

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

**Doc Type:** Working Group Document

**Title:** Optimizing Georgian representation in the BMP of the UCS

**Source:** Michael Everson, EGT (IE)

**Status:** Expert Contribution

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

**Date:** 1999-02-26

This document requests additional characters to be added to the UCS and contains the proposal summary form.

## A. Administrative

### 1. Title

Optimizing Georgian representation in the BMP of the UCS.

### 2. Requester's name

Michael Everson, EGT (WG2 member for Ireland).

### 3. Requester type

Expert contribution.

### 4. Submission date

1999-02-26.

### 5. Requester's reference

#### 6a. Completion

This is a complete proposal.

#### 6b. More information to be provided?

No.

## B. Technical -- General

### 1a. New script? Name?

Yes. Nuskhuri (38 characters).

### 1b. Addition of characters to existing block? Name?

Yes. Basic Georgian, Cyrillic, and Modifier Letters (5 + 1 + 7 characters).

### 2. Number of characters

51 (38 + 5 + 1 + 7)

### 3. Proposed category

Category A but not very rigorously so.

### 4. Proposed level of implementation and rationale

Level 1; no combining characters are used.

### 5a. Character names included in proposal?

Yes.

### 5b. Character names in accordance with guidelines?

Yes.

### 5c. Character shapes reviewable?

Yes (see below).

**6a. Who will provide computerized font?**

The late Anton Dumbadze via Iraki Garibashvili and Michael Everson. Anton was killed in a car accident in November 1998. His beautiful glyphs should be used in both ISO/IEC 10646 and Unicode.

**6b. Font currently available?**

Yes.

**6c. Font format?**

TrueType.

**7a. Are references (to other character sets, dictionaries, descriptive texts, etc.) provided?**

Yes, see bibliography below.

**7b. Are published examples (such as samples from newspapers, magazines, or other sources) of use of proposed characters attached?**

Yes, see below.

**8. Does the proposal address other aspects of character data processing?**

Yes, see Unicode properties below.

## **C. Technical -- Justification**

**1. Contact with the user community?**

Yes. Irakli Garibashvili, Jost Gippert, Carl-Martin Bunz.

**2. Information on the user community?**

Georgian specialists inside and outside Georgia, including members of the Georgian Academy of Sciences and library associations in Georgia. There is a resurgence of ecclesiastical use of Nuskhuri in Georgia.

**3a. The context of use for the proposed characters?**

Used to write Georgian and other minority languages in Georgia; some characters are also used in Georgian- and Latin-based transliteration of Caucasian languages.

**3b. Reference**

See bibliography.

**4a. Proposed characters in current use?**

Yes.

**4b. Where?**

In scholarly and ecclesiastical contexts.

**5a. Characters should be encoded entirely in BMP?**

Yes.

**5b. Rationale**

All Georgian characters should be encoded in the BMP.

**6. Should characters be kept in a continuous range?**

No. Nuskhuri is a new block; other symbols should be added in the appropriate blocks.

**7a. Can the characters be considered a presentation form of an existing character or character sequence?**

No.

**7b. Where?**

**7c. Reference**

**8a. Can any of the characters be considered to be similar (in appearance or function) to an existing character?**

No.

**8b. Where?**

**8c. Reference**

**9a. Combining characters or use of composite sequences included?**

No. But the characters at xx34–xx36 may need discussion (see below).

**9b. List of composite sequences and their corresponding glyph images provided?**

**10. Characters with any special properties such as control function, etc. included?**

No.

**E. Proposal**

**The Nuskhuri script needs to be added to the UCS.** Although the UCS currently treats the Nuskhuri and Mkhedruli scripts as “font variants” of an underlying “lower case” Georgian script, the Nuskhuri script should be disunified from the Mkhedruli script. Numerous manuscripts exist in which the three scripts are used on the same page (see Figure 1 below). Computer representation of these manuscripts should be able to encode them as plain text. If Nuskhuri enjoyed truly marginal use, possibly the unification might stand; but two facts mitigate against this: the historical origin of Nuskhuri and Mkhedruli, and the resurgence at the present time of Nuskhuri in ecclesiastical contexts in Georgia. It should be noted that Nuskhuri was used much more widely than Asomtavruli was, though Asomtavruli has already been encoded in the UCS. Irakli Garibashvili tells me that *most* documents in Georgian were written in Nuskhuri until the 19th century CE, and that “the church is publishing more (than several years ago) documents in Nuskhuri (‘for internal use’).”

It is easy to see the glyphic relationship between the Asomtavruli and Nuskhuri scripts (see the chart on page 12 below). Asomtavruli is suited to monumental inscriptions; Nuskhuri can be described as Asomtavruli shapes represented by simple penstrokes, suited to manuscript writing (*cf.* the development of Latin and Greek minuscules from their capitals). Georgian scholarship indicates that Mkhedruli was derived *not* from the Nuskhuri current hand, but *independently* from monumental Asomtavruli. Close analysis and a bit of imagination can show how the Asomtavruli letterforms were modified to produce Mkhedruli; a full analysis of this evolution may in fact be a part of Georgian alphabet scholarship, though I have not seen a set of comprehensive charts. A plausible description of a few characters may be ventured here, however:

- 10A1 Ⴀ GEORGIAN CAPITAL LETTER BAN has a greatly reduced upper bow and its lower right foot closed to a loop in 10D1 Ⴁ GEORGIAN LETTER BAN. (It might be suggested that Mkhedruli BAN was derived by turning Asomtavruli BAN 180°. I think the looping of the right foot in other letters (GAN and XAN) speaks against this, but one never knows.)
- 10A2 Ⴂ GEORGIAN CAPITAL LETTER GAN has its right angle transformed to a curve and its lower right foot closed to a loop in 10D2 Ⴃ GEORGIAN LETTER GAN.
- 10BE Ⴄ GEORGIAN CAPITAL LETTER XAN has its right angle transformed to a diagonal line and its lower right foot closed to a loop in 10EE Ⴅ GEORGIAN LETTER XAN.
- 10A4 Ⴆ GEORGIAN CAPITAL LETTER EN has its right angle transformed to a curve and its vertical bar extended and swung left to form a curve in 10D4 Ⴇ GEORGIAN LETTER EN.
- 10AF Ⴇ GEORGIAN CAPITAL LETTER ZHAR has a reduced upper bow and its vertical bar extended and swung left to form a curve in 10DF Ⴈ GEORGIAN LETTER ZHAR.
- 10B4 Ⴈ GEORGIAN CAPITAL LETTER PHAR has its bowls reanalysed somewhat and its vertical bar extended and swung left to form a curve in 10E4 Ⴉ GEORGIAN LETTER PHAR.
- 10B5 Ⴉ GEORGIAN CAPITAL LETTER KHAR has modifications to its head and crossbar and its vertical bar extended and swung left to form a curve in 10E5 Ⴊ GEORGIAN LETTER KHAR.
- 10B7 Ⴊ GEORGIAN CAPITAL LETTER QAR has a reduced upper bow and its vertical bar extended and swung left to form a curve in 10E7 Ⴋ GEORGIAN LETTER QAR.
- 10A3 Ⴋ GEORGIAN CAPITAL LETTER DON has been more or less turned on its side and its horizontal bar modified in 10D3 Ⴌ GEORGIAN LETTER DON.

It is clear from the penstrokes that Mkhedruli was derived from Asomtavruli without reference to Nuskhuri. It is also clear that the unification of Mkhedruli with Nuskhuri is a wholly different thing from

the font-based unification of German *Fraktur* or Irish *Cló Gaelach* with ordinary Latin. In the first place, a patient naïve reader of French can easily puzzle out, say, a French text written in *Fraktur* or *Cló Gaelach* (ლა ლანგუე გეორგიენე = la langue géorgienne), while a naïve reader of Georgian *cannot* read a Georgian text written in Nuskhuri without a key (ქართული ენა = ქართული ენა). The shapes are simply too different. Unification of Nuskhuri with Mkhedruli is somewhat analogous to a posited unification of Ogham with Latin, or of Phoenician with Hebrew, on purely structural grounds.

The Nuskhuri characters presented in xx00–xx2F below should be added to the UCS. The annotations (Mkhedruli) and (Nuskhuri) should be added to the relevant character names, and the annotation (Khutsuri) should be changed to (Asomtavruli). Western scholarship has treated Asomtavruli and Nuskhuri as casing pairs of a script called Khutsuri, but this classification is at odds with current Georgian terminology and classification. Manuscripts do use Asomtavruli as capitals with Nuskhuri, but there has also been a 20th-century attempt to use Asomtavruli as capitals with Mkhedruli. This was promulgated by Prof. A. G. Shanidze; it was not popularly accepted, but has been found in works by him and in works on the Georgian script by others honouring him (e.g. შანიძე 1970).

A particular style of Mkhedruli is called *mtavruli*. In this style, the distinction between letters with ascenders and descenders is not maintained, but all letters appear with an equal height standing on the baseline (example: ქართული ენა vs. მბრთული ენა). Mtavruli-style letters are never used as capitals; a word is always entirely presented in mtavruli or not. Mtavruli-style is used in titles and other kinds of headings. I suggest that non-plain text SMALL CAPS coding be used to represent Mkhedruli in mtavruli-style. (There is no other such small caps category category that would have an application in Georgian.) The use of Asomtavruli as capitals, whether with Nuskhuri or with Mkhedruli, is not casing as normally understood in Latin, Greek, Cyrillic, or Armenian; capitalization in Georgian is an unusual matter of mixing (related) scripts. Irakli Garibashvili pointed out to me that in documents with Asomtavruli capitals with Nuskhuri text, Asomtavruli capitals are always found at the beginning of a paragraph but for sentences within a paragraph they may be optional.

**Five Georgian characters and one Cyrillic character need to be added to the UCS.** The UCS character encoded at 10F6 GEORGIAN LETTER FI is used to represent /f/ in Laz and in Georgian dialects. In addition, 10F7 გ GEORGIAN LETTER SHVA and 10F8 ზ GEORGIAN LETTER ELIFI, used in transcriptions in Svan and Laz (see Figure 3 and Figure 2 below) should be added; these are widely attested and have been in use for at least the better part of a century; indeed it is not only transcriptions in which these characters appear: for Mingrelian, at least, publishing of books and newspapers was undertaken in the late 1920s and 1930s, (*cf.* Enwall 1992). 10F9 Ⴀ GEORGIAN LETTER TURNED GAN is used in transcriptions of some east Caucasian languages to indicate the voiced uvular stop (IPA /g/); it is used in transcriptions of Tabasaran to indicate the voiced velar fricative (IPA /ɣ/) (see Figure 7 below). 10FA <sup>6</sup> MODIFIER LETTER GEORGIAN NAR is widely used to indicate nasalization of a preceding vowel in many standard linguistic transcriptions of Caucasian languages. It is also necessary to add xx37 <sup>H</sup> MODIFIER LETTER CYRILLIC EN for the same reason: the nasal /ã/ is variously transcribed as *a<sup>n</sup>*, *ɶ<sup>6</sup>*, and *a<sup>H</sup>* in Latin, Georgian, and Cyrillic Caucasian linguistic texts (see Figure 6 below; 207F SUPERSCRIPT LATIN SMALL LETTER N is already encoded). 10FC Ⴁ GEORGIAN LETTER AIN (a name coined to reflect the letter's own derivation from ARABIC LETTER AIN) is used in transcriptions of Bats (see Figure 4 below).

**Five modifier letters need to be added to the UCS.** xx30 ~ MODIFIER LETTER PALATALIZATION, xx31 ̡ MODIFIER LETTER LATERAL AFFRICTION, and xx32 ̢ MODIFIER LETTER INTENSIVITY are used in many sources I have to hand (see Figures 5 and 7). They are used with either the Georgian or the Latin scripts in linguistic transcription. xx33 ̣ MODIFIER LETTER INTENSIVE PHARYNGEALIZATION is used following 10F0 GEORGIAN LETTER HAE or 0068 LATIN SMALL LETTER H to represent a pharyngealized *h* (IPA /ħ/) in Bats (see Figure 4 below).

**Three modifier letters may need to be added to the UCS.** Some work needs to be done with regard to xx34–xx36. Caucasian linguistic practice presents these sitting on the base line as well as raised like modifier letters; they also combine with one another in the vertical plane (which also happens with some Finno-Ugric Phonetic Alphabet (FUPA) characters).

NOTE: A comprehensive paper on the FUPA is being prepared by myself and Klaas Ruppel.

The UCS has 02DA MODIFIER LETTER RING ABOVE, but not MODIFIER LETTER LOW RING. Possibly xx35 <sup>◌̣</sup> MODIFIER LETTER PRERUPTIVITY and <sup>◌̤</sup> MODIFIER LETTER PHARYNGEALIZATION could be unified with other apostrophes or commas in the UCS (02BB, 02BC, 02BE, 02BF). I have seen xx35 written with a superscript <sup>◌̣</sup> and xx32 written with superscript schwa or baseline schwa in Klimov 1994 (but Gippert says that this was just a font convenience as he prepared the camera-ready copy). In Figure 7 below, ‘ is found; but in Figure 8 (provided by Jost Gippert) ‘ and ’ are distinguished from <sup>◌̣</sup> and <sup>◌̤</sup> (ღ<sup>◌̣</sup>/ღ‘ and ზ<sup>◌̤</sup>/ზ’).

As noted above, a further issue has to do with the possibility of combining xx32 ◌̣ MODIFIER LETTER INTENSIVITY and xx33 ◌̤ MODIFIER LETTER LABIALIZATION. It is possible that ◌̤ would be xx33 + 030A COMBINING RING ABOVE (or?? 02DA ◌̣ MODIFIER LETTER RING ABOVE + COMBINING INTENSIVITY BELOW); the height of ◌̣ MODIFIER LETTER INTENSIVITY seems to vary between low and high in sources to hand. This may argue for more than one of these to be added to the UCS. Actually, a useful exercise would be simply to list a number of characters shown in Klimov 1994 (Appendix 4) for whom the UCS representation needs to be worked out, comparing these with the samples from გუგუა 1971 which have low ◌̣. There is also a kind of LOW PRIME whose meaning is unknown to me. See Figure 8 below, which provides a repertoire of Georgian phonetic characters.

## Unicode Character Properties

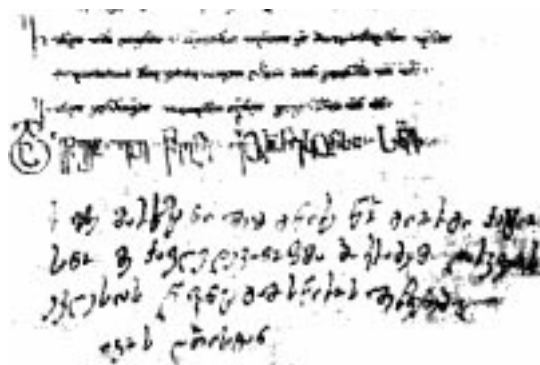
Spacing letters, category “Lo”, bidi category “L” (strong left to right)

10F7-10FA, xx00-xx25, xx30-xx37

## Bibliography

- Enwall, Joakim. 1992. "Some remarks on the language debate in the Mingrelian newspaper 'Qazaqishi Gazeti'", in: George Hewitt (ed.), *Caucasian Perspectives*. Unterschleissheim/München: Lincom. Pp. 278-284.
- Haarmann, Harald. 1990. *Die Universalgeschichte der Schrift*. Frankfurt: Campus Verlag. ISBN 3-593-34346-0
- Klimov, Georgij A. 1994. *Einführung in die kaukasische Sprachwissenschaft*. Tr. Jost Gippert. Hamburg: Buske Verlag. ISBN 3-87548-060-0
- Marr, N. & M. Brière. 1931. *La langue géorgienne*. Paris: Firmin-Didot et C<sup>ie</sup>.
- Имнайшвили, Давид Силибистрович (= დ. იმნაიშვილი). 1977. *Историко-сравнительный анализ фонетики нахских языков*. Тбилиси: Мецниереба.
- Кадагидзе, Давит, & Нико Кадагидзе. 1984. *Цова-тушинско-грузинско-русский словарь*. Тбилиси: Мецниереба.
- Магоматов, Александр Амарович (= ალ. მაჰომეტოვი). 1965. *Табасаранский язык: исследование и тексты* (= ტაბასარანუმი ენა: გამოკვლევა და ტექსტები). Тбилиси: Мецниереба.
- МаррЪ, Н. 1910. *Грамматика Чанского (лазского) языка: съ хрестоматією и словаремъ*. (Матеріалы по яфетичяскому языкознанію: II) С.-Петербургъ: Типографія Императорской Академіи Наукъ.
- ჭლექსანდრელი, ჭთანასი (= Афанасий Александрийский). 1970. *ჯხორეპაჲ წმიდისა ჭნტონისი*. (= Афанасий Александрийский) *Житие Святого Антония*. Eds. ჭ. შანიძე & ჭახტანგ შმნაიშვილი (= A. G. Шანიძე & B. I. Имнайшвили). ტბილისი: თბილისის უნივერსიტეტის გამომცემლობა.
- გუდავა, გ. (= Того Евстафьевич Гудава). 1971. *ბაგვაღური ენა: გრამატიკური ანალიზი და ტექსტები* (= *Багвалийский язык: грамматический анализ с текстами*). თბილისი: მეცნიერება.
- კარტოზია, გურამ. 1993. *ლაზური ტექსტები II*. თბილისი.
- ლომთათიძე, ქ. 1976. *აფხაზური და აბაზური ენების ისტორიულ-შედარებითი ანალიზი*. თბილისი: მეცნიერება.
- ჯონიძე მესხაშვიმ, არსენძე. (= Арсень Оніанъ изъ Сасāша). 1917. *ლუმნუ ამბჯარ ლცლშხუ შუმი ნინშჯ გიმ ი ხნეძარ, არკარ ი ციგინარ, შადრალ ი ლიდრალ, ლიმზერალ* (= *Сванскіе тексты на лāшскомъ нарѣчїи: страна и нравы, рассказы и сказки, стихи и пѣсни, заговоры*). Петроградъ: Типография Академіи Наукъ.
- შანიძე, ა., ვ. თოფურია, et al. 1939–1979. *სვანური პროზაული ტექსტები*. 4 vols. (1939, 1957, 1979, 1967), თბილისი.
- ჭარაია, პეტრე. 1997. *მეგრულ-ქართული ლექსიკონი*. თბილისი.
- ხუბუა, მაკარ. 1937. *მეგრული ტექსტები* = *Мегрельские тексты* = *Textes mégréliens*. თბილისი.

Annex



**Figure 1.** Sample of Nuskhuri, Asomtavruli, and Mkhedruli on one page in a manuscript. From a scan on Jost Gippert’s web site.

35.	ჰჳ	ჰ	ჰ	-[h]	h(±) ح
36.	ჴ	ჴ	ჴ	Φ	f ف
37.	ქლიჴი	ხ	ხ	-[ʰ]	ʰ(±) أ

**Figure 2.** The letter ELIFI shown with its name and its Arabic transliteration in Маррѣ 1910, a grammar of Laz. The sorting order shows that ELIFI following FI. The Mingrelian-Georgian dictionary of 1997 (Ch’araia 1997), ჲ ELIFI follows ყ QAR (which it looks like).

აღდურ-ბაღდურ,  
 ბახსან ქეღდურ.  
 ქელი ქობა ჭადარისა,  
 მაი ჭაკუ? და შალი შართქუმ,  
 ყაჭი თურქატ დაფანებე,  
 ტუფას ტუფას მეშხე თჯეთუნე  
 გეზალ ხორი აფუნსაში გკლასხანი.

**Figure 3.** Sample of Svan text showing SHVA in lines 1, 2, and 5, from ჳონბან მკსსჳშიჳ 1917.

ჳალღირ-ღ-არ ჰალ-  
 wir-d-ar ბ.ა.ა. ღირ-ღ-არ  
 wir-d-ar  
 ჳალღირ-ღ-ითარ ჰალ  
 wir-d-itar ბ.ა.ა. ღირ-  
 -ღ-ითარ wir-d-itar. ნე-

**Figure 4.** Sample Bats text showing MODIFIER LETTER INTENSIVE PHARYNGEALIZATION used with Georgian and Latin scripts, as well as GEORGIAN LETTER AYN, in Кадагидзе & Кадагидзе 1984. This dictionary, the only Bats dictionary published to date, was reproduced from a handwritten manuscript. The first two words read ჳალღირ-ღ-არ and *hʷalwir-d-ar*.



აფრიკატი		სპირანტი			
ყ რ ჳ		მჟღერი	ყ რ ჳ		მჟღერი
ფშენვ.	მევეთრი		ფშენვ.	მევეთრი	
		ჟ			
	ჭ	ჭ	სა	ს	ჭა
ჩ	ჭ	ჭ	შა	შ	ჭა
		ჩ			
				ბ	
		ღ	ბა	ბ	
ქ	ყ	ჭ	ჭა	ჭა	
	ც			ჭ	
	ტ				
	ტ				

Аффрикаты		Спиранты			
Глухие		Глухие			
Придых.	Абр. лт.	Звонкие	Придых.		Абр. лт.
			Сильн.	Слаб.	
		w			
	ჭ	z	sa	s	ჭა
ჯ	ჯ	ჯ	შა	შ	ჯა
		j			
			x		
		γ	xა	x	
q	q	ა		ha	
				h	
	ტ		lა		

Figure 5. Part of two tables in გულაზა 1971, showing Georgian and Latin transliteration of affricates and spirants in Bagulal phonology, including the characters ˘ MODIFIER LETTER PALATALIZATION, 6 MODIFIER LETTER LATERAL AFFRICATION, օ MODIFIER LETTER INTENSIVITY, and ჟ GEORGIAN LETTER ELIFI.

1. Чеченская			2. Ингушская			2. Батбийская	
Латинская	Грузинская	Лит. чеч.	Латинская	Грузинская	Лит. ингуш.	Латинская	Грузинская
a	ა	a	a	ა	a	a	ა
a	ა	-	a	ა	-	a	ა
ა	ა	ა	-	-	-	-	-
ა	ა	ა	ა	ა	ა	-	-
ა	ა	(ა <sup>n</sup> )	-	-	-	ა <sup>n</sup>	ა <sup>n</sup>
ა	ა	-	-	-	-	-	-
ბ	ბ	ბ	ბ	ბ	ბ	ბ	ბ
ბ	ბ	ბ	ბ	ბ	ბ	ბ	ბ
ბ	ბ	ბ	ბ	ბ	ბ	ბ	ბ
ბ	ბ	ბ	ბ	ბ	ბ	ბ	ბ
ბ	ბ	ბ	ბ	ბ	ბ	ბ	ბ
ბ	ბ	ბ	ბ	ბ	ბ	ბ	ბ
ბ	ბ	ბ	ბ	ბ	ბ	ბ	ბ

Figure 6. Sample of Latin and Georgian transcriptions of Chechen, Ingush, and Bats, with Cyrillic orthography for Chechen and Ingush, showing 207F SUPERSCRIPT LATIN SMALL LETTER N, MODIFIER LETTER SMALL CYRILLIC EN, and MODIFIER LETTER GEORGIAN NAR, in Имнайшвили 1977.



Латин.	Грузин.	Табасаран.	Латин.	Грузин.	Табасаран.
a	ა	а	s	ს	с
ä	ა̇	аь	so	სო	т
b	ბ	б	t	თ	т
w	ჴ	в	ṭ	თ̣	т; тт
g	გ	г	to	თო	у
g'	გ'	г'	u	უ	уь
d	დ	д	ü	უ̇	уь'
e	ე	е	f	ფ	ф
ž	ჯ	ж	fo	ფო	х
ž'	ჯ'	ж'	x	ხ	хь
z	ზ	з	x̣	ხ̣	ц
i	ი	и	co	ცო	ц
j	ი̇	й	č	ჩ	ц; цц
k	კ	к	č'	ჩ'	ч
k <sup>o</sup>	კ <sup>o</sup>	к <sup>o</sup> ; кк	č <sup>o</sup>	ჩ <sup>o</sup>	ч <sup>o</sup>
γ	ღ	гь	č <sup>o</sup> '	ჩ <sup>o</sup> '	ч <sup>o</sup> '
q	ღ̇	хь	č <sup>o</sup> ' <sup>o</sup>	ჩ <sup>o</sup> ' <sup>o</sup>	жь
q <sup>o</sup>	ღ <sup>o</sup>	кь; кьь	š	შ	ш
š	შ	кь	š <sup>o</sup>	შ <sup>o</sup>	ш
l	ლ	л	š <sup>o</sup> '	შ <sup>o</sup> '	шв
m	მ	м	ho	ხო	
n	ნ	н	o	ო	гь
p	პ	п	'	'	ь
p'	პ'	п			
p <sup>o</sup>	პ <sup>o</sup>	п; пп			
r	რ	р			

Figure 7. Transcription of Tabasaran in Latin and Georgian transcription and Cyrillic orthography, showing Ⴑ GEORGIAN LETTER INVERTED GAN, ˘ MODIFIER LETTER PALATALIZATION, Ⴑ MODIFIER LETTER INTENSIVITY, Ⴑ MODIFIER LETTER LABIALIZATION, ˘ MODIFIER LETTER TURNED COMMA and ˘ MODIFIER LETTER APOSTROPHE, in Магомедов 1965.

ა	ბ	გ	დ	ე	ვ	ზ	თ	ი	კ	ლ	მ	ნ	ო	პ	ჟ
რ	ს	ტ	ყ	შ	ჩ	ც	ძ	წ	ჭ	ხ	ძ	წ	ჭ	ხ	ძ
წ	ჭ	ხ	ძ	წ	ჭ	ხ	ძ	წ	ჭ	ხ	ძ	წ	ჭ	ხ	ძ
ჟ	რ	ს	ტ	ყ	შ	ჩ	ც	ძ	წ	ჭ	ხ	ძ	წ	ჭ	ხ
წ	ჭ	ხ	ძ	წ	ჭ	ხ	ძ	წ	ჭ	ხ	ძ	წ	ჭ	ხ	ძ
ჟ	რ	ს	ტ	ყ	შ	ჩ	ც	ძ	წ	ჭ	ხ	ძ	წ	ჭ	ხ
წ	ჭ	ხ	ძ	წ	ჭ	ხ	ძ	წ	ჭ	ხ	ძ	წ	ჭ	ხ	ძ
ჟ	რ	ს	ტ	ყ	შ	ჩ	ც	ძ	წ	ჭ	ხ	ძ	წ	ჭ	ხ
წ	ჭ	ხ	ძ	წ	ჭ	ხ	ძ	წ	ჭ	ხ	ძ	წ	ჭ	ხ	ძ
ჟ	რ	ს	ტ	ყ	შ	ჩ	ც	ძ	წ	ჭ	ხ	ძ	წ	ჭ	ხ
წ	ჭ	ხ	ძ	წ	ჭ	ხ	ძ	წ	ჭ	ხ	ძ	წ	ჭ	ხ	ძ
ჟ	რ	ს	ტ	ყ	შ	ჩ	ც	ძ	წ	ჭ	ხ	ძ	წ	ჭ	ხ
წ	ჭ	ხ	ძ	წ	ჭ	ხ	ძ	წ	ჭ	ხ	ძ	წ	ჭ	ხ	ძ
ჟ	რ	ს	ტ	ყ	შ	ჩ	ც	ძ	წ	ჭ	ხ	ძ	წ	ჭ	ხ
წ	ჭ	ხ	ძ	წ	ჭ	ხ	ძ	წ	ჭ	ხ	ძ	წ	ჭ	ხ	ძ
ჟ	რ	ს	ტ	ყ	შ	ჩ	ც	ძ	წ	ჭ	ხ	ძ	წ	ჭ	ხ
წ	ჭ	ხ	ძ	წ	ჭ	ხ	ძ	წ	ჭ	ხ	ძ	წ	ჭ	ხ	ძ
ჟ	რ	ს	ტ	ყ	შ	ჩ	ც	ძ	წ	ჭ	ხ	ძ	წ	ჭ	ხ
წ	ჭ	ხ	ძ	წ	ჭ	ხ	ძ	წ	ჭ	ხ	ძ	წ	ჭ	ხ	ძ
ჟ	რ	ს	ტ	ყ	შ	ჩ	ც	ძ	წ	ჭ	ხ	ძ	წ	ჭ	ხ
წ	ჭ	ხ	ძ	წ	ჭ	ხ	ძ	წ	ჭ	ხ	ძ	წ	ჭ	ხ	ძ
ჟ	რ	ს	ტ	ყ	შ	ჩ	ც	ძ	წ	ჭ	ხ	ძ	წ	ჭ	ხ
წ	ჭ	ხ	ძ	წ	ჭ	ხ	ძ	წ	ჭ	ხ	ძ	წ	ჭ	ხ	ძ
ჟ	რ	ს	ტ	ყ	შ	ჩ	ც	ძ	წ	ჭ	ხ	ძ	წ	ჭ	ხ
წ	ჭ	ხ	ძ	წ	ჭ	ხ	ძ	წ	ჭ	ხ	ძ	წ	ჭ	ხ	ძ

Figure 8. A complete set of Georgian letters with their diacritical marks, as used in linguistic transcriptions, provided by Jost Gippert. A RING for instance appears to sit on the baseline, the midline (encode two characters?), and a COMBINING RING appears also to be used. Note the distinction in rows 8 and 15 between <sup>◌̇</sup>, <sup>◌̈</sup>, <sup>◌̉</sup>, and <sup>◌̊</sup>.

- 1. — ღორო იგბოჲ ჟბანდჲს ზოჲდსჲ ვაგბ-  
 ზგჲს ზოჲბანდსდგინსდს. ვიგბდ <sup>◌̈</sup> ზგბბბბ  
 ბგჲინსდდ. დსდ ბბბბბბ ბბბბბბბბბბ  
 ბბბბბბბ ბბბბბბბ ბბ ბბბ ბბბ
- 2. — სდდდ ბბბ ბბბბბბ ბბბ ბბბ. ბბბბბბ  
 ბბბბბბბბბბ? ბბბბბბბბ ბბბბბბბ ბბბბბბბბ  
 ბბბბ ბბბ ბბბბბბბბბბ ბბ ბბბბბბბ ბბბბ-  
 ბბბბბბბ-ბბბბბბ ბბბ.

Figure 9. Sample of Asomtavruli-only text, from Marr & Brière 1931.



ASOMTAVRULI, MKHEDRULI, and NUSKHURI GEORGIAN

	10A	10B	10C	10D	10E	10F	xx0	xx1	xx2	xx3
0	Ⴒ	Ⴓ	Ⴔ	Ⴕ	Ⴖ	Ⴗ	Ⴘ	Ⴙ	Ⴚ	Ⴛ
1	Ⴜ	Ⴝ	Ⴞ	Ⴟ	Ⴚ	Ⴛ	Ⴜ	Ⴝ	Ⴞ	Ⴟ
2	Ⴝ	Ⴞ	Ⴟ	Ⴚ	Ⴛ	Ⴜ	Ⴝ	Ⴞ	Ⴟ	Ⴚ
3	Ⴛ	Ⴜ	Ⴝ	Ⴞ	Ⴟ	Ⴚ	Ⴛ	Ⴜ	Ⴝ	Ⴞ
4	Ⴟ	Ⴚ	Ⴛ	Ⴜ	Ⴝ	Ⴞ	Ⴟ	Ⴚ	Ⴛ	Ⴜ
5	Ⴚ	Ⴛ	Ⴜ	Ⴝ	Ⴞ	Ⴟ	Ⴚ	Ⴛ	Ⴜ	Ⴝ
6	Ⴜ	Ⴝ		Ⴟ	Ⴚ	Ⴛ	Ⴜ	Ⴝ		Ⴟ
7	Ⴝ	Ⴞ		Ⴚ	Ⴛ	Ⴜ	Ⴝ	Ⴞ		Ⴚ
8	Ⴞ	Ⴟ		Ⴚ	Ⴛ	Ⴜ	Ⴝ	Ⴞ		
9	Ⴟ	Ⴚ		Ⴜ	Ⴝ	Ⴞ	Ⴟ	Ⴚ		
A	Ⴚ	Ⴛ		Ⴜ	Ⴝ	Ⴞ	Ⴟ	Ⴚ		
B	Ⴛ	Ⴜ		Ⴝ	Ⴞ	Ⴟ	Ⴚ	Ⴛ		
C	Ⴜ	Ⴝ		Ⴞ	Ⴟ	Ⴚ	Ⴛ	Ⴜ		
D	Ⴝ	Ⴞ		Ⴟ	Ⴚ		Ⴛ	Ⴜ		
E	Ⴞ	Ⴟ		Ⴚ	Ⴛ		Ⴜ	Ⴝ		
F	Ⴟ	Ⴚ		Ⴜ	Ⴝ		Ⴞ	Ⴟ		

G = 00  
P = 00

**ASOMTAVRULI, MKHEDRULI, and NUSKHURI GEORGIAN**

dec	hex	Name
	A0	GEORGIAN CAPITAL LETTER AN (Asomtavruli)
	A1	GEORGIAN CAPITAL LETTER BAN (Asomtavruli)
	A2	GEORGIAN CAPITAL LETTER GAN (Asomtavruli)
	A3	GEORGIAN CAPITAL LETTER DON (Asomtavruli)
	A4	GEORGIAN CAPITAL LETTER EN (Asomtavruli)
	A5	GEORGIAN CAPITAL LETTER VIN (Asomtavruli)
	A6	GEORGIAN CAPITAL LETTER ZEN (Asomtavruli)
	A7	GEORGIAN CAPITAL LETTER TAN (Asomtavruli)
	A8	GEORGIAN CAPITAL LETTER IN (Asomtavruli)
	A9	GEORGIAN CAPITAL LETTER KAN (Asomtavruli)
	AA	GEORGIAN CAPITAL LETTER LAS (Asomtavruli)
	AB	GEORGIAN CAPITAL LETTER MAN (Asomtavruli)
	AC	GEORGIAN CAPITAL LETTER NAR (Asomtavruli)
	AD	GEORGIAN CAPITAL LETTER ON (Asomtavruli)
	AE	GEORGIAN CAPITAL LETTER PAR (Asomtavruli)
	AF	GEORGIAN CAPITAL LETTER ZHAR (Asomtavruli)
	B0	GEORGIAN CAPITAL LETTER RAE (Asomtavruli)
	B1	GEORGIAN CAPITAL LETTER SAN (Asomtavruli)
	B2	GEORGIAN CAPITAL LETTER TAR (Asomtavruli)
	B3	GEORGIAN CAPITAL LETTER UN (Asomtavruli)
	B4	GEORGIAN CAPITAL LETTER PHAR (Asomtavruli)
	B5	GEORGIAN CAPITAL LETTER KHAR (Asomtavruli)
	B6	GEORGIAN CAPITAL LETTER GHAN (Asomtavruli)
	B7	GEORGIAN CAPITAL LETTER QAR (Asomtavruli)
	B8	GEORGIAN CAPITAL LETTER SHIN (Asomtavruli)
	B9	GEORGIAN CAPITAL LETTER CHIN (Asomtavruli)
	BA	GEORGIAN CAPITAL LETTER CAN (Asomtavruli)
	BB	GEORGIAN CAPITAL LETTER JIL (Asomtavruli)
	BC	GEORGIAN CAPITAL LETTER CIL (Asomtavruli)
	BD	GEORGIAN CAPITAL LETTER CHAR (Asomtavruli)
	BE	GEORGIAN CAPITAL LETTER XAN (Asomtavruli)
	BF	GEORGIAN CAPITAL LETTER JHAN (Asomtavruli)
	C0	GEORGIAN CAPITAL LETTER HAE (Asomtavruli)
	C1	GEORGIAN CAPITAL LETTER HE (Asomtavruli)
	C2	GEORGIAN CAPITAL LETTER HIE (Asomtavruli)
	C3	GEORGIAN CAPITAL LETTER WE (Asomtavruli)
	C4	GEORGIAN CAPITAL LETTER HAR (Asomtavruli)
	C5	GEORGIAN CAPITAL LETTER HOE (Asomtavruli)
	C6	(This position shall not be used)
	C7	(This position shall not be used)
	C8	(This position shall not be used)
	C9	(This position shall not be used)
	CA	(This position shall not be used)
	CB	(This position shall not be used)
	CC	(This position shall not be used)
	CD	(This position shall not be used)
	CE	(This position shall not be used)
	CF	(This position shall not be used)
	D0	GEORGIAN LETTER AN (Mkhedruli)
	D1	GEORGIAN LETTER BAN (Mkhedruli)
	D2	GEORGIAN LETTER GAN (Mkhedruli)
	D3	GEORGIAN LETTER DON (Mkhedruli)
	D4	GEORGIAN LETTER EN (Mkhedruli)
	D5	GEORGIAN LETTER VIN (Mkhedruli)
	D6	GEORGIAN LETTER ZEN (Mkhedruli)
	D7	GEORGIAN LETTER TAN (Mkhedruli)
	D8	GEORGIAN LETTER IN (Mkhedruli)
	D9	GEORGIAN LETTER KAN (Mkhedruli)
	DA	GEORGIAN LETTER LAS (Mkhedruli)
	DB	GEORGIAN LETTER MAN (Mkhedruli)
	DC	GEORGIAN LETTER NAR (Mkhedruli)
	DD	GEORGIAN LETTER ON (Mkhedruli)
	DE	GEORGIAN LETTER PAR (Mkhedruli)
	DF	GEORGIAN LETTER ZHAR (Mkhedruli)
	E0	GEORGIAN LETTER RAE (Mkhedruli)
	E1	GEORGIAN LETTER SAN (Mkhedruli)
	E2	GEORGIAN LETTER TAR (Mkhedruli)
	E3	GEORGIAN LETTER UN (Mkhedruli)
	E4	GEORGIAN LETTER PHAR (Mkhedruli)
	E5	GEORGIAN LETTER KHAR (Mkhedruli)
	E6	GEORGIAN LETTER GHAN (Mkhedruli)
	E7	GEORGIAN LETTER QAR (Mkhedruli)
	E8	GEORGIAN LETTER SHIN (Mkhedruli)
	E9	GEORGIAN LETTER CHIN (Mkhedruli)
	EA	GEORGIAN LETTER CAN (Mkhedruli)
	EB	GEORGIAN LETTER JIL (Mkhedruli)
	EC	GEORGIAN LETTER CIL (Mkhedruli)
	ED	GEORGIAN LETTER CHAR (Mkhedruli)
	EE	GEORGIAN LETTER XAN (Mkhedruli)
	EF	GEORGIAN LETTER JHAN (Mkhedruli)
	F0	GEORGIAN LETTER HAE (Mkhedruli)
	F1	GEORGIAN LETTER HE (Mkhedruli)
	F2	GEORGIAN LETTER HIE (Mkhedruli)
	F3	GEORGIAN LETTER WE (Mkhedruli)
	F4	GEORGIAN LETTER HAR (Mkhedruli)
	F5	GEORGIAN LETTER HOE (Mkhedruli)
	F6	GEORGIAN LETTER FI (Mkhedruli)
	F7	GEORGIAN LETTER SHVA (Mkhedruli)
	F8	GEORGIAN LETTER ELIFI (Mkhedruli)

dec	hex	Name
	F9	GEORGIAN LETTER TURNED GAN (Mkhedruli)
	FA	MODIFIER LETTER GEORGIAN NAR (Mkhedruli)
	FB	GEORGIAN PARAGRAPH SEPARATOR
	FC	GEORGIAN LETTER AIN (Mkhedruli)
	FD	(This position shall not be used)
	FE	(This position shall not be used)
	FF	(This position shall not be used)
	00	GEORGIAN SMALL LETTER AN (Nuskhuri)
	01	GEORGIAN SMALL LETTER BAN (Nuskhuri)
	02	GEORGIAN SMALL LETTER GAN (Nuskhuri)
	03	GEORGIAN SMALL LETTER DON (Nuskhuri)
	04	GEORGIAN SMALL LETTER EN (Nuskhuri)
	05	GEORGIAN SMALL LETTER VIN (Nuskhuri)
	06	GEORGIAN SMALL LETTER ZEN (Nuskhuri)
	07	GEORGIAN SMALL LETTER TAN (Nuskhuri)
	08	GEORGIAN SMALL LETTER IN (Nuskhuri)
	09	GEORGIAN SMALL LETTER KAN (Nuskhuri)
	0A	GEORGIAN SMALL LETTER LAS (Nuskhuri)
	0B	GEORGIAN SMALL LETTER MAN (Nuskhuri)
	0C	GEORGIAN SMALL LETTER NAR (Nuskhuri)
	0D	GEORGIAN SMALL LETTER ON (Nuskhuri)
	0E	GEORGIAN SMALL LETTER PAR (Nuskhuri)
	0F	GEORGIAN SMALL LETTER ZHAR (Nuskhuri)
	10	GEORGIAN SMALL LETTER RAE (Nuskhuri)
	11	GEORGIAN SMALL LETTER SAN (Nuskhuri)
	12	GEORGIAN SMALL LETTER TAR (Nuskhuri)
	13	GEORGIAN SMALL LETTER UN (Nuskhuri)
	14	GEORGIAN SMALL LETTER PHAR (Nuskhuri)
	15	GEORGIAN SMALL LETTER KHAR (Nuskhuri)
	16	GEORGIAN SMALL LETTER GHAN (Nuskhuri)
	17	GEORGIAN SMALL LETTER QAR (Nuskhuri)
	18	GEORGIAN SMALL LETTER SHIN (Nuskhuri)
	19	GEORGIAN SMALL LETTER CHIN (Nuskhuri)
	1A	GEORGIAN SMALL LETTER CAN (Nuskhuri)
	1B	GEORGIAN SMALL LETTER JIL (Nuskhuri)
	1C	GEORGIAN SMALL LETTER CIL (Nuskhuri)
	1D	GEORGIAN SMALL LETTER CHAR (Nuskhuri)
	1E	GEORGIAN SMALL LETTER XAN (Nuskhuri)
	1F	GEORGIAN SMALL LETTER JHAN (Nuskhuri)
	20	GEORGIAN SMALL LETTER HAE (Nuskhuri)
	21	GEORGIAN SMALL LETTER HE (Nuskhuri)
	22	GEORGIAN SMALL LETTER HIE (Nuskhuri)
	23	GEORGIAN SMALL LETTER WE (Nuskhuri)
	24	GEORGIAN SMALL LETTER HAR (Nuskhuri)
	25	GEORGIAN SMALL LETTER HOE (Nuskhuri)
	26	(This position shall not be used)
	27	(This position shall not be used)
	28	(This position shall not be used)
	29	(This position shall not be used)
	2A	(This position shall not be used)
	2B	(This position shall not be used)
	2C	(This position shall not be used)
	2D	(This position shall not be used)
	2E	(This position shall not be used)
	2F	(This position shall not be used)
	30	MODIFIER LETTER PALATALIZATION
	31	MODIFIER LETTER LATERAL AFFRICTION
	32	MODIFIER LETTER INTENSIVITY
	33	MODIFIER LETTER INTENSIVE PHARYNGEALIZATION
	34	MODIFIER LETTER LABIALIZATION
	35	MODIFIER LETTER PRERUPTIVITY
	36	MODIFIER LETTER PHARYNGEALIZATION
	37	MODIFIER LETTER SMALL CYRILLIC EN
	38	(This position shall not be used)
	39	(This position shall not be used)
	3A	(This position shall not be used)
	3B	(This position shall not be used)
	3C	(This position shall not be used)
	3D	(This position shall not be used)
	3E	(This position shall not be used)
	3F	(This position shall not be used)