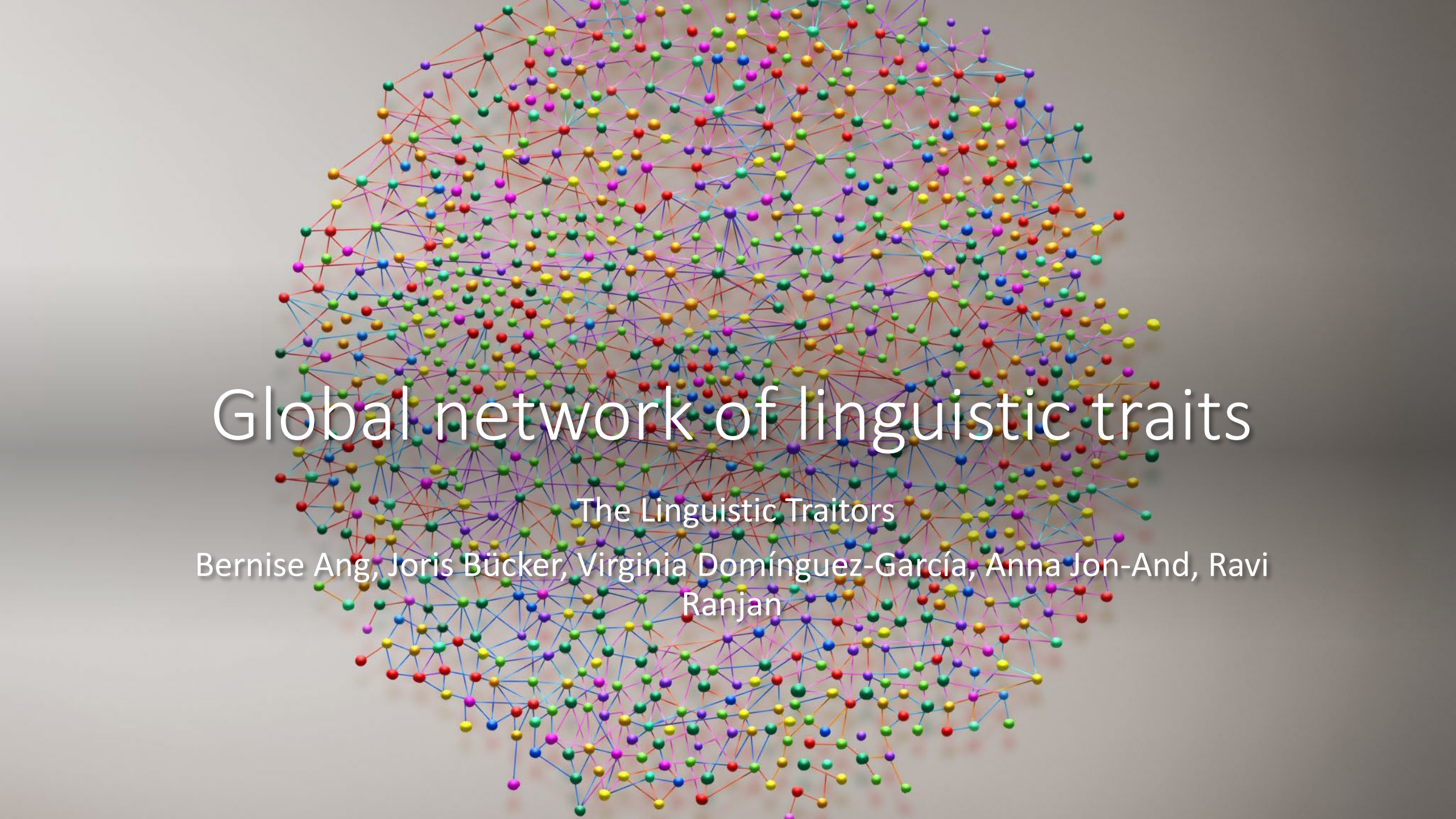


Global network of linguistic traits

The Linguistic Traitors

Bernise Ang, Joris Bücker, Virginia Domínguez-García, Anna Jon-And, Ravi Ranjan

Overall ambitions

Analyse relations between linguistic traits at a systems level.

Understand if and how a linguistic system (the grammar of a language) constrains or influences language change.

Linguistic traits

A structural linguistic trait is a grammatical property of a language.

Examples:

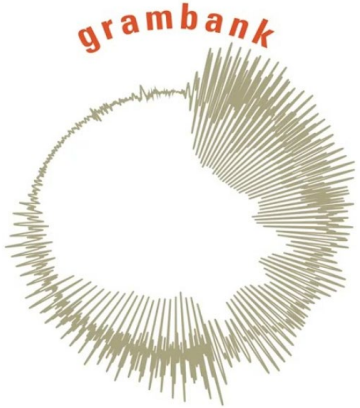
- What is the order of subject and verb in intransitive clauses?

  She reads.

- Are there morphological cases for pronominal core arguments?

  She knows her.

Each language is built up of a set of structural traits that form a system – the grammar of the language



GRAMBANK (Skirgård et al. 2023)
2467 languages
195 traits

Dependencies in language typology

- 45 linguistic universals (Greenberg, 1963):
  - Implicational/hierarchical: "If a language has x, then it also has y."
- Claimed to be pervasive in grammar and to be related to conceptual complexity or processing costs (Croft 2002, Hawkins 1980).
- Lack of larger quantitative studies of hierachical dependencies.

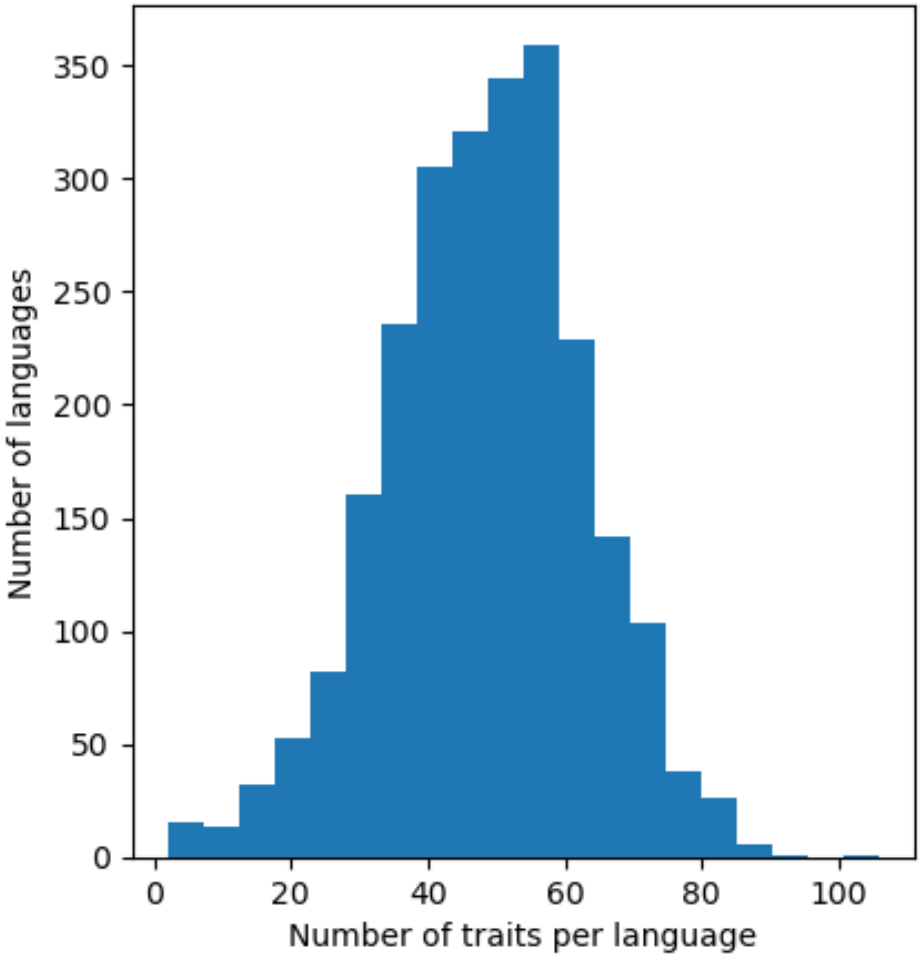
Selective pressures in language evolution

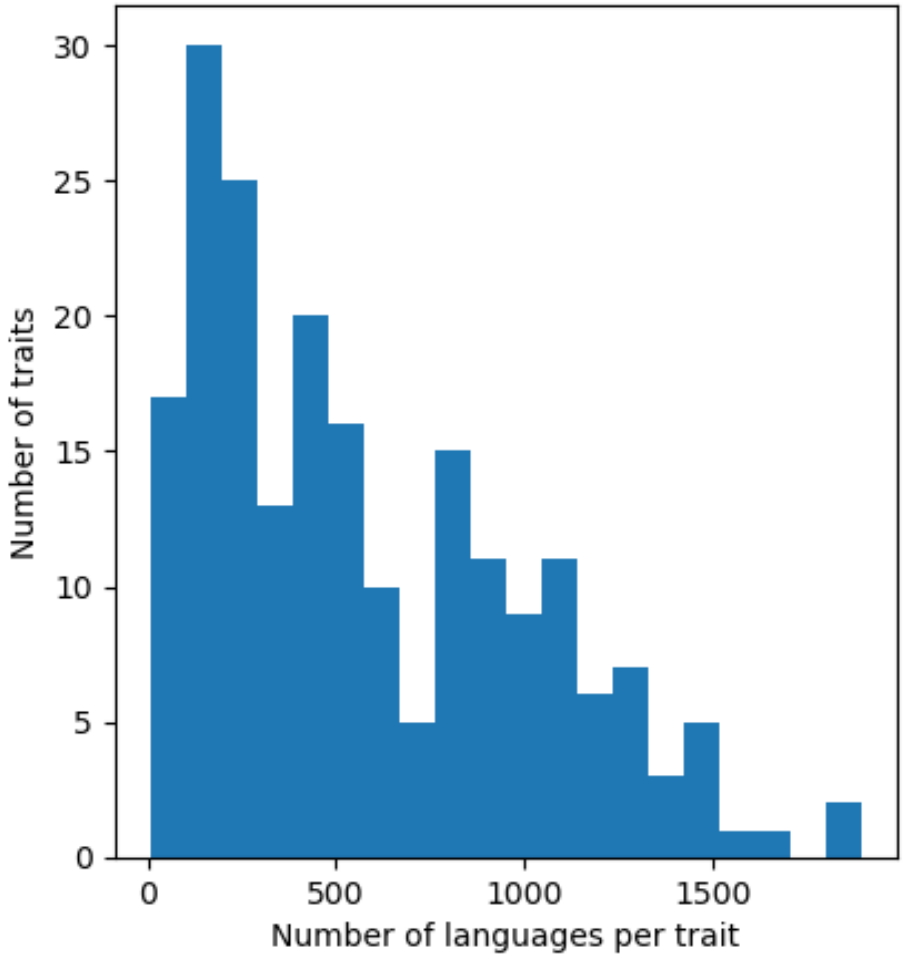
- Focus on universal cultural selective pressures: learnability, expressivity, ease of production/perception/processing.
- If fitness of traits is universal, why the vast grammatical variation?
- Several studies suggest that different linguistic subsystems may expose different dynamics, but do not develop this further.
- Path-dependency in language change is understudied.

Hypotheses on systems tendencies

- Complementarity or trade-offs between traits
- Hierarchy between grammatical traits

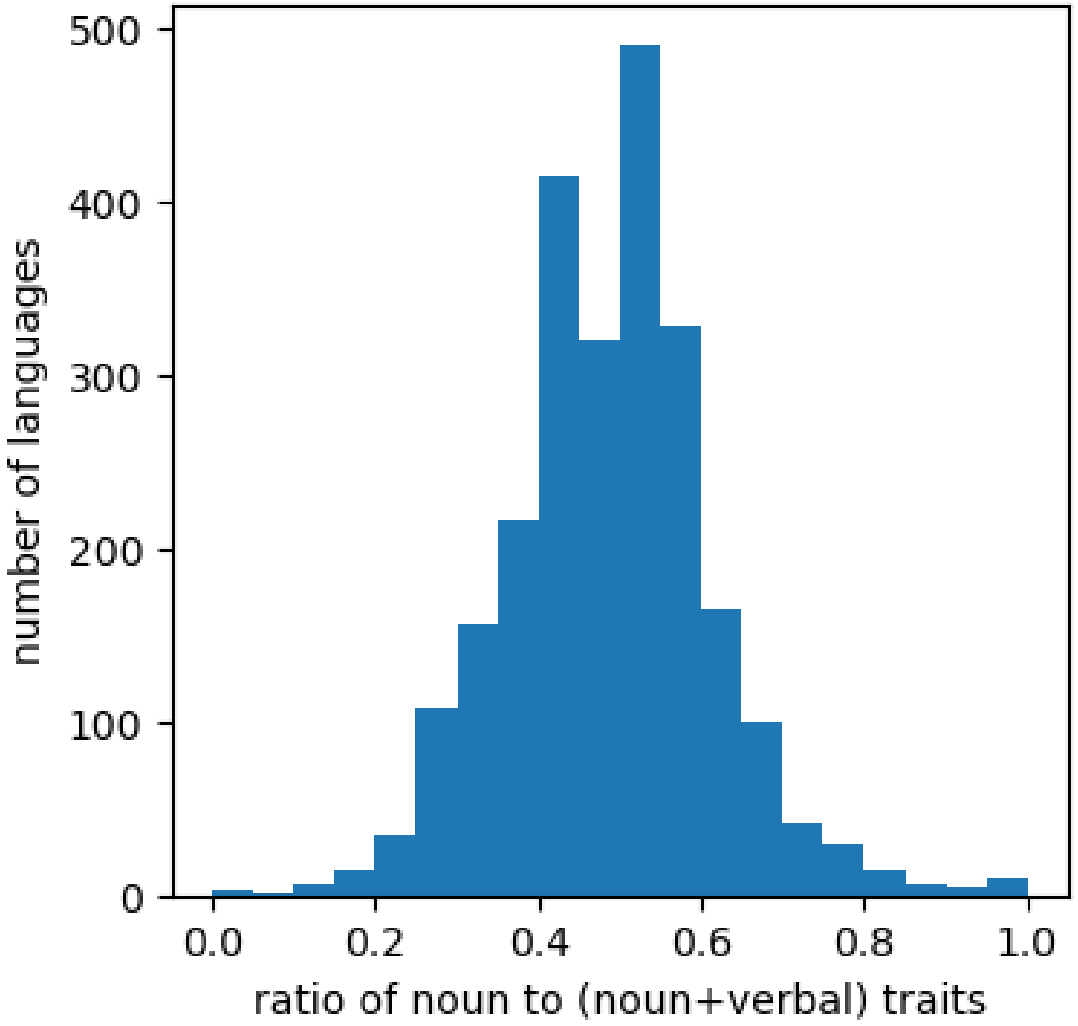
Traits show complementarity





Quantitative trade-off. Most languages have around 50 traits. Languages with few or many traits are unusual.

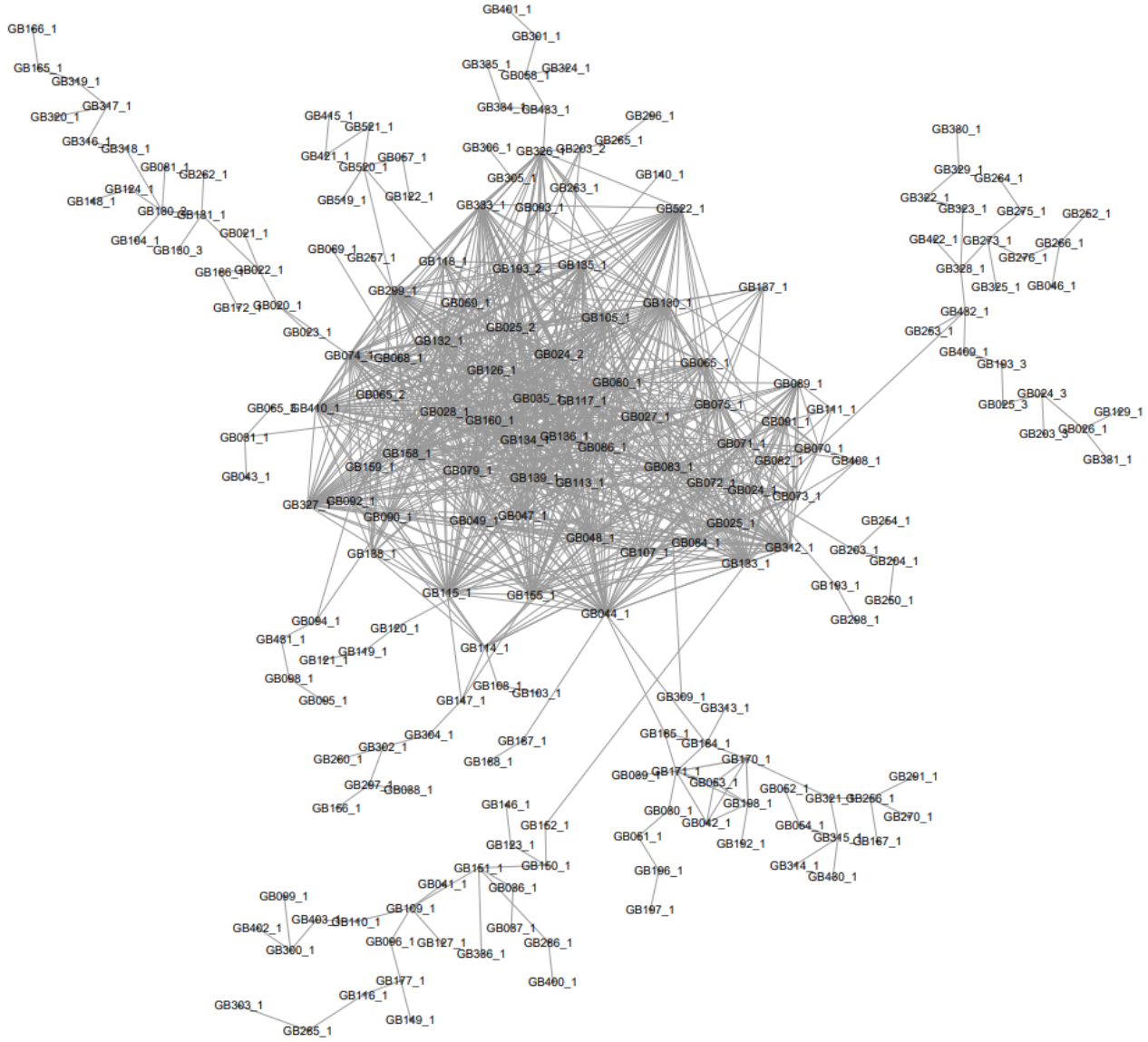
Morphology tends to be evenly distributed between the noun phrase and the verb phrase



Network of traits

- Edges: co-occurrence of traits in languages

- Network indicates a core periphery structure



Backbone of traits network

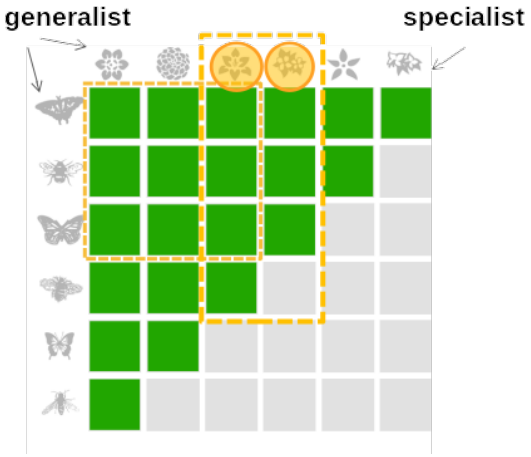
Hierarchy (Nestedness)

Hierarchy (Nestedness)

Languages (N=2647)

Traits (N=210)

Hierarchy (Nestedness)



Not significantly nested
(expect no hierarchy)

$$\eta \equiv \frac{\tilde{\eta}}{\tilde{\eta}_{conf}} = \frac{\langle k \rangle^2}{\langle k^2 \rangle N} \sum_{ij} \frac{(\hat{a}^2)_{ij}}{k_i k_j}. \quad = 1.03$$

NODF = 0.420

Languages

Traits

Hierarchical (Greenberg) Universals

Universal 29: "If a language has inflection, it always has derivation."

Inflection = YES	Derivation=Yes	Both= Yes
2333	1647	1594

Universal 30: "If the verb has categories of person-number or if it has categories of gender, it always has tense-mode categories."

PNG categories = YES	Tense-mode =Yes	Both= Yes
1845	2198	1687

Universal 34:"No language has a trial number unless it has a dual. No language has a dual unless it has a plural."

Trial = YES	Dual =Yes	Both= Yes
16	257	13

Dual = YES	Plural =Yes	Both= Yes
257	1667	235

Complementary Universals (CSSS universals)

Universal 1:"If a language does not have case marking is more likely to have a fixed word order for subject and objects"

Case marking= NO	Fixed word order =Yes	Both= Yes
1343	1245	726

Fixed order is slightly more common without case markings (54%) than considering all languages (50%)

Universal 2:"If the subject can be omitted there is likely person marking on verb"

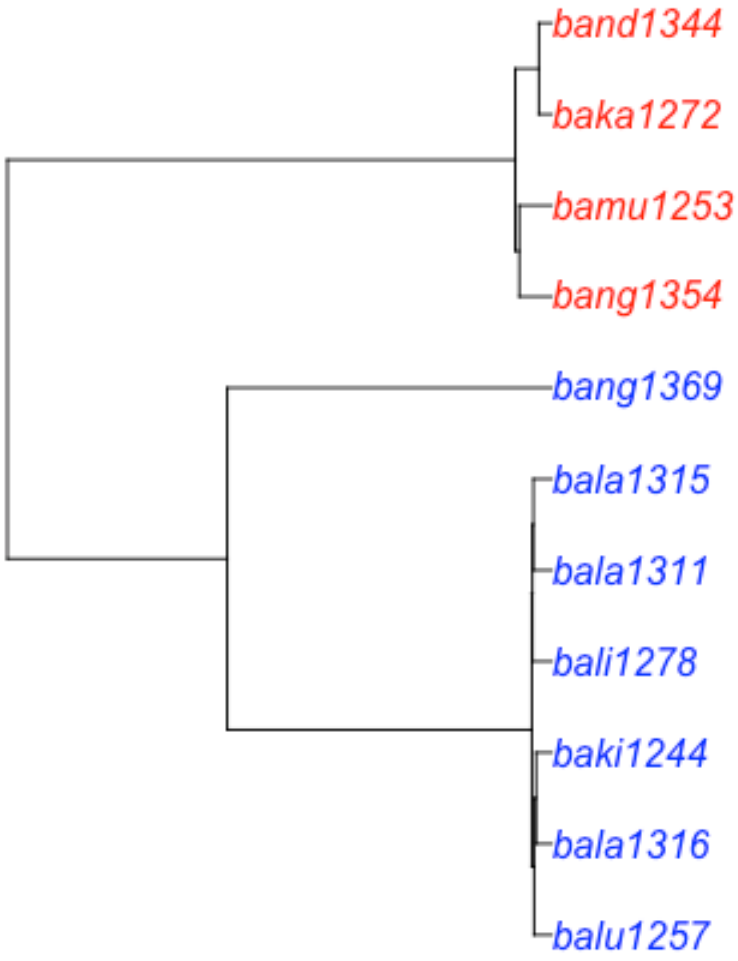
Omit Subj. = YES	Person marking =Yes	Both= Yes
1135	809	408

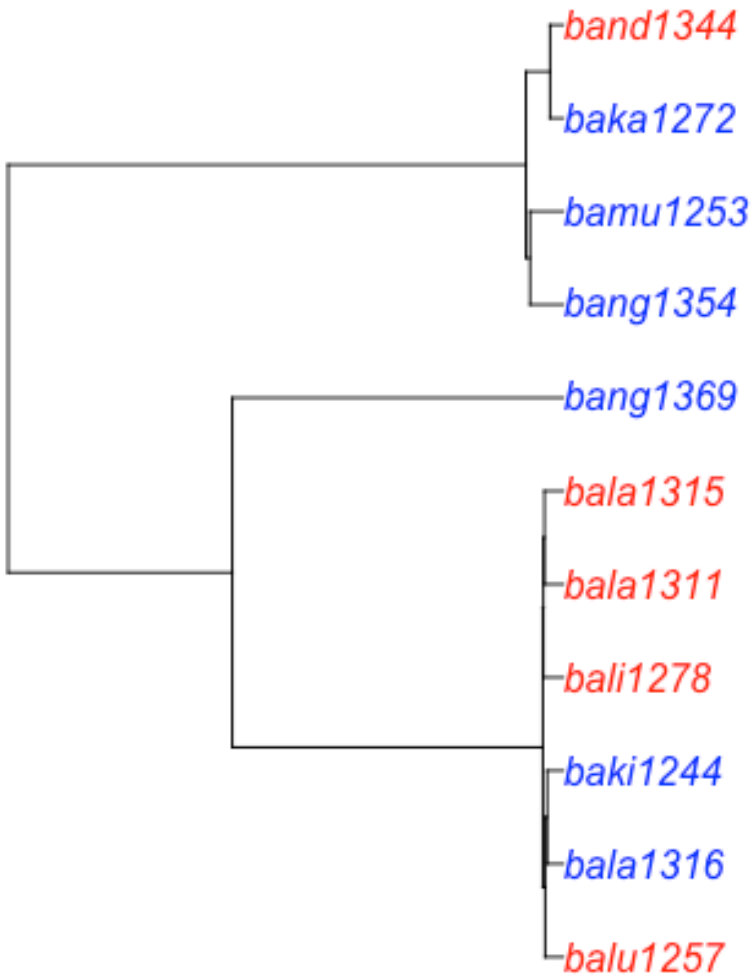
Most languages that omit Subj. do not have person marking.

Conclusions

- Quantitative trade-off confirmed
  - The number of traits in languages is normally distributed.
  - The proportion of verbal morphology and nominal morphology is normally distributed around 50/50
- Qualitative complementarity confirmed in few cases, others not.
- Hierarchical relations between traits are rare.

The road ahead: phylogenies





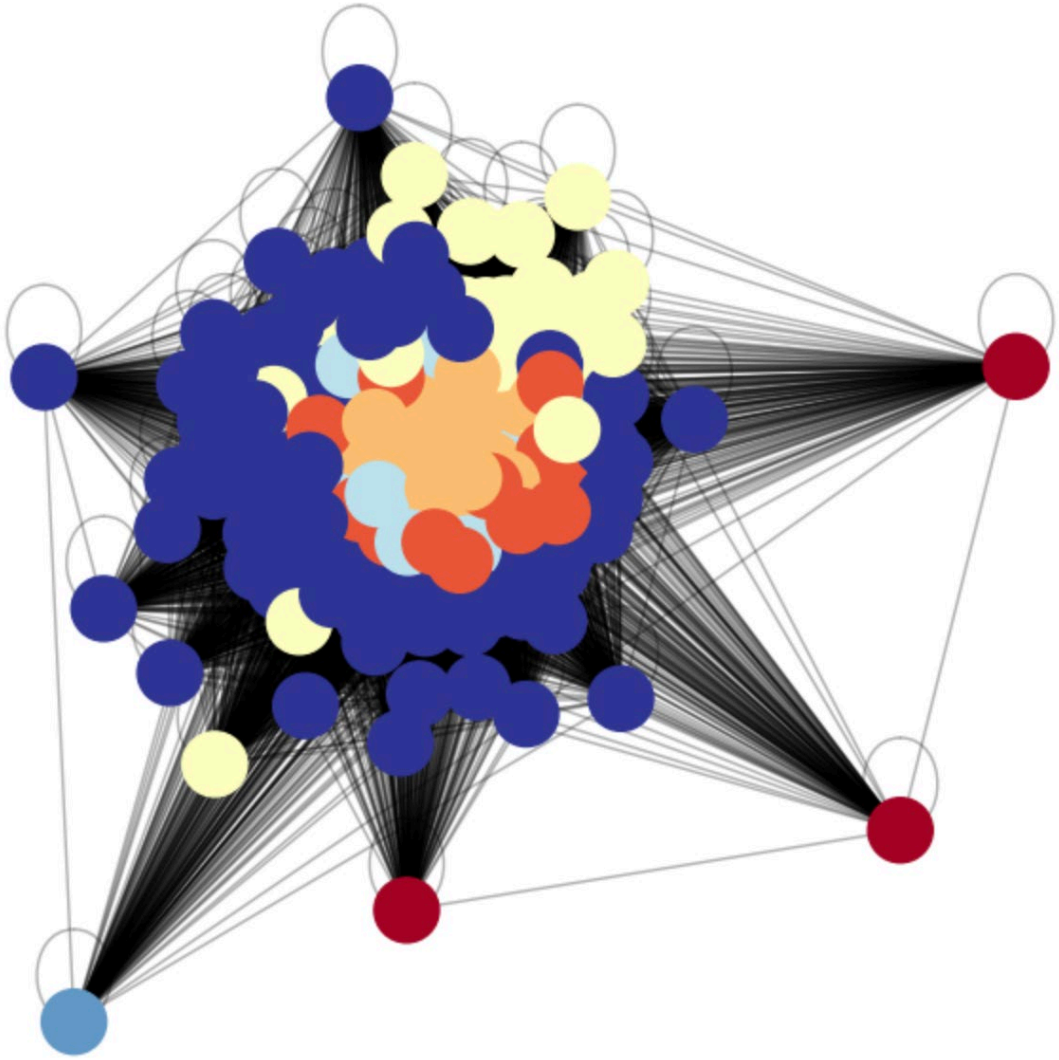
The road ahead: Community Detection

Methods used:

- Mod Opt — Louvain
- Detection of Modular Network structure (DEMON)

Results:

- Number of communities: 4; 1
- Modularity score: 0.06; 0.13



"We'd now like to open the floor to shorter speeches disguised as questions."

Icon art by Gregor Cesnar from the Noun Project