



# Gender Stability in Nakh-Daghestanian Languages

Kate Bellamy, Jesse Wichers Schreur & Gerd Carling

Languages, Dialects and Isoglosses of Anatolia, the Caucasus and Iran Webinar,  
08/03/2021, “Paris”

# Roadmap

1. Gender in Nakh-Daghestanian languages
2. The present study
  - a. Data collection
  - b. Coding
3. Results
4. Discussion and next steps

---

# 1. Gender in Nakh-Daghestanian

# Gender: A definition

“Genders are classes of nouns reflected in the behavior of associated words.”

e.g. *le* chat noir- $\emptyset$  chasse *la* souris grise

(Hockett, 1958: 231, cited in Corbett, 1991:1).

# 1. North Caucasian & Kartvelian: Overview (Caucasus area)



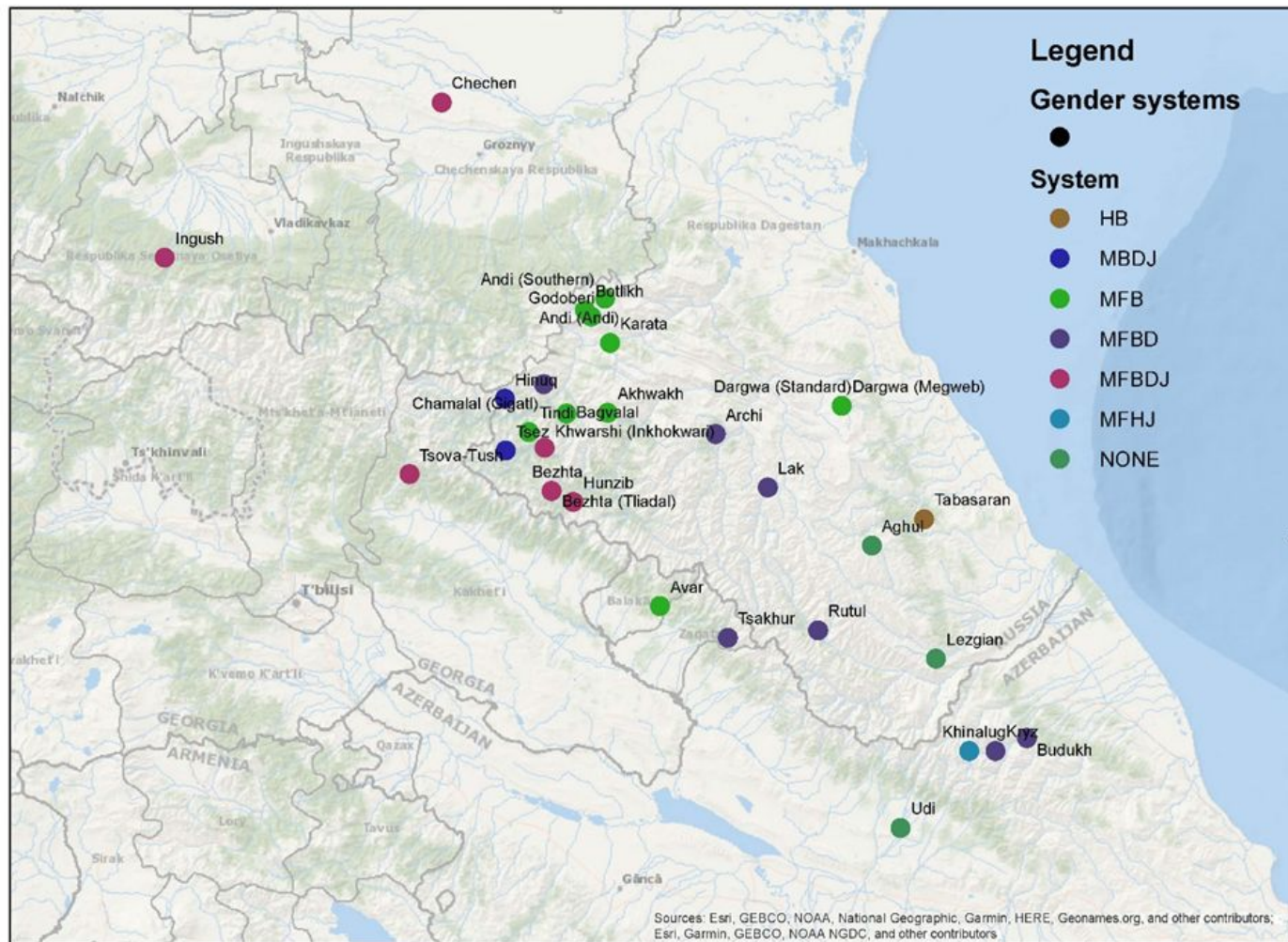
- |   |  |  |  |  |  |  |  |
|---|--|--|--|--|--|--|--|
| <p><b>NORTH CAUCASIAN {1}</b><br/> <b>ADYGHE-ABKHAZ {1A}</b><br/>         1 West Circassian {1A-aa}<br/>         2 Kabardian {1A-ab}<br/>         3 Abkhaz {1A-ca}<br/>         4 Abaza {1A-cc}</p> | <p><b>NAKH-DAGESTANIC {1B}</b><br/>         5 Ingush {1BA-aa}<br/>         6 Galanchozh {1BA-ab}<br/>         7 Chechen {1BA-ac}<br/>         8 Bats {1BA-b}</p> | <p>9 Avar {1BB-a}<br/>         10 Andi {1BB-b}<br/>         11 Botlikh {1BB-c}<br/>         12 Godoberi {1BB-d}<br/>         13 Karata {1BB-e}<br/>         14 Akhvakh {1BB-f}<br/>         15 Chamalal {1BB-g}<br/>         16 Bagvalal {1BB-h}<br/>         17 Tindi {1BB-i}</p> | <p>18 Wider Khvarshi {1BC-a}<br/>         19 Dido {1BC-b}<br/>         20 Hinukh {1BC-c}<br/>         21 Bezhta {1BC-d}<br/>         22 Hunzib {1BC-e}</p> | <p>23 Lak {1BD-a}<br/>         24 North Dargwa {1BE-1}<br/>         25 Wider Tsudakhar {1BE-2}</p> | <p>26 Kadar {1BE-3}<br/>         27 Muiirñ {1BE-4}<br/>         28 Megeb {1BE-5}<br/>         29 Sirwa {1BE-6}<br/>         30 Kunki {1BE-7}<br/>         31 Lower Vurqñi {1BE-8}<br/>         32 Kaitak {1BE-9}<br/>         33 Kubachi-Ashti {1BE-10}<br/>         34 Chirag-Amuq {1BE-11}</p> | <p>35 Archi {1BF-a}<br/>         36 Tsakhur {1BF-b}<br/>         37 Rutul {1BF-c}<br/>         38 Agul {1BF-d}<br/>         39 Tabasaran {1BF-e}<br/>         40 Lezgi {1BF-f}<br/>         41 Kryz {1BF-g}<br/>         42 Budukh {1BF-h}<br/>         43 Udi {1BF-i}</p> | <p>44 Khinalug {1BG-a}</p> <p><b>KARTVELIAN {2}</b><br/>         45 Mingrelian {a}<br/>         46 Laz {b}<br/>         47 Georgian {d}<br/>         48 Balian (Upper Svan) {e}<br/>         49 Lower Svan {f}</p> |
|---|--|--|--|--|--|--|--|

# Gender in Nakh-Daghestanian (1)

- Pervasive grammatical category
- Present in all languages except Aghul, Lezgian, Udi
- Between two and six gender classes
- Marked on certain verbs, adverbs (agreeing with the nominative argument), some adjectives, numerals and quantifiers (agreeing with their head)

# Gender in Nakh-Daghestanian (2)

- Systems comprise:
  - Male humans: M
  - Female humans: F
  - 1-3 'neuter' genders used for animals and inanimates, named for their agreement markers: B, D, J
  - Assignment principles are language-specific for each language, but are primarily based on semantic factors, with phonology a secondary influence
    - More fine-grained sub-divisions, e.g. animals third class and liquids in fourth (Klimov 1978: 67; Corbett 1991: 25ff)





# Examples

(1a) Sanzhi Dargwa (Forker, 2020)

*r-uqna*    *aba*                    *le-r=de*                    *di-la*  
F-old    mother(F)    exist-F=PST    1SG-GEN

‘My old mother was alive.’

(1b) Tsova-Tush (Hauk & Harris, forthc.)

*tiši<sup>n</sup> c’a*                    *dañ d-ox-d-Ø-o-t*                    *ve*  
old house(D)    PV    D-destroy-D-TR-PRS-PL    1PL.INCL

‘We are tearing down the old house.’

# Gender stability (1)

- A noun's gender is not necessarily constant. It can differ between languages of the same family/branch (van Epps et al., forthc.), both:
  - Within one concept set, e.g.: 'moon'
    - Masc. in Germanic languages, viz. *der Mond* in German.
    - Fem. in Romance languages, e.g. *la luna* in Spanish & Italian (Matasović, 2014).
  - Within one cognate set:
    - e.g. Sanskrit *cakrás* (masc.) vs. Proto-Germanic *\*hweywlán* (neut.) < same root 'wheel'.

# Gender stability (2)

- Motivations for these changes are still poorly understood
- Are they connected to:
  - A specific gender class?
  - Particular semantic domains or words?
  - Frequency and therefore usage?
  - Loanword status?
  - Or a combination of the above?

(e.g. Carling, 2020)

---

## 2. The present study

# Aim

To investigate the distribution and stability of gender classes across 18 Nakh-Daghestanian languages, in both inherited and borrowed lexemes, using a predefined list of cultural vocabulary items.

# Research questions

- How are the gender classes distributed in Nakh-Daghestanian languages?
  - Are certain semantic domains over-/underrepresented in particular gender classes?
  - What is the relationship between the gender of a concept with semantic categories such as animacy, abstractness, naturalness?
  - How stable are cognate sets in the different gender classes?

# Method

- Collect wordlists of 95 vocabulary items from dictionaries for 18 N-D languages.
- Transliterate into Latin script where necessary.
- Classify lexemes for semantic domain according to colexification and meaning change within the 104-lexeme culture list of the Mouton Atlas of Languages and Cultures (Carling, 2019).
- Coded for gender, improved cognate coding

# Domains and concepts used

Domain	Concepts						
CATTLE	BULL	CATTLE	COW				
CROPS	BARLEY	FLAX	GRAIN (GENERIC)	OATS	RYE	WHEAT	
DOMESTIC ANIMALS	CAT	DOG					
DOMESTIC INSECTS	BEE						
DRAFT ANIMALS	DONKEY	HORSE	OX				
DRINK AND DRUGS	MEAD	WINE					
GAME ANIMALS	BISON	HARE	RABBIT	WILD BOAR			
IMPLEMENTS	KNIFE	SAW	SCYTHE	SICKLE	SPADE		
MATERIALS	FUR	GREASE	LEATHER	MEAT	STONE	WOOD	
METALS	COPPER	GOLD	IRON	SILVER			
PIG RAISING	PIG	PIGLET					
POULTRY	CHICKEN	DUCK	HEN	ROOSTER			
PREDATOR ANIMALS	BEAR	FOX	JACKAL	LEOPARD	LION	SNAKE	WOLF
PREDATOR BIRDS	EAGLE	RAVEN					
PRODUCTS	HONEY	HOPS	MILK	SALT	WAX (BEES)	WOOL	
SEASONS	AUTUMN	HARVEST	SPRING	SUMMER	WINTER		
SMALL CATTLE	GOAT	LAMB	RAM	SHEEP			
TILLAGE	CULTIVATED FIELD	PLOW (N)					
VEGETABLES & FRUIT	APPLE	GRAPE	TURNIP				
VEHICLES	WAGON	WHEEL	YOKE				
WEAPONS	ARMY	ARROW	AXE	BOW	SHIELD	SPEAR	SWORD



Lexeme_ID	Lang_Name	Word_Transcription	Word_Meaning	Gender_Normalised	Gender_LanguageSpecific	Gender_Marker	Classification_Stats_Name	Word_List_Item	J+K_Cognacy	Borrowed	Source_Language
90798	Aghul	k'ur	wood	X			MATERIALS	WOOD		No loan	
91699	Andi	l'udi	wood	B	IV	b/b	MATERIALS	WOOD	wood-4	No loan	
91599	Avar	c'ul	wood	B	III	b/l,r	MATERIALS	WOOD		No loan	
198	Batsbi	xe <sup>n</sup>	wood	B	B	b/d	MATERIALS	WOOD		No loan	
90556	Beshta	hudo	wood	?			MATERIALS	WOOD		No loan	
90557	Beshta	hæžæ	wood	?			MATERIALS	WOOD		No loan	
90558	Beshta	ažbar	wood	?			MATERIALS	WOOD		No loan	
89973	Budukh	t <sup>h</sup> axt <sup>h</sup> y	wood	?			MATERIALS	WOOD		Loan	Azerbaijani
89974	Budukh	k'əna	wood	?			MATERIALS	WOOD	wood-3	No loan	
703	Chechen	deč <sup>h</sup> ig	wood	J	J	j/j	MATERIALS	WOOD	wood-1	No loan	
91906	Dargwa	galga	wood	B	III	b/d	MATERIALS	WOOD		No loan	
91907	Dargwa	urc <sup>h</sup> ul	wood	B	III	b/d	MATERIALS	WOOD		No loan	
91135	Ingush	daxč <sup>h</sup> ä	wood	D	D	d/d	MATERIALS	WOOD	wood-1	No loan	
90210	Khinalug	odun	wood	J	IV	j/j	MATERIALS	WOOD		Loan	Azerbaijani
90211	Khinalug	t <sup>h</sup> axt <sup>h</sup> a	wood	J	IV	j/j	MATERIALS	WOOD		Loan	Persian
90435	Khwarshi	lida	wood	?			MATERIALS	WOOD	wood-4	No loan	
90109	Kryz	zabaæ	wood	D	IV	d/d	MATERIALS	WOOD		No loan	
90110	Kryz	k'əna	wood	D	IV	d/d	MATERIALS	WOOD	wood-3	No loan	
91793	Lak	t'ama	wood	D	IV	d/d	MATERIALS	WOOD		No loan	
284	Lezgian	k'aras	wood	X			MATERIALS	WOOD		No loan	
91041	Rutul	us	wood	?			MATERIALS	WOOD	wood-2	No loan	
91042	Rutul	xuk <sup>h</sup>	wood	B	III	b/d	MATERIALS	WOOD		No loan	
90661	Tabasaran	gak'ul	wood	B	NH		MATERIALS	WOOD		No loan	
90927	Tsakhur	os	wood	D	IV		MATERIALS	WOOD	wood-2	No loan	
90928	Tsakhur	jiv	wood	B	III		MATERIALS	WOOD		No loan	
90321	Udi	durut'	wood	X			MATERIALS	WOOD		No loan	

# Standardisation process

- Gender class standardised across the languages, based on historical development of gender markers in N-D.

ID	Language	Word	Word	Gender	Gender	Gender
	Name	Transcription	Meaning	Normalised	Language specific	Markers
91732	Andi	berkʰa	snake	B	III	b/j
91655	Andi	beli	spade	B	IV	b/b
91716	Andi	ribu	spring	D	V	r/r
5763	Tsova-Tush	pʰɯ	dog	D	D	d/d

- Removed languages that have just one gender for all non-human nouns

# Cognacy and borrowings

- Majority of lexemes also coded either for cognacy or as loanwords, with the provenance of the latter indicated where known.
- Some lexemes do not have known cognates within the dataset.
- Result: about half of the lexemes have both cognacy and gender coding (however, many cognate sets have only two lexemes).

---

## 3. Results

# The data in raw numbers (1)

Language	Lexemes	Comment
Aghul	149	No gender
Andi	93	
Avar	110	MFB
Batsbi	97	
Beshta	131	
Budukh	118	
Chechen	80	
Dargwa	122	MFB
Ingush	86	
Khinalug	113	
Khwarshi	88	
Kryz	133	
Lak	93	
Lezgian	99	No gender
Rutul	107	
Tabasaran	107	
Tsakhur	124	
Udi	135	No gender
<b>Total</b>	<b>1985</b>	

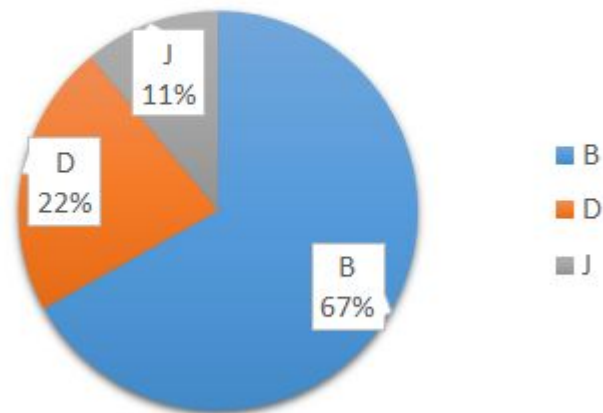
Language	Lexemes
Andi	93
Batsbi	97
Beshta	131
Budukh	118
Chechen	80
Ingush	86
Khinalug	113
Khwarshi	88
Kryz	133
Lak	93
Rutul	107
Tabasaran	107
Tsakhur	124
<b>Total</b>	<b>1370</b>

Removal of Aghul, Lezgian and Udi (no gender) and Avar, Dargwa (with M/F/B system)

# The data in raw numbers (2)

Gender	Lexemes
B	819
D	271
J	135
<b>Total</b>	<b>1225</b>

Distribution of B, D, J

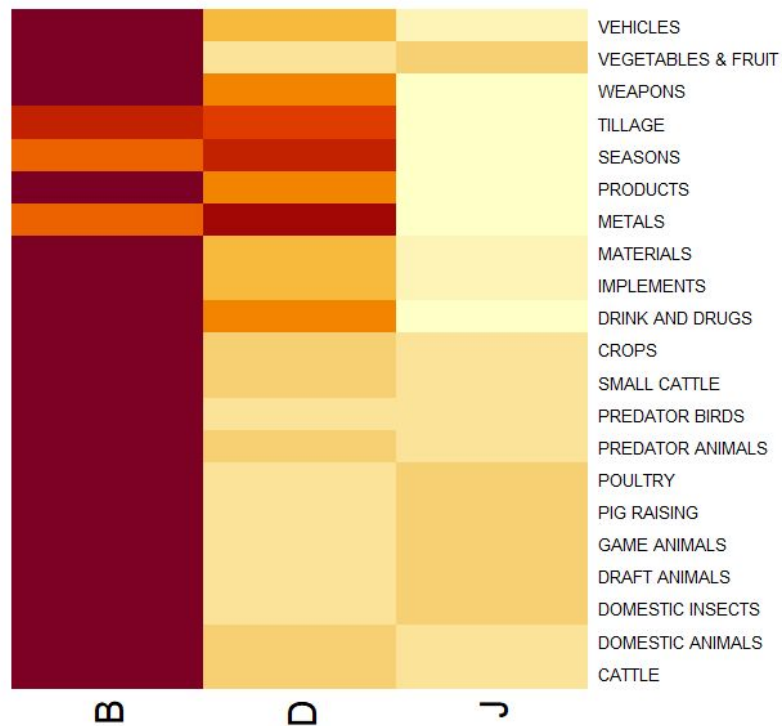


Removal of X (no gender available), ? (gender not known) and M (masculine)

# Gender/semantic domain (1)

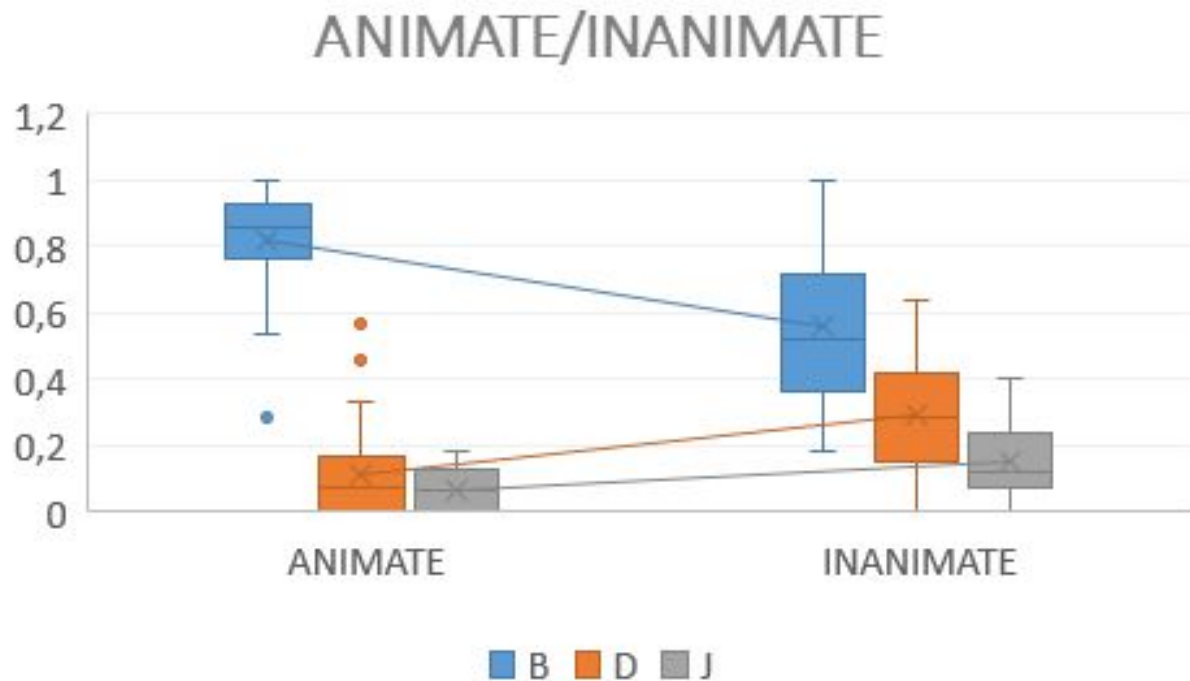
CLASS	B	D	J
DOMESTIC INSECTS	92,31%	0,00%	7,69%
SMALL CATTLE	91,36%	6,17%	2,47%
GAME ANIMALS	89,19%	2,70%	8,11%
DRAFT ANIMALS	87,72%	3,51%	8,77%
CATTLE	87,27%	10,91%	1,82%
DOMESTIC ANIMALS	85,37%	14,63%	0,00%
PREDATOR BIRDS	84,62%	7,69%	7,69%
VEGETABLES & FRUIT	82,14%	7,14%	10,71%
PREDATOR ANIMALS	80,22%	14,29%	5,49%
CROPS	79,52%	13,25%	7,23%
PIG RAISING	77,78%	7,41%	14,81%
POULTRY	76,47%	9,80%	13,73%
IMPLEMENTS	64,44%	27,78%	7,78%
MATERIALS	59,43%	26,42%	14,15%
DRINK AND DRUGS	53,13%	34,38%	12,50%
WEAPONS	51,72%	34,48%	13,79%
PRODUCTS	51,16%	33,72%	15,12%
VEHICLES	43,48%	30,43%	26,09%
TILLAGE	41,38%	37,93%	20,69%
METALS	38,33%	48,33%	13,33%
SEASONS	37,14%	41,43%	21,43%

# Gender/semantic domain (2)



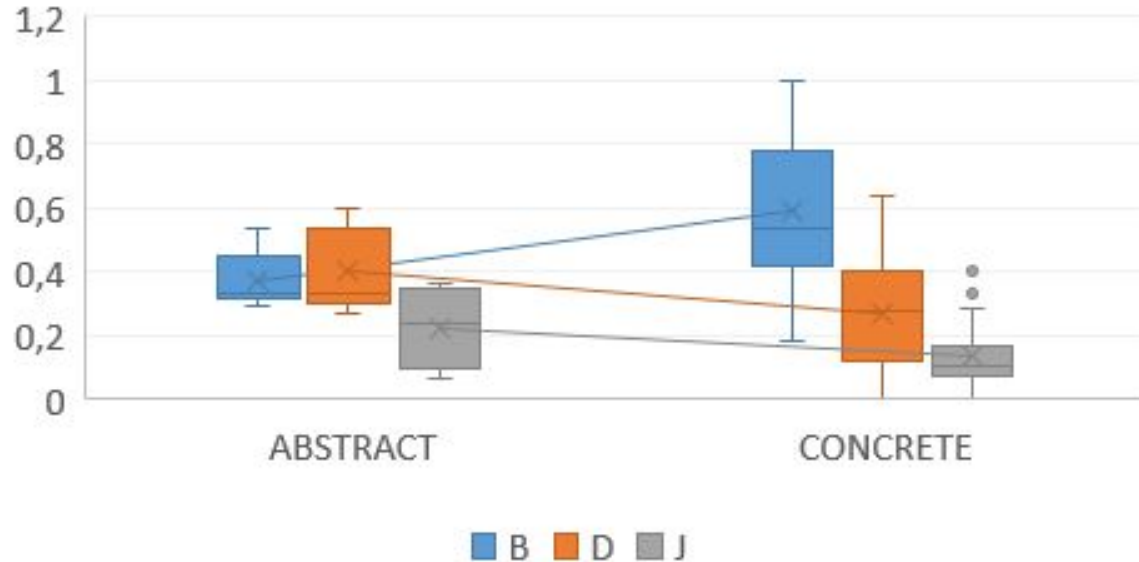


# Gender/concept (1): Animacy

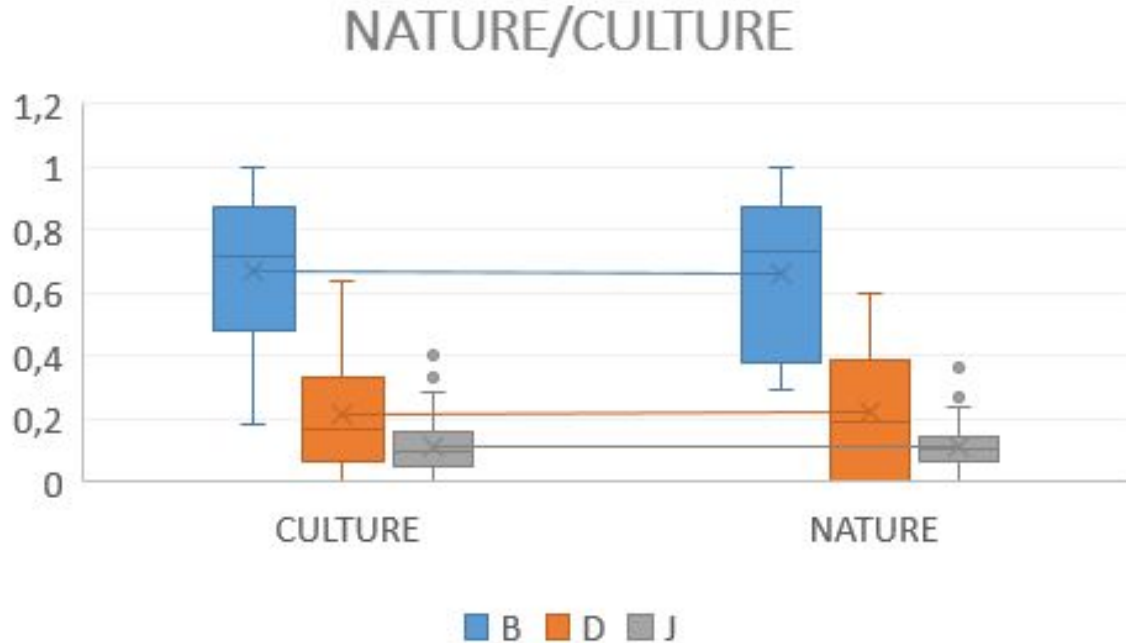


# Gender/concept (2): Abstractness

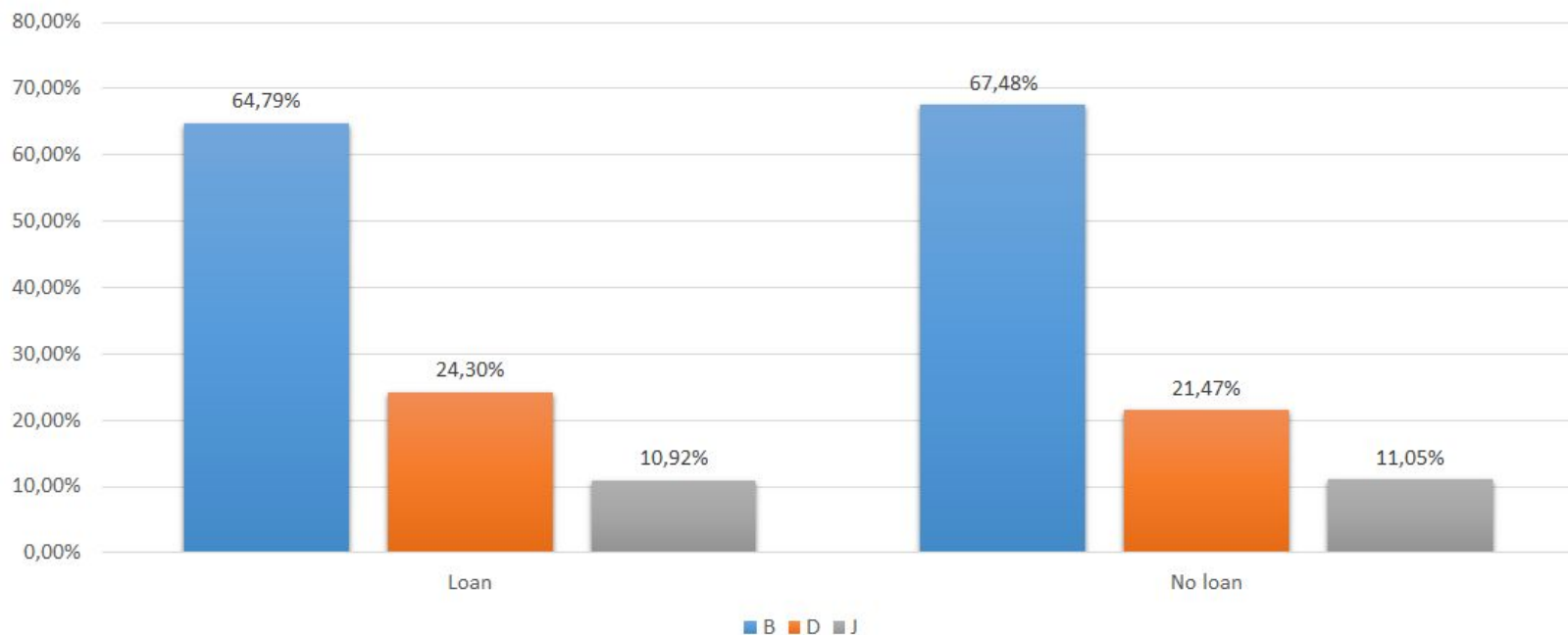
ABSTRACT/CONCRETE (INANIMATE only)



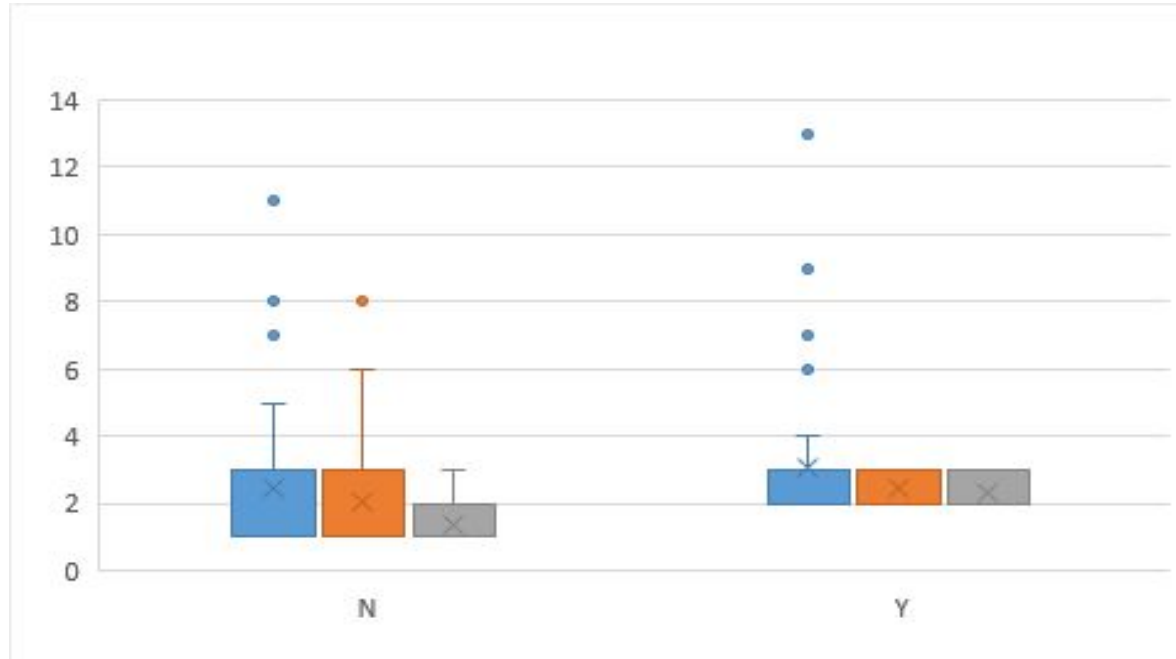
# Gender/concept (3): Nature vs culture



# Inherited vs. borrowed lexemes



# Cognacy



Gender distribution of change (N) versus stable cognates (Y) (in absolute numbers)

---

## 4. Discussion and next steps

# Main findings (1)

- B is most represented gender (c. 67% of lexemes), even with languages that have B as only 'neuter' gender removed.
  - Nonetheless, some languages have B+D or B+J (or B+D+J), so B is still over-represented.
- All genders can mark animate and inanimate concepts, but there are internal differences.
  - B is over-represented in animate groups, e.g. cattle, domestic animals.
  - D is over-represented in inanimates, with metals especially being an interesting case for further research.

## Main findings (2)

- Within inanimates, B gender correlated with concreteness, D (not B) is over-represented in abstract nouns (notably, seasons)
  - Falls in line with established hierarchy:  
*Masculine > Feminine > Animal > Concrete object > Abstract notion*
- No correlation between gender and:
  - Nature vs. culture
  - Borrowedness
  - Stability in cognate set



# Next steps

- Complete cognate coding
- Expand number of languages
- Expand number of concepts
- Possible avenues:
  - Comparison between languages with precisely the same gender system
  - Comparison with Indo-European
  - Stability of loanwords

# Thank you for your attention!

[kate.bellamy@cncrs.fr](mailto:kate.bellamy@cncrs.fr)

[wichersschreur@em.uni-frankfurt.de](mailto:wichersschreur@em.uni-frankfurt.de)

[gerd.carling@ling.lu.se](mailto:gerd.carling@ling.lu.se)

---



Language	Reference
Budukh	Mejlanova, U. A. 1984. Buduxsko-Rusko Slovar'. Moscow: Nauka
Tsakhur	Kibrik, A. E. & Testelec, Ya. G. (eds.) 1999. Ėlementy caxurskogo jazyka v tipologiĉeskom osveŝĉenii. Moscow: Nasledie
Kryz	Hümətov, İ. Ə. & Rind-Pawlowski, M. 2020. Qrızca-azərbaycanca-İngiliscə lüğət. Baku
Rutul	Alisultanov, A. C. & Suleymanova, T. A. 2019. Rutul'sko-Russkij Slovar'. Makhachkala: Alef
Tabasaran	Alekseev, M. & Shikhalieva, S. 2003. Tabasaranskij jazyk. Moscow
Tsova-Tush	Kadagidze, D. & Kadagidze, N. 1984. ts'ova-tushur-kartul-rusul leksik'oni. Tbilisi: Mecniereba Fieldwork by: J. Wichers Schreur
Chechen	Matsiev, A. G. 1961. Āĉĉensko-russkij slovar'. Moscow
Ingush	Nichols, J. 2015. Ingush-English and English-Ingush dictionary. London: Taylor and Francis
Khinalug	Fieldwork by: M. Rind-Pawlowski
Andi	Fieldwork by: S. Verhees
Avar	Forker, D. (forthc.) Avar grammar sketch. In: Handbook of Caucasian languages, Polinsky, M. (ed.). Oxford: Oxford
Bezhta	Khalilov, M. Sh. 1995. Beŝtinsko-russkij slovar'. Makhachkala
Dargwa (Standard)	Van den Berg, H. 2001. Dargi Folktales (and Grammar). Leiden: CNWS
Lak (Standard)	Khajdakov, S. M. 1962. Laksko-russkij slovar'. Moscow