

Vazgen Ghazaryan. *A Proposal for Pan-Caucasian Alphabet*.  
Yerevan, 2012, ISBN 978-9939-0-0603-1.

## **A PROPOSAL FOR PAN-CAUCASIAN ALPHABET**

(Synopsis)

The Caucasus or *Jabal Al-Asun* ("Mountain of Tongues"), as it was named in the annals of medieval Arab geographers, is one of the most prominent places on Earth by the diversity and density of concentration of languages. The region is home to more than 50 nations, which belong to five language families. The three of those, namely, the *Northwest Caucasian* (West-Caucasian or Abkhaz-Adyghean), *Northeast Caucasian* (East Caucasian or Nakh-Dagestanian) and *South Caucasian* (Kartvelian) families are confined within the limits of the Greater and Lesser Caucasus Mountain Ranges and currently comprise 41 languages considered native to the region for at least four thousand years.

Located amidst the sphere of distribution, contiguity and mutual influence of the Indo-European, Altaic and Semitic language families, the Caucasus, presumably due to the peculiarities of the landform and relative isolation of mountainous regions, was able to preserve a certain degree of representativeness and sufficient diversity of the three autochthonous language families, preventing their extinction or reduction to a language isolate.

Nevertheless, the historical and political realities, established in the region during the latest 200 years, gravely threaten the means of development of these languages and even their future existence. For instance, from the mid-19<sup>th</sup> century the majority of

native speakers of the NWC languages settled beyond the limits of the Caucasus, being almost completely deported in consequence of imperial expansion. As a result, this once multilingual family that was widely and homogeneously spread all over the Western Caucasus, has presently diminished to a few isolated enclaves with a shrinking population. The situation is somewhat different in the Eastern Caucasus, but there also a number of NEC languages presently count a few thousand of even a few hundred speakers, and being enclosed within one or a few villages, are reduced to a definitely or severely endangered state.

The uncertainty of the future of those North Caucasian languages (NWC and NEC together), which are endangered or under an imminent threat of extinction, is not only stipulated by the historical, cultural, political and confessional realities of the Caucasus, but also by the absence of one of the most essential components that shape the identity of ethnic communities – a centuries-old written tradition and heritage, that places these generally small groups into a vulnerable position against cultural and especially political and religious expansion.

Among the languages of three native families of the Caucasus, Georgian is the only one to have an uninterrupted centuries-old written tradition, which originated, by the prevailing opinion, in the 5th century by virtue of the invention of an authentic alphabet. From other Caucasian languages, the presently extinct language of Caucasian Albania – the so-called "Caucasian Albanian" or "Aluanian", "Aghuan", "Aghvan", which belongs to the Lezgian branch of NEC languages and is considered to be the direct ancestor of modern Udi, have had an original written tradition and literature (at least in translation). During the Middle Ages – in the period of the spread of Islam, an attempt has been made to adapt a modified version of the Arabic script to some NEC languages. However, due to various reasons it did not turn fully successful in this linguistic area and the Arabic script gained limited use.

Standardized writing systems for the North Caucasian languages have been implemented only in the 20<sup>th</sup> century. Initially based upon the Latin script, the adapted alphabets have been shifted to Cyrillic-shaped graphics during the mid 30s. Unfortunately, both versions do not match in practice with the phonetics of the Caucasian languages since these writing systems originated or have been modified to express totally unlike phonetic structures of the source languages. They are extremely deficient for the purposes they were intended for, and present serious difficulties not only for graphical harmony and aesthetics of composition, but also for their practical application and functionality. More specifically, the limited number of characters and their insufficient flexibility in graphical modifications make these writing systems incapable to represent, in an unambiguous way, the phonetics of the North Caucasian languages, which in their turn possess an outstanding feature of having one of the richest consonant inventories among all the languages of the world.

For instance, the language of the Ubykhs (extinct since 1992) has 86 consonants and two vowels; the Archi language, presently reduced to 1200 speakers, distinguishes 81 consonants and 26 vowels (many of the former do not have exact correspondences in other languages); the consonant inventory of the Bzyp dialect of Abkhaz includes 68 phonemes, etc.

All Caucasian languages have a regular three-level phonation for stops and affricates (voiced, ejective and aspirated voiceless), whereas the Cyrillic script distinguishes only two levels of phonation (voiced and unaspirated voiceless) in case of stops and one (aspirated voiceless) in case of affricates. The Latin alphabet does not represent affricates at all. However, if the Latin alphabet, with all its drawbacks, acquired a more or less conventional set of auxiliary signs and supplementary characters due to a wide usage among a vast number of languages during two millennia, the Cyrillic script until the 20<sup>th</sup> century has chiefly been confined to

a limited range of a few Slavic languages sharing similar phonetic traits. And, since writing systems containing nearly 30 graphemes are essentially unfit for full phonemic rendition with simultaneous preservation of the alphabetic principle (i.e. one-to-one correspondence between graphemes and phonemes) in case of languages possessing 50 phonemes in average, dissimilar sets of symbols often with illogical combinations were introduced into the alphabets of particular North Caucasian languages during the adaptation of the Russian script.

Another important reason why these adaptations fail to be consistent is that this script has evolved in the environment of Russian language and therefore is specifically intended to express its pronunciation peculiarities alone. It possesses only two means of alteration of articulation – palatalization, expressed by the sign *ь*, and depalatalization, which is mostly omitted in current orthography, but if necessary, can be marked by the sign *ѣ*. Meanwhile, the North Caucasian languages feature four different types of coarticulation – labialization, palatalization, pharyngealization and labiopharyngealization. Among these only the palatalization can regularly be expressed using the Russian script. As a result, the other specifics of articulation are either expressed by graphemes, which in their turn have individual phonetic values (such as the use of the letters *в* /v/ and *у* /u/ to mark the labialization in most Caucasian languages), or totally omitted in the script.

For instance, the series of Ubykh postalveolar affricates and fricatives numbers 22 distinct phonemes, for which there is only one (!) correspondence in the Cyrillic script – the letter *ц* (tc). Obviously, this one sign alone is quite insufficient to express the overall phonemic diversity of this extensive series by means of Cyrillic graphics. Set aside the letter *х*, there is no other direct or indirect graphical correspondence for uvular, pharyngeal, epiglottal and glottal stops, affricates, fricatives and sonorants, which in Caucasian languages abound.

As one can see, the quantity of phonemes of these languages by far exceeds the graphical capabilities of all alphabetic systems that have previously been proposed for them or are currently in use. Such a vast phoneme inventory significantly hinders the possibilities of adaptation of any actual script and constitutes the prime reason for the current project.

So, it becomes evident that the structural deficiency of the Latin and Cyrillic writing systems in regard to their adaptation to the North Caucasian languages had to cause unnecessary functional complications in expressing these languages by means of the adapted alphabets.

This can clearly be noticed from the practice of overuse of constrained di-, tri-, tetra- and pentagraphs (*xxьIв, κIκIв, чIчIв, xxьI, лIьв, зьIв, кхьу, κκв, ххв, κьь, зьу, хьу, κIу, жьу* etc.), which not only greatly confuse the orthography, but also undermine the very base of the alphabetic principal – the systematic and predictable relationship between written letters, symbols and spoken words.

Note, that a moderate number of di- and trigraphs does not raise many difficulties for the practical use of a particular writing system. In German, for instance, there is one trigraph (*sch*) and even one tetragraph (*tsch*), both of which do not perplex the perception of a written text.

However, a certain limit exists that, if passed, may lead to a serious distortion or even a violation of the alphabetic principle.

This is the case in almost all North Caucasian Cyrillic alphabets, where the quantity of compound combinations often exceeds 50% of the total letter inventory. For example, 24 out of 69 characters of Adyghe alphabet are double, and 11 – triple, making in sum 35 compounds, which is more than half of the total listing with 69:35 ratio.

The similar statistics of the other Caucasian Cyrillic alphabets is as follows: Abaza (74:40), Kabardian (55:25), Abkhaz (64:24),

Akhvakh (56:30), Aghul (69:32), Avar (53:37), Lak (59:25), Tabasaran (59:25), Tsez (40:14), Chechen (45:16) etc.

The Chechen alphabet alone having merely 45 characters in the presence of 44 authentic vowels and diphthongs in the language itself, clearly convinces one in the fact that even at the cost of universal violation of the alphabetic principle and inappropriate complication of orthography, the given alphabets are unable to express the phonemic structure of the Caucasian languages even in the least satisfactory manner.

Hence, it is no surprise that the superposition of the Russian script over the Caucasian languages resulted in a number of grammatical discrepancies. One can also add that during the adaptation procedure an essentially substandard approach was shown in choosing the characters and their combinations that fails to represent in a similar way the orthographies of even closely related languages. The inconsistent differences are such that a few years ago a proposal to unify the orthographies of two literary versions of Circassian – Adyghe and Kabardian, was presented. The Parliament of the Kabardino-Balkar Republic discussed the issue, but with no progress whatsoever.

All above-mentioned complications essentially limited the means of graphical expression of these languages and led to a point, where, set aside rare dialectal phonemes, a series of sounds of literary languages were omitted in a number of alphabets. In many cases, these very same circumstances also defined the selection of dialects upon which the literary versions of some Caucasian languages were subsequently based: neither the extent of geographical distribution, nor the greater number of speakers was the decision criterion, but the minimal consonant inventory.

In summary, we may conclude that presently for the languages of both North Caucasian families there are practically no alphabets with a satisfactory level of phonematicity. Moreover, in

Cyrillic script we deal with a quite inconsistent system of symbols, the potential of which in respect of grapheme morphology and structural correspondences is extremely low and insufficient not only for a simple, practical and phonemically complete rendition of the North Caucasian languages, but also for their aesthetic graphical representation.

The latter is of paramount importance for psychology of writing and determines the representativeness and competitiveness of an authentic language under the dominance of another – an official language with an identical writing system. This can be exemplified in an Adyghe text that looks as follows:

*УбланІэм щегъэжъагъэу а ГущыІэр Тхъэм ыдэжъ  
щыІагъ. Тхъэм а ГущыІэм зэкІэри къыригъэгъэхъ-  
угъ... Тхъэм къыгъэхъугъэ пстэуми ащыщэу а Гу-  
щыІэм къыримыгъэгъэхъугъэ зи щыІэн.*

There is no need to prove the obvious – that a similar graphic portrayal of a language can never assist its literary development.

Besides the imperfection of the writing system or even its absence, the official status and the cultural dominance of Russian has a no less considerable impact on the marginalization of the spheres of usage of the native Caucasian languages, promoting their gradual extinction. A large number of languages, among which were examples unique by their grammatical and phonetic properties, either completely died out or are on the brink of extinction being reduced to a few hundred or thousand of speakers, such as Archi, Akhvakh, Khinalug, Khwarshi, Tsez, Hinukh, Hunzib, Bezhta, Ghodoberi, Kryts, Budukh, Udi and many others.

Granting these languages a new impetus and means for development may be a possible break through the mentioned difficulties. To achieve this goal we suggest a scientifically valid development of completely new and well-adapted Caucasian al-

phabets that are adjusted by the morphology of symbols and the logic of their modification. At the same time, they must correspond to the complex intrinsic phonetic features of these languages, render them with full phonemic representativeness and alphabetical unambiguity, and at the same time be free from any irrelevant political or cultural context and enforcing influence.

The best illustrations to the aforesaid are the 1600-year time-tested alphabets of once dominant languages of the South Caucasus and the Armenian Highlands – Armenian, Georgian and the extinct and only recently deciphered Caucasian Albanian. Among the writing systems of the world, these three alphabets are distinguished by their highest level of grapheme-to-phoneme correspondence and are listed among the phonetically most perfect. Caucasian Albanian, *inter alia*, was the only language of the South Caucasus possessing phonetic features similar to the North Caucasian languages and an ancient alphabet adapted to it.

Throughout the centuries, the viability of the Armenian and Georgian alphabets, alongside with the power of tradition and some peculiarities of the identity of these two nations, was supported by the fundamental fact: they reflect the phonetics of these languages with the highest level of perfection. These alphabets were never superseded by the Greek, Aramaic, Arabic, Russian or Latin scripts, as they weren't adopted or modified, but from the beginning developed on the basis of a meticulous scientific analysis of the phonetics of Armenian and Georgian (as well as Caucasian Albanian).

In their triunity, these remarkable examples of cultural heritage of the contemporary Caucasian region once unified the three nation-states of the South Caucasus and the Armenian Highlands in a common but mutually independent and unique writing tradition. The current project is an attempt to restore former triunity, which has been interrupted after one of its essential parts – the Caucasian Albanian alphabet became obsolete and fell into oblivion.



Hence, we set forth the idea to introduce completely original, easily legible, and most importantly – phonetically perfect and grammatologically thorough writing systems for the North Caucasian linguistic area, based on the character forms and graphical principles of construction of the alphabets of geographically adjacent and historically akin South Caucasus.

If formulated properly, this problem, however complicated it may seem, does not cause principal difficulties. With an appropriate political will and flexible approach, it might have had a satisfactory solution as early as in the beginning of the 20<sup>th</sup> century.

Thus, pooling together the available scientific research on the Caucasian phonology and summarizing the general pattern of the overall phonetic diversity of these families within single terminology, we have constructed *phonetically complete featural alphabets*, based on a peculiar usage of the morphological principles of construction of millennial South Caucasian alphabets.

For this purpose the whole grammatological and constructive potential of these writing systems was revealed and analyzed, and, based on their generalized grapheme matrix, a solid and thoroughly arranged regular system of characters was proposed. It has exceptional properties for precise and alphabetically unequivocal rendition of all phonetic peculiarities of the North Caucasian languages possessing the largest consonant inventory among all the languages of the world. The set of distinct rules for grapheme formation enables one to remember a large number of visual forms with no difficulty whatsoever.

To achieve maximum efficiency, two similar, but mutually independent generalized sets of characters for NWC and NEC language families are introduced, wherefrom the specific alphabets for the particular languages are subsequently deduced.

Additionally, for the Ossetic language alike, which is an integral part of the Caucasian heritage, an independent alphabet shar-

ing the features of both the newly developed North Caucasian and the ancient Transcaucasian alphabets has been created.

Thus, we attempt to give a new and historically sound unity and continuity to the millennial writing culture of the Caucasian region, a new impulse to the development of greater speech communities, as well as viability for preservation and future revival to smaller ones currently on the verge of extinction. Unlike all previous attempts, the developed featural alphabets are not adopted, and they do not copy the writing of any linguo-cultural, religious or political order directly or indirectly reducing to a secondary position the original and unique Caucasian languages.

They are aimed to renew and reinterpret the distinct historical, linguistic, cultural and philosophical heritage of the South and North Caucasus in a united and coherent Caucasian context.

---

To illustrate the essential ideas, we present the generalized alphabets for the NWC and NEC language families in a specially developed typeface titled *Aria Caucasica Unicode* (based on the isometric features of the Arial family) alongside with IPA transcriptions.

The specific alphabets, derived to fit the phonetics of particular Caucasian languages, can be found in the book, published in Armenian and Russian languages:

Վազգեն Դազարյան, *Համակրոկայսյան այբուբենի նախագիծ*.

Вазген Казарян. *Проект общекавказского алфавита*.

Vazgen Ghazaryan. *A Proposal for Pan-Caucasian Alphabet*. Yerevan, 2012, ISBN 978-9939-0-0603-1.



**The Generalized Northeast Caucasian Alphabet**  
*(the tense consonants and vowel articulations are not listed)*

U	Ա	Մ	Բ	Բ	Ը	Դ	Դ	Դ	Գ	Գ	Ձ	Չ
a	æ	a	p	b	p'	t	d	t'	t <sup>w</sup>	d <sup>w</sup>	t <sup>w</sup> '	c
Չ	Չ	Կ	Կ	Կ	Գ	Գ	Գ	Ը	Ը	Կ	Կ	Կ
j	c'	k	g	k'	k <sup>w</sup>	g <sup>w</sup>	k <sup>w</sup> '	ε	æ	q	G	q'
Գ	Գ	Գ	Գ	Գ	Գ	Գ	Գ	Գ	Գ	Գ	Գ	Գ
q <sup>w</sup>	G <sup>w</sup>	q <sup>w</sup> '	ʔ	ʔ	i	y	ts	dz	ts'	ts <sup>w</sup>	dz <sup>w</sup>	ts <sup>w</sup> '
Չ	Չ	Չ	Չ	Չ	Չ	Չ	Չ	Չ	Չ	Չ	Չ	Չ
tʃ	tʃ'	tʃ <sup>w</sup>	tʃ <sup>w</sup> '	tʃ	dʒ	tʃ	tʃ <sup>w</sup>	dʒ <sup>w</sup>	tʃ <sup>w</sup> '	kx	gɣ	kx'
Ի	Օ	Բ	Բ	Ը	Բ	Բ	Բ	Բ	Բ	Փ	Փ	Ն
i	u	qχ	GԽ	qχ'	qχ <sup>w</sup>	GԽ <sup>w</sup>	qχ <sup>w</sup> '	ə	ø	f	v	s
Ճ	Ճ	Ճ	Ճ	Ճ	Ճ	Ճ	Ճ	Ճ	Ճ	Ճ	Ճ	Ճ
z	s <sup>w</sup>	z <sup>w</sup>	t	ʒ	t <sup>w</sup>	u	ʃ	ʒ	ʃ <sup>w</sup>	ʒ <sup>w</sup>	ç	j
Խ	Խ	Խ	Խ	Օ	Լ	Ե	Ե	Ե	Յ	Թ	Յ	Ի
x	ɣ	x <sup>w</sup>	ɣ <sup>w</sup>	o	χ	ɛ	χ <sup>w</sup>	ɛ <sup>w</sup>	H	ɤ	H <sup>w</sup>	h
Կ	Ի	Բ	Լ	Շ	Լ							
m	n	w	r	l	j							