

Supplementary Materials for

The evolutionary dynamics of how languages signal who does what to whom

Olena Shcherbakova, Damián E. Blasi, Volker Gast, Hedvig Skirgård, Russell D. Gray, and Simon J. Greenhill

*Corresponding author. Email: olena_shcherbakova@eva.mpg.de

This PDF file includes:

Figs. S1 to S6
Tables S1 to S2

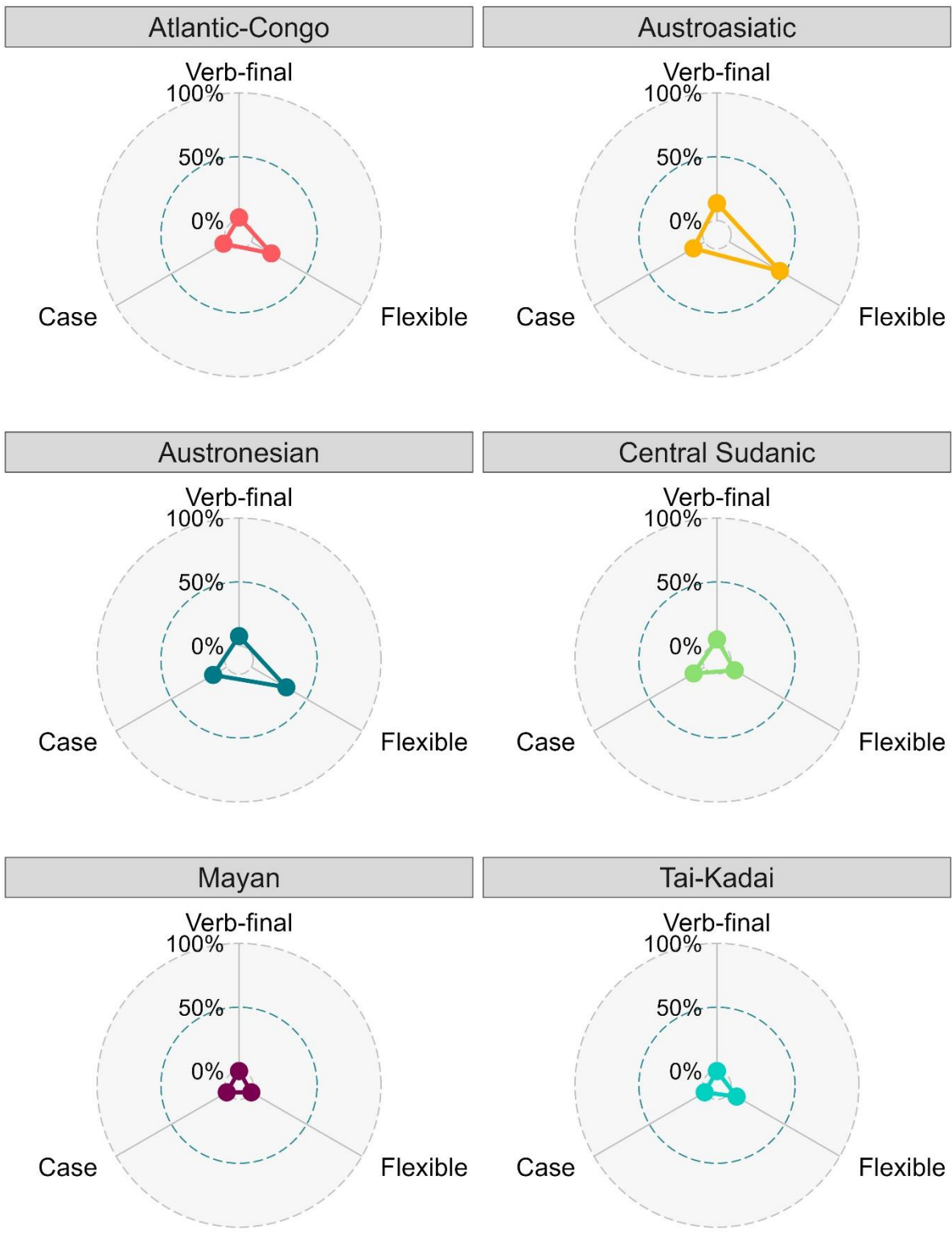


Fig. S1. The radar chart illustrating percentages of presence of three grammatical features in languages of six language families: Atlantic-Congo, Austroasiatic, Austronesian, Central Sudanic, Mayan, and Tai-Kadai. The features are predominantly absent.

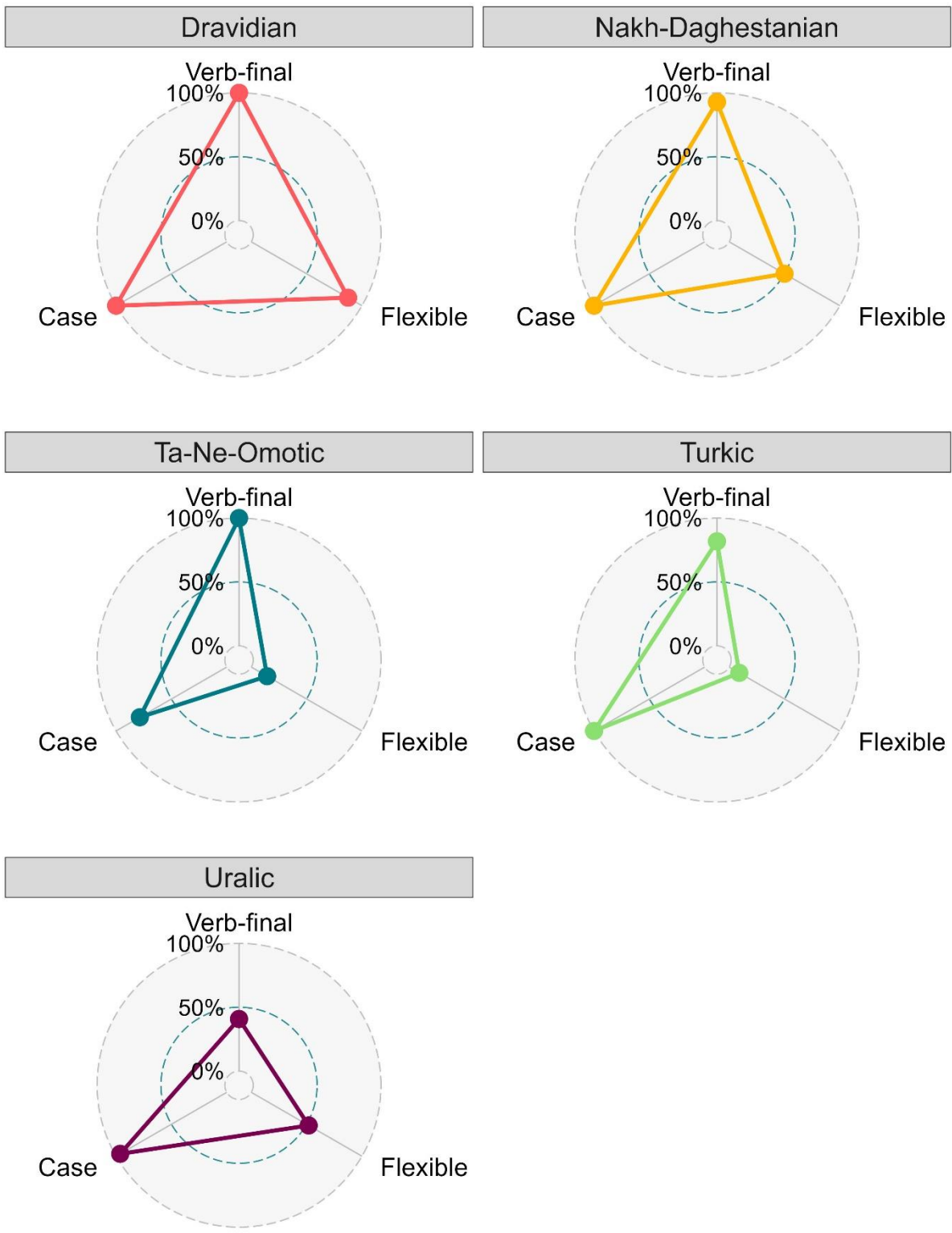
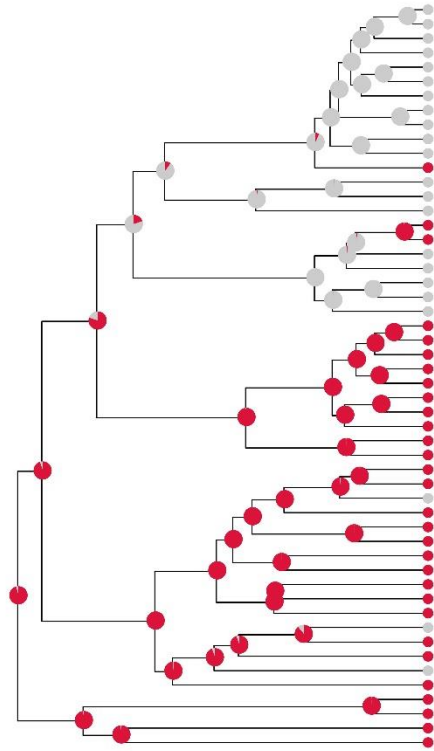


Fig. S2. The radar chart illustrating percentages of presence of three grammatical features in languages of five language families: Dravidian, Nakh-Daghestanian, Ta-Ne-Omotc, Turkic, and Uralic. In these language families, case is present at least in the majority of the languages, and verb-final word order is found in most languages of all but Uralic language families. Most

Dravidian languages and around half of Nakh-Daghestanian and Uralic languages have flexible word order.

Case

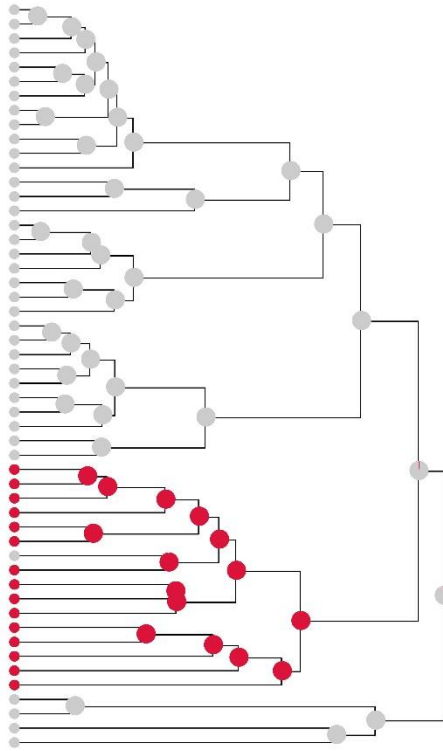
- absent
- present



- Portuguese
- Galician
- Catalan
- Occitan
- Romansh
- French
- Lombard
- Sicilian
- Italian
- Corsican
- Campidanese Sardinian
- Aromanian
- Welsh
- Breton
- Irish
- Icelandic
- Faroese
- Danish
- Swedish
- English
- Western Frisian
- Dutch
- Ukrainian
- Belarusian
- Russian
- Polish
- Czech
- Slovenian
- Serbian-Croatian-Bosnian
- Macedonian
- Lithuanian
- Latvian
- Darai
- Bhojpuri
- Maijthili
- Odia
- Konkan Marathi
- Marathi
- Kashmiri
- Eastern Panjabi
- Lambadi
- Awadhi
- Hindi
- Kumzari
- Western Farsi
- Southern Pashto
- Gurani
- Prasun
- Gheg Albanian
- Northern Tosk Albanian
- Eastern Armenian
- Modern Greek

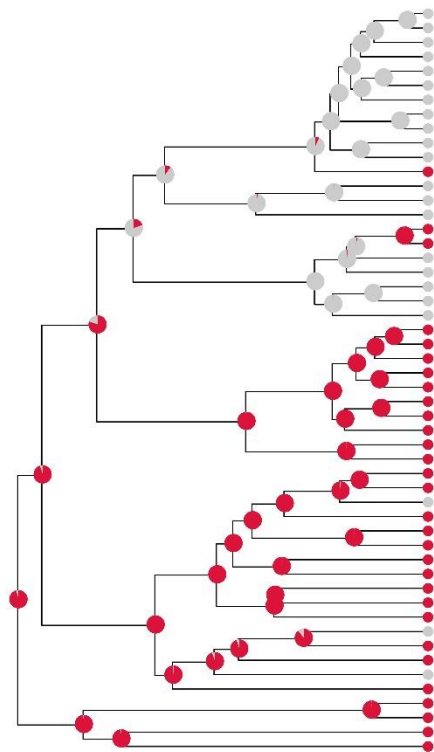
Verb-final word order

- absent
- present



Case

- absent
- present



- Portuguese
- Galician
- Catalan
- Occitan
- Romansh
- French
- Lombard
- Sicilian
- Italian
- Corsican
- Campidanese Sardinian
- Aromanian
- Welsh
- Breton
- Irish
- Icelandic
- Faroese
- Danish
- Swedish
- English
- Western Frisian
- Dutch
- Ukrainian
- Belarusian
- Russian
- Polish
- Czech
- Slovenian
- Serbian-Croatian-Bosnian
- Macedonian
- Lithuanian
- Latvian
- Darai
- Bhojpuri
- Maijthili
- Odia
- Konkan Marathi
- Marathi
- Kashmiri
- Eastern Panjabi
- Lambadi
- Awadhi
- Hindi
- Kumzari
- Western Farsi
- Southern Pashto
- Gurani
- Prasun
- Gheg Albanian
- Northern Tosk Albanian
- Eastern Armenian
- Modern Greek

Flexible word order

- absent
- present

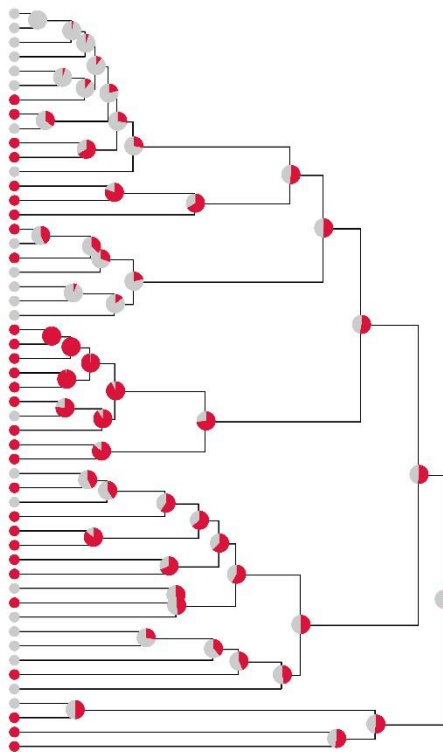
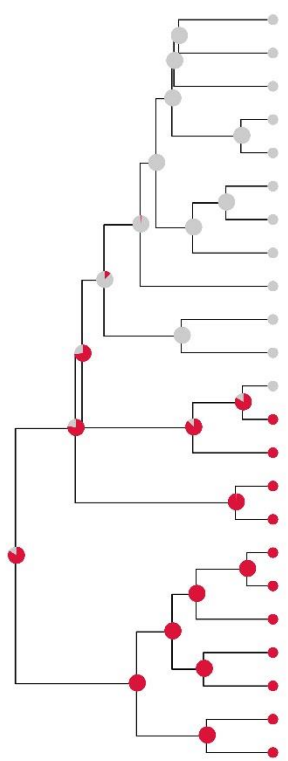


Fig. S3. The facing trees of the Indo-European language family with values of nominal case (on the left) and word order: verb-final (top right) and flexible word order (bottom right). Red stands for the presence of the grammatical feature, while gray indicates its absence. The languages from the Indic branch preserve both case and verb-final word order, whereas many languages from the Italic and Germanic branches lost both of these features. In the Slavic branch, the presence of case was preserved along with flexible word order.

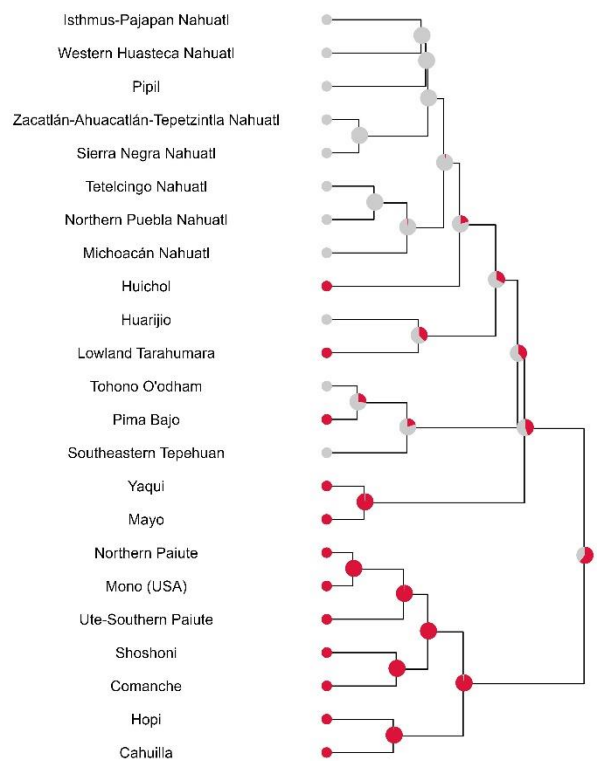
Case

- absent
- present



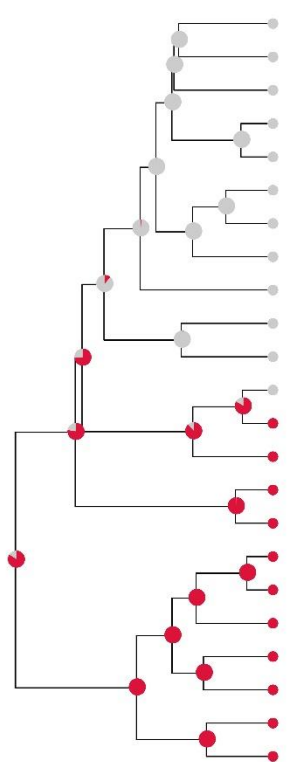
Verb-final word order

- absent
- present



Case

- absent
- present



Flexible word order

- absent
- present

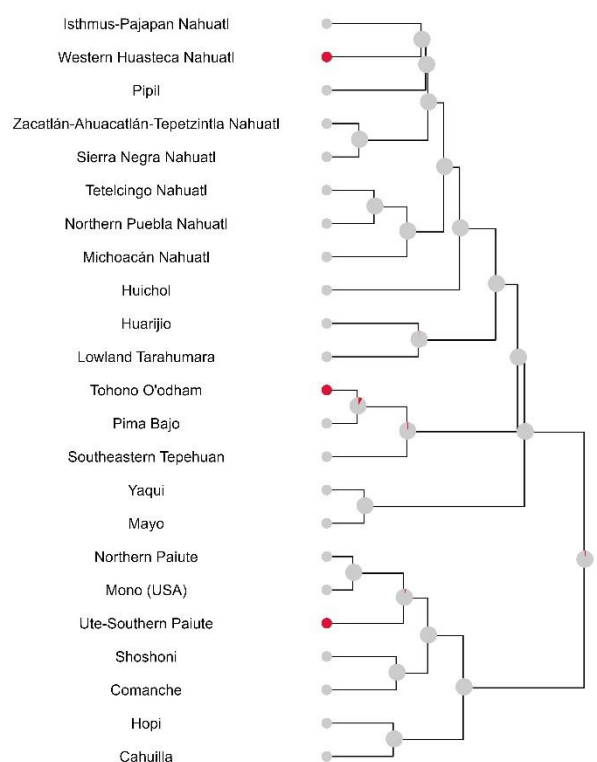


Fig. S4. The facing trees of the Uto-Aztecan language family with values of nominal case (on the left) and word order: verb-final (top right) and flexible word order (bottom right). Red stands for the presence of the grammatical feature, while gray indicates its absence. Verb-final word order and case have been faithfully preserved in Northern Uto-Aztecan languages, whereas many Northern Uto-Aztecan lost both of the features.

Case

- absent
- present

Verb-final word order

- absent
- present

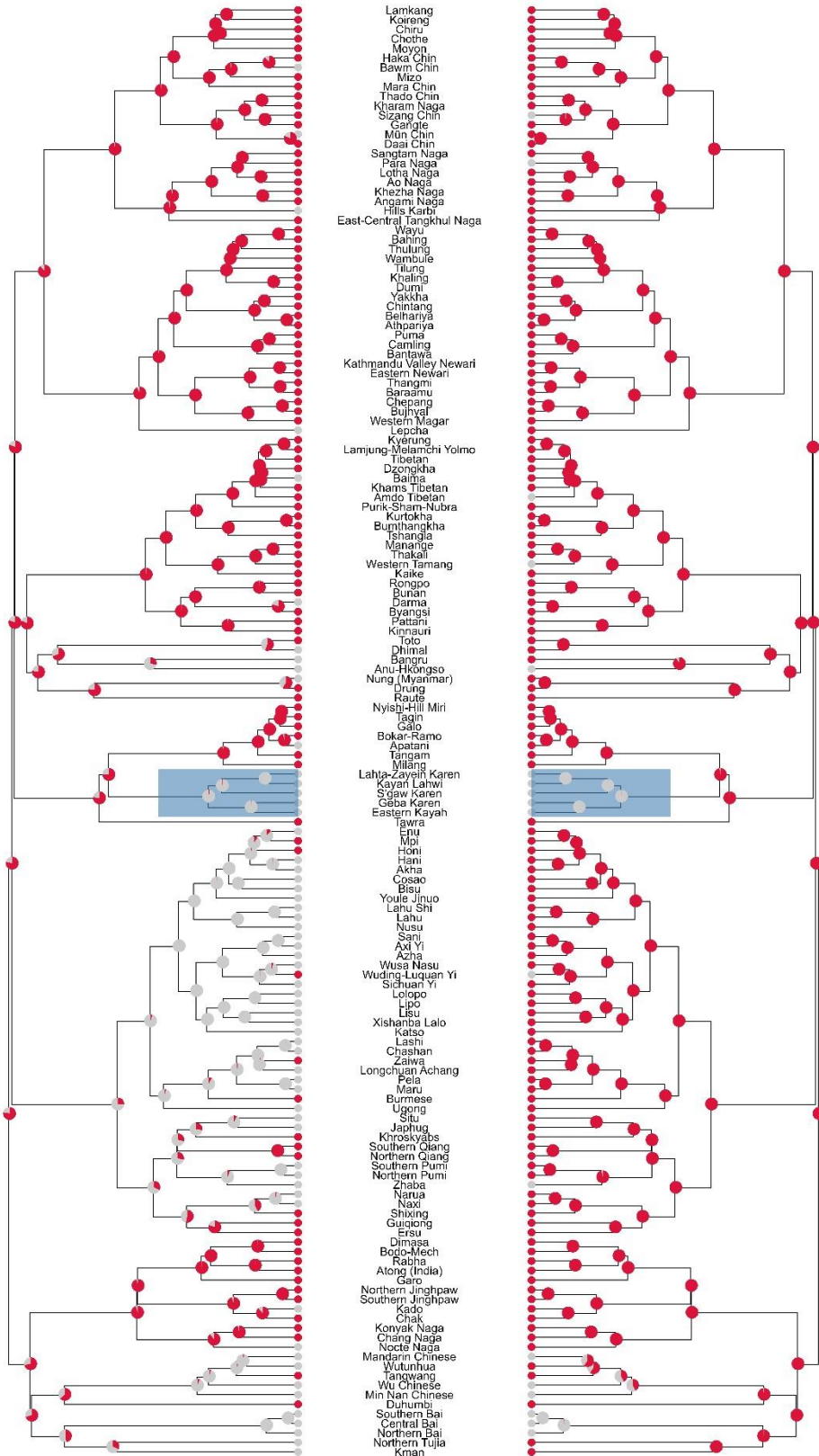


Fig. S5. The facing trees of the Sino-Tibetan family with values of nominal case (on the left) and verb-final word order (on the right). Red stands for the presence of the grammatical feature, while gray indicates its absence. Both case and verb-final word order are absent in the languages of the Karenic branch (Lahta-Zayein Karen, Kayan Lahwi, S'gaw Karen, Geba Karen, and Eastern Kayan), whereas other clades possess both of these features.

Case

- absent
- present

Flexible word order

- absent
- present

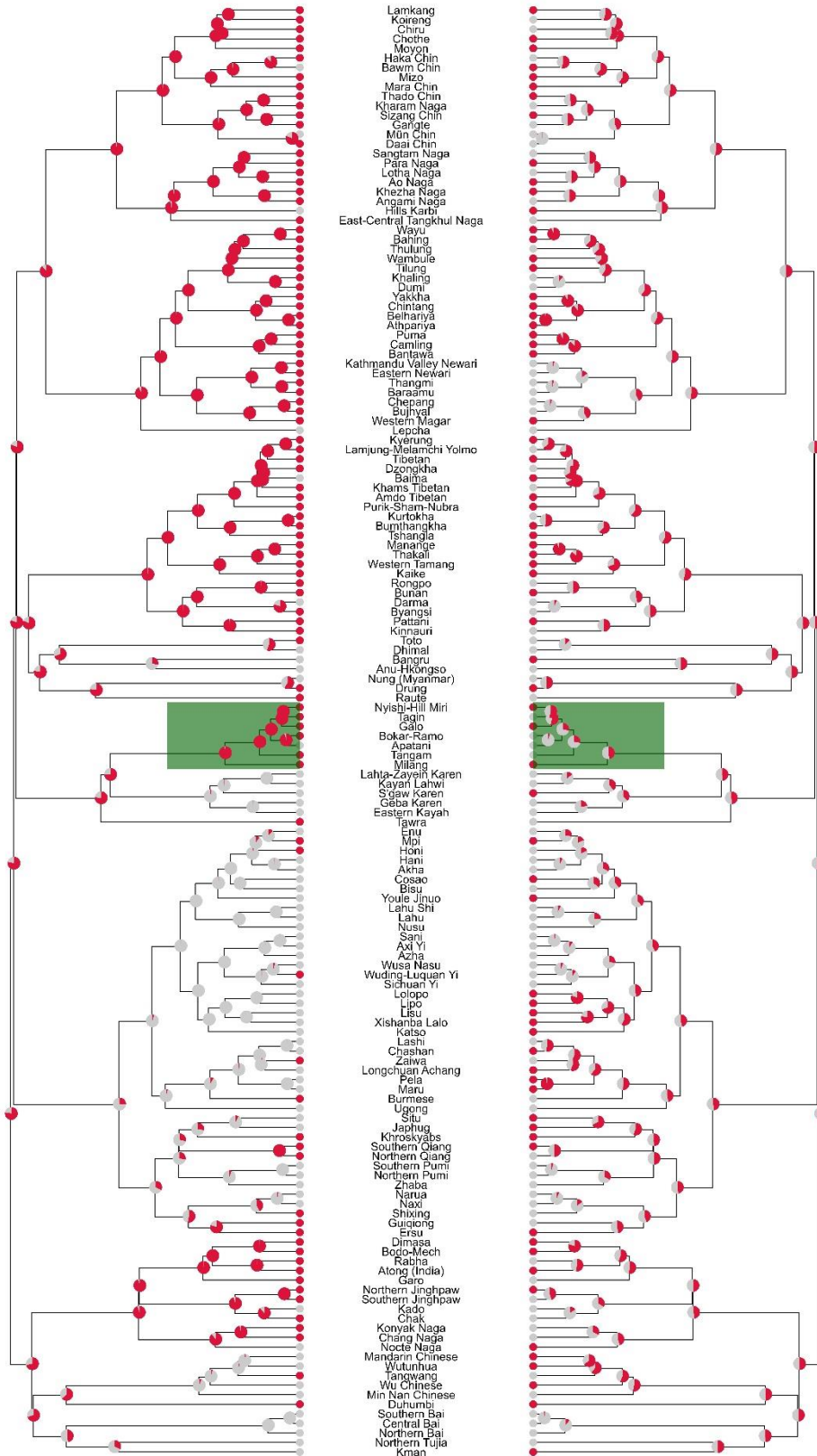


Fig. S6. The facing trees of the Sino-Tibetan family with values of nominal case (on the left) and flexible word order (on the right). Red stands for the presence of the grammatical feature, while gray indicates its absence. Almost all languages from the Macro-Tani branch in our sample (except for Apatani) possess case, and some of these languages, such as Nyishi-Hill Miri, Galo, and Milang have likely developed flexible word order.

Table S1. The results of cross-validation of three models fit with *brms*, which incorporate random effects and predict the distribution of case, verb-final word order and flexible word order.

Model	Response variable	Δ ELPD	Δ SE	Cross-validation type
spatial + phylogenetic	Case	0.00	0.00	LOO-CV
phylogenetic		-0.25	1.60	
spatial		-138.38	17.47	
phylogenetic		0.00	0.00	LOO with moment matching
spatial + phylogenetic		-3.27	1.78	
spatial		-135.52	17.80	
phylogenetic		0.00	0.00	K-fold CV
spatial		-0.23	0.92	
spatial + phylogenetic		-1.05	1.02	
phylogenetic	Verb-final word order	0.00	0.00	LOO-CV
spatial + phylogenetic		-13.04	5.05	
spatial		-215.52	20.01	
phylogenetic		0.00	0.00	LOO with moment matching
spatial + phylogenetic		-15.34	5.18	
spatial		-207.27	20.06	
spatial		0.00	0.00	K-fold CV
phylogenetic		-0.37	1.03	
spatial + phylogenetic		-0.38	1.19	
spatial + phylogenetic	Flexible word order	0.00	0.00	LOO-CV
phylogenetic		-5.82	6.65	
spatial		-26.73	8.34	

spatial + phylogenetic		0.00	0.00	LOO with moment matching
phylogenetic		-4.73	6.65	
spatial		-26.82	8.35	
spatial + phylogenetic		0.00	0.00	K-fold CV
spatial		-1.20	0.77	
phylogenetic		-2.31	0.82	

Table S2. The coefficients of the fixed effects (with the 95% credible intervals) in three best-supported models (*b*, *c*, and *d*) by Phylogenetic Path Analysis fit with *brms* to test the robustness of the results after incorporating spatial random effect to predict Flexible word order.

Response	Predictor	Model	Estimate	Est.Error	Q2.5	Q97.5
Flexible word order	Case	model b	0.38	0.18	0.02	0.72
Verb-final	Case		1.76	0.59	0.67	3.00
Flexible word order	Case	model c	0.37	0.18	0.02	0.73
Case	Verb-final		1.49	0.33	0.85	2.17
Verb-final	Case	model d	1.75	0.59	0.66	2.99
Case	Flexible word order		0.25	0.22	-0.18	0.68