

# Phonological and morphological variation in Botlikh: Comparing two dictionaries

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

24.03.2020

# Botlikh

- ▶ Botlikh < Andic < Avar-Andic-Tsezic < East Caucasian
- ▶ Spoken by ~5,000-8,000 speakers
- ▶ Three villages in the Botlikh district of the Republic of Daghestan: Botlikh, Miarso, and Ashino
- ▶ Unwritten and mostly spoken at home; the Cyrillic script of Avar functions as an ad hoc writing system on social media
- ▶ Evaluated as “threatened” by Ethnologue ([Simons & Fenning 2018](#)), but many children still speak the language and attitudes are positive
- ▶ Heavy influence from Avar and Russian

# Botlikh

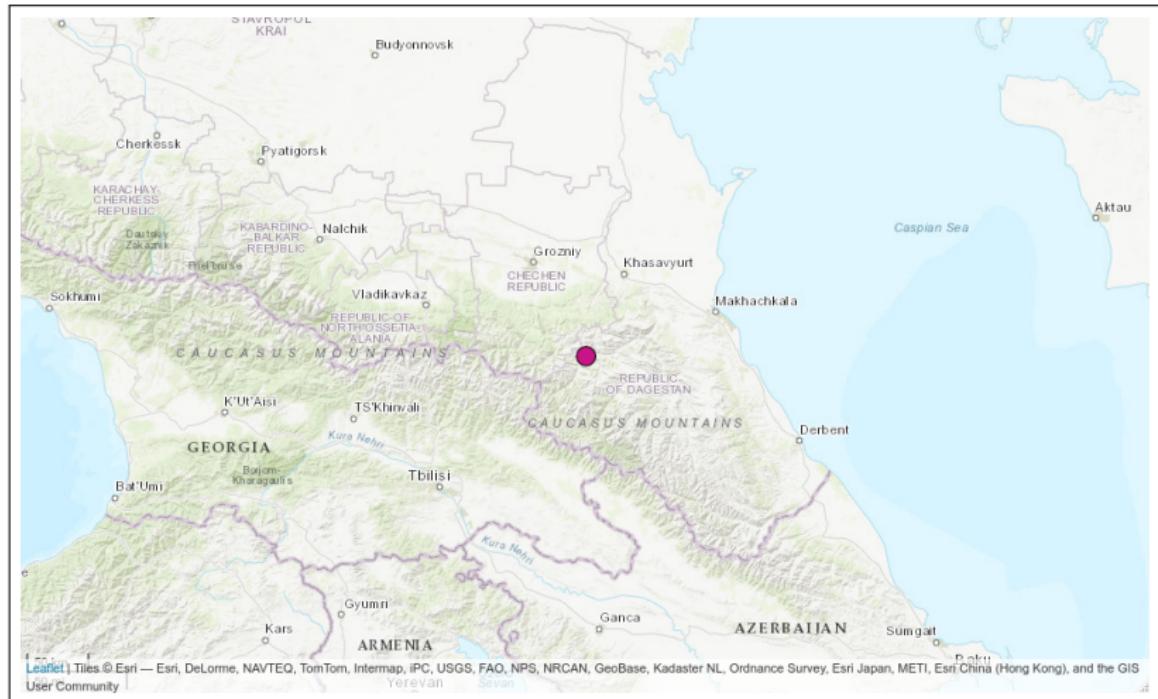


Figure 1: Botlikh on the map

# Botlikh



Figure 2: The old village

## Botlikh literature

- ▶ One full reference grammar in Georgian ([Gudava 1962](#))
- ▶ Several short sketches mostly based on information contained in the grammar by Togo E. Gudava ([Gudava 1967](#), [Azaev 2000](#), [Saidova 2001](#), [Magomedbekova 2001](#), [Xalidova 2017](#), [Alekseyev & Verhees Forth.](#))
- ▶ Several works on the lexicon and word formation ([Azaev 1975](#), [Sulejmanova 2013](#), [Alekseyev 2016](#))
- ▶ In general poorly described compared to other Andic languages like Godoberi or Bagvalal, BUT two Botlikh-Russian dictionaries are available to date ([Saidova & Abusov 2012](#), [Alekseev & Azaev 2019](#))

## Two dictionaries



Figure 3: Two Botlikh-Russian dictionaries

## Two dictionaries

- ▶ Saidova & Abusov (2012) compiled in the 2000s by a native speaker of Botlikh (Magomed G. Abusov) and an experienced linguist (Patimat A. Saidova)
- ▶ Alekseev & Azaev (2019) compiled in the 1960s/1970s by a native speaker of Botlikh and philologist (Xalil G. Azaev), later (in the 2000s) systematized by an experienced linguist (Mixail E. Alekseev), and published posthumously after the editing by Timur A. Maisak

## Two dictionaries

- ▶ Comparable both quantitatively and qualitatively
- ▶ ~8,000 headwords for Saidova & Abusov (2012) vs. ~9,000 words and expressions for Alekseev & Azaev (2019)
- ▶ Although the data in Alekseev & Azaev (2019) were collected several decades earlier, Magomed G. Abusov also consulted elderly speakers with the aim of collecting archaic vocabulary
- ▶ Saidova & Abusov (2012) also contains some notes on Miarso; reference to Miarso variants is not explicit in Alekseev & Azaev (2019), but it seems that such variants are occasionally reported in this dictionary too
- ▶ No metadata on the speakers consulted
- ▶ At first glance, the two resources seemed to display variation

## Our research

- ▶ Comparison of the two resources
- ▶ A quantitative investigation of an understudied language
- ▶ Provide numerical approximations for the impressionistic observations available in the existing literature
- ▶ Analysis of both phonological and morphological features
- ▶ Detect patterns of systematic variation within these two areas

# Outline

- ▶ Data
  - ▶ merging
  - ▶ extracting grammatical information
  - ▶ pairing and annotation
- ▶ Analysis
  - ▶ phonology (George Moroz)
  - ▶ nominal morphology (Chiara Naccarato)
  - ▶ verbal morphology (Samira Verhees)
- ▶ Results and discussion
- ▶ Methodological remarks

# Merging (George Moroz)

A	B
АБАБЕ фон, пурдакъ вхудоплотной буктроном; ингр. <i>тапану</i>	авадабирабыс масаты буквъ говорили о стю яңа нареджет
АБАЛАДЫР перс. <i>шанав</i>	АБАЛАДЫР III (-тэй) веңность, асемлек, аспарханостык; <i>райнде</i> аваданымсыз түсген жаңырлыктың человек (бүгүн именоцый в сердце веңность человек) аванызын түгөлб түмүрү бәрәдештөн жынын; <i>раскрычу</i> аваданымсыз түсген жынын; <i>шам</i> илемнегиздиктөн; <i>аваданы</i> в'ятка пекедистек устриң веңность
АБАЛДЫРЫЛГЫЛТ то же, что авадан	АВАЛ III (-акты, -ад) 1) кирады; тый ишиң авал иша этина кирады; ишиң авал маникен ишү ви шашем кирады построил; мечеть 2) скрут; тый авале куйнештесе жана кирады; тый ви этом округе из бояндан корынды чекшесе
АБАЛ-АНГР III (-алы) ишады и көндөн; авал-ангар түләб миру работа без начала и конца (о запущенной работе, о безконечном деле, берегупланни девелопменте)	АВАЛАНЫШ III (-саны) ишады; авансай аренда гильбады или получены деньги в шаше; аванс ишки дакат ишады
АБАРАТ I (-акты, -абалы) ред. пророк; авариялык айрын бику хана просит; авариялык пророк; пророк; авариялык миссионер хана просит или по пути, указанному пророком;	АВАРА I (-тэй) проприетар; помеха; тый музүл авара ишу этому делу помехами; тый авара ишим мин кра түнгизчел? ты принес, что помеха
АБАРДАЧКА III (-аль, -аль) авторчук; авторчуканы таңырып иштеп ишилти за авторчук; авторчуканы шашы бишил ви авторчук кончины чернила	АВАРА III (-тэй) авария; авария ишилти сокиртти авария; авария руки
АБАТЕКА III (-тын, -былын) авторке; авторчуканы таңырып иштеп ишилти за авторчук; авторчуканы шашы бишил ви авторчук кончины чернила	авайдай ишоды стя обязательно, естественно, конечно; <b>авайдай дин</b> гынбы дүй санду конечно, я тобой с тобой
АБАДАЦИА никон образом	авадаңдаца
АБАДАЙ-Б вечный	авада-абад, ре.-иль уст. вечность
АБАДАУ-АБАД, ре.-иль оборонять; аварона биккій оборонять; авароналдык оборонительный	аварона, ин.-4, ре.-иль обороне; аварона биккій оборонять; авароналдык оборонительный
АБДИКАЦИЯ 1) сдача, сдасть	абдидай
АБДИКАЦИЯ 2) сдача, сдасть	авадаңдаилянни кий дайтты объявление в газете; авадаңдаилян бүххай, авадаңдаилян күбәнди, авадаңдаилян тыйтай пояснять объявление
АБОРТ, ин.-иль, ре.-иль аборт; аборт ишле салып аборт	аборт, ин.-иль, ре.-иль аборт; аборт ишле салып аборт; Гый тәншікін азаматы аборт иту Эта женщина сделала аборт
АВАДАНЫБ (просторный); аваданыб айыр майдан; аваданыб райкес	аваданыб просторный; аваданыб айыр майдан; аваданыб райкес
АВАДАНЫЛТ 1. простор, 2. щедрость	аваданылты
АВАЛ-АХИР конец света; авал-ахир зәнді дандылар кібін на том свете встретимся; авал-ахир түтүл заман вечность	аваира, ин.-былы, ре.-иль помеха, преграда
АВАРАГ ин.-иль, ре.-иль пророк;	авараг ин.-иль, ре.-иль пророк; Муса авараг пророк Муса (Монсе); авараг решинкай бүрәл мінхым а) раздуть; б) Млечный путь
АВАРИ, ин.-иль, ре.-иль авария; аварий ишле совершил авария	авари, ин.-иль, ре.-иль авария; аварий ишле совершил авария
АВАШЧУ, ин.-иль, ре.-иль авишчук; Ишкүр кыгыт авашчукки бите! Я посып письмо авишчукоттой	аваашчук, ин.-иль, ре.-иль авишчук; бүлил аваашчук рұту сунт познаны на пастбище; бите аваашчук равнина
АВАЛЫК, ин.-иль, ре.-иль белое место; бүлил аваалық жағынан	авадебонгри ин.-иль, ре.-иль автобиография; Деб авадебонгрииалык беки

Figure 4: Merging the dictionaries

# Merging (George Moroz)

From two .doc files to one .xls:

- ▶ .doc preprocessing: convert files to .html, change non-standard symbols, solve several tag problems ( $I$ ,  $H$ , etc.)
- ▶ extract first bold entrance, parse parenthesis for grammar information

# Extracting grammatical information

- ▶ Total number of lexemes extracted: 8,464 from Saidova & Abusov (2012) and 6,821 from Alekseev & Azaev (2019)
- ▶ Nouns: 2,871 from Saidova & Abusov (2012) and 3,097 from Alekseev & Azaev (2019)
  - ▶ grammatical information: genitive and plural
- ▶ Verbs: 1,504 from Saidova & Abusov (2012) and 1,640 from Alekseev & Azaev (2019)
  - ▶ grammatical information: habitual and aorist

# Extracting grammatical information

lemma	pos	noun_gen	noun_pl	verb_prs	verb_pst	reference
арбагзыв	noun	-лъи				Saidova, Abusov
арджан	noun	-алъи	-			Alekseev 2006
арёнда	noun	-лъи				Saidova, Abusov
аржаж/áй	verb			-аймалé	-ó	Saidova, Abusov
аржái	verb			-é	-o	Alekseev 2006
арж/áй	verb			-áймале	-ó	Saidova, Abusov
áржар	noun	-лъи	-дé			Alekseev 2006
аржíй	verb			-é	-a	Alekseev 2006
арж/и	verb			-e	-o	Saidova, Abusov
.						- - -

Figure 5: The database

## Pairing

We manually checked for lexemes represented in both dictionaries to carry out phonological and morphological analysis

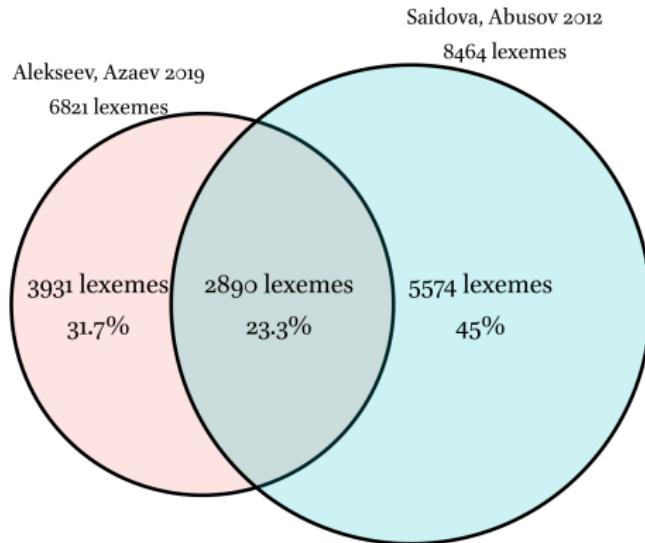


Figure 6: The database

# Pairing

We manually checked for lexemes represented in both dictionaries to carry out phonological and morphological analysis

- ▶ Those 2,890 include all POS
- ▶ There are multiple entries for the same root  
*Alekseev & Azaev (2019)*: *besqχe* ‘behind’, *besqχéku* ‘from behind’,  
*besqé-ssu-b* ‘back/hind (adjective)’, *héči besqérussubla* ‘in the end’
- ▶ There are also some sayings and proverbs (about 670) that constitute separate entries in *Alekseev & Azaev (2019)*  
*héči besqχérussubla* ‘in the end’

# Annotation

- ▶ Manual correction of automatically extracted information about grammatical features
- ▶ Addition of further annotation (for features that appeared to be potentially relevant for our research):
  - ▶ masdars
  - ▶ borrowings
- ▶ (In progress) Filter out superfluous entries for multi-word expressions and inflected forms

# Phonology

- ▶ There are 1,996 lexemes which look phonetically the same, while 909 are different (31%)
- ▶ If we remove the stress sign, there are 2,449 lexemes which look phonetically the same, and 456 are different (16%)
  - ⇒ 15% of lexemes have different stress pattern?...

# Phonology

- ▶ There are 1,996 lexemes which look phonetically the same, while 909 are different (31%)
- ▶ If we remove the stress sign, there are 2,449 lexemes which look phonetically the same, and 456 are different (16%)
  - ⇒ 15% of lexemes have different stress pattern?... Yes, but including 265 (9%) cases where the stress is present in one dictionary and absent in the other

# Phonology

- ▶ What causes the difference between the two dictionaries?
  - ▶ Stress pattern differences in 188 lexemes (about 6%)
  - ▶ Multiple cases where there is a small difference that could be explained either as a typo or in terms of phonological variation: *čuhí* 'to run' [aa] vs. *čūhí* [sa], *kusu* 'cherry plum' [aa] vs. *kussu* [sa]
  - ▶ Multiple cases where Russian borrowings were adopted differently: *awtobus* 'bus' [aa] vs. *abtabus* [sa], *biton* 'milk can' [aa] vs. *bitun* [sa], *apteka* 'pharmacy' [aa] vs. *abteka* [sa]
  - ▶ Morphological preferences: *dinija=w* 'pious' [aa] vs. *dinija=b* [sa]

# Phonology: segments

Alekseev & Azaev (2019)

Saidova & Abusov (2012)

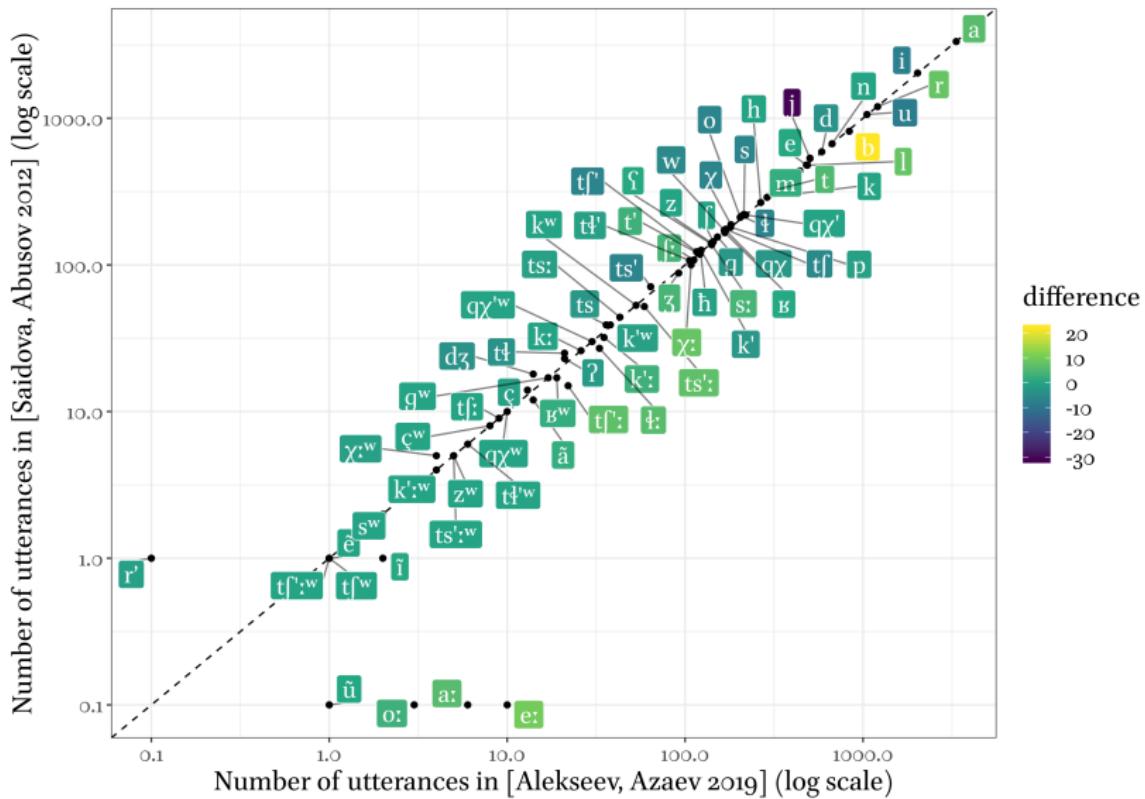
About 25 cases:

<i>ãhajr</i>	<i>ãhar</i>	'message'
<i>bezajr</i>	<i>bezir</i>	'roasting'
<i>mik'kujr</i>	<i>mik':ur</i>	'swallowing'
<i>reqχujr</i>	<i>reqχ&gt;wir</i>	'fight'
<i>reʃkujr</i>	<i>reʃkur</i>	'overnight stay'
<i>rikʷajr</i>	<i>rikʷar</i>	'lighting'
<hr/>		
<i>χwardar</i>	<i>χwardir</i>	'digging'
<i>mi?ar</i>	<i>mi?ar</i>	'nose'
...	...	...

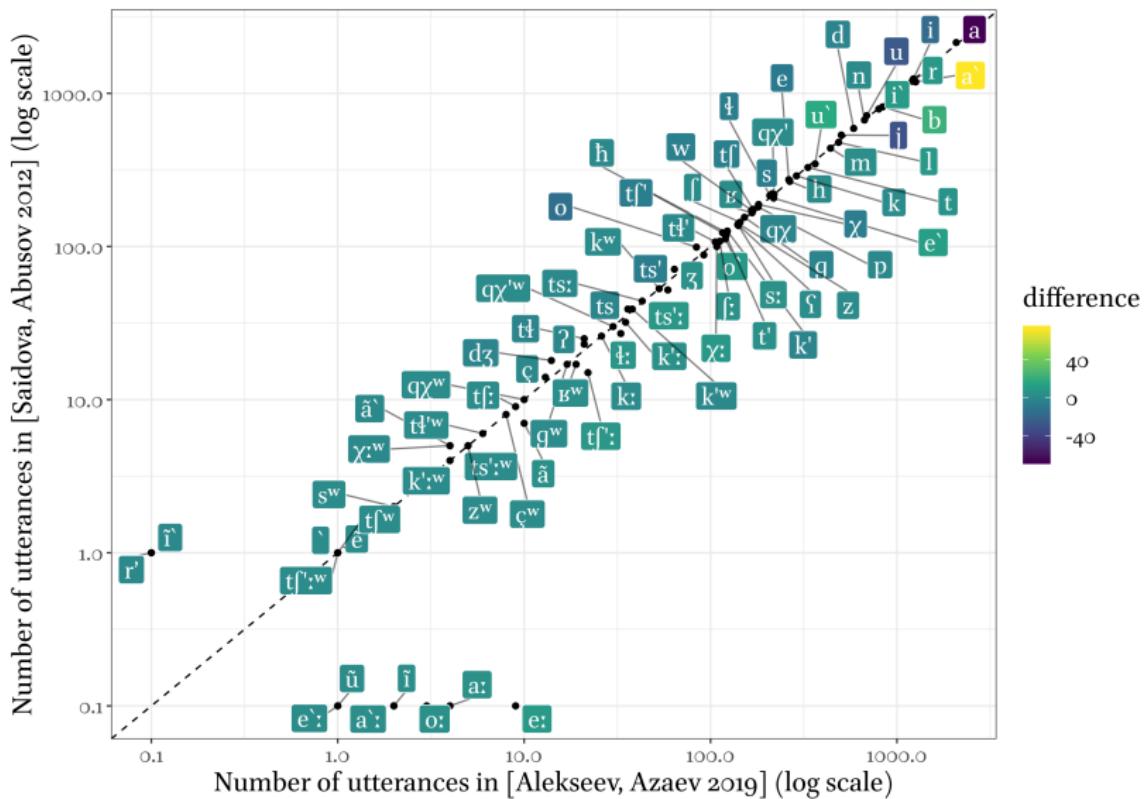
About 6 cases:

<i>fjalaj</i>	<i>fjallaj</i>	'silt'
<i>inuʃala</i>	<i>inuʃalla</i>	'everywhere'
<i>ʃila</i>	<i>ʃilla</i>	'reason'
...	...	...

# Phonology: compare Botlikh segments



# Phonology: compare Botlikh segments

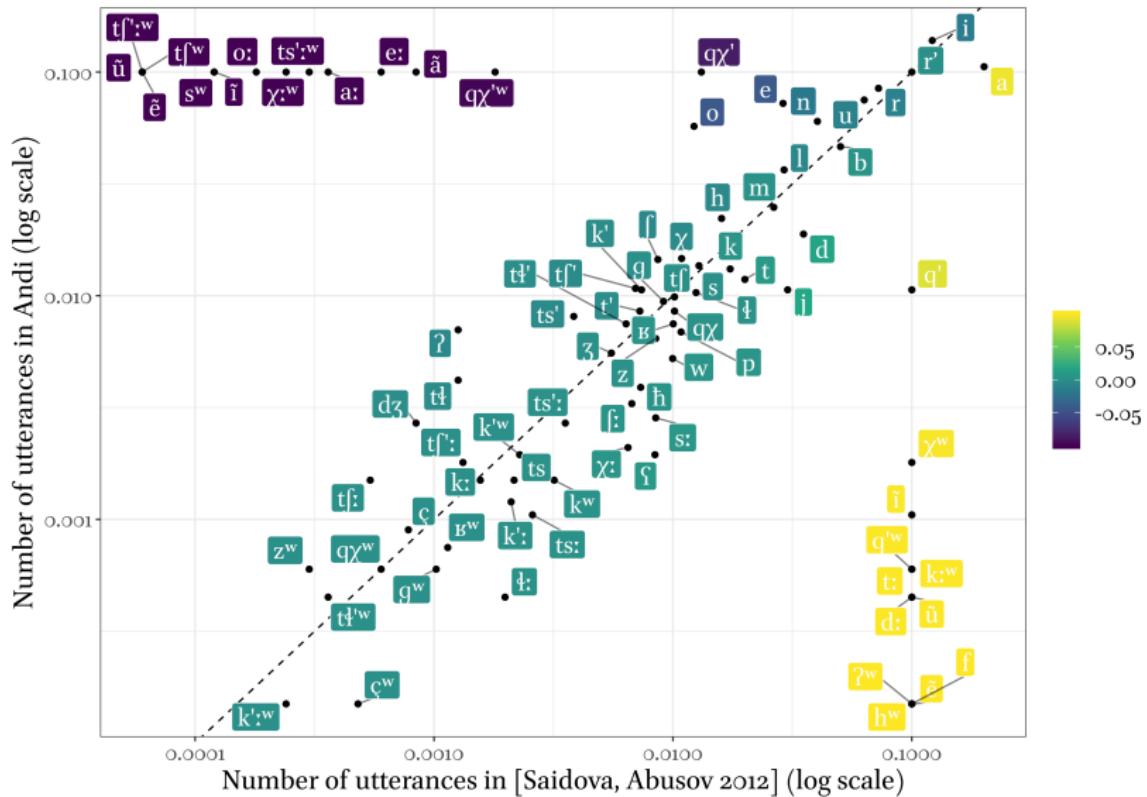


## Zilo Andi data

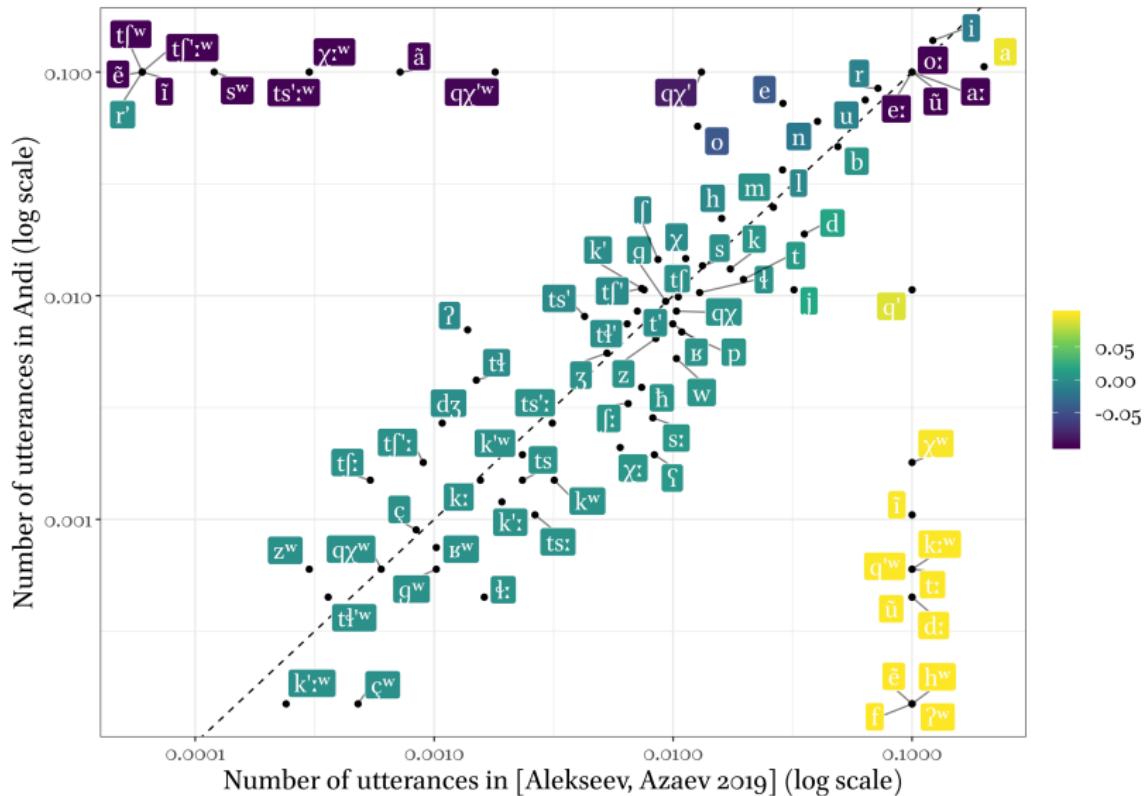
Dictionary data for Zilo were collected during fieldtrips to Zilo in 2016–2019 with N. Rochant, S. Verhees, A. Martynova and A. Zakirova who contributed to the same FieldWorks project

- ▶ Contain morphological affixes
- ▶ Do not contain additional affixes in a lemma form
- ▶ Contain different stems of the same lexeme (e.g. SG.ABS, SG.OBL, PL.ABS, PL.OBL, PST, NPST); those forms were removed during the analysis
- ▶ No information about stress

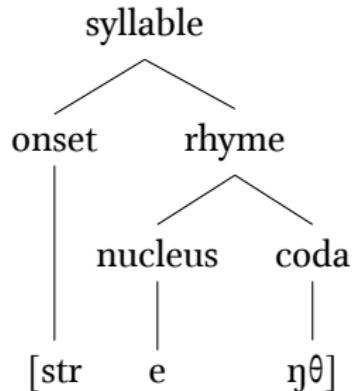
# Phonology: compare Botlikh and Zilo segments



# Phonology: compare Botlikh and Zilo segments

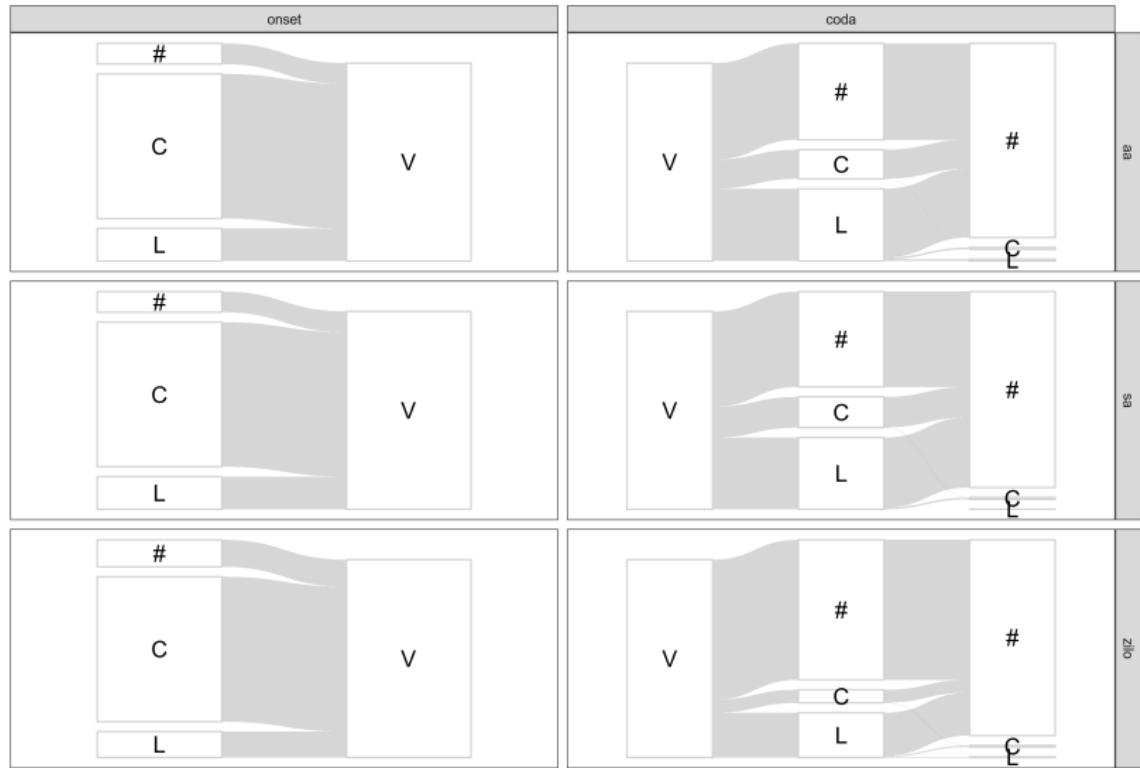


# Phonology: data–driven analysis of syllables



- ▶ Analyze all onsets of initial syllables in the corpus
- ▶ Analyze all codas of final syllables in the corpus
- ▶ Generalize obtained initials and codas into a syllable model
- ▶ Check, whether this model describes all intervocal consonant clusters

# Phonology: data–driven analysis of syllables



C — obstruent, L — sonorant, V — vowel, # — syllable boundary

# Nominal morphology

Two topics investigated:

- ▶ Formation of the plural
  - ▶ to check the productivity of different suffixes
- ▶ Formation of the genitive
  - ▶ to study alternations in the formation of oblique stems

Comparison of the two resources to look for possible variation in such areas of nominal morphology

(based on 1,072 pairs retrieved during the first annotation round)

# Plural formation in Botlikh

- ▶ A suffix is attached to the absolute stem:  
*na* 'thing' < *na-bati* 'things'
- ▶ With stems ending in a consonant, the vowel *-a-* is often inserted before the suffix:  
*majmalak* 'monkey' < *majmalak-a-bati* 'monkeys'
- ▶ With stems ending in a vowel, alternation can occur:  
*ruša* 'tree' < *ruši-bati* 'trees',  
*salu* 'tooth' < *sala-bati* 'teeth',  
*burači* 'pitcher' < *burača-bati* 'pitchers'

# Plural formation in Botlikh

Among the most common suffixes are:

- ▶ *-baṭi* and allomorphs (*-maṭi* for stems ending in a nasal, *-wabaṭi* for stems ending in *-u*, etc.), the variant *-zabaṭi* (mostly with borrowings)  
*apicer* ‘officer’ < *apicer-zabaṭi* ‘officers’
- ▶ *-de* (mostly for stems ending in a sonorant)  
*ambur* ‘roof’ < *ambur-de* ‘roofs’
- ▶ *-e* and its variant *-we* (for stems ending in *-u*)  
*čan* ‘deer’ < *čan-e* ‘deers’

Other, less common, suffixes are: *-(b)daṭi*, *-(b)diṭi*, *-(a)l*, *-rdi*, *-bala(l)*

# Plural formation in Botlikh

Can our dictionary data help us be more precise about the distribution/frequency/productivity of plural suffixes in Botlikh?

...

Do the two dictionaries show any variation in these respects?

# Plural suffixes in the dictionaries

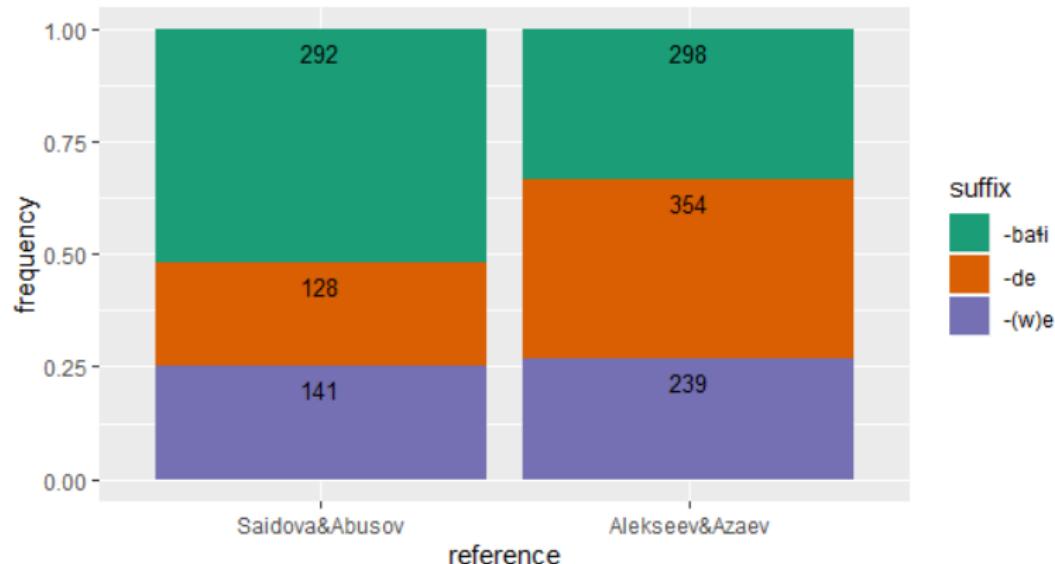
- ▶ Plural suffixes are not reported for all nouns, cf. *singularia tantum* and plural entries (nationalities, *pluralia tantum*)
- ▶ Sometimes more than one variant is reported (so, some nouns were counted twice)

Table 1: Plural suffixes in the dictionaries

	Saidova & Abusov (2012)	Alekseev & Azaev (2019)
- <i>(x)ati</i>	292	298
<i>-de</i>	128	354
- <i>(w)e</i>	141	239
<b>other</b>	24	21
<b>no plural</b>	499	193

# Plural suffixes in the dictionaries

Figure 7: Plural suffixes in the dictionaries



$$(\chi^2 = 59.368, \text{df} = 2, \text{p-value} = 1.283e-13)$$

# Plural suffixes in the dictionaries

- ▶ Preference for *-(x)batı* over *-de* in Saidova & Abusov (2012) vs. the opposite trend in Alekseev & Azaev (2019)
- ▶ The higher frequency of *-de* in Alekseev & Azaev (2019) is partly due to masdars
  - ▶ Saidova & Abusov (2012) almost never report the plural form for such nouns, whereas Alekseev & Azaev (2019) consistently report *-de*
- ▶ Variation often involves (but is not restricted to) borrowings
  - ▶ *birgadir* ‘foreman’ < *birgadir-zabatı* vs. *birgadir-de*
  - ▶ *kassir* ‘cashier’ < *kassir-zabatı* vs. *kassir-de*
- ▶ The frequent mentioning of more than one variant might suggest idiosyncratic variation

# Case declension in Botlikh

## Two declension types

- ▶ 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)

# Case declension in Botlikh

Can our dictionary data help us be more precise about the patterns of case declension (oblique stem formation) and their frequencies?

...

Do the two dictionaries show any variation in these respects?

# Oblique stem formation in the dictionaries

We used the grammatical information included in the dictionaries (genitive suffix) to investigate oblique stem formation in Botlikh

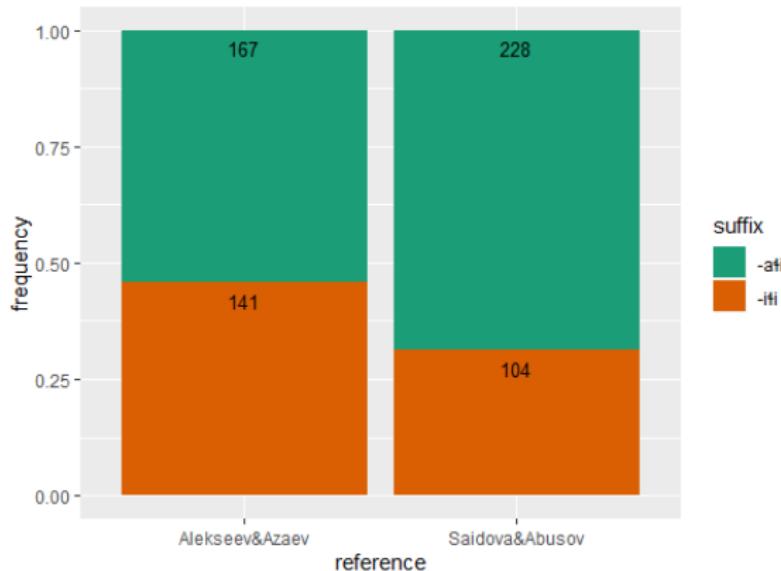
Table 2: Oblique stem formation in the dictionaries

	- <i>ti</i>		- <i>a-ti</i>		- <i>i-ti</i>		- <i>u-ti</i>	
consonant	232	266	228	167	104	141	-	1
- <i>a</i>	182	167	-	-	3	13	15	14
- <i>i</i>	143	151	10	4	-	-	3	2
- <i>u</i>	81	78	1	-	-	3	-	-
- <i>e</i>	6	6	-	-	-	-	-	-
- <i>o</i>	7	7	-	-	-	-	-	-

Saidova & Abusov (2012) vs. Alekseev & Azaev (2019)

# Oblique stem formation in the dictionaries

Figure 8: Oblique stem formation for stems ending in a consonant



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

## Oblique stem formation in the dictionaries

- ▶ Significant variation between the two dictionaries in the formation of oblique stems for nouns ending in a consonant
- ▶ This again involves (but is not restricted to) borrowings (a general preference for *-a-* over *-i-* in [Saidova & Abusov \(2012\)](#))
  - ▶ *dakument* 'document' < *dakument-a-ti* vs. *dakument-i-ti*
  - ▶ *kassir* 'cashier' < *kassir-a-ti* vs. *kassir-i-ti*
  - ▶ *adijal* 'blanket' < *adijal-a-ti* vs. *adijal-i-ti*
- ▶ Different variants for one and the same noun are reported far less frequently as compared to plural suffixes

# Verbal morphology

Table 3: Basic inflectional paradigm of verbs

<b>Infinitive</b>	-i
<b>Habitual</b>	-e
<b>Aorist / General converb</b>	-a / -u / -iw
Progresive converb	-ata / -a
<b>Future</b>	-ik'a-CM
<b>Imperative</b>	-a / -i / -j

# Verbal morphology

Three types of infinitives / inflection

- ▶ Basic verbs (infinitive in *-i*)
- ▶ Derived verbs (infinitive in *-ti*)
- ▶ Causative verbs (infinitive in *-a-j*)

# Verbal morphology

- ▶ Compare inflectional forms of 554 pairs of verbs retrieved during the first annotation round
- ▶ Distribution of aorist suffixes
- ▶ Correlation of infinitive type with habitual and aorist inflection
- ▶ Agreement between sources

# Basic verbs

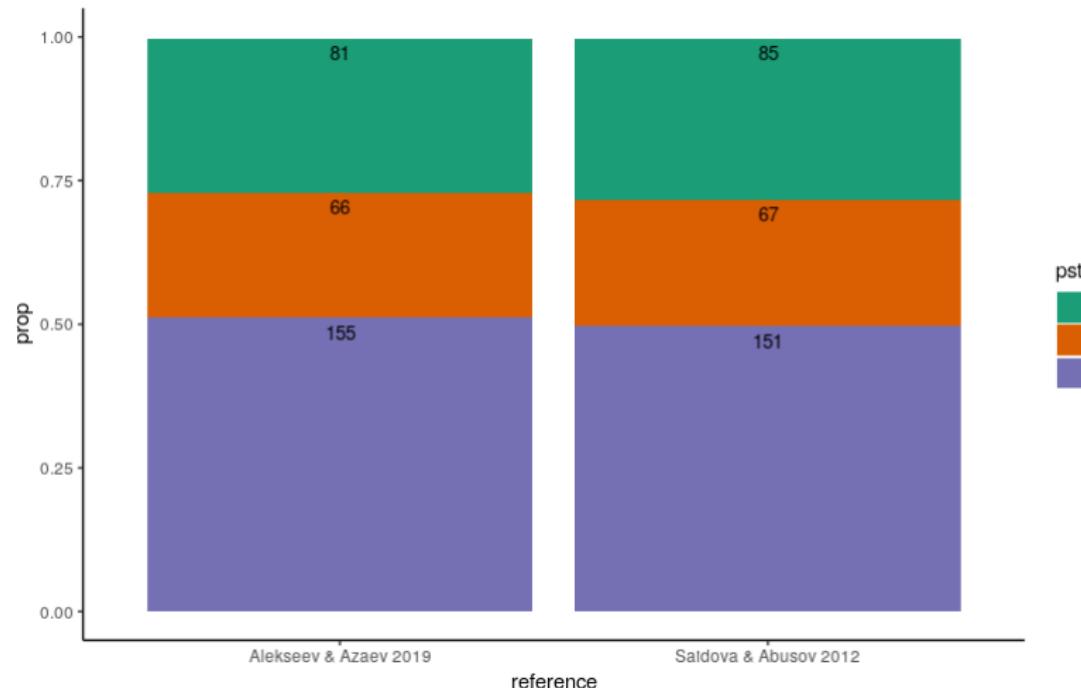
- ▶ Habitual: *-e*
- ▶ Aorist: *-a / -u / -iw*

Table 4: Basic verbs: inflection

	Infinitive	Habitual	Aorist
see	<i>haꝝ-i</i>	<i>haꝝ-e</i>	<i>haꝝ-a</i>
do	<i>ih-i</i>	<i>ih-e</i>	<i>ih-u</i>
be able	<i>baꝝar-i</i>	<i>baꝝar-e</i>	<i>baꝝar-iw</i>

# Basic verbs: Aorist

Figure 9: Aorist suffixes in the dictionaries



## Basic verbs: Aorist

- There is more variation than the plot shows

Table 5: Aorist 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

## Basic verbs: Aorist

- There is more variation than the plot shows

Table 5: Aorist suffixes in both dictionaries

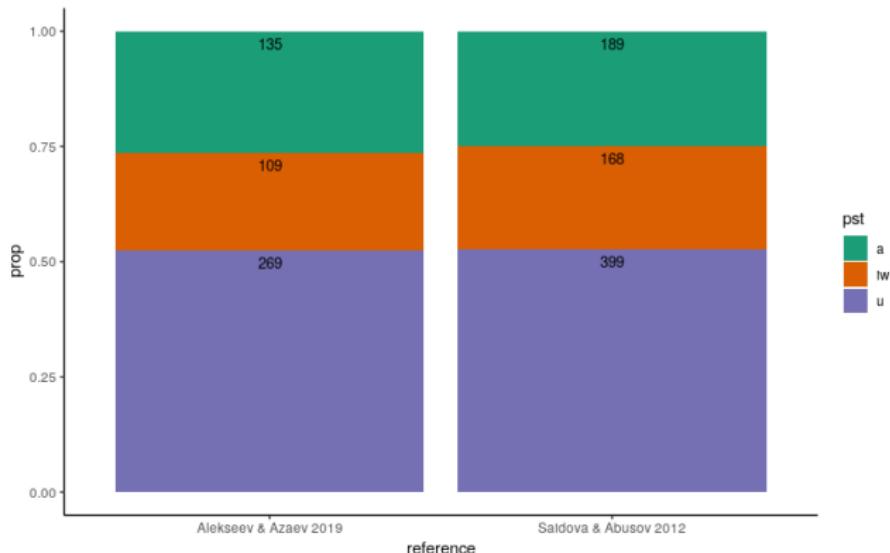
S&A / A&A	a	iw	u	other
a	72	6	3	4
iw	3	55	6	3
u	3	4	142	2
other	3	1	4	0

- But still not a lot
- And no systematic correspondence

# Verbal morphology

Same proportions if we use all data (not just pairs)

Figure 10: Past tenses in all dictionary data



( $\chi^2 = 0.347$ ,  $df = 2$ ,  $p\text{-value} = 0.8407$ )

# Derived verbs

- ▶ Derive verbs from adverbs, converbs, and...?
- ▶ Stem typically ends in *-a*
- ▶ Analytic inflection: stem + 'be' *b-uk'*- or 'become' *b-ah-*

Table 6: Derived verbs: inflection

	Infinitive	Habitual	Aorist
roar	<i>buda-ti</i>	<i>buda b-uk'-e</i>	<i>buda b-uk'-a</i>
bleat	<i>baʃada-ti</i>	<i>baʃada b-ah-e</i>	<i>baʃada b-ah-u</i>

# Causative verbs

- ▶ mixed paradigm: synthetic and semi-periphrastic forms
- ▶ stem + *-a* + infinitive + inflection / stem + inflection

Table 7: Causative verbs: inflection

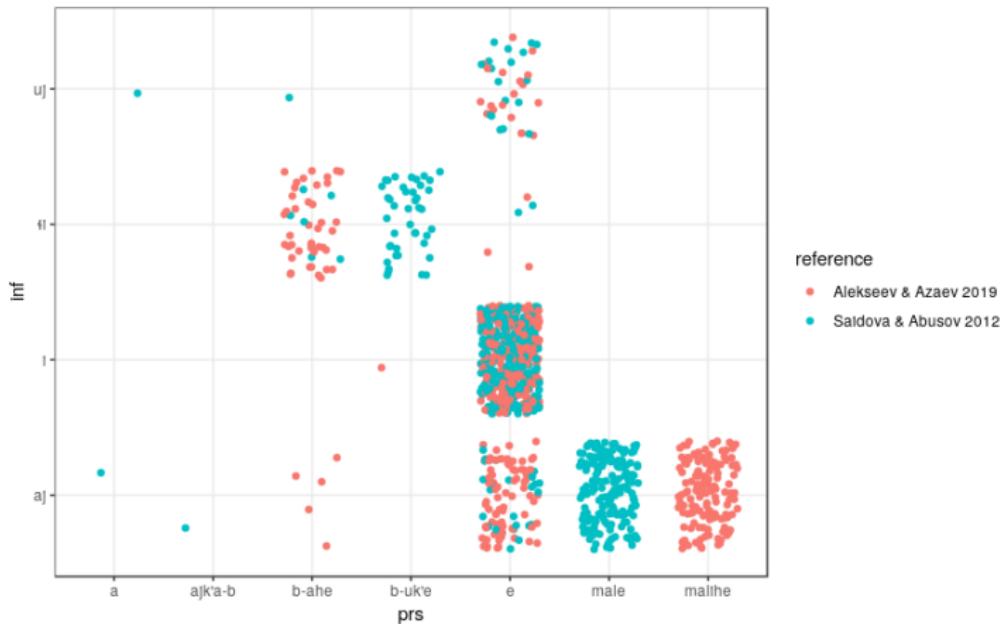
	Infinitive	Habitual	Aorist
resettle	<i>guč-a-j</i>	<i>guč-e</i>	<i>guč-o</i>
roast	<i>žad-a-j</i>	<i>žad-a-j-mal-e</i>	<i>žad-a-j-mal-o</i>
sew up	<i>mik'-a-j</i>	<i>mik'-a-j-mal-e</i>	<i>mik'-o</i>
sew up	<i>mik'-a-j</i>	<i>mik'-a-j-malih-e</i>	<i>mik'-a-j-malih-u</i>

## Causative verbs

- ▶ -*mal-* is likely a reduced form of -*malih-*
- ▶ Both are of unknown origin
- ▶ -*malih-* possibly contains the verb *ih-* ‘make, do’
- ▶ -*o* \*< -*a-u*
- ▶ Relationship between the synthetic and semi-periphrastic inflection is unclear: competing constructions or truncation/reduction?

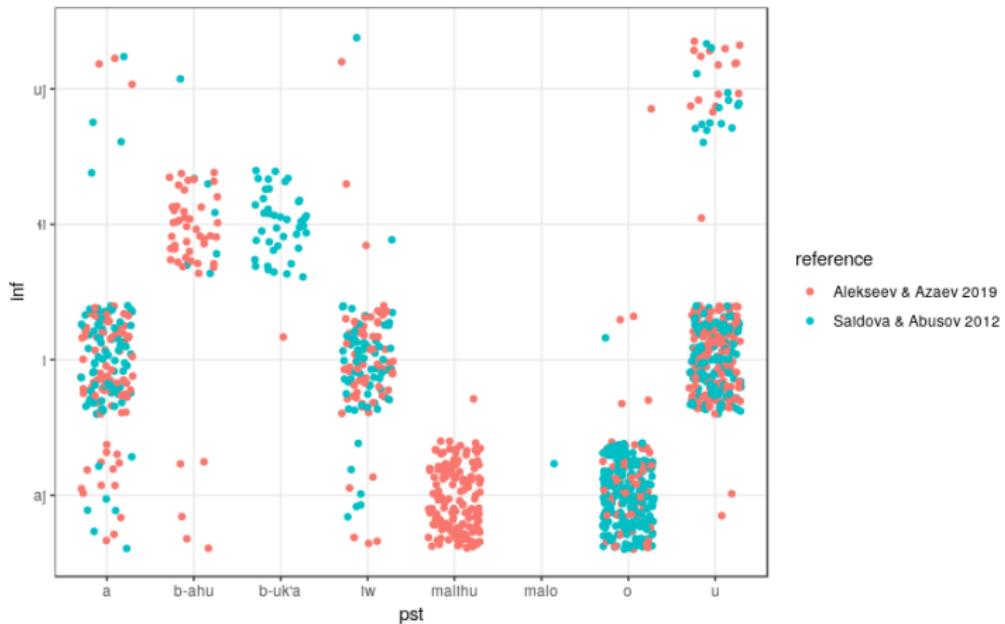
# Infinitive and habitual

Figure 11: Habitual



# Infinitive and aorist

Figure 12: Aorist



# Derived verb inflection

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.'

# Derived verb inflection

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'

# Causative verb inflection

Table 8: Causative paradigms in two Botlikh dictionaries

	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)

# Causative verb inflection

Table 8: Causative paradigms in two Botlikh dictionaries

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

- ▶ The bulk of material in A&A was collected several decades earlier with respect to S&A
- ▶ A&A in many cases note synthetic inflection as a variant

# Variation

- ▶ Distribution of aorist suffixes is very consistent across the data
- ▶ Derived and causative verbs show some variation:
  - ▶ Derived verbs: preference for auxiliary 'be' in Saidova & Abusov (2012) vs. 'become' in Alekseev & Azaev (2019)
  - ▶ This seems to be a matter of personal taste
  - ▶ Causative verbs: full (older?) forms *-malih-e* and *-malih-u* in Alekseev & Azaev (2019) vs. reduced form *-mal-e* and synthetic *-o* in Saidova & Abusov (2012), reflecting diachronic developments

# Variation

- ▶ Phonological variation
  - ▶ stress patterns
  - ▶ syllable structure
- ▶ Morphological variation
  - ▶ nouns: plural suffixes and oblique stem formation
  - ▶ verbs: present (habitualis) and past (aorist) forms

Variation seems to affect (to a greater extent) specific groups of words

- ▶ nouns: borrowings and masdars (both at the phonological and at the morphological level)
- ▶ verbs: causative (and derived – with respect to the choice of a specific auxiliary)

# Variation

- ▶ In general more variation in the nominal than in the verbal paradigm
- ▶ This might be due to the higher percentage of borrowed nouns with respect to verbs (it seems that verbs are rarely borrowed)
- ▶ And to the fact that in general nouns are more numerous in the dictionaries
- ▶ Also, in the nominal paradigm there are more variants

# Variation

How do we explain the variation observed?

- ▶ Idiolectal variation?
- ▶ Diachronic variation?
- ▶ Personal preferences of the author?
- ▶ ...?

# Methodological remarks

- ▶ Variation — is something that is present in most languages, but it is not clear
  - ▶ how to measure the degree of variation in languages
  - ▶ how to make cross-linguistic comparison of the degree of variability

# Methodological remarks

- ▶ Variation — is something that is present in most languages, but it is not clear
  - ▶ how to measure the degree of variation in languages
  - ▶ how to make cross-linguistic comparison of the degree of variability
- ▶ Two similar datasets collected independently in a small language community approximately in the same time period can nevertheless display considerable variation
  - ▶ does not necessarily reflect variation, but shows where to look for it
- ▶ This demonstrates the importance of transparency in data collection i.e. metadata on the speakers consulted
- ▶ And methodological decisions
  - e.g. what is included or not as a separate dictionary entry

## Methodological remarks

- ▶ The availability of comparable material on which quantitative investigations can be conducted is a rare luck for such small languages like Botlikh
- ▶ This precious information can be used to provide numerical approximations for impressionistic observations reported in the available literature on the language
- ▶ This is exactly the phylosophy of data-driven phonology

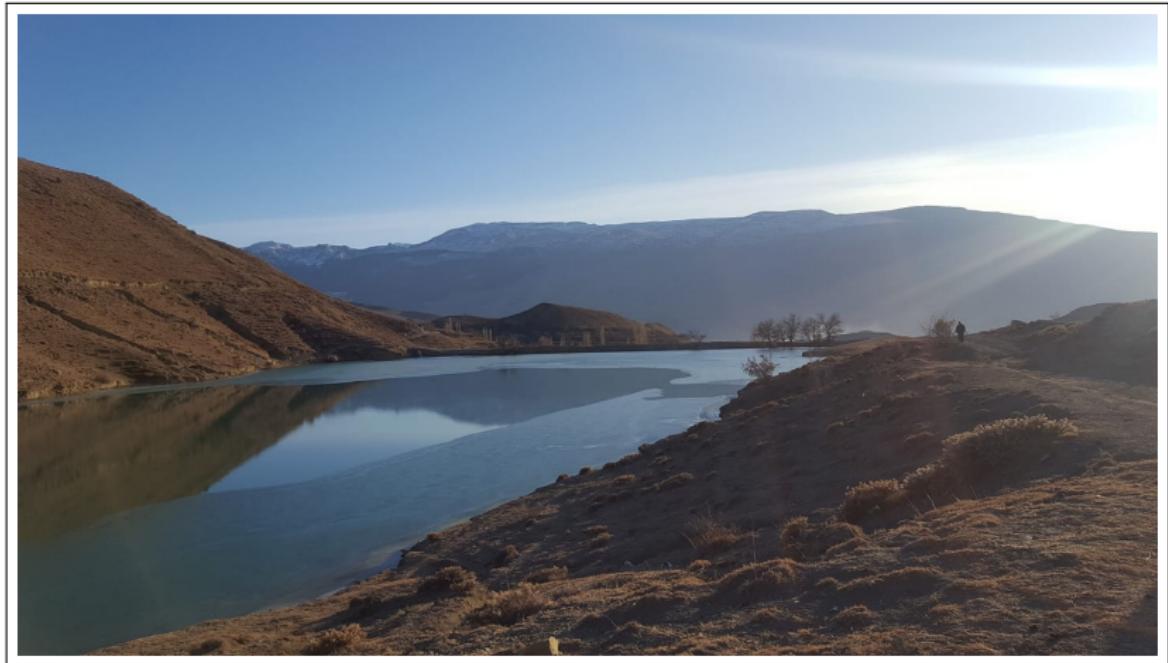
## Further research

- ▶ Cleaning the data by removing superfluous entries (multi-word expressions, inflected forms)
- ▶ Completing the annotation of masdars and borrowings to analyze them separately
- ▶ How much variation is left if we do not consider such “hot spots”?

## Further research

- ▶ Other Avar-Andic dictionaries are being added (Avar, Godoberi, Karata, Tindi, Chamalal, Bagvalal, Akhvakh)
- ▶ The same procedure of data processing and extraction of grammatical information is being carried out
- ▶ This will give us the opportunity to conduct comparative analysis at different linguistic levels (phonology, morphology, lexicon) of the different languages of the Avar-Andic branch of EC languages

The end



# References I

-  Alekseev, Mixail E. & Xalil G. Azaev. 2019. *Botlixsko-russkij slovar'* [*Botlikh-Russian dictionary*]. Moscow: Academia.
-  Alekseyev, Mikhail E. 2016. 'Botlikh'. In Peter O. Müller, Ingeborg Ohnheiser, Susan Olsen & Franz Rainer (eds.), *Word-formation. An International Handbook of the Languages of Europe. Volume 5*. Berlin/New York: Mouton de Gruyter.
-  Alekseyev, Mikhail E. & Samira Verhees. Forth. 'Botlikh'. In Yuri Koryakov & Timur Maisak (eds.), *The Caucasian Languages. An International Handbook. Volume 2*. Berlin/New York: Mouton de Gruyter.
-  Azaev, Xalil G. 1975. *Leksika i slovoobrazovanie botlixskogo jazyka* [*Lexicon and word formation in Botlikh*]. Ph.D. dissertation. Tbilisi State University.
-  Azaev, Xalil G. 2000. 'Botlixskij jazyk [Botlikh]'. In Gadzhi G. Gamzatov (ed.), *Jazyki Dagestana* [*Languages of Daghestan*], 346–359. Moscow/Makhachkala: Vostok (Oriens).

## References II

-  Gudava, Togo E. 1962. *Botlixuri ena [The Botlikh language]*. Tbilisi: Mecniereba.
-  Gudava, Togo E. 1967. 'Botlixskij jazyk [Botlikh]'. In Evgenij A. Bokarev & Ketevan V. Lomatidze (eds.), *Jazyki narodov SSSR tom 4. Iberijsko-kavkazskie jazyki [Languages of the peoples of USSR volume 4. Ibero-caucasian languages]*, 293–306. Moscow: Nauka.
-  Magomedbekova, Zagidat M. 2001. 'Botlixskij jazyk [Botlikh]'. In Mixail E. Alekseev, Georgij A. Klimov, Sergej A. Starostin & Jakov G. Testelec (eds.), *Jazyki mira. Kavkazskie jazyki [Languages of the world. Caucasian languages]*, 228–236. Moscow: Academia.
-  Saidova, Patimat A. 2001. 'Botlixskij jazyk [Botlikh]'. In Viktorija N. Jarceva & Viktor A. Vinogradov (eds.), *Jazyki Rossijskoj Federacii i sosednix gosudarstv. Ènciklopedija v 3-x tomax [The languages of Russia and adjacent states. Encyclopaedia in 3 volumes]*, 222–228. Moscow: Nauka.

## References III

-  Saidova, Patimat A. & Magomed G. Abusov. 2012. *Botlixsko-russkij slovar'* [Botlikh-Russian dictionary]. Makhachkala: IJaLI.
-  Simons, Gary F. & Charles D. Fenning. 2018. *Ethnologue: Languages of the World, 21st edition*. Dallas/Texas: SIL International.
-  Sulejmanova, Džavgarat I. 2013. *Otraslevaja leksika botlixskogo jazyka* [Thematic lexicon of Botlikh]. Ph.D. dissertation. Makhachkala: IJaLI.
-  Xalidova, Rašidat Š. 2017. Lingvisticheskiy portret botlixskogo jazyka [A linguistic portrait of the Botlikh language]. *Vestnik Akademii nauk Čečenskoj Respubliki* 34(1). 77–81.