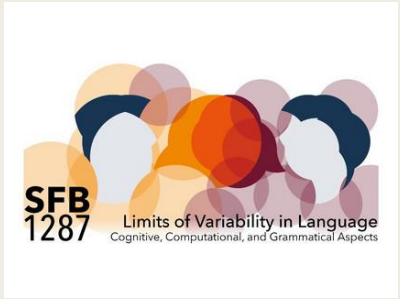
PaVeDa workshop: Verbal valency, valency databases and Pavia Verbs Database
Univeristy of Pavia, 15 May 2024

Complexity in valency class systems and its typological correlates

Sergey Say
sergey.say@uni-potsdam.de
University of Potsdam



Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 317633480 – SFB 1287

Valency classes: intro

(1) *Sie* NOM *wartet* *auf ihren Bruder* auf+ACC

'She is waiting for her brother.'

(2) *Mir* DAT *fehlt* *ein Euro* NOM

'I am one Euro short.'

Valency encoding devices

- are related to the meaning of verbs
- contain cues that can be used in sentence perception
- **are complex**

Valency classes: intro

(1) *Sie* NOM *wartet* *auf ihren Bruder* auf+ACC

'She is waiting for her brother.'

(2) *Mir* DAT *fehlt* *ein Euro* NOM

'I am one Euro short.'

Valency encoding devices

- but are they **equally** complex across languages?
- if not, how can these differences be captured?
- how can they be interpreted?

Structure of the talk

- Background & goals
- Data & methods
- Results: valency class system complexity and…
  - its components
  - its cross-linguistic variation
  - its typological correlates
  - word order
- Summary and discussion

Structure of the talk

- **Background & goals**
- Data & methods
- Results: valency class system complexity and…
  - its components
  - its cross-linguistic variation
  - its typological correlates
  - word order
- Summary and discussion

Background

- The valency of a verb = “the list of its arguments with their coding properties” (Malchukov et al. 2015: 30)

- Coding properties (devices)
  - flagging: cases & adpositions
  - indexing: agreement, cross-referencing
  - word order (rarely)

Background

- (mainly) flagging: trivial examples, e.g. (1)-(2)

- (mainly) indexing

Abaza (< Northwest Caucasian)

(3) *fatíma murád jə-z-qá-l-ç-əj-t*
PN PN [3SG.M.IO-BEN]-LOC-[3SG.F.ERG]-believe-PRS-DCL
'Fatima trusts Murad.'

- (mainly) word order: *Mary kissed Peter*

Background

- The **discriminatory** function of case and other argument-coding devices (Comrie 1989: 124-127; Dowty 1991; Levshina 2021: 4 ; Seržant 2019)

=> facilitates establishing the role-reference associations in discourse

Background

- A minimal system that would fulfil that goal
  - All monovalent verbs are uniform (e.g. NOM)
  - All bivalent verbs are uniform (e.g. NOM-ACC)*
  - All trivalent verbs are uniform (e.g. NOM-ACC-DAT)*

  – such a system would facilitate production
  – but encoding devices would be poor in semantic content => fewer cues for the hearer

* These systems can be reduced even further, e.g. via DOM

Background

- More complex systems could provide additional cues for the hearer?

(4) *Teilhabe darf nicht vom Alter …*

Background

- More complex systems could provide additional cues for the hearer?

(4) *Teilhabe darf nicht vom Alter abhängen* (www)

‘Participation must not depend on age’

Background

- Hypothetical maximally complex system: each verb is associated with a unique argument-encoding scheme

=> such a system would maximize cue reliability

=> but would be too costly for the speaker

Background

In reality

- the transitive class universally constitutes the core of bivalent verbs (Tsunoda 1985; Næss 2007)
- all languages possess verbs deviating from the transitive class
- bivalent verbs are especially prone to display deviant valency behaviour (Bickel et al. 2014)

=> all languages find a **point of equilibrium**

Goals

- To propose a technique that can be used to measure the degree of heterogeneity: entropy
- Identify the limits of cross-linguistic variation
- Detect correlations with other parameters
- To (try to) explain them by appealing to efficiency constraints

Structure of the talk

- Background & goals
- **Data & methods**
- Results: valency class system complexity and…
  - its components
  - its cross-linguistic variation
  - its typological correlates
  - word order
- Summary and discussion

Data

- Say, Sergey (ed.). 2020--... BivalTyp: Typological database of bivalent verbs and their encoding frames. (Available online at https://www.bivaltyp.info)

Data

- Questionnaire with 130 verbs given in context

  => “probes” in the infinite semantic space

- First-hand data provided by language experts

- Disclaimer: types (in the lexicon) not tokens (in discourse)

Data

#21 (Peter was crossing the river in a boat)

‘Peter **reached** the bank’

X Y

#22 (The wall was covered with fresh paint)

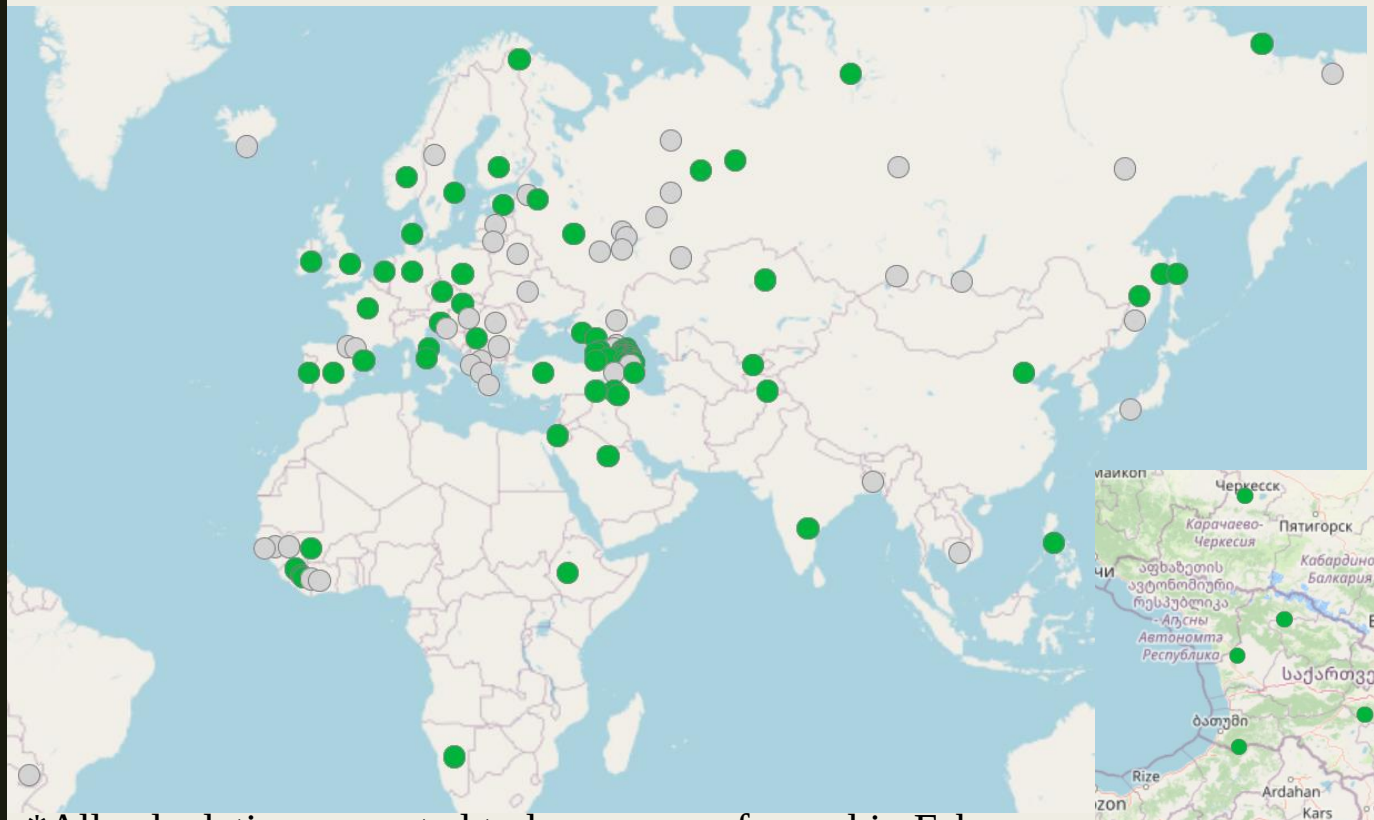
‘Peter **touched** the wall’ (and got dirty)

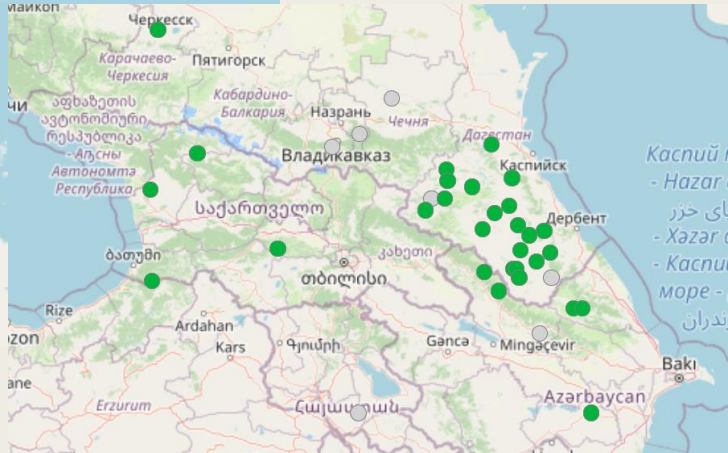
X Y

=> Two pre-defined arguments (X, Y) for each predicate

Data

- The sample: currently 124* languages, mainly spoken in Northern Eurasia (75 “published” lgs are in green)





*All calculations reported today were performed in February 2024. Since then, the sample size has reached 127.

Data

- Each construction is tagged for its (language-specific) valency pattern: encoding of X and Y

- The pattern is considered transitive iff its X and Y arguments are coded like the two arguments of the sentence with ‘kill’, see also (Haspelmath 2015: 136)

Data

- Some patterns in Kina Rutul (Nakh-Daghestanian)

31	call	rasul-a ramazan-ɨ-χda ses ha<w>ɨ-r PN-ERG PN-OBL-SUB sound <3>do.PFV-CVB 'Rasul called Ramazan.'	ERG_SUB
32	get_to_know	rasul ramazan-ɨ-kan sa sen xura gi-na taniš jiši-d PN(NOM) PN-OBL-COM one year in front SUB.be-CVB acquainted 1.become.PFV-ATTR 'Rasul got to know Ramazan a year ago.'	NOM_COM
33	know	rasul-u-s hac'a-r=a ramazan PN-OBL-DAT 1.know.IPFV-CVB=be PN(NOM) 'Rasul knows Ramazan.'	DAT_NOM
34	play#instrument#	rasul-a saz wɨ-rɨˁχɨˁ-r=a PN-ERG saz(NOM) 3-hit.IPFV-CVB=be 'Rasul is playing the saz.'	TR
35	avoid	rasul ramazan-ɨ-la: gʲa-ʁaˁbʁu-r=a PN(NOM) PN-OBL(SUP)-EL PV-1.avoid.PFV-CVB=be 'Rasul avoids Ramazan.'	NOM_SUP.EL
36	make	rasul-a bomba ha<w>ɨ-r PN-ERG bomb(NOM) <3>do.PFV-CVB 'Rasul made a bomb.'	TR
37	make_fun	rasul-a ramazan-ɨ jiʁ laɢa-r=a PN-ERG PN-OBL(SUP) mock LV.IPFV-CVB=be 'Rasul is making fun of Ramazan.'	ERG_SUP
38	have	rasul-ɨ-χda mašin a-ni w-iʔi PN-OBL-SUB car(NOM) be-CVB 3-COP 'Rasul has a car.'	SUB_NOM

Methods: introducing entropy

- Casual observations: languages differ

Table 1. Selected verbs and valency patterns in Joola-Fonyi and Khwarshi.

	Joola-Fonyi		Khwarshi	
meaning	verb	pattern	verb	pattern
'be afraid'	*kólí*	TR	*j/uƛ'a*	ABS_CONT
'avoid'	*ŋom*	TR	*j/iča*	ABS_CONT.EL
'wait'	*kob*	TR	*gic'a*	ABS_CONT.LAT
'attack'	*lóúm*	TR	*k'oƛa*	ABS_SUPER
'win, beat'	*ŋoolen*	TR	*j/iža*	ABS_SUPER.EL
'see'	*juk*	TR	*j/akʷa*	DAT_ABS
'touch'	*gor*	TR	*j/etaχa*	ERG_CONT
'bite'	*rum*	TR	*hana*	ERG_GEN1
'be angry'	*leet*	TR	*semi mak'a*	GEN1_CONT.LAT
'eat'	*ri*	TR	*j/ac'a*	TR (ERG_ABS)

Khwarshi clearly is more complex / conveys more information

Methods: introducing entropy

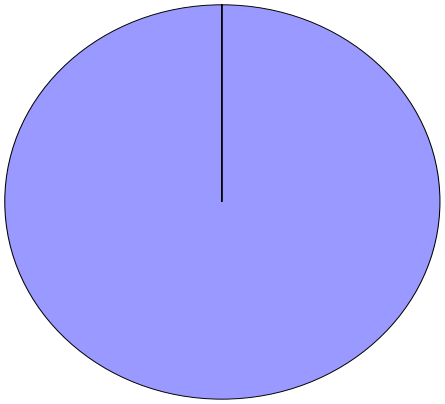
- Entropy: informal introduction

A mathematical tool aimed at quantizing the amount of information associated with a certain discrete variable

- Technicalities: Shannon's entropy (in nats)

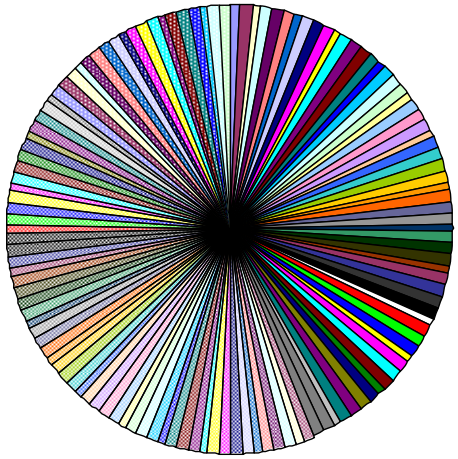
Methods: introducing entropy

$$H\ (x) = -\sum_{i=1}^{k} p(x_i) \cdot \log(p(x_i))$$



Hypothetical Language 1: All verbs belong to the same class

$$H = 0$$

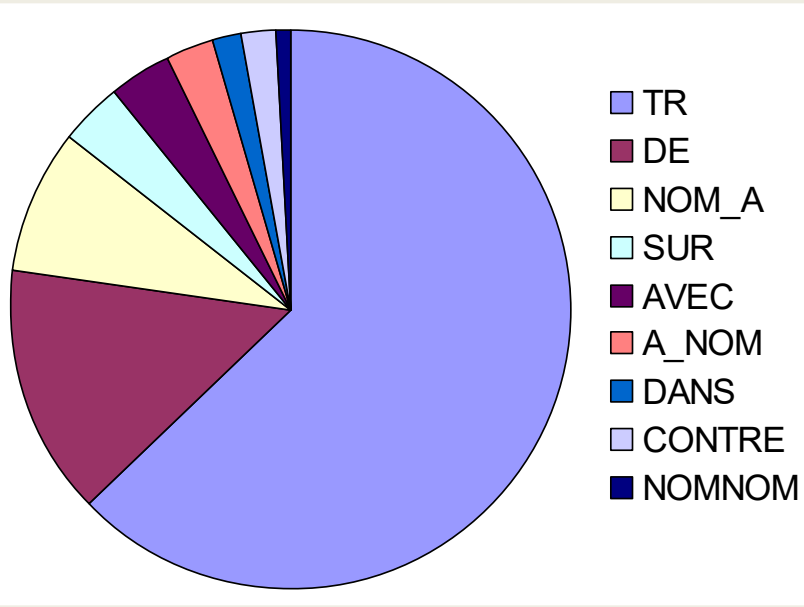


Hypothetical Language 2: 130 verb classes

$$H = \log(\frac{1}{130}) \approx 4{,}87$$

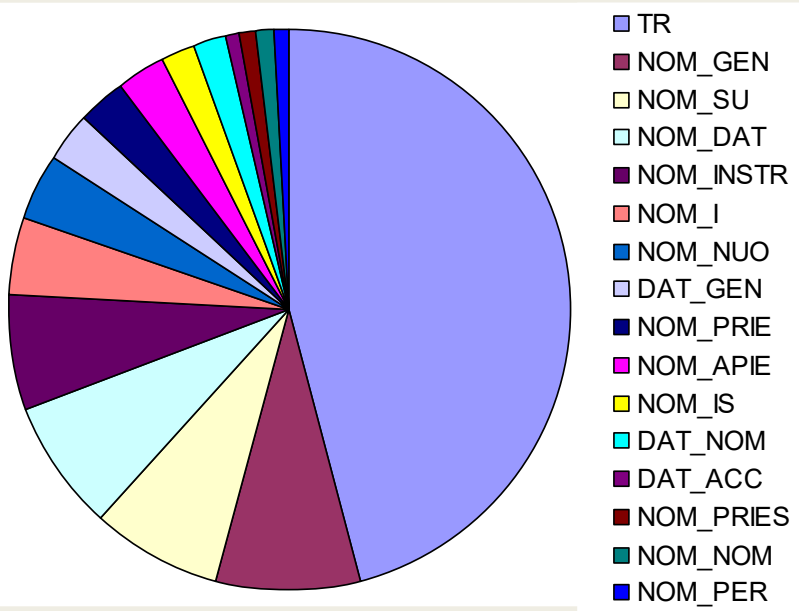
Methods: introducing entropy

French



H ≈ 1.31

Lithuanian



H ≈ 2.02

Correction

- Problem: observed Shannon's entropy (slightly) depends on the number of datapoints in the dataset
- The number of datapoints depends on non-essential factors (e.g. speaker proficiency)
- Solution: randomly select 90 datapoint for each languages, run 100 iterations, calculate average observed H
- Side-effect: disregard 4 languages with <90 datapoints
- Henceforth, "H" stands for "corrected H"

Correction

- Problem: observed Shannon's entropy (slightly) depends on the number of datapoints in the dataset
- The number of datapoints depends on non-essential factors (e.g. speaker proficiency)
- Solution: randomly select 90 datapoint for each languages, run 100 iterations, calculate average observed H
- Side-effect: disregard 4 languages with <90 datapoints
- Henceforth, "H" stands for "corrected H"

Correction

No	language	overall N	entropy	
			observed	corrected
67	Macedonian	129	1.84	1.81
68	Kina Rutul	116	2.69	2.62
69	Adyghe	125	2.73	2.65
70	Goo	92	2.27	2.27
71	Chuvash	128	1.67	1.64

Entropy of intransitives

- The value of entropy, H, captures the distribution of **all** verbs among all valency classes
- Inevitably, it largely depends on the prevalence of transitive verbs (see below)
- I am mainly interested in the distribution of verbs among intransitive classes
- Solution: calculate H_{intr}, the entropy observed in the distribution of intransitive verbs
- Correction was also necessary (this time, 24 is used as the limit for resampling procedures)

Structure of the talk

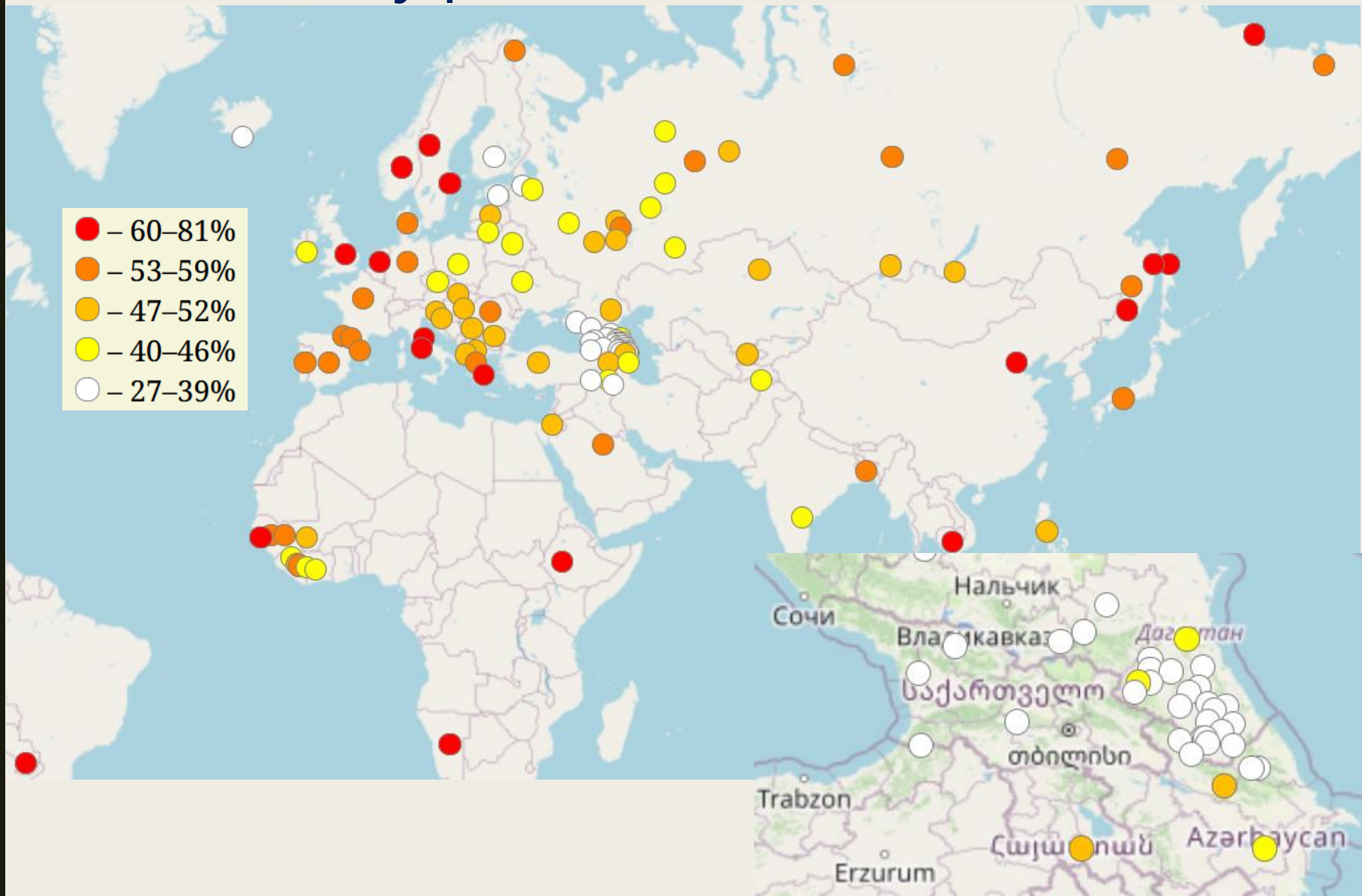
- Background & goals
- Data & methods
- **Results: valency class system complexity and…**
  - **its components**
  - its cross-linguistic variation
  - its typological correlates
  - word order
- Summary and discussion

Transitivity prominence

- Transitivity prominence = the number of transitive entries divided by the total number of entries in the dataset, see also (Haspelmath 2015)

- Transitivity prominence displays robust areal patterning

Transitivity prominence

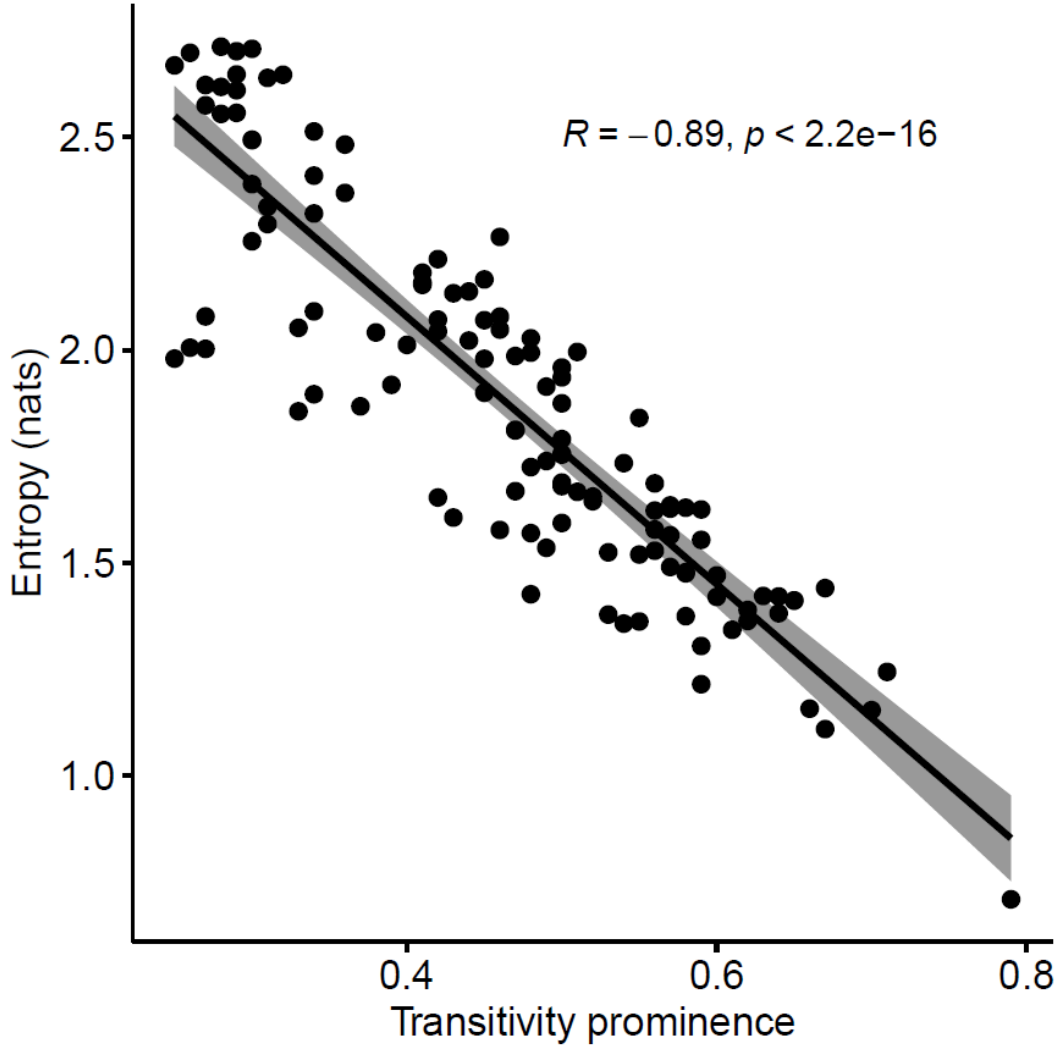


Transitivity prominence & H

- H (entropy) has a strong negative correlation with the Transitivity prominence

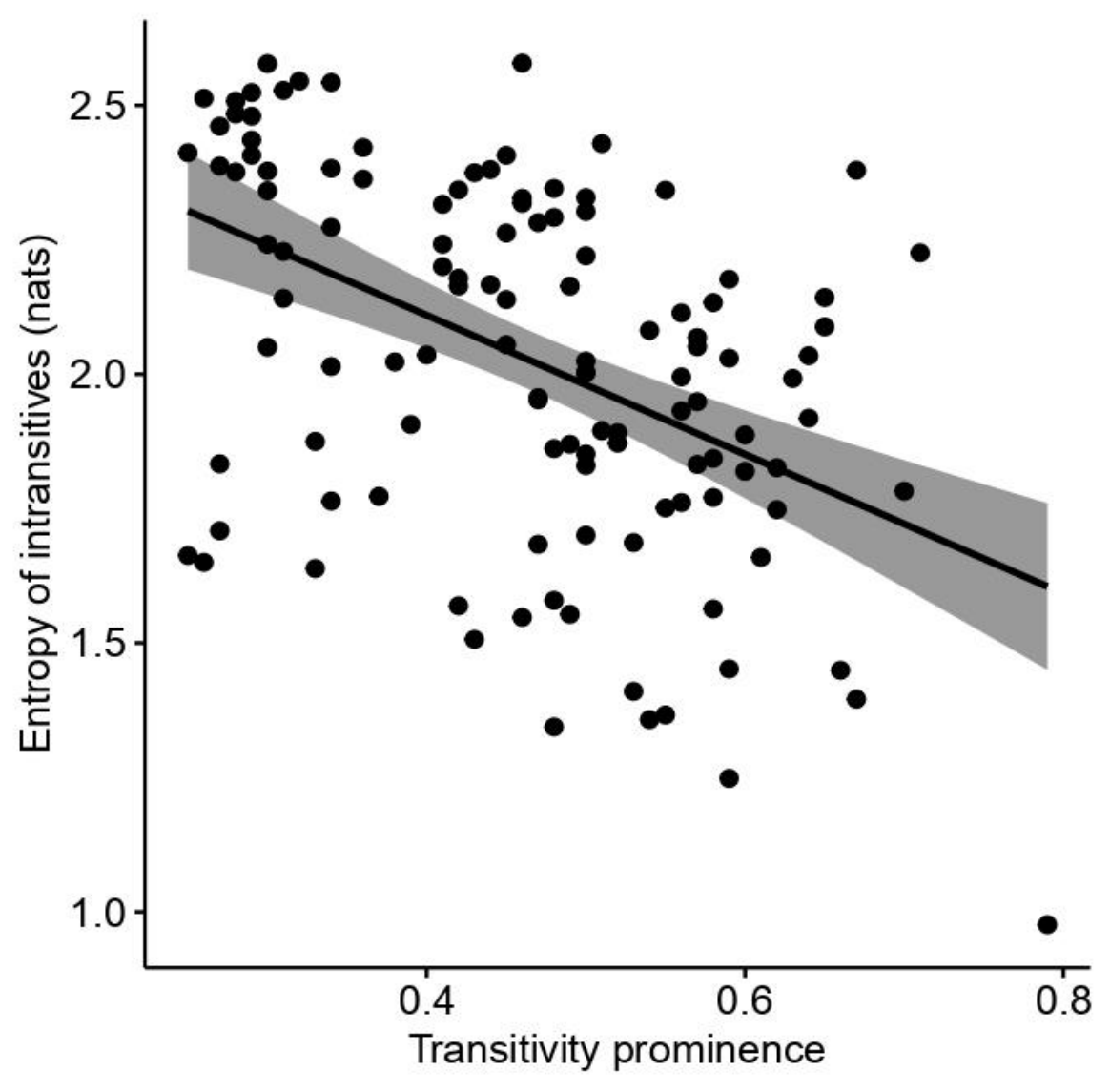
  <= Obvious reasons: the transitive class is also the biggest class and does not contribute much to the overall entropy

Transitivity prominence & H



This and similar plots below are built in R using the package `ggpubr` (Kassambara 2023)

Entropy of intransitives & H_{intr}



Transitivity prominence and H_{intr}

- H_{intr} also has a strong negative correlation with the Transitivity prominence
    - **NB: these two parameters are logically independent of each other**

=> Languages that have a more restricted transitive class also tend to make finer semantic distinctions *between* intransitive classes.

Structure of the talk

- Background & goals
- Data & methods
- **Results: valency class system complexity and…**
  - its components
  - **its cross-linguistic variation**
  - its typological correlates
  - word order
- Summary and discussion

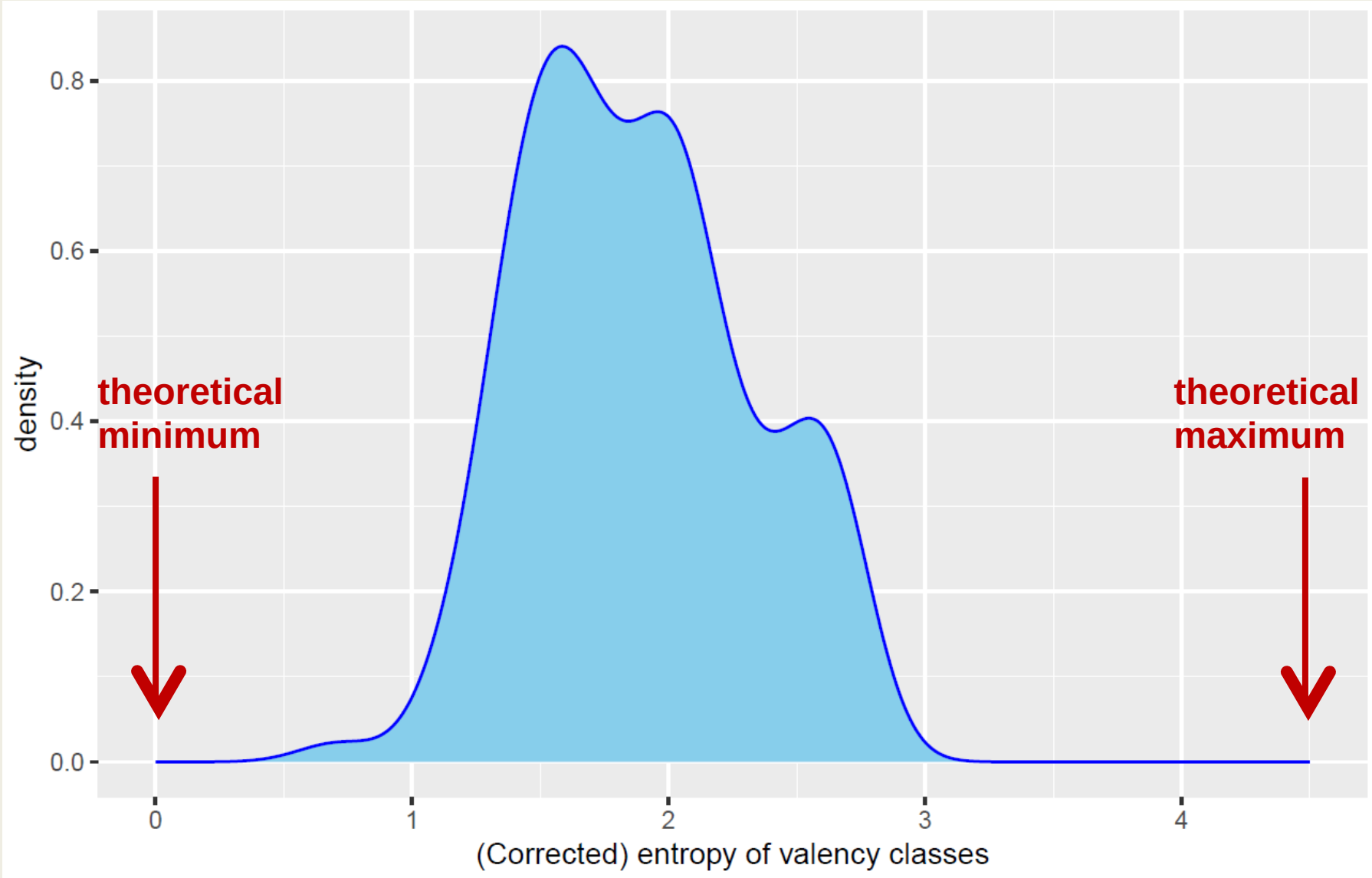
Cross-linguistic variation

- Min = 0.71 (Joola-Fonyi; Atlantic-Congo)
- Max = 2.71 (Khrwarshi; Nakh-Daghestanian)

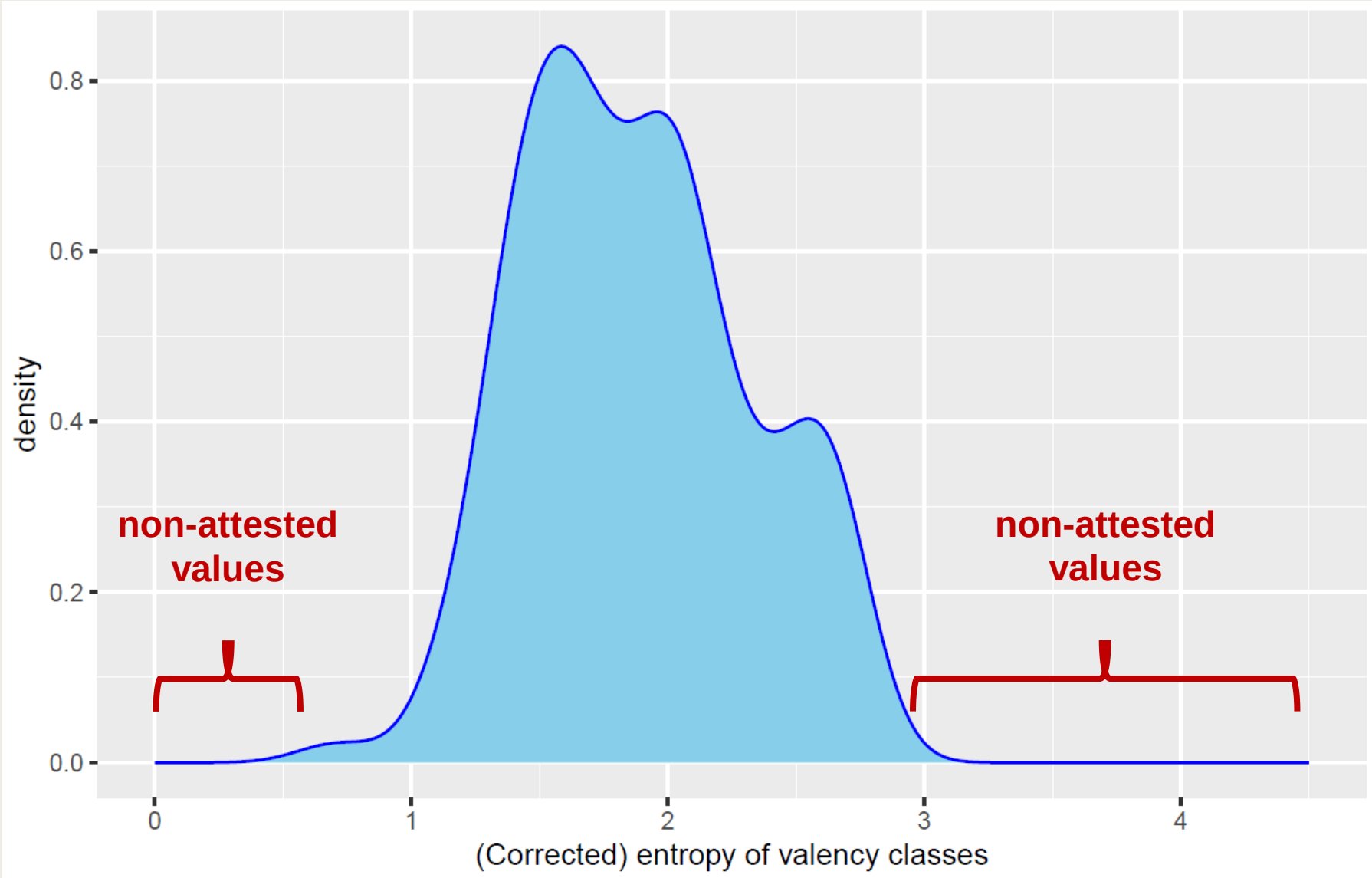
Table 1. Selected verbs and valency patterns in Joola-Fonyi and Khwarshi.

	Joola-Fonyi		Khwarshi	
meaning	verb	pattern	verb	pattern
'be afraid'	*kóli*	TR	*j/uƛ'a*	ABS_CONT
'avoid'	*ŋom*	TR	*j/iča*	ABS_CONT.EL
'wait'	*kob*	TR	*gic'a*	ABS_CONT.LAT
'attack'	*lóúm*	TR	*k'oƛa*	ABS_SUPER
'win, beat'	*ŋoolen*	TR	*j/iža*	ABS_SUPER.EL
'see'	*juk*	TR	*j/akʷa*	DAT_ABS
'touch'	*gor*	TR	*j/etaχa*	ERG_CONT
'bite'	*rum*	TR	*hana*	ERG_GEN1
'be angry'	*leet*	TR	*semi mak'a*	GEN1_CONT.LAT
'eat'	*ri*	TR	*j/ac'a*	TR (ERG_ABS)

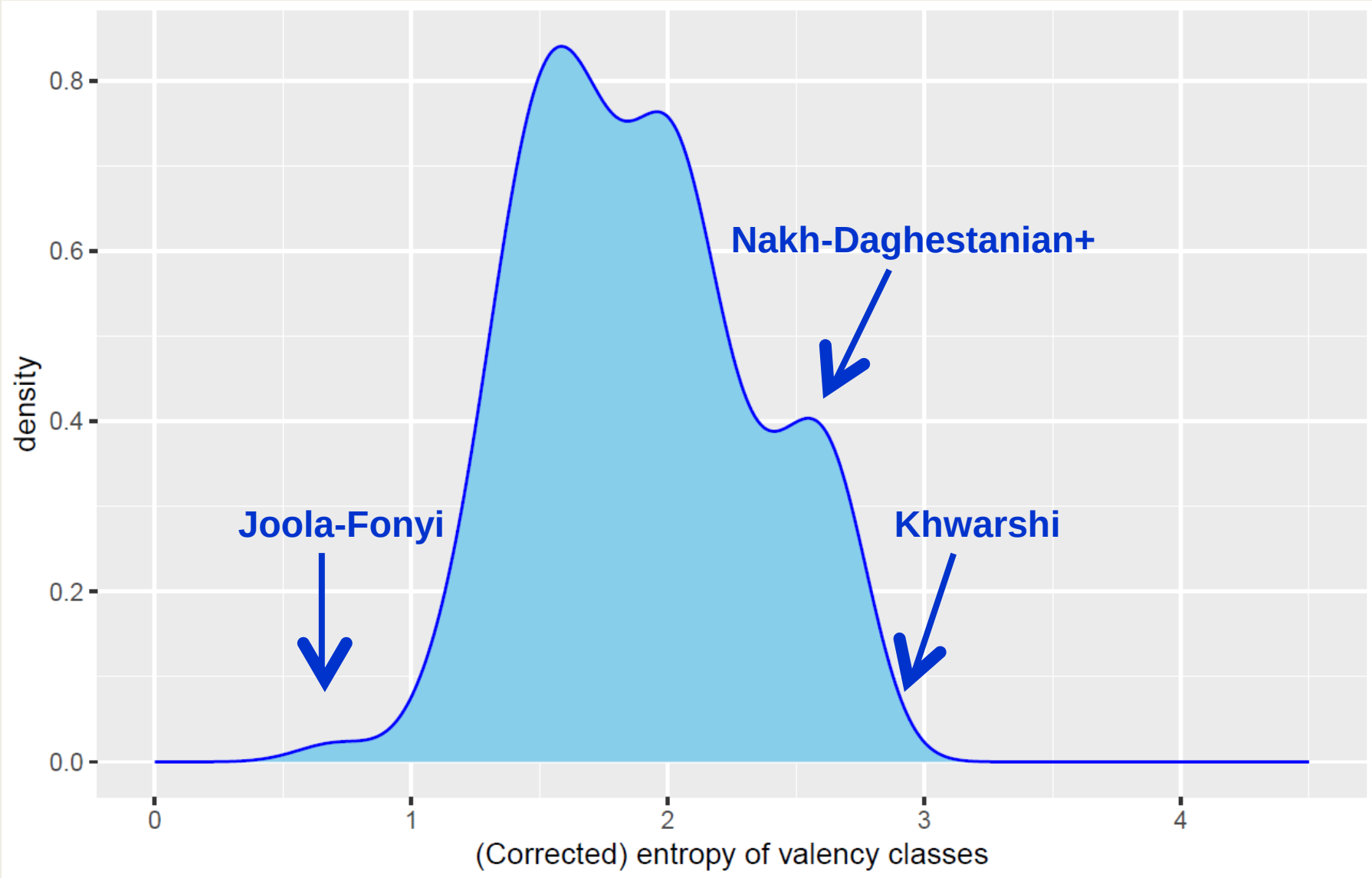
Cross-linguistic variation of entropy



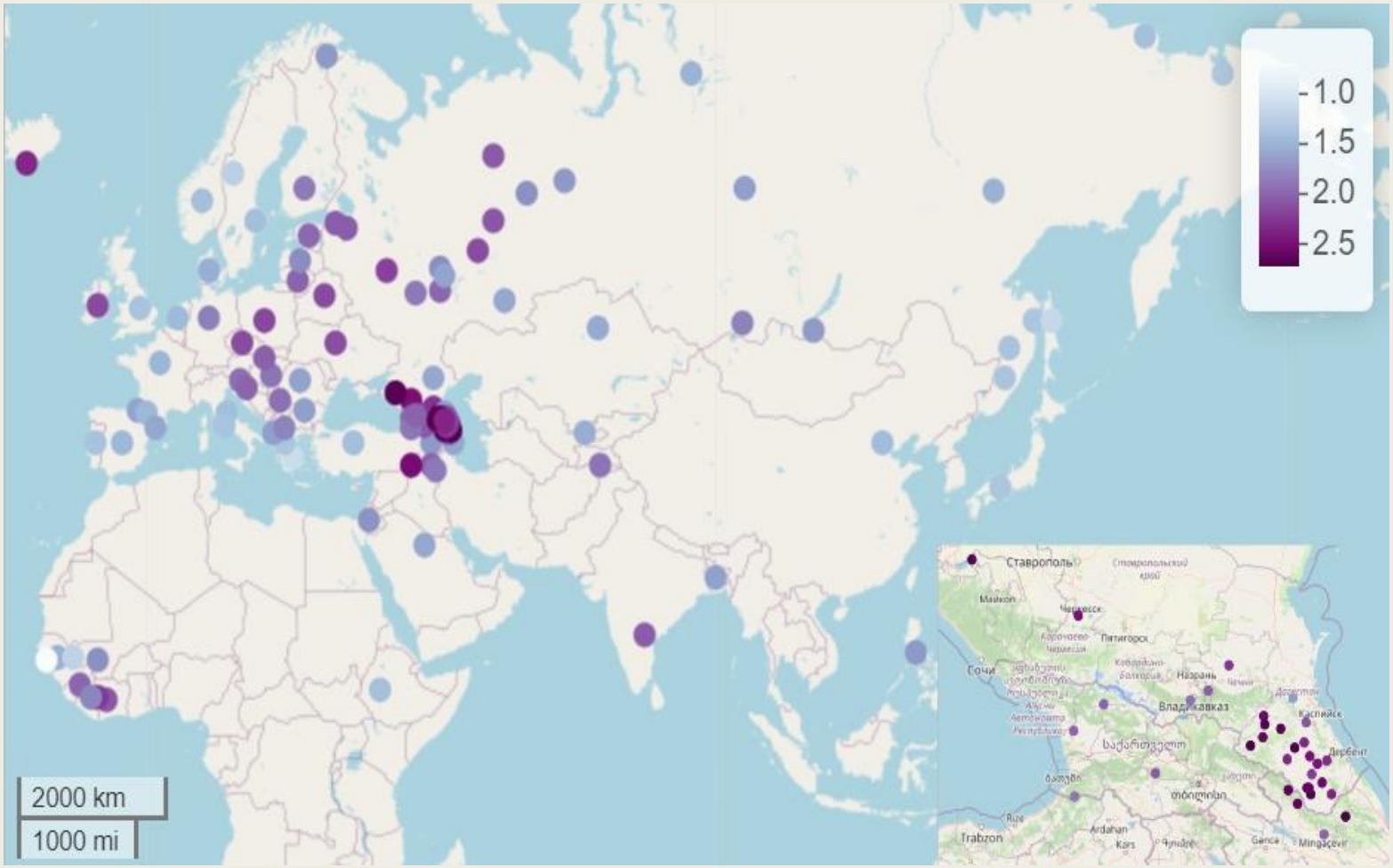
Cross-linguistic variation of entropy



Cross-linguistic variation of entropy



Areal patterns



This map was built using package `lingtypology` (Moroz 2017) in R

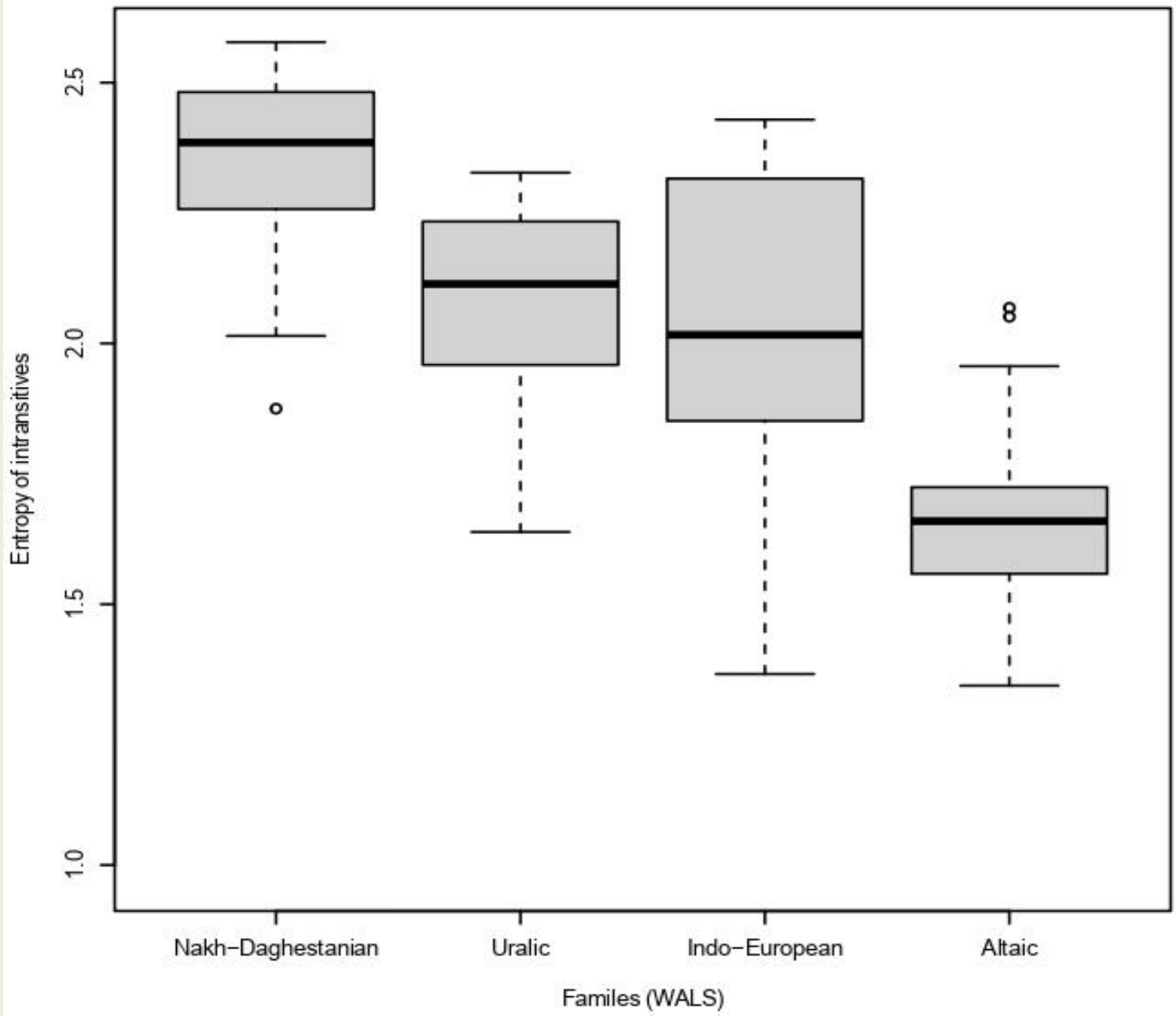
Genealogy

- The entropy of intransitives is largely conditioned by the genealogical factor

- Tentative hierarchy of some Eurasian families:

  Nakh-Daghestanian >
  Uralic, Indo-European >
  Altaic*

"Families" are taken from WALS. I remain agnostic with respect to the genealogical validity of "Altaic"


Entropy of intransitives
1.0
1.5
2.0
2.5
Nakh−Daghestanian
Uralic
Indo−European
Altaic
Familes (WALS)

Structure of the talk

- Background & goals
- Data & methods
- Results: valency class system complexity and…
  - its components
  - its cross-linguistic variation
  - **its typological correlates**
  - word order
- Summary and discussion

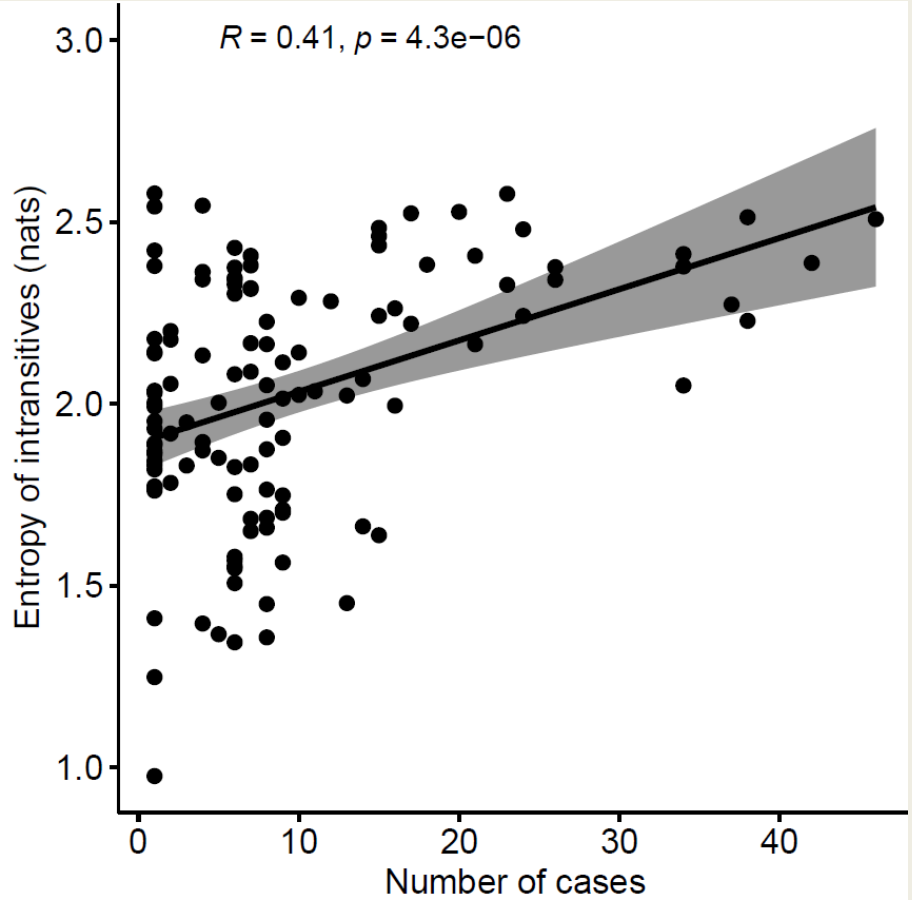
Typological correlates

Hypotheses

- The size of case inventory
- Prevalence of non-verbal predicates
- Satellite- vs. verb-framing
- Colexification patterns

The size of case inventory

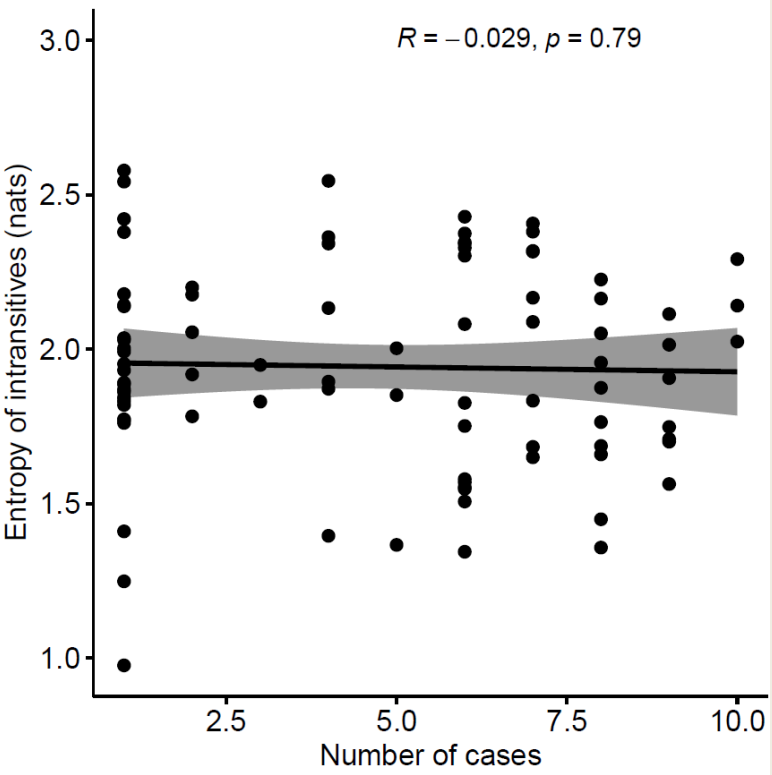
- Is H_{intr} mainly determined by the size of the case inventory? Probably, yes?..



The size of case inventory

- Is H_{intr} mainly determined by the size of the case inventory? Probably, yes… **BUT!**

The correlation does not exist in languages with <11 cases



The size of case inventory

- Is H_{intr} mainly determined by the size of the case inventory? Probably, yes… **BUT!**

The number of cases has very low impact in the linear regression model that also takes transitivity prominence into account and views linguistic family as a random variable

```
Fixed effects:
                                          Estimate Std. Error         df t value Pr(>|t|)
(Intercept)                               2.252490   0.170756  67.358731  13.191  < 2e-16 ***
languages_everything$transitivity.ratio  -0.760989   0.282055 112.177382  -2.698  0.00805 **
languages_everything$number_nominal_cases 0.007032   0.003910 115.873556   1.798  0.07472 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

The size of case inventory

- Is H_{intr} mainly determined by the size of the case inventory?

  => Probably not!

Non-verbal predicates

- Shinaz Rutul (Nakh-Daghestanian)

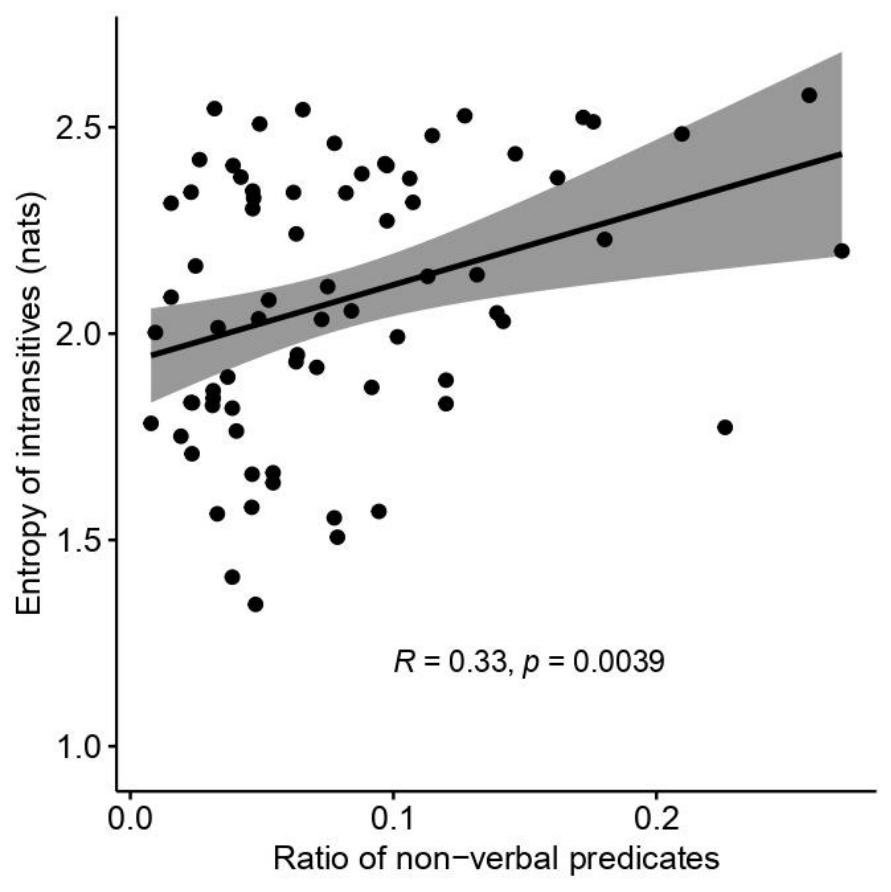
Basir-ɬis *k'vač'* *ij* *Karam*
PN-DAT hated COP PN
'Basir hates Karam.'

- Disclaimer: the distinction between verbal and non-verbal predicates has been annotated for only 75 "published" languages

- Non-verbal predicates are almost never transitive (Dixon 2004: 5; but see Lowe 2017) => strong correlation with the overall entropy and transitivity prominence

Non-verbal predicates

- Less trivially, the prevalence of non-verbal predicates also displays a positive correlation with H_{intr}



Satellite- vs. verb-framing

- **Canonical examples (Talmy 2000, etc.)**

  - Satellite-framed

English (Beavers et al. 2010: 333)

John limped into the house.

  - Verb-framed

French (Beavers et al. 2010: 333)

Je suis entré dans la maison en boitant.

Literally, 'I entered the house limping.'

Satellite- vs. verb-framing

Entropy of intransitive verbs (H_{intr}) in presumed satellite- vs. verb-framed languages

Satellite-framed	H_{intr}	Verb-framed	H_{intr}
English	1.92	French	1.83
German	2.34	Spanish	1.84
other Germanic	1.89-2.36	other Romance	1.41-2.18
Russian	2.34	Turkish	1.34
other Slavic	1.89-2.43	Japanese	1.36
Mandarin Chinese	2.38	Hebrew	1.86

- Languages presumed to be “satellite-framed” display greater values of H_{intr}, the entropy of intransitive verbs
- This can be related to the polysemy/grammaticalization patterns associated with spatial expressions

Colexification patterns

- Many dateset entries are headed by identical lexical items used in non-identical patterns

- Eastern Maninka (Mande)

a. *Sékù bára bɔ́ à lá só` lá*
PN PRF exit 3SG POSS village\ART at
'Seku left his village.': SBJ_lá pattern

b. *Sékù bára à lá dìriki` bɔ́*
PN PRF 3SG POSS shirt\ART exit
'Seku has taken off his shirt.': transitive pattern

Colexification patterns

- The correlation between the incidence of colexification patterns and H in the whole sample s not statistically significant (R = 0.11, *p* = 0.35)
- Arguably, this covariance exists only in morphologically poor languages

The prevalence of colexified patterns and entropy of valency classes in 5 languages from Africa

language	family (WALS)	% colexifications	entropy (H)
Eastern Maninka	Mande	0.15	1.98
Mano	Mande	0.14	1.79
Bambara	Mande	0.10	1.74
West Central Oromo	Afro-Asiatic	0.07	1.41
Khoekhoe	Khoe-Kwadi	0.00	1.15

Colexification patterns

- The correlation between the incidence of colexification patterns and H in the whole sample s not statistically significant (R = 0.11, *p* = 0.35)
- Arguably, this covariance exists only in morphologically poor languages

=> Possible explanation: complexity trade-off scenario; less information in the verb itself (colexifications) is more tolerable if it is compensated for by morphosyntactic complexity

Typolofical correlates: summary

Possible typological correlates of the complexity of valency class systems

	high entropy	**low entropy**
the size of case inventory	high	low
prevalence of non-verbal predicates	high	low
prevalence of preverbs	high	low
verb-framing	satellite-framed	verb-framed
prevalence of verb colexifications	high	low
and a possible generalization…		
language type* ()	"nouny"	"verby"

* See (Polinsky 2012; Qiu and Winsler 2017)

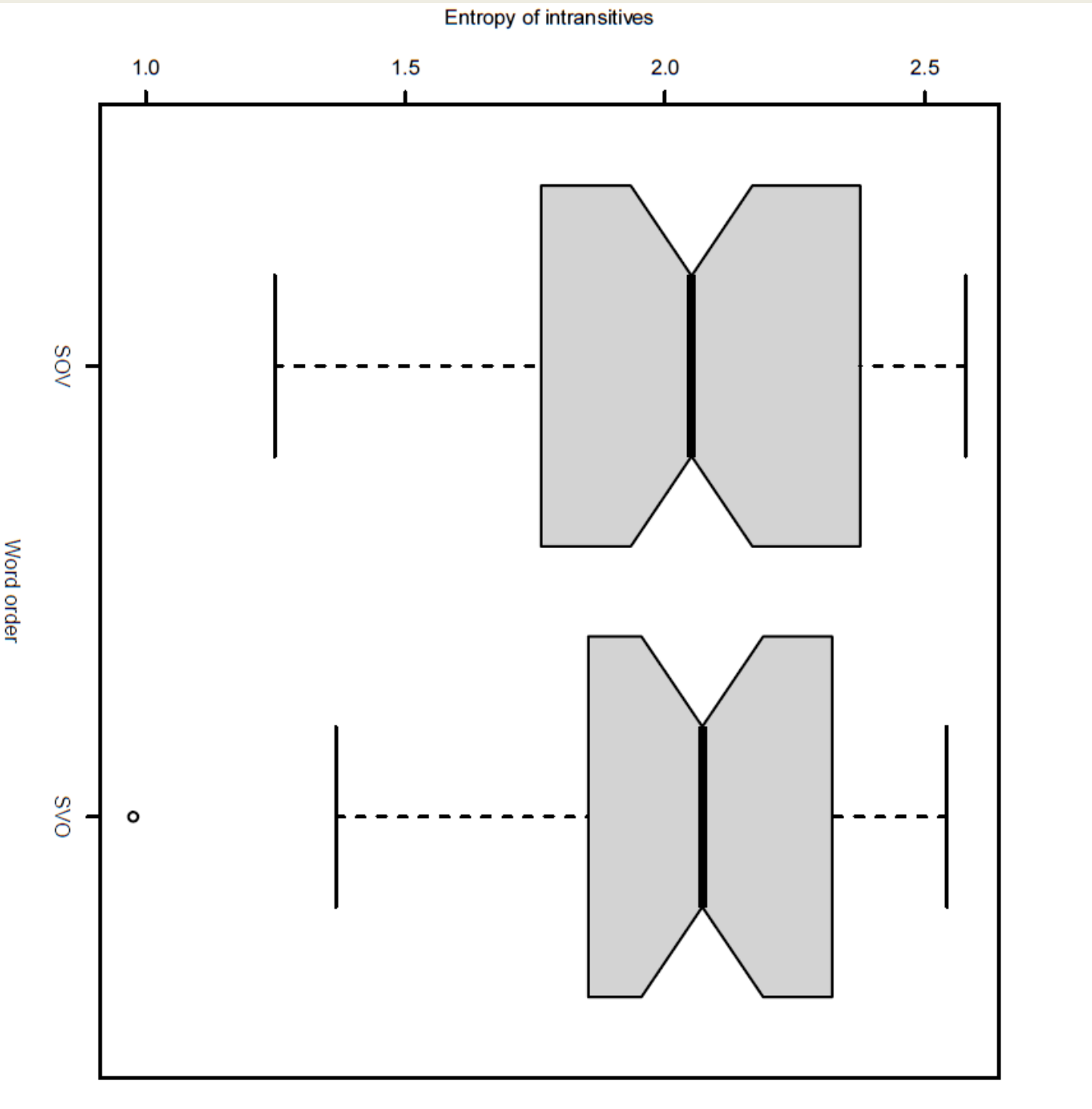
Structure of the talk

- Background & goals
- Data & methods
- Results: valency class system complexity and…
  - its components
  - its cross-linguistic variation
  - its typological correlates
  - **word order**
- Summary and discussion

SOV vs. SVO

- A well-established idea: SOV favor case contrasts (Greenberg 1966: 96; Maion 2018; Sinnemäki 2010)

- But overall, there is no robust difference between SOV and SVO languages in terms of the observed H (and H_{intr}) values


Entropy of intransitives
1.0
1.5
2.0
2.5
Word order
SOV
SVO

Complexity in postverbal positions

- However, more complex valency encoding is associated with postverbal positions

- Five pieces of evidence

Complexity in postverbal positions (1)

- Universal prevalence of preverbal subjects (Tomlin 1986, Dryer 2013) with maximal role neutralization (Kibrik 1997, Van Valin & LaPolla 1997: 251ff.)

Complexity in postverbal positions (2)

- Clause-fronted topics without overt encoding of their thematic relation to the verb (Lambrecht 2001: 1069-1070), as opposed to antitopics

Occitan (ibid.)

a. *Lo cinema, i vau sovent*
the cinema there I.go often
'The movies, I goe there often'

b. *I vau sovent,* ***al*** *cinema*
there I.go often to.the cinema
'I go there often, to the movies'

Complexity in postverbal positions (3)

- Non-canonical A's (typically preverbal) are usually less variegated than non-canonical O's (often postverbal) (Bickel et al. 2014: 496-500; Say 2018: 565-566)

Complexity in postverbal positions (4)

- (S)OVX patterns are significantly more widespread than (S)XVO patterns (Dryer and Gensler 2013; Hawkins 2008: 170)

Bambara (< Mande; Vydrin 2023)

a. *Sékù ye nàmasà dún*
PN PFV.TR banana eat
'Seku ate a banana.'

b. *Sékù b'í túlomajɔ̀ àrajɔ̀ fɛ̀*
PN IPFV.REFL listen radio\ART by
'Seku is listening to the radio.'

Complexity in postverbal positions (5)

- Preliminary token-based evidence

Entropy (H) of flagging patterns in pre- and postverbal positions: data from a spoken corpus of a North-Eastern Neo-Aramaic dialect of Urmiya (Ovsjannikova & Say 2023)

	preverbal	postverbal
all arguments	0.64	1.51
non-subjects only (O & E)	1.24	1.70

Structure of the talk

- Background & goals
- Data & methods
- Results: valency class system complexity and…
  - its components
  - its cross-linguistic variation
  - its typological correlates
  - word order
- **Summary and discussion**

Summary and discussion

- Valency class system complexity
  - can be captured in terms of entropy
  - displays high cross-linguistic variation
  - is a robust typological feature: the more intransitive verbs, the more distinctions in them
  - is a largely neglected area in typology
  - is diachronically stable

Summary and discussion

- Languages favor mid-range value in terms of their valency class system complexity

- Overly complex systems are avoided…
  - because of high processing costs
  - and learnability issues
- Overly simple systems are avoided…
  - because they would lack semantic cues
  - and would require excessive guesswork on the hearer's part

Summary and discussion

- Possible correlates of high valency class system complexity
  - large case inventory
  - high prevalence of non-verbal predicates
  - satellite-framed patterns
  - verbal colexifications

  $\Rightarrow$ “nouny” properties.
- This aligns with the idea that peripheral arguments favor dependent marking (Nichols 1986)

Summary and discussion

- Complex argument encoding systems are associated with postverbal positions
  - A possible explanation: avoiding severe forms of “looking ahead”, see also the Maximize Online Processing principle in (Hawkins 2014: 28ff.)
  - This preference arguably outranks other components of efficiency in the domain of choosing between argument encoding patterns, see also Seržant & Moroz (2022)
  - Convenient for the speaker, somewhat redundant for the hearer

Broader generalization

- (The limits on) typological distributions are largely shaped by efficiency-related processing constraints

THANK YOU!

References

- Ackerman, Farrell & Robert Malouf. 2013. Morphological organization: The low conditional entropy conjecture. *Language* 89(3). 429–464.
- Beavers, John, Beth Levin & Shiao Wei Tham. 2010. The typology of motion expressions revisited. *Journal of Linguistics* 46. 331–377.
- Bickel, Balthasar, Taras Zakharko, Lennart Bierkandt & Alena Witzlack-Makarevich, 2014. Semantic role clustering: An empirical assessment of semantic role types in non-default case assignment. Studies in language, 38 (3). Advances in research in semantic roles. 485–511.
- Comrie, Bernard. 1989. Language universals and linguistic typology. 2nd edition. Chicago: CUP.
- Creissels, Denis. 2018. Transitivity prominence in typological perspective: The case of Basque. Anuario del Seminario de Filología Vasca Julio de Urquijo 52(1–2). 175–187.
- Dahl, Östen. 2004. *The growth and maintenance of linguistic complexity*. Amsterdam: John Benjamins.

References

- Dryer, Matthew S. 2013. Order of Subject, Object and Verb. In Dryer, Matthew S.; Haspelmath, Martin (eds.). The World Atlas of Language Structures Online. Leipzig: Max Planck Institute for Evolutionary Anthropology.
- Dowty, David. 1991. Thematic proto-roles and argument selection. *Language* 67(3). 547–619.
- Drossard, Werner. 1991. Transitivitat (vs. Intransitivitat) und Transitivierung (vs. Intransitivierung) unter typologischem Aspekt. In Hansjakob Seiler & Waldfried Premper (eds.), *Partizipation: Das sprachliche Erfassen von Sachverhalten*, 408–445. Tübingen: Narr.
- Fenk-Oczlon, Gertraud & August Fenk. 2008. Complexity trade-offs between the subsystems of language. In Matti Miestamo, Kaius Sinnemaki & Fred Karlsson (eds.), *Language complexity: Typology, contact, change*, 43–65. Amsterdam & Philadelphia: John Benjamins.

References

- Fenk-Oczlon, Gertraud & August Fenk. 2014. Complexity trade-offs do not prove the equal complexity hypothesis. *Poznan Studies in Contemporary Linguistics* 50(2). 145–155.
- Greenberg, Joseph H. 1966. Some universals of grammar with particular reference to the order of meaningful elements. In Joseph H. Greenberg, *Universals of language* (2nd edn), 73–113. Cambridge, MA: MIT Press.
- Haspelmath, Martin. 2015. Transitivity prominence. In: Malchukov, Andrej & Comrie, Bernard (eds.). 2015. Valency classes in the world's languages. Vol. 1-2. Berlin, Boston: De Gruyter Mouton. Proch. P. 131–147.
- Haspelmath, Martin & Iren Hartmann. 2015. Comparing verbal valency across languages. In In Andrej Malchukov & Bernard Comrie, (eds.). *Valency classes in the world's languages. Vol. 1. Introducing the framework, and case studies from Africa and Eurasia*, 41–71. Berlin & Boston: De Gruyter Mouton.
- Hawkins, John A. 1986. *A comparative typology of English and German: Unifying the contrasts*. Austin: University of Texas Press.

References

- Hawkins, John A. 2008. An asymmetry between VO and OV languages: The ordering of obliques. In Greville G. Corbett & Michael Noonan (eds.), *Case and grammatical relations: Studies in honor of Bernard Comrie*, 167–190. Amsterdam: Benjamins.
- Hawkins, John A. 2014. Cross-linguistic variation and efficiency. Oxford: Oxford University Press.
- Juola, Patrick. 1998. Measuring linguistic complexity: the morphological tier. *Journal of Quantitative Linguistics* 5. 206–213.
- Kibrik, Aleksandr E. 1997. Beyond subject and object: toward a comprehensive relational typology. Linguistic typology 1.3: 279-346.
- Lambrecht, K. 2001. Dislocation. In M. Haspelmath, E. Konig, W. Oesterreicher, and W. Raible (Eds.) Typology and Language Universals. Berlin: Walter de Gruyter. pp. 1050–1078.
- Lazard, Gilbert. 1994. *L'actance*. Paris: Presses Universitaire de France.
- Levshina, Natalia. 2019. Token-based typology and word order entropy: A study based on Universal Dependencies. *Linguistic typology* 23(3). 533–572.

References

- Lowe, John J. 2017. *Transitive nouns and adjectives. Evidence from Early Indo-Aryan*. Oxford: Oxford University Press.
- Maion, Fabio. 2018. Do certain word orders attract case marking? A typological survey on the dependency of syntax and morphology. Tübingen: University of Tübingen BA thesis.
- Malchukov, Andrej and Leipzig Valency Classes Project team. 2015. Leipzig Questionnaire on valency classes. In: Malchukov, A. & B. Comrie (eds.). Valency classes in the world's languages. Vol. I. Berlin: Mouton de Gruyter. 27-40.
- Miestamo, Matti, Kaius Sinnemaki & Fred Karlsson (eds.). 2008. *Language complexity: Typology, contact, change*. Amsterdam & Philadelphia: John Benjamins
- Moroz, G. 2017. lingtypology: easy mapping for Linguistic Typology. https://CRAN.R-project.org/package=lingtypology.
- Müller-Gotama, Franz. 1994. *Grammatical relations: A cross-linguistic perspective on their syntax and semantics*. Berlin: Mouton de Gruyter.
- Næss, Å. 2007: *Prototypical Transitivity*. Amsterdam, Philadelphia: John Benjamins.

References

- Nichols, Johanna. 1986. Head-marking and dependent-marking grammar. *Language* 62(1). 56–119.
- Ovsjannikova, Maria & Sergey Say. 2023. Valency-encoding devices in a spoken Northeastern Neo-Aramaic corpus. Paper presented at the Grammatical Relations in Spoken Corpora conference. Paris, INALCO, 15-16 June 2023.
- Polinsky, Maria. 2012. Headedness, again. In Thomas Graf, Denis Paperno, Anna Szabolcsi & Jos Tellings (eds.), *Theories of everything. In honor of Ed Keenan*, 348–359. Los Angeles: UCLA Department of Linguistics.
- Qiu, Chen & Adam Winsler. 2017. Language use in a 'one parent–one language' Mandarin–English bilingual family: Noun versus verb use and language mixing compared to maternal perception. *International journal of bilingual education and bilingualism,* 20(3). 272–291.
- Say, Sergey. 2014. Bivalent verb classes in the languages of Europe: A quantitative typological study. In *Language dynamics and change* 4(1). 116–166.

References

- Seržant, Ilja A. 2019. Weak universal forces: The discriminatory function of case in differential object marking systems. In Karsten Schmidtke-Bode, et al. (eds.), Explanation in typology: Diachronic sources, functional motivations and the nature of the evidence, 149–178. Berlin: Language Science Press.
- Seržant, I. A. & Moroz, G. 2022. Universal attractors in language evolution provide evidence for the kinds of efficiency pressures involved, Humanities & Social Sciences Communications 9. Article 58.
- Shcherbakova, Olena, Volker Gast, Damián E. Blasi, Hedvig Skirgård, Russell D. Gray & Simon J. Greenhill. 2022. A quantitative global test of the complexity trade-off hypothesis: the case of nominal and verbal grammatical marking. *Linguistics Vanguard*, 9(s1), 155–167. doi:10.1515/lingvan-2021-0011 (accessed 11 May 2024).
- Shosted, Ryan K. 2006. Correlating complexity: A typological approach. *Linguistic typology* 10(1). 1–40.
- Sinnemäki, Kaius. 2008. Complexity trade-offs in core argument marking. In Matti Miestamo, Kaius Sinnemaki & Fred Karlsson (eds.), *Language complexity: Typology, contact, change*, 67–88. Amsterdam & Philadelphia: John Benjamins.

References

- Sinnemäki, Kaius. 2009. Complexity in core argument marking and population size. In Geoffrey Sampson, David Gil & Peter Trudgill (eds.), *Language complexity as an evolving variable*, 126–140. Oxford: Oxford University Press.
- Sinnemäki, Kaius. 2010. Word order in zero-marking languages. *Studies in language* 34(4), 869–912.
- Sinnemäki, Kaius. 2014. Global optimization and complexity trade-offs. *Poznan studies in contemporary linguistics* 50(2). 179–195.
- Talmy, Leonard. 2000. *Toward a cognitive semantics. Vol. 2. Typology and process in concept structuring*. Cambridge, MA: MIT Press.
- Tomlin, R.S. 1986. *Basic Word Order: Functional Principles.* London: Routledge.
- Tsunoda, Tasaku. 1985. Remarks on transitivity. *Journal of Linguistics* 21. 385–396.
- Van Valin, Robert D. Jr. & Randy LaPolla. 1997. Syntax. Structure, meaning & function. Cambridge: CUP.