

## The Parametric Comparison Method (PCM)

The PCM aims to reconstruct the **historical evolution of human languages** by pursuing a **phylogenetics of grammars** (Longobardi & Guardiano 2009).

While classical **phylogenetic analysis** identifies etymological relations by uncovering systematic regularities in morphological structures and sound change, the PCM compares languages and captures their historical relations using **formal syntactic rules** universally available to the human language faculty (*parameters*).

Language taxonomies based on parameter states largely correspond to established traditional genealogical groups, while also revealing potentially deeper historical relations (Ceolin et al. 2020, 2021).

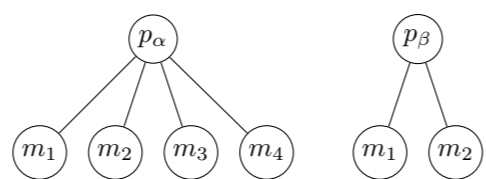
### Parameters and surface patterns

**Parameters:** underlying combinatorial abstract rules encoded in each speaker's individual grammar (I-language, Chomsky 1986)

**Manifestations** of a parameter: the set of different **surface patterns** observed in a language when that parameter is active, i.e. has been *added* to the speaker's I-language during acquisition (Crisma et al. 2020)

parameters → abstract rules

≠  
manifestations → observable structure(s) generated from a rule



**Symbolism** →  
[+] the parameter is active in the grammar (at least one manifestation is observed)  
[-] the parameter is absent from the grammar (none of its manifestations is observed)

By analysing the crosslinguistic distribution of parameter states rather than surface patterns, the PCM aims to capture historical signals that may remain hidden when looking at directly observable structures.

Different observable patterns (e.g. word order)	One abstract rule
Greek: <i>N Gen</i> to thavmastó portréto tis kopélas 'the girl's beautiful portrait'	Lithuanian: <i>Gen N</i> (šis) juodas Reginos automobilis '(this) black car of Regina's'
	Both: <i>GenL</i> Low structural Genitive (Crisma et al. 2024)

### Crossparametric dependencies and parameter networks

Parameters form a dense system of **internal dependencies**: one state of parameter  $\alpha$  may render the state of parameter  $\beta$  predictable. In this case, parameter  $\beta$  is *neutralised*, i.e. contains no information relevant for comparison (encoded as [0], Guardiano & Longobardi 2017).

**Crossparametric dependencies** are **pervasive**: in PCM datasets, ~45% of the total number of states is neutralised.

The system of crossparametric dependencies in a modularised dataset of parameters can be represented as a **directed network graph** (Fig. 1, from the 94 parameters described in Crisma et al. 2025):

parameters → nodes  
dependencies → edges

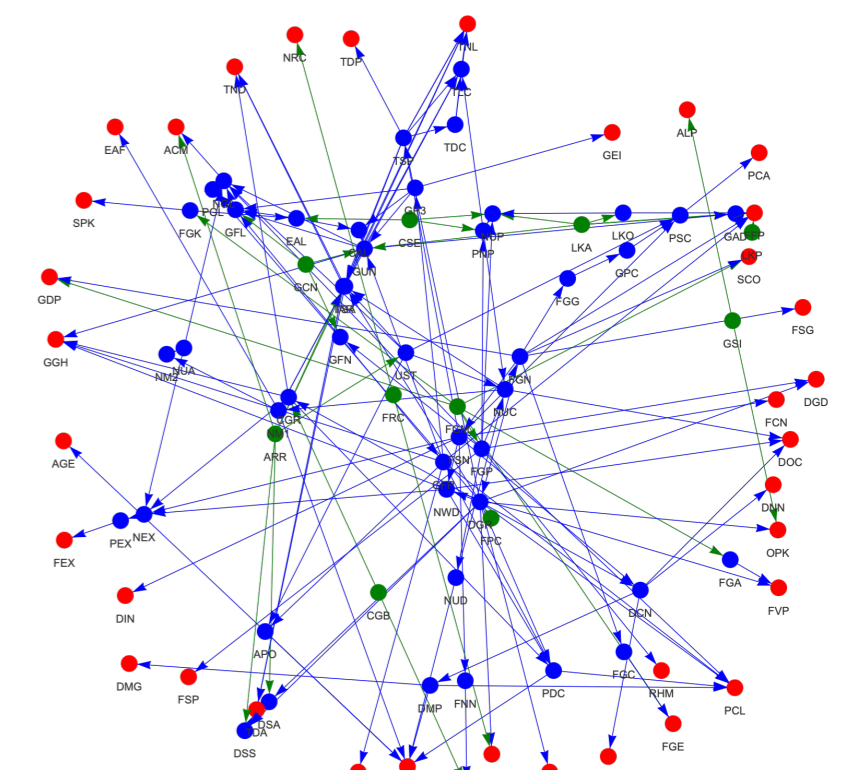


Fig. 1 Dependency structure of the parameter system

Using tools provided by **Social Network Analysis** (SNA, Prell 2012), **structural prominence** of each parameter within the network is quantified in terms of:

- **closeness centrality**, to identify which parameters are more relevant in activating others
- **betweenness centrality**, to identify which parameters act as brokers in developing network chains

### An example: parameter FSN ( $\pm$ Number spread to N)

**Parameter FSN** → distinguishes languages that may mark Number distinctions on nouns (e.g., English, Italian, Arabic) from languages that mark Number distinctions only on determiners (e.g., Basque, Wolof)

**Manifestations** →  
a. the language has nouns that bear variable number morphology (it. *il gatto* 'the.SG cat.SG' vs. *i gatti* 'the.PL cat.PL')  
b. one finds bare nouns in (at least some) argument function (it. *ho bevuto acqua* 'I drank water')

**Dependency** → Parameter FSN has the implicational condition **+FGN** ⇒ parameter FSN can be assigned a state if (and only if) parameter FGN ( $\pm$  grammaticalised Number, active in languages that obligatorily express at least singular/plural distinctions in nominal phrases) has the state [+]. If this condition is not satisfied, parameter FSN is neutralised → [0] (Crisma et al. 2025)

### Languages as networks of parameter states

1. We selected **six languages** from Ceolin et al.'s 2021 dataset: Mandarin (Man, Fig. 2) [Sinitic]; Italian (It), French (Fr), and Romanian (Rm) [Indo-European, Romance]; Bulgarian (Blg, Fig. 3) and Serbo-Croatian (SC) [Indo-European, Slavic]
2. We represented each language as a **network of parameter states**, instantiating the dependency structure in Fig. 1
3. We estimated **pairwise network similarity** (ranging from 0 to 1) as the combination of two equally weighted components:
  - **node attribute distance** → proportion of nodes whose attribute values differ
  - **edge set distance** → Jaccard distance between the edge sets

	Man	It	Fr	Rm	Blg	SC
Man	1	0.697	0.718	0.691	0.681	0.761
It	0.697	1	0.931	0.936	0.878	0.835
Fr	0.718	0.931	1	0.904	0.851	0.824
Rm	0.691	0.936	0.904	1	0.904	0.803
Blg	0.681	0.878	0.851	0.904	1	0.856
SC	0.761	0.835	0.824	0.803	0.856	1

**Similarity levels** range between 0.681 (Man~Blg) and 0.931 (Fr~It)

Results are consistent with the **Anti-Babelic Principle** (Guardiano & Longobardi 2005), which states that genealogically unrelated languages (e.g. Sinitic and Indo-European) are expected to display values closer to the mean than to 0

4. We assessed the **impact** of structurally prominent nodes in determining similarity between networks through a **node-by-node removal** procedure, based on closeness (Fig. 4) and betweenness (Fig. 5) centrality

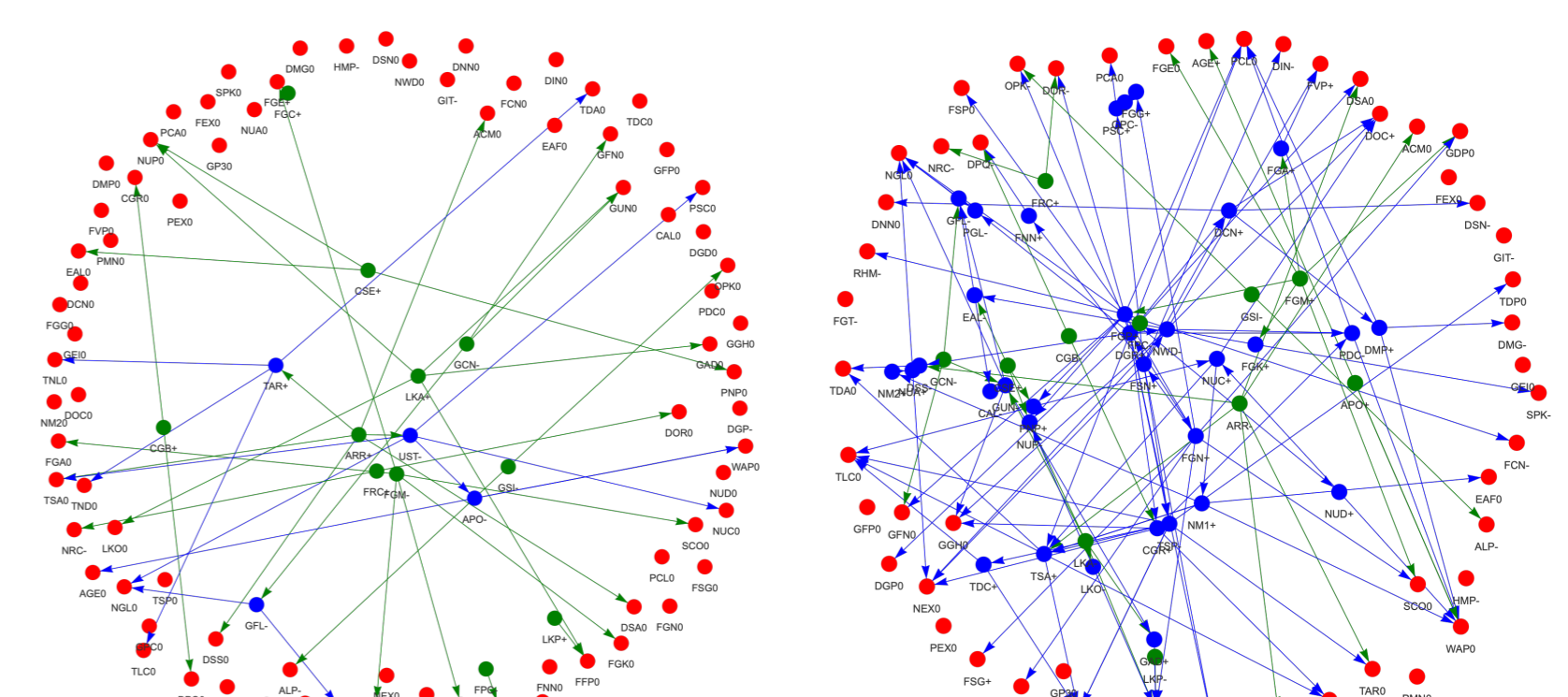


Fig. 2 Mandarin

Fig. 3 Bulgarian

### Network comparison through node removal

Our findings show **distinct correlation trajectories** for closely related versus genealogically unrelated pairs:

- pairs including **Mandarin** (the only non Indo-European language) display **parallel trajectories**, regardless of which parameter is removed
- pairs of languages belonging to the **same subgroup** display **parallel trajectories**, regardless of which parameter is removed, even when their overall similarity differs
- although the pair Romanian~Bulgarian shows some marked proximity, reasonably due to well-known secondary convergence effects (*Balkan Sprachbund*, Tomić 2006), the corresponding curve still follows **the same trajectory as other Romance~Slavic pairs**

**TAKE-HOME:** Absolute correlation ranges (including *distance* measures classically adopted in PCM works) may **lack in resolution** especially when considering languages with a history of prolonged contact, like Bulgarian and Romanian. In this respect, SNA may provide worth-exploring criteria to discriminate between deep structural similarities and (accidental) convergencies due to recent events.

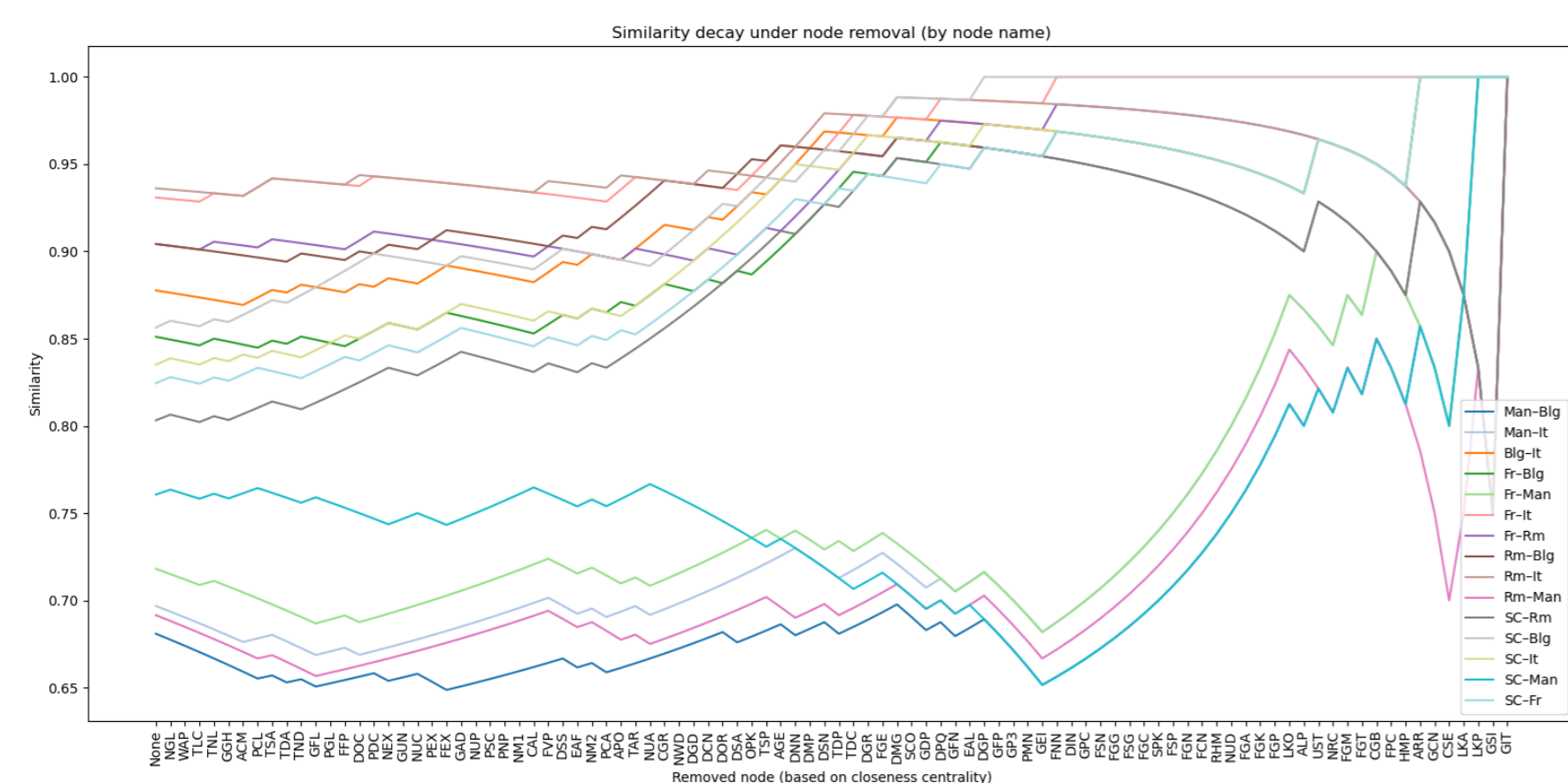


Fig. 4 Network similarity and node removal based on closeness centrality

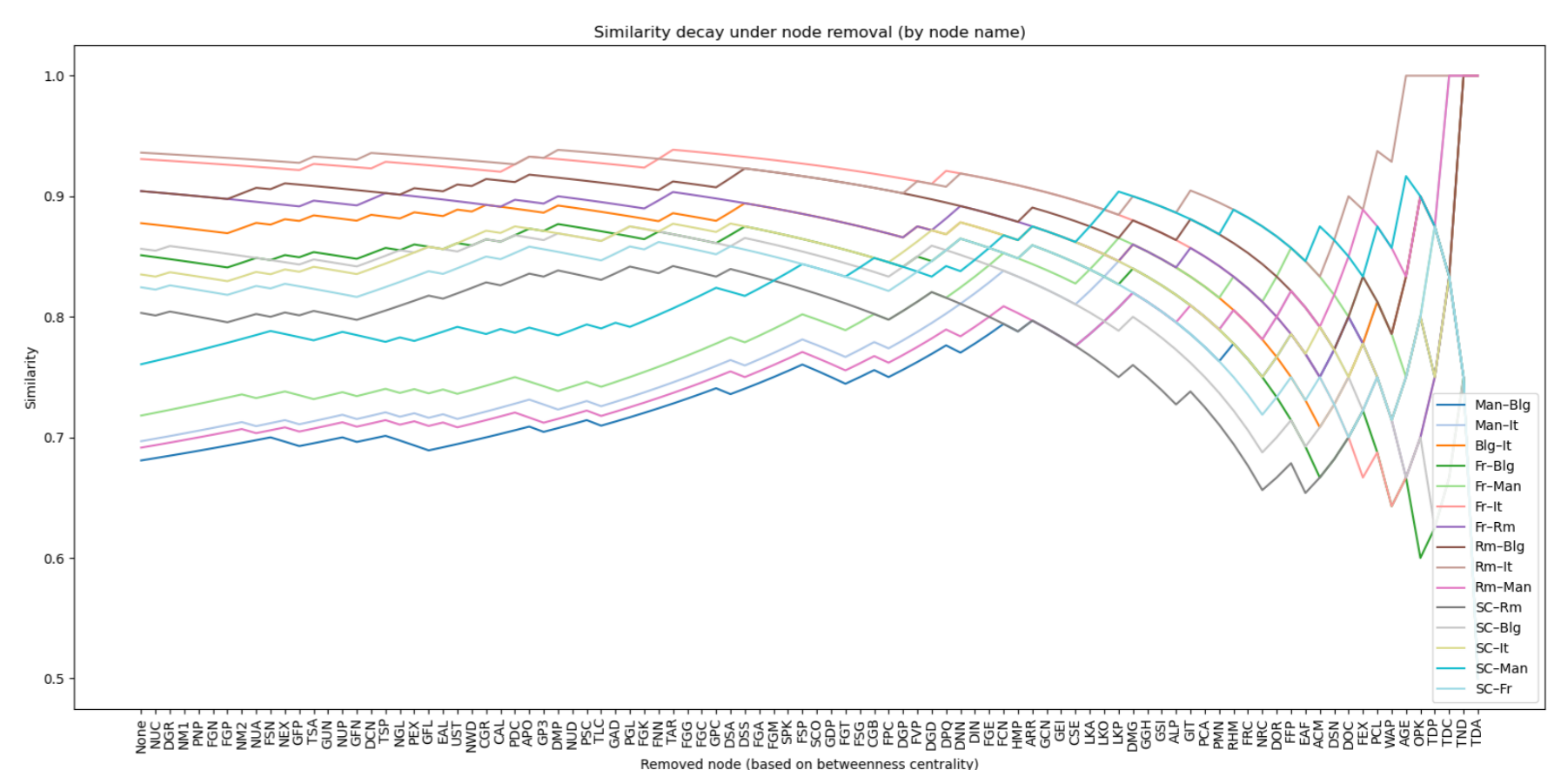


Fig. 5 Network similarity and node removal based on betweenness centrality

### Analytical outlook

- Languages with the same degree of historical relatedness exhibit similar node-centrality configurations
- SNA may enable the integration of dependency structure into parametric phylogenetic modeling
- Extending the analysis to more diverse language samples could validate network-structure analysis as a method for quantifying historical relatedness and modeling parameter change

### References

Ceolin, A., C. Guardiano, M. A. Irimia & G. Longobardi. 2020. Formal syntax and deep history. *Frontiers in Psychology* 11, 488871.  
 Ceolin, A., C. Guardiano, G. Longobardi, M. A. Irimia, L. Bortolussi & A. Sgarro. 2021. At the boundaries of syntactic prehistory. *Philosophical Transactions of the Royal Society B* 376, 20200197.  
 Chomsky, N. 1986. *Knowledge of language. Its nature, origin, and use*. New York: Praeger.  
 Crisma, P., G. Fabbris, G. Longobardi & C. Guardiano. 2025. What are your values? Default and asymmetry in parameter states. *Journal of Historical Syntax* 9, 1-26.  
 Crisma, P., C. Guardiano & G. Longobardi. 2020. Syntactic diversity and language learnability. *Studi e Saggi Linguistici* 58(2), 99-130.  
 Crisma, P., C. Guardiano & G. Longobardi. 2024. A unified theory of Case form and Case meaning. Genitives and parametric syntax. In ed. by C. Sevdali et al., *The Place of Case in Grammar*, 427-466. Oxford: OUP.  
 Guardiano, C. & G. Longobardi. 2005. Parametric comparison and language taxonomy. In ed. by M. Batllori et al., *Grammaticalization and parametric variation*, 149-174. Oxford: OUP.  
 Guardiano, C. & G. Longobardi. 2017. Parameter theory and parametric comparison. In ed. by I. Roberts, *The Oxford Handbook of Universal Grammar*, 377-398. OUP.  
 Longobardi, G. & C. Guardiano. 2009. Evidence for syntax as a signal of historical relatedness. *Lingua* 119, 1679-1706.  
 Prell, C. 2012. *Social Network Analysis. History, theory and methodology*. London: SAGE Publications.  
 Tomić, O. M. 2006. *Balkan Sprachbund morpho-syntactic features*. Dordrecht: Springer.