIC LETTER YENISEI AR;Lo;0;R;;;;N;;;;;
10C3A;OLD TURKIC LETTER ORKHON AER;Lo;0;R;;;;;N;;;;;
10C3B;OLD TURKIC LETTER ORKHON AS;Lo;0;R;;;;;N;;;;
10C3C;OLD TURKIC LETTER ORKHON AES;Lo;0;R;;;;N;;;;
10C3D; OLD TURKIC LETTER ORKHON ASH; Lo; 0; R;;;;; N;;;;;
10C3E;OLD TURKIC LETTER YENISEI ASH;Lo;0;R;;;;;N;;;;;
10C3F;OLD TURKIC LETTER ORKHON ESH;Lo;0;R;;;;;N;;;;;
10C40;OLD TURKIC LETTER YENISEI ESH;Lo;0;R;;;;;N;;;;;
10C41;OLD TURKIC LETTER ORKHON AT;Lo;0;R;;;;;N;;;;;
10C42;OLD TURKIC LETTER YENISEI AT;Lo;0;R;;;;;N;;;;;
10C43;OLD TURKIC LETTER ORKHON AET;Lo;0;R;;;;;N;;;;;
10C44;OLD TURKIC LETTER YENISEI AET;Lo;0;R;;;;;N;;;;;
10C45; OLD TURKIC LETTER ORKHON OT; Lo; 0; R;;;;; N;;;;;
10C46; OLD TURKIC LETTER ORKHON BASH; Lo; 0; R;;;;; N;;;;;
```

Bibliography

Erdal, Marcel. 2004. A grammar of Old Turkic. Leiden & Boston: Brill.

Faulmann, Carl. 1990 (1880). Das Buch der Schrift. Frankfurt am Main: Eichborn. ISBN 3-8218-1720-8
Kara, György. 1996. "Aramaic scripts for Altaic languages", in Peter T. Daniels and William Bright, eds. The world's writing systems. New York; Oxford: Oxford University Press. ISBN 0-19-507993-0
Róna-Tas, A. 1991. An introduction to Turkology. Szeged.

Scharlipp, Wolfgang Ekkehard. 2000. Eski Türk run yazıtlarına giriş: ders kitabı = An introduction to the Old Turkish Runic inscriptions: A textbook in English and Turkish. Engelschoff: Auf dem Ruffel. ISBN 3-933847-00-X

Tekin, Talât. [1968]. A grammar of Orkhon Turkic. Bloomington: Indiana University.

von Gabain, A. 1941. *Alttürkische Grammatik mit Bibliographie, Lesestücken und Wörterverzeichnis, auch Neutürkisch*. Mit vier Schrifttafeln und sieben Schriftproben. (Porta Linguarum Orientalium; 23) Leipzig: Otto Harrassowitz.

Малов, С.Е. 1951, Памятники Древнитюркской Письменности, Москва & Ленинград.

10. Acknowledgements. This project was made possible in part by a grant from the U.S. National Endowment for the Humanities, which funded the Script Encoding Initiative in respect of the Old Turkic encoding.

Figures

The characters of the old Turkic script are stable in form. Some of them, however, have varieties which slightly differ from each other.

The characters are given in the following table.

Characters of the Old Turkic Alphabet

Characters	Transliteration	Transcription
1	a	<u>a, ä</u>
1	i	<u>i, ï</u>
>}	u	<u>u</u> , <u>o</u>
h h	ü	<u>ü, ö</u>
(T, Kč را)	b ¹	<u>ab</u> , <u>b</u>
X (T, O, KČ)) b ²	<u>äb</u> , <u>b</u>
(T, KČ J) (T, o, kč) (KČ)	d^1	ad, d
X	d^2	<u>äd</u> , <u>d</u>
) ₁ ((T 1 ₁ 1)	Υ	<u>а</u> у, у
(T, 0), Kč	g	äg, g
ήλ	,\ P	aq, q
(T, 0), KČ	/) k	äk, k
1	11	<u>al</u> , <u>l</u>
Y	1 ²	<u>äl</u> , <u>1</u>
· ·)	n^1	an, n
# (T 4 7), KČ	n ²	än, n
44	r^1	ar, r
7	r²	är, r
\$ (I) KČ Z) s ¹	as, s

Figure 1a. Table of Old Turkic letters, from Tekin1968.

Characters	Transliteration	Transcription
	s²	ās, s
(T, o, Kč 🖒, o	⟨ t¹	at, t
hh	t²	<u>āt, t</u>
D	y¹	<u>ay</u> , <u>y</u>
D) (T), 09, Kč, ↓ (T, 0B, Kč R) (T, 0D)	y²	<u>ä</u> y, y
1	8	oq, uq, qo, qu, q
Ä (T, OB, KČ Å)	8	ök, ük, kö, kü, k
√ (T, O ▷)	$\mathbf{\dot{q}}$	<u>ïव, वृï, व</u>
+	<u>ič</u>	ič
* 人 》(エ〉) う 1 * (T.O. Kč 4)	č	<u>ač, ač, č</u>
(∀ T) ≪	m	am, äm, m
3	ń	<u>ań, äń, ń</u>
J	ŋ	<u>aŋ, äŋ, ŋ</u>
1	, P	ap, äp, p
Y	š	as, äš, š
f (т.ο, κč 4)	z	az, äz, z
M }	<u>lt</u>	alt, lt
}	nč	anč, anč, nč
(∵) (Т (), О (⊙, КС́	①) <u>nt</u>	ant, ant, nt
1	aš	aš
\bowtie	baš	baš²

² Apart from the letters discussed above and given in the table, the following signs occur in the Uigur and

Figure 1b. Table of Old Turkic letters, from Tekin1968.

1.12. Origin of the Alphabet

Various theories have so far been put forward on the srigin of the Turkic 'runic' script. Otto Donner, still before the script was deciphered, drew attention to the general conformity between the characters of the Yenisei script, that is, the Turkic 'runic' script, and the Lycian and Carian alphabets of Asia Minor of the antiquity, and assumed that the Yenisei script might have been derived from these alphabets. Thomsen, who deciphered the Turkic 'runic' script, did not accept this assumption and put forward the theory that the Turkic alphabet was derived from the late Semitic (Aramaic) alphabet, with or without an Iranian intermediary. Aristov and Mallitskiy suggested that the letters of the Turkic 'runic' script might have been developed from the Turkic tamgas. According to Polivanov,

Tenisei inscriptions, and in the manuscripts found in **Eastern** Turkestan: $Y = s^1$ (in back-vocalic words), $Y = s^1$ (closed e), $Y = s^1$, $Y = s^1$ (in the manuscripts), $Y = s^1$ (in $Y = s^1$), $Y = s^1$ (in $Y = s^1$), $Y = s^1$

Figure 1c. Notes on additional Old Turkic letters, from Tekin1968.

³ <u>Inscriptions de l'Orkhon</u>, p. XLXXX ff.

⁴ Inscriptions de l'Orkhon déchiffrées, p. 44. ff.

⁵ N. N. Aristov, 'Zametki ob etničeskom sostave prkskix plemen i narodov,' in: <u>Živaya Starina</u> III-IV, 1896 (I quote this from A. Caferoğlu, <u>Türk Dili Tarihi</u> I, 1958, p. 115).

N. G. Mallitskiy, 'O svyazi tyurkskix tamga s orxonmini pismenami,' in: Protok. sozed. i soobšč. Turkesmskogo kružka lyub. arxeologii, 1897-1898 (Quoted from A. Caferoğlu, Türk Dili Tarihi I, İstanbul 1958, p.

Die türkische Runenschrift

Orchon I, II Orchon III Jenissei [O, T] [Hss.] (Ta)	Orchon I, II Orchon III Jenissei [O, T] [Hss.] [Ta]
$a = \ddot{a}$ f = $[=\mathcal{F}]$ = 1 S; oder f	2n H =[= 4h] = 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
e 1 — H	${}^{0}y$ ${}^{\bullet}$ $[=\}]$ ${}^{=}$ oder ${}^{\bullet}$ ${}^{\bullet}$ ${}^{1}y$ and ${}^{\bullet}$ ${}^{\bullet}$ ${}^{\circ}$
	⁰ ny → = 为 [= 为] -;XXXVII €
o=u > }	${0n\ddot{c}\atop -0n\ddot{c}} \left\{ \left[\text{OT} \right\} \right] = \left[= \right\} = \left\langle \right. \right\} \left(\right. \right\}$
	= 11g
$\begin{bmatrix} 1y & D & = [=D] & = 0 & 0 \\ 2y & 9 & [=Q] & [Q] & = P \end{bmatrix}$	⁰ nt ♥ [O ♥; = ⁰ nd T ⊙] • [=] -
¹ b δ J δ [= δ]	⁰ p 1 = =
2b 文 其 [O 其 克, 本 [余 会] · 本	op = up [
- X J	^{1}q of A = $[=g()$ = N
${}^{0}\check{c} = {}^{0}\check{g} \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \$	<i>iq</i> = <i>qi</i> 4 [0 ▶] -[= △] -; XXXVII ▶
$i\check{c}=i\check{g}$ Y [=]	00 = 40
¹d % = = ₹ }}	$=q_0=q_u \qquad -[=\downarrow] \qquad \uparrow$
³ d × =[=†} =	¹ т Ч = = = 1 4 И
¹γ ¥ [Τ ነነ] ≖ ነነ ነነ ¼ Ж Ƴ	²r Υ =[=
2g f [0 f]=7[7f]= f f f 2k 7 =>(55]=> 7	18 (4.4.) 4 (4.4.) 1 (4.1.0 8) 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
arif black	
$\begin{vmatrix} \ddot{o}k = \ddot{u}k \\ = k\ddot{o} = k\ddot{u} & \text{FR[B] [B]} & \text{B} & \text{B} \end{vmatrix}$	28 = [= 1] =
1/	% ¥[Ÿ Y] = Y[= 大] ハリコー
² ℓ Υ =[=Υ] =	¹t \$[O 数; OT ô] ô @[=※] ≪[♥]
$\left \begin{array}{cc} {}^{0}ld \\ {}^{\bullet}{}^{0}lt \end{array} \right $ [=] $\left \begin{array}{cc} {}^{\bullet}XXXIII, \\ XXXVII \end{array} \right $	ot=ut [A]
0m >> [O×)] = [= → = [>=(⊙]	²thh = =4
¹ n) = =	⁰ 2 # # ፟ ፟፟፟፟፟፟፟፟፟

Figure 2. Table of Old Turkic letters in Orkhon and Yenisei, taken from von Gabain 1941. This summary table is the source for the inherent vowel quality taken for the character names. It contains all of the characters in the proposal.

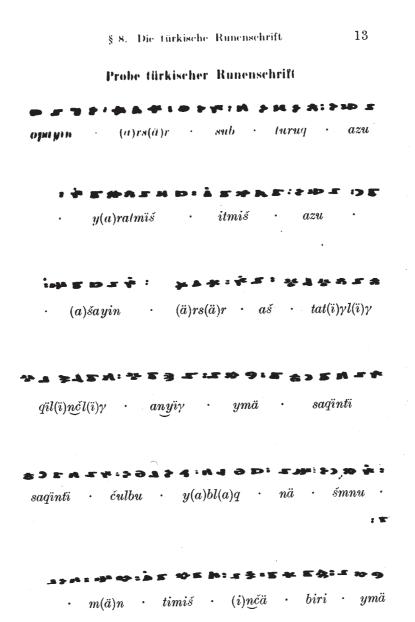


Figure 3. Sample of Old Turkic text from a manuscript, taken from von Gabain 1941.

Figure 4. Sample of Old Turkic text.

ئورقۇن-يېنىسەي يېزىقىنىڭ ئېلىپبەسى

No	ھەر پلەر	ئوقۇلۇشى	N ₀	ھەر پلەر	وقؤلؤشى	5
1	↑ 1X	ئې/ئا a//e	21	444	n يۈمشاق	ن
2	111	ئى/ئە a//i	22	444	a قاتتىق	ز
3	> >	ئۆ/ئو o//u	23	T	يۇمشاق r	١,
4	K WH	Ø//y ٷ/ٷ	24	1 <	قاتتىق 1	J
5	4 4	ق قاتتىق q	25	Y	يۇمشاق 1	J
6	4 4	ق a بىلەن كېلىدۇ q	26	8	قاتتىق s	س
7	→ ↑	o//u بىلەن كېلىدۇ ق q	27	1	يۇمشاق 5	ٍ س
8	44 6	ن يۇمشاق k	28	3	اق تاۋۇشى	دىما
9	Fi BB	k يىلىن كېلىدۇ ك %	29	1	p	Ÿ
10	36 it it it	غ قاتتىق كا	30	1 ry	ŋ	لڠ
11	E L	ت يۇمشاق g	31	≯ ८	m	٠
12	🕏 ጵስጵ	ت قاتتىق t	32	λ	t∫	٤
13	hhh	ت يۇمشاق t	33	Ψ	it∫	ئىچ
14	¾ ₩ 33	د قاتتىق d	34	X Y FYIN	l	ش
15	X X	د يۇمشاق d	35	4458	z	ز
16	3 d3	ب قاتنىق b	36	M	lt//ld	لت
17	₹ ₹ Œ	ب يۇمشاق b	37	७ ତ⊕	nt//nd	نت
18	D O D	ي قاتتىق ز	38	3	nt//nət	نەت
19	9 P	ي ي ۇ مشاق ز	39	M	rt	وت
20)	ن قاتتىق n	40	:	ى بەلگىسى	ئاير ىن

Figure 5. Table of Old Turkic characters from an Uighur source.

قەدىمكى تۈرك (ئورمۇن) يېزىقساڭ ئېلىپبەسى

_		-		_			
Nº	تەلەپپۇ زى	ئورغۇن	يېنسەي	Nº	تەلەپ چۇزى	ئورطۇن	يېنسەي
1	A .a	1	1 X I	21	ny	3	8
2	i.ĭ	1	4	22	ng	4	Y
3	o,u	> } }		23	n')	
4	θ,ü	74	μ	24	n²	*	2 4
5	Қ	4 4		25	m	≫	
6	K_{Ki}^{iK}	4	DΧ	26	r'	4	44
7	оқ,иқ Қо.қи	1	↑	27	rz	~	
8	o)(ም የ ት ቾ	28	1'	1	~
9	K	7	99	29	12	Y	1
10	θk,ük Kθ∙kü	Fi	В	30	9	\	λ
н	g	E	£ € £	31	19	Ψ.	
12	f,	\$ \$ 6	â \$	32	5'	5	
13	t²	hh		33	S²	1	
14	d۱	33	¾ (⊙ ◊)	34	X	¥	メイヤン
15	d²	X		35	Z	44	५ % % የ
16	P	1		36	nt	ఆ	0
17	Ь'	9	9920	37	nq	3	23{}
18	Ь²	交叉	余 久	38	14	M	
19	У'	D	00	39	ax	1	
20	У²	9	P (P)	40	bax	\bowtie	

Figure 6. Table of Old Turkic characters from an Uighur source.

RÜNİK TÜRK ALFABESI

ORHUN I,II	ORH. III	YENISEY	ORHUN · I , II	ORH	YENISEY
1 a, e		1 X(X = e)	9 2		P (P)
) + 0, 4	}	,	3 y		(Exxxvii)
PPö. ä	4	h) n	,	
44 ¥		(N	µ n ≯m	-	11
♦ k (v-dan evelve sora		() xxxv11)	9 m		44
↓ k(o, a-dan evelve sora		↑	Υ .		
of gelin.)		th XXXX	1 1		·
9 k	7	3 4	Y		2 1
R kio ii - den evel ve sona gelis,		В	Ŷ iq		>
€ ĕ		₹ € ₹	4.		
I	ဂ်င်္ဂ	~	1 0		
hh ^t 3 d		% 33.⊙ ♦	¥ ?	۴	*
X d			₩ ₩ ≈	4	45525
1 ^p		50 00	nd o	0	0
286	ద		3 nç		2311
Dy		00	Med		

Figure 7. Table of Old Turkic letters showing Orkhon and Yenisei varieties, from "Iski Türk Yazitlari", Yusayin Namiq Orqun, Istanbul.

ያ የ የ ነ የ ነ የ ነ የ ነ የ ነ የ ነ የ ነ የ ነ የ ነ የ	, а 61 г, а г	a, ā b¹ b², ā ' g d¹	Начертание	й нт нч о, у ö, ў	100
\$ ጵ ፯ የ ነገር ነ የ አ የ ነገር ነ የ አ የ ነገር ነ የ ነ የ ነ የ ነ የ ነ የ ነ የ ነ የ ነ የ	6 ¹ ² , ä ⁵ r	b1 b2, ă ; g d1	γηη >> Θ⊙Θ÷⊗	нт нч о, у	nt
\$ ጵ ፯ የ ነገር ነ የ ር ፡፡ ፡፡ ፡፡ ፡፡ ፡፡ ፡፡ ፡፡ ፡፡ ፡፡ ፡	6 ¹ ² , ä ⁵ r	b1 b2, ă ; g d1	γηη >> Θ⊙Θ÷⊗	нт нч о, у	nt nč o, u
\$ \$ ፬ የ ነገ ነር ነላ የ አ * * * * * * * * * * * * * * * * * * *	2, ä 5 r	b*, ä ; g d¹	ЬИ Н >>	нч о, у	nč o, u
¥''' እ(አ & ደ ፠፠፠ × ዜ ተ '' ዜ ተ	r A ¹	g d ¹	ЬИ Н >>	o, y	o, u
€¢ ¾¾¾ × ₩₩₩\$8	r A ¹	g d¹	Lh h	100000	100
*}} × ዓለት ዓ	A 1	d1		ö, ÿ	ö, ü
× ₩₩₩\$8			1		100
4 ተዛታሪ	Α1			п	р
		d =	44	p1	r 1
	а	z	4	p 2	r2
11	, i	у, і	4	c 1	s ¹
D	j1	j', ī	1	c 2	s 1
9	j ^a	j², ī	\$ ☆	т1	t1
HADTT	k	q	h	T2	t2
ARTTBE	ĸ	k	A.Y	4	ě
1	1	11	** # *	m	š
Y	:	12	M		
M	AT	lt			
& &	N	m			
)	t1.	n t			
\ \\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\	12	n ²			
414	ų	ц			

Figure 8. Table of the "Old Turkic Runic Alphabet", from Малов 1951.



Figure 9. An Old Turkic manuscript.



Figure 10. Monument with an Old Turkic inscription

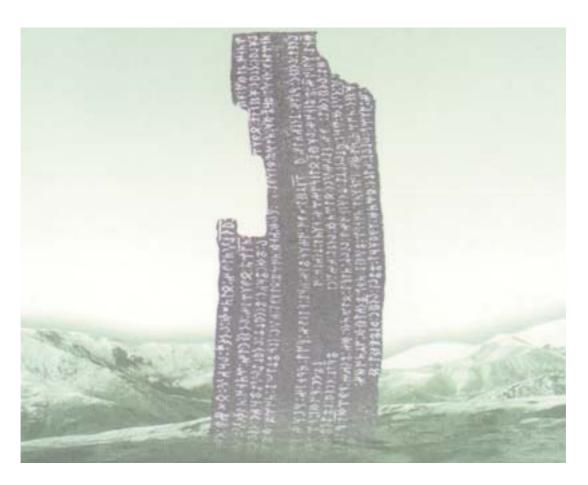




Figure 11. Old Turkic monuments.

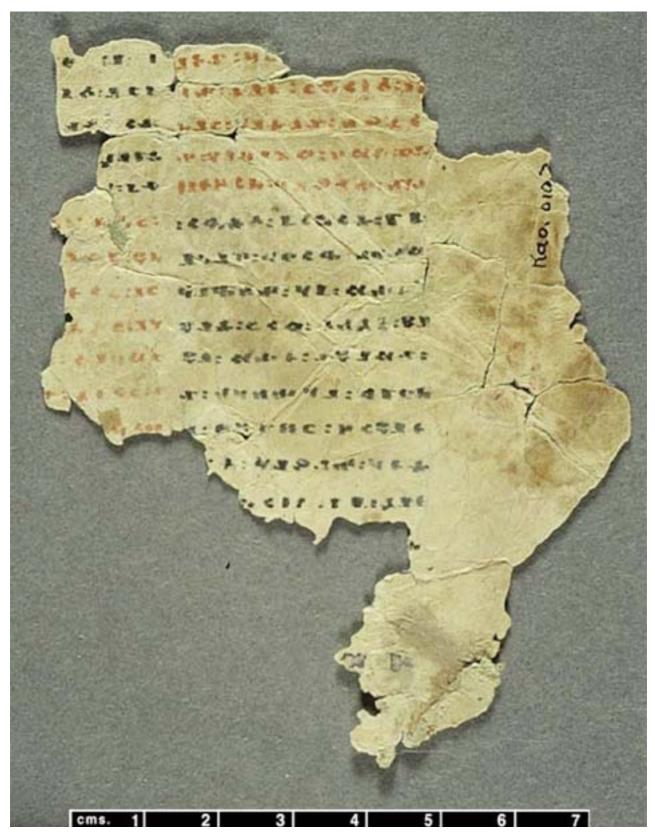


Figure 12. An Old Turkic manuscript.

```
1⊬1/491: IdX×4/H: 1</br>
1/1/494/H: 41: D1
```

Памятник хану Могиляну

```
ልዓለታ : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha : \alpha
```

Памятники рунического письма Монголии VII-VIII вв

14

```
18; $1744; F179$CY; .... (0b) .... YHPCH4; (%1); H1741; H2741; H274; (%1); H274; (%1); H274; H2
```

Figure 13. Selections from several different editions of Old Turkic texts.

A. Administrative

1. Title

Proposal for encoding the Old Turkic script in the BMP of the UCS

Requester's name

China, Ireland, the UK, and UC Berkeley Script Encoding Initiative (Universal Scripts Project)

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

National Body and Liaison contribution.

4. Submission date

2008-01-25

- 5. Requester's reference (if applicable)
- 6. Choose one of the following:

6a. This is a complete proposal

No

6b. More information will be provided later

Yes.

B. Technical - General

1. Choose one of the following:

1a. This proposal is for a new script (set of characters)

Yes.

1b. Proposed name of script

Old Turkic.

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

No.

- 1d. Name of the existing block
- 2. Number of characters in proposal

71.

3. 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 or Ideographic; G-Obscure or questionable usage symbols)

Category E.

4a. Is a repertoire including character names provided?

Ves

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? **Wushour Silamu and Michael Everson.**

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

Wushour Silamu and Michael Everson, Fontographer.

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?

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 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 Unicode Character Database http://www.unicode.org/Public/UNIDATA/UnicodeCharacterDatabase.html and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.

See above.

C. Technical – Justification

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

No.

2a. Has contact been made to members of the user community (for example: National Body, user groups of the script or characters, other experts, etc.)?

Yes.

2b. If YES, with whom?

Experts at Xinjiang University and Xinjiang Museum in China, Marcel Erdal (France), Mehmet Olmez (Turkey)

2c. 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?

Turkologists and other scholars.

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

Rare; historical use.

4b. Reference

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

Yes.

5b. If YES, where?

Scholarly and popular publications.

6a. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP?

No.

6b. If YES, is a rationale provided?

6c. If YES, reference

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

Yes.

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

No.

8b. If YES, is a rationale for its inclusion provided?

8c. If YES, reference

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

No.

9b. If YES, is a rationale for its inclusion provided?

9c. If YES, reference

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

No.

10b. If YES, is a rationale for its inclusion provided?

10c. If YES, reference

11a. Does the proposal include use of combining characters and/or use of composite sequences (see clauses 4.12 and 4.14 in ISO/IEC 10646-1: 2000)?

No.

11b. If YES, is a rationale for such use provided?

11c. If YES, reference

11d. Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?

No.

11e. If YES, reference

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

No.

12b. If YES, describe in detail (include attachment if necessary)

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

No.

13b. If YES, is the equivalent corresponding unified ideographic character(s) identified?

	10C0	10C1	10C2	10C3	10C4
0	10000	10010	10C20	ل)
1	10001	>> 10C11	10C21	λ	\$
2	10002	10C12	10C22	10032	10C42
3	10003	10C13	10023	10033	10C43
4	10004	10014	10024	10034	Ч
5	Ŏ	4	⊙	Þ	10044
6	10C05	10C15	10C25	10C35	10C45
7	10006	10C16	10C26	10C36	10C46
8	10007	7	3	10C37	
9	10C08	10C18	10C28	10C38	
Α	10009	10019	10C29	10C39	
В	10C0A	10C1A	10C2A	10C3A	
С	10C0B	10C1B	10C2B	10C3B	
D	10C0C	10010	10C2C	10C3C	
E	10C0D	10C1D	10C2D	10C3D	
F	10C0E	10C1E	10C2E	10C3E	
r	10C0F	10C1F	10C2F	10C3F	

Date: 2008-01-25

Vowels

10C00	1	OLD TURKIC LETTER ORKHON A
10C01	1	OLD TURKIC LETTER YENISEI A
10C02	χ	OLD TURKIC LETTER YENISEI AE
10C03	1	OLD TURKIC LETTER ORKHON I
10C04	١	OLD TURKIC LETTER YENISEI I
10C05	Ŏ	OLD TURKIC LETTER YENISEI E
10C06	>	OLD TURKIC LETTER ORKHON O
10C07	۲	OLD TURKIC LETTER ORKHON OE
10008	н	OLD TUDVIC LETTED VENICELOE

10C08 OLD TURKIC LETTER YENISEI OE **Consonants** 10C09 δ OLD TURKIC LETTER ORKHON AB 10C0A → OLD TURKIC LETTER YENISEI AB 10C0B \$ OLD TURKIC LETTER ORKHON AEB 10C0C ♦ OLD TURKIC LETTER YENISEI AEB 10C0D Y OLD TURKIC LETTER ORKHON AG 10C0E '\' OLD TURKIC LETTER YENISEI AG OLD TURKIC LETTER ORKHON AEG 10C10 £ OLD TURKIC LETTER YENISEI AEG 10C11 » OLD TURKIC LETTER ORKHON AD 10C12 * OLD TURKIC LETTER YENISEI AD 10C13 X OLD TURKIC LETTER ORKHON AED 10C14 # OLD TURKIC LETTER ORKHON EZ 10C15 4 OLD TURKIC LETTER YENISEI EZ 10C16 DOLD TURKIC LETTER ORKHON AY 10C17 O OLD TURKIC LETTER YENISEI AY 10C18 9 OLD TURKIC LETTER ORKHON AEY 10C19 ▶ OLD TURKIC LETTER YENISEI AEY 10C1A 7 OLD TURKIC LETTER ORKHON AEK 10C1B Y OLD TURKIC LETTER YENISEI AEK 10C1C B OLD TURKIC LETTER ORKHON OK 10C1D J OLD TURKIC LETTER ORKHON AL 10C1E V OLD TURKIC LETTER YENISEI AL 10C1F Y OLD TURKIC LETTER ORKHON AEL 10C20 M OLD TURKIC LETTER ORKHON ELT 10C21 冷 OLD TURKIC LETTER ORKHON EM 10C22 > OLD TURKIC LETTER ORKHON AN 10C23 🖟 OLD TURKIC LETTER ORKHON AEN 10C24 № OLD TURKIC LETTER YENISEI AEN 10C25 ♥ OLD TURKIC LETTER ORKHON ENT 10C26 > OLD TURKIC LETTER ORKHON ENC 10C27 → OLD TURKIC LETTER YENISEI ENC 10C28 3 OLD TURKIC LETTER ORKHON ENY 10C29 6 OLD TURKIC LETTER YENISEI ENY 10C2A ♦ OLD TURKIC LETTER YENISEI ANG 10C2B 1 OLD TURKIC LETTER ORKHON ENG 10C2C F OLD TURKIC LETTER YENISEI ENG 10C2D 1 OLD TURKIC LETTER ORKHON EP 10C2E # OLD TURKIC LETTER ORKHON OP 10C2F Y OLD TURKIC LETTER ORKHON IC 10C31 λ OLD TURKIC LETTER YENISEI EC 10C32 d OLD TURKIC LETTER ORKHON AQ 10C33 ► OLD TURKIC LETTER YENISEI AQ 10C34 OLD TURKIC LETTER ORKHON IQ 10C35 ▶ OLD TURKIC LETTER YENISEI IQ 10C36 ↓ OLD TURKIC LETTER ORKHON OQ 10C37 ↑ OLD TURKIC LETTER YENISEI OQ

10C38 4 OLD TURKIC LETTER ORKHON AR

Date: 2008-01-25

10C39 1 OLD TURKIC LETTER YENISEI AR 10C3A Y OLD TURKIC LETTER ORKHON AER 10C3B Y OLD TURKIC LETTER ORKHON AS 10C3C | OLD TURKIC LETTER ORKHON AES 10C3D ☐ OLD TURKIC LETTER ORKHON ASH 10C3E ^ OLD TURKIC LETTER YENISEI ASH 10C3F Y OLD TURKIC LETTER ORKHON ESH OLD TURKIC LETTER YENISEI ESH 10C41 ₺ OLD TURKIC LETTER ORKHON AT 10C42 ☆ OLD TURKIC LETTER YENISEI AT 10C43 h OLD TURKIC LETTER ORKHON AET 10C45 OLD TURKIC LETTER ORKHON OT 10C46 M OLD TURKIC LETTER ORKHON BASH