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

```
Doc Type:Working Group DocumentTitle:Proposal for encoding the Todhri script in the SMP of the UCSSource:UC Berkeley Script Encoding Initiative (Universal Scripts Project)Authors:Michael EversonStatus:Liaison ContributionAction:For consideration by JTC1/SC2/WG2 and UTCDate:2020-10-01
```

1. Introduction. A number of 18th-and 19th-century orthographies exist for Albanian; one of the more significant ones given its corpus is called Todhri (/'toðri/), from the nickname of the man who devised it, Dhaskal Todhri 'teacher Todhri', whose name was Theodor Haxhifilipi (§1730–1805). Unfortunately, I did not get a chance to work with the late Robert Elsie on this proposal, which never benefitted from his expertise due to his untimely death. I give below here the introduction to the script from Elsie 2017:

The Todhri alphabet is an original writing system that was used for Albanian in the central Albanian region of Elbasan in the eighteenth and nineteenth centuries. It consists of 36 letters (plus nine ligatures and six rare graphemes). This script is attributed to one Dhaskal Todhri ("teacher Todhri"), whose full name was Theodor Haxhifilipi (ca. 1730-1805), and who stemmed from Elbasan.

The biblical texts in this Todhri alphabet were discovered by the Austrian consul in Janina, Johann Georg von Hahn (1811-1869), who announced his finding in an article entitled *Bericht über die Auffindung eines uralten albanesischen Alphabetes* (Report on the Discovery of an Ancient Albanian Alphabet), submitted to the Austrian Academy of Sciences in November 1849. Hahn published a small portion of this material, in what he regarded as "the original" Albanian alphabet, in his seminal three-volume *Albanesische Studien* (Albanian Studies), Jena 1854, and regarded it, rather naively, as stemming from ancient Phoenician script. On the author and origin of the material in Todhri script, Hahn explains:

"As far as I have been able to ascertain, this alphabet seems to be native only to the town of Elbasan. It is unknown in Durres, Shkodra and in all of Toskeria (Epirus), but it seems to have been understood and used a bit in the neighbouring town of Berat to the south.

In Elbasan, usage of this alphabet is not limited to ecclesiastical texts. It is used there by several merchants for correspondence with their compatriots abroad. Some of them, indeed, keep their accounts in this script, but I doubt whether there are more than fifty people who know this script nowadays.

This alphabet is traditionally attributed to a certain Theodor, a teacher at the Greek school there, who was not only a town preacher but also a well-educated man. It cannot be ascertained, however, whether he invented the alphabet or was simply the first person to bring it to Elbasan. He is said to have translated into Albanian not only the New Testament, but also the Old Testament (from the Septuagint [Greek translation} because he did not know Hebrew). Unfortunately all of his writings were burnt by his relatives during a major outbreak of the plague, out of fear of infection. Judging from the age of his living relatives, he probably died around the end of the last century. He had studied in Moschopolis [Voskopoja], ten hours to the east of Berat. [...]

I searched in vain in Elbasan for manuscripts in this Albanian alphabet. All I could find were two copybooks, one of seven pages and the other of twelve pages in quarto; fragments of a Gheg translation of the Greek Horologion; and a Gheg translation of a fragment of the Gospel of St John written on a folio sheet. The man who wrote the older of these works in his youth was still alive. As such, none of these manuscripts was over fifty years old. "

Todhri stemmed from a family of Orthodox silversmiths. For periods in the second half of the eighteenth century, he apparently kept the register of the Orthodox Church of Saint Mary's in the fortress of Elbasan, though he was not a clergyman himself. In these records, he called himself Theodor, son of Haxhi, or Theodor Dhaskalos (teacher). He seems to have died around 1805.

After Hahn, the study of Todhri script was taken up by Leopold Geitler (1847–1885), professor at the University of Zagreb, in his work *Die albanesischen und slavischen Schriften* (Albanian and Slavic Scripts), Vienna 1883, who regarded the Todhri alphabet as derived from Greek and Roman cursive. He had access to the copybooks that Hahn had found in Elbasan, and in 1883 he deposited them in the Imperial Library, now the Austrian National Library, in Vienna. The thirty sheets (46 pages) are still preserved in the manuscript division of the Library. This manuscript, kindly made available by the manuscript division of the said library, forms the basis of this publication. Other texts in Todhri script, of which there may have been many, seem, tragically, to have been lost.

In 1901, the Albanian linguist Gjergj Pekmezi (1872-1938), in his *Vorläufiger Bericht über das Studium des albanischen Dialektes van Elbasan* (Preliminary Report on the Study of the Albanian Dialect of Elbasan) sent to the Austrian Academy of Science, reported that he had come across texts in Todhri script on his journey to Elbasan in 1900:

"It was quite by coincidence that during my stay in Elbasan, an over eighty-year-old monk called Parthenius came to see me for health reasons from the Monastery of St. John Vladimir which is situated an hour away from Elbasan. He is no doubt the oldest man alive who knows how to write in this alphabet. He told me that he had learned it from a student of the Theodor who was said to have invented it. This student, called Camblak, succeeded Theodor as a teacher in Elbasan and later became Hegumen [abbot] of the said monastery. According to him, Theodor invented the script and brought it with him from Moschopolis where he had studied. The script is now dying out. It was pushed aside by the Kristoforidhi alphabet, whereas now it is the Fashëri alphabet that is being used. Very few people (some merchants and a woman who gave me a sample of her writing) still use the old script for their letters. Church hymns can also be found written in it. Not too long ago, it was being used by Demetrios Andreas Pina, a native of Elbasan and one-time bishop of central Albania. [...]

Parthenius told me that there was a chest in the church choir of the monastery with some unused church utensils, including books and manuscripts, in it. He thought there might still be something of Theodor's there, even though, according to Hahn, almost everything was burnt. I took up his suggestion and found the chest, but it was in a terrible state of neglect. There and in the side altar to the left, there were not only sheets of Greek parchment and hymnbooks, but also a few pages and copybooks written in this Elbasan script. They seemed to stem from various ages because, among them, were some of Pina's sermons.

There were three copybooks, containing 6-20 sheets, of the size of those discovered by Hahn, which were probably the oldest – but just how old is difficult to say. They were written in various hands. One of them with some church hymns probably stemmed from the eighteenth century. A second one contained some ancient Greek fables with interlinear modern Greek and Albanian translations. There were also pages from the New Testament in other handwriting. The third copybook and various separate sheets were much the same. The most interesting of all was a rather dirty page of thick paper,

torn in the middle, that seemed to contain the oldest Albanian fragment. It was most certainly from the eighteenth century, and Parthenius seemed to believe it was written by Theodor. In the corner of the front page was the number 55 and on the back page was 56, an indication that it derived from a larger body of writing. The contents also made this clear - two sections from the First Epistle of Paul to the Romans. [...]"

Together with this information, Pekmezi published two facsimiles of the aforementioned manuscripts in Todhri script.

In 1918, the first six pages of the manuscript of the Divine Liturgy of Saint John Chrysostom were then published by the Elbasan scholar Lef Nosi (1877–1946). Nosi's substantial archives and collections were confiscated by the communists after his execution and some of the material found its way into the state archives. It was the Elbasan scholar of the communist period, Dhimitër Shuteriqi (1915–2003), who made an inventory of Todhri writings and published several articles in the late 1940s and 1950s. He was the first to make this fascinating component of Albania's cultural heritage widely known in Albania.

Todhri script was also the subject of an article by the Slovenian scholar Rajko Nahtigal (1877–1958) published in 1923 and entitled *O elbasanskem pismu in pismenstvu na njem* (On the Elbasan Script and Writing in it), and was touched upon briefly by the German scholar Armin Hetzer (b. 1941) in 1987 in another article entitled *Armenier und Albaner* (Armenians and Albanians).

Todhri script seems to have been used sporadically in Elbasan well into the twentieth century. The family chronicle of Popaj family of Elbasan, kept from 1785 to 1933, was written in Todhri. There are also indications that the last person to write in Todhri script was a woman in Elbasan called Dona Xhuvani who died in the 1930s. However, nothing can be said for certain until we have a full inventory of all texts written in Todhri, and this is unlikely to occur anytime soon. Most of the material has vanished. For the full text, see: Robert Elsie, *Early Albanian Bible Translations in Todhri Script*. Albanian Studies, Vol. 30. (London: Centre for Albanian Studies, 2016) 162 pp.

2. Processing. Todhri is a simple, unicameral alphabetic script written from left to right horizontally. There is no real ligation, though Elsie describes some of its characters as "ligatures" but this seems more a likely description of their derivations than of a typographic quality.

2.1 Diacritical marks. A number of diacritical marks can be observed to be used in the Todhri texts. Some are straight, some curved, some angled to the right or to the left. They are used only with vowels, and remind one of Greek accents, but in Elsie's transcriptions, they are uniformly represented in Latin by the ACUTE ACCENT, and are used to indicate stress. There might be a tendency to use one with narrower characters and the other with wider ones, but in any of the manuscripts there are exceptions. If a concordance of the manuscripts' texts were made, a frequency of choice could be established. In the meantime, however, it is prudent to recommend that U+0301 COMBINING ACUTE ACCENT be used for the vertical mark and that U+0304 COMBINING MACRON be used for the horizontal. Where a macron appears over two vowels (as it does in a number of instances) then U+035E COMBINING DOUBLE MACRON can be used. The remaining mark seems to be found only over the letter OO, which derives from Greek OMEGA; so far a smooth breathing (for Greek U+0313 COMBINING COMMA ABOVE) has been found over 00 when used as a vocative particle; this mark is sometimes omitted, and may or may not have a stress mark above it. Fonts should support the following repertoire. In accordance with what is seen in the manuscript, the smooth breathing looks best when represented by U+0311 COMBINING INVERTED BREVE. (Some unusual combinations (which can't be breathings) occur (Figure 4): $\Box \delta \dot{\delta} \dot{\delta} d$ prûûm, $\nabla \bar{\delta} \nu \dot{\gamma} i \bar{a} n \ddot{e}$ 'they are'; I list them in parentheses below. Perhaps these are just $\bar{000}$ prūūm and $\bar{000}$ jánë; in Figure 6 we find $\bar{000}$ jánë.)

 $\acute{v}\ \bar{v}\ \bar{v}\ \acute{i}\ \ddot{i}\ \ddot{i}\ (i\ \bar{i}\ \bar{i}\ \bar{i}\ (i\ \bar{v}\ \dot{o})$

2.2 Albanian vowels. Albanian has seven vowel phonemes: a/a/, $e/\epsilon/$, $\ddot{e}/\vartheta/$, i/i/, $o/\vartheta/$, u/u/, and y/y/. The vowels *i* and *o* are represented with a vertical line I and a circle ϑ ; the vowels *e* and *u* have the same base shape with a dot above: i and ϑ . It has been suggested that options for representing these be presented to SC2 and to the UTC.

2.2.1 Atomic encoding. Encode v a, i e, $\neq \ddot{e}$, $\mid i$, $\circ/\cup o/\bar{o}$, $\circ u$, and $\flat y$ individually. This is the simplest model, and it is one which has been used successfully with a great many scripts. For a historical script like Todhri, the simplest model is the best, particularly as diacritics may be used to indicate stress.

2.2.2 Atomic encoding with canonical decomposition. Encode i e and $\circ u$ atomically, but make them canonically equivalent to i and $\circ o$ followed by COMBINING DOT ABOVE. There does not seem to be much rationale for this. Just because a variety of diacritics are used for stress does not mean that diacritics should also be permitted spellings for e and u. In fact for textual analysis the stress diacritics might be considered "noise" and it could be preferred to strip them out. In such a case the distinction between e and i on the one hand and o and u on the other could be lost. This encoding would work, but it would makes font development and inputting needlessly complicated.

2.2.3 Decomposed encoding. This proposal would have only five vowels encoded, and require the use of COMBINING DOT ABOVE to indicate two of the seven Albanian vowels. There is little to recommend this. In light of the use of a range of diacritics with all of the vowels, adding combining dot above for e and u would just complicate text representation and analysis, and invite multiple or erroneous spellings. This is by no means the right way to encode Todhri.

3.1 Repertoire. The repertoire of letters represents not only the basic phonemes of Albanian, but also a number of sequenced phonemes, such as letters for prenasalized consonants. In the table below, Group 1 gives the basic Albanian phonemes as expressed in the modern orthography; in Group 2 are given the prenasalized equivalents, and in Group 3 are Greek letters, and in Group 4 are other additions to the repertoire.

1.	А	В	С	Ç	D	DH	Е	Ë	F	G	GJ	Н	Ι	J	Κ	L	LL	М	Ν	NJ
	۷	M	L	q	Λ	6	i	7	в	h	ե	e	I)	(þ	Ħ	ç	V	б
2.	_	MB	_	_	ND	_	_	_	_	NG	NGJ	_	_	_	_	_	_	_	_	_
	_	b	_	_	ж	_	_	_	_	ĥ	К	_	_	_	_	_	_	_	_	_
3.	_	_	_	_	_	_	_	_	_	GH	_	_	_	_	KH	_	_	_	_	_
	_	_	_	_	_	_	_	_	_	v	_	_	_	_	Х					
4.	AS	_	_	_	_	_	_	_	_	_	_	HJ	_	_	_	_	_	_	_	_
	γ.	_	_	_	_	_	_	_	_	_	_	Э	_	_	_					
1.	0	Р	Q	R	RR	S	SH	Т	TH	U	V	Х	XH	Y	Ζ	ZH				
	0	И	%	5	ſ	s	8	q	Ş	ò	l	7	g	q	V	8				
2.	_	_	_	_	_	_	_	_	_	_	_	NX	NXH		_	_				
	_	_	_	_	_	_	_	_	_	_	_	Z	¥	_	_	_				
3.	00	PS	_	_	_	(SK)	_	ST	_	_	_	_	_	_	_	_				
	W	4	_	_	_	(8)	_	5	_	_	_	_	_	_	_	_				
4.	_	_	_	_	_	_	SHT	_	_	_	_	_	_	JY	_	_				
	_	_	_	_	_	_	R	_	_	_	_	_	_	ĥ	_	_				

3.1.1 Group 2. The prenasalized characters are by and large modifications of the base letters. Prenasalized *b* looks a bit like *m* ($d + m = \ell MB$); prenasalized *d* fuses both ($v + \Lambda D = w ND$);

prenasalized g and gj both share an extension (v N + h G = K NG; v N + h GJ = h NG); prenasalized x and xh share a similar extension (v N + \mathbf{y} X = \mathbf{z} NX; v N + \mathbf{g} XH = \mathbf{g} NXH).

3.1.2 Group 3. The 6 letters here answer to Greek characters: \mathbb{V} GH = γ GAMMA; X KH = χ CHI; ω OO = ω OMEGA; $\underline{\varphi}$ PS = ψ PSI; 5 ST = ζ STIGMA. The letter 8 SK may or may not be related to ξ XI but certainly was /sk/ and not /ks/—compare the Greek loanwords zcvvlvq'z *skandhalis* 'to stumble' ($\sigma \kappa \alpha \nu \delta \alpha \lambda \eta \zeta$) and Lulv'81 *dhidháski* 'teaching' ($\delta \iota \delta \alpha \kappa \alpha \lambda o \zeta$). Note that despite the fact that the printed source charts (Figures 1, 2, and 3) show the OO ω with a horizontal bar above ($\bar{\omega}$), a variety of shapes are found in the actual manuscript, including as mentioned above. The form without any diacritic does occur more than once in the manuscripts, so it is recommended that this be the form used in the code charts. The phrase *o zot* 'O God' can be seen as $\hat{\omega}$ Vól at the beginning of line 15 of Figures 4 and 5; it is given as ω Vól \bar{o} zót in Figure 6 and as o Vól *o zót* without an omega in Figure 8.

3.1.3 Group 4. The 4 letters here are a little less clear in their derivation. \mathfrak{V} AS might simply be $\mathfrak{v} + \mathfrak{e}$; \mathfrak{I} HJ is a turned \mathfrak{E} H (von Hahn gives X "dark ch" [x] and \mathfrak{I} "light ch" [ç]); \mathfrak{A} SHT is clearly derived from \mathfrak{A} SH + \mathfrak{I} T and is attested many times throughout the texts; and \mathfrak{H} JY (Elsie writes *ju/jy*; $\mathfrak{e} \mathfrak{H} \mathfrak{I}$ *sjýt* appears on p. 29 of the MS) would appear to be $\iota \mathfrak{I} + \mathfrak{d}$ Y, and is common enough in the texts. It is impossible to talk about "typographic ligatures" for Todhri as there was no traditional typography.

Note that Elsie's chart omits von Hahn's \ni HJ, though this may be an oversight (modifying ϵ [h] to \ni [ç] is logical). Elsie's chart also includes \Downarrow JY which von Hahn does not give, but von Hahn gives a letter ϑ *të* which is also clearly ϑ T + ϑ EH but is similar to that in shape (see 3.2 below). The text which contains \land SHT usually writes *të* as ϑ (not ϑ); see Figures 8 and 9. As to whether SHTA should be encoded on its own, the recommendation is that it should be. Certainly both \land SHA and ϑ TA both exist; I have not yet seen an instance of a sequence " $\land \vartheta$ SHT" though the "ligature" is very common (*e.g.*, $\epsilon_{\Im} \land \vartheta$ *hrishtë*, $\epsilon_{\Box} \land \vartheta$ *krishtë* 'Christ'). Compare Glagolitic which has III SHA, ID TE, and " SHTA (also clearly a ligature in origin). Users of Todhri are scholars, and they can deal with searching issues revolving around $\land \vartheta$ and \land , if any, given the relatively small size of the corpus.

3.2 Ligature $t\ddot{e}$. A ligature Ψ $t\ddot{e}$ is given in von Hahn's chart. Perhaps von Hahn saw an example of a logograph for $t\ddot{e}$ 'of'. It has not been seen in any of the manuscripts I have access to, where the unligated form is common. Fonts may choose to include an optional ligature for Ψ $t\ddot{e}$.

4. Character names. Letter-names end in -A following von Hahn 1854; two of von Hahn's letters have -AN and this form has been retained. The spellings used for the characters here are based on those of the modern Albanian alphabet, with c written CH, e written E, and \ddot{e} written EH. Here are the correspondences:

Todhri	UCS	von Hahn	Todhri	UCS	von Hahn	Todhri	UCS	von Hahn	Todhri	UCS	von Hahn
v	А	а	К	NGA	ngha	v	NA	na	A	SHTA	schta
v	AS	as	b	GJA	gja	v	NJAN	njan	q	TA	ta
M	BA	ba	К	NGJA	ngja	0	0	0	2	THA	θa
ß	MBA	mba	v	GHA	γa	w.	00	ω	ò	U	и
۹.	CA	tsa	e	HA	ha	И	PA	ра	e	VA	wa
q	CHA (ÇA)	tscha	э	HJA	chja	4	PSA	psa	7	XA	dsa
Ň	DA	da	I	Ι	i	%	QA	kja	Z	NXA	ndsa
xx	NDA	nda)	JA	ja (deutsch. j)	5	RA	ra	g	XHA	dscha
L	DHA	δa	C	KA	ka	ŝ	RRA	rra	ġ	NXHA	ndscha
i	Е	e (offen e)	Х	KHA	cha	s	SA	sa	å,	Y	ü
÷	EH (Ë)	e (stumm e)	þ	LA	lja	8	SKAN	xan	y,	JY	*jü
6	FA	fa	Ħ	LLA	la	5	STA	sta	V	ZA	ζα
h	GA	gha	ç	MA	ma	x	SHA	scha	8	ZHA	ja (franz. j)

5. Numerals and punctuation. One supposes that European digits and punctuation are to be used with Todhri; COMMA and FULL STOP have been seen in the manuscripts.

6. Ordering. As for the Vithkuqi script, it is likely that modern Albanians will wish the alphabetical order for Todhri to be similar to that of the Latin Albanian alphabet. (The order given by von Hahn is not particularly practical for modern users.) This order has been followed in the code chart, with prenasalized consonants following their bases and other characters inserted where appropriate.

7. Unicode Character Properties

/. Unicouc	Chara	ter i roperties
105C0;TODHRI	LETTER	A;Ll;0;L;;;;;N;;;;;
105C1;TODHRI	LETTER	AS;Ll;0;L;;;;;N;;;;;
105C2;TODHRI	LETTER	BA;Ll;0;L;;;;;N;;;;;
105C3;TODHRI	LETTER	MBA;Ll;0;L;;;;;N;;;;;
105C4;TODHRI	LETTER	CA;Ll;0;L;;;;;N;;;;;
105C5;TODHRI	LETTER	CHA;Ll;0;L;;;;;N;;;;;
105C6;TODHRI	LETTER	DA;Ll;0;L;;;;;N;;;;;
105C7;TODHRI	LETTER	NDA;Ll;0;L;;;;;N;;;;;
105C8;TODHRI	LETTER	DHA;Ll;0;L;;;;;N;;;;;
105C9;TODHRI	LETTER	E;Ll;0;L;;;;;N;;;;;
105CA;TODHRI	LETTER	EH;Ll;0;L;;;;;N;;;;;
105CB;TODHRI	LETTER	FA;Ll;0;L;;;;;N;;;;;
105CC;TODHRI	LETTER	GA;Ll;0;L;;;;;N;;;;;
105CD;TODHRI	LETTER	NGA;Ll;0;L;;;;;N;;;;;
105CE;TODHRI	LETTER	GJA;Ll;0;L;;;;;N;;;;;
105CF;TODHRI	LETTER	NGJA;L1;0;L;;;;;N;;;;;
105D0;TODHRI	LETTER	GHA;Ll;0;L;;;;;N;;;;;
105D1;TODHRI	LETTER	HA;Ll;0;L;;;;;N;;;;;
105D2;TODHRI	LETTER	HJA;Ll;0;L;;;;;N;;;;;
105D3;TODHRI	LETTER	I;Ll;0;L;;;;;N;;;;;
105D4;TODHRI	LETTER	JA;Ll;0;L;;;;;N;;;;;
105D5;TODHRI	LETTER	KA;Ll;0;L;;;;;N;;;;;
105D6;TODHRI	LETTER	KHA;Ll;0;L;;;;;N;;;;;
105D7;TODHRI	LETTER	LA;Ll;0;L;;;;;N;;;;;
105D8;TODHRI	LETTER	LLA;Ll;0;L;;;;;N;;;;;
105D9;TODHRI	LETTER	MA;Ll;0;L;;;;;N;;;;;
105DA;TODHRI	LETTER	NA;Ll;0;L;;;;;N;;;;;
105DB;TODHRI	LETTER	NJAN;L1;0;L;;;;;N;;;;;
105DC;TODHRI	LETTER	O;Ll;0;L;;;;;N;;;;;
105DD;TODHRI	LETTER	00;Ll;0;L;;;;;N;;;;;
105DE;TODHRI	LETTER	PA;Ll;0;L;;;;;N;;;;;
105DF;TODHRI	LETTER	PSA;Ll;0;L;;;;;N;;;;;
105E0;TODHRI	LETTER	QA;Ll;0;L;;;;;N;;;;;
105E1;TODHRI	LETTER	RA;L1;0;L;;;;;N;;;;;
105E2;TODHRI	LETTER	RRA;L1;0;L;;;;;N;;;;;
105E3;TODHRI	LETTER	SA;L1;0;L;;;;;N;;;;;
105E4;TODHRI	LETTER	SKAN;Ll;0;L;;;;;N;;;;;
105E5;TODHRI	LETTER	STA;L1;0;L;;;;;N;;;;;
105E6;TODHRI	LETTER	SHA;L1;0;L;;;;;N;;;;;
105E7;TODHRI	LETTER	SHTA;L1;0;L;;;;;N;;;;;
105E8;TODHRI	LETTER	TA;Ll;0;L;;;;;N;;;;;
105E9;TODHRI	LETTER	THA;Ll;0;L;;;;;N;;;;;
105EA; TODHRI	LETTER	U;Ll;0;L;;;;;N;;;;;
105EB;TODHRI	LETTER	VA;L1;0;L;;;;;N;;;;;
105EC;TODHRI	LETTER	XA;Ll;0;L;;;;;N;;;;;
105ED;TODHRI	LETTER	NXA;Ll;0;L;;;;;N;;;;;
105EE;TODHRI	LETTER	XHA;Ll;0;L;;;;N;;;;;
105EF; TODHRI	LETTER	NXHA;Ll;0;L;;;;;N;;;;;
105F0;TODHRI	LETTER	Y;Ll;0;L;;;;N;;;;;
105F1;TODHRI	LETTER	JY;L1;0;L;;;;;N;;;;;
105F2;TODHRI	LETTER	ZA;L1;0;L;;;;;N;;;;;
105F3;TODHRI	LETTER	<pre>ZHA;L1;0;L;;;;;N;;;;;</pre>

8. Bibliography

Diringer, David. 1949. The Alphabet.

Elsie, Robert. 2016. *Early Albanian Bible translations in Todhri Script*. (Albanian Studies; 30) London: Centre for Albanian Studies. ISBN 978-1-53970333-4.

Elsie, Robert. 2017. "The alphabet of Dhaskal Todhri", in *Albanian Alphabets: Borrowed and invented*. (Albanian Studies; 35) London: Centre for Albanian Studies. ISBN 978-1-54429409-4.

Faulmann, Karl. 1880. Das Buch der Schrift. Wien.

Skendi, Stavro. 1960. "The history of the Albanian alphabet: a case of complex cultural and political development", in *Südost-Forschungen: Internationale Zeitschrift für Geschichte, Kultur und Landes-kunde Südosteuropas* 19:263-284.

Trix, Frances. 1997. "Alphabet conflict in the Balkans: Albanian and the congress of Monastir", in *International Journal of the Sociology of Language* 128:1-23.

Десницкая, Агния Васильевна. 1968. *Албанский язык и его диалекты*. Leningrad: Nauka. Десницкая, Агния Васильевна. 1987. *Албанская литература и албанский язык*. Leningrad: Nauka.

9. Acknowledgements. This project was made possible in part by a grant from the U.S. National Endowment for the Humanities, which funded the Universal Scripts Project (part of the Script Encoding Initiative at UC Berkeley) in respect of the Todhri encoding. Any views, findings, conclusions or recommendations expressed in this publication do not necessarily reflect those of the National Endowment of the Humanities.

Figures.

Nr.	Zeichen.	Benen- nung.	Werth.	Nr.	Zeichen.	Benen- nung.	Werth.	Nr.	Zeichen.	Benen- nung.	Werth.
1.	۷	A	a	19.	5	Ra	r	37.	3	θa	priech 13
2.	i	E	stes C	20.	r	Rra	rr	38.	M	Ba	Ь
3.	1	I	i	21.	в	Fa	f	39.	B	Мва	mb
4.	0	0	0	22.	64	∆a	griech o	40.	Ы	Pa	Р
5.	0	U	u	23.	5	Ma	m	41.	v	Na	n
6.	6	Ũ	ű	24.)	Ja	dentsch. j	42.	9	Tscha	tsch
7.))	E	siumm O	25.	h	Gha	gh	43.	g	Dscha	dsch
8.	s	Sa	S	26.	K	Ngha	ngh	44.	g	Ndscha	ndsch
9.	8	ζα	griech 5	27.	5	Gja	gj	45.	5	Sta	st
10.	Í	Tsa	ts	28.	5	Ngja	ngj	46.	X	Scha	sch
11.	7	Dsa	ds	29.	V	Γα	griech. Y	47.	\$	Ja	tun j
12.	Z	Ndsa	nds	30.	ч	Psa	ps	4 8.	R	Schta	scht
13.	l	Wa	w	31.	٤	Ha	h	49.	Ŷ	Te	te
14.	H	La	l	32.	X	Cha	dunkei Ch	50.	¥	Njan	nj
15.	þ	Lja	lj	33.	3	Chja	ben ch	51.	*	As	25
16.	8	Kja	kj	34.	99	Ta	t	52.	Ū	Q	griech (J)
17.	¢	Ka	k	35.	۸	Da	a				
18.	8	Xan	x	·36.	x	Nda	nd				

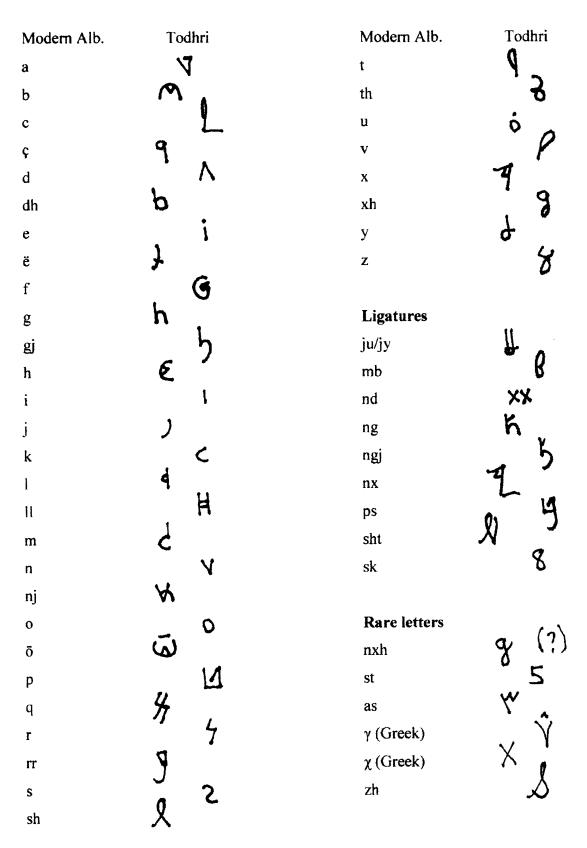
§. 1. Das Alphabet.

Figure 1. Todhri characters as given in von Hahn 1854. See §4 above.

		TH	IE P	LB	ANIAN	ALF	PHAE	BET	S.		
Values.	Büthakukye.	Elbasan.	Probable Greek Prototypes.	Values.	Büthak ukye	Elbasan.	Probable Greek Prototypes.	Values.	Büthakukye.	Elbasan.	Probable Greek Prototypes.
a	(S a	۷	a	n	՝ ՝	v	ν	v	Цų	ℓ	β
b	ર દ દ	M	β	x		8	Ę	d	tı tı	٨	δ
$\left \begin{array}{c} g \end{array} \right $	సి సి	V	γ	0	C o	0	o	y)	E
dh	h h	٨Р	δ	р	3 2	Ы	യ	$\begin{vmatrix} v \\ y \end{vmatrix}$	fin		η
е	K e))	£	r	6	5	ρ				.,
z	22	V	ζ	8	υυ	s	σ	88	ឞ៑		σ
ē	N n	2	η	t	4 f	9 9	τ	u	γ. μ		υ
$\left \begin{array}{c} th \\ i \end{array} \right $	l l :	\$	θ	ü	Ϋ́ Ϋ́	4	υ	gh		b	к
$\begin{bmatrix} \iota \\ k \end{bmatrix}$	š i		ι	f	પર ∾	G	$\boldsymbol{\phi}$	e		i	L
$\begin{vmatrix} n \\ l \end{vmatrix}$	& a V 7 V s	¢ ∦	κ λ	kh	8 8	x	x	u		Ō	o
$\begin{bmatrix} v \\ m \end{bmatrix}$	v 7 7 7 5 5	н . ८	π μ	ō	// /	Ū	л w	rr		A	ρ
			$\frac{r}{\mathrm{IV.}}$		11.	III.	IV.	<u> </u>	II.		
				Т	HE LIGAT	URES					
			Elbas	an.					Büthaku	kye.	
as	v =	a+		dsh	9 =	$\theta + \sigma$	σ	ky	Cé ei =	κ+u	
ps	<u>y</u> =	π+ σ	1	ıdsh	<i>Y</i> =	$\nu + \theta e$	σσ				
mb	θ =	$\mu + \beta$		nd		ν+δ		ks	efc of =	κ+a	r
st	5 =	$\sigma + \tau$		ds		δ+σ		dz	(O) =	δ+0	r
sh	Å =	σ+σ		nds ,		v+80	7				
shy		σσ+		ngh		ν+κ		ds	8 5 =	$\theta + \phi$	r
sht	0	σσ+΄	-	gy		κ+ε		+	& & =	_ 1	
te	(1)	$\tau + \epsilon$		ngy		v+ке	E	tzy	ω ω =	τ+0	+1
	a	•		ly In:	•	λ+ε		ng	શ n =	η+1	,
ts	ι=	$\tau + \sigma$		ky	8 =	κ+ι	and the second				

Figure 2. Todhri ("Elbasan") and Vithkuqi ("Büthakukye") characters as given in Taylor 1883. Taylor's "ligatures" seem to be his (not entirely successful) attempts to show the derivation of the letters: $v as (v a + \varepsilon s), \forall ps (\sqcup p + \varepsilon s), \& mb (d m + \cap b), 5 st (\varepsilon s + \P t$ —correctly Greek ζ as derived from ζ), $\lambda sh (\varepsilon s + \varepsilon s), \& zh (\lambda sh + \exists e), \& sht (\lambda sh + \P t), & \forall te (\P t + \exists e), \P c (\P t + \varepsilon s), g xh (q d + \lambda sh), g nxh (v n + q xh), w nd (v n + \wedge d), \forall x (\wedge d + \varepsilon s), \forall nx (v n + \forall x), h ng (v n + c k$ —correctly v n + h g), $h gj (c k + \exists e), H c (H + \exists e), H c (H + \exists e), H c (k + 1)$.

THE ALBANIAN ALPHABETS.



The Todhri Alphabet

Figure 3. Todhri alphabet chart from Elsie 2016:155.

1 Whis In white Drink of iter iter I bill 2 Billviv soli i and alloviv soli, 21 Dyin wi Bil bill god 3 1027 9242 nigeiv, invo vii 201, xi xri vii Goril love zi 4 wive wiise Explose love i dir ci diverse de Marine 5 for no sorre vii usis Dig veira dios 250 6 ei solis of Brilling with signed wind ingist wi 7 i huish litel vadi, icisno, xidil? Prileil. Robert Staroor 8 60 dr 1251. 12-9 Sale spostovisoury 10 Papilici, il o Guaida, Brill unstaine low? 11 Deplici in Graice Gill Dister vis 2000 for 12 Papilies ile Graide vin Gille Brile unstand low?. 13 as a yas To 14 Gourselil, Gost Prile sind covier ilio usis 15 CARDizichove ANTREHIVISTOPPINATI ANIGHIC-SANDEN ICI 16 xi Abli it obloov 22015 Sain ibra?. isigobo Sil 17 Aishi Ma Cili 12/4, xi loslott sobor ilu ivoc conci 18 Bisor wost il. Loyic evil ilon of Bigirlif og 19 ic ministry is sind. Nocar wir all gi connew 20 Not Perspirazet. plice bily Visileil 21

Figure 4. The Trisagion Prayers; The Lord's Prayer; Psalms 84: 1-4 from Elsie 2016:98. Circled is ŵ Vớl ō zót 'O God'' (in green). In the Greek heading (in red) above Psalm 84 reads ὡς ἀγαπιτὰ (evidently for ὡς ἀγαπητὰ 'as dear'); the scribe writes both rough and smooth breathings identically. In purple are two instances of *kríshtit* showing the letter SHTA; in cyan are two instances of *jánë* showing written as ν̄ψν + jānë and one instance of *prūūm* written as Δόόc *prûûm*.

01. [...] 02. babáj ýnë qí jée më qíejtë u shenjtënóftë émëni ýt, árthtë 03. mbretënía jóte, e u báftë dashúnía jóte, sí më qíejtë nde mbíi dhéet, búkën 04. tónë të përdíçen, épna née sót, nde ndjé née fájet tóna, sí 05. nde ná ndjéejmë fajtórët tónë e mós kíi me na prúum më të kandá-06. tun, pó rújna née prej cë kéqi, amín [...] 07. se jóteja áshë mbretënía nde fuqía nde láfta e átit e e bírit nde 08. i shpírtit shénjt njimé e kúrdó, nde më jétë të jétvet [...] 09. zót mallëngjé 10. [...] éni të lútemi e ti fálemi, mbrétit perëndíis tónë, 11. 12. éni të lútemi e ti fálemi, Kríshtit mbrétit perëndíis tónë 13. éni të lútemi e ti fálemi, atíi Kríshtit mbrétit perëndíis tónë. 14. [...] 15. o zót) fuqíavet, fort te dáshme jánë konákët e túu prej 16. mállit e sevdáas ndë avllíat zótit gjíndet shpíirti ím zémbra íme 17. nde shtáti ím, u gëzúanë më perëndiinë e gjállë e se zógu ngjít 18. çérdhe për vétet të tíi, nde túrtulli zógun e tíi e núk káa me 19. i thérun zóqt e tíi, thugím hánet e túa, zót í fuqíavet, mbréti ím nde perëndía íri jánë të lúmun jánë tá qí kondísin 20. 21. më shtëpii táte, më jétat të jétvet dot të vargëzójnë.

Figure 5a. Transcription of Figure 4 above from Elsie 2016:99.

- 1 Πάτερ ήμῶν
- 2. $mm'_{2} \overline{J}_{1} \overline{J}_{2} \gamma_{1} \gamma_{2} \gamma_{1} \gamma_{2} \gamma_{1} \gamma_{2} \gamma_{2} \gamma_{1} \gamma_{2} \gamma_{2} \gamma_{1} \gamma_{2} \gamma_{2}$
- 3. lsi92viv sóli, i o mý G92 Avloviv sóli, zí d2 gist2 xi li Lii, mó <math>2v
- 5. \sqrt{x} xy \sqrt{y} \sqrt{y} i \sqrt{y} i \sqrt{y} i \sqrt{y} ix \sqrt{y} i
- 6. $1 \circ v, \Box o \circ s \circ v v v i \Box u s i v 1 + i s v 1 + i s i v 1 + i s i v 1 + i s i v 1 + i s v$
- 7. ri $\sqrt{3}$ ri
- 8. Ι λμίνι λ_i νι ci i cosΛό, xi c? σίη? η? σίνηι [κύριε ἐλέισον (for ἐλέησον)]
- 9. Vól dvy-Kí
- 10. [Greek]
- 11. ivi 47 ajici i 4 Gvaici, 6sique istante i 4 often i
- 12. ivi 47 ao4ici i 4 Gvaici, Gran esita uistavire 40v7
- 13. ivi 97 aólici i 9 Graici, voli csi And Bsi 9 dis?xi 12 dov?.
- 14. [ώς ἀγαπητὰ]
- 15. $\hat{\omega}$ Với i Gözívlii, Gósi i Avzci $\bar{v}v$? covýc? i $\hat{v}o$ dsi
- 17. xi $\Re \sqrt{4}$ i (c, o h) $\sqrt{6}$ xv) c) \square is $\Re \sqrt{4}$ i zi $\sqrt{6}$ ho $\Im \sqrt{6}$
- 18. qīski \square γs (í) 97 (í), xi (όs) ομι V όhov i (í) i voc $v\overline{v}v$ ci
- 19. $13\overline{1}$ isov $\sqrt[3]{6}$ i 1/1, $3\overline{6}$ i $26\overline{1}$ i $10\overline{6}$, $\sqrt[3]{6}$ i $3\overline{6}$

Figure 5b. Acute-accented vowels are marked in magenta; vowels with macron are marked in blue.

Nosig dian weasin los entione abli nosig los 8 1 Wizz soli ciinillor Wikis I've icles, der Purste 2 và usento sicre esizor raiszair pistair. 3 higidi Bulocia sation whe remainer the # 19 wash 4 she ivers? ilili sici dir 14109 isi 129 242 11 5 Toriasikin doll rosiyisiis? waring my por 6 Badethe inder as port in Point de ind gen 7 Poeir woll si ví lig ligió ilsi 924 Hoy vác busé en 8 outofelpis IP & XX 255 orchow Ector 2% AI HUgo 2 9 143 425 Visixcior als son 0.89 88058 10 con dichi go ve 110091 225092 - x150 wish 11 Stanto via site star lic infile AIH 12 2292 do fior 2/14 900 1 118 200 10 Lover 13 Gel SOI crait 14 Hidobing Gosto 60 di euste gi alle 15 x to vel visio disi 12 low work 16 Sister when is all his first som his soin 17 ai provi xxi 22 sist. il use do need 20 18 Zi ve wigo all had hog poet dist 19 2312 at all cit? ci 12 24 shear do st? 1484 20 VI visite no Gh' VO HI VOL CE VOSOCI VIDOD 21 21. Englines comos com dim ut al 22

Figure 6. Hebrews 1:8–12; 2 Timothy 2:1–6 from Elsie 2016:30. Circled in red are $\cup \forall \diamond | \bar{o} z \delta t$ 'O God', $\cup \ell \neq | \psi \rangle \vee \bar{o} v \bar{e} l l \dot{a} z \bar{e} n$ 'O brothers', and $\bar{\omega} \cap \omega \dot{o} \delta l \bar{i} r i$ 'O my son'; in green di elis $\psi \neq \omega \ell = h i r t \bar{e} q i$ ashtë 'in the grace that is' with TA followed by EH (ψ) and SHTA followed by EH (λ ?). Circled in cyan is $\nu \psi \neq j a n \bar{e}$ 'they are' and circled in magenta is Lub $\psi \vee v d h d h a z h$ 'teaching' (recte Lub $\psi \otimes v d h d h a s h r h$?).

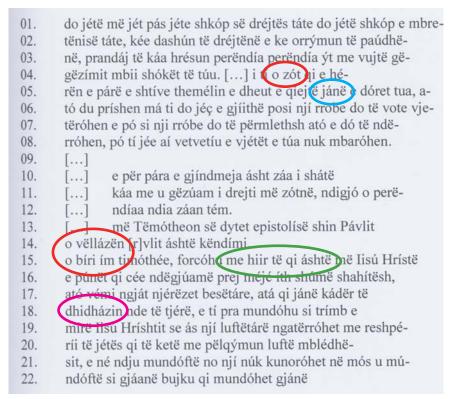


Figure 7. Transcription of figure 6 above from Elsie 2016:31. In line 18 Elsie transcribed אולעיע as *dhidházin* but this may be a scribal error for *dhidháskin* (not אולטע).

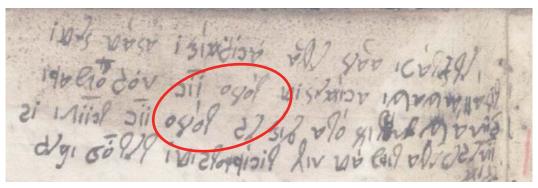


Figure 8. Hebrews 13:7–16 from Elsie 2016:20. Circled are two instances of $\circ \sqrt[5]{6} \circ z \circ t$ 'O God' with an omicron rather than an omega.

01.	e për páara e gjíndmeja ashtë záa i katërtë.
02.	e lavdómun je o zot perëndíja i baballáret
03.	se i d[r]éejt jee o zót n'ë gjith ató qi bán më nee.
04.	më çifútët epistolijet shin Pávlit ashtë këndimi:

Figure 9. Transcription of figure 8 above from Elsie 2016:21.

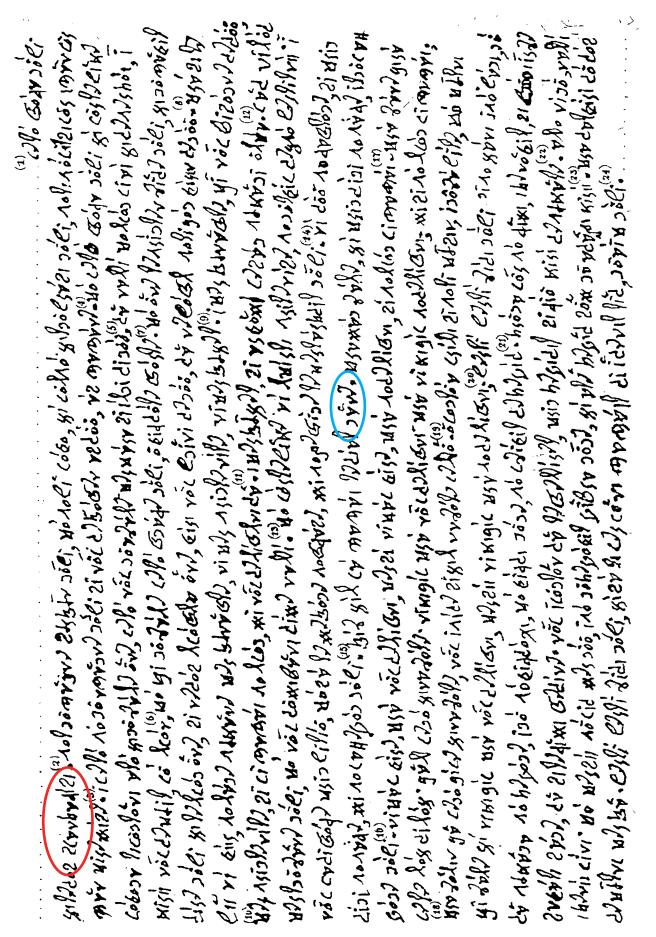


Figure 10. The text of John 16:1–23 transcribed by von Hahn, from Elsie 2016:104. Circled is the Greek loanword ecvvLvqíei skandalísi 'to stumble'. Also circled is ⊃vv-j jânë which might have a sort of breathing over it.

1.	këtó fóla júve
2.	qi të mos skandalisi, qo tju bájnë syrgjýn júve, po do ví kóho, qí
	kushdó qi uju vrási juve, do ti dúket sikur i bán kur-
3.	bán perëndísë, e këtó do jua bájnë júve se nuk më ngjófën ás múu, as
	babánë, pó këtó fóla júve qí kur të vínjë
4.	kóhoja ti kujtóni ató qi ju tháshë únë, këtó núk jua tháshë për pára se
	íshje me júu, ma nashtí po shkój ke aí qi më dërgój, e
5.	njeríi nuk më pyet, kú shkon, po pse ju tháshë këto fjálë júve, u
	helmútë fórtë, pó únë të dréjtën thémë júve, qi ju báhet
6.	mírë júve qi të shkój únë, se në mós shkófsha únë, híri nuk vjéne më
	júu, má në vofsh do te çój hírin më júu, pra si të
7.	víi aí híir, do shájë dynjánë për gjynáftë, ne për drejtënítë, ne për
	gjýqët, e për gjynáftë, psé núk mbesójnë më múu,
8.	për drejtëníte, se ke babái do shkój, nde núk më shífshni ma, e për
	gjýqët, se arhóndi kësáj dynjáje u sháa, kám ne shúm
9.	për tju thánë júve, po núk mundi mbáni méndë nashtí, pó kúr të vínjë
	aí shpírti drejtënësë, do ju t hék më çdó vërtetníi e
10.	núk ka mé fólë prej vetíu, pó sá të ndëgjojë do flásë nde do rrëféjë të
	për t árdhmet júve, aí múu do laftójë, se prej
11.	méje do dálë, nde do kallëzój júve, gjíth qísh ká babái të míat janë,
	prandáj tháshë qí prej méje do dálë e tju kall-
12.	ëzójë júve, ne pák hérë, prá núk më shífni, përsí ne pák hérë, prá do
	më shífni, se do shkój ke babái, prá thánë çirá-
13.	këtë shóq me shóq ç asht këjó qi na thótë, ne nji çík prá nuk më shífni,
	prá ne nji çík do më shífni, nde se dó shkój ke babái,
14.	pra thóshin ça[shtë] këjó çikë që na thótë, nuk e dimë se qish na thótë
	kështú. u kujtúa Kríshti se do te pýsin e ju tha vetë, po pýtni
15.	pse tháshë qí ne nji çík prá núk më shífni, përsíi ne nji çík prá do më
	shifni; vërté, vërté thémi júve, se do qáni e dó váji jú,
16.	má dynjája dó gëzójë, e jú dó helmóchi, pó hélmi jújë dó kthéhet më
	gëzím, grúja kúr dó líndi, idhënóhet si i érthë
17.	saháti sajë, ma si të líndi fëmínë, núk e kujtón ma të fështirët, prej
	gëzímit se léu njerí më dynjátë, ashtu ne ju, nashtí

18. idhëníi kéni, po përsíi dukem ndër júu, e dó ju gëzóhet zémbra jújë, qí atë gëzím súnd ju a márri njerií, pra m at hére mú mós

Figure 11. Transcription of figure 10 above from Elsie 2016:105.

105C0

Todhri

	105C	105D	105E	105F
0	V 105C0	1 05D0	% 105E0	Ј 105F0
1	V≁ 105C1	E 105D1	5 105E1	L 105F1
2	\bowtie	3	ſ	8
3	105C2	105D2	105E2 2	105F2
4	105C3	105D3	105E3	105F3
5	105C4 G 105C5	105D4	^{105E4}	
6	105C5	105D5	105E5	
7	105C6	105D6	105E6	
8	105C7	105D7	105E7	
9	105C8	105D8	105E8 ?	
A	105C9	105D9	105E9	
	105CA	105DA	105EA	
В	о 105СВ h	105DB	105EB	
С	105CC	0 105DC	105EC	
D	К 105CD	105DD	Z 105ED	
E	b 105CE	105DE	9 105EE	
F	Б 105СF	y 105DF	Y 105EF	

105C0	٧	TODHRI LETTER A
105C1	v -	TODHRI LETTER AS
105C2	M	TODHRI LETTER BA
105C3	b	TODHRI LETTER MBA
105C4	L	TODHRI LETTER CA
105C5	q	TODHRI LETTER CHA
105C6	Ň	TODHRI LETTER DA
105C7		TODHRI LETTER NDA
	xx	
105C8	Ŀ	TODHRI LETTER DHA
105C9	i	TODHRI LETTER E
105CA	2	TODHRI LETTER EH
105CB	G	TODHRI LETTER FA
105CC		
	h	TODHRI LETTER GA
105CD	ĥ	TODHRI LETTER NGA
105CE	Ь	TODHRI LETTER GJA
105CF	Ĕ	TODHRI LETTER NGJA
105D0	Ŵ	TODHRI LETTER GHA
10300	v	
		\rightarrow 03B3 γ greek small letter gamma
105D1	e	TODHRI LETTER HA
105D2	3	TODHRI LETTER HJA
105D3	I.	TODHRI LETTER I
105D4)	TODHRI LETTER JA
105D5	C	TODHRI LETTER KA
105D6	Х	TODHRI LETTER KHA
10000		
		\rightarrow 03C7 χ greek small letter chi
105D7	þ	TODHRI LETTER LA
105D8	Ĥ	TODHRI LETTER LLA
105D9	ď	TODHRI LETTER MA
105DA	v	TODHRI LETTER NA
105DB	r	TODHRI LETTER NJAN
105DC	0	TODHRI LETTER O
105DD		
105DD	W	TODHRI LETTER OO
		TODHRI LETTER OO $\rightarrow 03C9 \omega$ greek small letter omega
105DE		TODHRI LETTER OO \rightarrow 03C9 ω greek small letter omega TODHRI LETTER PA
105DE	⊌ ⊌	TODHRI LETTER OO \rightarrow 03C9 ω greek small letter omega TODHRI LETTER PA
	W	TODHRI LETTER OO \rightarrow 03C9 ω greek small letter omega TODHRI LETTER PA TODHRI LETTER PSA
105DE 105DF	ы М М	TODHRI LETTER OO \rightarrow 03C9 ω greek small letter omega TODHRI LETTER PA TODHRI LETTER PSA \rightarrow 03C8 ψ greek small letter psi
105DE 105DF 105E0	ы М М	TODHRI LETTER OO \rightarrow 03C9 ω greek small letter omega TODHRI LETTER PA TODHRI LETTER PSA \rightarrow 03C8 ψ greek small letter psi TODHRI LETTER QA
105DE 105DF	⊌ ⊌	TODHRI LETTER OO \rightarrow 03C9 ω greek small letter omega TODHRI LETTER PA TODHRI LETTER PSA \rightarrow 03C8 ψ greek small letter psi
105DE 105DF 105E0 105E1	⊌ ⊌ %	TODHRI LETTER OO \rightarrow 03C9 ω greek small letter omega TODHRI LETTER PA TODHRI LETTER PSA \rightarrow 03C8 ψ greek small letter psi TODHRI LETTER QA TODHRI LETTER RA
105DE 105DF 105E0 105E1 105E2	ы М М М М	TODHRI LETTER OO \rightarrow 03C9 ω greek small letter omega TODHRI LETTER PA TODHRI LETTER PSA \rightarrow 03C8 ψ greek small letter psi TODHRI LETTER QA TODHRI LETTER RA TODHRI LETTER RA
105DE 105DF 105E0 105E1 105E2 105E3	ы м м м м м м м м	TODHRI LETTER OO \rightarrow 03C9 ω greek small letter omega TODHRI LETTER PA TODHRI LETTER PSA \rightarrow 03C8 ψ greek small letter psi TODHRI LETTER QA TODHRI LETTER RA TODHRI LETTER RRA TODHRI LETTER SA
105DE 105DF 105E0 105E1 105E2	ы М М М М	TODHRI LETTER OO \rightarrow 03C9 ω greek small letter omega TODHRI LETTER PA TODHRI LETTER PSA \rightarrow 03C8 ψ greek small letter psi TODHRI LETTER QA TODHRI LETTER RA TODHRI LETTER RRA TODHRI LETTER SA TODHRI LETTER SKAN
105DE 105DF 105E0 105E1 105E2 105E3	ы м м м м м м м м	TODHRI LETTER OO \rightarrow 03C9 ω greek small letter omega TODHRI LETTER PA TODHRI LETTER PSA \rightarrow 03C8 ψ greek small letter psi TODHRI LETTER QA TODHRI LETTER RA TODHRI LETTER RRA TODHRI LETTER SA TODHRI LETTER SKAN
105DE 105DF 105E0 105E1 105E2 105E3 105E4	し い い い い い い い い い い い い い い い い い い い	TODHRI LETTER OO \rightarrow 03C9 ω greek small letter omega TODHRI LETTER PA TODHRI LETTER PSA \rightarrow 03C8 ψ greek small letter psi TODHRI LETTER QA TODHRI LETTER RA TODHRI LETTER RAA TODHRI LETTER SA TODHRI LETTER SKAN \rightarrow 03BE ξ greek small letter ksi
105DE 105DF 105E0 105E1 105E2 105E3	ы м м м м м м м м	TODHRI LETTER OO \rightarrow 03C9 ω greek small letter omega TODHRI LETTER PA TODHRI LETTER PSA \rightarrow 03C8 ψ greek small letter psi TODHRI LETTER QA TODHRI LETTER RA TODHRI LETTER RAA TODHRI LETTER SA TODHRI LETTER SKAN \rightarrow 03BE ξ greek small letter ksi TODHRI LETTER STA
105DE 105DF 105E0 105E1 105E2 105E3 105E4 105E5	₩ ₩ ₩ ₩ ₩ ₩ ₩ ₩ ₩	TODHRI LETTER OO \rightarrow 03C9 ω greek small letter omega TODHRI LETTER PA TODHRI LETTER PSA \rightarrow 03C8 ψ greek small letter psi TODHRI LETTER QA TODHRI LETTER RA TODHRI LETTER RAA TODHRI LETTER SAA TODHRI LETTER SKAN \rightarrow 03BE ξ greek small letter ksi TODHRI LETTER STA \rightarrow 03DB ζ greek small letter stigma
105DE 105DF 105E0 105E1 105E2 105E3 105E4	し い い い い い い い い い い い い い い い い い い い	TODHRI LETTER OO \rightarrow 03C9 ω greek small letter omega TODHRI LETTER PA TODHRI LETTER PSA \rightarrow 03C8 ψ greek small letter psi TODHRI LETTER QA TODHRI LETTER RA TODHRI LETTER RAA TODHRI LETTER SA TODHRI LETTER SKAN \rightarrow 03BE ξ greek small letter ksi TODHRI LETTER STA
105DE 105DF 105E0 105E1 105E2 105E3 105E4 105E5	₩ ₩ ₩ ₩ ₩ ₩ ₩ ₩ ₩	TODHRI LETTER OO \rightarrow 03C9 ω greek small letter omega TODHRI LETTER PA TODHRI LETTER PSA \rightarrow 03C8 ψ greek small letter psi TODHRI LETTER QA TODHRI LETTER RA TODHRI LETTER RAA TODHRI LETTER SAA TODHRI LETTER SKAN \rightarrow 03BE ξ greek small letter ksi TODHRI LETTER STA \rightarrow 03DB ζ greek small letter stigma
105DE 105DF 105E0 105E1 105E2 105E3 105E4 105E5 105E6	₩ Ч Ч У У Р г 8 5 Х	TODHRI LETTER OO \rightarrow 03C9 ω greek small letter omega TODHRI LETTER PA TODHRI LETTER PSA \rightarrow 03C8 ψ greek small letter psi TODHRI LETTER QA TODHRI LETTER RA TODHRI LETTER SA TODHRI LETTER SKAN \rightarrow 03BE ξ greek small letter ksi TODHRI LETTER STA \rightarrow 03DB ζ greek small letter stigma TODHRI LETTER SHA TODHRI LETTER SHA
105DE 105DF 105E0 105E1 105E2 105E3 105E4 105E5 105E6 105E7	₩ 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	TODHRI LETTER OO \rightarrow 03C9 ω greek small letter omega TODHRI LETTER PA TODHRI LETTER PSA \rightarrow 03C8 ψ greek small letter psi TODHRI LETTER QA TODHRI LETTER RA TODHRI LETTER RAA TODHRI LETTER SAA TODHRI LETTER SKAN \rightarrow 03BE ξ greek small letter ksi TODHRI LETTER STA \rightarrow 03DB ζ greek small letter stigma TODHRI LETTER SHA TODHRI LETTER SHA TODHRI LETTER SHA
105DE 105DF 105E0 105E1 105E2 105E3 105E4 105E5 105E6 105E7 105E8	₩	TODHRI LETTER OO \rightarrow 03C9 ω greek small letter omega TODHRI LETTER PA TODHRI LETTER PSA \rightarrow 03C8 ψ greek small letter psi TODHRI LETTER QA TODHRI LETTER RA TODHRI LETTER RA TODHRI LETTER SA TODHRI LETTER SKAN \rightarrow 03BE ξ greek small letter ksi TODHRI LETTER STA \rightarrow 03DB ζ greek small letter stigma TODHRI LETTER SHA TODHRI LETTER SHA TODHRI LETTER SHA TODHRI LETTER SHTA • in origin a ligature of sha and ta TODHRI LETTER TA
105DE 105DF 105E0 105E1 105E2 105E3 105E4 105E5 105E6 105E7	₩ 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	TODHRI LETTER OO \rightarrow 03C9 ω greek small letter omega TODHRI LETTER PA TODHRI LETTER PSA \rightarrow 03C8 ψ greek small letter psi TODHRI LETTER QA TODHRI LETTER RA TODHRI LETTER RAA TODHRI LETTER SAA TODHRI LETTER SKAN \rightarrow 03BE ξ greek small letter ksi TODHRI LETTER STA \rightarrow 03DB ζ greek small letter stigma TODHRI LETTER SHA TODHRI LETTER SHA TODHRI LETTER SHA
105DE 105DF 105E0 105E1 105E2 105E3 105E4 105E5 105E6 105E7 105E8 105E9	₩ Ч У У У 8 8 5 8 8 5 8 8 5 8 8 5 8 8 5 8 8 5 8 8 5 8 8 8 5 8 8 8 5 8 8 8 5 8 8 8 8 9 8 9	TODHRI LETTER OO \rightarrow 03C9 ω greek small letter omega TODHRI LETTER PA TODHRI LETTER PSA \rightarrow 03C8 ψ greek small letter psi TODHRI LETTER QA TODHRI LETTER RA TODHRI LETTER RA TODHRI LETTER SA TODHRI LETTER SKAN \rightarrow 03BE ξ greek small letter ksi TODHRI LETTER STA \rightarrow 03DB ζ greek small letter stigma TODHRI LETTER SHA \rightarrow 03DB ζ greek small letter stigma TODHRI LETTER SHA \rightarrow 010HRI LETTER SHA \bullet in origin a ligature of sha and ta TODHRI LETTER TA TODHRI LETTER THA
105DE 105DF 105E1 105E2 105E3 105E4 105E5 105E6 105E7 105E8 105E8 105E9 105EA	₩ ч ч ч ч ч ч ч ч ч ч ч ч ч	TODHRI LETTER OO \rightarrow 03C9 ω greek small letter omega TODHRI LETTER PA TODHRI LETTER PSA \rightarrow 03C8 ψ greek small letter psi TODHRI LETTER QA TODHRI LETTER RA TODHRI LETTER RA TODHRI LETTER SA TODHRI LETTER SA \rightarrow 03BE ξ greek small letter ksi TODHRI LETTER STA \rightarrow 03DB ζ greek small letter stigma TODHRI LETTER SHA TODHRI LETTER SHA TODHRI LETTER SHA \rightarrow 01 a ligature of sha and ta TODHRI LETTER TA TODHRI LETTER THA TODHRI LETTER THA TODHRI LETTER THA
105DE 105DF 105E1 105E2 105E3 105E4 105E5 105E6 105E7 105E8 105E9 105EA 105EB	Ш у у у у у у у у у у у у у	TODHRI LETTER OO \rightarrow 03C9 ω greek small letter omega TODHRI LETTER PA TODHRI LETTER PSA \rightarrow 03C8 ψ greek small letter psi TODHRI LETTER QA TODHRI LETTER RA TODHRI LETTER RA TODHRI LETTER SA TODHRI LETTER SA \rightarrow 03BE ξ greek small letter ksi TODHRI LETTER STA \rightarrow 03DB ζ greek small letter stigma TODHRI LETTER SHA TODHRI LETTER SHA \rightarrow 01B ζ greek small letter stigma TODHRI LETTER SHA \rightarrow 01B ζ greek small letter stigma TODHRI LETTER SHA \rightarrow 01B ζ greek small letter stigma \rightarrow 01B ζ greek small ζ star ζ s
105DE 105DF 105E1 105E2 105E3 105E4 105E5 105E6 105E7 105E8 105E9 105EA 105EB 105EB	₩ Ц Ч 5 8 5 8 8 5 8 8 7 8 8 7 8 8 8 8 8 5 8 8 8 8 8 8 8 8 8 8 8 8 8	TODHRI LETTER OO \rightarrow 03C9 ω greek small letter omega TODHRI LETTER PA TODHRI LETTER PSA \rightarrow 03C8 ψ greek small letter psi TODHRI LETTER QA TODHRI LETTER RA TODHRI LETTER RA TODHRI LETTER SA TODHRI LETTER SKAN \rightarrow 03BE ξ greek small letter ksi TODHRI LETTER STA \rightarrow 03DB ζ greek small letter stigma TODHRI LETTER SHA TODHRI LETTER SHA \rightarrow 01RI LETTER SHA \rightarrow 01RI LETTER SHA \rightarrow 01RI LETTER TA \rightarrow 01RI LETTER TA \rightarrow 01RI LETTER TA \rightarrow 01RI LETTER TA \rightarrow 00HRI LETTER THA \rightarrow 00HRI LETTER VA TODHRI LETTER VA TODHRI LETTER XA
105DE 105DF 105E1 105E2 105E3 105E4 105E5 105E6 105E7 105E8 105E9 105EA 105EB 105EC 105ED	Ш у у у у у у у у у у у у у	TODHRI LETTER OO \rightarrow 03C9 ω greek small letter omega TODHRI LETTER PA TODHRI LETTER PSA \rightarrow 03C8 ψ greek small letter psi TODHRI LETTER QA TODHRI LETTER RA TODHRI LETTER RA TODHRI LETTER SA TODHRI LETTER SKAN \rightarrow 03BE ξ greek small letter ksi TODHRI LETTER STA \rightarrow 03DB ζ greek small letter stigma TODHRI LETTER SHA TODHRI LETTER SHA \rightarrow 010HRI LETTER SHA TODHRI LETTER SHA TODHRI LETTER TA \rightarrow 010HRI LETTER TA \rightarrow 010HRI LETTER VA TODHRI LETTER VA TODHRI LETTER VA TODHRI LETTER XA
105DE 105DF 105E1 105E2 105E3 105E4 105E5 105E6 105E7 105E8 105E9 105EA 105EB 105EB	₩ Ц Ч 5 8 5 8 8 5 8 8 7 8 8 7 8 8 8 8 8 5 8 8 8 8 8 8 8 8 8 8 8 8 8	TODHRI LETTER OO \rightarrow 03C9 ω greek small letter omega TODHRI LETTER PA TODHRI LETTER PSA \rightarrow 03C8 ψ greek small letter psi TODHRI LETTER QA TODHRI LETTER RA TODHRI LETTER RA TODHRI LETTER SA TODHRI LETTER SKAN \rightarrow 03BE ξ greek small letter ksi TODHRI LETTER STA \rightarrow 03DB ζ greek small letter stigma TODHRI LETTER SHA TODHRI LETTER SHA \rightarrow 01RI LETTER SHA \rightarrow 01RI LETTER SHA \rightarrow 01RI LETTER TA \rightarrow 01RI LETTER TA \rightarrow 01RI LETTER TA \rightarrow 01RI LETTER TA \rightarrow 00HRI LETTER THA \rightarrow 00HRI LETTER VA TODHRI LETTER VA TODHRI LETTER XA
105DE 105DF 105E1 105E2 105E3 105E4 105E5 105E6 105E7 105E8 105E9 105EA 105EB 105EC 105ED	ال بال بال بال بال بال بال بال بال بال ب	TODHRI LETTER OO \rightarrow 03C9 ω greek small letter omega TODHRI LETTER PA TODHRI LETTER PSA \rightarrow 03C8 ψ greek small letter psi TODHRI LETTER QA TODHRI LETTER RA TODHRI LETTER RA TODHRI LETTER SA TODHRI LETTER SKAN \rightarrow 03BE ξ greek small letter ksi TODHRI LETTER STA \rightarrow 03DB ζ greek small letter stigma TODHRI LETTER SHA TODHRI LETTER SHA \rightarrow 010HRI LETTER SHA TODHRI LETTER SHA TODHRI LETTER TA \rightarrow 010HRI LETTER TA \rightarrow 010HRI LETTER VA TODHRI LETTER VA TODHRI LETTER VA TODHRI LETTER XA
105DE 105DF 105E1 105E2 105E3 105E4 105E5 105E6 105E7 105E8 105E9 105EA 105EB 105EC 105ED 105EE 105EF	ال بال بال بال بال بال بال بال بال بال ب	TODHRI LETTER OO \rightarrow 03C9 ω greek small letter omega TODHRI LETTER PA TODHRI LETTER PSA \rightarrow 03C8 ψ greek small letter psi TODHRI LETTER QA TODHRI LETTER RA TODHRI LETTER RA TODHRI LETTER SA TODHRI LETTER SKAN \rightarrow 03BE ξ greek small letter ksi TODHRI LETTER STA \rightarrow 03DB ζ greek small letter stigma TODHRI LETTER SHA TODHRI LETTER SHA \rightarrow 03DB ζ greek small letter stigma TODHRI LETTER SHA \rightarrow 03DB ζ greek small letter stigma TODHRI LETTER SHA \rightarrow 03DB ζ greek small letter stigma TODHRI LETTER SHA \rightarrow 010HRI LETTER SHA \rightarrow 010HRI LETTER VA TODHRI LETTER VA TODHRI LETTER VA TODHRI LETTER NXA TODHRI LETTER NXA TODHRI LETTER XHA TODHRI LETTER XHA
105DE 105DF 105E1 105E2 105E3 105E4 105E5 105E6 105E7 105E8 105E9 105EA 105EB 105EC 105ED 105EC 105ED 105EE 105EF	ш Ц У у к в х х х х х х х х х х х х х	TODHRI LETTER OO \rightarrow 03C9 ω greek small letter omega TODHRI LETTER PA TODHRI LETTER PSA \rightarrow 03C8 ψ greek small letter psi TODHRI LETTER QA TODHRI LETTER RA TODHRI LETTER RA TODHRI LETTER SA TODHRI LETTER SA \rightarrow 03BE ξ greek small letter ksi TODHRI LETTER STA \rightarrow 03DB ζ greek small letter stigma TODHRI LETTER SHA TODHRI LETTER SHA TODHRI LETTER SHA TODHRI LETTER SHTA • in origin a ligature of sha and ta TODHRI LETTER THA TODHRI LETTER TA TODHRI LETTER TA TODHRI LETTER VA TODHRI LETTER VA TODHRI LETTER NXA TODHRI LETTER NXA TODHRI LETTER NXA TODHRI LETTER NXHA TODHRI LETTER NXHA TODHRI LETTER NXHA TODHRI LETTER NXHA
105DE 105DF 105E0 105E1 105E2 105E3 105E4 105E5 105E6 105E7 105E8 105E8 105E8 105E8 105E8 105EB 105EC 105ED 105EE 105EF 105F0 105F1	шЦ %5f28 5 XX q20C Яхд9ды	TODHRI LETTER OO \rightarrow 03C9 ω greek small letter omega TODHRI LETTER PA TODHRI LETTER PSA \rightarrow 03C8 ψ greek small letter psi TODHRI LETTER QA TODHRI LETTER RA TODHRI LETTER RA TODHRI LETTER SA TODHRI LETTER SA \rightarrow 03BE ξ greek small letter ksi TODHRI LETTER STA \rightarrow 03DB ζ greek small letter stigma TODHRI LETTER SHA TODHRI LETTER SHA TODHRI LETTER SHA TODHRI LETTER SHTA • in origin a ligature of sha and ta TODHRI LETTER THA TODHRI LETTER TA TODHRI LETTER TA TODHRI LETTER VA TODHRI LETTER VA TODHRI LETTER NXA TODHRI LETTER NXA TODHRI LETTER NXHA TODHRI LETTER NXHA TODHRI LETTER NXHA TODHRI LETTER Y TODHRI LETTER JY
105DE 105DF 105E0 105E1 105E2 105E3 105E4 105E5 105E6 105E7 105E8 105E9 105EA 105EB 105EC 105ED 105EE 105EF 105F0 105F1 105F2	しい いちょうしょう ひょう しょうしょうしょう	TODHRI LETTER OO \rightarrow 03C9 ω greek small letter omega TODHRI LETTER PA TODHRI LETTER PSA \rightarrow 03C8 ψ greek small letter psi TODHRI LETTER QA TODHRI LETTER RA TODHRI LETTER RA TODHRI LETTER SA TODHRI LETTER SA \rightarrow 03BE ξ greek small letter ksi TODHRI LETTER STA \rightarrow 03DB ζ greek small letter stigma TODHRI LETTER SHA TODHRI LETTER SHA TODHRI LETTER SHA TODHRI LETTER SHTA • in origin a ligature of sha and ta TODHRI LETTER TA TODHRI LETTER TA 100HRI LETTER TA TODHRI LETTER VA TODHRI LETTER VA TODHRI LETTER XA TODHRI LETTER NXA TODHRI LETTER NXA TODHRI LETTER NXHA TODHRI LETTER NXHA TODHRI LETTER Y TODHRI LETTER JY TODHRI LETTER ZA
105DE 105DF 105E0 105E1 105E2 105E3 105E4 105E5 105E6 105E7 105E8 105E8 105E8 105E8 105E8 105EB 105EC 105ED 105EE 105EF 105F0 105F1	шЦ %5f28 5 XX q20C Яхд9ды	TODHRI LETTER OO \rightarrow 03C9 ω greek small letter omega TODHRI LETTER PA TODHRI LETTER PSA \rightarrow 03C8 ψ greek small letter psi TODHRI LETTER QA TODHRI LETTER RA TODHRI LETTER RA TODHRI LETTER SA TODHRI LETTER SA \rightarrow 03BE ξ greek small letter ksi TODHRI LETTER STA \rightarrow 03DB ζ greek small letter stigma TODHRI LETTER SHA TODHRI LETTER SHA TODHRI LETTER SHA TODHRI LETTER SHTA • in origin a ligature of sha and ta TODHRI LETTER THA TODHRI LETTER TA TODHRI LETTER TA TODHRI LETTER VA TODHRI LETTER VA TODHRI LETTER NXA TODHRI LETTER NXA TODHRI LETTER NXHA TODHRI LETTER NXHA TODHRI LETTER NXHA TODHRI LETTER Y TODHRI LETTER JY

Printed using UniBook[™] (http://www.unicode.org/unibook/)

Printed: 01-Oct-2020

17

A. Administrative

1. Title

Proposal for encoding the Todhri script in the SMP of the UCS

2. Requester's name

UC Berkeley Script Encoding Initiative (Universal Scripts Project)

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

Liaison contribution.

4. Submission date

2020-10-01

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

Todhri.

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

53.

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?

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? **Jaon Glavy via Michael Everson.**

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

Michael Everson, FontLab.

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 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, in L2/09-328.

2b. If YES, with whom?

Robert Elsie, Stiven Llupa, Edon Muhaxheri.

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?

See above.

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

To write the Albanian language.

4b. Reference

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

Yes.

5b. If YES, where?

In scholarly 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?