

Comparing phonological systems and ~~syllable structure~~ of Botlikh and Zilo Andi: a data-driven analysis

George Moroz

Linguistic Convergence Laboratory, NRU HSE, Moscow, Russia

24 October 2019

“Caucasian Languages: Typology and Diachrony” in honor of M. E. Alekseev,
Institute of Linguistics RAS, Moscow

Presentation is available here: <https://tinyurl.com/y5xoks3l>



Phonological description: data-driven analysis

Traditional analysis

Data-driven analysis

Phonological description: data-driven analysis

	Traditional analysis	Data-driven analysis
1.	Done by trained linguist	Evaluated by trained linguist
2.	Can be done from scratch	Previous description needed (or at least prior expectations)
3.	Doesn't care about amount of data	Care more about amount of data
4.	Less reproducible	More reproducible
5.	Can not be automated	Can be automated

Phonological description: data-driven analysis

	Traditional analysis	Data-driven analysis
1.	Done by trained linguist	Evaluated by trained linguist
2.	Can be done from scratch	Previous description needed (or at least prior expectations)
3.	Doesn't care about amount of data	Care more about amount of data
4.	Less reproducible	More reproducible
5.	Can not be automated	Can be automated

Data-driven approach to phonological description and syllable structure analysis:

- was proposed in [[Moroz 2018](#)]
- was applied to syllable structure in [[Moroz 2019](#)] to Adyghe data
- was applied to syllable structure in [[Romanova 2019](#)] to Russian and Macedonian data

Phonological description: data-driven analysis

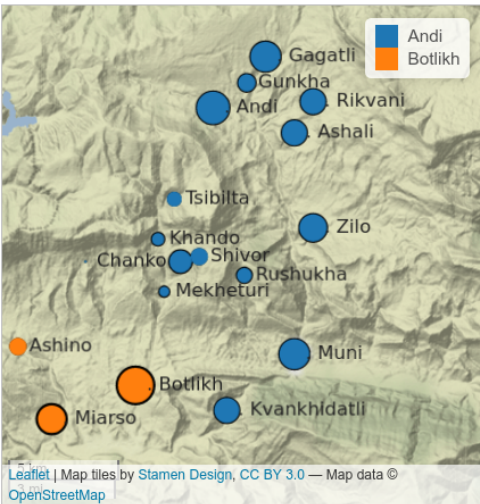
	Traditional analysis	Data-driven analysis
1.	Done by trained linguist	Evaluated by trained linguist
2.	Can be done from scratch	Previous description needed (or at least prior expectations)
3.	Doesn't care about amount of data	Care more about amount of data
4.	Less reproducible	More reproducible
5.	Can not be automated	Can be automated

Data-driven approach to phonological description and syllable structure analysis:

- was proposed in [[Moroz 2018](#)]
- was applied to syllable structure in [[Moroz 2019](#)] to Adyghe data
- was applied to syllable structure in [[Romanova 2019](#)] to Russian and Macedonian data

I will present an application of this method to Botlikh and Zilo Andi data

Andi and Botlikh villages, created with lingtypology package [[Moroz 2017](#)]



- Botlikh < Andic group < EC
- Unwritten (can be written with extended Cyrillic script for Avar)
- ~5,000–8,000 speakers
- Mostly spoken in 3 villages in northwestern Daghestan (Russian Federation): Botlikh, Miarso, Ashino, (Ankho); minor dialectal differences
- One full reference grammar in Georgian [[Gudava 1962](#)]
- Two dictionaries: [[Saidova and Abusov 2012](#)], [[Alekseev and Azaev 2019](#)]

- Andi < Andic group < EC
- Unwritten (can be written with extended Cyrillic script for Avar)
- ~16,500 speakers
- About 14 villages; There are two main dialect groups: Lower Andi (Muni, Kvankhidatli) and Upper Andi (the rest)
- Several reference grammars [[Suleymanov 1957](#)] (Rikvani), [[Salimov 1968](#)] (Gagatli), [[Tsertsvadze 1965](#)] (Andi)
- No dictionary except [[Kibrik and Kodzasov 1988](#)]

Comparing two Botlikh dictionaries

[Saidova and 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

Comparing two Botlikh dictionaries

[Saidova and 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 and 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

Comparing two Botlikh dictionaries

Summary:

- Dictionaries were compiled independently of each other
- with no metadata on the speakers consulted
- data collection was separated with several decades break

A	A
АБАБЕ <i>бот.</i> пузырник вздутойный (<i>косточка</i>); <i>шарс.</i> Глнлрф	аваданабирля <i>масита</i> бука говорили о его жизнерадостности
АБАДАДИ <i>шарс.</i> Глнлрф	АВАДАНЫТЪИ III (-лыи) весёлость, веселее; <i>жизнерадостность</i> ; рактваре <i>аваданыткъылыи гъебля</i> жизнерадостный человек (<i>буик</i> имеющий в сердце веселость человек); аваданылы гучюб гъумру бедрастная жизнь; рактварчы <i>аваданылы гучюб</i> адам искжизнерадостный человек; аваданылы итъя веселиться, устроить веселье
АБАДАЛ-ЦЛАГЪИЛ то же, что абадаи	АВАЛ III (-алты, -де) 1) квартал; гъаб ишиб <i>авад</i> или это наш квартал; ишиб <i>авад</i> манкыл игъу в нашем квартале построили мечеть 2) округ; гъаб <i>аваде</i> кучычча гъумтув кылов гъекля в этом округе не было такого хорошего человека
АБАЗЕНАВ I (-чуб) абазинен	АВАЛ-АХИР III (-алты) начало и конец; авад-ахир гучяб муху работа без начала и конца (<i>о занятииной работе, о безконечности деле, безрезультатной деятельности</i>)
АБАЗЕНАЙ II (-лыи) абазиня	АВАНС III (-алты) аванс; авансей <i>арси гъыбгъа</i> или получили деньги в аванс; аванс <i>иски</i> давать аванс
АБАЗЕНАЛ (-уб) абазини; абаиналуз <i>миншыл</i> абазинский язык	АВАРАГ I (-алты, -абалыи) <i>рес.</i> пророк; аваригалыи <i>амру</i> великие пророка; аваригалыи плитери <i>биксу</i> хачи просить именем пророка; аваригитъяду <i>ради</i> пророка; аваригити гытамалаб <i>мыслили</i> хъади идти по пути, указанному пророком;
АБАЙ (<i>шарс.</i>) см. гъабай	АВАРА II III (-алты) пренепствие; <i>загруженное</i> ; помеха ; гъаб <i>мухуй</i> авара <i>игъу</i> этому <i>делу</i> помешали; гъаб <i>авара</i> <i>игъыма</i> <i>мини</i> <i>ара гынгъа?</i> ты пришел, чтобы чинить мне пренепствия?
АБИ (<i>шарс.</i>) см. гъабин, гъаб5	АВАРИ III (-алты) авария; авари <i>игъы</i> совершить аварию; аварилья <i>рукко</i>
АБИР (<i>шарс.</i>) см. гъабир	
АБИРХА (<i>шарс.</i>) см. гъабСргъа	
АБТАБУС III (-алты, -абалыи) автобус; абтабусавальтыи <i>истанин</i> автобусная станция; абтабусчы гъа <i>ив</i> схать на автобусе; абтабусавальты гыргъа автобусы прибыли	
АБТАМАТ III (-алты, -абалыи) автомат (<i>фр.</i>) <i>автоматического стрелкового оружия</i> ; абтаматильну <i>бахай/абтаматиди</i> <i>кыва</i> гъадай стрелять из автомата	
АБТАРЧКА III (-алты, -балыи) авторучка; абтаручкыи <i>хъарди</i> писать авторучкой; абтаручкыи <i>арси</i> <i>иски</i> платить за авторучку; абтаручкыи шыкы <i>бичча</i> в авторучке кончились чернила	
АБТЁКА III (-алты, -балыи) аптека; абтекалы <i>дарубалты</i> гъыбгъа в аптеку поступили лекарства; абтекалыну <i>дарубалты</i> <i>иски</i> купить лекарства из	

A	A
абай <i>аводи.</i> сл. обязательно, естественно, конечно; абай ден гыкто дуй <i>санду</i> конечно, я пойду с тобой	
абадйыла <i>никонм</i> образом	
абадйыла =6 вечный	
абдул-абал , <i>рп.</i> -илыи <i>уст.</i> вечность	
аварона , <i>ин.</i> -ё, <i>рп.</i> -алты оборона; аварона <i>бикси</i> оборонять; авароналыи оборонительный	
абды , <i>ин.</i> абд. , <i>ис.</i> абду <i>прилечь</i> , лечь ненадолго; полежать ; абду <i>букай</i> прислонить; абду <i>ингы</i> прислониться	
абвалёни , <i>ин.</i> -йабалыи, <i>рп.</i> -алты <i>объявление</i> ; гъагъитыи абвалёни <i>объявление</i> в газете; гъагъитыи абвалёни <i>кни</i> дать объявление в газете; абвалёни <i>буххай</i> , абвалёни <i>кйиванди</i> , абвалёни <i>чйабай</i> повесить объявление	
аборт , <i>ин.</i> -абалыи, <i>рп.</i> -алты <i>аборт</i> ; аборт <i>игъы</i> сделать аборт; Гъай <i>гъешикъа</i> <i>адамиди</i> аборт <i>игъу</i> Эта женщина сделала аборт	
авадйыб <i>просторный</i> ; авадйыб <i>ахи</i> просторный сад; авадйыб <i>рактва</i> шедрая душа	
авадйылы I , простор; 2) шедрость	
авад-ахир <i>концы</i> света; авад-ахир <i>жылди</i> <i>дандалты</i> <i>кйбли</i> на том свете встретим; авад-ахир гучяб <i>заман</i> вечность	
авара , <i>ин.</i> -балыи, <i>рп.</i> -алты <i>помеха</i> , <i>преграда</i>	
аварат <i>ин.</i> -илыи, <i>рп.</i> -де <i>пророк</i> ; Муса аварат <i>пророк</i> Муса (Моисей); аварат <i>решинаккя</i> бурдаб <i>мыкы</i> а) радуга; б) Млечный путь	
авари , <i>ин.</i> -йабалыи, <i>рп.</i> -алты <i>авария</i> ; авари <i>игъы</i> совершить аварию	
авануч , <i>ин.</i> -ё, <i>рп.</i> -илыи <i>аванпюта</i> ; Ишкур кагъат аванпючиди битю Я послал письмо <i>аванпютой</i>	
авалъх , <i>ин.</i> -абалыи, <i>рп.</i> -алты <i>безопасное место</i> ; буйи <i>авалъхалда</i> <i>рйту</i> скот погнали на пастбище; бити <i>авалъх</i> <i>равнина</i>	
автабиограпи <i>ин.</i> -йабалыи, <i>рп.</i> -йалыи <i>автобиография</i> ; Диб <i>автабиограпиныи</i> <i>беси</i>	

● Automatically merge two .doc file into one unified .xls file, ...

А

АБАБЕ *бот.* пухляк издутоплотный (*кустарник*); *шмарс.* **Һаһаһу**

АБАДАДИ *нареч.* 1) никогда; **абадади гынгыһаи гучиһи ден биһтиха** никогда я к вам не приду 2) ни за что; **абадади ньевчиһи гыһуди гыб** ни за что не делает он этого

АБАДАЛ-ИЛЛАГЫЛ то же, что **абадади**

АБАЗНАВИ I (-шуб) абазинец

АБАЗНАИ II (-лыһы) абазинка

АБАЗНАЛ (-уб) абазин; **абазналуб миһиһи** абазинский язык

АБАЙ (*шмарс.*) см. **гыбай**

АБИ (*шмарс.*) см. **гыби, гыбS**

АБИР (*шмарс.*) см. **гыбир**

АБИРХА (*шмарс.*) см. **гыбSргха**

АБТАБУС III (-алы, -абалы) автобус; **абтабусабалыһи иһаһиһи** автобусная станция; **абтабусичуһи гыа** нан ехать на автобусе; **абтабусабалы гыргха** автобусы прибыли

АБТАМАТ III (-илы, -абалы) автомат (*род автоматического стрелкового оружия*); **абтаматлыку бахай!** **абтаматчи кыва гыа дай** стрелять из автомата

АБТАРЧКА III (-лы, -балы) авторучка; **абтарчкати хыарди** писать авторучкой; **абтарчкати ари** икки платить за авторучку; **абтарчкати шыкы бичи** в авторучке кончились чернила

АБТЕКА III (-лы, -балы) аптека; **абтекалы дарубалы гынгыһаи** в аптеку поступили лекарства; **абтекалыку дарубалы нхи** купить лекарства из

аваданабирлы масата букйа говорили о его жизнерадостности

АВАДАНТЫ III (-лы) веселые, веселье; **жизнерадостность**; **ракваре авадаһылы гыкына** жизнерадостный человек (*бук.* имеющий в сердце веселость человек) **аваданы гучыб гымуру** безрадостная жизнь; **ракварчы авадаһы гучыа** адам нежизнерадостный человек; **аваданы и гты** веселиться, устроить веселье

АВАЛ III (-алы, -лэ) 1) квартал; **гыб иһиб авал** ида это наш квартал; **иһиб авале** маһкид игы в нашем квартале построили мечеть 2) округ; **гыб авале нукиһа гыһугу хыов гыкына** в этом округе не было такого хорошего человека

АВАЛ-АХР III (-алы) начало и конец; **авал-ахир гучыаб муху** работа без начала и конца (*о записочной работе, о бесконечном деле, безрезультатной деятельности*)

АВАНС III (-алы) аванс; **аванси ари** гыһыһа ида получили деньги в аванс; **аванс икки** давать аванс

АВАРАГ I (-лы, -абалы) *рел.* пророк; **аварагылы амуру** веселение пророка; **аварагылы иһтери биксу хад** просить именем пророка; **аварагылы ради пророка**; **аварагыди гыһамалаб миксыди хыади** иди по пути, указанному пророком;

АВАРА III (-лы) препятствие; затруднение; помеха; **гыб муху аvara игы** этому делу помешали; **гыб авара игыһа миһи ара гынгыһаи гты** пришел, чтобы чинить мне препятствия?

АВАРИ III (-лы) авария; **авари игы** совершить аварию; **авариһа руки**

А

абабай *вводн. сл.* обязательно, естественно, конечно; **абабай ден гыкыло дуй санду** конечно, я пойду с тобой

абадаһила инкомом образом

абадаһа=б вечный

абадул-абад, *рп.* -илы *уст.* вечность

абарона, *мн.* -э, *рп.* -лы оборона; **абарона бикки** оборонять; **абароналы** оборонительный

абди, *мн.* **абдэ**, *пс.* **абду** прилечь, лечь ненадолго; полежать; **абду букйай** прислонить; **абду вигы** прислониться

абйавлэни, *мн.* -йабалы, *рп.* -лы объявление; **газэтиһа абйавлэни** объявление в газете; **газэтиһа абйавлэни кки** дать объявление в газету; **абйавлэни буххай**, **абйавлэни кыкындий**, **абйавлэни чйабдай** повестить объявление

аборт, *мн.* -абалы, *рп.* -алы аборт; **аборт игы** сделать аборт; **Гыай гыһиһаи адамиһи аборт игы** Эта женщина сделала аборт

аваданаб просторный; **аваданаб ахи** просторный сад; **аваданаб ракына** шедрая душа

аваданылы 1. простор; 2. шедрость

авал-ахир конец света; **авал-ахир зивди даһдаһи кыблн** на том свете встретимся; **авал-ахир гучыаб** заман вечность

авара, *мн.* -балы, *рп.* -лы помеха, преграда

авараг *мн.* -илы, *рп.* -де пророк; **Муса авараг** пророк Муса (Моисей); **авараг реһиһаиһа бурдаб микыи а)** радуга; б) Млечный путь

авари, *мн.* -йабалы, *рп.* -лы авария; **авари игы** совершить аварию

авиануч, *мн.* -е, *рп.* -лы авианочка; **Ишкур кагыат авианучди битю** Я послал письмо авианочкой

авлахь, *мн.* -абалы, *рп.* -алы безлесое место; **буһи авлахьада рыту** скот погнали на пастбище; **битн авлахь** равнина

автабиһраһи *мн.* -йабалы, *рп.* -йалы автобиография; **Диб автабиһраһиһалы бехи**

- Automatically merge two .doc file into one unified .xls file, ...
- Manually check for similarities (S. Verhees, C. Naccarato and me)

А

АБАБЕ *бот.* пухляк издутоплотный (*кустарник*); *шмарс.* **Панапу**

АБАДАДИ *нареч.* 1) никогда; **абадади гынгытви гучи** ден **биштиха** никогда я к вам не приду 2) ни за что; **абадади ивеччи** гышуди **губ** ни за что не делает он этого

АБАДАЛ-ИЛЛАГЫЛ то же, что **абадади**

АБАЗНАВИ I (-шуб) абазинец

АБАЗНАИ II (-лыты) абазинка

АБАЗНАЛ (-уб) абазин; **абазналу** **минчи** абазинский язык

АБАЙ (*шмарс.*) *см.* **гъабай**

АБИ (*шмарс.*) *см.* **гъаби, гъабС**

АБИР (*шмарс.*) *см.* **гъабир**

АБИРХА (*шмарс.*) *см.* **гъабСргъа**

АБТАБУС III (-алы, -абалы) автобус; **абтабусабалы** **ишани** автобусная станция; **абтабусичу** **гъа** **наи** ехать на автобусе; **абтабусабалы гьиргъа** автобусы прибыли

АБТАМАТ III (-илы, -абалы) автомат (*род автоматического стрелкового оружия*); **абтаматилыку** **бахай/абтаматчи** **къвагъадай** стрелять из автомата

АБТАРЧКА III (-лы, -балы) авторучка; **абтарчкани хъварди** писать авторучкой; **абтарчкани арси** **икки** платить за авторучку; **абтарчканы шакы бичи** в авторучке кончились чернила

АБТЭКА III (-лы, -балы) аптека; **абтекалы дарубалы гынгытви** в аптеку поступили лекарства; **абтекалыку дарубалы** **ибхи** купить лекарства из

аваданабирлы **масата** **букIа** говорили о его жизнерадостности

АВАДАНТЫ III (-лы) веселое, веселье; **жизнерадостность**; **ракIаре** **аваданытлы гъекIа** **жизнерадостный** человек (*бук.* имеющий в сердце веселое сердце) **аваданы гучIаб** **гIумру** безрадостная жизнь; **ракIарчы** **аваданы гучIав** **адам** неизнерадостный человек; **аваданы итIы** **веселиться**, устроить веселье

АВАЛ III (-алы, -лэ) 1) квартал; **гъаб ичиб** **авал** **ида** это наш квартал; **ичиб** **авале** **машкид** **игъу** в нашем квартале построили мечеть 2) округ; **гъаб** **авале** **пукIчиIа** **гыштIув** **хъовон** **гъекIа** в этом округе не было такого хорошего человека

АВАЛ-АХРР III (-алы) начало и конец; **авал-ахир гучIаб** **муху** **работа** без начала и конца (*о записанной работе, о бесконечном деле, безрезультатной деятельности*)

АВАНС III (-алы) аванс; **аванси арси** **гынгытви** **ида** получили деньги в аванс; **аванс** **икки** **давать** **аванс**

АВАРАГ I (-лы, -збалы) *рел.* пророк; **аварагилы** **амру** **веселые** пророка; **аварагилы ишIери** **биксу** **хад** **просить** именем пророка; **аварагигъалу** **ради** пророка; **аварагиди гъагъамалаб** **миксиди** **хъоиди** **иди** по пути, указанному пророком;

АВАРА III (-лы) препятствие; затруднение; помеха; **гъаб** **муху** **авара игъу** **этому** **делу** **помешали**; **гъаб** **авара игъыма** **мини** **ара** **гынгытви** **ты** **пришел**, чтобы **чинить** **мне** **препятствия?**

АВАРИ III (-алы) авария; **авари игъи** совершить аварию; **авариIа** **рукки**

А

аббай *вводи.* *сл.* обязательно, естественно, конечно; **аббай ден гынIо дуй** **санду** конечно, я пойду с тобой

абададиIа **инком** **образом**

абадина=б вечный

абадул-абад, *рп.* -илы *уст.* вечность

абарона, *мн.* -э, *рп.* -лы **оборона**; **абарона** **бикки** **оборонять**; **абароналы** **оборонительный**

абди, *мн* **абдэ**, *пс* **абду** **прилечь**, **лечь** **ненадолго**; **полежать**; **абду** **букIай** **прислонить**; **абду** **вингы** **прислониться**

абиявлени, *мн.* -йабалы, *рп.* -лы **объявление**; **газетилы** **абиявлени** **объявление** в газете; **газетилы** **абиявлени** **кки** **дать** **объявление** в газету; **абиявлени** **буххай**, **абиявлени** **кIкIанди**, **абиявлени** **чиIабдай** **повестить** **объявление**

аборт, *мн.* -абалы, *рп.* -алы **аборт**; **аборт** **игъи** **сделать** **аборт**; **Гъай** **гъеникIа** **адамиди** **аборт** **игъу** Эта женщина сделала аборт

аваданаб **просторный**; **аваданаб** **ахи** **просторный** сад; **аваданаб** **раIкIа** **щедрая** душа

аваданыIы 1. простор; 2. щедрость

авал-ахир **конец** света; **авал-ахир** **зиди** **дандалы** **кIбIа** **на** **том** **свете** **встретимся**; **авал-ахир** **гучIаб** **заман** **вечность**

авара, *мн.* -балы, *рп.* -лы **помеха**, **преграда**

авараг *мн.* -илы, *рп.* -ае **пророк**; **Муса** **авараг** **пророк** **Муса** (**Моисей**); **авараг** **решинакIа** **бурдаб** **микIы** а) радуга; б) Млечный путь

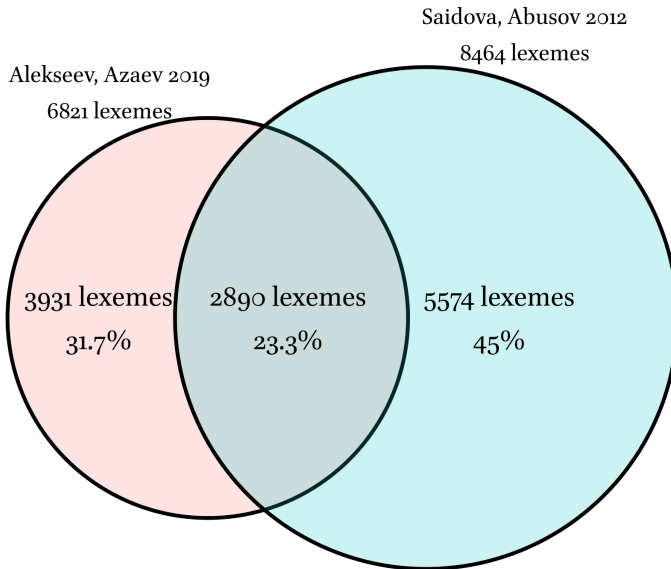
авари, *мн.* -йабалы, *рп.* -лы **авария**; **авари** **игъи** **совершить** **аварию**

авиануч, *мн.* -е, *рп.* -илы **авианочка**; **Ишкур** **кагъат** **авианучиди** **битIо** Я послал письмо авианочкой

авлахъ, *мн.* -абалы, *рп.* -алы **безлесое** место; **буни** **авлахъадда** **рIту** **скот** **погнали** **на** **пастбище**; **битIи** **авлахъ** **равнина**

автабиографIи *мн.* -йабалы, *рп.* -йалы **автобиография**; **Диб** **автабиографIи** **илы** **бехи**

Comparing two Botlikh dictionaries



Comparing two Botlikh dictionaries

- If we remove the stress sign, there are only 2495 lexemes which look phonetically the same, and 395 are different (14%)
 - If we don't remove the stress sign, there are 2027 lexemes which look phonetically the same, and 863 are different (30%)
- ⇒ 16% of lexemes have different stress pattern?

Comparing two Botlikh dictionaries

- If we remove the stress sign, there are only 2495 lexemes which look phonetically the same, and 395 are different (14%)
 - If we don't remove the stress sign, there are 2027 lexemes which look phonetically the same, and 863 are different (30%)
- ⇒ 16% of lexemes have different stress pattern? Including cases where stress is present in one dictionary and absent in the other.

Comparing two Botlikh dictionaries

- If we remove the stress sign, there are only 2495 lexemes which look phonetically the same, and 395 are different (14%)
- If we don't remove the stress sign, there are 2027 lexemes which look phonetically the same, and 863 are different (30%)
- ⇒ 16% of lexemes have different stress pattern? Including cases where stress is present in one dictionary and absent in the other.
- What causes the difference between dictionaries?

- Stress pattern differences in 317 lexemes (about 11%)
- Multiple cases where there is a small difference that could be explained either as a typo or in terms phonological variation

čuhí 'to run' [Alekseev and Azaev 2019] vs. *čŭhí* [Saidova and Abusov 2012],
kusu 'cherry plum' [Alekseev and Azaev 2019] vs. *kus:u* [Saidova and Abusov 2012]

- Multiple cases where Russian borrowings were adopted differently
awtobus 'bus' [Alekseev and Azaev 2019] vs. *abtabus* [Saidova and Abusov 2012],
biton 'milk can' [Alekseev and Azaev 2019] vs. *bitun* [Saidova and Abusov 2012],
apteka 'pharmacy' [Alekseev and Azaev 2019] vs. *abteka* [Saidova and Abusov 2012]

- Morphological preferences

dinija=w 'pious' [Alekseev and Azaev 2019] vs. *dinija=b* [Saidova and Abusov 2012]

Comparing two Botlikh dictionaries

About 25 cases:

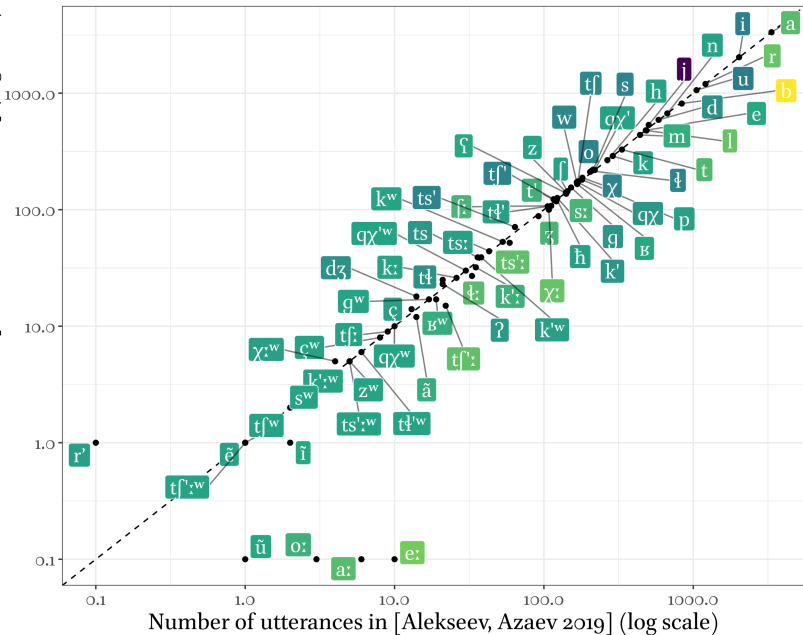
[Alekseev and Azaev 2019]	[Saidova and Abusov 2012]	
<i>āhajr</i>	<i>āhar</i>	'message'
<i>bezajr</i>	<i>bezir</i>	'roasting'
<i>mik'kujr</i>	<i>mik':ur</i>	'swallowing'
<i>reqχujr</i>	<i>reqχ^wir</i>	'fight'
<i>reſkujr</i>	<i>reſkur</i>	'overnight stay'
<i>rik^wajr</i>	<i>rik^war</i>	'lighting'
<hr/>		
<i>χwardar</i>	<i>χwardir</i>	'digging'
<i>miʔar</i>	<i>miʔar</i>	'nose'
...

About 6 cases:

<i>ʃ:alaj</i>	<i>ʃ:allaj</i>	'silt'
<i>inuʃala</i>	<i>inuʃalla</i>	'everywhere'
<i>ʃila</i>	<i>ʃilla</i>	'reason'
...

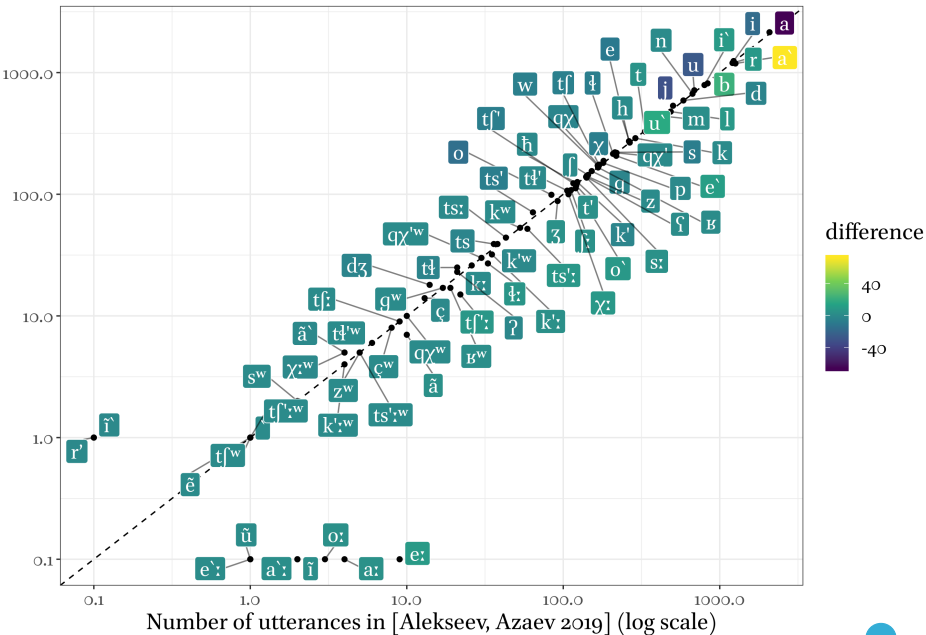
Comparing two Botlikh dictionaries (without stress patterns)

Number of utterances in [Saidova, Abusov 2012] (log scale)



Comparing two Botlikh dictionaries (with stress patterns)

Number of utterances in [Saidova, Abusov 2012] (log scale)

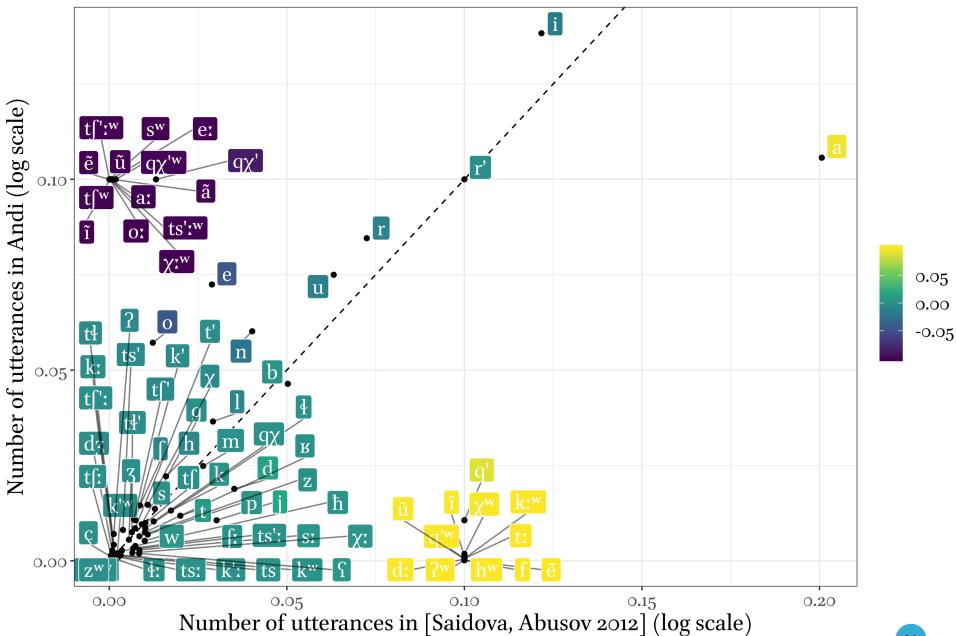


Zilo Andi data

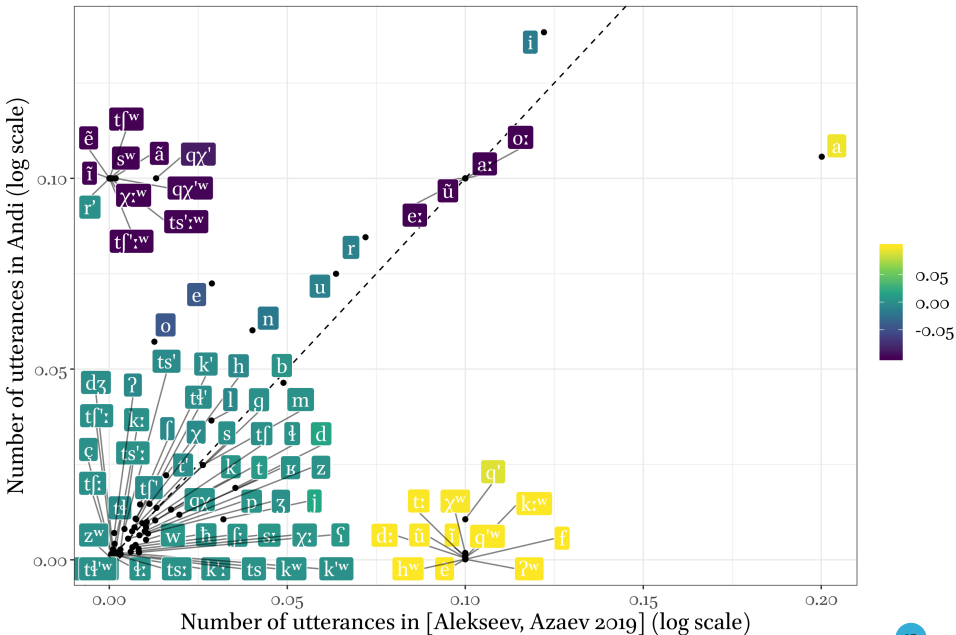
Dictionary data for Zilo were collected during fieldtrips to Zilo in 2016–2019 with N. Rochant, S. Verhees, A. Martynova and A. Zakirova who contributed to the same FieldWorks project.

- Contain morphological affixes
- Doesn't 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)
- No information about stress

Comparing [Saidova, Abusov 2012] and Zilo data



Comparing [Alekseev, Azaev 2019] and Zilo data



Discussion:

- Botlikh dictionaries were specially selected for shared meaning, the same procedure for the Andic dictionary was not done

Discussion:

- Botlikh dictionaries were specially selected for shared meaning, the same procedure for the Andic dictionary was not done
- Botlikh dictionaries contain a lot of borrowings, this is not true for the Andic dictionary

Discussion:

- Botlikh dictionaries were specially selected for shared meaning, the same procedure for the Andic dictionary was not done
- Botlikh dictionaries contain a lot of borrowings, this is not true for the Andic dictionary
- Lemmata are not the same as wordforms, so the model should be checked with the wordform material

Discussion:

- Botlikh dictionaries were specially selected for shared meaning, the same procedure for the Andic dictionary was not done
- Botlikh dictionaries contain a lot of borrowings, this is not true for the Andic dictionary
- Lemmata are not the same as wordforms, so the model should be checked with the wordform material
- Lemmata can contain some affix that will shift all frequencies (e. g. INF, PL or =CL) for some types of phonological units

Discussion:

- Botlikh dictionaries were specially selected for shared meaning, the same procedure for the Andic dictionary was not done
- Botlikh dictionaries contain a lot of borrowings, this is not true for the Andic dictionary
- Lemmata are not the same as wordforms, so the model should be checked with the wordform material
- Lemmata can contain some affix that will shift all frequencies (e. g. INF, PL or =CL) for some types of phonological units
- It would be nice to compare the obtained models with the models built on corpora data

Discussion:

- Botlikh dictionaries were specially selected for shared meaning, the same procedure for the Andic dictionary was not done
- Botlikh dictionaries contain a lot of borrowings, this is not true for the Andic dictionary
- Lemmata are not the same as wordforms, so the model should be checked with the wordform material
- Lemmata can contain some affix that will shift all frequencies (e. g. INF, PL or =CL) for some types of phonological units
- It would be nice to compare the obtained models with the models built on corpora data , when/if it will be available

Discussion:

- Botlikh dictionaries were specially selected for shared meaning, the same procedure for the Andic dictionary was not done
- Botlikh dictionaries contain a lot of borrowings, this is not true for the Andic dictionary
- Lemmata are not the same as wordforms, so the model should be checked with the wordform material
- Lemmata can contain some affix that will shift all frequencies (e. g. INF, PL or =CL) for some types of phonological units
- It would be nice to compare the obtained models with the models built on corpora data , when/if it will be available
- Model computed using row frequencies should be extended with Markov Chains and vector models (like in a Distributional semantics)

Discussion:

- Botlikh dictionaries were specially selected for shared meaning, the same procedure for the Andic dictionary was not done
- Botlikh dictionaries contain a lot of borrowings, this is not true for the Andic dictionary
- Lemmata are not the same as wordforms, so the model should be checked with the wordform material
- Lemmata can contain some affix that will shift all frequencies (e. g. INF, PL or =CL) for some types of phonological units
- It would be nice to compare the obtained models with the models built on corpora data , when/if it will be available
- Model computed using row frequencies should be extended with Markov Chains and vector models (like in a Distributional semantics): only in this way it will be possible to compare **paradigmatic** and **syntagmatic** relations within the phonological systems and across languages

Send me a letter!
agricolamzgmail.com

Presentation is available here:
tinyurl.com/y3wtkcbq



References

- Alekseev, M. and X. Azaev (2019). *Botlixsko-russkij slovar'* [*Botlikh-Russian dictionary*]. Moscow: Academia.
- Gudava, Togo, E. (1962). *Botlixuri ena* [*The Botlikh language*]. Tbilisi: Mecniereba.
- Kibrik, A. E. and S. V. Kodzasov (1988). *Sopostavitelnoye izuchenije dagestanskikh yazykov* [*Comparative study of Dagestani languages*]. Moscow: Moscow State University.
- Moroz, G., A. (2018). *lingphonology: automatic phonological description*. R package draft.
- Moroz, G., A. (2019). *Slogovaya struktura adygeyskogo yazika: ot dannyx k obosheniyam* [*Adyge syllable structure: From empirical data to generalizations*]. *Voprosy Jazykoznanija* 2, 82–95.
- Moroz, G. (2017). *lingtypology: easy mapping for Linguistic Typology*.
- Romanova, K., I. (2019). *Automatic Syllable Structure Extracting From Dictionaries: Slavic Data*. Term paper.
- Saidova, P. A. and M. G. Abusov (2012). *Botlixsko-russkij slovar'* [*Botlikh-Russian dictionary*]. Makhachkala: IJaLI.
- Salimov, X. S. (2010 (1968)). *Gagatlinskij govor andijskogo jazyka* [*The Gagatli dialect of the Andi language*]. Makhachkala.
- Suleymanov, J. G. (1957). *Grammatičeskij očerk andijskogo jazyka (po dannim govora s. Rikvani)* [*Grammar sketch of the Andi language (based on material from the dialect of the village Rikvani)*]. Ph. D. thesis, Institut Jazykoznanija AN SSSR.
- Tsertsvadze, I. I. (1965). *Andiuri Ena*. Tbilisi: Metsniereba.