

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

**Doc Type: Working Group Document****Title: Proposal to add Georgian and other characters to the BMP of the UCS****Source: NSAI/ICTSCC/SC4 (Ireland) and GSDIT (Georgia)****Status: Member Body Contribution****Action: For consideration by JTC1/SC2/WG2 and UTC****Date: 2003-08-24**

This document requests additional characters to be added to the UCS and contains the proposal summary form. It corrects errors in the N2608 published on 2003-08-13.

**A. Administrative****1. Title**

Proposal to add Georgian and other characters to the BMP of the UCS.

**2. Requester's name**

NSAI/ICTSCC/SC4 (Ireland) and GSDIT (Georgia)

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

Member body contribution.

**4. Submission date**

2003-08-24

**5. Requester's reference (if applicable)**

N1962 (1999-02-26), L2/03-087 (2003-03-03)

**6. Choose one of the following:****6a. This is a complete proposal**

Yes.

**6b. More information will be provided later**

No.

**B. Technical – General****1. Choose one of the following:****1a. This proposal is for a new script (set of characters)**

NO.

**Proposed name of script****1b. The proposal is for addition of character(s) to an existing block**

Yes.

**1b. Name of the existing block**

Georgian, Phonetic Extensions, Cyrillic, and a new block, Georgian Supplement.

**2. Number of characters in proposal**

(43) 3 + 1 + 1 + 38

**3. Proposed category (see section II, Character Categories)**

Category B.1.

**4a. Proposed Level of Implementation (1, 2 or 3) (see clause 14, ISO/IEC 10646-1: 2000)**

All these characters are spacing alphabetic characters and can be used at level 1.

**4b. Is a rationale provided for the choice?**

No.

**4c. If YES, reference****5a. Is a repertoire including character names provided?**

Yes.

**5b. If YES, are the names in accordance with the character naming guidelines in Annex L of ISO/IEC 10646-1: 2000?**

Yes.

**5c. Are the character shapes attached in a legible form suitable for review?**

Yes.

**6a. Who will provide the appropriate computerized font (ordered preference: True Type, or PostScript format) for publishing the standard?**

The late Anton Dumbadze via the Georgian State Department of Information Technology 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 the Unicode Standard.

**6b. If available now, identify source(s) for the font (include address, e-mail, ftp-site, etc.) and indicate the tools used:**

Michael Everson, Fontographer.

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

Yes, see bibliography below.

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

Yes.

**8. 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, see below.

**9. 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.**

Yes, see Unicode properties below.

## **C. Technical – Justification**

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

Yes, N1962 and a number of documents submitted to UTC, including L2/03-087.

**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?**

David Tarkhan-Mouravi and Irakli Baramidze (Georgian State Department of Information Technology), Irakli Garibashvili (Georgian State Library), Jost Gippert (TITUS Projekt)

**2c. If YES, available relevant documents**

L2/03-087

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

Scholarly, ecclesiastical, and library communities. 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.

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

Originally used to write Georgian, particularly in ecclesiastical and historical documents.

**4b. Reference**

See below.

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

Yes.

**5b. If YES, where?**

By scholars, ecclesiastical researchers, and librarians.

**6a. After giving due considerations to the principles in Principles and Procedures document (a WG 2 standing document) must the proposed characters be entirely in the BMP?**

Yes.

**6b. If YES, is a rationale provided?**

Yes.

**6c. If YES, reference**

Accordance with the Roadmap, and casing relationship of Khutsuri Nuskhuri with Khutsuri Asomtavruli.

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

Yes, they should be encoded in a single block as presented here.

**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**

**12a. Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?**

No.

**12b. If YES, reference**

**13a. Does the proposal contain characters with any special properties such as control function or similar semantics?**

No.

**13b. If YES, describe in detail (include attachment if necessary)**

**14a. Does the proposal contain any Ideographic compatibility character(s)?**

No.

**14b. If YES, is the equivalent corresponding unified ideographic character(s) identified?**

**14c. If YES, reference**

## **D. Proposal**

**The Nuskhuri case-form of the Khutsuri 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” of the “Georgian” script, this is incorrect, and Nuskhuri should be disunified from Mkhedruli.

There are two related scripts in use in Georgia: Mkhedruli, which is caseless, and is used for ordinary Georgian text of all kinds, and Khutsuri, which is obsolete but for ecclesiastical purposes, and is largely unknown and unrecognizable to most Georgians. Khutsuri has two case forms: the older Asomtavruli or “monumental” writing is used as an uppercase form of the Nuskhuri miniscule.

Georgian is *not* a single “tricameral” script as some have suggested, though in Unicode and ISO/IEC 10646 the word is used to subsume the two related scripts. The Mkhedruli script is caseless, and the Khutsuri script is casing. It is accidental that the Unicode Standard and ISO/IEC 10646 subsume the two under the name of “the Georgian script”. Because “Georgian” is the script name used in the character names, we propose the preservation of this anomaly, so Mkhedruli = GEORGIAN LETTER, Asomtavruli = GEORGIAN CAPITAL LETTER, and Nuskhuri = GEORGIAN SMALL LETTER. We do however propose that the correct case-mapping behaviour be implemented, as can be seen below in the section on Unicode properties.

Monumental Asomtavruli (Figures 1, 3) was originally carved in stone, like monumental Greek or Latin. It is analogous to Greek and Latin capitals. A cursive variety of this script, called Nuskhuri (Figures 2, 3), was derived from Asomtavruli in handwritten manuscripts. This is analogous to Greek and Latin

minuscules. Just as with Greek and Latin, Asomtavruli and Nuskhuri letters came to be used together as a regularly casing pair.

Mkhedruli, however, is caseless, and no casing behaviour is expected or permitted by Georgian users. The *mtavruli* titling style of Mkhedruli (Figure 4) is not case; it is a style analogous to small caps or bold or italic. 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. ძველი 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, newspaper headlines, and other kinds of headings. It might be suggested that non-plain text SMALL CAPS coding be used to represent Mkhedruli in *mtavruli*-style. (There is no other such small caps category that could be applied to Mkhedruli.)

Manuscripts exist in which the three scripts are used on the same page (see Figure 5 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 as lower-case Khutsuri 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 has informed us that that *most* documents in Georgian were written in Nuskhuri until the 19th century CE, and that currently “the church is publishing more (than several years ago) documents in Nuskhuri (‘for internal use’).” Garibashvili has also pointed out 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. (This is also not dissimilar to what one finds in ecclesiastical Latin and Greek from certain periods.) The user inputs what is required.

L2/03-087 is an additional request from the Georgian State Department of Information Technology for disunification of Nuskhuri from Mkhedruli is proposed. This document contains the letter, dated January 2001, from Ilia II, the Catholicos Patriarch of all Georgia, which indicates clearly that Georgians do not consider Nuskhuri to be a font variant of Mkhedruli. It indicates that the two scripts do not stand in one-to-one relationship to each other insofar as they are used differently in terms of orthography. A Nuskhuri text may contain abbreviations which make no sense in Mkhedruli, as well as ligatures which do not exist in Mkhedruli. The Patriarch’s letter indicates that the Georgian Orthodox Church is the primary user group for this script, and that the Georgian Church – in addition to Jost Gippert – consider it important to be able to represent a single text with all three scripts in plain text.

Georgian users working on ecclesiastical texts in Asomtavruli and Nuskhuri expect casing behaviour no different from that of Latin and Greek. Currently the Unicode Standard does not provide this functionality because Asomtavruli and Mkhedruli do not stand in a casing relationship with one another (as they should not).

The Unicode Standard states:

The Unicode Standard regards Glagolitic as a separate script from Cyrillic, not as a font change from Cyrillic. This position is taken primarily because Glagolitic appears unrecognizably different from Cyrillic, and secondarily because Glagolitic has not grown to match the expansion of Cyrillic.

This is exactly the same situation with Nuskhuri. It is unrecognizably different from Mkhedruli, and it has not grown to match the expansion of Mkhedruli – a number of new letters have been added to Mkhedruli (and not all of them are yet encoded in the Unicode Standard) and, while it is possible to devise Nuskhuri letterforms for them, this is entirely artificial and may even be considered incorrect by Georgians. (It is true that there was an early 20th-century attempt to use Asomtavruli as capitals with Mkhedruli. This was promulgated by Prof. A. G. Shanidze; it was not popularly accepted, though it has been found in works by him and in works on the Georgian script by others honouring him (e.g.

ჯეოსკანდრელი 1970). It is an anomaly, and does not affect the true casing relation of Asomtavruli and Nuskhuri in the Khutsuri script.)

Character names should follow existing practice. For instance, Mkhedruli letters are named GEORGIAN LETTER AN and Asomtavruli Khutsuri letters are named GEORGIAN CAPITAL LETTER AN; Nuskhuri Khutsuri letters should be named GEORGIAN SMALL LETTER AN, and so on.

**Three Mkhedruli characters need to be added to the UCS.** 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]). 10FA Ⴁ GEORGIAN LETTER AIN (a name coined to reflect the letter’s own derivation from ARABIC LETTER AIN) is used in transcriptions of Bats. 10FC Ⴃ MODIFIER LETTER GEORGIAN NAR is widely used to indicate nasalization of a preceding vowel in many standard linguistic transcriptions of Caucasian languages. Glyphs and proposed code positions:

- 10F9    Ⴀ    GEORGIAN LETTER TURNED GAN (Figs. 8, 9)
- 10FA    Ⴁ    GEORGIAN LETTER AIN (Fig. 6)
- 10FC    Ⴃ    MODIFIER LETTER GEORGIAN NAR (Figs. 7, 9)

**One Latin character needs to be added to the UCS.** It is also necessary to add 1D6C Ⴄ LATIN SMALL LETTER TURNED G as this is used in Latin transliteration of the GEORGIAN LETTER TURNED GAN.

- 1D6C    Ⴄ    LATIN SMALL LETTER TURNED G (Fig. 8)

**One Cyrillic character needs to be added to the UCS.** It is also necessary to add 04CF <sup>h</sup> MODIFIER LETTER CYRILLIC EN: the nasal [ã] is variously transcribed as *a<sup>n</sup>*, *Ⴅ<sup>6</sup>*, and *a<sup>n</sup>* in Latin, Georgian, and Cyrillic Caucasian linguistic texts (207F SUPERSCRIPT LATIN SMALL LETTER N is already encoded). Glyph and proposed code position:

- 04CF    <sup>h</sup>    MODIFIER LETTER CYRILLIC EN (Fig. 7)

## Unicode Character Properties

```
04CF;MODIFIER LETTER SMALL CYRILLIC EN;Lm;0;L;<super> 041D;;;N;;;;;
10F9;GEORGIAN LETTER TURNED GAN;Lo;0;L;;;;N;;;
10FA;GEORGIAN LETTER AIN;Lo;0;L;;;;N;;;
10FC;MODIFIER LETTER GEORGIAN NAR;Lo;0;L;;;;N;;;
1D6C;LATIN SMALL LETTER TURNED G;Ll;0;L;;;;N;;;
2D00;GEORGIAN SMALL LETTER AN;Ll;0;L;;;;N;Khutsuri;;
2D01;GEORGIAN SMALL LETTER BAN;Ll;0;L;;;;N;Khutsuri;;
2D02;GEORGIAN SMALL LETTER GAN;Ll;0;L;;;;N;Khutsuri;;
2D03;GEORGIAN SMALL LETTER DON;Ll;0;L;;;;N;Khutsuri;;
2D04;GEORGIAN SMALL LETTER EN;Ll;0;L;;;;N;Khutsuri;;
2D05;GEORGIAN SMALL LETTER VIN;Ll;0;L;;;;N;Khutsuri;;
2D06;GEORGIAN SMALL LETTER ZEN;Ll;0;L;;;;N;Khutsuri;;
2D07;GEORGIAN SMALL LETTER TAN;Ll;0;L;;;;N;Khutsuri;;
2D08;GEORGIAN SMALL LETTER IN;Ll;0;L;;;;N;Khutsuri;;
2D09;GEORGIAN SMALL LETTER KAN;Ll;0;L;;;;N;Khutsuri;;
2D0A;GEORGIAN SMALL LETTER LAS;Ll;0;L;;;;N;Khutsuri;;
2D0B;GEORGIAN SMALL LETTER MAN;Ll;0;L;;;;N;Khutsuri;;
2D0C;GEORGIAN SMALL LETTER NAR;Ll;0;L;;;;N;Khutsuri;;
2D0D;GEORGIAN SMALL LETTER ON;Ll;0;L;;;;N;Khutsuri;;
2D0E;GEORGIAN SMALL LETTER PAR;Ll;0;L;;;;N;Khutsuri;;
2D0F;GEORGIAN SMALL LETTER ZHAR;Ll;0;L;;;;N;Khutsuri;;
2D10;GEORGIAN SMALL LETTER RAE;Ll;0;L;;;;N;Khutsuri;;
```

2D11;GEORGIAN SMALL LETTER SAN;Ll;0;L;;;;N;;Khutsuri;;;
2D12;GEORGIAN SMALL LETTER TAR;Ll;0;L;;;;N;;Khutsuri;;;
2D13;GEORGIAN SMALL LETTER UN;Ll;0;L;;;;N;;Khutsuri;;;
2D14;GEORGIAN SMALL LETTER PHAR;Ll;0;L;;;;N;;Khutsuri;;;
2D15;GEORGIAN SMALL LETTER KHAR;Ll;0;L;;;;N;;Khutsuri;;;
2D16;GEORGIAN SMALL LETTER GHAN;Ll;0;L;;;;N;;Khutsuri;;;
2D17;GEORGIAN SMALL LETTER QAR;Ll;0;L;;;;N;;Khutsuri;;;
2D18;GEORGIAN SMALL LETTER SHIN;Ll;0;L;;;;N;;Khutsuri;;;
2D19;GEORGIAN SMALL LETTER CHIN;Ll;0;L;;;;N;;Khutsuri;;;
2D1A;GEORGIAN SMALL LETTER CAN;Ll;0;L;;;;N;;Khutsuri;;;
2D1B;GEORGIAN SMALL LETTER JIL;Ll;0;L;;;;N;;Khutsuri;;;
2D1C;GEORGIAN SMALL LETTER CIL;Ll;0;L;;;;N;;Khutsuri;;;
2D1D;GEORGIAN SMALL LETTER CHAR;Ll;0;L;;;;N;;Khutsuri;;;
2D1E;GEORGIAN SMALL LETTER XAN;Ll;0;L;;;;N;;Khutsuri;;;
2D1F;GEORGIAN SMALL LETTER JHAN;Ll;0;L;;;;N;;Khutsuri;;;
2D20;GEORGIAN SMALL LETTER HAE;Ll;0;L;;;;N;;Khutsuri;;;
2D21;GEORGIAN SMALL LETTER HE;Ll;0;L;;;;N;;Khutsuri;;;
2D22;GEORGIAN SMALL LETTER HIE;Ll;0;L;;;;N;;Khutsuri;;;
2D23;GEORGIAN SMALL LETTER WE;Ll;0;L;;;;N;;Khutsuri;;;
2D24;GEORGIAN SMALL LETTER HAR;Ll;0;L;;;;N;;Khutsuri;;;
2D25;GEORGIAN SMALL LETTER HOE;Ll;0;L;;;;N;;Khutsuri;;;

In addition, add Lowercase mappings from Georgian capital letters (10A0-10C5) to Georgian small letters (2D00-2D25), and add Uppercase and titlecase mappings for Georgian small letters (2D00-2D25) to Georgian capital letters (10A0-10C5).

In **ScriptNames.txt**, add the script Khutsuri, and make the following specific annotation in order to make clear the name versus script identity anomaly built into Georgian in ISO/IEC 10646 and the Unicode Standard:

The term “Georgian” in the Unicode Standard refers to two distinct but related scripts. The older Khutsuri script is composed of characters named “Georgian capital letters” also known as “Asomtavruli”, and of “Georgian small letters” also known as “Nuskhuri”. The script used for ordinary modern Georgian is called “Mkhedruli” and is composed of “Georgian letters” in the Unicode Standard.

Normative casemapping must be backrevved into **UnicodeData.txt** for Asomtavruli, despite its removal by explicit decision of the UTC for Unicode 3.0. (Note that the removal set the stage for the separate encoding of Nuskhuri, by removing the mappings to what is now only to be caseless Mkhedruli, so this backrevving makes very good sense.) In **CaseFolding.txt**, the Georgian capital letters should fold to the Georgian small letters.

## Bibliography

- Haarmann, Harald. 1990. *Die Universalgeschichte der Schrift*. Frankfurt: Campus Verlag. ISBN 3-593-34346-0
- 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.
- Магометов, Александр Амарович (= ალ. მაჰომეტოვი). 1965. *Табасаранский язык: исследование и тексты* (= ტაბასარანუმი ენა: გამოკვლევა და ტექსტები). Тбилиси: Мецниереба.
- ლომთათიძე, ქ. 1976. *აფხაზური და აბაზური ენების ისტორიულ-შედარებითი ანალიზი*. თბილისი: მეცნიერება.
- ჭლექსანდრელი, ოთახასი (= Афанасий Александрійский). 1970. *სწორებაჲ წმიდისა წმკონისი*. (= Афанасий Александрійский) *Житие Святого Антония*). Eds. ო. შანიძე & ო. შანიძე (შანიძე იმნაიშვილი) (= A. G. Shanidze & V. I. Imnaishvili). თბილისი: თბილისის უნივერსიტეტის გამომცემლობა.

# Figures

1. — ღოზო უნოა ქათუნებს შოყდნე ყაფ-  
 უნებს უნოქათუნდენენენ. თოთოქ უნოქოთი  
 კოფუნდოქ. ზუნე კოლოაენი ღოქოსენენენი  
 კოქიფუნ უნოქათუნდენენ თუნე უნოქოთი :

2. — ნდუნე ღოქ კოქოქი უნე უნენ. კოქოქი  
 10 უნოქათუნდენ? კოქოქოქი უნენენი უნქათუნ-  
 უნენი კინე ღოქოსენენენი თუნე კოქიფუნ-  
 უნქათუნდენენ-გოქოქი კინე.

Figure 1. Sample of Asomtavruli-only text, from Marr & Brière 1931.

11. — შოქიში უნქოქი ზუნე თოქოქი : კუნენე უნქი უნქი  
 თი თი.

12. — ზუნე თოქოქი უნქი თი ში ში თი თი თი თი :  
 ში თი. ში თი თი თი თი თი თი თი თი თი თი თი თი  
 თი თი თი თი თი თი თი თი თი თი თი თი თი თი თი

Figure 2. Sample of Nuskhuri-only text, from Marr & Brière 1931.

ბოქი ქოქი თი თი თი თი თი თი თი თი თი თი თი თი თი  
 თი თი თი თი თი თი თი თი თი თი თი თი თი თი თი თი  
 თი თი თი თი თი თი თი თი თი თი თი თი თი თი თი თი  
 თი თი თი თი თი თი თი თი თი თი თი თი თი თი თი თი  
 თი თი თი თი თი თი თი თი თი თი თი თი თი თი თი თი  
 თი თი თი თი თი თი თი თი თი თი თი თი თი თი თი თი

Figure 3. Sample of Nuskhuri text with Asomtavruli capitals, from Haarmann 1990.

შინაარსი	
შესავალი	3
I. აფხაზურისა და აბაზურის ურთიერთობა და დიალექტური უბედ- გენილობა	7
II. აფხაზური და აბაზური სამწერლობო ენები	15
1. აფხაზური სამწერლობო ენა	15
2. აბაზური სამწერლობო ენა	21

Figure 4. Sample of Mkhedruli text in plain and mtavruli style, from ლომთათიძე 1976.

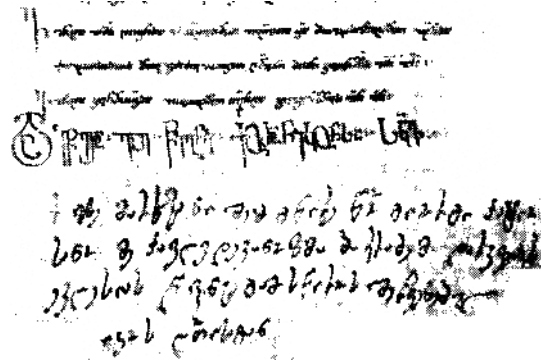


Figure 5. Sample of Nuskhuri, Asomtavruli, and Mkhedruli on one page in a manuscript. From a scan on the TITUS Projekt web site.

ჰალდარ-დარ ჰალ-  
 wir-d-ar ო.ს. სობ-დარ  
 wir-d-ar.  
 ჰალდარ-დ-ითარ ჰალ  
 wir-d-itare ო.ს. სობ-  
 დ-ითარ wir-d-itare. ნუ-

Figure 6. Sample Bats text showing MODIFIER LETTER INTENSIVE PHARYNGEALIZATION used with Georgian and Latin scripts, as well as GEORGIAN LETTER AIN, in Кадагидзе & Кадагидзе 1984. This dictionary, the only Bats dictionary published to date, was reproduced from a handwritten manuscript. The first two words read ჰალდარ-დარ and *hʷalwir-d-ar*.

1. Чеченская			2. Ингушская			2. Бацбийская	
Латинская	Грузинская	Лит. чеч.	Латинская	Грузинская	Лит. ингуш.	Латинская	Грузинская
a	ა	a	a	ა	a	a	ა
ა	ა	-	ა	ა	-	ა	ა
ä	ა̇	ä	-	-	-	-	-
ä̇	ა̇̇	ä̇	ä̇	ა̇̇	ä̇	-	-
a <sup>n</sup>	ა <sup>ნ</sup>	(a <sup>n</sup> )	-	-	-	a <sup>n</sup>	ა <sup>ნ</sup>
a <sup>n</sup>	ა <sup>ნ</sup>	-	-	-	-	-	-
b	ბ	б	b	ბ	б	b	ბ
g	გ	г	g	გ	г	g	გ
d	დ	д	d	დ	д	d	დ
e	ე	e	e	ე	e	e	ე
e <sup>n</sup>	ენ	-	-	-	-	e <sup>n</sup>	ენ

Figure 7. 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 Имнашвили 1977.



Латин.	Грузин.	Табасаран.	Латин.	Грузин.	Табасаран.
a	ა	а	s	ს	с
ä	ა̇	аь	sa	სა	
b	ბ	б	t	თ	т
w	უ	в	ṭ	ტ	т̣
g	გ	г	ta	თა	т; тт
g'	გ'	г	u	უ	у
d	დ	д	ü	უ̇	уь
e	ე	е	ü'	უ̇'	уь'
ž	ჯ	ж	f	ფ	ф
ž'	ჯ'	жв	fa	ფა	
ž̇	ჯ̇	ж̇	x	ხ	х
ž̇'	ჯ̇'	ж̇ь	xa	ხა	
z	ზ	з	x'	ხ'	хь
		з	x'̇	ხ'̇	
i	ი	и	e	ე	ц
j	ი̇	й	ė	ე̇	ц̇
k	კ	к	ca	ცა	ц; цц
ḳ	კ̣	к̣	č	ჩ	ч
ka	კა	к; кк	č'	ჩ'	ч'
γ	ღ	гь	ča	ჩა	ч̇
q	ყ	хь	č'	ჩ'	ч̇'
qa	ყა	кь	č̇	ჩ̇	ж̇
qa'	ყა'	кь; кьь	č̇'	ჩ̇'	ш̇
s	ს	кь	š	შ	ш
l	ლ	л	ša	შა	ш̇
m	მ	м	š'	შ'	ш̇'
n	ნ	н	ṧ	შ̇	
p	პ	п	ṧ'	შ̇'	
p̣	პ̣	п̣	ha	ჰა	
pa	პა	п; пп	h'	ჰ'	гь
r	რ	р	ḣ	ჰ̇	ь

**Figure 8.** Transcription of Tabasaran in Latin and Georgian transcription and Cyrillic orthography, showing Ⴑ GEORGIAN LETTER INVERTED GAN, Ⴒ MODIFIER LETTER PALATALIZATION, Ⴓ MODIFIER LETTER INTENSIVITY, Ⴔ MODIFIER LETTER LABIALIZATION, Ⴕ MODIFIER LETTER TURNED COMMA and Ⴖ MODIFIER LETTER APOSTROPHE, in Магомедов 1965.



GEORGIAN · GEORGIAN SUPPLEMENT · PHONETIC EXTENSIONS · CYRILLIC

	10A	10B	10C	10D	10E	10F	2D0	2D1	2D2	1D6C
0	Ⴒ	Ⴓ	Ⴔ	Ⴕ	Ⴖ	Ⴗ	Ⴘ	Ⴙ	Ⴚ	Ⴛ
1	Ⴜ	Ⴝ	Ⴞ	Ⴟ	Ⴀ	Ⴁ	Ⴂ	Ⴃ	Ⴄ	Ⴅ
2	Ⴆ	Ⴇ	Ⴈ	Ⴉ	Ⴊ	Ⴋ	Ⴌ	Ⴍ	Ⴎ	04CF
3	Ⴏ	Ⴐ	Ⴑ	Ⴒ	Ⴓ	Ⴔ	Ⴕ	Ⴖ	Ⴗ	H
4	Ⴘ	Ⴙ	Ⴚ	Ⴛ	Ⴜ	Ⴝ	Ⴞ	Ⴟ	Ⴀ	
5	Ⴁ	Ⴂ	Ⴃ	Ⴄ	Ⴅ	Ⴆ	Ⴇ	Ⴈ	Ⴉ	
6	Ⴊ	Ⴋ		Ⴌ	Ⴍ	Ⴎ	Ⴏ	Ⴐ		
7	Ⴑ	Ⴒ		Ⴓ	Ⴔ	Ⴕ	Ⴖ	Ⴗ		
8	Ⴘ	Ⴙ		Ⴚ	Ⴛ	Ⴜ	Ⴝ	Ⴞ		
9	Ⴟ	Ⴀ		Ⴁ	Ⴂ	Ⴃ	Ⴄ	Ⴅ		
A	Ⴆ	Ⴇ		Ⴈ	Ⴉ	Ⴊ	Ⴋ	Ⴌ		
B	Ⴏ	Ⴐ		Ⴒ	Ⴓ	Ⴔ	Ⴕ	Ⴖ		
C	Ⴘ	Ⴙ		Ⴚ	Ⴛ	Ⴜ	Ⴝ	Ⴞ		
D	Ⴟ	Ⴀ		Ⴁ	Ⴂ	Ⴃ	Ⴄ	Ⴅ		
E	Ⴆ	Ⴇ		Ⴈ	Ⴉ	Ⴊ	Ⴋ	Ⴌ		
F	Ⴏ	Ⴐ		Ⴒ	Ⴓ	Ⴔ	Ⴕ	Ⴖ		

## GEORGIAN • GEORGIAN SUPPLEMENT • PHONETIC EXTENSIONS • CYRILLIC

hex	Name	hex	Name
10A0	GEORGIAN CAPITAL LETTER AN (Khutsuri)	10F9	GEORGIAN LETTER TURNED GAN
10A1	GEORGIAN CAPITAL LETTER BAN (Khutsuri)	10FA	GEORGIAN LETTER AIN
10A2	GEORGIAN CAPITAL LETTER GAN (Khutsuri)	10FB	GEORGIAN PARAGRAPH SEPARATOR
10A3	GEORGIAN CAPITAL LETTER DON (Khutsuri)	10FC	MODIFIER LETTER GEORGIAN NAR
10A4	GEORGIAN CAPITAL LETTER EN (Khutsuri)	10FD	(This position shall not be used)
10A5	GEORGIAN CAPITAL LETTER VIN (Khutsuri)	10FE	(This position shall not be used)
10A6	GEORGIAN CAPITAL LETTER ZEN (Khutsuri)	10FF	(This position shall not be used)
10A7	GEORGIAN CAPITAL LETTER TAN (Khutsuri)		
10A8	GEORGIAN CAPITAL LETTER IN (Khutsuri)	2D00	GEORGIAN SMALL LETTER AN (Kutsuri)
10A9	GEORGIAN CAPITAL LETTER KAN (Khutsuri)	2D01	GEORGIAN SMALL LETTER BAN (Kutsuri)
10AA	GEORGIAN CAPITAL LETTER LAS (Khutsuri)	2D02	GEORGIAN SMALL LETTER GAN (Kutsuri)
10AB	GEORGIAN CAPITAL LETTER MAN (Khutsuri)	2D03	GEORGIAN SMALL LETTER DON (Kutsuri)
10AC	GEORGIAN CAPITAL LETTER NAR (Khutsuri)	2D04	GEORGIAN SMALL LETTER EN (Kutsuri)
10AD	GEORGIAN CAPITAL LETTER ON (Khutsuri)	2D05	GEORGIAN SMALL LETTER VIN (Kutsuri)
10AE	GEORGIAN CAPITAL LETTER PAR (Khutsuri)	2D06	GEORGIAN SMALL LETTER ZEN (Kutsuri)
10AF	GEORGIAN CAPITAL LETTER ZHAR (Khutsuri)	2D07	GEORGIAN SMALL LETTER TAN (Kutsuri)
10B0	GEORGIAN CAPITAL LETTER RAE (Khutsuri)	2D08	GEORGIAN SMALL LETTER IN (Kutsuri)
10B1	GEORGIAN CAPITAL LETTER SAN (Khutsuri)	2D09	GEORGIAN SMALL LETTER KAN (Kutsuri)
10B2	GEORGIAN CAPITAL LETTER TAR (Khutsuri)	2D0A	GEORGIAN SMALL LETTER LAS (Kutsuri)
10B3	GEORGIAN CAPITAL LETTER UN (Khutsuri)	2D0B	GEORGIAN SMALL LETTER MAN (Kutsuri)
10B4	GEORGIAN CAPITAL LETTER PHAR (Khutsuri)	2D0C	GEORGIAN SMALL LETTER NAR (Kutsuri)
10B5	GEORGIAN CAPITAL LETTER KHAR (Khutsuri)	2D0D	GEORGIAN SMALL LETTER ON (Kutsuri)
10B6	GEORGIAN CAPITAL LETTER GHAN (Khutsuri)	2D0E	GEORGIAN SMALL LETTER PAR (Kutsuri)
10B7	GEORGIAN CAPITAL LETTER QAR (Khutsuri)	2D0F	GEORGIAN SMALL LETTER ZHAR (Kutsuri)
10B8	GEORGIAN CAPITAL LETTER SHIN (Khutsuri)	2D10	GEORGIAN SMALL LETTER RAE (Kutsuri)
10B9	GEORGIAN CAPITAL LETTER CHIN (Khutsuri)	2D11	GEORGIAN SMALL LETTER SAN (Kutsuri)
10BA	GEORGIAN CAPITAL LETTER CAN (Khutsuri)	2D12	GEORGIAN SMALL LETTER TAR (Kutsuri)
10BB	GEORGIAN CAPITAL LETTER JIL (Khutsuri)	2D13	GEORGIAN SMALL LETTER UN (Kutsuri)
10BC	GEORGIAN CAPITAL LETTER CIL (Khutsuri)	2D14	GEORGIAN SMALL LETTER PHAR (Kutsuri)
10BD	GEORGIAN CAPITAL LETTER CHAR (Khutsuri)	2D15	GEORGIAN SMALL LETTER KHAR (Kutsuri)
10BE	GEORGIAN CAPITAL LETTER XAN (Khutsuri)	2D16	GEORGIAN SMALL LETTER GHAN (Kutsuri)
10BF	GEORGIAN CAPITAL LETTER JHAN (Khutsuri)	2D17	GEORGIAN SMALL LETTER QAR (Kutsuri)
10C0	GEORGIAN CAPITAL LETTER HAE (Khutsuri)	2D18	GEORGIAN SMALL LETTER SHIN (Kutsuri)
10C1	GEORGIAN CAPITAL LETTER HE (Khutsuri)	2D19	GEORGIAN SMALL LETTER CHIN (Kutsuri)
10C2	GEORGIAN CAPITAL LETTER HIE (Khutsuri)	2D1A	GEORGIAN SMALL LETTER CAN (Kutsuri)
10C3	GEORGIAN CAPITAL LETTER WE (Khutsuri)	2D1B	GEORGIAN SMALL LETTER JIL (Kutsuri)
10C4	GEORGIAN CAPITAL LETTER HAR (Khutsuri)	2D1C	GEORGIAN SMALL LETTER CIL (Kutsuri)
10C5	GEORGIAN CAPITAL LETTER HOE (Khutsuri)	2D1D	GEORGIAN SMALL LETTER CHAR (Kutsuri)
10C6	(This position shall not be used)	2D1E	GEORGIAN SMALL LETTER XAN (Kutsuri)
10C7	(This position shall not be used)	2D1F	GEORGIAN SMALL LETTER JHAN (Kutsuri)
10C8	(This position shall not be used)	2D20	GEORGIAN SMALL LETTER HAE (Kutsuri)
10C9	(This position shall not be used)	2D21	GEORGIAN SMALL LETTER HE (Kutsuri)
10CA	(This position shall not be used)	2D22	GEORGIAN SMALL LETTER HIE (Kutsuri)
10CB	(This position shall not be used)	2D23	GEORGIAN SMALL LETTER WE (Kutsuri)
10CC	(This position shall not be used)	2D24	GEORGIAN SMALL LETTER HAR (Kutsuri)
10CD	(This position shall not be used)	2D25	GEORGIAN SMALL LETTER HOE (Kutsuri)
10CE	(This position shall not be used)	2D26	(This position shall not be used)
10CF	(This position shall not be used)	2D27	(This position shall not be used)
10D0	GEORGIAN LETTER AN	2D28	(This position shall not be used)
10D1	GEORGIAN LETTER BAN	2D29	(This position shall not be used)
10D2	GEORGIAN LETTER GAN	2D2A	(This position shall not be used)
10D3	GEORGIAN LETTER DON	2D2B	(This position shall not be used)
10D4	GEORGIAN LETTER EN	2D2C	(This position shall not be used)
10D5	GEORGIAN LETTER VIN	2D2D	(This position shall not be used)
10D6	GEORGIAN LETTER ZEN	2D2E	(This position shall not be used)
10D7	GEORGIAN LETTER TAN	2D2F	(This position shall not be used)
10D8	GEORGIAN LETTER IN		
10D9	GEORGIAN LETTER KAN	1DC6	LATIN SMALL LETTER TURNED G
10DA	GEORGIAN LETTER LAS		
10DB	GEORGIAN LETTER MAN	04CF	MODIFIER LETTER SMALL CYRILLIC EN
10DC	GEORGIAN LETTER NAR		
10DD	GEORGIAN LETTER ON		
10DE	GEORGIAN LETTER PAR		
10DF	GEORGIAN LETTER ZHAR		
10E0	GEORGIAN LETTER RAE		
10E1	GEORGIAN LETTER SAN		
10E2	GEORGIAN LETTER TAR		
10E3	GEORGIAN LETTER UN		
10E4	GEORGIAN LETTER PHAR		
10E5	GEORGIAN LETTER KHAR		
10E6	GEORGIAN LETTER GHAN		
10E7	GEORGIAN LETTER QAR		
10E8	GEORGIAN LETTER SHIN		
10E9	GEORGIAN LETTER CHIN		
10EA	GEORGIAN LETTER CAN		
10EB	GEORGIAN LETTER JIL		
10EC	GEORGIAN LETTER CIL		
10ED	GEORGIAN LETTER CHAR		
10EE	GEORGIAN LETTER XAN		
10EF	GEORGIAN LETTER JHAN		
10F0	GEORGIAN LETTER HAE		
10F1	GEORGIAN LETTER HE		
10F2	GEORGIAN LETTER HIE		
10F3	GEORGIAN LETTER WE		
10F4	GEORGIAN LETTER HAR		
10F5	GEORGIAN LETTER HOE		
10F6	GEORGIAN LETTER FI		
10F7	GEORGIAN LETTER SHVA		
10F8	GEORGIAN LETTER ELIFI		