

Variation in two dictionaries of Botlikh

George Moroz, Chiara Naccarato, Samira Verhees
Linguistic Convergence Laboratory at NRU HSE Moscow

Документирование языков и диалектов коренных
малочисленных народов России
14–16.10.2019 - St.Petersburg

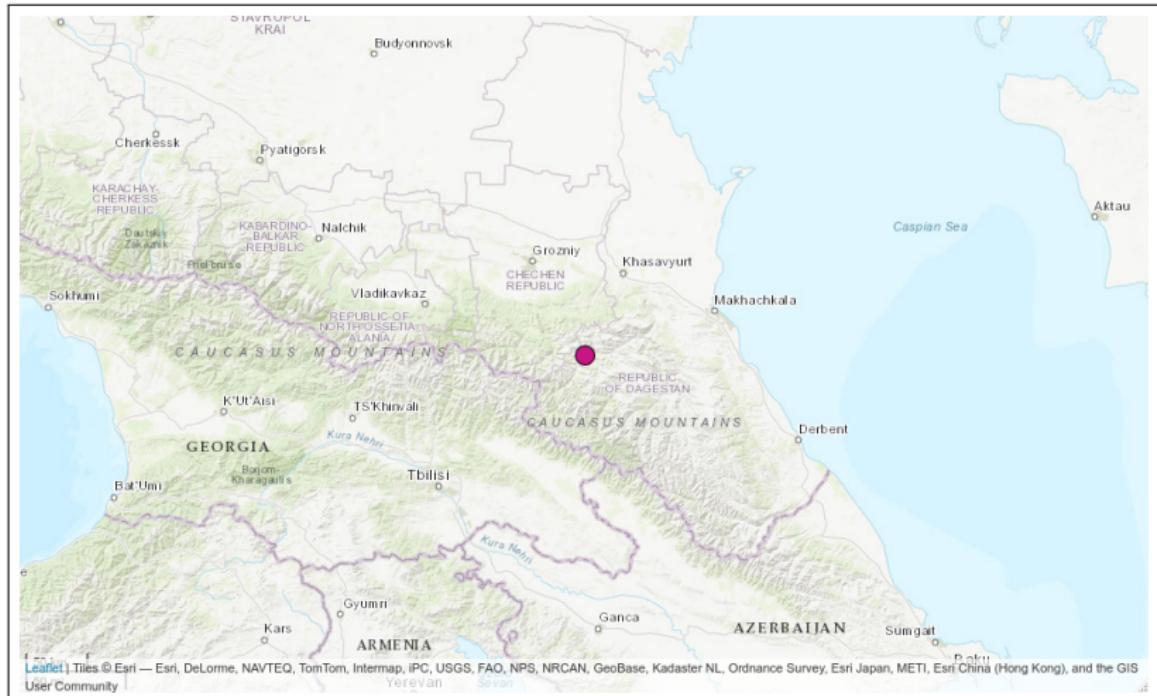
Botlikh

- ▶ Botlikh < Andic group < East Caucasian language family
- ▶ Unwritten (can be written with extended Cyrillic script for Avar)
- ▶ ~5000-8000 speakers
- ▶ Mostly spoken in 3 villages in northwestern Daghestan (Russian Federation): Botlikh, Miarso, Ashino, (Ankho); minor dialectal differences
- ▶ Opinions vary on the language's vitality — it is still passed on to children and spoken at home, but some families are shifting to Russian

Botlikh

- ▶ Botlikh < Andic group < East Caucasian language family
- ▶ Unwritten (can be written with extended Cyrillic script for Avar)
- ▶ ~5000-8000 speakers
- ▶ Mostly spoken in 3 villages in northwestern Daghestan (Russian Federation): Botlikh, Miarso, Ashino, (Ankho); minor dialectal differences
- ▶ Opinions vary on the language's vitality — it is still passed on to children and spoken at home, but some families are shifting to Russian
- ▶ One full reference grammar in Georgian ([Gudava 1962](#))
- ▶ Two dictionaries:
[\(Saidova & Abusov 2012\)](#) and [\(Alekseev & Azaev 2019\)](#)

Botlikh



Two dictionaries

(Saidova & Abusov 2012)

- ▶ Compiled in the 2000s by a native speaker (M.G. Abusov) and an experienced linguist (P.A. Saidova)
- ▶ Mostly Botlikh with some notes on Miarso

Two dictionaries

(Saidova & Abusov 2012)

- ▶ Compiled in the 2000s by a native speaker (M.G. Abusov) and an experienced linguist (P.A. Saidova)
- ▶ Mostly Botlikh with some notes on Miarso

(Alekseev & Azaev 2019)

- ▶ Compiled in the 1960s / 1970s by a native speaker / philologist (X.G. Azaev) and later (in the 2000s) systematized by an experienced linguist (M.E. Alekseev)
- ▶ Subsequently edited by T.A. Maisak and scheduled for posthumous publication later this year
- ▶ Botlikh only

Two dictionaries

- ▶ Dictionaries were compiled *independently* of each other
- ▶ ~8,000 headwords ([Saidova & Abusov 2012](#)) vs. ~9,000 ‘words and expressions’ ([Alekseev & Azaev 2019](#))
- ▶ No metadata on the speakers consulted
- ▶ Data from ([Alekseev & Azaev 2019](#)) was collected several decades earlier, but M.G. Abusov also consulted elderly speakers with the aim of collecting archaic vocabulary (p.c.)

Two dictionaries

- ▶ Dictionaries were compiled *independently* of each other
- ▶ ~8,000 headwords ([Saidova & Abusov 2012](#)) vs. ~9,000 ‘words and expressions’ ([Alekseev & Azaev 2019](#))
- ▶ No metadata on the speakers consulted
- ▶ Data from ([Alekseev & Azaev 2019](#)) was collected several decades earlier, but M.G. Abusov also consulted elderly speakers with the aim of collecting archaic vocabulary (p.c.)
- ▶ At first glance, the two resources seemed to display variation

Two dictionaries

Aim of the present investigation

Compare the two resources, see to what extent they overlap and diverge,
and verify whether they display systematic variation

...

Two dictionaries

Aim of the present investigation

Compare the two resources, see to what extent they overlap and diverge,
and verify whether they display systematic variation

...

and *hopefully*, explain the kind of variation we found
(idiolectal, diachronic, etc.)

Data

Merge dictionaries (doc to xls table) through a painstaking process of unification ([George Moroz](#)).

Data

Nouns (genitive, plural) and verbs (habitual and aorist)

lemma	pos	noun_gen	noun_pl	verb_prs	verb_pst	reference
арбаглзыв	noun	-лъи				Saidova, Abusov
арджан	noun	-алъи	-			Alekseev 2006
арёнда	noun	-лъи				Saidova, Abusov
аржаж/áй	verb			-аймалé	-ó	Saidova, Abusov
аржáй	verb			-é	-o	Alekseev 2006
арж/áй	verb			-аймале	-ó	Saidova, Abusov
аржар	noun	-лъи	-дé			Alekseev 2006
аржí	verb			-é	-a	Alekseev 2006
арж/и	verb			-e	-o	Saidova, Abusov

Data

Nouns (genitive, plural) and verbs (habitual and aorist)

lemma	pos	noun_gen	noun_pl	verb_prs	verb_pst	reference
арбаглзыв	noun	-лъи				Saidova, Abusov
арджан	noun	-алъи	-			Alekseev 2006
арёнда	noun	-лъи				Saidova, Abusov
аржаж/áй	verb			-аймалé	-ó	Saidova, Abusov
аржáй	verb			-é	-o	Alekseev 2006
арж/áй	verb			-аймале	-ó	Saidova, Abusov
áржар	noun	-лъи	-дé			Alekseev 2006
аржí	verb			-é	-a	Alekseev 2006
арж/и	verb			-e	-o	Saidova, Abusov

From ~15,000-17,000 lexical entries, we automatically extracted 7,583 nouns and verbs, further manual correction

Parameters

- ▶ Phonology: word similarity, stress patterns
- ▶ Nominal morphology: formation of genitive case and plural forms
- ▶ Verbal morphology: formation of basic tenses

Phonology

For phonological analysis we used only those words which were found in both dictionaries (1,688).

- ▶ 1,309 words are the same
- ▶ 254 words have difference in stress: *ans:í* ‘warm up’ (A&A) vs. *áns:i* (S&A)
- ▶ and only 125 words have some segmental differences
 - ▶ Some differences could be typos (*kalχoznik* ‘kolkhoz member’ (A&A) vs. *kalχoznik* (S&A), *lah* (лагъ) ‘slave’ (A&A) 2006 vs. *laх* (лагъ) (S&A))
 - ▶ Multiple cases where there is a small difference that could be explained either as a typo or in terms phonological variation (*čuhí* ‘to run’ (A&A) vs. *čuhí* (S&A), *kusu* ‘cherry plum’ (A&A) vs. *kus:u* (S&A))
 - ▶ Multiple cases where Russian borrowings were adopted differently (*pojiz* ‘train’ (A&A) vs. *poezd* (S&A), *biton* ‘milk can’ (A&A) vs. *bitun* (S&A))

Nominal morphology

Formation of the plural

- ▶ A suffix is attached to the absolute stem: *na* ‘thing’ < *na-bałi* ‘things’
- ▶ With stems ending in a consonant the vowel *-a-* is often inserted before the suffix: *majmalak* ‘monkey’ < *majmalak-a-bałi* ‘monkeys’
- ▶ With stems ending in a vowel alternation can occur: *ruša* ‘tree’ < *ruši-bałi* ‘trees’, *salu* ‘tooth’ < *sala-bałi* ‘teeth’, *burałi* ‘pitcher’ < *burała-bałi* ‘pitchers’
- ▶ Among the most common suffixes are: *-bałi* (and variants, e.g. *-zabałi*, *-małi*, etc.), *-de*, *-(w)e*

Nominal morphology

Case declension (core cases)

- ▶ I type — the stem does not change when a suffix is attached (mostly stems ending in a vowel and masdars): *babu* ‘mom’ < *babu-ti* (genitive), *masir* ‘measurement’ < *masir-ti* (genitive)
- ▶ II type — case suffixes are attached to the oblique stem of the noun (mostly stems ending in a consonant, sometimes stems ending in a vowel): *askar* ‘army’ < *askar-a-ti* (genitive), *din* ‘religion’ < *din-i-ti* (genitive), *ima* ‘father’ < *imu-ti* (genitive)

Nominal morphology

- ▶ Both dictionaries report both the **genitive** and the **plural** suffix
- ▶ We used this information to study the productivity of such suffixes, and to see whether the two dictionaries display variation

Nominal morphology

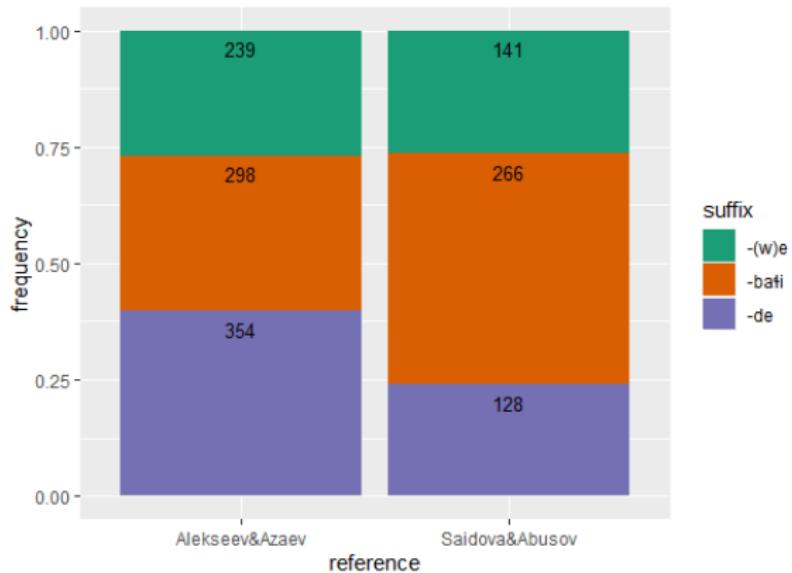
- ▶ Both dictionaries report both the **genitive** and the **plural** suffix
- ▶ We used this information to study the productivity of such suffixes, and to see whether the two dictionaries display variation
- ▶ We retrieved 2,965 nouns for A&A, 2,952 for S&A
- ▶ 1,072 pairs were present in both dictionaries and had information about inflectional forms

Nominal morphology

- ▶ Both dictionaries report both the **genitive** and the **plural** suffix
- ▶ We used this information to study the productivity of such suffixes, and to see whether the two dictionaries display variation
- ▶ We retrieved 2,965 nouns for A&A, 2,952 for S&A
- ▶ 1,072 pairs were present in both dictionaries and had information about inflectional forms
- ▶ The genitive suffix is almost always reported in both dictionaries: for all 1,072 words in S&A, for 1,066 words in A&A
- ▶ The plural is not reported for all words: for 571 words in S&A, for 879 words in A&A

Nominal morphology

Figure 1: Variation: plural suffixes



(χ^2 -squared = 47.118, df = 2, p-value = 5.869e-11)

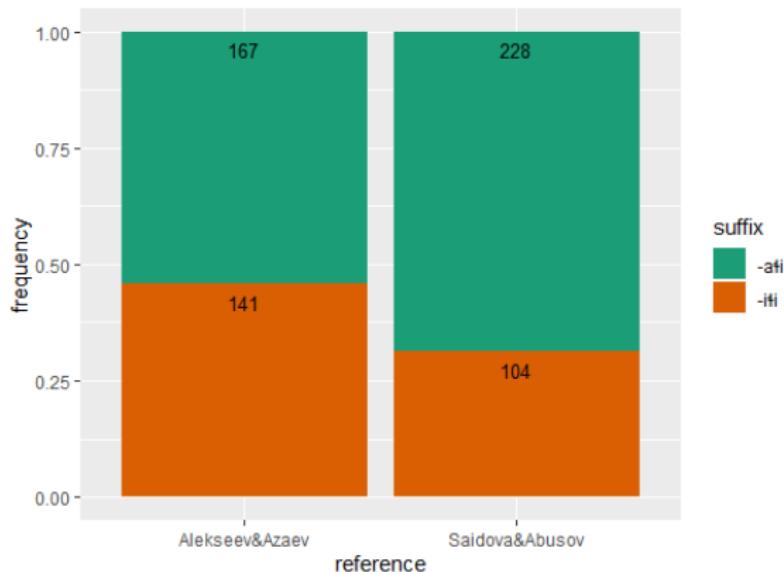
Nominal morphology

Table 1: Stem endings and genitive suffix

Suffix	-ti		-ati		-iti		-uti	
Stem	S&A	A&A	S&A	A&A	S&A	A&A	S&A	A&A
consonant	232	266	228	167	104	141	-	1
-a	182	167	-	-	3	13	15	14
-i	143	151	10	4	-	-	3	2
-u	81	78	1	-	-	3	-	-
-e	6	6	-	-	-	-	-	-
-o	7	7	-	-	-	-	-	-

Nominal morphology

Figure 2: Variation: genitive in *-ati* vs. *-iti* with stems ending in a consonant



($\chi^2 = 13.523$, $df = 1$, $p\text{-value} = 0.0002357$)

Nominal morphology

Variation: examples

- ▶ It seems that variation frequently involves loanwords:
 - ▶ *dakument* ‘document’ < *dakument-atī* (genitive, S&A) vs. *dakument-iṭi* (genitive, A&A)
 - ▶ *birgadir* ‘foreman’ < *birgadir-zabaṭi* (plural, S&A) vs. *birgadir-de* (plural, A&A)
 - ▶ *kassir* ‘cashier’ < *kassir-atī* (genitive, S&A), *kassir-zabaṭi* (plural, S&A) vs. *kassir-iṭi* (genitive, A&A), *kassir-de* (plural, A&A)

Nominal morphology

Variation: examples

- ▶ It seems that variation frequently involves loanwords:
 - ▶ *dakument* ‘document’ < *dakument-atı* (genitive, S&A) vs. *dakument-ići* (genitive, A&A)
 - ▶ *birgadir* ‘foreman’ < *birgadir-zabați* (plural, S&A) vs. *birgadir-de* (plural, A&A)
 - ▶ *kassir* ‘cashier’ < *kassir-atı* (genitive, S&A), *kassir-zabați* (plural, S&A) vs. *kassir-ići* (genitive, A&A), *kassir-de* (plural, A&A)
- ▶ But not only:
 - ▶ *gač'a* ‘hoe’ < *gač'a-tı* (genitive, S&A), *gač'-ibatı* (plural, S&A) vs. *gač'-ići* (genitive, A&A), *gač'-e* (plural, A&A)

Verbal morphology

- ▶ We retrieved 1,570 verbal entries for A&A and 1,475 for S&A
- ▶ Only 554 pairs were present in both dictionaries and had information about inflectional forms

Verbal morphology

- ▶ We retrieved 1,570 verbal entries for A&A and 1,475 for S&A
- ▶ Only 554 pairs were present in both dictionaries and had information about inflectional forms
- ▶ Botlikh has only one verb stem; suffixes are attached directly
- ▶ Infinitive and Aorist are indicative for conjugation:

Verbal morphology

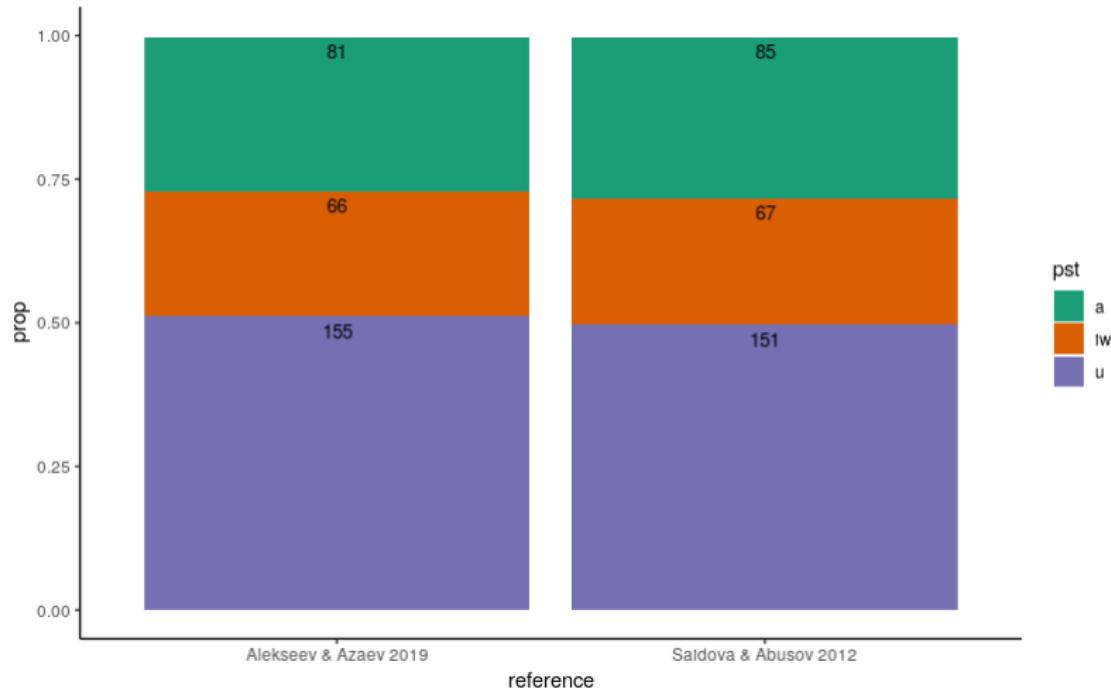
- ▶ We retrieved 1,570 verbal entries for A&A and 1,475 for S&A
- ▶ Only 554 pairs were present in both dictionaries and had information about inflectional forms
- ▶ Botlikh has only one verb stem; suffixes are attached directly
- ▶ Infinitive and Aorist are indicative for conjugation:

Table 2: Basic verb inflection in Botlikh

	infinitive	present	past
Basic	-i	-e	-a / -u / -iw
Derived	-hi	b-ah-e	b-ah-u 'become'
Causative	-a-j	b-uk'-e mal-e / malih-e	b-uk'-a 'be' -o / malih-u

Verbal morphology

Figure 3: Past tense suffixes in the sample of common verbs



Verbal morphology

- There is more variation than the plot shows

Table 3: Past suffixes in both dictionaries

S&A / A&A	a	iw	u	other
a	72	6	3	4
iw	3	55	6	3
u	3	4	142	2
other	3	1	4	0

Verbal morphology

- ▶ There is more variation than the plot shows

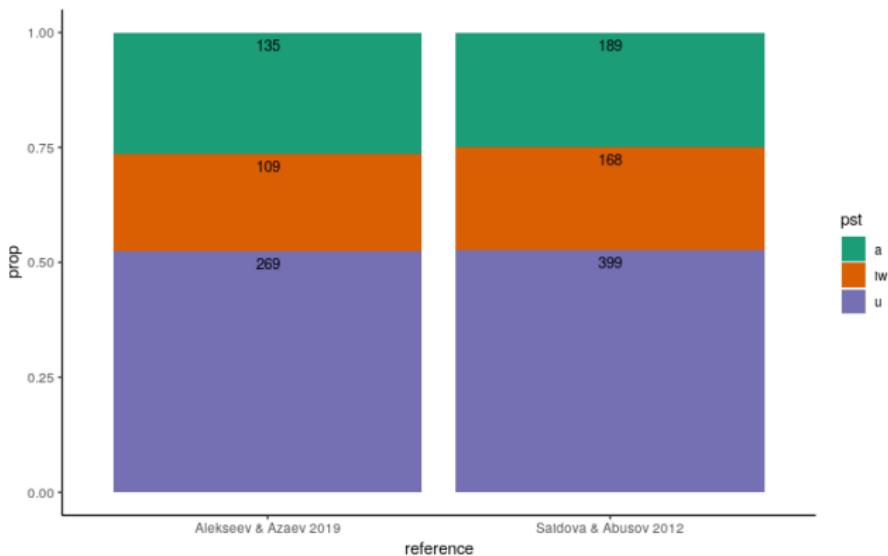
Table 3: Past suffixes in both dictionaries

S&A / A&A	a	iw	u	other
a	72	6	3	4
iw	3	55	6	3
u	3	4	142	2
other	3	1	4	0

- ▶ But still not a lot
- ▶ And no systematic correspondence

Verbal morphology

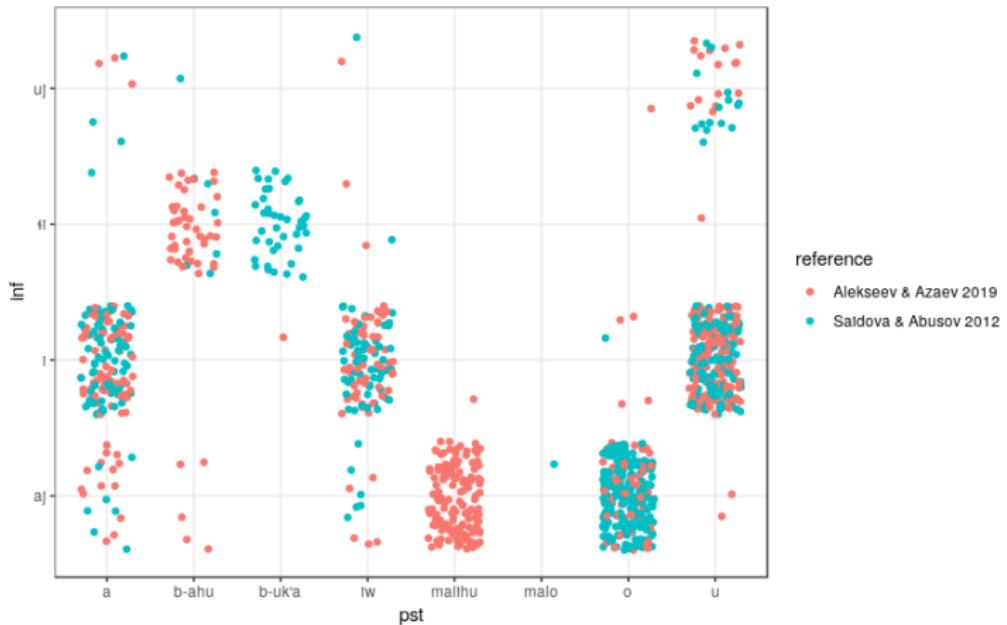
Figure 4: Past tenses in all dictionary data



(χ^2 -squared = 0.347, df = 2, p-value = 0.8407)

Verbal morphology

Figure 5: Infinitive vs. past tense



Verbal morphology

Table 4: Causative paradigms in Botlikh

	infinitive	present	past
A&A	-a-j	malih-e	malih-u
		-e	-o
S&A	-a-j	mal-e	-o
		-e	-o
		mal-e	mal-o (ix)

Verbal morphology

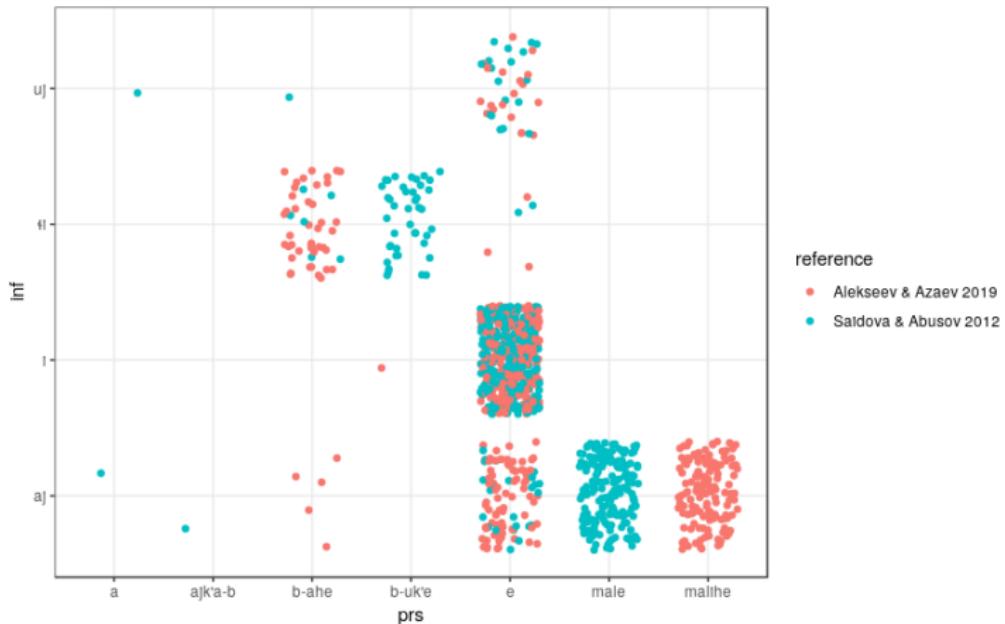
Table 4: Causative paradigms in Botlikh

	infinitive	present	past
A&A	-a-j	malih-e	malih-u
		-e	-o
S&A	-a-j	mal-e	-o
		-e	-o
		mal-e	mal-o (ix)

- ▶ The bulk of material in A&A was collected several decades earlier with respect to S&A
- ▶ It seems that in contemporary Botlikh, the present form *malih-e* has been reduced to *mal-e* and the past form *malih-u* has been replaced by *-o*, formerly reconstructed as <^{*}-a-u [-CAUS-PST]

Verbal morphology

Figure 6: Infinitive vs. present tense



Verbal morphology

Correspondence between infinitive and present tense form

- ▶ Illustrates the division in causative forms (cf. Table 4): full form *malihe* in A&A, shorter form *male* in S&A

Verbal morphology

Correspondence between infinitive and present tense form

- ▶ Illustrates the division in causative forms (cf. Table 4): full form *malihe* in A&A, shorter form *male* in S&A
- ▶ Preference for a specific auxiliary to inflect derived verbs: *b-ahe* 'become' in A&A, *b-uk'a* 'be' in S&A
- ▶ This is also the case in the past

Verbal morphology

ENTRY ‘BLEAT’ **baʃada/tí**, *prs.* **bahé**, *pst.* **báhu** ([Alekseev & Azaev 2019](#))

- (1) ha-b λ'et'ír **baʃáda** b-uk'-é
DEM-N lamb bleat n-be-HAB
'That lamb often bleats.'

Verbal morphology

ENTRY ‘BLEAT’ **baʃada/ʃí**, *prs.* **bahé**, *pst.* **báhu** ([Alekseev & Azaev 2019](#))

- (1) ha-b ɬ'et'ír **baʃáda b-uk'-é**
DEM-N lamb bleat n-be-HAB
'That lamb often bleats.'

ENTRY ‘ROAR’ **bu:da/ʃí**, *prs.* **buké**, *pst.* **búk'a** ([Saidova & Abusov 2012](#))

- (2) únsa **bu:da b-áh-u**
ox roar N-become-AOR
'The ox started to roar.'

Summary

- ▶ Little phonological variation besides different stress patterns in 254 words from a 1,688 word sample
- ▶ Nouns show some variation both in the genitive and in the plural
 - ▶ A general preference for *-bałi* over *-de* in the plural and for *-ałi* over *-ili* in the genitive in S&A
 - ▶ The opposite trends in A&A
 - ▶ Variation frequently involves loanwords
 - ▶ In some cases the dictionaries report more than one variant for both the genitive and the plural, so we might conclude that idiosyncratic variation is at play

Summary

- ▶ Little phonological variation besides different stress patterns in 254 words from a 1,688 word sample
- ▶ Nouns show some variation both in the genitive and in the plural
 - ▶ A general preference for *-bałi* over *-de* in the plural and for *-ałi* over *-ili* in the genitive in S&A
 - ▶ The opposite trends in A&A
 - ▶ Variation frequently involves loanwords
 - ▶ In some cases the dictionaries report more than one variant for both the genitive and the plural, so we might conclude that idiosyncratic variation is at play
 - ▶ But more research is needed!

Summary

- ▶ Verbs show diachronic and idiosyncratic variation
 - ▶ The auxiliary used to derive the present tense of causatives has undergone morphologization: *malihe* — *male*
 - ▶ Some causatives (can) have a regular suffix *-e*
 - ▶ The past tense of the causative is formed with a suffix *-o*, which replaced an auxiliary *malihu*
 - ▶ It remains unclear how *-o* emerged

Summary

- ▶ Verbs show diachronic and idiosyncratic variation
 - ▶ Each dictionary prefers a different auxiliary as citation form for the inflection of *-ti* forms (derived verbs): 'become' in A&A, 'be' in S&A
 - ▶ This is a matter of personal taste: examples in the entries of these verbs show that both are possible
 - ▶ It seems the auxiliaries form a kind of aspectual opposition: inchoative ('become') vs. durative ('be')

Summary

- ▶ Verbs show diachronic and idiosyncratic variation
 - ▶ Each dictionary prefers a different auxiliary as citation form for the inflection of *-ti* forms (derived verbs): 'become' in A&A, 'be' in S&A
 - ▶ This is a matter of personal taste: examples in the entries of these verbs show that both are possible
 - ▶ It seems the auxiliaries form a kind of aspectual opposition: inchoative ('become') vs. durative ('be')
 - ▶ Limited variation in other parts of the verbal domain

Спасибо



Abbreviations

AOR aorist 39, 40

DEM demonstrative 39, 40

HAB habitual 39, 40

N neuter 39, 40

References

-  Alekseev, M.E. & X.G. Azaev. 2019. *Botlixsko-russkij slovar'* [Botlikh-Russian dictionary]. Moscow: Academia.
-  Gudava, E., Togo. 1962. *Botlixuri ena* [The Botlikh language]. Tbilisi: Mecniereba.
-  Saidova, Patimat A. & Magomed G. Abusov. 2012. *Botlixsko-russkij slovar'* [Botlikh-Russian dictionary]. Makhachkala: IJaLI.