

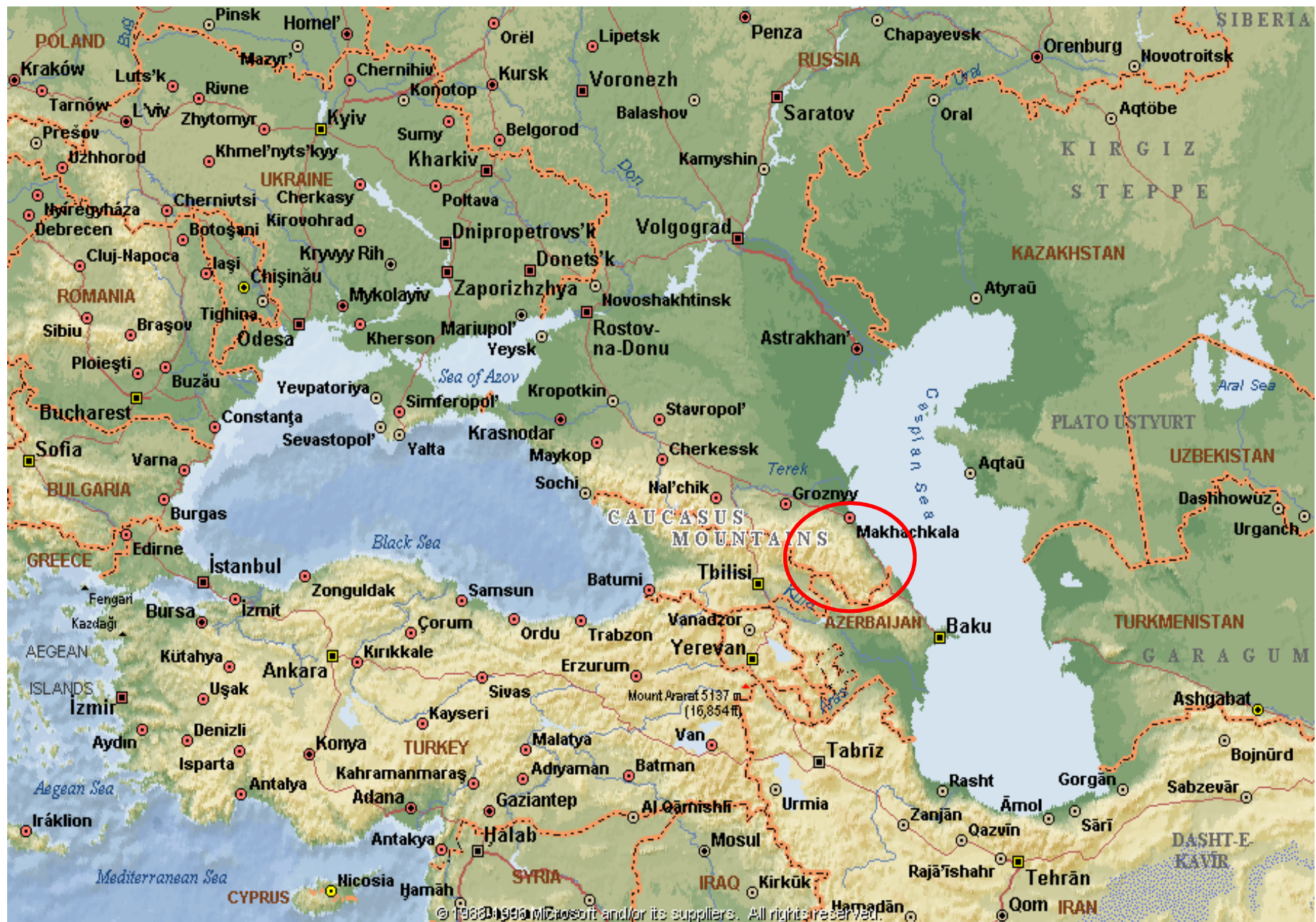




Multilingualism in Daghestan: a retrospective field research

Nina Dobrushina
Aleksandra Kozhukhar'
26/09/2017
Linguistic Convergence Lab, HSE

- **Introduction – Daghestan and its multilingualism**
- Problem setting
- Method and data
- Aims
- Some results
- Problems and restrictions



Daghestan

- Northeast Caucasus
- Russian Federation, border with Chechnya, Georgia and Azerbaijan
- Highlands and ridges
- Over 40 languages on a territory of ~ 50 000 km²
- Diverse geography and economy (plains fertile and rich, mountains scarce and poor)
- Population almost exclusively Muslims
- Three language families (Nakh-Daghestanian, Turkic, Indo-European)
- Even related languages are quite distant

What is special about multilingualism in Daghestan?

- **Diversity:** Villages with different local languages are often within walking distance
- **Marrying-in:** Mixed marriages not encouraged, wives most often taken from the same village
- **One native language per village:** Mountain villages were ethnically and linguistically homogenous
- **Vitality:** Local languages are still spoken in highland villages
- **Multilingualism:** Traditional patterns of language contact are still observable

- Introduction – Daghestan and its multilingualism
- **Problem setting**
- Method and data
- Aims
- Some results
- Problems and restrictions

Research problem

- How did neighbors with different languages communicate between each other?

Some questions to be answered

- What configurations of bilingualism (lingua franca, asymmetrical bilingualism, symmetrical bilingualism etc.) were more typical of the area?
- Why?
- Which language is chosen as lingua franca?
- Which language is chosen as the dominant in the case of asymmetric contact and why?
- How multilingualism was socially distributed in the community?

- Introduction – Daghestan and its multilingualism
- Problem setting
- **Method and data**
- Aims
- Some results
- Problems and restrictions

The method of retrospective family interviews

As a proxy for the intensity of language contact, we take the rate of bilingualism at the community level

- Short interviews about language repertoire of the villagers are taken
- The respondent reports the data not only about himself but also about all her elder relatives whom (she thinks) she remembers

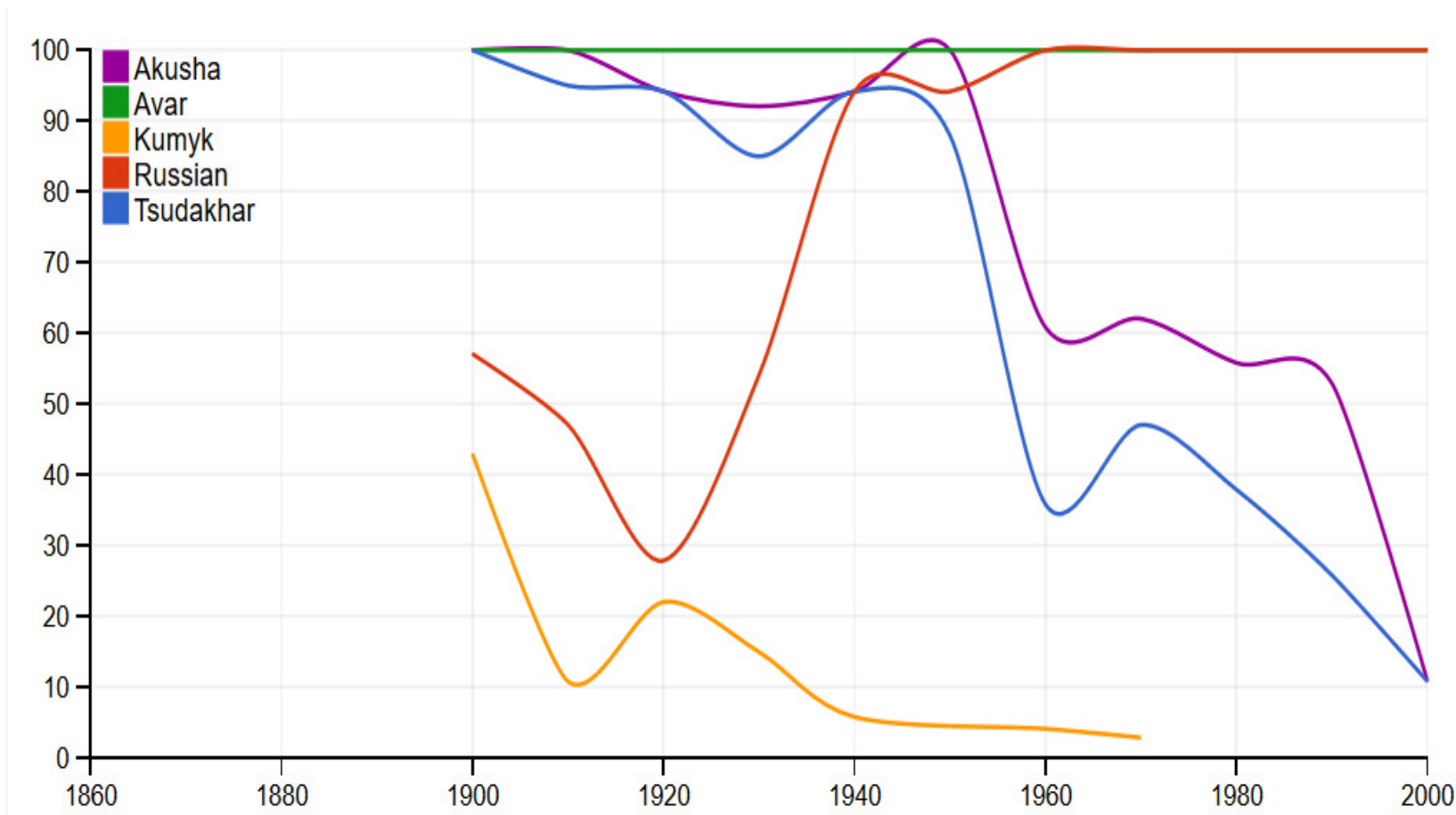
<i>Name</i>	Akaj
<i>Born in</i>	Chabanmakhi
<i>The interviewer was talking to</i>	Umaidat
<i>Family relation to the respondent</i>	father of Umaidat
<i>Years of birth and death</i>	1900 – 1973
<i>Native language</i>	Kadar Dargwa
<i>Education and living outside the village</i>	worked as a mason, also in other villages
<i>Did he read the Koran?</i>	yes, but he could not translate
<i>Did he speak Avar?</i>	yes
<i>Did he speak Kumyk?</i>	yes
<i>Did he speak Russian?</i>	yes
<i>Did he speak any other languages?</i>	no
<i>Literate in</i>	Arabic, Cyrillic



Why retrospective?

- From the establishment of the Soviet schools in the 1930s, Russian quickly spread over Daghestan as L2
- Traditional patterns of language contact have been almost completely substituted by Russian as a lingua franca

The dynamics of multilingualism in Chuni (Avar) (years of birth)

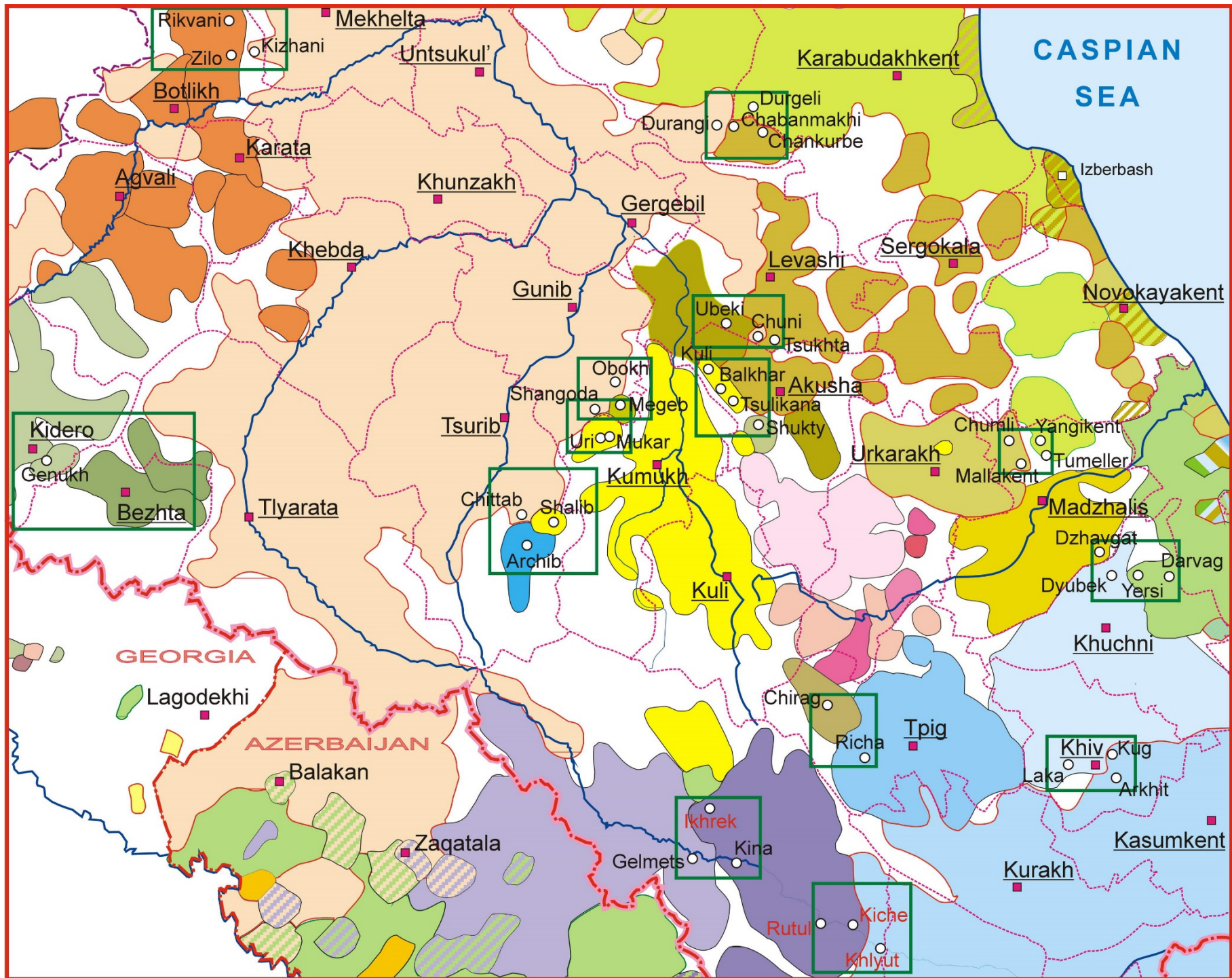


Why retrospective?

- Presumably, people born before 1919 were not (heavily) affected by Russification
- People born in the 30s to 40s usually remember language repertoire of their parents and grandparents
- ...these people are passing away

Data

- fieldwork every year since 2009
- collective fieldtrips with students since 2012
- fieldtrips to 14 clusters of villages
(2 to 4 villages in each cluster)
- 47 villages
- 28 lects (languages or dialects)



E7 X ✓ fx бабушка Шамиля (ск23), мать Халмиат (ск24)

Table with columns A-AM and rows 1-66. Contains personal data, education, and language skills for individuals like 'бабушка Шамиля' and 'мать Халмиат'.

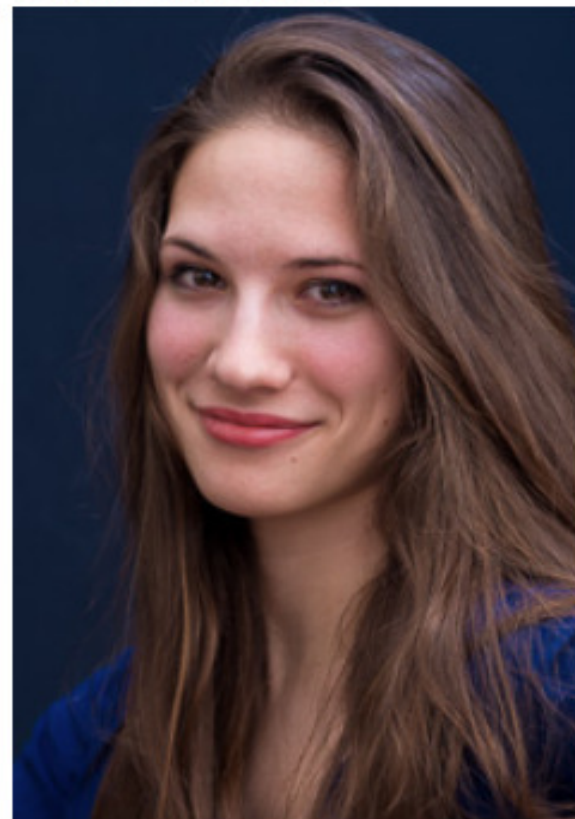
Database online from 2017

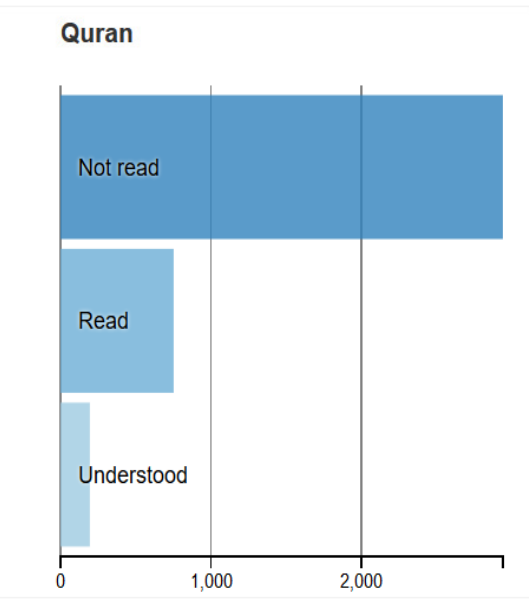
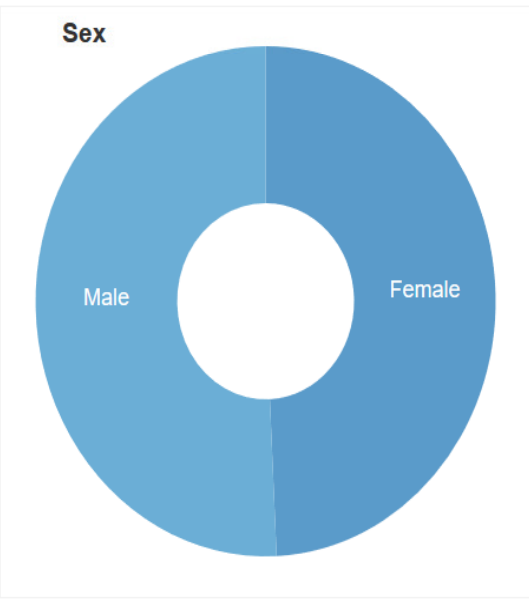
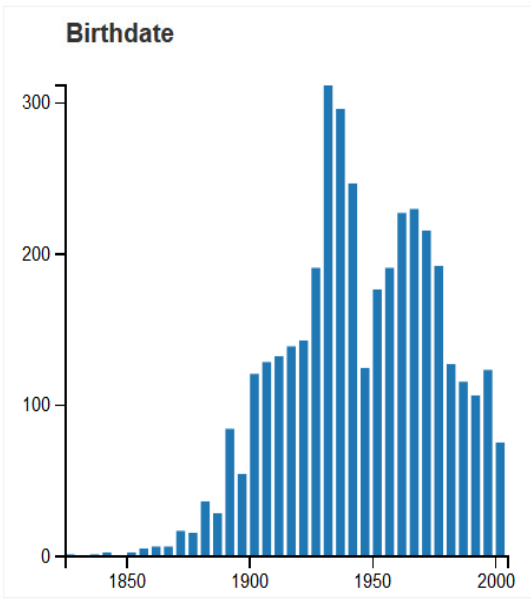
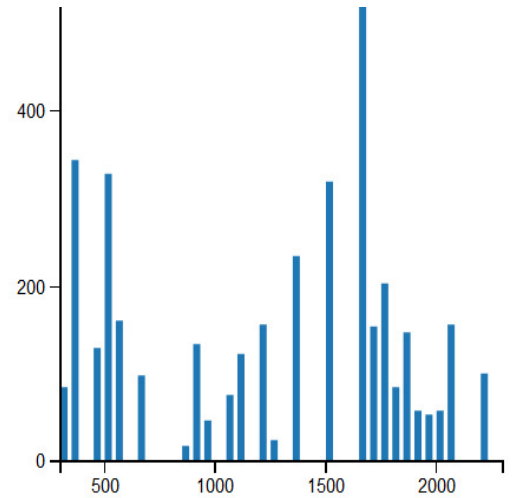
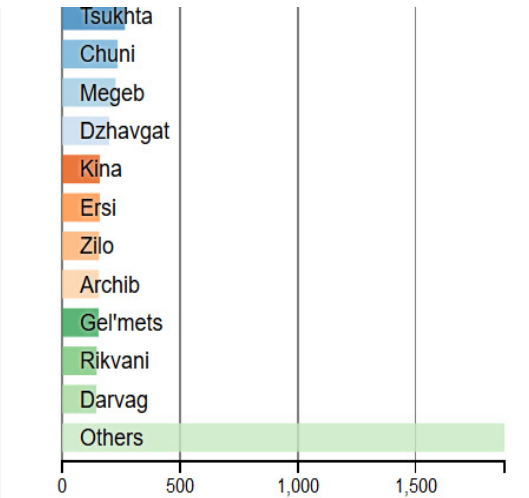
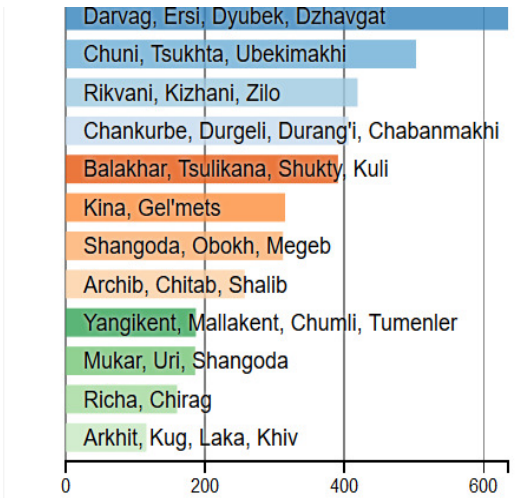
<http://multidagestan.com/>

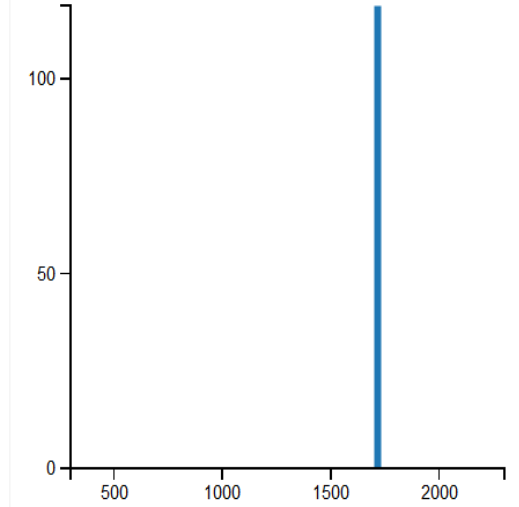
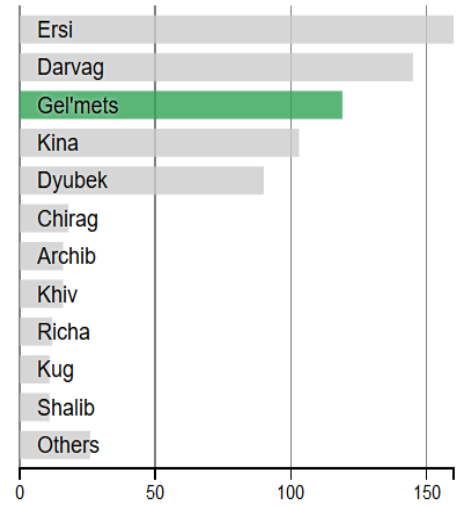
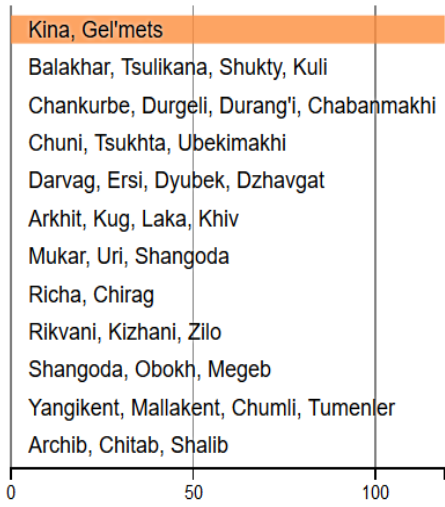
Alexander Belokon



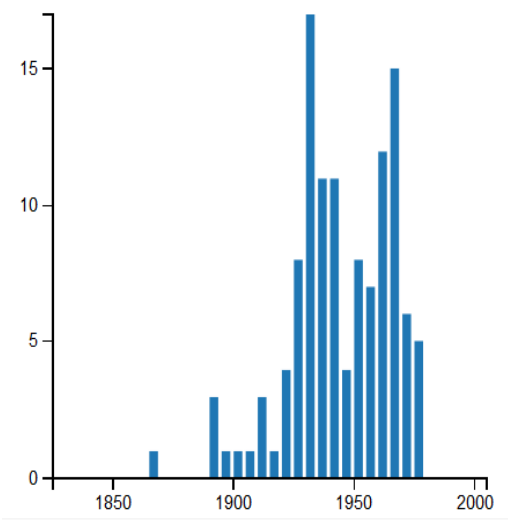
Daria Staferova



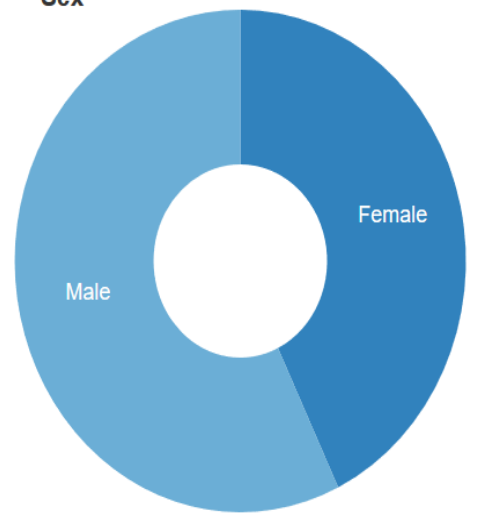




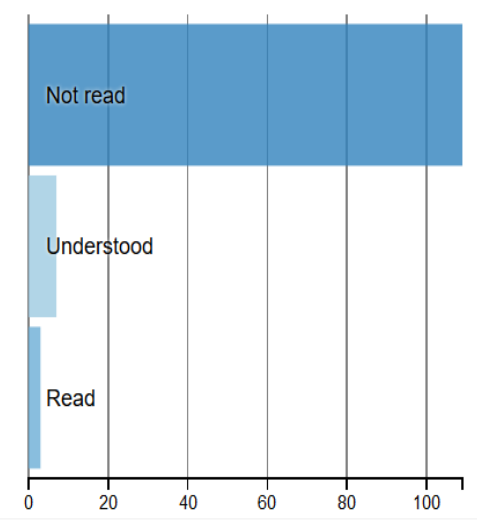
Birthdate

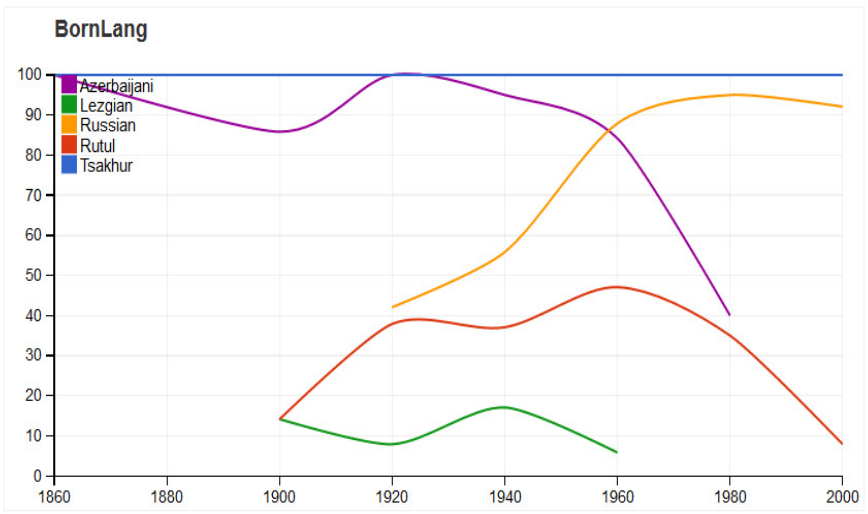
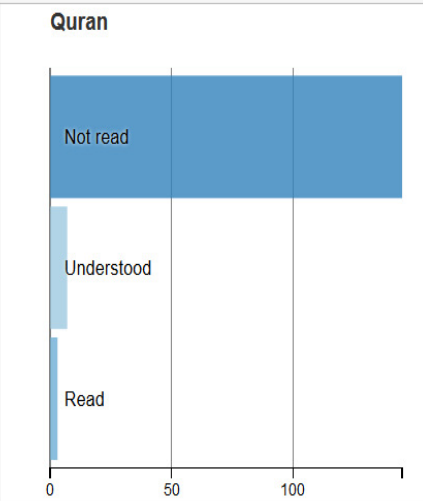
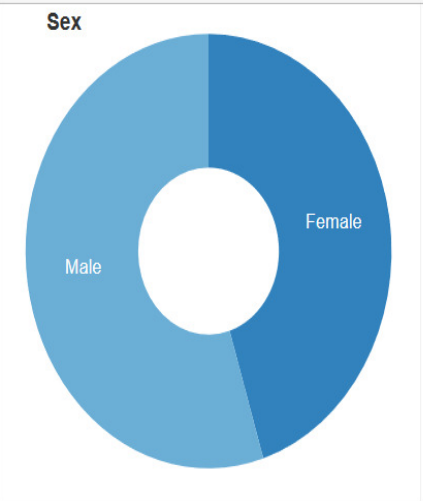
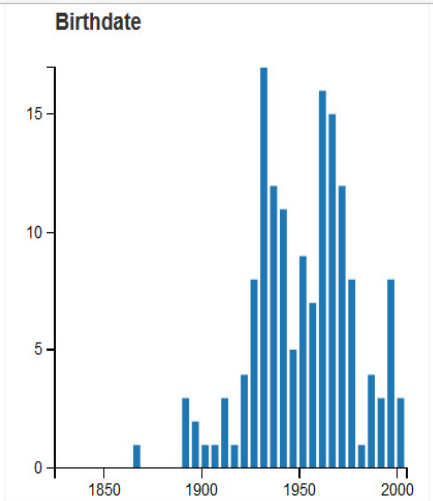
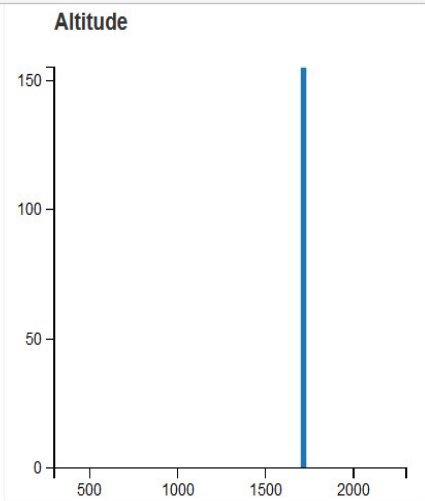


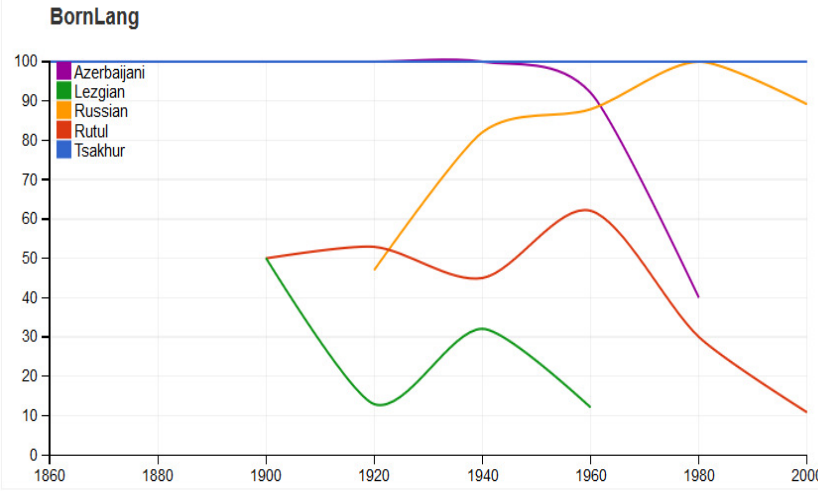
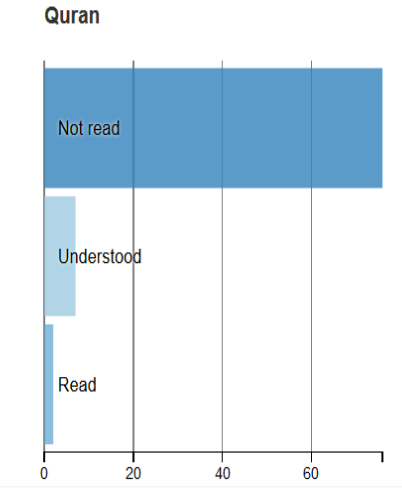
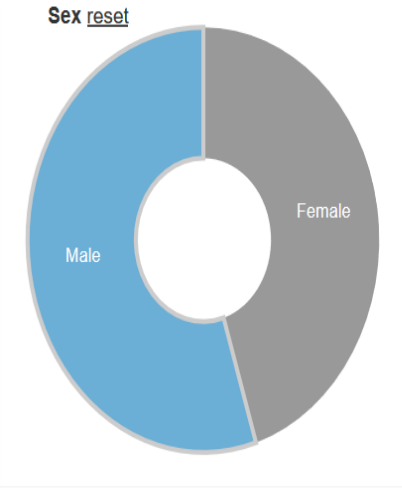
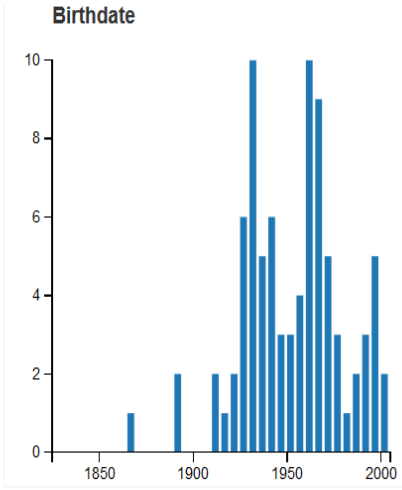
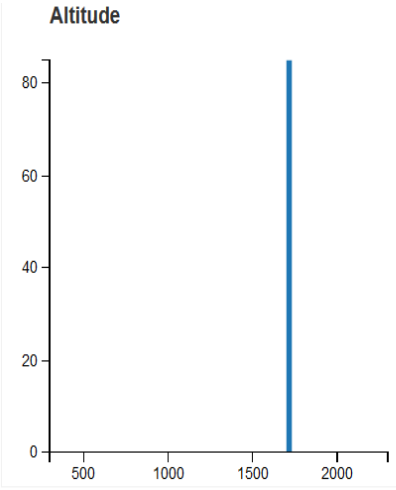
Sex

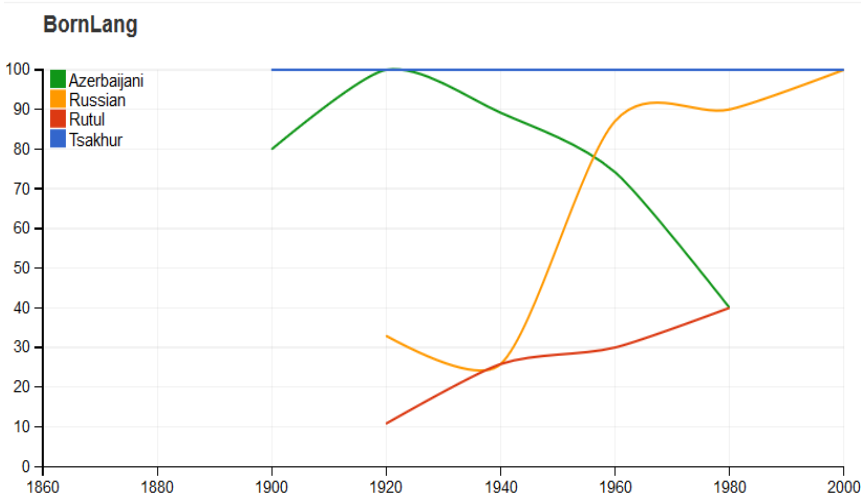
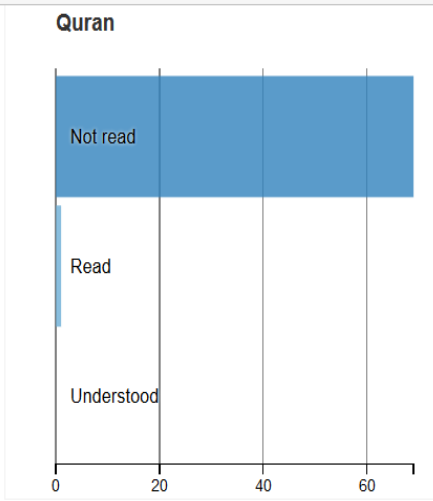
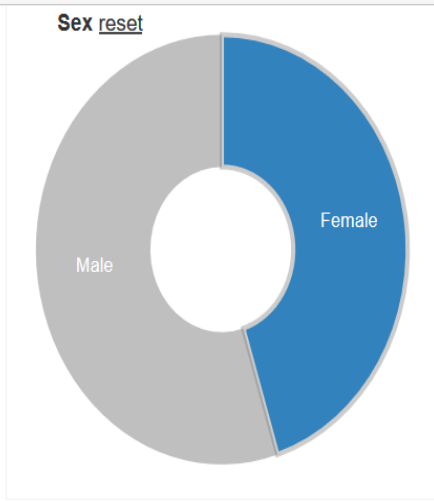
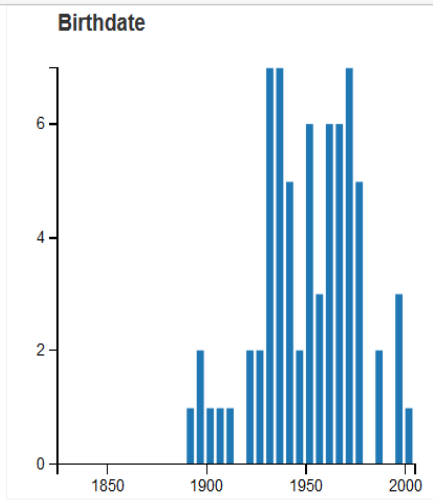
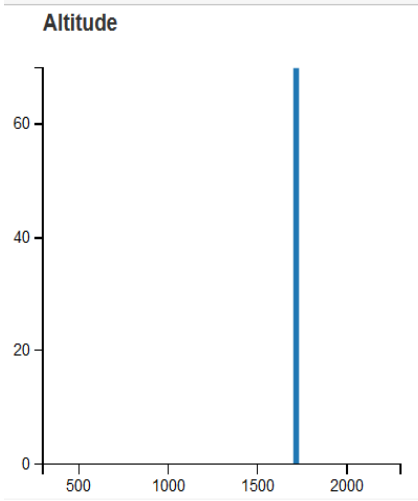


Quran

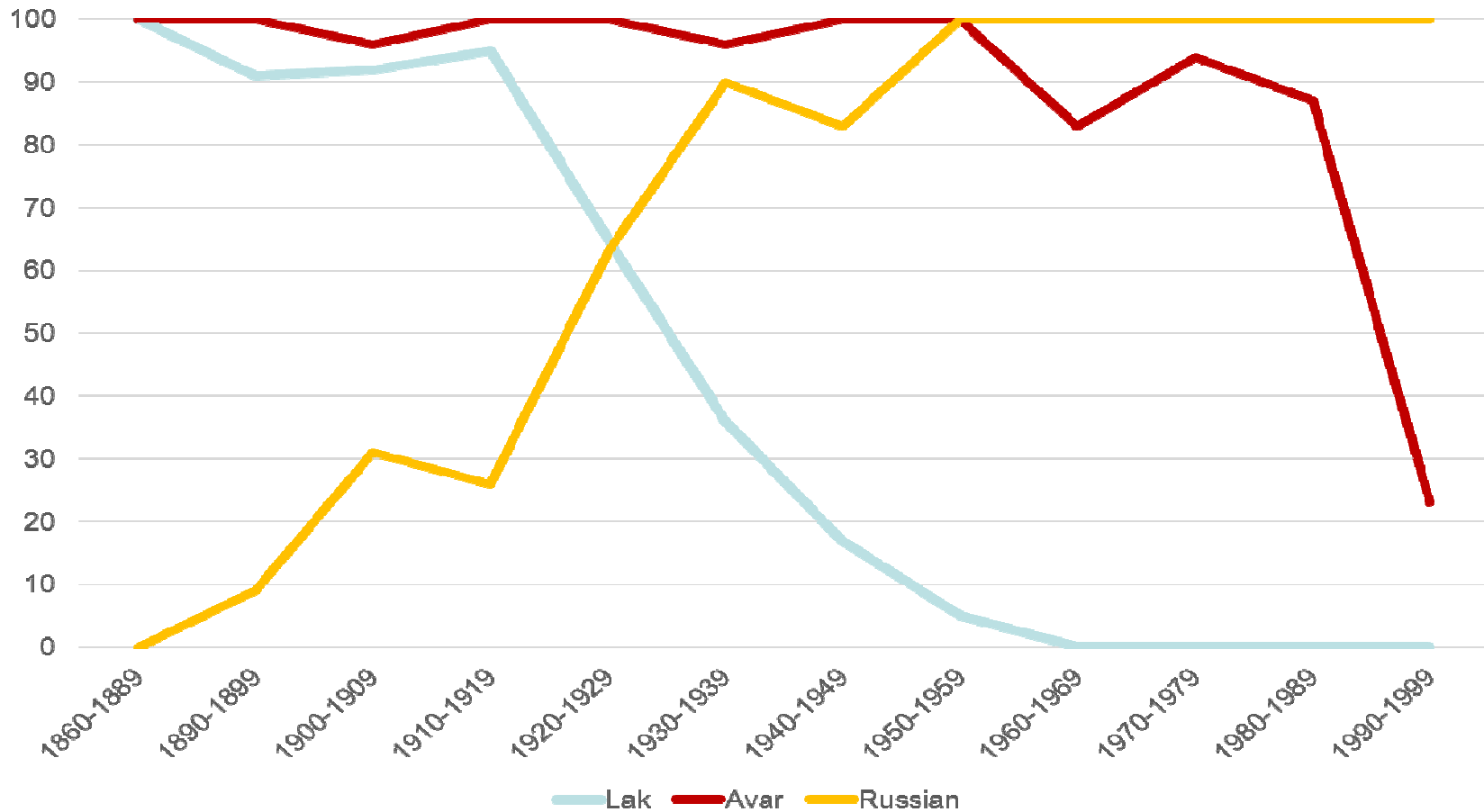




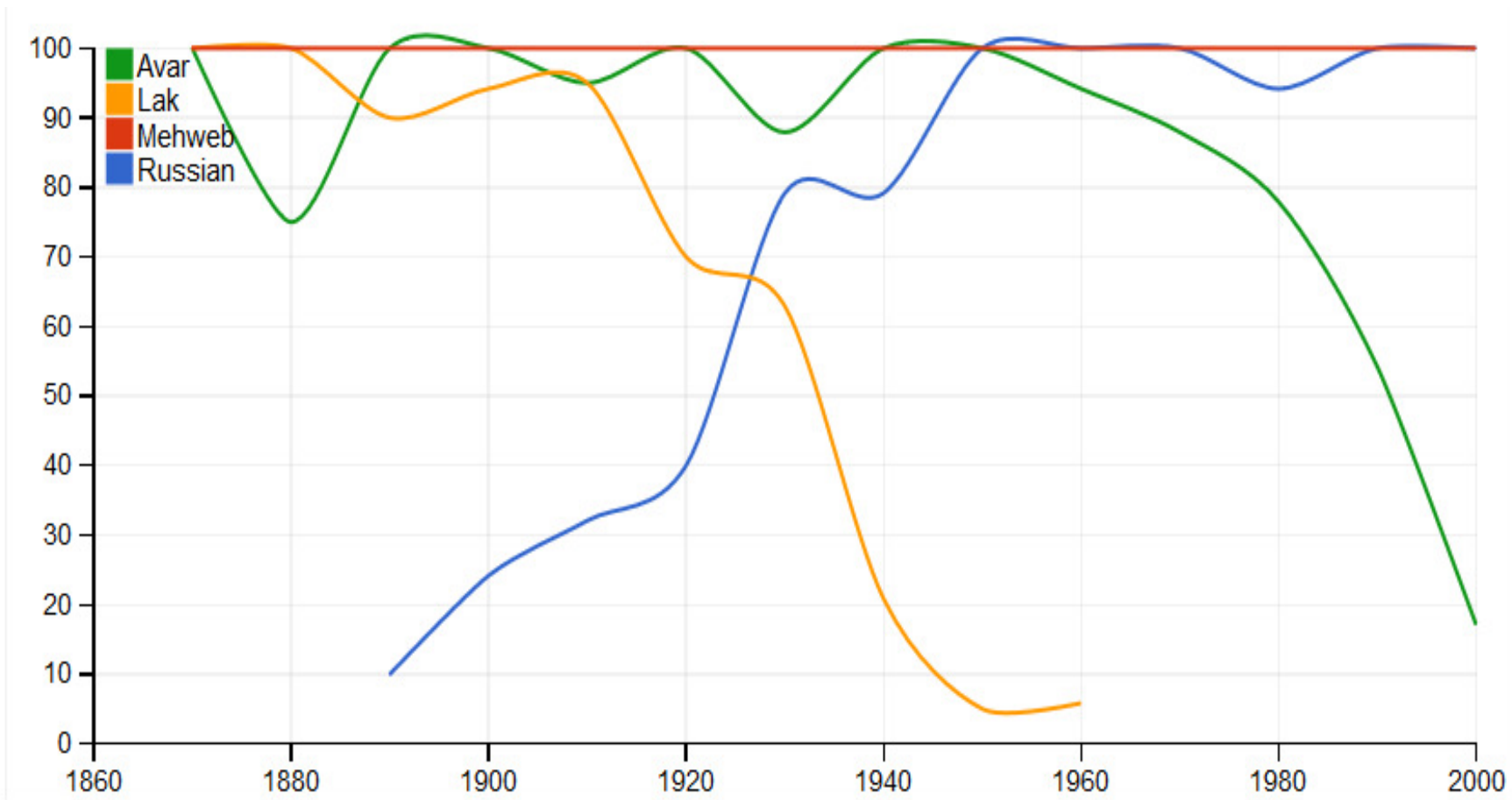




Second languages in Mehweb hand-made in Excel, October 2017 (60 minutes spent)



Multilingualism in Mehweb automatically made in May 2017 (1 minute spent)



- Introduction – Daghestan and its multilingualism
- Problem setting
- Method and data
- Aims
- **Some results**
- Problems and restrictions

- Distant VS Neighboring languages
- Neighbors VS Second neighbors
- Distribution of second languages across genders

Distant VS neighboring languages

- **Distant languages** are spoken beyond the neighborhood, acquired in the course of seasonal activities
- **Neighboring languages** are spoken in the neighborhood, acquired in the course of everyday activities

Distant languages

- **Azerbaijani** in Kug, Khiv, Arkhit, Laka
 - Seasonal oil extraction
- **Avar** in Balkhar, Tsulikana
 - Selling Balkhar pottery
- **Kumyk** in Chuni, Tsukhta, Verkhnie Ubeki
 - Seasonal shepherding

Why do we need to distinguish between distant VS neighbor multilingualism?

- They differ in respect of the number of bilingual people
- They differ in terms of genders of speakers

What do we know about genders and multilingualism?

In many places, **multilingualism** is practiced almost exclusively **by the male part** of the community. Women are considered guardians of the home language and culture, and speak only the local language (Hill 1987, Kulick 1992, Herbert 2002, Hoffman 2006: 150)

Most of these works are studies of bilingual communities with the interplay between first (local) language and second (dominant and / or urban) language

Females speak only the local language, men speak both

In highland Daghestan, multilingual repertoire includes three, four, five languages

We consider the dynamics of shift between different second languages

Social distribution of multilingualism: **genders**

According to t-test results,
the **difference between
female and male
multilingualism is
significant** (p-value =
2.2e-16)

Dataset contains 4000
observations

(run by Sasha Kozhukhar')



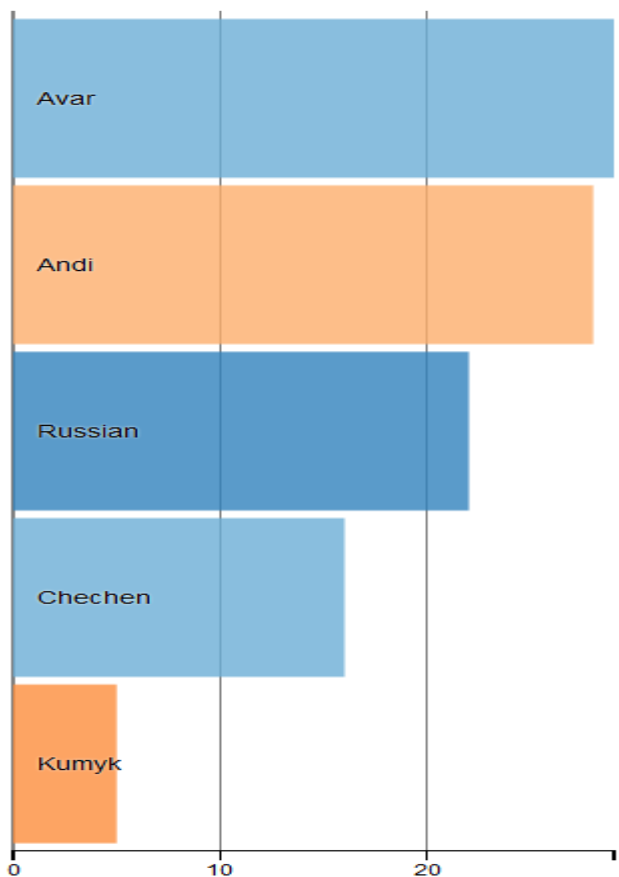
Second languages are gendered

Distant languages were spoken almost exclusively by men

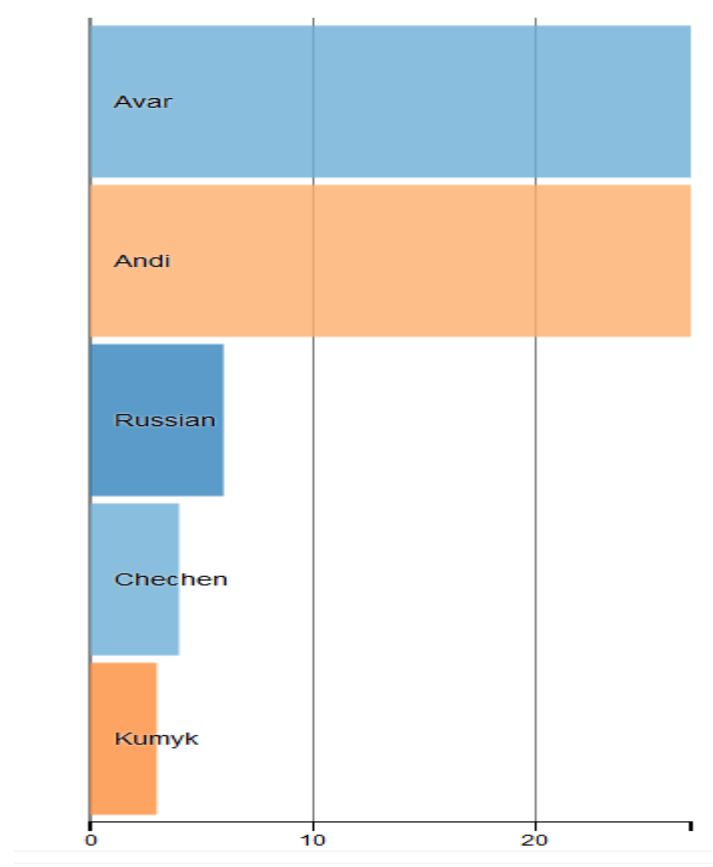
Neighboring languages were gendered to a smaller extent

Rikvani 1889 – 1940 y.b.

Men

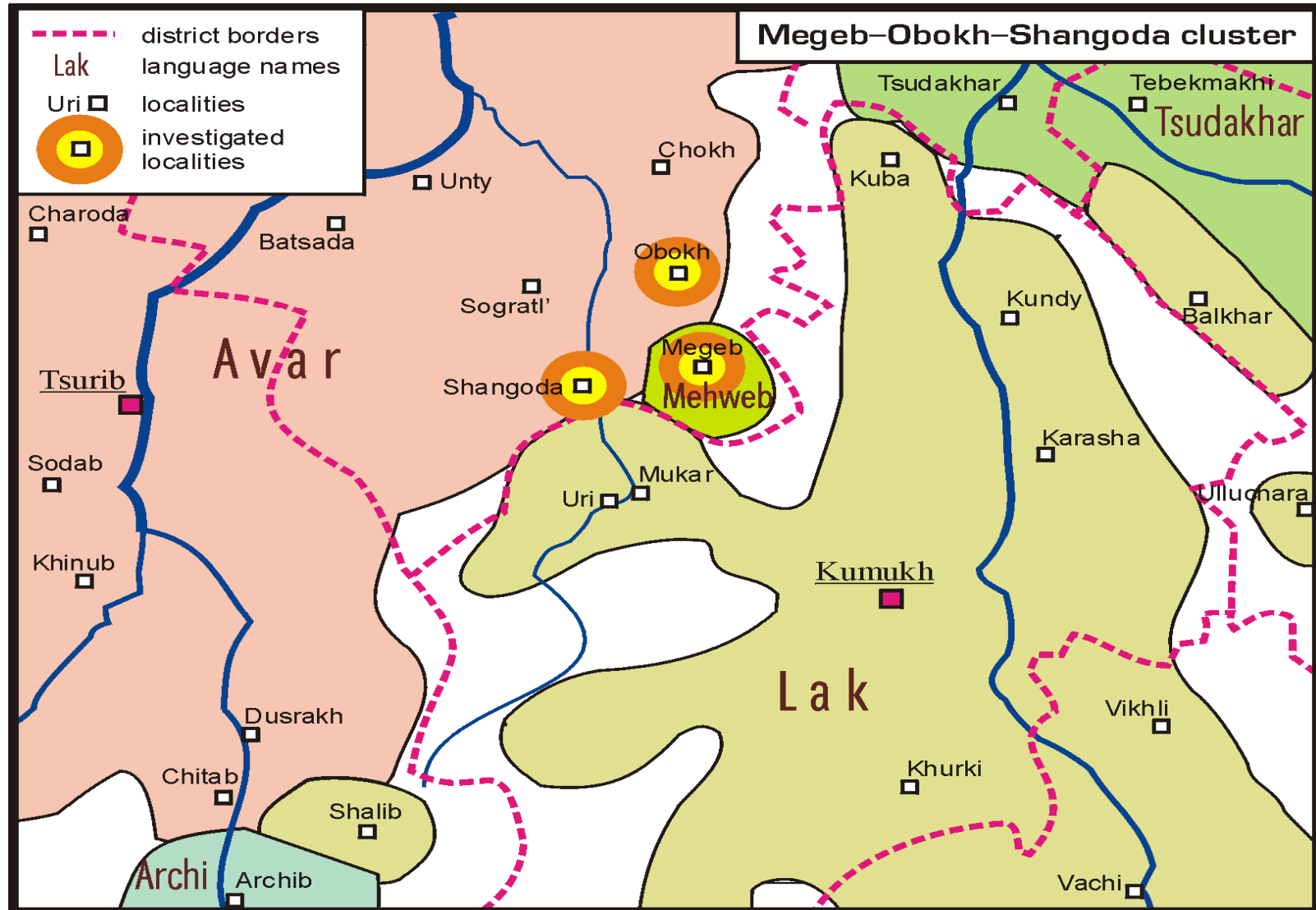


Women

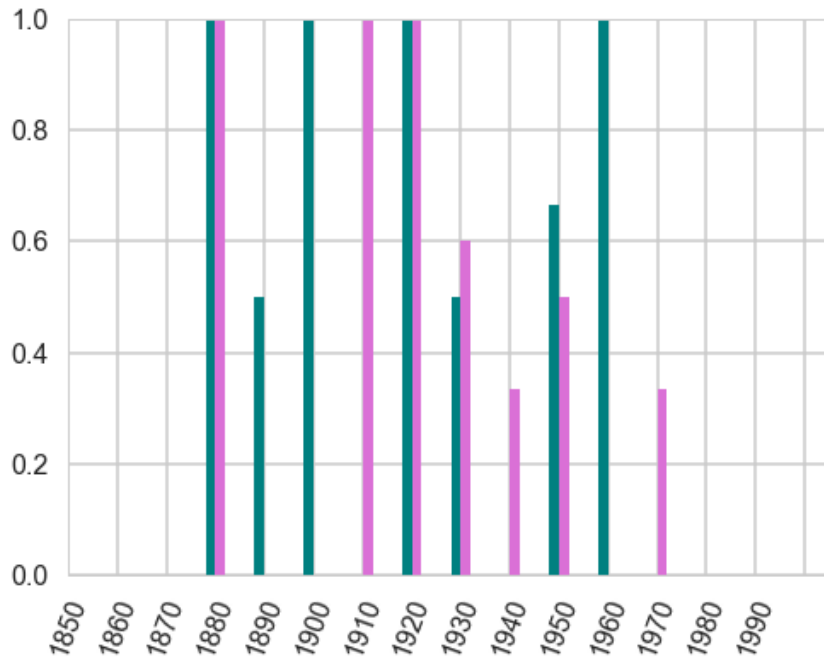


What about neighbors of
neighbors (second neighbors)?

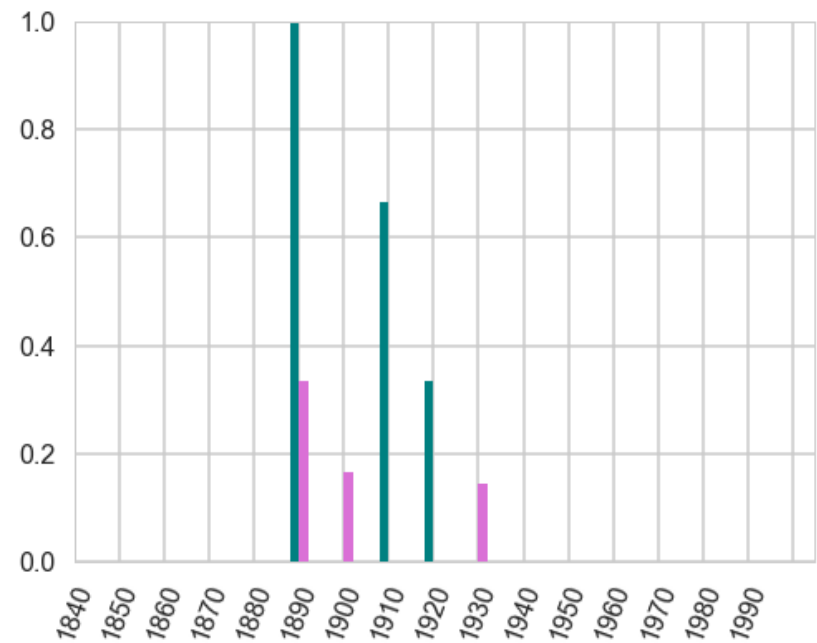
Lak in Shangoda - 92%, Lak in Obokh - 24%



Second neighbor language: Lak

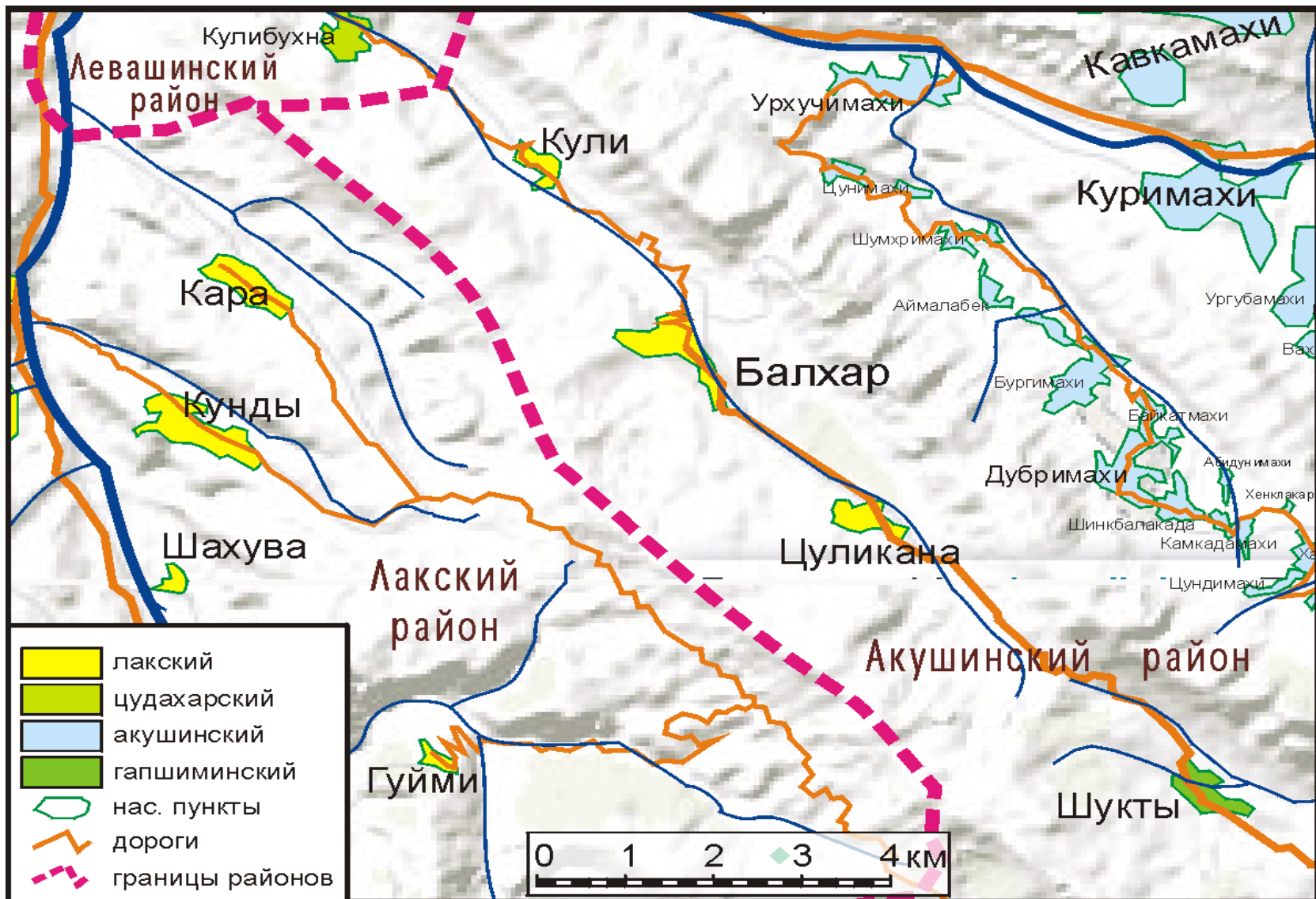


Lak by gender in neighboring villages
(Shangoda)

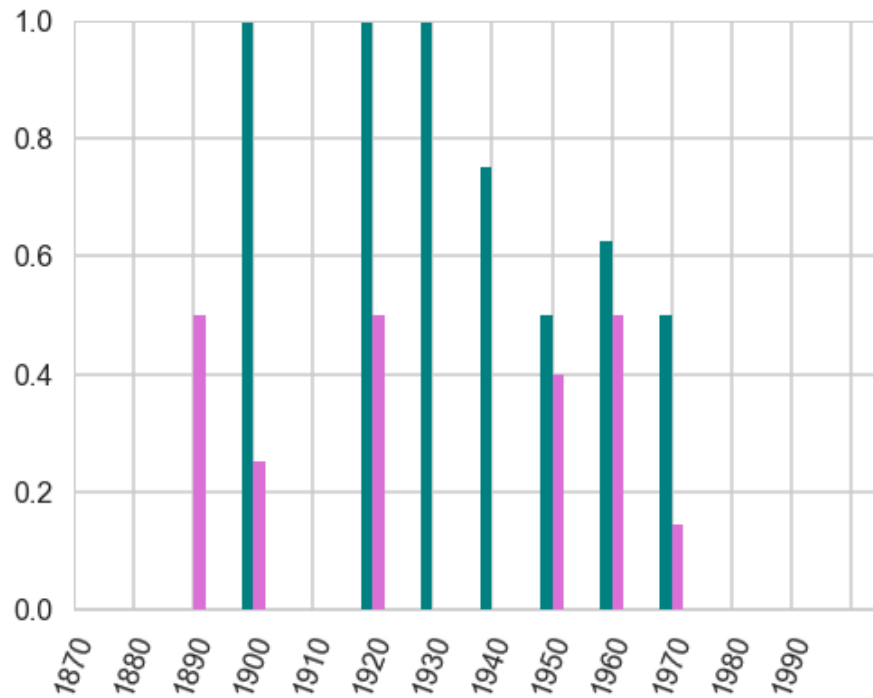


Lak by gender in second neighbor villages
(Obokh)

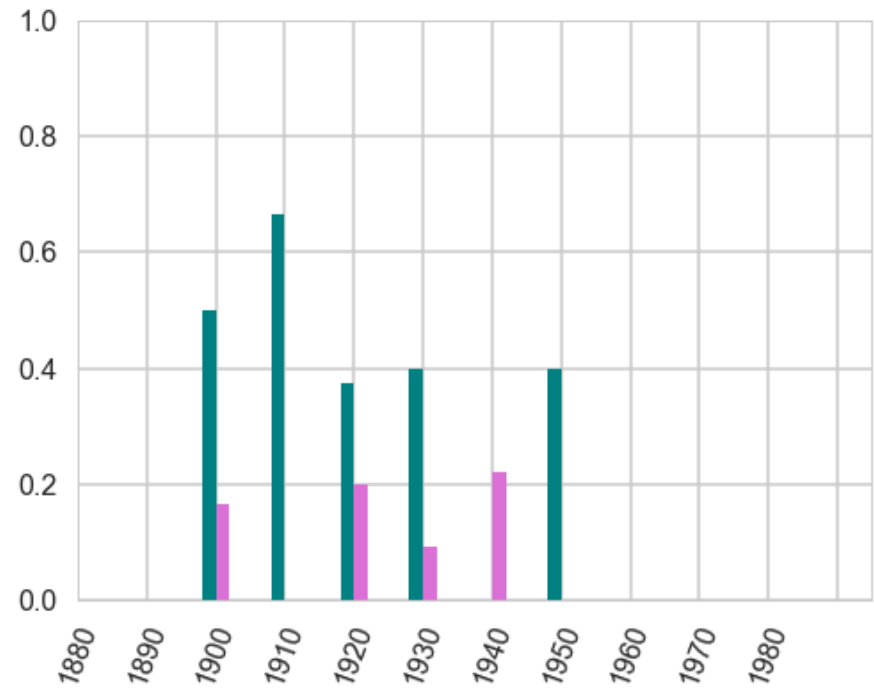
Dargwa in Tsulikana – 63%,
Dargwa in Balkhar - 32%



Second neighbor language: Dargwa



Dargwa by gender in neighboring villages
(Tsulikana)



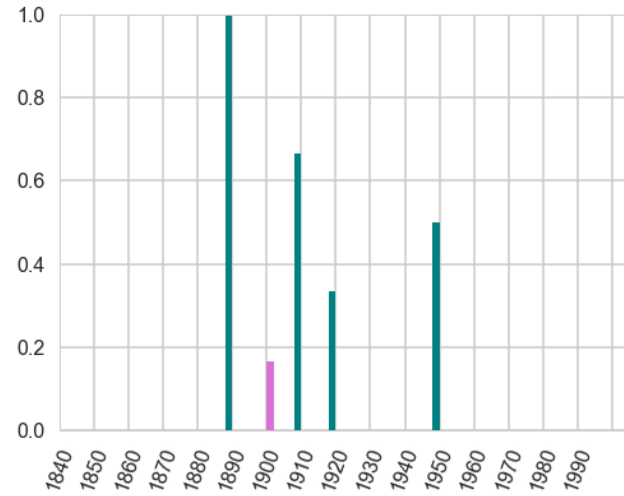
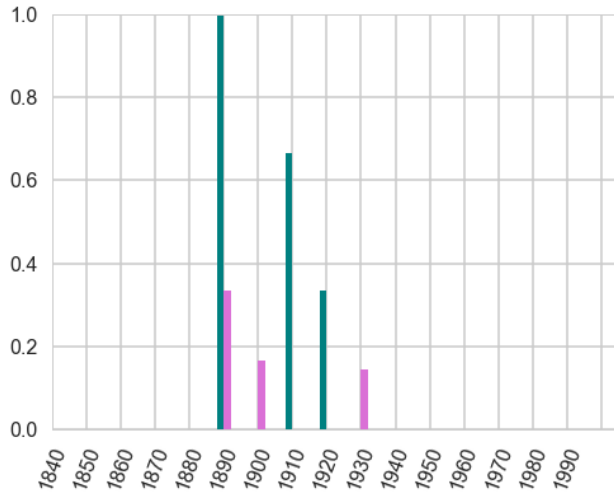
Dargwa by gender in second neighbor villages
(Balkhar)

Second neighbor languages

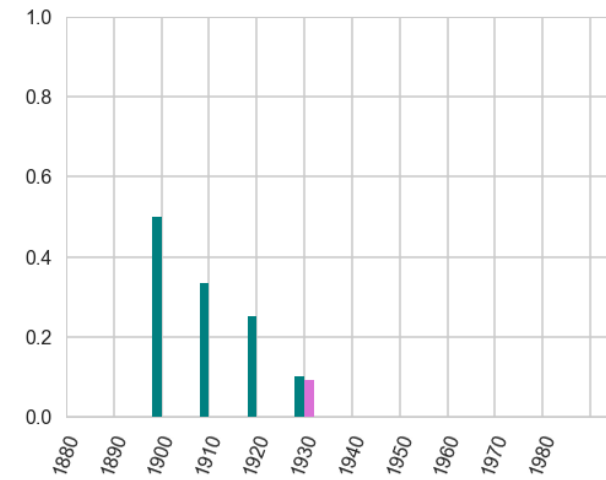
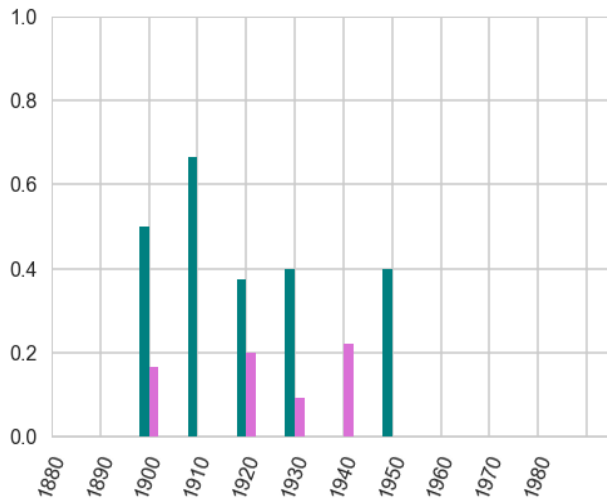
- Behave as if they are distant #1: female respondents do not speak them
- Behave as if they are distant #2: people stop to use them earlier (1930s)

SNL vs DL (Kumyk)

Obokh



Balkhar



Second neighbor languages: better sampling

- We need to expand sample
- We need to think of a better way of encoding SN-languages in database
- ... and a better way of encoding SN villages too

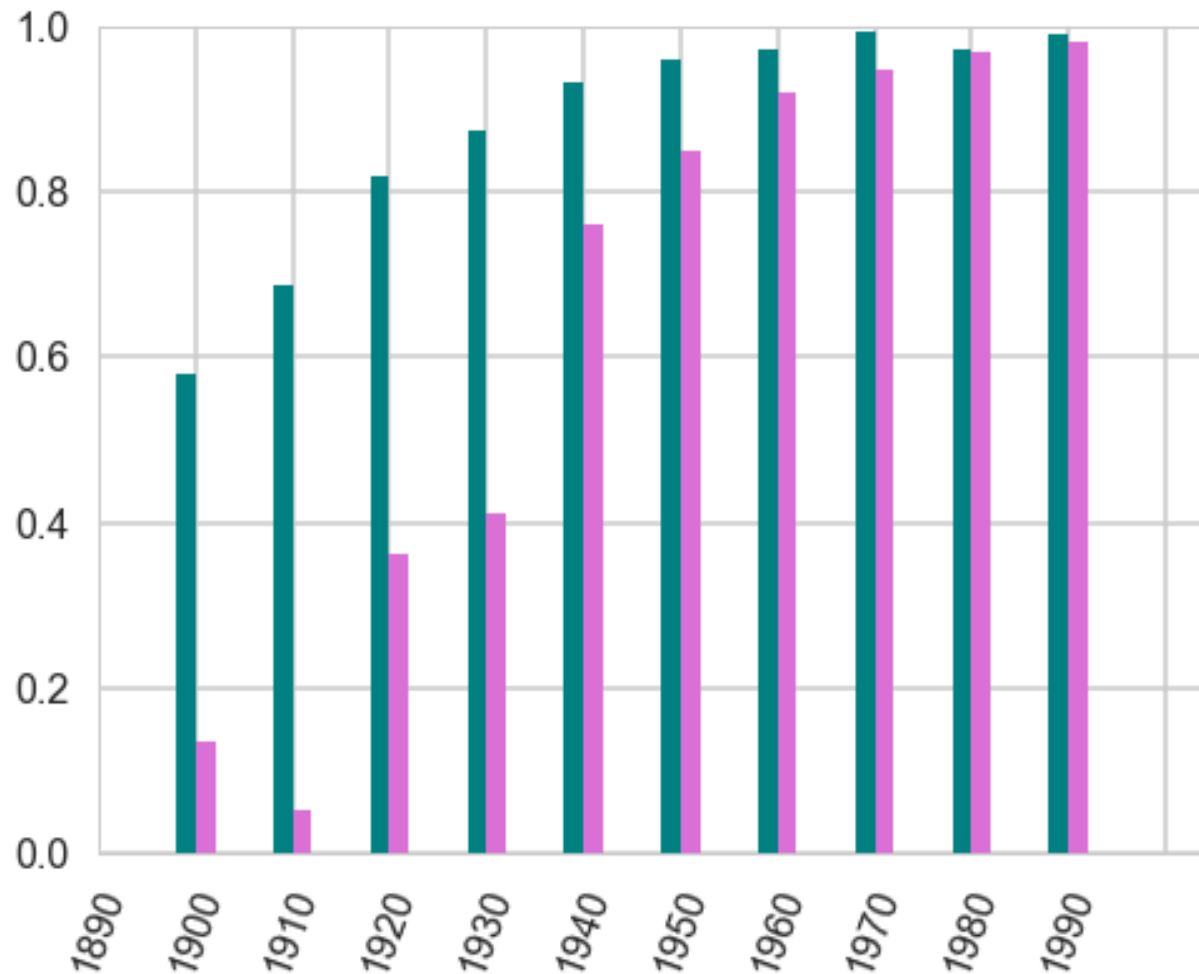
Distant languages were gendered because they were acquired through **gendered practices**: men were involved in migratory jobs, while women stayed at home with children and cattle

Russian

Russian was acquired during military service, contacts with administration, trips to Russian-speaking areas

In the generations born before 1930s, the distribution of Russian was very similar to that of distant languages: it was spoken almost exclusively by men

Russian competence among men and women through whole time period



In the 1930s, in most Daghestanian villages
Soviet schools were opened

Parents were forced to bring their children
to school (fines for non-schooling)

Parents were especially reluctant to bring
girls to school

But they had to concede



Archis building the road and Russian teacher (1930 y.b.)



People born in the 1930s went through Soviet school

80% of men, 30% of women

80% of men, 70% of women

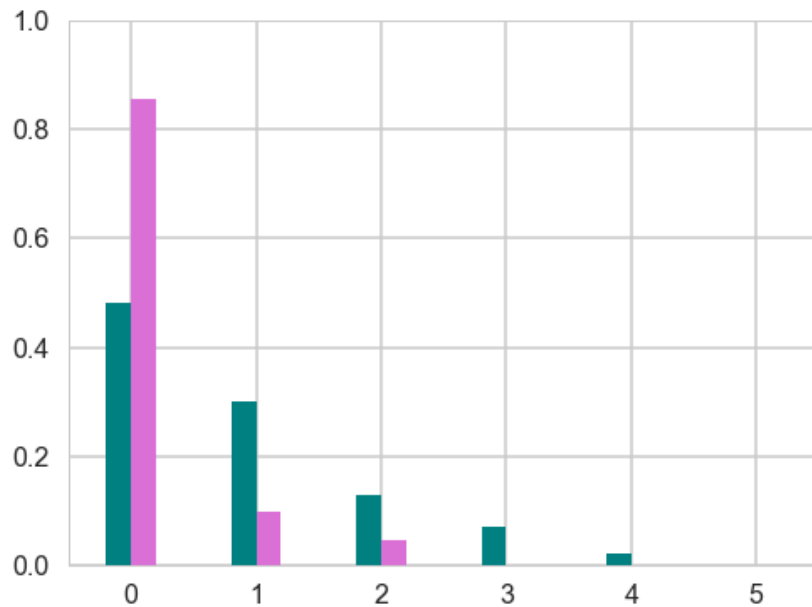


- After the 1950s years of birth, everyone speaks Russian
- Command of Russian is not gendered anymore
- Spread of Russian influenced the command of other languages

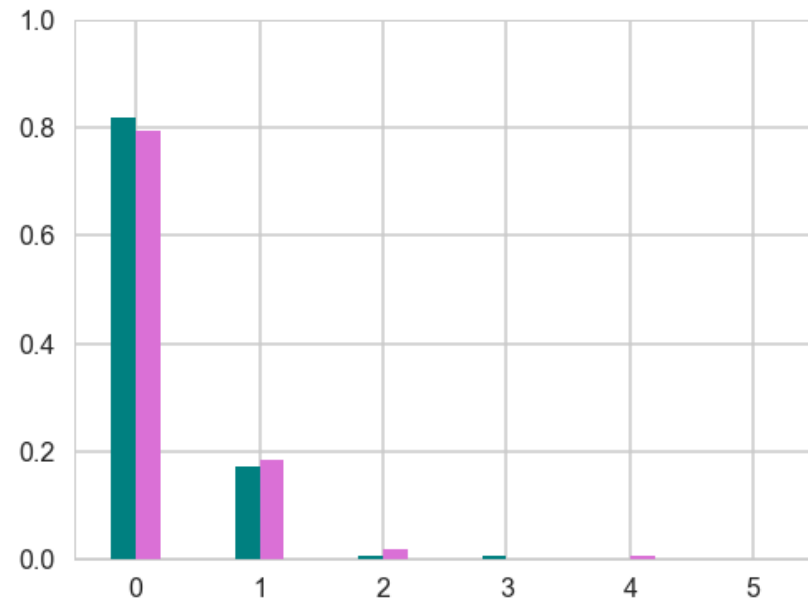
With the spread of Russian, DLs were lost

**40% of men and 10% of women spoke one DL,
10% of men and 0% of women spoke two or three DL**

10% of men and women spoke one DL, nobody spoke more than one



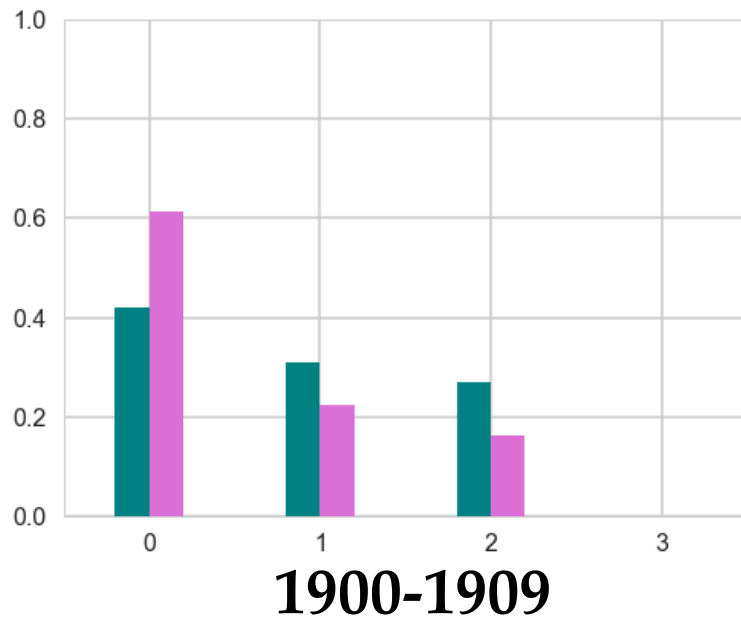
1900-1909



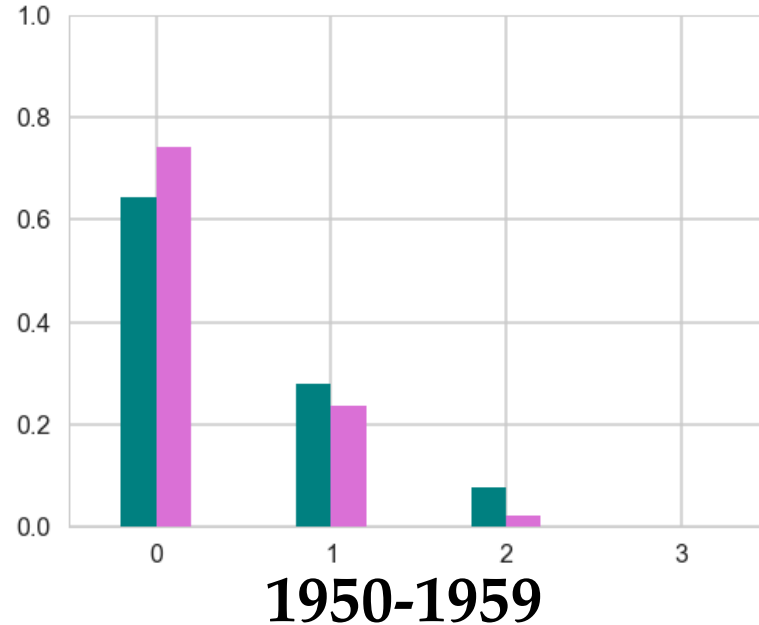
1950-1959

With the spread of Russian, NLs also declined

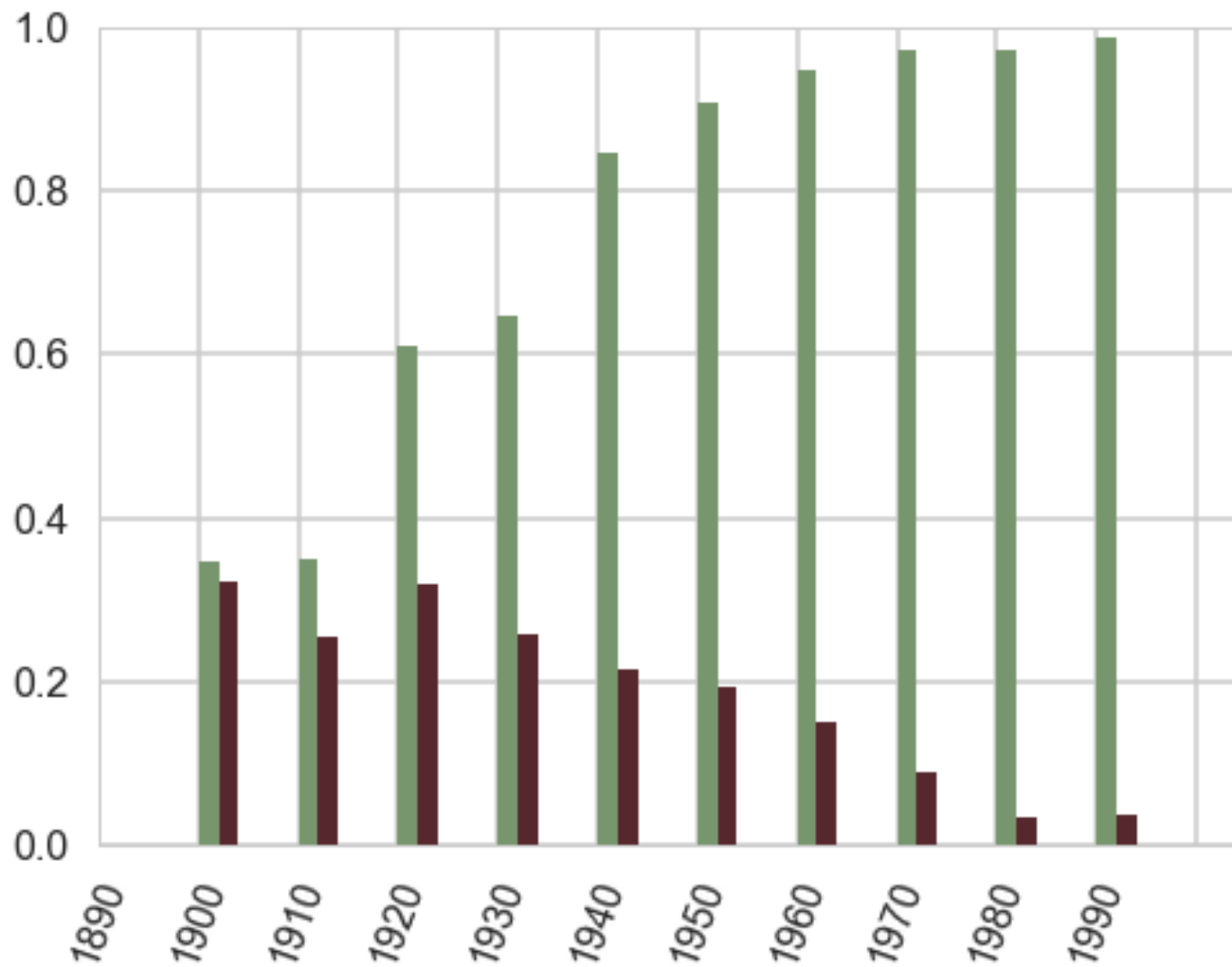
**45% of men and 40% of women spoke 1 NL,
30% of men and 15% of women spoke two NLs**



**30% of men and women spoke one NL,
5% of men and women spoke two NLs**



Knowledge of Russian and knowledge of distant languages among males



Very quickly Russian became a
lingua franca of Daghestan

Changes

In 1890 – 1910 y.b.

Distant Ls are exclusively
men-spoken

Russian is exclusively
men-spoken

Neighboring Ls are
slightly men-spoken

In 1950s y.b.

Distant Ls are almost lost

Russian is spoken by
everyone

Neighboring Ls have
declined and ceased to
be gendered

Why Daghestanian multilingualism has lost its gendered character?

Not because practices became less gendered
(transhuman shepherding continues to be the
occupation of men),

...but because a new non-gendered universal
practice evolved – school education

Through the schools, Russian spread all over
Daghestan and became a lingua franca

- Introduction – Daghestan and its multilingualism
- Problem setting
- Method and data
- Aims
- Some results
- Problems and restrictions

Problems and restrictions

The method fully relies

- on our respondents' self-assessment
- and on our respondents' memories about the multilingualism of their elder relatives

Problems with respondents' self-assessment

- Respondents may overestimate their language repertoires (boasting) or underestimate it (modesty)
- Respondents may mean different things when saying “I speak Dargwa/ Kumyk” etc.

Problems with respondents' memories about the multilingualism of their elder relatives

- Respondents often rely on community stereotypes, and extend them to their relatives (“It is known that all people of that generation spoke Lak, so my grandparents did, too.”)
- Even if they rely on their own memories, respondents remember their elder relatives at their later age; the person could have acquired a language at 50 or 60, together with the younger generation (especially important for 'waxing' languages - Russian)

Problems with our respondents' memories about their elder relatives: solutions

- Consistency of data
- Cross-checking with historical documents
- Cross-checking within the database:
self-reports VS indirect reports

Consistency: no “strange” results

For example, from 7 situations with three villages,

- **Six** show significant distance-related differences in the level of bilingualism - first neighbor always exhibit higher level of bilingualism than second neighbor
- In **one** case the distance doesn't play a role
- **Zero** cases when second neighbor shows higher level of bilingualism when first neighbor

Comparison with historical documents

Data is very sparse – several anecdotal reports from different villages by ethnographers, linguists, historians

Some comparison – roughly holds

Лавров 1953, Никольская
1952

- People from **Ikhrek**, **Mikik**, **Tsakhur** and **Gelmets** went for seasonal jobs to Georgia and some of them could speak Georgian

Our data (born before 1919)

- **Georgian** is mentioned twice in our **Gelmets** data, twice in **Ikhrek** and never in neighboring **Kina**

Some comparison – does not hold

**Genko, Tabassaran
fieldtrip of 1934**

«СКОЛЬКО-НИБУДЬ
УДОВЛЕТВОРИТЕЛЬНО
ГОВОРЯЩИХ ПО-
РУССКИ В 1934 Г. НЕ
ВСТРЕЧАЛОСЬ СРЕДИ
ТАБАСАРАНЦЕВ»

Our data

before 1915 y.b. level of
Russian

–in Khiv – 33% (from
9),

–in Laka – 30% (from
10),

–in Dyubek – 13%
(from 15)

Some comparison

Sergeeva 1967

- In the past, there were more people who spoke Lak than those who spoke Avar in Archi (Sergeeva 1967)
- According to a study in Archib by Sergeeva, in 1960
- 88% spoke Avar,
- 36 % Russian,
- 28% Lak,
- 6% - Azerbaijani

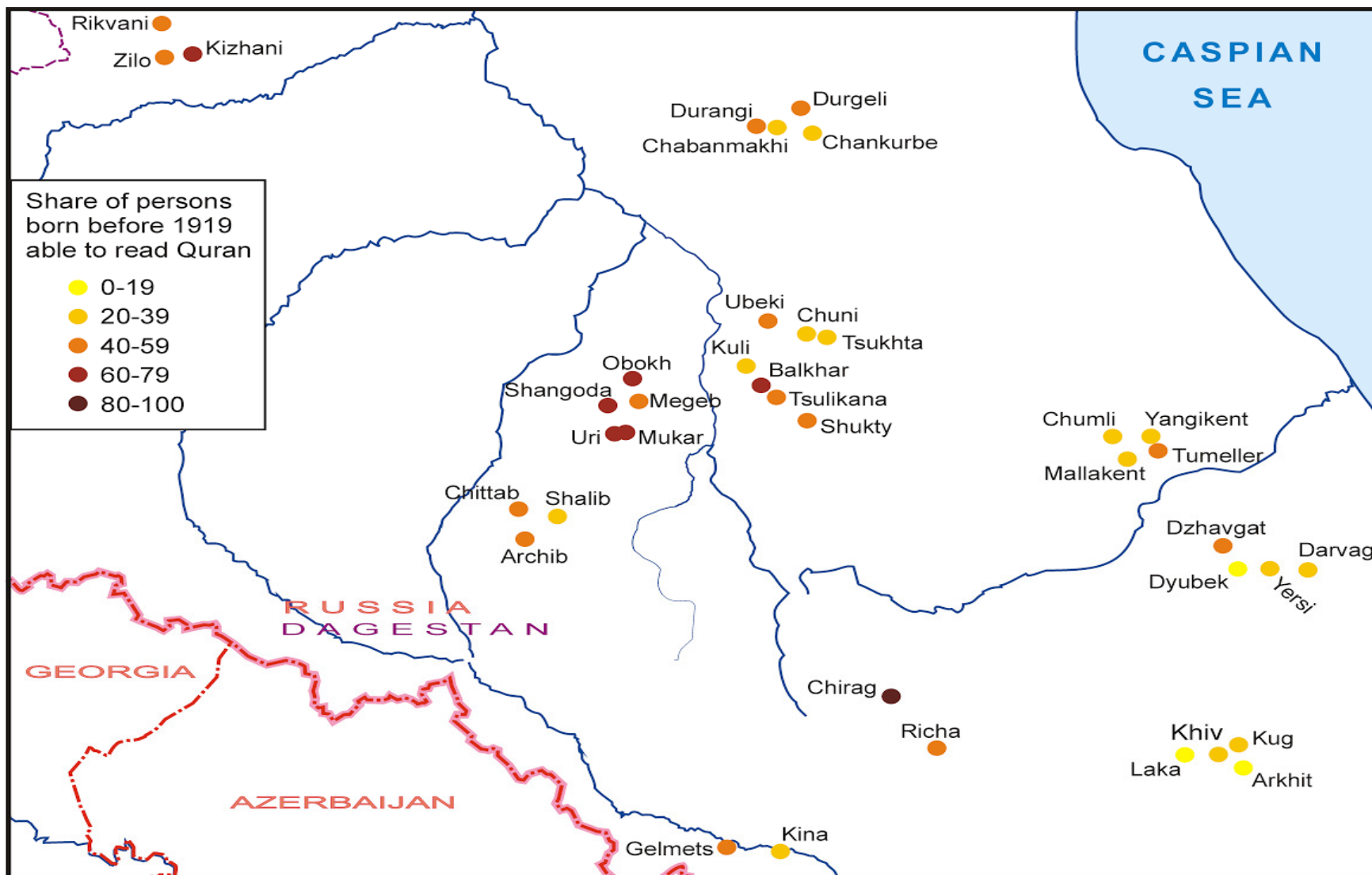
Our data

- According to our data, among people born in Archi between 1849 and 1900, 89% spoke Avar and 89% spoke Lak
- According to our data, among people born in Archi between 1890 and 1940
- 96% spoke Avar,
- 30% spoke Russian,
- 59% spoke Lak,
- 17% spoke Azerbaijani

Genko, Tabassaran fieldtrip of 1934

- «По грамотности (в то время только на азербайджанском языке) табасаранцы стояли в 1926 году на одном из последних мест среди народностей Дагестана (1,6% грамотных)»
- Before 1910 y.b. the number of literate people
 - in Khiv – 21% (14),
 - In Laka - 18% (from 11),
 - In Dubek – 16% (19)

Arabic literacy: reading Quran + reading and understanding

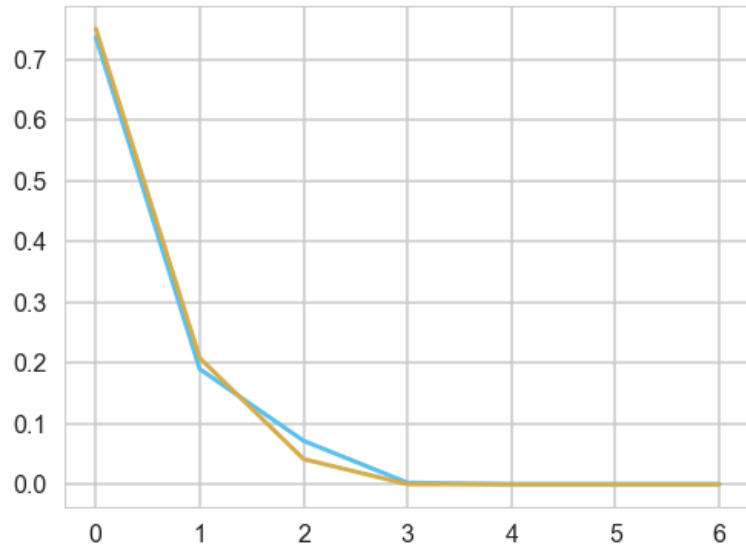


Cross-checking within database: self-reports VS indirect reports

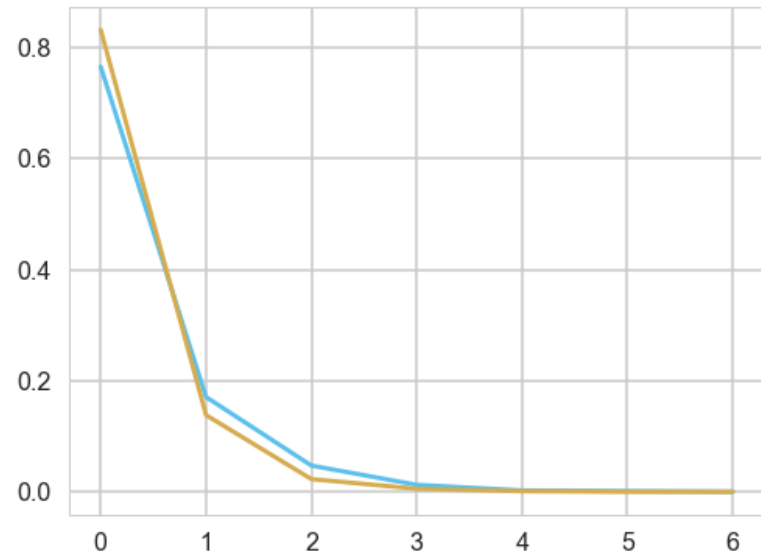
The entire population is divided in two parts: those who were approached personally (2,500 people) and those who were coded through approaching their relatives (1,000 people)

Is there any difference with respect to their assessments of language repertoire?

Self-reports vs. Indirect reports



DL



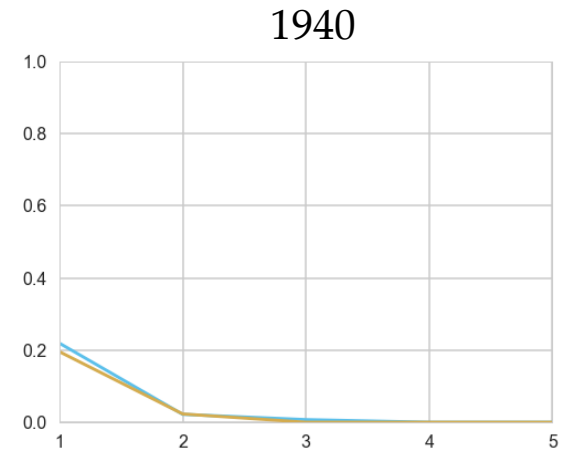
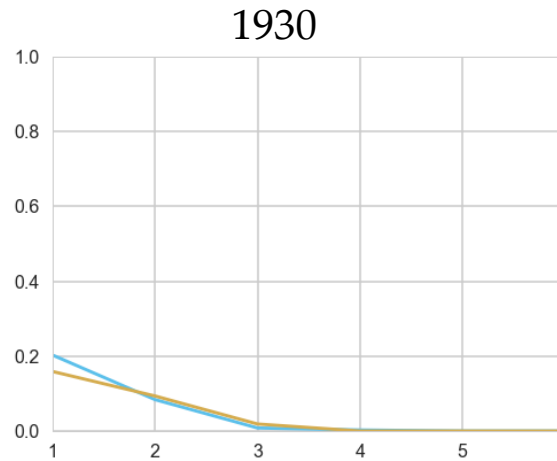
NL

Better cross-checking?

- Reduce or even drop earliest periods ✓
1000 respondents from 1930s and 1940s with 4:6 proportion
- Cross-validate each decade separately ✓
- Use some other statistic test
- Cross-check within interviews of one type

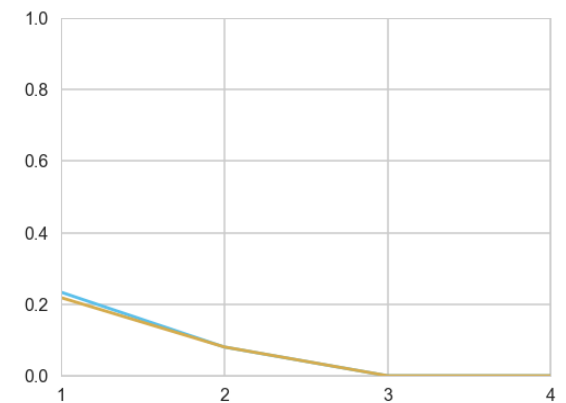
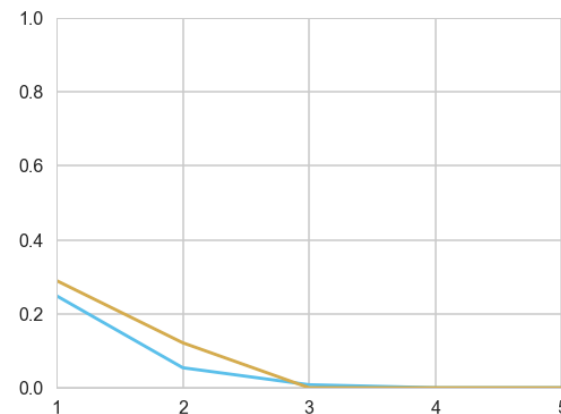
Distant
multilingualism
(1930-1940)

p-value = 3.532e-08



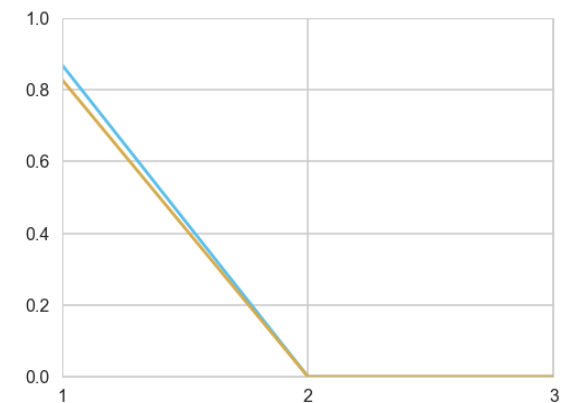
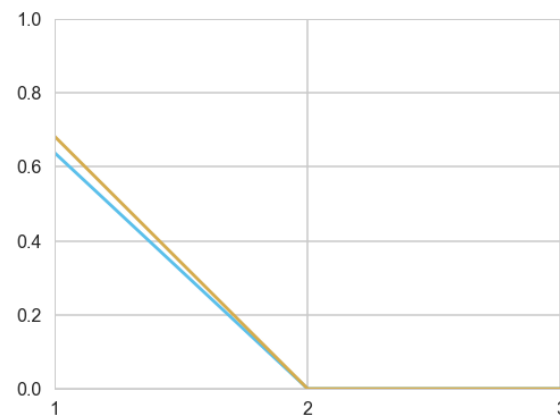
Neighbor
multilingualism
(1930-1940)

p-value = 0.01247



Russian
competence
(1930-1940)

p-value < 2.2e-16



Restrictions

The collected data show tendencies,
tendencies are meaningful, but the particular
numbers are not

Solution:

Indexes instead of numbers?

0 – 20% - low

21% - 40% - low-middle

40% - 60% - middle....

This is a collective project

- **Fieldwork:** Michael Daniel, Darya Baryl'nikova, Ilja Chechuro, Anna Djachkova, Aleksej Fedorenko, Dmitry Ganenkov, Aleksandra Khadzhiskaya, Marina Korshak, Aleksandra Kozhukhar', Marina Kustova, Yevgenij Lapin, Aleksandr Letuchiy, Aleksandra Martynova, Stepan Mikhailov, Yevgenij Mozhaev, Rasul Mutalov, Elena Nikishina, Olga Shapovalova, Semen Sheshenin, Aleksandra Sheshenina, Maria Shejanova, Anastassija Vasilisina
- **Hospitality:** Karim Musaev and his family, Anwar and Maisarat Musaevy, Kamil and his family, Hamzat and his family, Ibadulla, Akhmed, Dzhalil, Ramazan, Said Sulejmanov and his family, Sejdul, Khadizhat
- **Organization:** Rasul Mutalov
- **Maps:** Yuri Koryakov
- **Editing:** Samira Verhees
- **Everything:** Mikhail Daniel



References:

- Campbell, Lyle & Verónica Grondona. (2010). Who speaks what to whom? Multilingualism and language choice in Misión La Paz. *Language in Society* 39 (5): 617-646.
- Chirikba, Viacheslav A. (2008). The problem of the Caucasian Sprachbund. In Pieter Muysken (ed.), *From linguistic areas to areal linguistics*. 25–94. Amsterdam, Philadelphia: John Benjamins Publishing Company.
- Dobrushina, Nina. (2013). How to study multilingualism of the past: Investigating traditional contact situations in Daghestan. *Journal of sociolinguistics*, 17(3), 376-393.
- Epps, Patience. (2015). Amazonian linguistic diversity and its sociocultural correlates. In Mily Crevels & Pieter Muysken (eds), *Language Dispersal, Diversification, and Contact: A Global Perspective*, Oxford: Oxford University Press.
- Nichols, Johanna. (2013). The vertical archipelago: adding the third dimension to linguistic geography. In Peter Auer, Martin Hilpert, Anja Stukenbrock & Benedikt Szmrecsanyi (eds.), *Space in Language and Linguistics. Geographical, Interactional, and Cognitive Perspectives*. Berlin: Mouton de Gruyter.
- Thomason, Sarah G. (2001). *Language contact: an introduction*. Washington: Georgetown University Press.
- Weinreich, Uriel. (1979). *Languages in Contact: Findings and Problems*. Mouton publishers. Berlin. The Hague. New York.
- Wixman, Ronald. (1980). *Language Aspects of Ethnic Patterns and Processes in the North Caucasus*. Chicago: University of Chicago.

A photograph of a group of children and a woman sitting on the back of a truck on a dirt road. The truck is moving away from the camera, and the children are waving. The woman is standing and waving. The background shows a mountainous landscape with green hills and a stone wall on the right. The license plate of the truck is E524AT05.

EURIAS Fellowship Programme
Co-funded by Marie Skłodowska-Curie Actions,
under the 7th Framework Programme

Science Foundation of the National Research
University Higher School of Economics

Barkalla!
(Thank you)