Efficiency in grammar: Patterns and explanations
University of Freiburg
July 5, 2023

Cross-linguistic variability in complexity of valency class systems: implications for efficiency

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

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: introducing entropy
- Results
  - cross-linguistic variation of entropy
  - entropy, case and transitivity prominence
  - entropy & genealogical signal
  - entropy and word order
- Implications for efficiency

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-ṭ*
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; 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

Data and methods

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

Data and methods

- 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 and methods

#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 and methods

- The sample: currently 99 languages, mainly spoken in Northern Eurasia

Data and methods

- 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 and methods

- Some patterns in Rutul

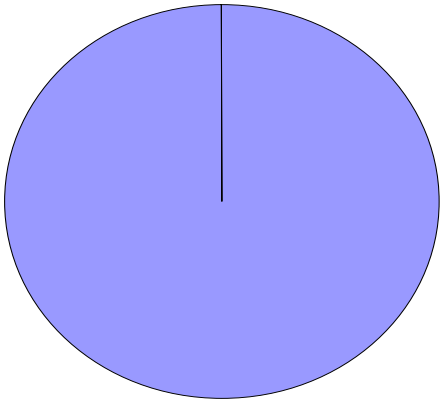
31	call	rasul-a ramazan-ɨ-χda ses ha<w>ɨ-r PN-ERG PN-OBL-SUB sound <3>do.PFV-CVB 'Rasul called Ramazan.'	ERG_SUB
32	get_to_know	rasul ramazan-ɨ-kan sa sen xura gi-na taniš jiši-d PN(NOM) PN-OBL-COM one year in front SUB.be-CVB acquainted 1.become.PFV-ATTR 'Rasul got to know Ramazan a year ago.'	NOM_COM
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-rɨˤχɨˤ-r=a PN-ERG saz(NOM) 3-hit.IPFV-CVB=be 'Rasul is playing the saz.'	TR
35	avoid	rasul ramazan-ɨ-la: gʲa-ʁaˤbʁu-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 laca-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

Data and methods

- Capturing complexity of valency class system: Shannon's entropy (in nats)

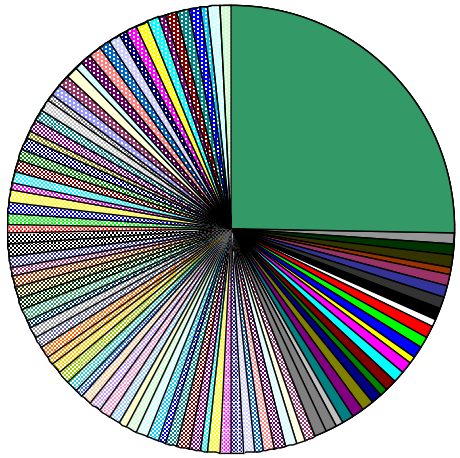
Data and methods

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



Hypothetical Language 1: All verbs belong to the same class

$$H = 0$$

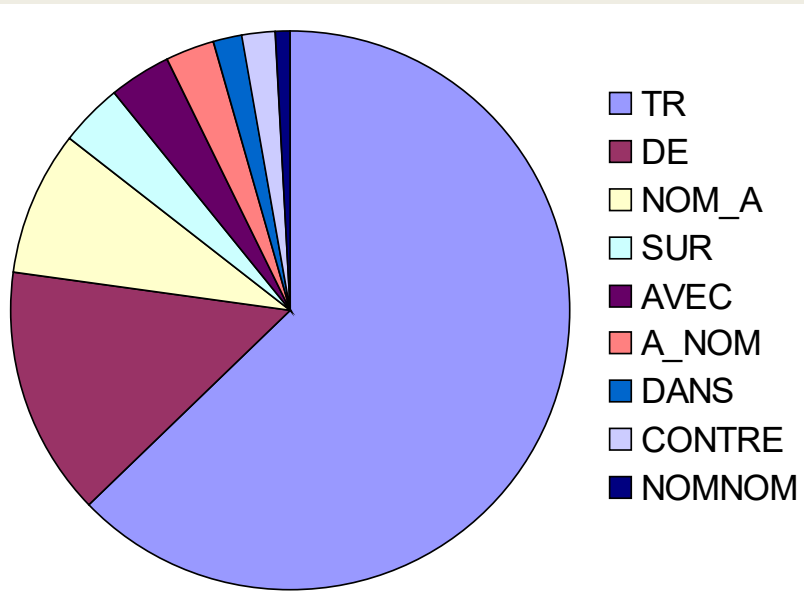


Hypothetical Language 2: 130 verb classes

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

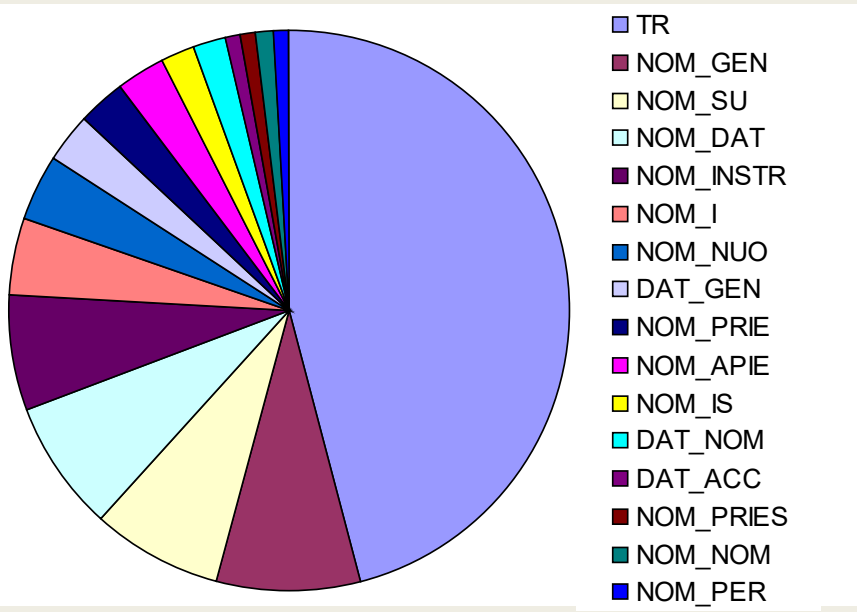
Data and methods

French



H ≈ 1.31

Lithuanian



H ≈ 2.02

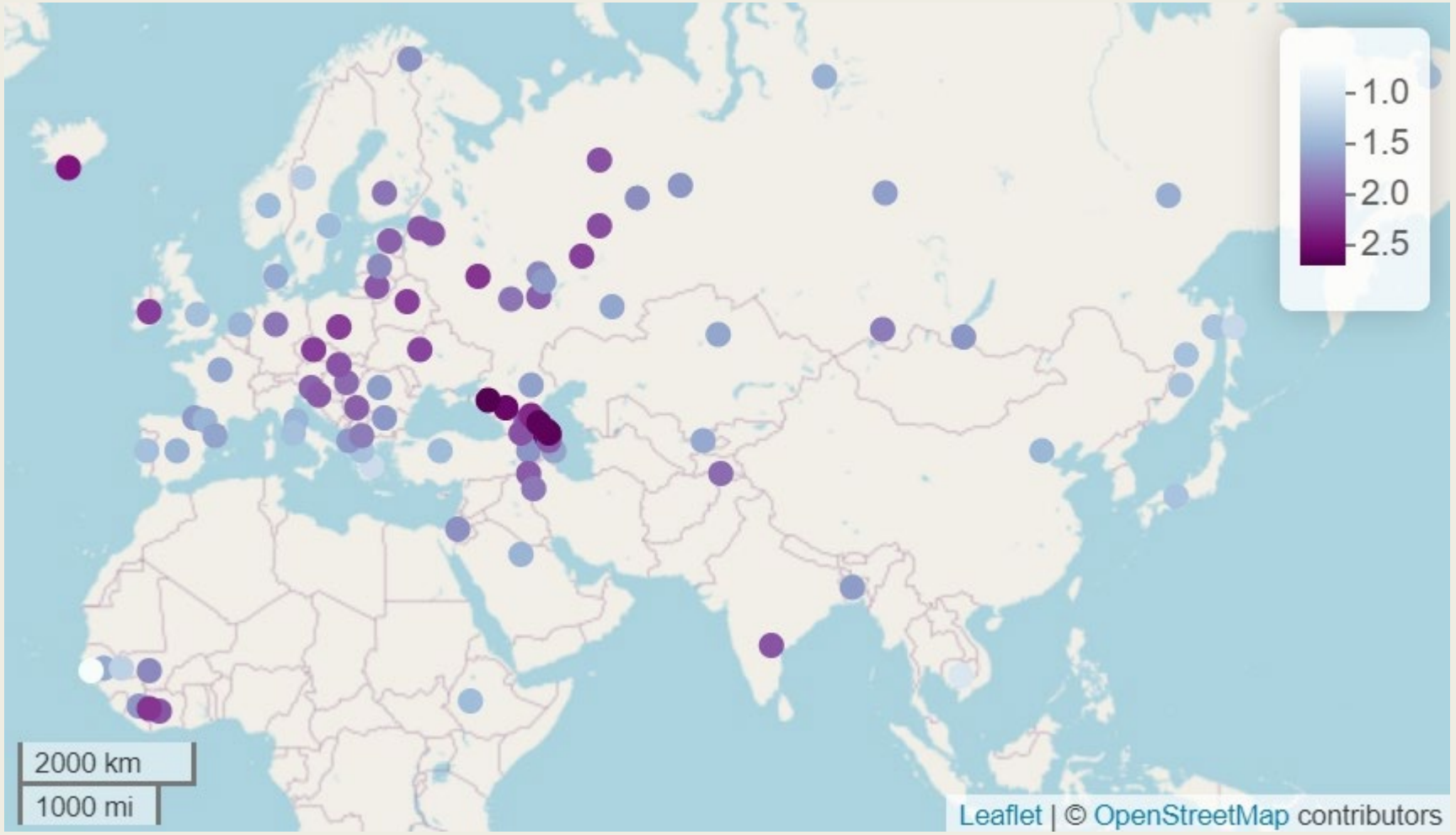
Results

Cross-linguistic variation of entropy

- The values of H vary in the range between 0.72 (Joola-Fonyi < North-Central Atlantic < Atlantic-Congo) and 2.69 (Rutul < Nakh-Daghestanian)

		Joola-Fonyi	Rutul
31	call	TR	ERG_SUB
32	get_to_know	TR	NOM_COM
33	know	TR	DAT_NOM
34	play#instrument#	TR	TR
35	avoid	TR	NOM_SUP.EL
36	make	TR	TR
37	make_fun	TR	ERG_SUP
38	have	TR	SUB_NOM

Cross-linguistic variation of entropy



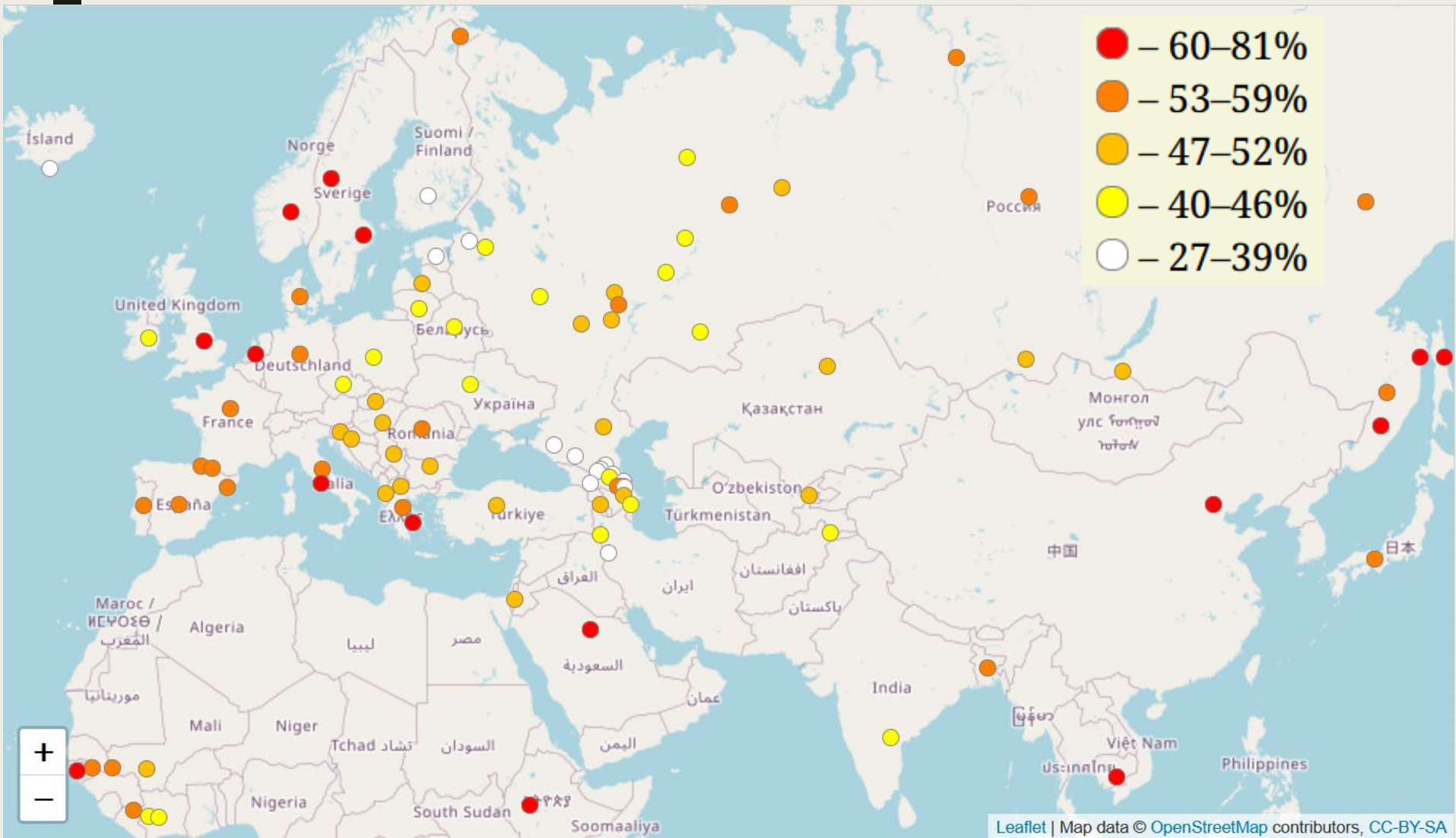
This map was built using package `lingtypology` (Moroz 2017) in R

Entropy, case & 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 also displays robust areal patterning

Transitivity prominence

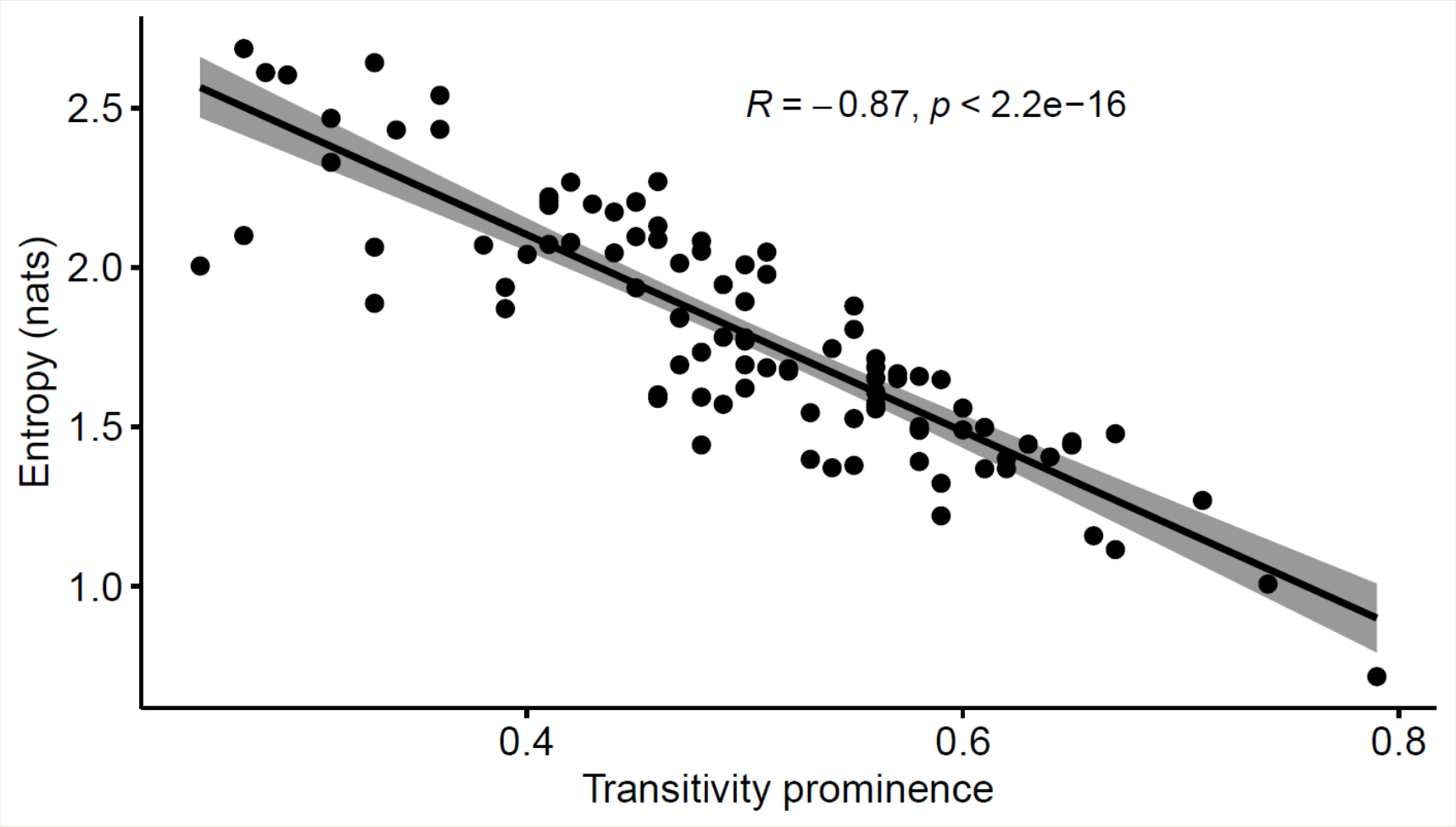


Entropy, case & transitivity prominence

- 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

Entropy, case & transitivity prominence



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

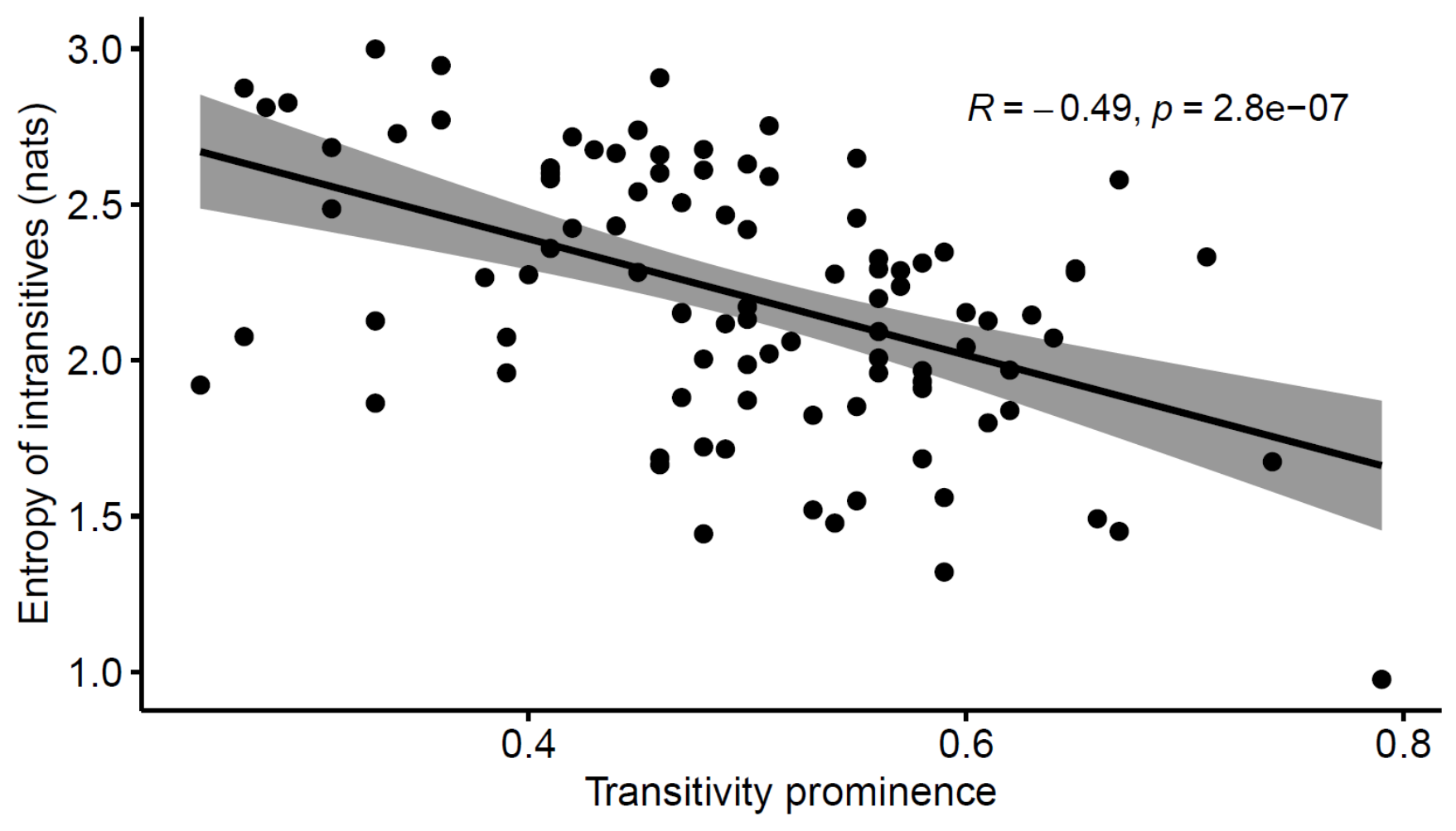
Entropy, case & transitivity prominence

- What about the verbs that are not transitive?

$\Rightarrow$ Calculate H_{intr} = the entropy of intransitive valency classes

  - The rationale: to capture the degree of complexity of intransitive valency classes

Entropy, case & transitivity prominence



Entropy, case & transitivity prominence

- 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.

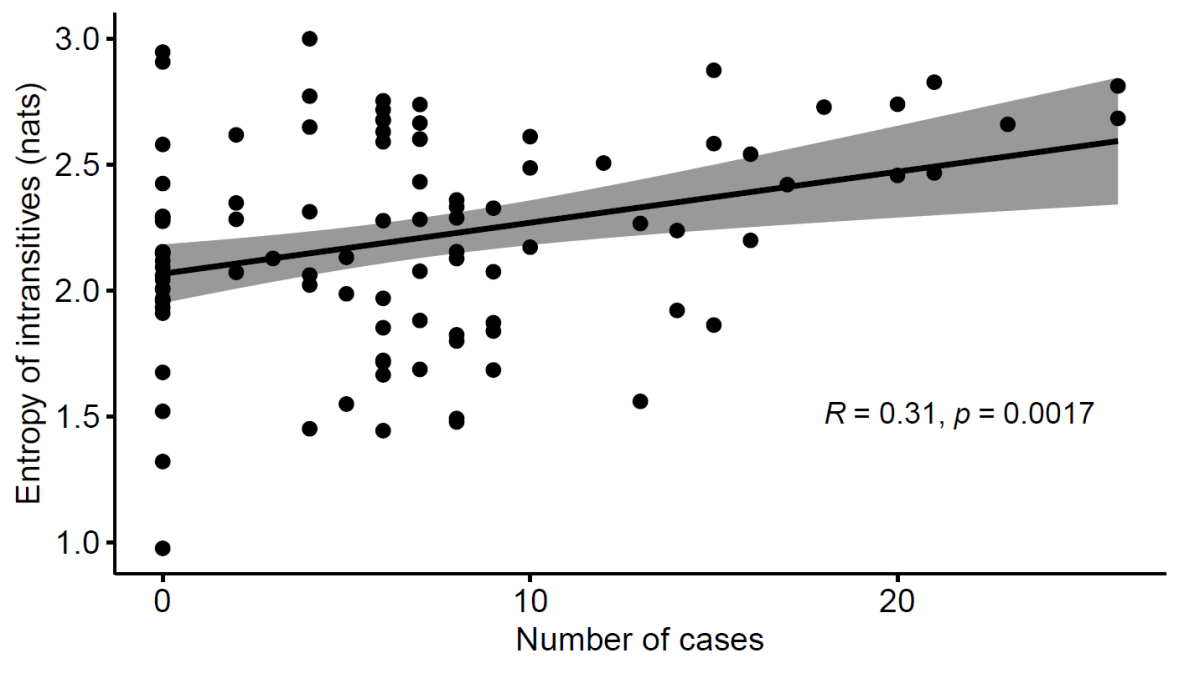
Entropy, case & transitivity prominence

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

Entropy, case & transitivity prominence

- Is H_{intr} mainly determined by the size of the case inventory? No!

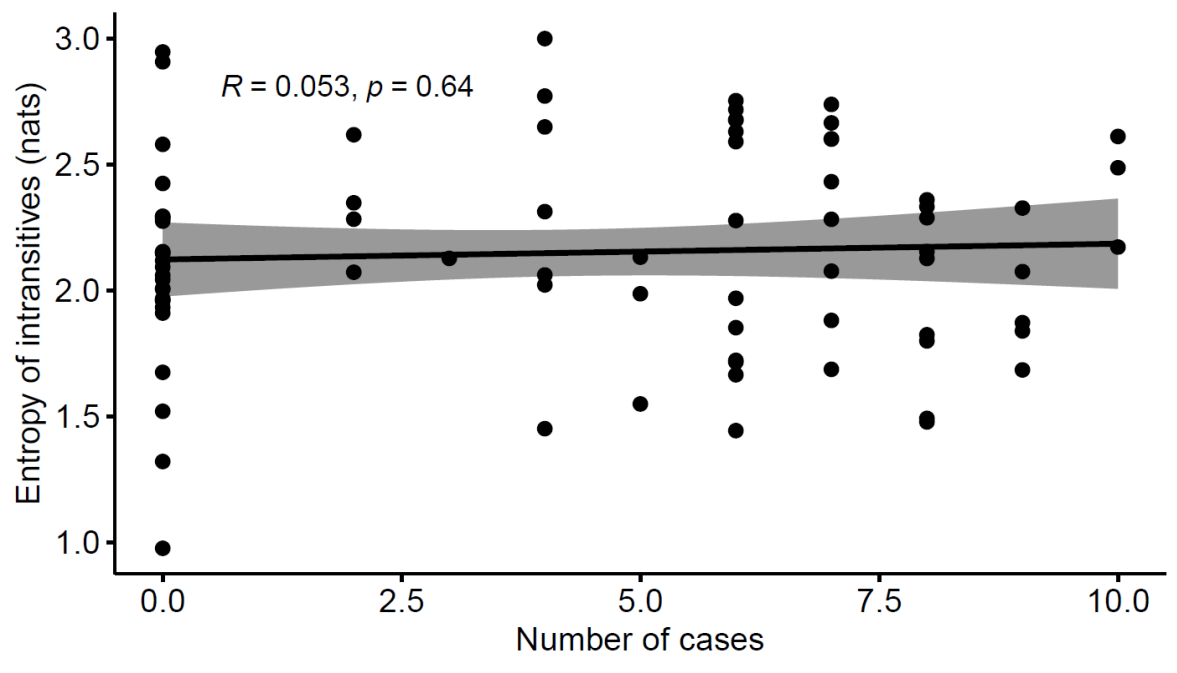
  <= 1) The correlation is not very strong…



Entropy, case & transitivity prominence

- Is H_{intr} mainly determined by the size of the case inventory? No!

  <= 2) The correlation does not exist in languages with <11 cases



Entropy, case & transitivity prominence

- Is H_{intr} mainly determined by the size of the case inventory? No!

  <= 3) The number of cases has very low impact in the linear regression model that also takes transitivity prominence into account

```
Coefficients:
                                Estimate Std. Error t value Pr(>|t|)
(Intercept)                     2.980651   0.212420  14.032  < 2e-16 ***
languages$transitivity.ratio   -1.663381   0.373800  -4.450 2.31e-05 ***
languages$number_nominal_cases  0.007952   0.006370   1.248    0.215
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Entropy & genealogical signal

- Is the entropy of intransitives conditioned by the genealogical factor?

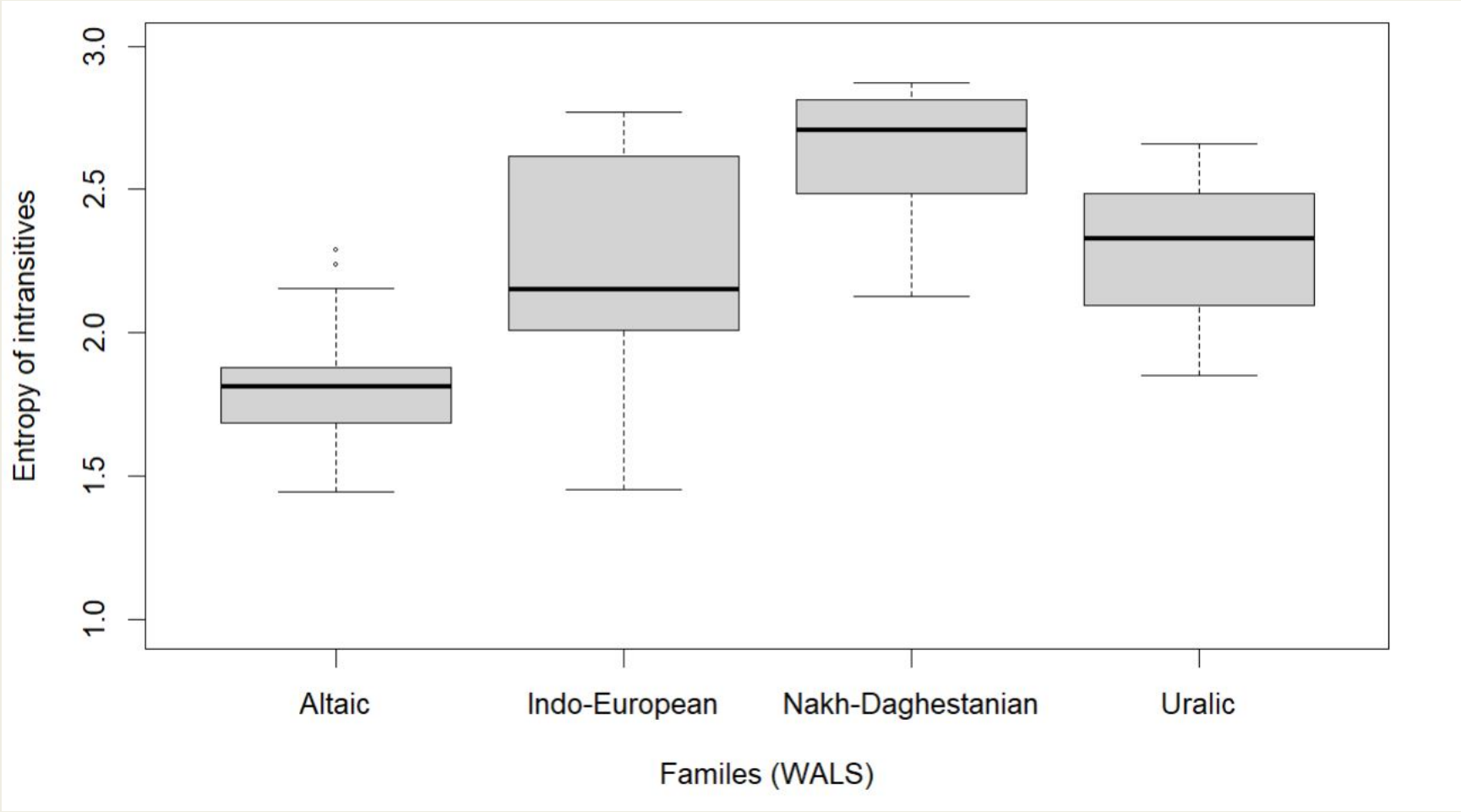
- Yes! Tentative hierarchy of some Eurasian families:

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

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

Entropy & genealogical signal

- Is the entropy of intransitives conditioned

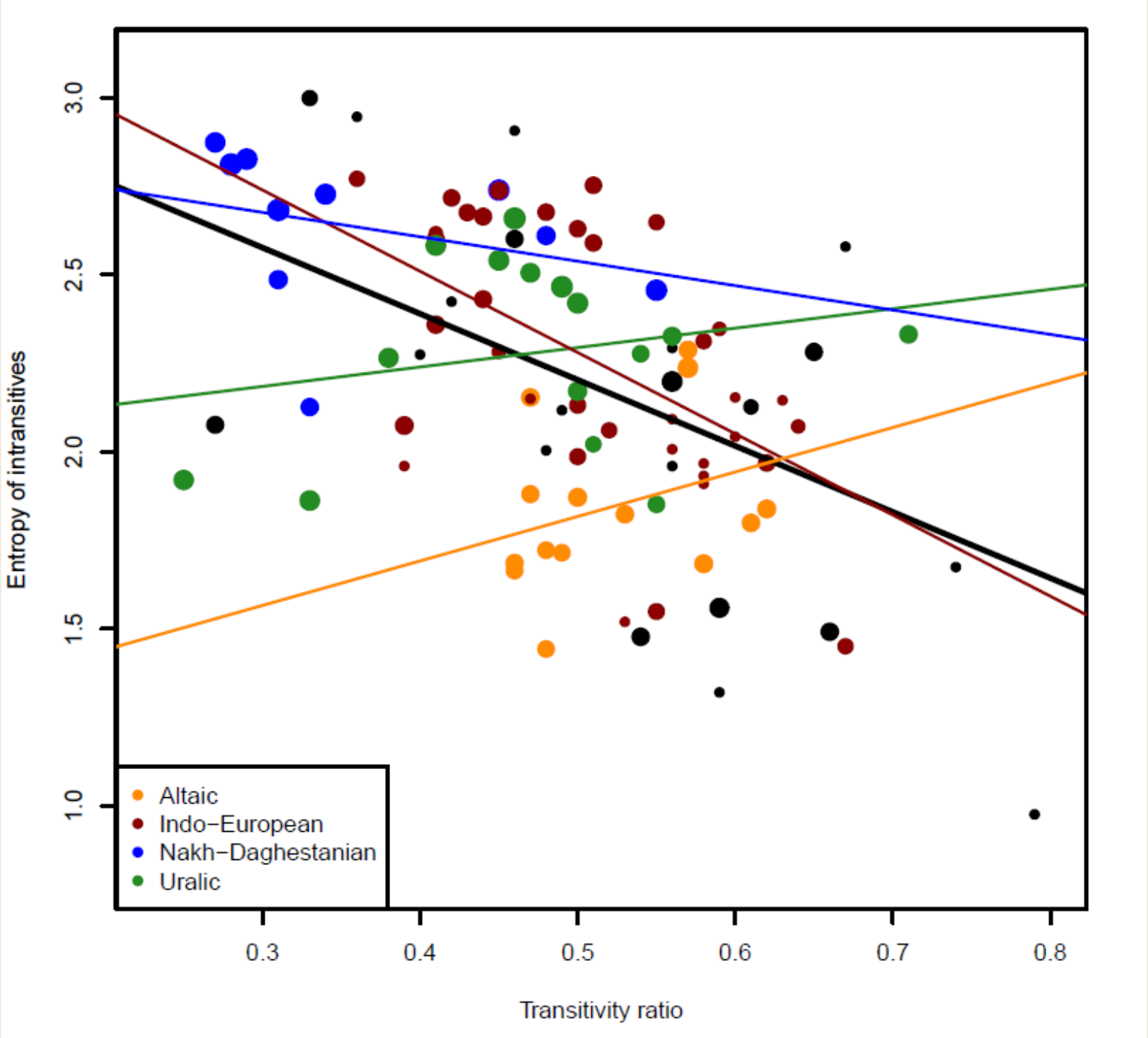


Entropy & genealogical signal

- Is the entropy of intransitives conditioned by the genealogical factors?

- Yes! The (negative) correlation between transitivity prominence and the entropy is found in e.g. Indo-European and Nakh-Daghestanian but not in Uralic and Altaic*

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

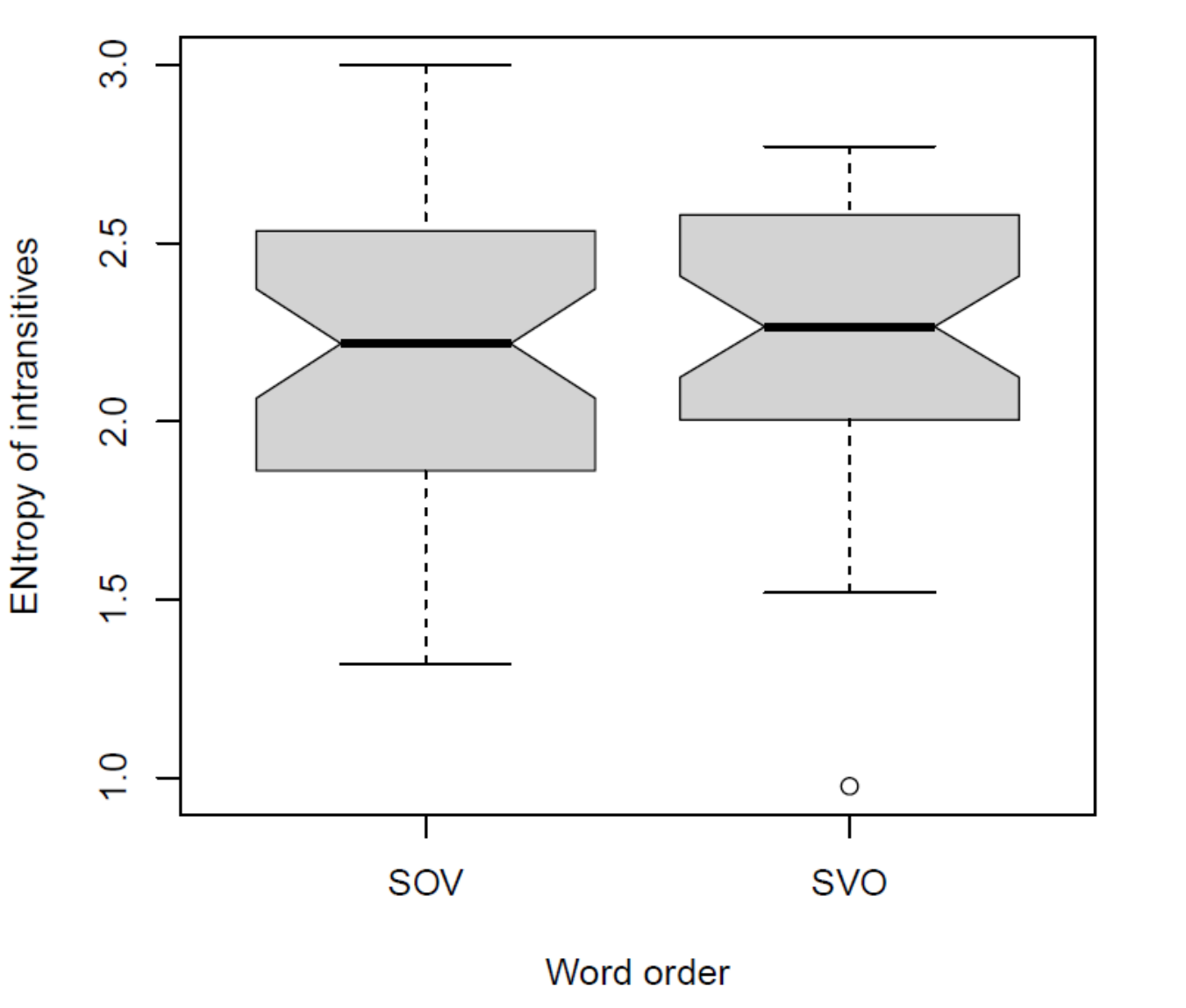

Entropy of intransitives
3.0
2.5
2.0
1.5
1.0
0.3
0.4
0.5
0.6
0.7
0.8
Transitivity ratio
Altaic
Indo−European
Nakh−Daghestanian
Uralic

Entropy & word order

- Overall, there is no robust difference between SOV and SVO languages

  (SVO-languages have only slightly more complex valency systems?)

Entropy & word order



Entropy & word order

- However, more complex valency encoding is associated with postverbal positions
  - Universal prevalence of preverbal subjects (Tomlin 1986, Dryer 2013) with maximal role neutralization (Kibrik 1997, Van Valin & LaPolla 1997: 251ff.)

Entropy & word order

- However, more complex valency encoding is associated with postverbal positions
  - Clause-fronted topics without overt encoding of their thematic relation to the verb (Lambrecht 2001: 1069-1070), as opposed to antitopics

Occitan (ibid.)

(5a) *Lo cinema, i vau sovent*
the cinema there I.go often
'The movies, I goe there often'

(5b) *I vau sovent, **al** cinema*
there I.go often to.the cinema
'I go there often, to the movies'

Entropy & word order

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

Entropy & word order

- However, more complex valency encoding is associated with postverbal positions
  - Languages with the SOVX pattern, e.g. Mande

Bambara (< Mande; Vydrin 2023)

(6a) *Sékù ye nàmasà dún*
PN PFV.TR banana eat
'Seku ate a banana.'

(6b) *Sékù b'í túlomajɔ̀ àrajɔ̀ fɛ̀*
PN IPFV.REFL listen radio\ART by
'Seku is listening to the radio.'

Entropy & word order

- However, more complex valency encoding is associated with postverbal positions
  - Preliminary token-based evidence

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

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

Conclusions and implications

- Complexity of valency classes / systems of valency-encoding patterns
  - 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

Conclusions and implications

- Complex valency class systems
  - are cognitively demanding for the speaker
  - provide redundant but useful cues for the hearer

Entropy & word order

- 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

Conclusions and implications

- To-do list
  - focus on word order and processing
  - use token-based approach
  - test whether more complex systems are associated with satellite-framed as opposed to verb-framed languages
  - get more data, especially for head-marking languages

THANK YOU!

References

- 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.
- 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.

References

- 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.
- Hawkins, John A. 2014. Cross-linguistic variation and efficiency. Oxford: Oxford University Press.
- Kassambara A (2023). *ggpubr: 'ggplot2' Based Publication Ready Plots*. R package version 0.6.0, https://rpkgs.datanovia.com/ggpubr/.
- Kibrik, Aleksandr E. Beyond subject and object: toward a comprehensive relational typology. Linguistic typology 1.3: 279-346.

References

- 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.
- 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.
- Moroz, G. 2017. lingtypology: easy mapping for Linguistic Typology. https://CRAN.R-project.org/package=lingtypology.
- Næss, Å. 2007: *Prototypical Transitivity*. Amsterdam, Philadelphia: John Benjamins.
- 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.

References

- Seržant, Ilja A. 2019. Weak universal forces: The discriminatory function of case in differential object marking systems. In Karsten Schmidtke-Bode, Natalia Levshina, Susanne Maria Michaelis & Ilja A. Seržant (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.
- Tomlin, R.S. 1986. *Basic Word Order: Functional Principles.* London: Routledge.

References

- 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.
- Vydrin, Valentin. 2023. Bivalent patterns in Bambara. In: Say, Sergey (ed.). BivalTyp: Typological database of bivalent verbs and their encoding frames. (Data first published on April 19, 2023; last revised on May 8, 2023.) (Available online at https://www.bivaltyp.info, Accessed on 3 July 2023.