On phonology of East Caucasian languages

G. Moroz

## Introduction

There are a lot of work about phonology of indigenous languages of the Caucasus (Catford 1977; Job and Smeets 1994; Smeets 1994; Alekseev et al. 2001; Hewitt 2004; Grawunder 2017; Beguš 2021; Boris 2021a, 2021b; Koryakov and Maisak Fort) and particular dedicated East Caucasian languages (A. E. Kibrik and Kodzasov 1990). There are also a lot of work on historical-comparative phonetics (Бокарев 1960, 1981; Гудава 1964; Имнайшвили 1977; Акиев 1977; Гигинейшвили 1977; Талибов 1980; Nikolayev and Starostin 1994; Nichols 1994; Ardoteli 2009; Мудрак 2019, 2020) and many others. Fortunately the amount of grammatical description of particular languages is increasing, so we have a lot of more new and detailed information about phonological inventories of particular villages and do not need to extrapolate our knowledge of standard languages onto all villages, where the language is spoken. Even though we have a lot of works on different East Caucasian languages, in order to proceed with the analysis we still need a unified description of those inventories. In order to solve this task, I compiled the database of East Caucasian languages that can be downloaded from [here](https://raw.githubusercontent.com/agricolamz/new_caucasian_phonology_dataset/master/database.csv). The database has the following variables:

* id – unique id for each segment;
* glottocode – language identifier after Glottolog database (Hammarström, Forkel, and Haspelmath 2017);
* language – language name;
* idiom;
* type;
* segment — segment in IPA notation;
* segment\_source — segment as it was presented in the source;
* segment\_type — annotation of type of the segment (consonant, vowel or diphthong);
* source — bibtex entry of the source of the information;
* page;
* comment;
* contributor;
* date – date of the contribution.

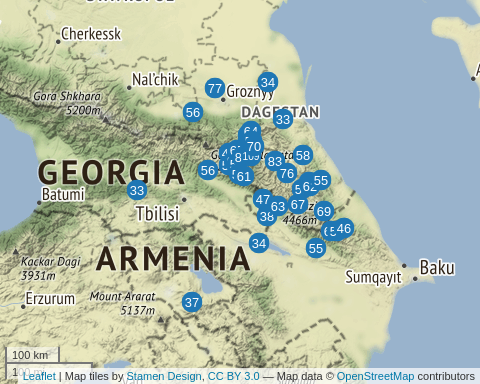
The database is compiled from different sources and therefore some languages can be present multiple times: in some cases it is different idioms (e. g. Darwa dialects), in other cases it is different sources of the same languages (e. g. Khinalug).

On average East Caucasian languages (and other indigenous languages of the Caucasus) have more consonants and vowels then other languages of the world. The main reasons for this are the following

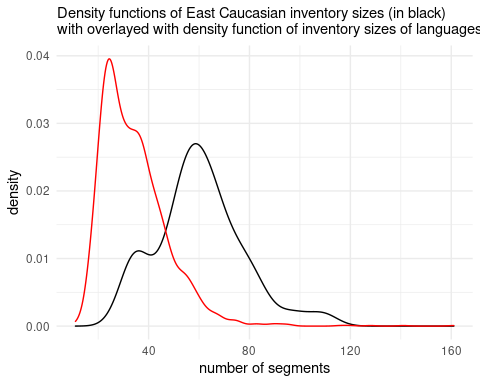
* as other indigenous languages of the Caucasus East Caucasian languages have **ejective consonants** (except Udi);
* as other indigenous languages of the Caucasus East Caucasian languages have **uvular consonants**;
* Andic and Tsezic languages have **lateral obstruents**;
* **labialization**, gemination and fortis/lenis destinction is widespread across Dagestan;
* common among languages of the world triangle vowel system (*i*, *e*, *a*, *o*, *u*) is complicated with **nasalization** (mostly Andic and some Tsezic languages), **long vowels** (Nakh, Andic and Tsezic), **pharyngealisation** and **umlaut vowels** (on the South, Azerbaijanian influence?).

## Inventory size

Here is the map of inventory sizes of East Caucasian languages:



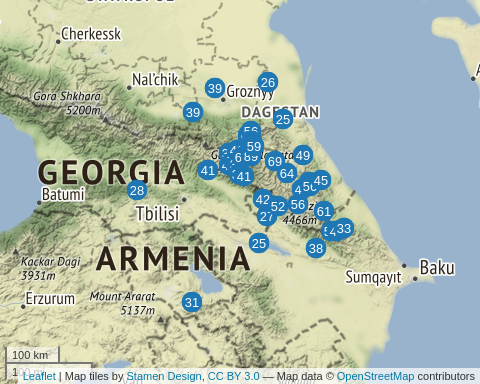
As we can see inventory size of languages in our dataset are ranged from 33 (Georgian and Kumyk) to 109 (Northern Akvakh). We can compare the obtained numbers with the PHOIBLE database (Moran, McCloy, and Wright 2014) which contains inventories of the languages of the world[[1]](#footnote-23):



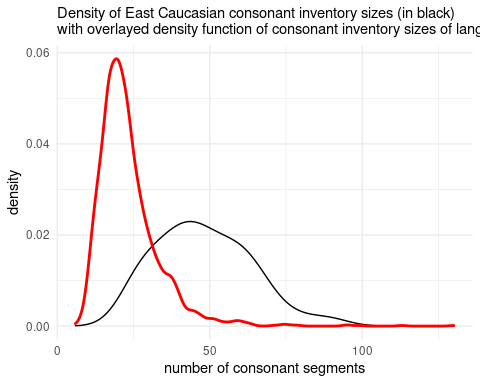
As demonstrated on the plot, East Caucasian languages in general have big segment inventories (with mean, median and mode near 60 segments) comparing to the rest languages of the world. A small peak around 40 can be explained by non-indigenous languages present in our dataset. From the next sections we will see that overall large inventories are mostly caused by large consonant inventories, but vowels and diphthongs also play a significant role.

### Consonant inventory size

We can do the same comparison for the consonant inventories:



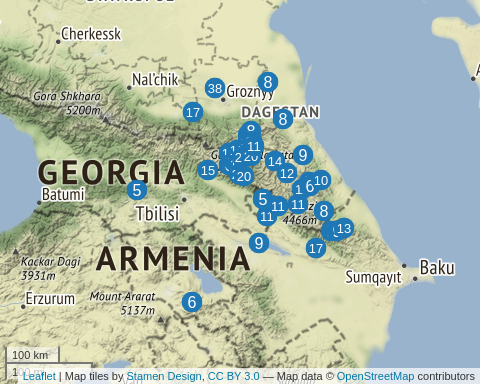
As we see inventory size are differe from 25 (Azerbaijani) to 89 (Northern Akvakh). We can compare the obtained numbers with the PHOIBLE database (Moran, McCloy, and Wright 2014):



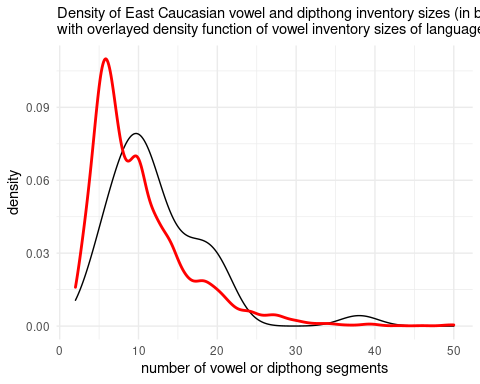
As demonstrated on the plot, the majority of languages from PHOIBLE has less consonants then East Caucasian languages. This is result caused by different subsystems of East Caucasian languages like ejectives, labialized consonants, uvulars and post uvulars. More or less the same phonological profile can be found in other indigenous language family of the Caucasus — North-East Caucasian languages with one of the world consonant champions Ubykh (Fenwick 2011: 16–17).

### Vowel inventory size

We can do the same comparison for the consonant inventories:



As we see inventory size are differ from 5 (Avar) to 38 (Chechen). We can compare the obtained numbers with the PHOIBLE database (Moran, McCloy, and Wright 2014):



As demonstrated on the plot, vowel/diphthong inventory sizes of East Caucasian languages are slightly bigger than the average language of the world. As expected PHOIBLE data reveal average vowel/diphthong inventory size around 5, while the East Caucasian dataset show the mean, median an mode around 10 vowels/diphthongs. Diphthongs are present only in Nakh languages and Hinuq (Forker 2013). However it is worth mentioning that the distinction between diphthong and and combination of vowel and semivowels *j* and *w* is not clear in East Caucasian languages. There is a tendency to have closing diphthongs or combinations of vowel and semivowel in the end of the syllable (like *ai*/*aj*, *eu*/*ew*) and opening diphthongs or combinations of vowel and semivowel in the beginning of the syllable (like *ia*/*ja*, *ue*/*we*), and, as far as I know, there is no phonological difference between diphthongs and combinations of vowel and semivowel in any East Caucasian language. I can stipulate that Nakh scholars tend to describe those unites as diphthongs and Dagestanian scholars tend to describe them as combinations of vowel and semivowel.

## Contrasts in voiceless obstruents

The noticeable feature of all indigenous languages of the Caucasus is the presence of ejective consonants that usually present in most possible places of articulation (see chapter on ejective pʼ) and not restricted to obstruents, but also affricates (and sometimes fricatives, see chapter on ejective fricatives). This feature is so widespread that even non-indigenous languages of the Caucasus like Qax dialect of Azerbaijani (Асланов 1974; Daniel 2021), Kaytag dialect of Kumyk (Дмитриев 1940: 32), Ossetic (Abaev 1964: 6) borrowed some ejective consonants due to the contact or a East Caucasian substrat. It is also worth mentioning that Udi, indigenous East Caucasian language, lost ejectives evidently because of the Azerbaijani and replaced them with non-aspirated obstruents.

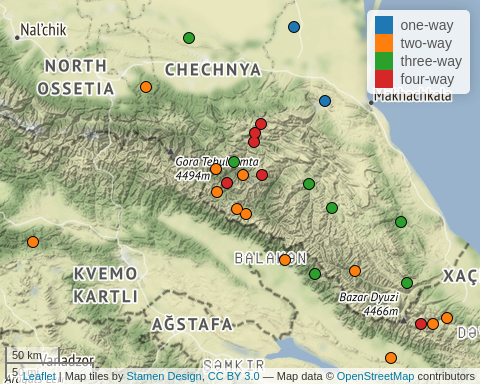
Voiceless stops of all indigenous languages of the Caucasus tend to be slightly aspirated. In some languages there is even opposition of aspirated vs. non-aspirated (sometimes in Russian linguistic the term *преруптивный* is used) stops: Udi, Lezgian, Khinalug.

And the last important contrast that is present in East Caucasian languages is gemination/fortis: that is produced with the prolongation of the closure part and reduction of aspiration in stops and affricates and with the prolongation of the fricative part in fricatives. Different scholars treat them differently in different languages, so I decided to use gemination through all phonological chapters. It is important to notice that unlike North-East Caucasian and Kartvelian languages, some Andic languages and Avar have a unique geminated ejectives (e. g. *kʼː*, *tsʼː*) that are produced with prolongation of the burst part of the consonant.

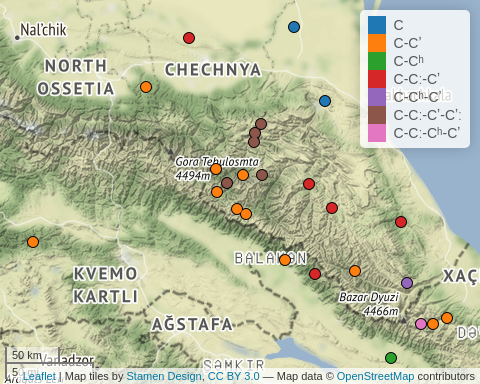
All those series (unaspirated voiceless - aspirated voiceless - geminates - ejectives - geminated ejectives) can be easily distinguished with the analysis of parts of consonants: closure and post burst region, usually called Voice Onset Time or VOT (Catford 1977; Cho and Ladefoged 1999; Grawunder, Simpson, and Khalilov 2010; Grawunder et al., n.d.).

In order to analyse these data we decided to use two features:

* type of distinctions:



* content of distinctions that is denoted as a consonant symbol C:

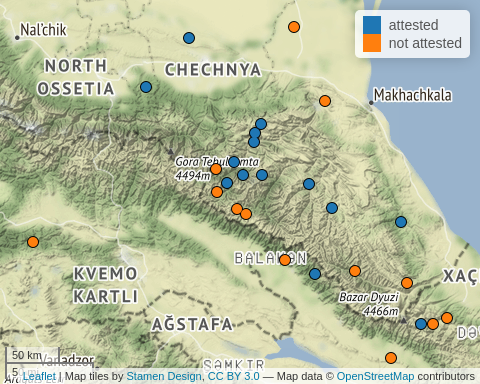


In the following table we can see two features together:

|  |  |  |
| --- | --- | --- |
| type | content | languages |
| one-way | C | Azerbaijani, Kumyk, Nogai, Tat |
| two-way | C-Cʼ | Avar, Bagvalal, Bezhta, Budukh, Chamalal, Dargwa, Georgian, Hinuq, Hunzib, Ingush, Inkhoqwari, Khwarshi, Kryz, Rutul, Tsez, Tsova-Tush |
| two-way | C-Cʰ | Armenian, Udi |
| three-way | C-Cː-Cʼ | Agul, Archi, Chechen, Dargwa, Lak, Lezgian, Tabasaran, Tindi, Tsakhur |
| three-way | C-Cʰ-Cʼ | Khinalug, Lezgian |
| three-way | C-Cʼ-Cʼː | Chamalal |
| four-way | C-Cː-Cʼ-Cʼː | Akhvakh, Andi, Avar, Botlikh, Godoberi, Karata, Tindi |
| four-way | C-Cː-Cʰ-Cʼ | Khinalug |

Even thought it is theoretically possible to have five-way distinction it is not present in our dataset. One-way contrast is present in non-indigenous languages (Azerbaijani, Kumyk, Nogai, Tat). The most frequent is two-way voiceless-ejective distinction that is present in majority of languages. The most frequent three-way distinction is voiceless-geminated-ejective distinction. The most frequent four-way distinction is voiceless-geminated-ejective-geminated ejective that is also attributed to the protolanguage (Nikolayev and Starostin 1994; Nichols 1994). This voiceless-geminated-ejective-geminated is present exclusively in Avar and Andic languages.

It is also interesting to create the map of just geminates:

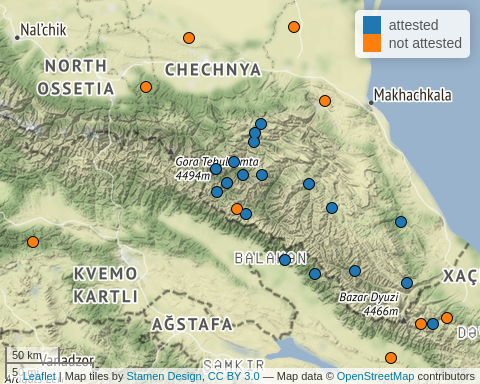


Geminated of fortis/lenis destinction attested through all branches of East Caucasian languages except Tsesic. It looks like languages near the border lose gemination contrast.

## Labialization

Labialization is widely attested in East Caucasian languages, however there are some languages that lost this feature (e. g. in Budukh, (Талибов 2007)). Unfortunately, sometimes scholars do not put the whole list of labialized consonants attested in a language, so it is hard to obtain the whole list. After (Catford 1972) it is worth mentioning that there are several types of labialization:

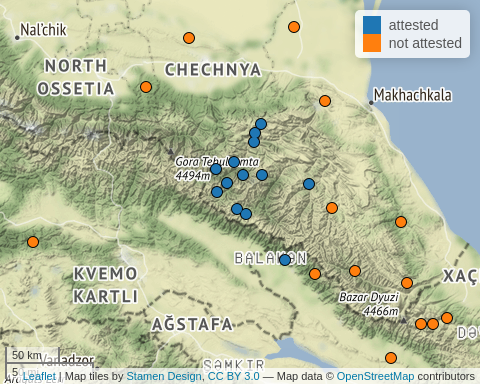
* /w/-labialization — this kind of labialization is more or less independent of the main place of articulation. This kind of labialization is typical for velar and uvular places fricatives and stops and widely attested in East Caucasian languages;
* /ɥ/-labialization — this kind of labialization is typical for pharyngeals. It looks like the tongue body is retracted (due to pharyngeal place of articulation) and raised, that results /ɥ/-like sound that is not present in East Caucasian languages;
* /y/-labialization — this kind of labialization is typical only for post-alveolars in Abkhaz and Abaza (North Caucasian), but also may be a case for Tabasaran (Кибрик et al. 1982: 7–8; A. E. Kibrik and Kodzasov 1990: 337) and Aghul (A. E. Kibrik and Kodzasov 1990: 338).



As we can see labialization is lost in Bezhta, Udi, Budukh, Khinalug. I can stipulate that the absence of labialised consonants in Nakh languages can be just reinterpreted by scholars as part of diphthong. Georgian probably also has labialized consonants, but those consonants are analyzed as a combination of the consonant and *w*.

## Presence of laterals

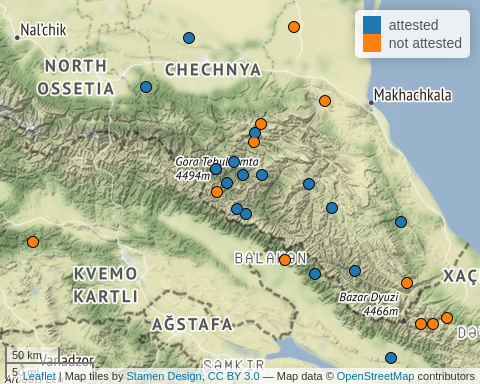
Laterals is the significant feature of Avar, Andic and Tsezic languages, however there is one Lezghian language, Archi, where they also were attested. In most cases lateral system is present by fricative ɬ and two affricates tɬ and tɬ’. Sometimes the subsystem even bigger when geminated version are present.



Lateral inventory can be differ from 1 (Tsova-Tush, ɬ) to 12 (Akhvakh, ɬ, tɬ, tɬ’ all long, labialised and long labialized).

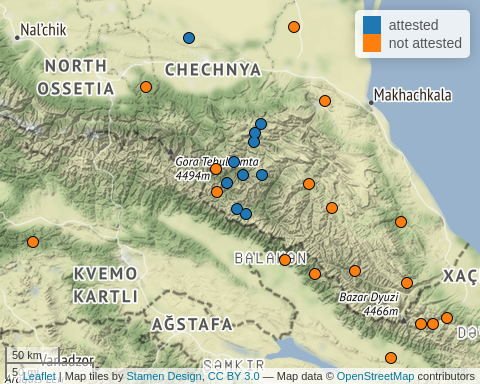
## Presence of long vowels

Long vowels are grouped in the North Dagestan, Chechnya and Ingushetia. It is not present in Avar and and South Dagestan. Hovewer there are some cases when they were reported as a result of some intervocal consonant delitions.



## Presence of nasal vowels

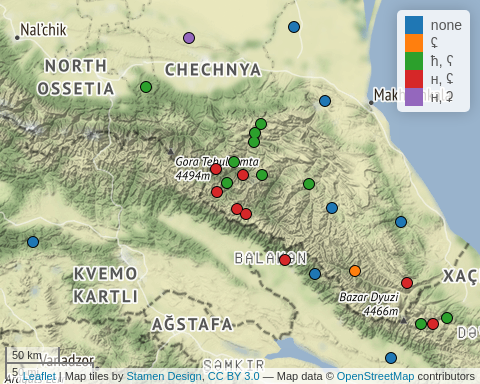
Nasalised vowels is a comon feature of Andic and Tsezic languages, hovewer some researchers report sporadic final **n**-delition that leads to development of nasalised vowels in other branches (e. g. Bezhta, (Талибов 2007)).



## Pharyngeals

Pharyngeals in the Caucasus has a huge coverage in literature (Catford 1983; Кодзасов 1986; Kodzasov 1987; Colarusso 2013; Arkhipov et al. 2019; Беляев 2021). Even thought there is a new model of laryngeal articulation (Esling 1996, 2005), it is hard to adjust data created within the old model to the new approach without new acoustic study, therefor I will use standrad IPA model that distinguishes pharyngeal (ħ, ʕ), epiglottal (ʜ, ʢ, ʡ).

ЭТОЙ КАРТЕ НЕЛЬЗЯ ВЕРИТЬ, ОНА ПРЯЧЕТ НЕКОТОРЫЕ ЗНАЧЕНИЯ

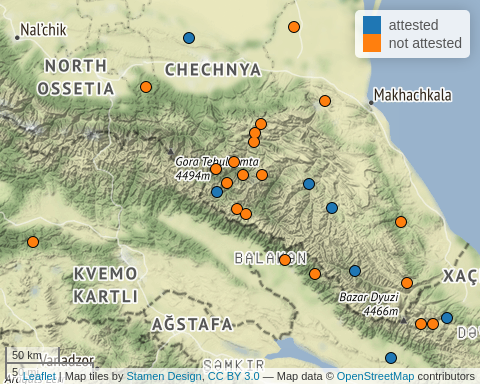


Most common scenario is one place of articulation (or none): either pharyngeal or epiglottal with voced and voiceless consonant. The rest zoo of systems is conected to epiglottal stop that is merged with different subsystems. There are also rare cases with just one pharyngeal consonant: ħ or ʢ.

|  |  |
| --- | --- |
| fricative inventory | languages |
| none | Armenian, Georgian, Karata, Kumyk, Lak, Lezgian, Nogai, Tabasaran, Tsakhur, Udi |
| ʜ, ʢ | Agul, Akhvakh, Andi, Archi, Avar, Bagvalal, Bezhta, Chamalal, Dargwa, Hinuq, Hunzib, Inkhoqwari, Khinalug, Kryz, Lezgian, Tindi, Tsez |
| ħ, ʕ | Akhvakh, Andi, Archi, Bezhta, Botlikh, Budukh, Chamalal, Dargwa, Godoberi, Hunzib, Ingush, Karata, Khinalug, Khwarshi, Kryz, Tindi, Tsova-Tush |
| ħ, ʡ | Hinuq, Rutul, Tat |
| ħ, ʕ, ʡ | Avar, Dargwa |
| ʜ, ʡ, ʢ | Agul, Budukh |
| ħ | Azerbaijani, Tsez |
| ħ, ʜ, ʡ | Dargwa |
| ʜ, ʡ | Chechen |
| ʢ | Rutul |

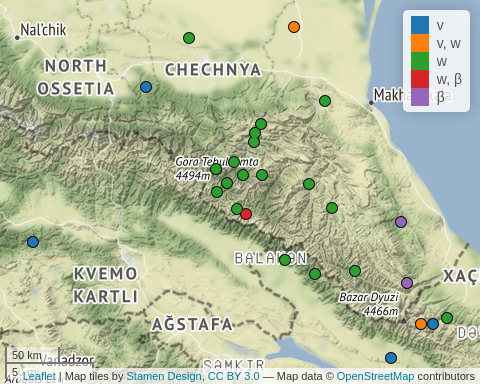
## Presence of pharyngealized segments

Pharyngealization — is a complex feature that phonetically realized as a constriction in the pharynx or epiglottis zone. Different scholars used different strategy in describing locus of the pharyngealization: sometimes it is analysed as a vowel feature (Forker 2013), sometimes as a consonant feature (Khalilova 2009) and sometimes as a suprasegmental feature (Aleksandr E. Kibrik 1994; Moroz 2019). It is common that pharyngealisation spreads through the word, hovewer most researchers distinguish core pharyngealized syllable (Aleksandr E. Kibrik 1994; Талибов 2007; Moroz 2019; Беляев 2021). Everything become even more complicated since some East Caucasian languages developed two different types of pharyngealisation (e. g. Беляев 2021). As we can see from the map, pharyngealization is common for central and south Daghestan for languages of Dargic, Lak, Tsezic and Lezgian branches.



## Inventory of v-like consonants

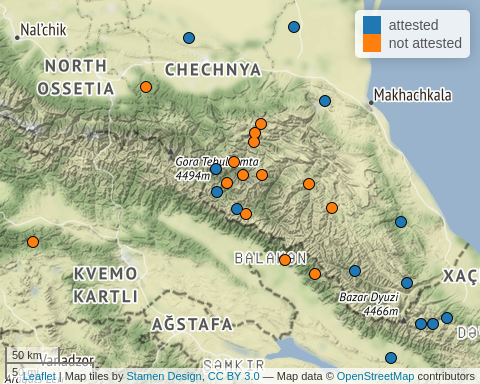
The most common labial sound in East Caucasian languages is bilabial **w** that is present in most languages, however in the South Dagestan approximant become a frecative **v**. I also visited the Dargwa village Ayalakab, where I witnessed the bilabial fricative realisation **β**.



|  |  |
| --- | --- |
| fricative inventory | languages |
| w | Agul, Akhvakh, Andi, Archi, Avar, Bagvalal, Bezhta, Botlikh, Budukh, Chamalal, Chechen, Dargwa, Godoberi, Hinuq, Hunzib, Inkhoqwari, Karata, Khwarshi, Kumyk, Lak, Lezgian, Rutul, Tabasaran, Tindi, Tsakhur, Tsez |
| v | Azerbaijani, Georgian, Ingush, Khinalug, Kryz, Rutul, Tat, Tsova-Tush, Udi |
| v, w | Armenian, Budukh, Khinalug, Nogai |
| w, β | Dargwa, Hunzib |
| β | Lezgian, Tabasaran |

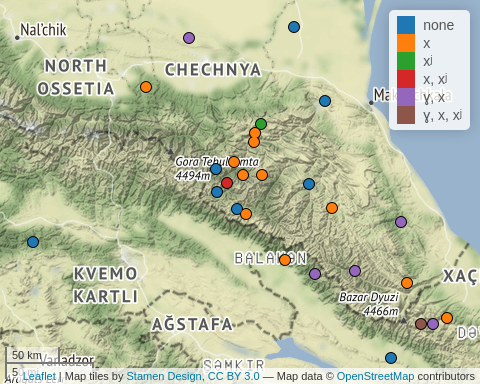
## Presence of umlaut vowels

Umlaut vowels are present in the South Dagestan and in Nakh languages. I can speculatively propose that they historically developed from pharyngealization, but this hypothesis needs to be accurately checked with shared lexicon.



## Presence of velar fricatives

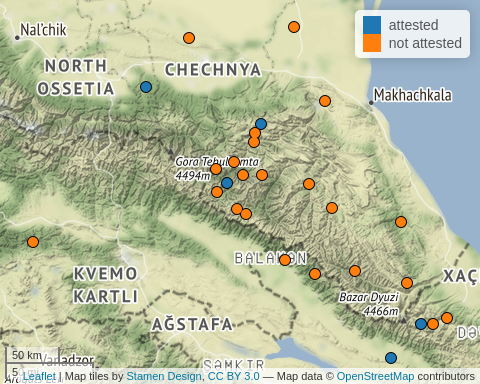
Presence of velar fricatives is a hard feature, since a lot of Russian scholars tend to merge voiced velar fricative *ɣ* and voiced velar stop *ɡ*. Sometimes voiced velar fricative is mixed (or positionally distributed) with voiced uvular fricative *ʁ*, and, as a result, scholars provide merged velar-uvular place of articulation or need to choose one.



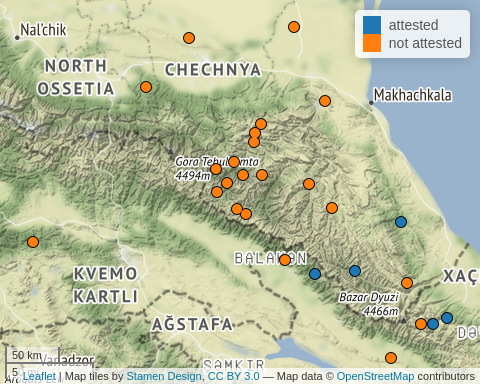
In order to make it easier I collected data in the table. As we see both from the map and the table the most common inventory is just *x*. There are 11 languages that lack velar fricatives including all non-indigenous languages (except Azerbaijani). There are 8 languages that has both voiced and voiceless fricatives. The rest systems are rare.

|  |  |
| --- | --- |
| fricative inventory | languages |
| none | Archi, Armenian, Bezhta, Georgian, Hinuq, Inkhoqwari, Kumyk, Nogai, Tat, Tsez, Udi |
| x | Agul, Akhvakh, Andi, Avar, Bagvalal, Botlikh, Budukh, Chamalal, Dargwa, Godoberi, Hunzib, Ingush, Khwarshi, Lak, Lezgian |
| xʲ | Andi, Azerbaijani |
| x, xʲ | Agul, Karata, Tindi, Tsakhur |
| ɣ, x | Budukh, Chechen, Dargwa, Kryz, Rutul, Tabasaran, Tsakhur, Tsova-Tush |
| ɣ, x, xʲ | Karata, Khinalug |
| ɣ, ɣʲ, x, xʲ | Khinalug |

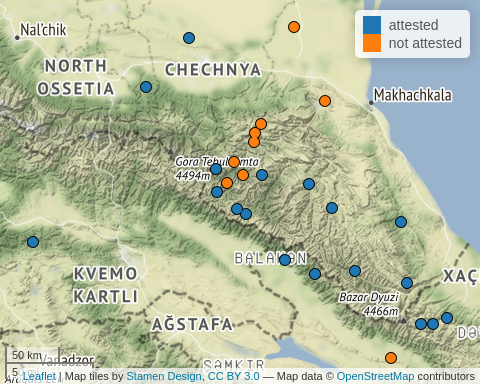
## Presence of palatalized consonants



## Presence of uvular voiced stop



## Presence of ejective pʼ



Abaev, V. I. 1964. *A Grammatical Sketch of Ossetic*. Bloomington: Indiana Univesrity Press.

Alekseev, M. E., S. A. Starostin, G. A. Klimov, and Ya. G. Testelets. 2001. *Yazyki Mira. Kavkazskiye Yaziki*. Academia.

Ardoteli, Nodar. 2009. *Xunjuri-Andiuri-Didouri Enat’a Istoriul-šedarebit’i p’onetika*. Gamomc’emloba Universali.

Arkhipov, A., M. Daniel, O. Belyaev, G. Moroz, and J. H. Esling. 2019. “A Reinterpretation of Lower-Vocal-Tract Articulations in Caucasian Languages.” In *Proceedings of the 19th International Congress of the Phonetic Sciences*, 5–9.

Beguš, Gašper. 2021. “Segmental Phonetics and Phonology.” In *The Oxford Handbook of Languages of the Caucasus*, edited by Maria Polinsky, 1–40. Language Science Press.

Boris, Lena. 2021a. “Tone and Intonation in Languages of the Caucasus.” In *The Oxford Handbook of Languages of the Caucasus*, edited by Maria Polinsky, 1–30. Oxford University Press.

———. 2021b. “Word Stress in the Languages of the Caucasus.” In *The Oxford Handbook of Languages of the Caucasus*, edited by Maria Polinsky, 1–30. Language Science Press.

Catford, J. C. 1972. “Labialisation in Causasian Languages, with Special Reference to Abkhaz.” In *Proceedings of the Seventh International Congress of Phonetic Sciences/Actes Du Septième Congrès International Des Sciences Phonétiques*, 679–82. De Gruyter Mouton.

———. 1977. *Fundamental Problems in Phonetics*. Edinburgh: Edinburgh University Press.

———. 1983. “Pharyngeal and Laryngeal Sounds in Caucasian Languages.” In *Vocal Fold Physiology: Contemporary Research and Clinical Issues*, 344–50. San Diego, CA: College Hill Press.

Cho, T., and P. Ladefoged. 1999. “Variation and Universals in VOT: Evidence from 18 Languages.” *Journal of Phonetics* 27: 207–29.

Colarusso, J. 2013. “The Typology of Gutturals.” In *Base Articulatoire Arrière: Backing and Backness*, edited by Jean Léo Léonard and Samia Naïm, 93–109. Munich: LINCOM Europa.

Daniel, Michael. 2021. “Ejectives in the Azerbaijani of Ilisu.”

Esling, J. H. 1996. “Pharyngeal Consonants and the Aryepiglottic Sphincter.” *Journal of the International Phonetic Association* 26 (2): 65–88.

———. 2005. “There Are No Back Vowels: The Larygeal Articulator Model.” *Canadian Journal of Linguistics/Revue Canadienne de Linguistique* 50 (1-4): 13–44.

Fenwick, R. S. H. 2011. *A Grammar of Ubykh*. Lincom Europa.

Forker, D. 2013. *A Grammar of Hinuq*. De Gruyter Mouton.

Grawunder, S. 2017. “The Caucasus.” In *The Oxford Handbook of Languages of the Caucasus*, edited by R. Hickey, 356–95. Cambridge University Press.

Grawunder, S., G. Moroz, V. Zhigulskaya, and M. Daniel. n.d. “VOT in Two East Caucasian Languages.”

Grawunder, S., A. Simpson, and M. Khalilov. 2010. “Phonetic Characteristics of Ejectives - Samples from Caucasian Languages.” In *Turbulent Sounds - an Interdisciplinary Guide, Number 21 in Interface Explorations*, edited by Jean Léo Léonard and Samia Naïm, 356–95. Berlin: De Gruyter Mouton.

Hammarström, Harald, Robert Forkel, and Martin Haspelmath. 2017. “Glottolog 4.4.” *Max Planck Institute for the Science of Human History*.

Hewitt, George. 2004. *Introduction to the Study of the Languages of the Caucasus*. Lincom.

Job, D. M., and R. Smeets, eds. 1994. *The Indigenous Languages of the Caucasus*. Vol. 3. Caravan Books.

Khalilova, Zaira. 2009. “A Grammar of Khwarshi.” PhD thesis, Leiden University.

Kibrik, A. E., and S. V. Kodzasov. 1990. *Sopostavitelnoye Izucheniye Dagestanskix Yazykov: Imya. Fonetika*. Vol. 2. Moskovskij Gosudarstvennyj Universitet.

Kibrik, Aleksandr E. 1994. “Archi.” In *The Indigenous Languages of the Caucasus. Vol. 4: The Northeast Caucasian Languages. Part 2*, edited by Rieks Smeets, 297–365. Delmar, NY: Caravan Press.

Kodzasov, S. V. 1987. “Pharyngeal Features in the Daghestan Languages.” In *Proceedings of the XIth International Congress of Phonetic Sciences*, 2:142–44.

Koryakov, Yu., and T. Maisak, eds. Fort. *Handbook of Caucasian Languages*. Vol. 2. Berlin/New York: De Gruyter Mouton.

Moran, Steven, Daniel McCloy, and Richard Wright, eds. 2014. *PHOIBLE Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://phoible.org/>.

Moroz, George. 2019. “Phonology of Mehweb.” In *The Mehweb Language: Essays on Phonology, Morphology and Syntax*, edited by Michael Daniel, Nina Dobrushina, and Dmitry Ganenkov, 17–39. Language Science Press.

Nichols, Johanna. 1994. “The Nakh-Daghestanian Consonant Correspondences.” In *Current Trends in Caucasian, East European, and Inner Asian Linguistics : Papers in Honor of Howard i. Aronson*, edited by Dee Ann Holisky and Kevin Tuite, 207–65. John Benjamins Publication Company.

Nikolayev, S. L., and S. A. Starostin. 1994. *A North Caucasian Etymological Dictionary*. Asterisk Press.

Smeets, R., ed. 1994. *The Indigenous Languages of the Caucasus*. Vol. 4. Caravan Books.

Акиев, А. Ш. 1977. “Историко-сравнительная фонетика даргинского и лакского языков: Система консонантизма.” *Дагучпедгиз*.

Асланов, А. М. 1974. “Иберийско-кавказские абруптивы в закатальско-кахских говорах азербайджанского языка.” *Ежегодник иберийско-кавказского языкознания* 1: 224–28.

Беляев, О. И. 2021. “Два типа фарингализации в ширинском даргинском: Предварительные наблюдения.” In *ДУРХЪАСИ ХАЗНА*, edited by Т. А. Майсак, Н. Р. Сумбатова, and Я. Г. Тестелец, 18–29.

Бокарев, Евгений Алексеевич. 1960. *Основные вопросы исторической фонетики дагестанских языков*. Издательство восточной литературы.

———. 1981. *Сравнительно-историческая фонетика восточнокавказских языков*. Наука.

Гигинейшвили, Б. К. 1977. “Сравнительная фонетика дагестанских языков.” *Тбилиси: Издательство Тбилисского университета*.

Гудава, Того Евстафьевич. 1964. *Консонантизм андийских языков: Историко-сравнительный анализ*. Изд-во Академии наук Грузинской ССР.

Дмитриев, Н. К. 1940. *Грамматика кумыкского языка*. АН СССР.

Имнайшвили, Д. С. 1977. *Историко-сравнительный анализ фонетики нахских языков*. Мецниереба.

Кибрик, Бергельсон, А. Е., КК Богатырев, ОЮ Богуславская, Анна А Зализняк, СВ and Кодзасов, ИА Муравьева, ВА Плунгян, and ЕВ Рахилина. 1982. “Табасаранские этюды.”

Кодзасов, С. В. 1986. “Фаринго-ларингальное сужение в дагестанских языках.” In *Актуальные проблемы дагестанско-нахского языкознания*, 344–50. Махачкала: Даг. фил. АН СССР.

Мудрак, Олег Алексеевич. 2019. “Аварские основы.”

———. 2020. “Андийские основы.”

Талибов, Б. Б. 1980. *Сравнительная фонетика лезгинских языков*. Наука.

———. 2007. *Будухский язык*. Academia.

1. Here and further performing such a comparison of our dataset and one, provided by PHOIBLE, arise a question: whether we need to exclude East Caucasian and non-indigenous languages of the Caucasus from the PHOIBLE subsample that we use? In this text we decided to exclude them for the sake of the comparability. However, groups that we get are far from to be comparable in the proper way: different language families are overrepresented in both groups; the dataset sizes are different (PHOIBLE’s 2169 languages vs our dataset of 50 idioms) etc. For those who are against of the exclusion we assure that this operation changes shapes of the density plots that we obtained, but the change is extremely small. [↑](#footnote-ref-23)