Small-scale multilingualism through the prism of lexical borrowing

Ilia Chechuro*, Michael Daniel*, Samira Verhees*
*Linguistic Convergence Laboratory, HSE University (Moscow)

One of the goals of contact linguistics is to offer tools to reconstruct unknown historical patterns of interaction between ethnic groups based on the observed structural outcomes of language contact. In this reconstruction, a necessary intermediate stage is modelling the historical patterns of bilingualism. Based on a database of loanwords in minority languages of Daghestan, we show how simple counts of lexical borrowings may shed light on patterns of small-scale multilingualism in the past.

Language contact is channelled through people, via their use of multiple languages (Milroy 1997: 311). Patterns of bilingualism vary considerably from society to society and to a great extent shape the outcomes of language contact. It is assumed that the amount of structural influence from one language to another is a function of the intensity of bilingualism (Thomason & Kaufman 1992: 74-76). When lexical influence is at work, the more intense the bilingualism, the higher the amount of loanwords from the donor to the recipient language is supposed to be (e.g. Scotton & Okeju 1973, Watson 2018).

The Republic of Daghestan, the area our data comes from, features high language density, diverse patterns of bilingualism (Dobrushina 2011, 2013) and widespread feature sharing (Comrie 2008, Klimov 1978, Chirikba 2008, Tuite 1999). Local languages include languages from various branches of East Caucasian, two Turkic and one Iranian language. Historically, Avar (East Caucasian) and, to a lesser extent, Kumyk (Turkic) were used as lingua francas in the northwestern part of the area, while Azerbaijani (Turkic) was used in the south. These L2s are in the process of being ousted by the use of Russian in interethnic communication. The dominance of two specific lingua francas in the respective areas is clearly seen in Table 1. We counted the amount of loanwords within a fixed and relatively short list of concepts, DagLoans (160 concepts, see Authors (in preparation), cf. (Haspelmath & Tadmor 2009)). We elicited the list from speakers of different languages, often interviewing several speakers of the same language and even of the same village variety of the language.

Numbers provided in Table 1 are median percentages of Turkic loanwords in the elicitations of the DagLoans list from villagers of the Rutul (south) and Botlikh, Tsunta and Akhvakh (north-west) districts of Daghestan. As the table shows, the Rutul district lacks loanwords from Avar completely, while the amount of Turkic (probably Kumyk[1]) loanwords in the Botlikh, Akhvakh and Tsunta districts is clearly lower than in the Rutul district. This suggests that bilingualism in one of the Turkic languages, Azerbaijani in the south, was stronger than bilingualism in either Azerbaijani or Kumyk in the north-west.

However, the table may be misleading as to how drastic this difference was. Indeed, some Turkic loans may have arrived to the minority languages via a major local language - Avar in the north-west or Lezgian in the south - rather than directly from Azerbaijani or Kumyk. To investigate this possibility, we distinguished between those Turkic loans that are attested in a minority language alone vs. those that are also attested in Avar or Lezgian, respectively. The results are shown in Table 2. As the table suggests, in the south, the part of the Turkic loans that is not shared with Lezgian is much higher than the part of the Turkic loans that is not shared with Avar in the north-west. The contact with Turkic in the south must have been direct, while the contact with Turkic in the north-west might have been mostly mediated by Avar.

We cannot know the details of the individual loanwords' journeys to the recipient languages. However, based on our counts, we conclude that not only the amount of Turkic loans is much lower in the north-west than it is in the south, but also, at least for some loans, the donor language could have been Avar, rather than Kumyk or Azerbaijani. The spread of Turkic bilingualism across Andic and Tsezic villages is likely to have

---

[1] It can be difficult to tell apart Azerbaijani from Kumyk borrowings based on the appearance of the word, so they are not distinguished in the counts (similar to the approach in, e.g., Comrie & Khalilov (2009) and Chumakina (2009)).

been even lower than the counts in Table 1 suggest — if present at all. This conclusion is in full accordance with an independent reconstruction of the traditional small-scale bilingualism in (Dobrushina et al. 2017).

Appendix

Table 1. Median percentages of borrowings in the four districts

District	Lingua franca	No. of villages	Languages	No. of speakers	Turkic loans (median %)	Avar loans (median %)
Botlikh	Avar	2	Andi (Andic)	3	12	23
Akhvakh		3	Karata, Tukita, Tadmagitl (Andic)	7	9	27
Tsunta		2	Tsez and Bezhta (Tsezic)	4	10	20
Rutul	Azerbaijani	5	Rutul and Tsakhur (Lezgic)	14	25	0

Table 2. Turkic influence: mediated or not?

	a. Districts in the north-west (Botlikh, Akhvakh, Tsunta)			b. District in the south (Rutul)		
	not in Avar	also in Avar			not in Lezgian	also in Lezgian
Rikvani	1 (5,6%)	17 (94,4%)		Ikhrek	14 (35%)	26 (65%)
Zilo	4 (22,2%)	14 (77,8%)		Kina	13 (32,5%)	27 (67,5%)
Karata	1 (6,7%)	14 (93,3%)		Kiche	12 (32,4%)	25 (67,6%)
Tadmagitl	3 (20%)	12 (80%)		Helmets	21 (42,8%)	28 (57,2%)
Tukita	1 (6,7%)	14 (93,3%)		Mikik	28 (50%)	28 (50%)
Kidero	0 (0%)	16 (100%)				
Bezhta	2 (18,2%)	9 (81,8%)				

References

Chirikba, V. A. (2008). The problem of the Caucasian Sprachbund. in Muysken, P. (Ed.). (2008). *From linguistic areas to areal linguistics* (Vol. 90). John Benjamins Publishing, 25-93.

Chumakina, M. (2009). Loanwords in Archi, a Nakh-Daghestanian language of the North Caucasus. *Loanwords in the world's languages. A comparative handbook*, 430-443.

Comrie, B. (2008). Linguistic diversity in the Caucasus. *Annual Review of Anthropology*, *37*, 131-143.

Comrie, B., & Khalilov, M. (2009). Loanwords in Bezhta, a Nakh-Daghestanian language of the North Caucasus. *Loanwords in the World's Languages: A Comparative Handbook. Berlin: De Gruyter Mouton*, 414-429.

Dobrushina, N. (2011). Mnogojazychije v Dagestane konca XIX - nachala XXI veka: popytka kolichestvennoj ocenki. Voprosy jazykoznanija, (4), 61-80.

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., Staferova D., Belokon A. (Eds.). (2017). Atlas of Multilingualism in Daghestan Online. Linguistic Convergence Laboratory, HSE. (Available online at https://multidagestan.com, accessed on 2018-11-28.)

Haspelmath, M., & Tadmor, U. (Eds.). (2009). *Loanwords in the world's languages: a comparative handbook*. Walter de Gruyter.

Klimov, G. (1978). *Strukturnyje osobennoste kavkazskix jazykov*. Moskva, Nauka.

Milroy, J. (1997). Internal vs external motivations for linguistic change. *Multilingua — Journal of Cross-Cultural and Interlanguage Communication*, *16*(4), 311-324.

Scotton, C. M., Okeju, J. (1973). Neighbors and lexical borrowings. *Language*, 871-889.

Thomason, S. G., Kaufman, T. (1992). *Language contact, creolization, and genetic linguistics*. University of California Press.

Tuite, K. (1999). The myth of the Caucasian Sprachbund: the case of ergativity. *Lingua*, *108*(1), 1-29.

Watson, R. (2018). Patterns of lexical correlation and divergence in Casamance. *Language & Communication.*