

The speakers of minority languages are more multilingual

Nina Dobrushina and George Moroz

Linguistic Convergence Laboratory, NRU HSE, Moscow, Russia

16 April 2019

Typology of small-scale multilingualism
Laboratoire Dynamique du Langage, Lyon, France

Presentation is available here: tinyurl.com/y6jjp38y



Problem part

Data

Data obtained during interviews on language usage from about 15 fieldtrips (see [Dobrushina 2013] for methodology details) and collected into **Atlas of Multilingualism in Daghestan** [Dobrushina et al. 2017]:

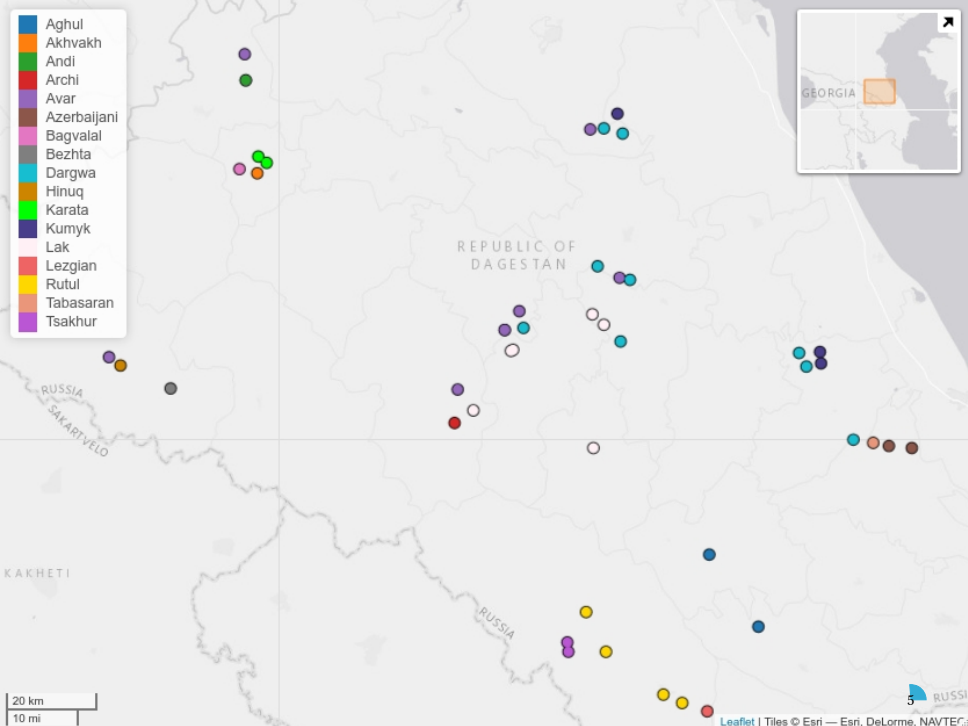
- 46 Daghestanian villages
- 24 languages (Russian excluded)

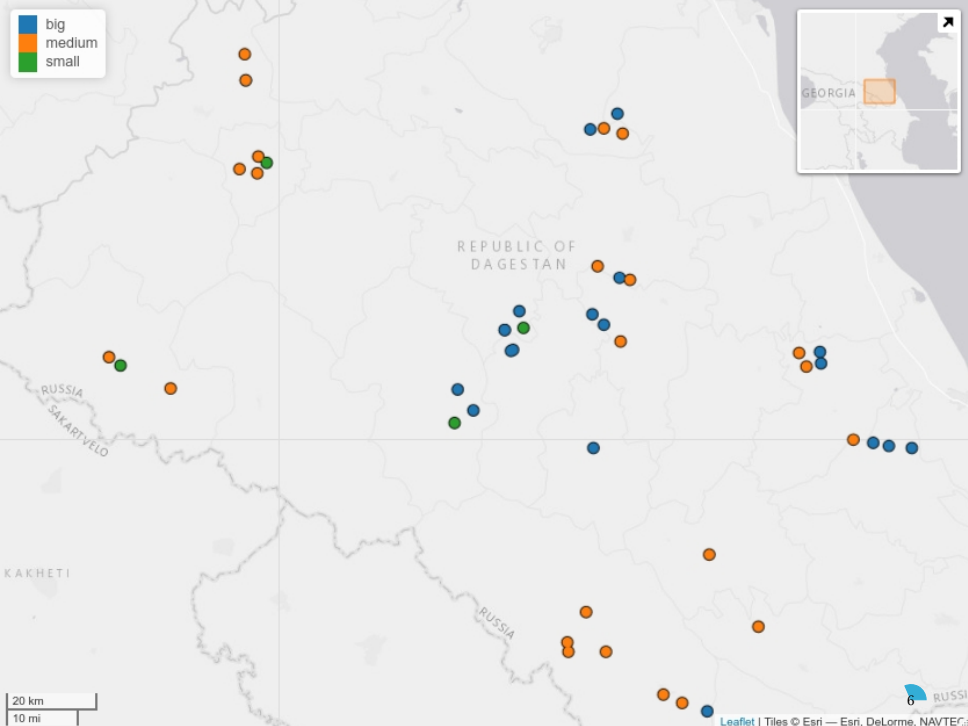
Data obtained during interviews on language usage from about 15 fieldtrips (see [Dobrushina 2013] for methodology details) and collected into **Atlas of Multilingualism in Daghestan** [Dobrushina et al. 2017]:

- 46 Daghestanian villages
- 24 languages (Russian excluded)
- 29 860 people born between 1900 and 1959
 - 14 410 females (48.2%)
 - 15 450 males (51.7%)
- variable containing number of known languages

Data obtained during interviews on language usage from about 15 fieldtrips (see [Dobrushina 2013] for methodology details) and collected into **Atlas of Multilingualism in Daghestan** [Dobrushina et al. 2017]:

- 46 Daghestanian villages
- 24 languages (Russian excluded)
- 29 860 people born between 1900 and 1959
 - 14 410 females (48.2%)
 - 15 450 males (51.7%)
- variable containing number of known languages
- we grouped all languages into three categories according to the nowadays number of speakers
 - **big** — 100 000 speakers and more
 - **medium** — 10 000–30 000 speakers
 - **small** — one village languages, 1 000–2 000 speakers





1900



REPUBLIC OF
DAGESTAN

RUSSIA
SAKARTVELO

KAKHETI

RUSSIA



1910



REPUBLIC OF
DAGESTAN

RUSSIA
SAKARTVELO

KAKHETI

RUSSIA

20 km

10 mi

1920



REPUBLIC OF
DAGESTAN

RUSSIA
SAKARTVELO

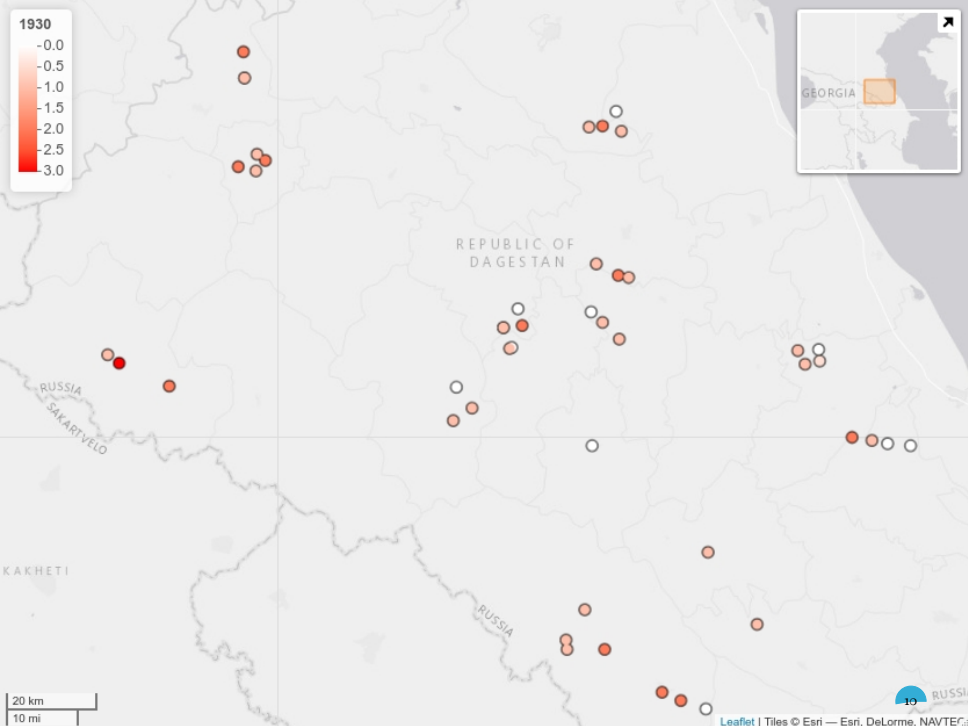
KAKHETI

RUSSIA

20 km

10 mi

1930



1940



REPUBLIC OF
DAGESTAN

RUSSIA
SAKARTVELO

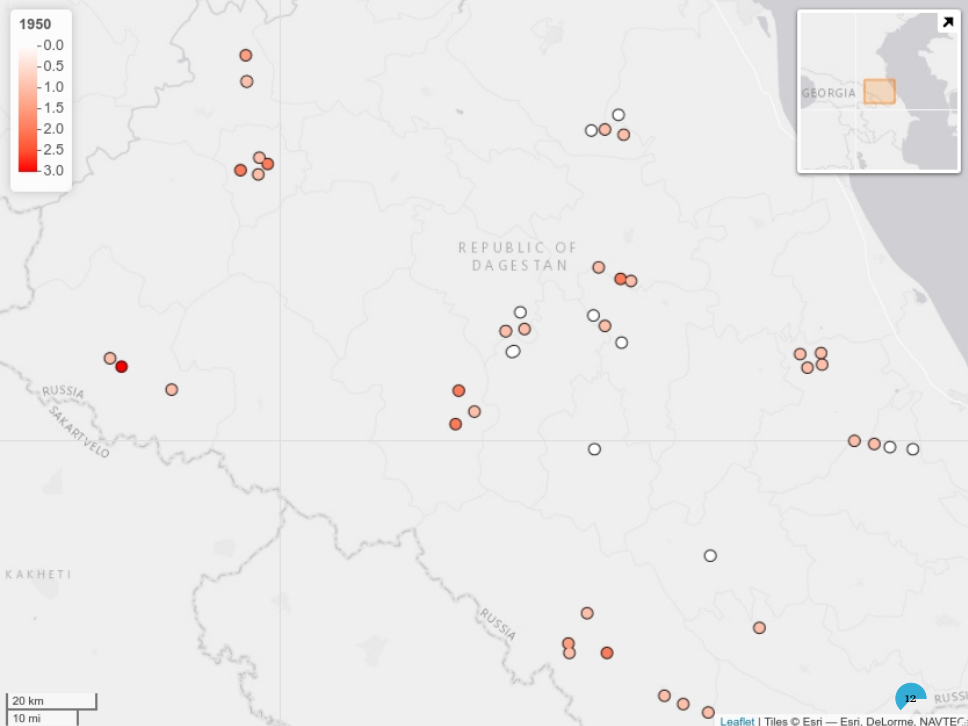
KAKHETI

RUSSIA



1950

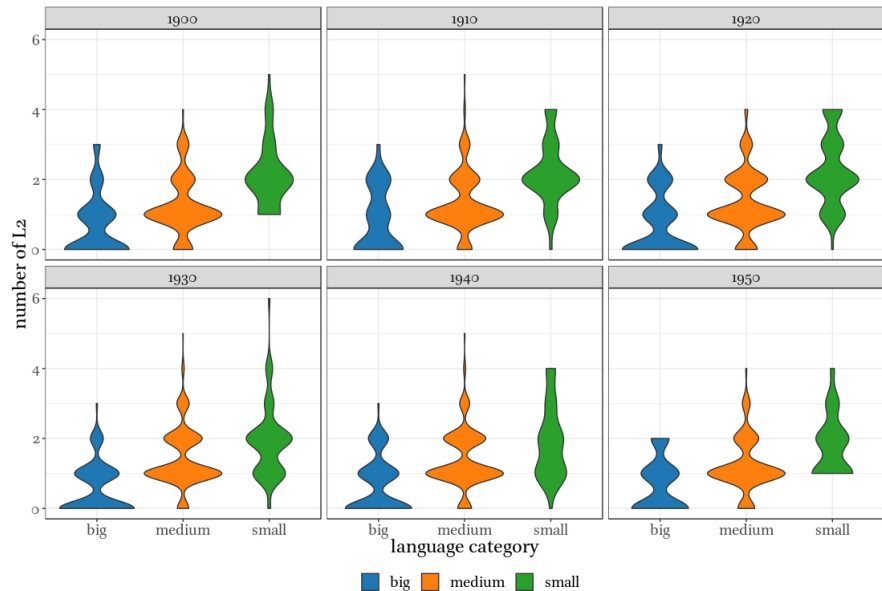
-0.0
-0.5
-1.0
-1.5
-2.0
-2.5
-3.0



20 km
10 mi



Number of L2 in each village by decade and language category



Poisson Mixed Effects Model

Generalized linear mixed model fit by maximum likelihood (Laplace Approximation)

Family: poisson (log)

Formula: sum_langs ~ status + (1 | residence.en) + (1 | decade)

Data: df

AIC	BIC	logLik	deviance	df.resid
7502.1	7532.5	-3746.1	7492.1	3195

Scaled residuals:

Min	1Q	Median	3Q	Max
-1.4215	-0.4501	-0.1821	0.3202	3.7685

Random effects:

Groups	Name	Variance	Std.Dev.
residence.en	(Intercept)	0.3679222	0.60657
decade	(Intercept)	0.0004188	0.02046

Number of obs: 3200, groups: residence.en, 46; decade, 6

Fixed effects:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	0.7151	0.3052	2.343	0.0191 *
statusmedium	-0.5011	0.3311	-1.514	0.1301
statusbig	-1.5692	0.3412	-4.599	0.00000424 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:

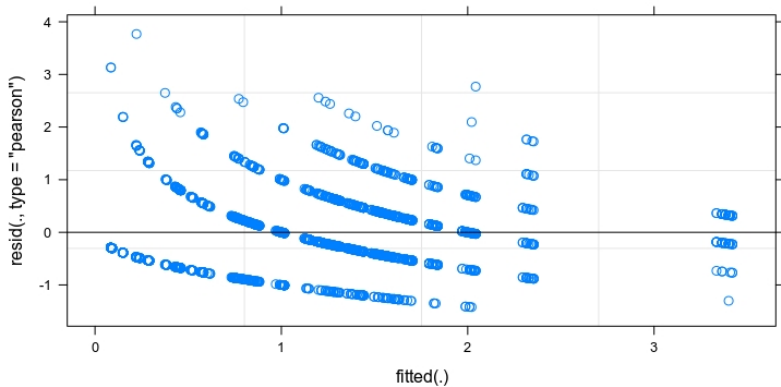
	(Intr)	sttsmd
statusmedim	-0.921	
statusbig	-0.893	0.824

Overdispersion test

dispersion ratio	= 0.4329
Pearson's Chi-Squared	= 1383.2471
p-value	= 1.0000

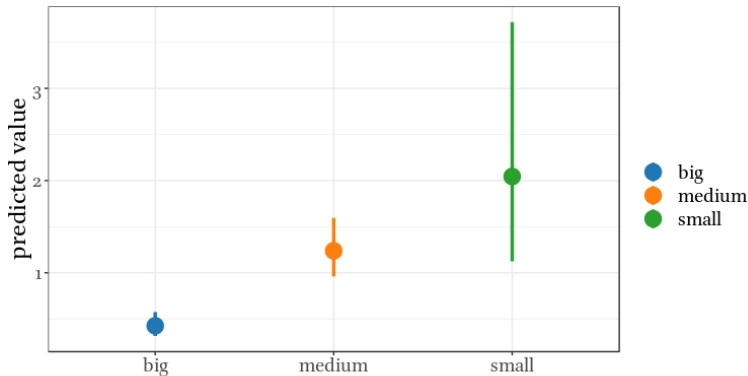
No overdispersion detected.

Poisson Mixed Effects Model: Residuals



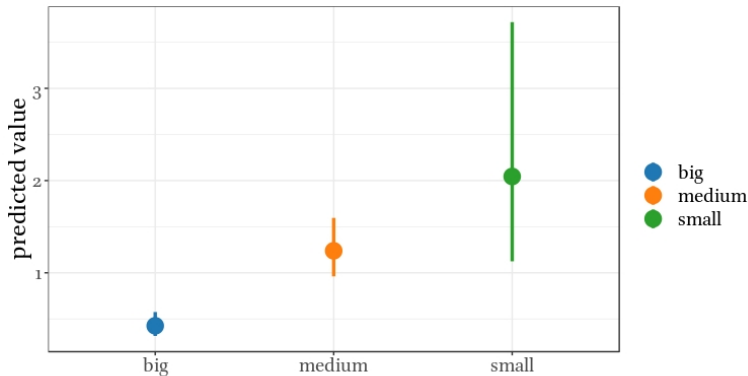
Conclusions:

- The variable language size is statistically significant.
- Obtained coefficients could be interpreted as following:



Conclusions:

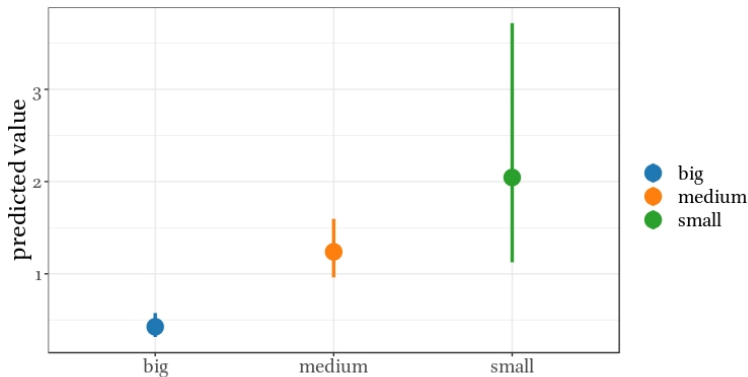
- The variable language size is statistically significant.
- Obtained coefficients could be interpreted as following:



- Special case: Chuni

Conclusions:

- The variable language size is statistically significant.
- Obtained coefficients could be interpret as following:



- Special case: Chuni
- It is not the only case of Daghestanian languages:
 - Circassians in Arabic communities in Israel [Kreindler et al. 1995]
 - Abaza in Circassian communities in Russia (personal observations)

Send us a letter!
nina.dobrushina@gmail.com
agricolamz@gmail.com

Presentation is available here: tinyurl.com/y6jjp38y



All visualisation and statistical analysis were made in R version 3.5.3 [R Core Team 2019] with packages ggplot2 [Wickham 2016], lme4 [Bates et al. 2015],
lingtypology [Moroz 2017]

References

- Bates, D., M. Mächler, B. Bolker, and S. Walker (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1), 1–48.
- Dobrushina, N. (2013). How to study multilingualism of the past: Investigating traditional contact situations in daghestan. *Journal of sociolinguistics* 17(3), 376–393.
- Dobrushina, N., D. Staferova, and A. Belokon (2017). Atlas of Multilingualism in Daghestan Online. <https://multidagestan.com/about>. Accessed: 2019-04-14.
- Kreindler, I., M. Bensoussan, E. Avinor, and C. Bram (1995). Circassian israelis: Multilingualism as a way of life. *Language, Culture and Curriculum* 8(2), 149–162.
- Moroz, G. (2017). *lingtypology: easy mapping for Linguistic Typology*.
- R Core Team (2019). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.