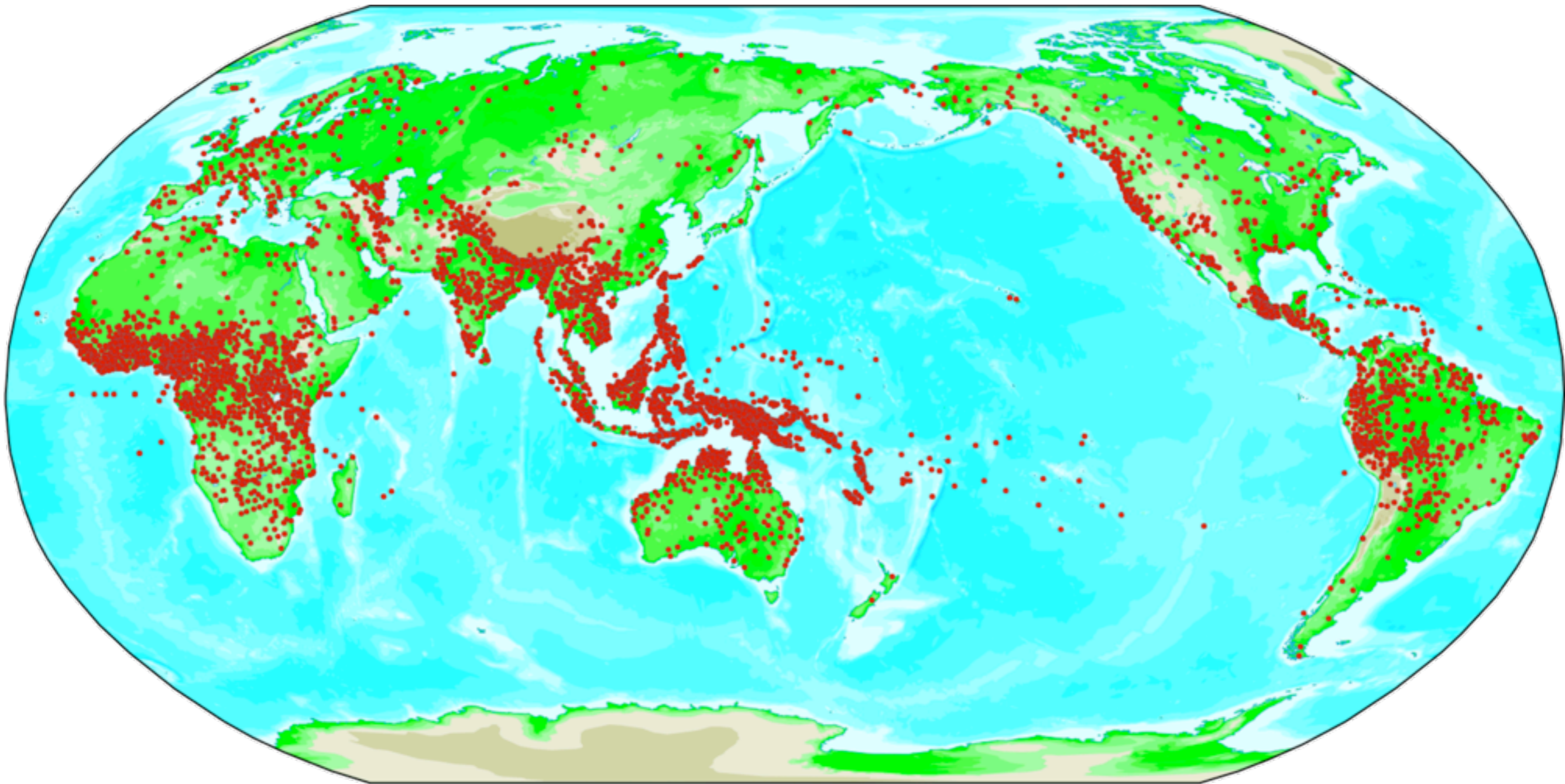


Geographic Distribution of Linguistic Diversity

A map of Southeast Asia and Oceania showing linguistic diversity hotspots. The map is color-coded with green for land and light blue for water. Colored dots are placed on various islands and regions: yellow dots in the Philippines, Sumatra, Java, and New Guinea; red dots in Sulawesi, Irian Jaya, and the Moluccas; and purple dots in Australia. The text 'Geographic Distribution of Linguistic Diversity' is overlaid in large black font.

Michael Cysouw
LMU München

about 7K languages



World Atlas of Language Structures (WALS)

- 142 world maps with structural linguistic information about hundreds of languages
- Phonology, Morphology, Syntax, Lexicon
- Somewhat simplistic and at times even redundant classifications

The World Atlas of Language Structures Online

[Home](#)[Features](#)[Languages](#)[References](#)[Authors](#)[Newsblog](#)[Contact](#)[Help](#)[Download WALs data](#) [errata](#) [credits](#) [legal](#)

Welcome to WALs Online

The data and the texts from *The World Atlas of Language Structures*, published as a book with CD-ROM in 2005 by [Oxford University Press](#), are now freely available online.

WALS Online is a joint project of the [Max Planck Institute for Evolutionary Anthropology](#) and the [Max Planck Digital Library](#). It is a separate publication, edited by Martin Haspelmath, Matthew S. Dryer, David Gil and Bernard Comrie (Munich: Max Planck Digital Library, 2008).

What is WALs?


WALS is a large database of structural (phonological, grammatical, lexical) properties of languages gathered from descriptive materials (such as reference grammars) by a team of more than 40 authors (many of them the leading authorities on the subject).

WALS consists of 141 maps with accompanying texts on diverse features (such as vowel inventory size, noun-genitive order, passive constructions, and "hand"/"arm" polysemy), each of which is the responsibility of a single author (or team of authors). Each map shows between 120 and 1370 languages, each language being represented by a symbol, and different symbols showing different values of the feature. Altogether 2,650 languages are shown on the maps, and more than 58,000 datapoints give information on features in particular languages.

WALS thus makes information on the structural diversity of the world's languages available to a large audience, including interested nonlinguists as well as linguists who would not normally read grammars of exotic languages or specialized works by comparative linguists. Although endangered languages are not particularly emphasized, they are automatically foregrounded because of the large sample of languages represented on each map, where each language (independently of its number of speakers) is shown by a single symbol.

Interactive Reference Tool (WALS program)

The World Atlas of Language Structures was published as a book with a CD-ROM in summer 2005. The CD-ROM contains the "Interactive Reference Tool (WALS program)" as a standalone application for Mac OS X, Mac OS 9.2 and Windows 2000, XP written by [Hans-Jörg Bibiko](#). To download the "Interactive Reference Tool (WALS program)" please follow the link <http://www.eva.mpg.de/lingua/research/tool.php>.

 powered by 

WALS News

Scheduled Server Downtime

by robert - May 05, 2009

To fulfill the test regulations of the Land Brandenburg, the data center where the WALs servers are hosted will have to shut ...

WALS Online not reachable

by robert - Apr 14, 2009

From April 13, 16:58 CEST until this morning, April 14, 7:46 CEST WALs Online was not reachable, due to an unplanned downtime ...

Location for Mehri updated

by robert - Mar 04, 2009

The geo-coordinates for Mehri have been corrected - since Mehri is spoken mostly in Yemen, it should be located there. Find the ...

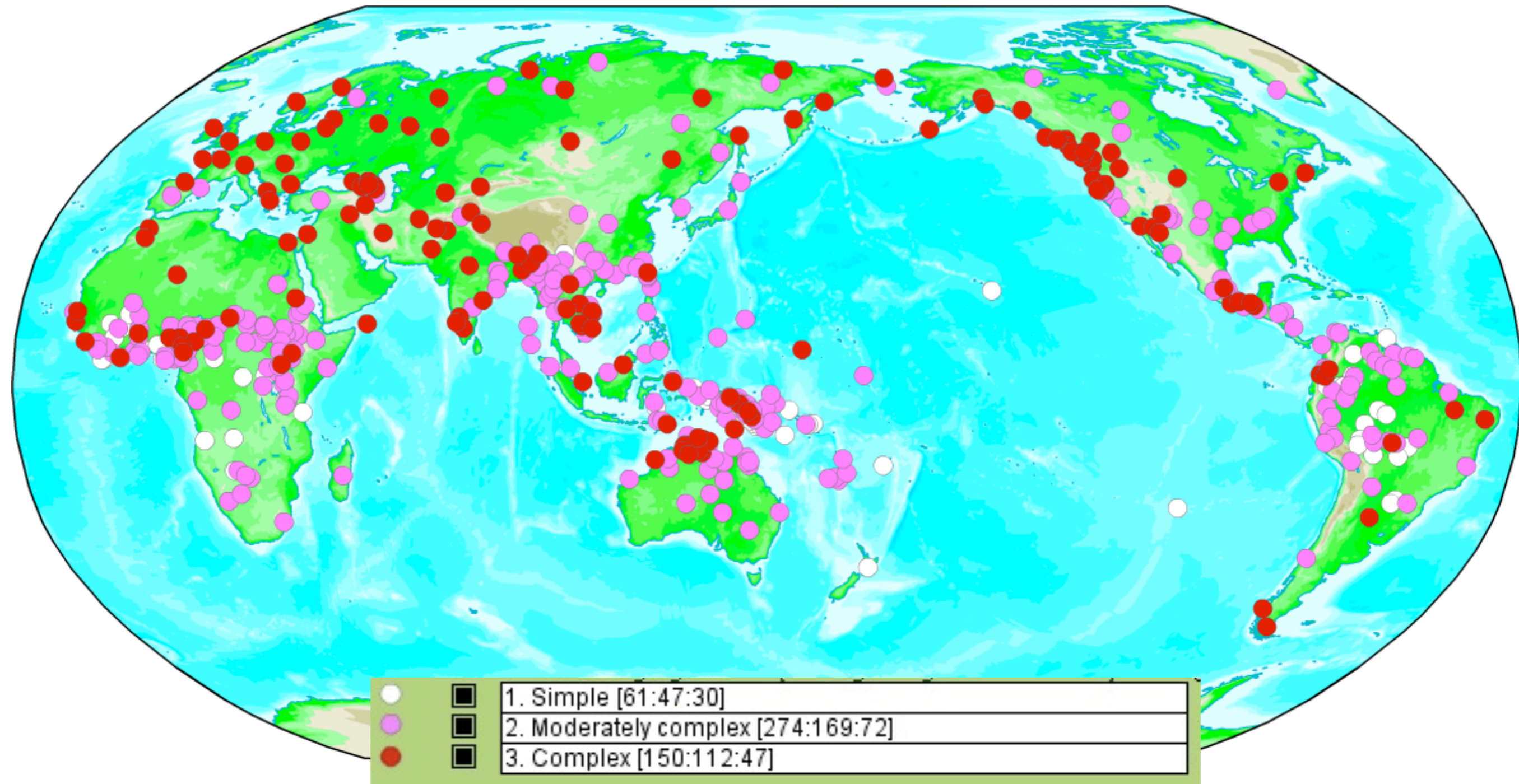
Map for Chapter 141

by robert - Feb 03, 2009

Today, we got a step further in filling the gap between WALs Online and the printed edition of 2005. We incorporated data to be ...

wals.info

Syllable Structure



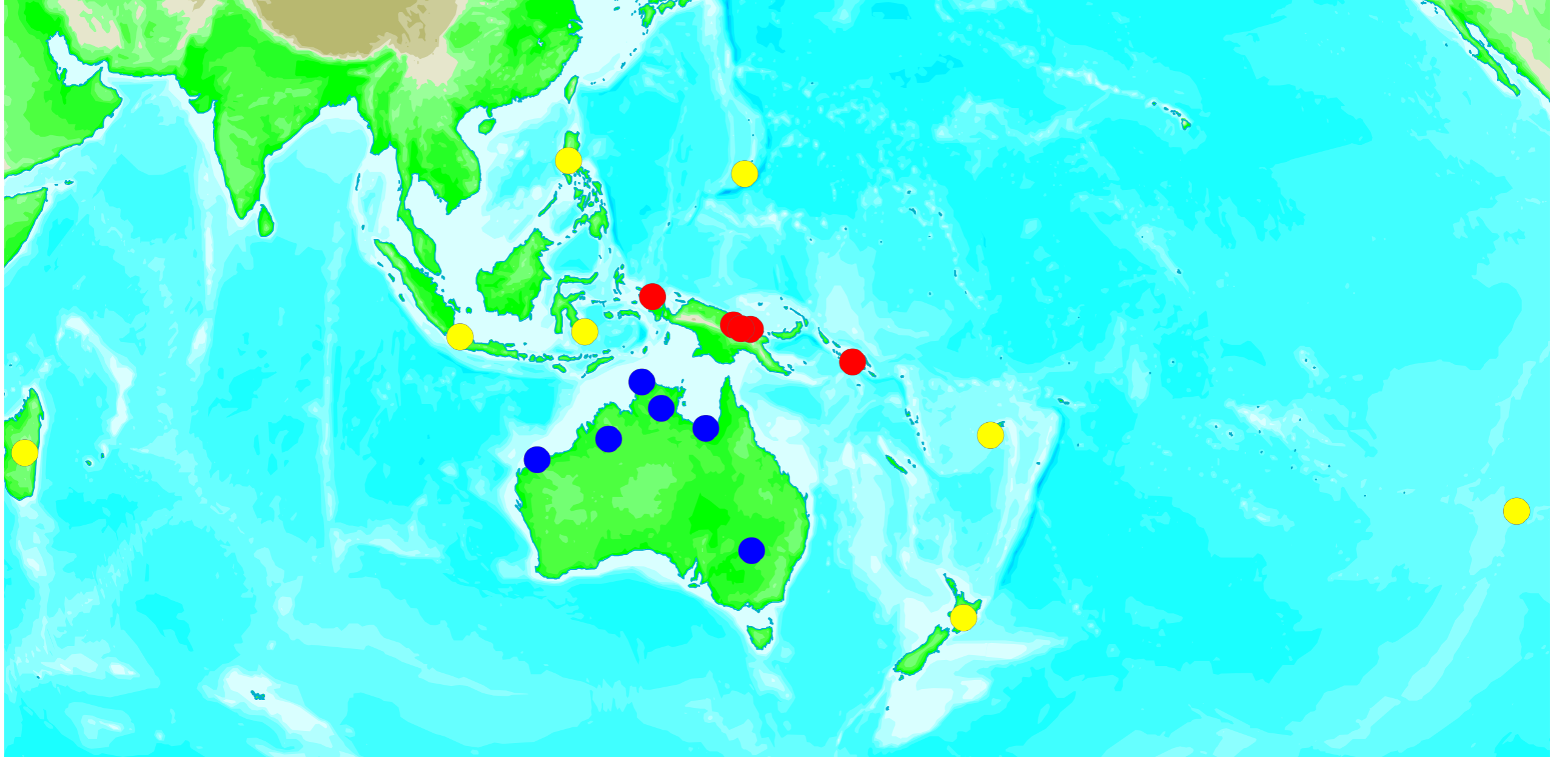
Maddieson, Ian (2005) 'Syllable structure' in: Martin Haspelmath, Matthew S. Dryer, David Gil, & Bernard Comrie (eds.) *World Atlas of Language Structures*. Oxford: Oxford University Press, 54-57.

Questions

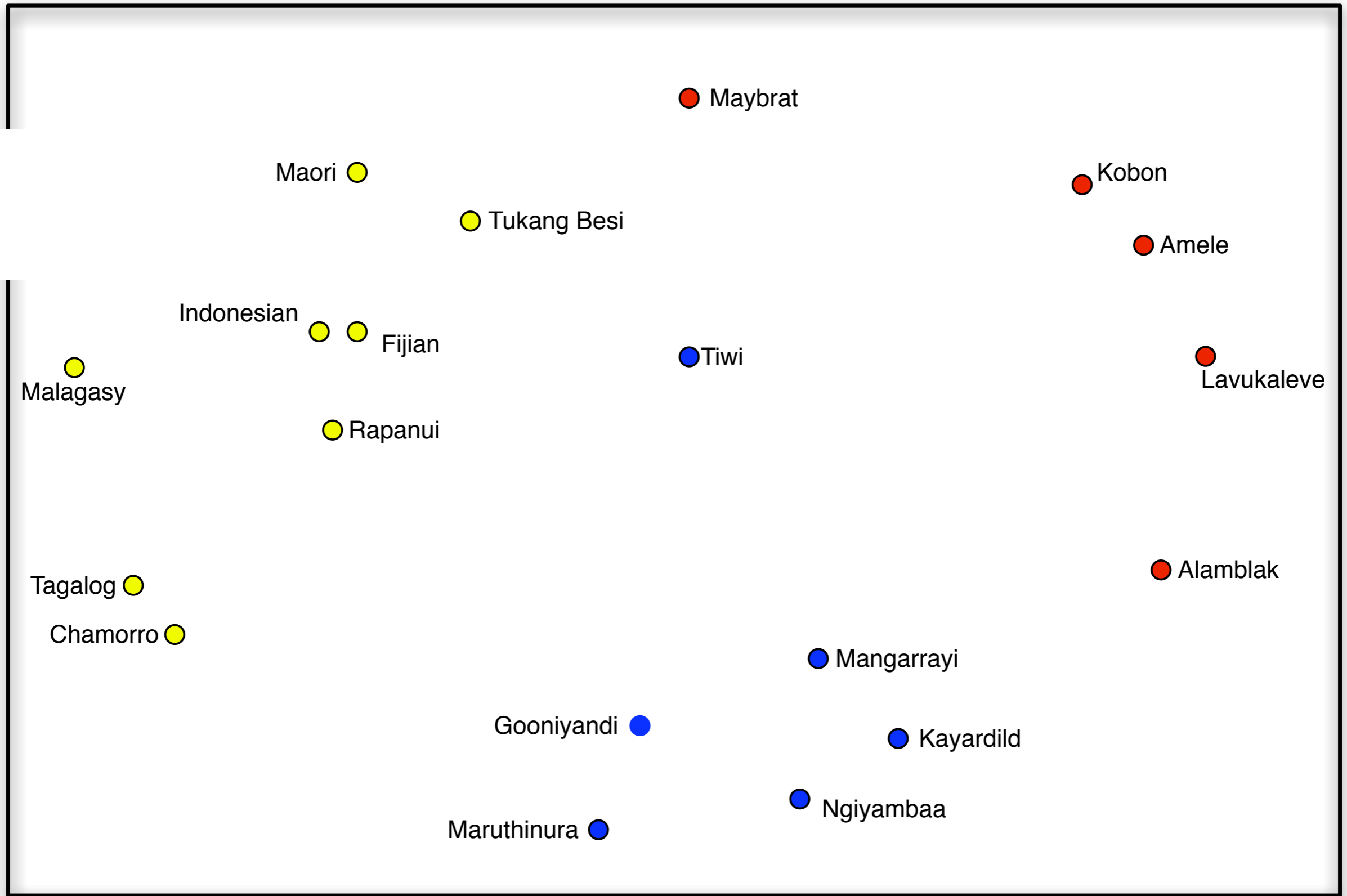
- How strong is the influence of geographic distribution of languages on their linguistic profile?
- How can the geographical distribution of language be suitably displayed?

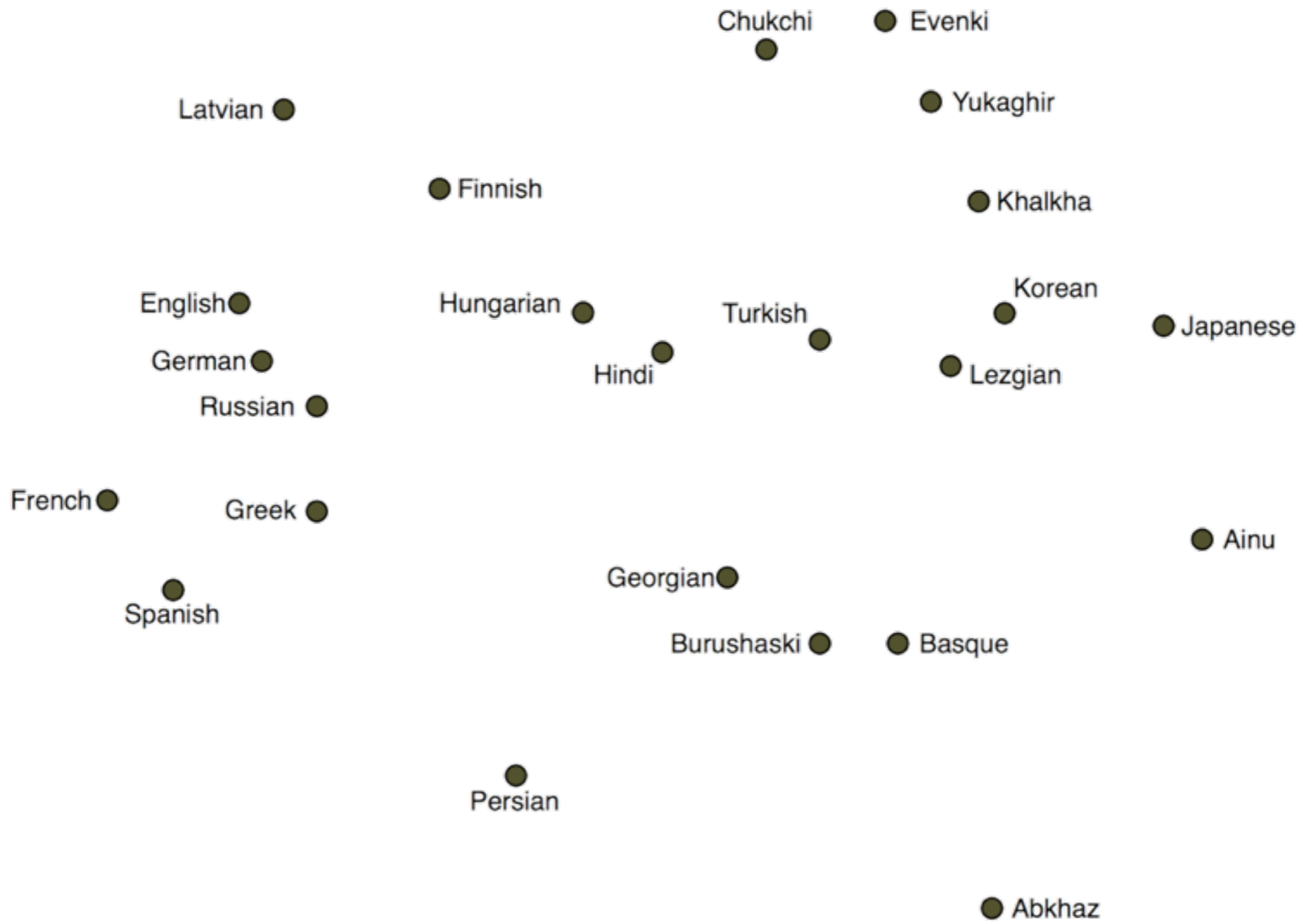
Questions

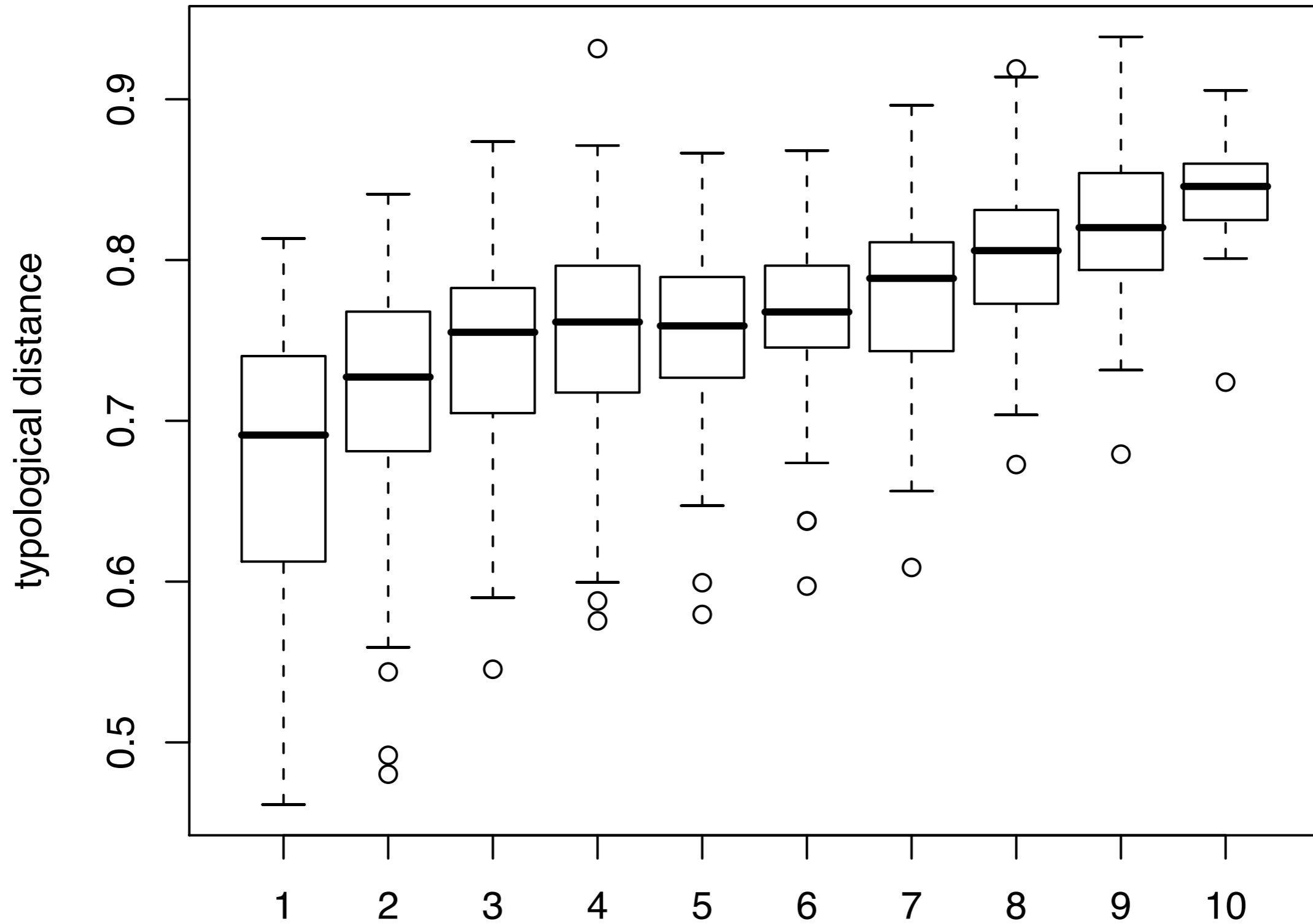
- How strong is the influence of geographic distribution of languages on their linguistic profile?
- How can the geographical distribution of language be suitably displayed?





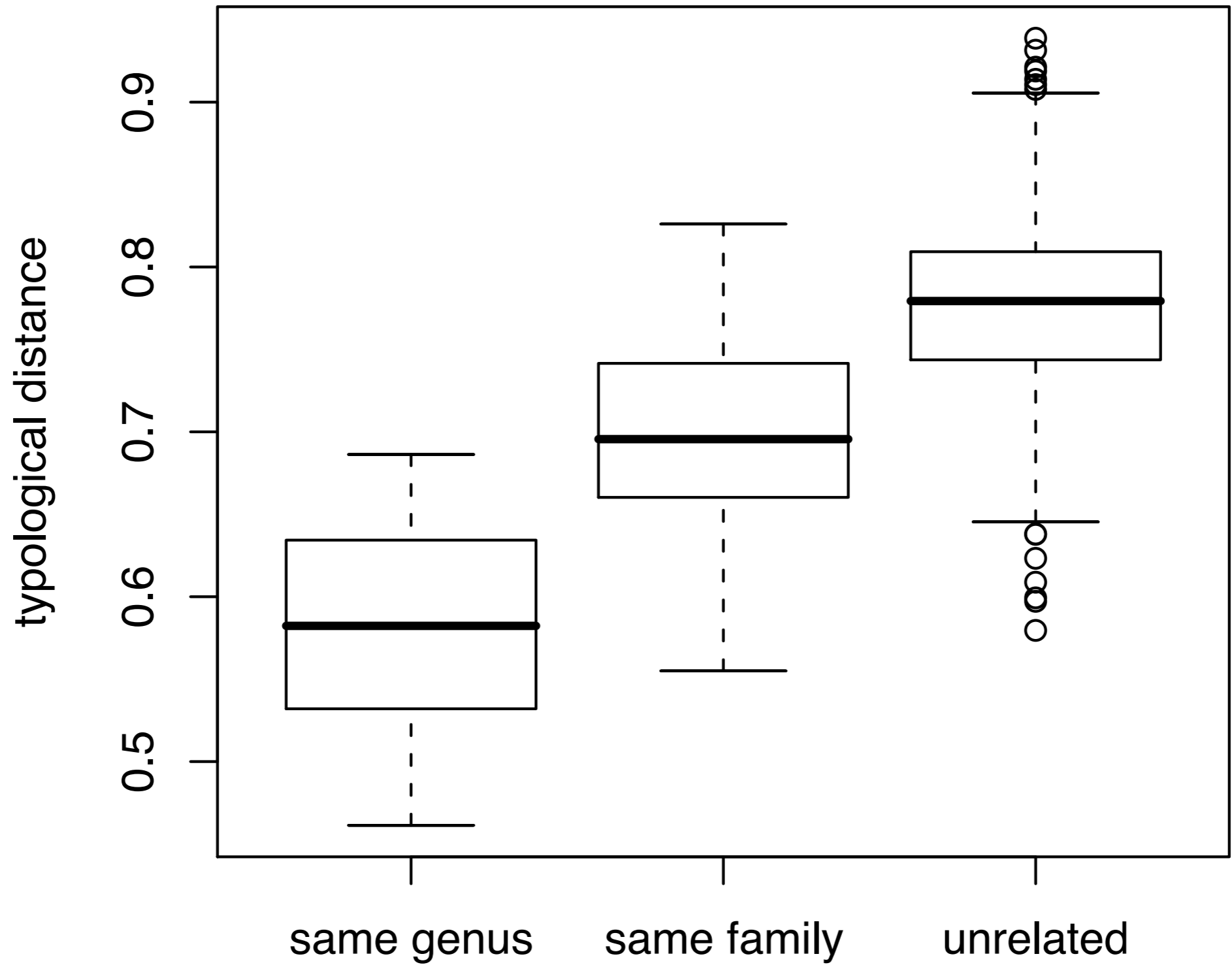






geographical distance (in bins of 1000 km)

Pearson $r = .52$
Mantel $p < .001$



Pearson $r = .61$
Mantel $p < .001$

Mantel Test

| Correlation | Pearson r | Mantel p |
|-----------------------|-------------|------------|
| Typology ~ Geography | 0.52 | < .001 |
| Typology ~ Genealogy | 0.61 | < .001 |
| Geography ~ Genealogy | 0.33 | < .001 |

Partial Mantel Test

| Correlation | Pearson r | Mantel p |
|----------------------------------|-------------|------------|
| Typology ~ Geography + Genealogy | 0.42 | < .001 |
| Typology ~ Genealogy + Geography | 0.54 | < .001 |

Multivariate Matrix Regression

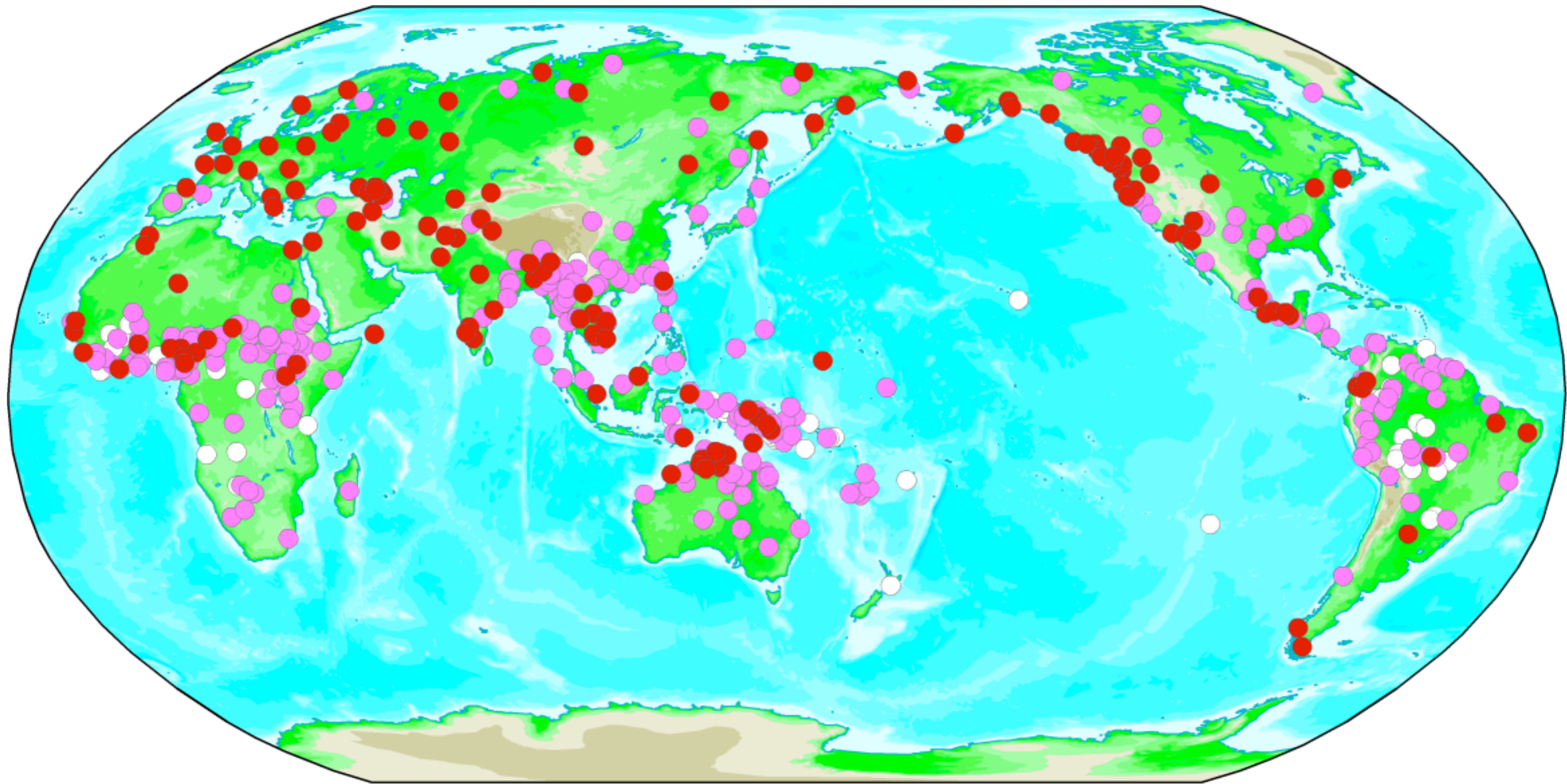
| | Sums of Sqs | Mean Sqs | F Model | R ² |
|----------------------|-------------|----------|---------|----------------|
| family : genus | 13.65 | 0.02 | -3.29 | 0.77 |
| latitude : longitude | 0.28 | 0.28 | -39.01 | 0.02 |
| Residuals | 3.82 | -0.01 | | 0.22 |

Zapala, M.A. and J. Schork (2006) Multivariate regression analysis of distance matrices for testing associations between gene expression patterns and related variables. PNAS 103(51): 19430–19435

Questions

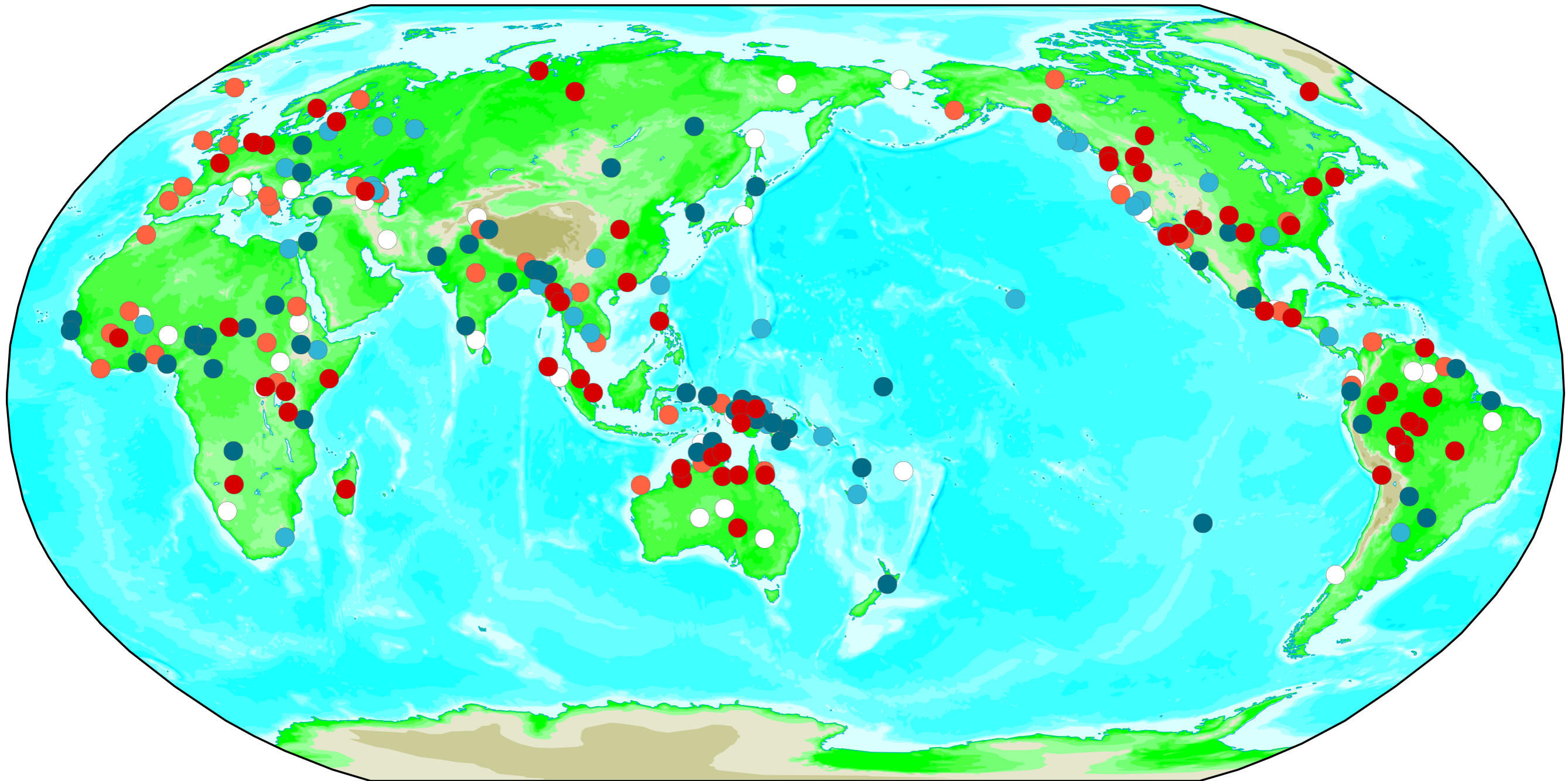
- How strong is the influence of geographic distribution of languages on their linguistic profile?
- How can the geographical distribution of language be suitably displayed?

Syllable Structure



Maddieson, Ian (2005) 'Syllable structure' in: Martin Haspelmath, Matthew S. Dryer, David Gil, & Bernard Comrie (eds.) *World Atlas of Language Structures*. Oxford: Oxford University Press, 54-57.

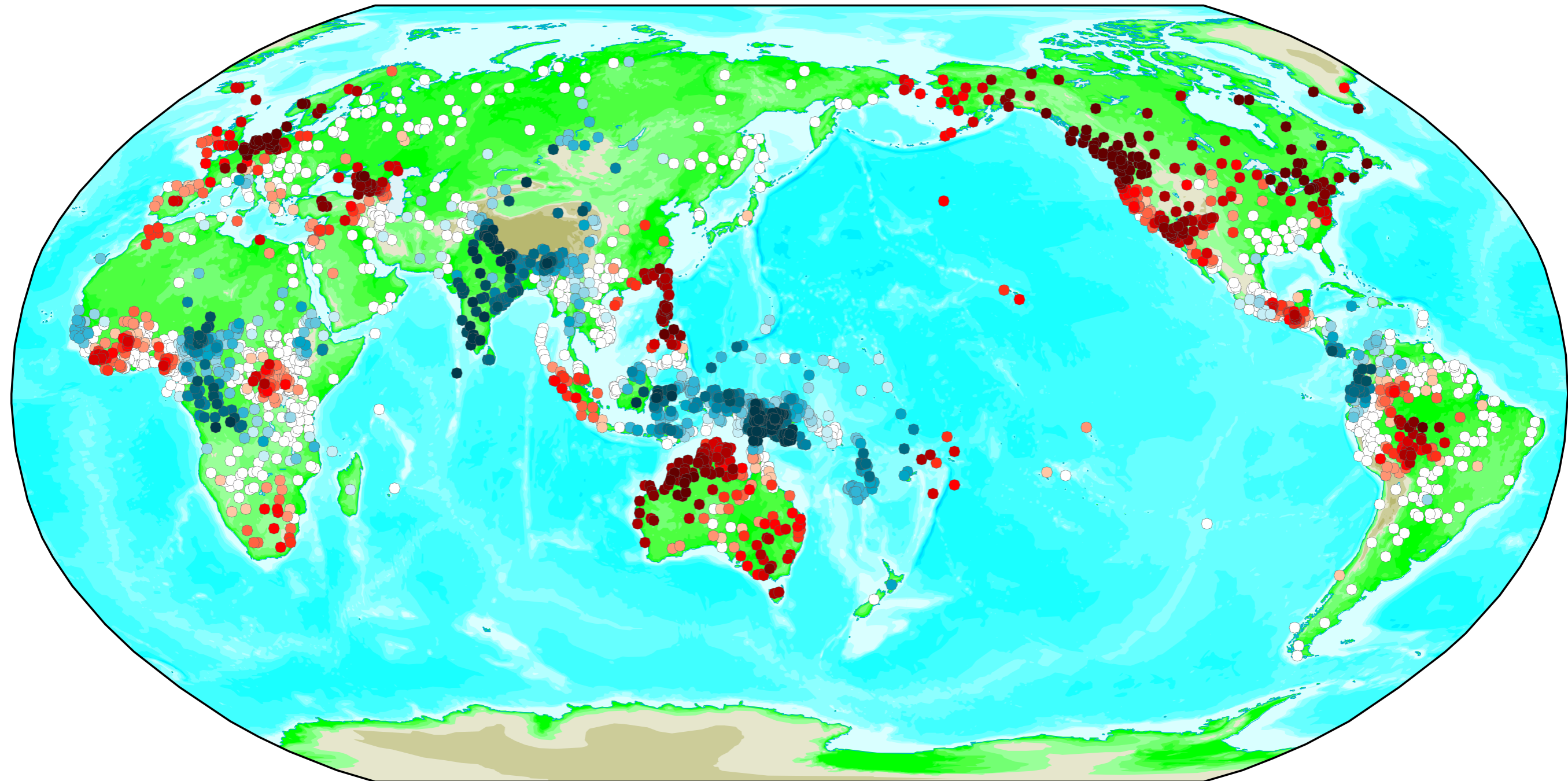
Average Rarity



few rare characteristics

many rare characteristics

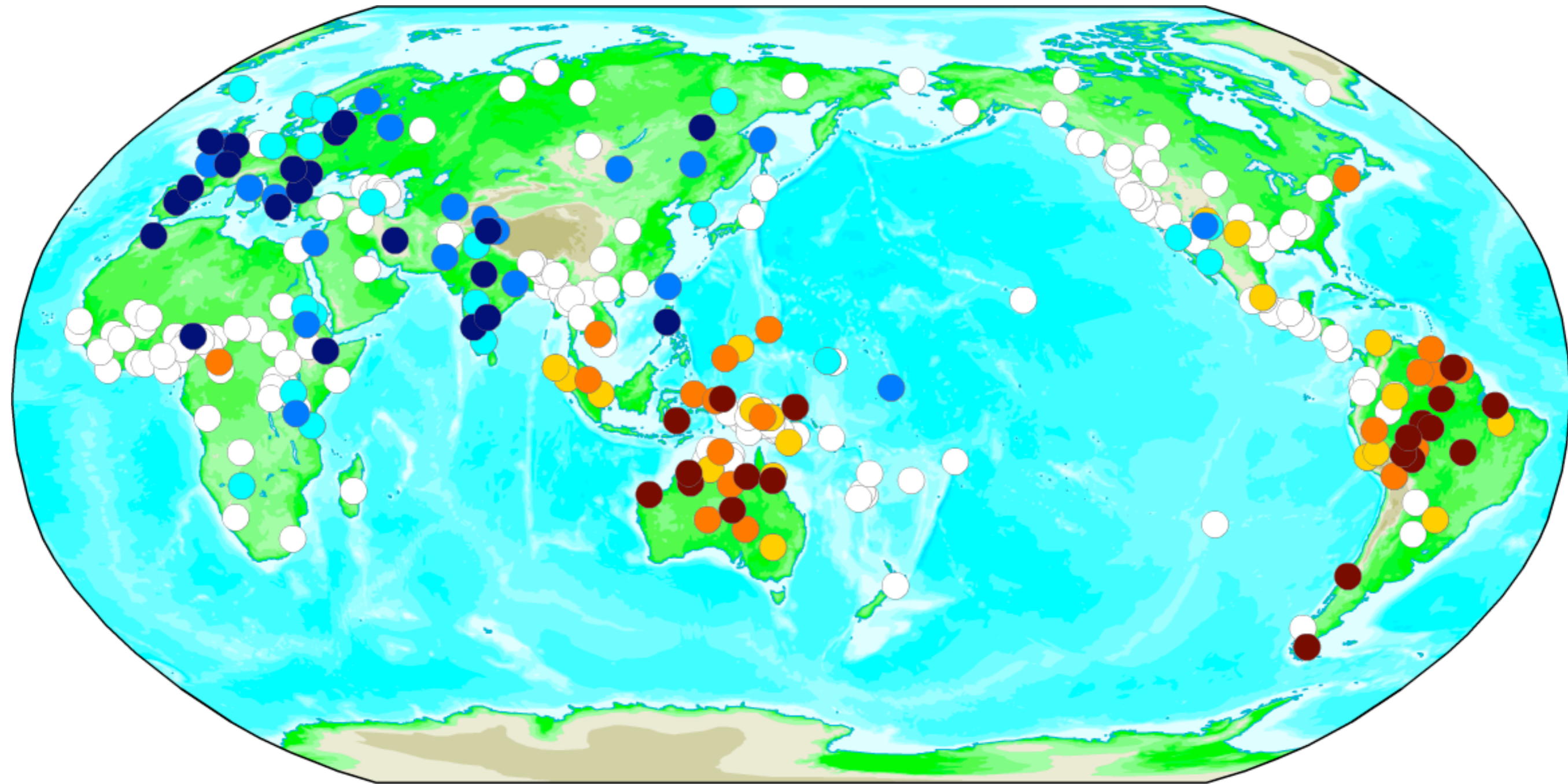
Average Rarity



few rare characteristics

many rare characteristics

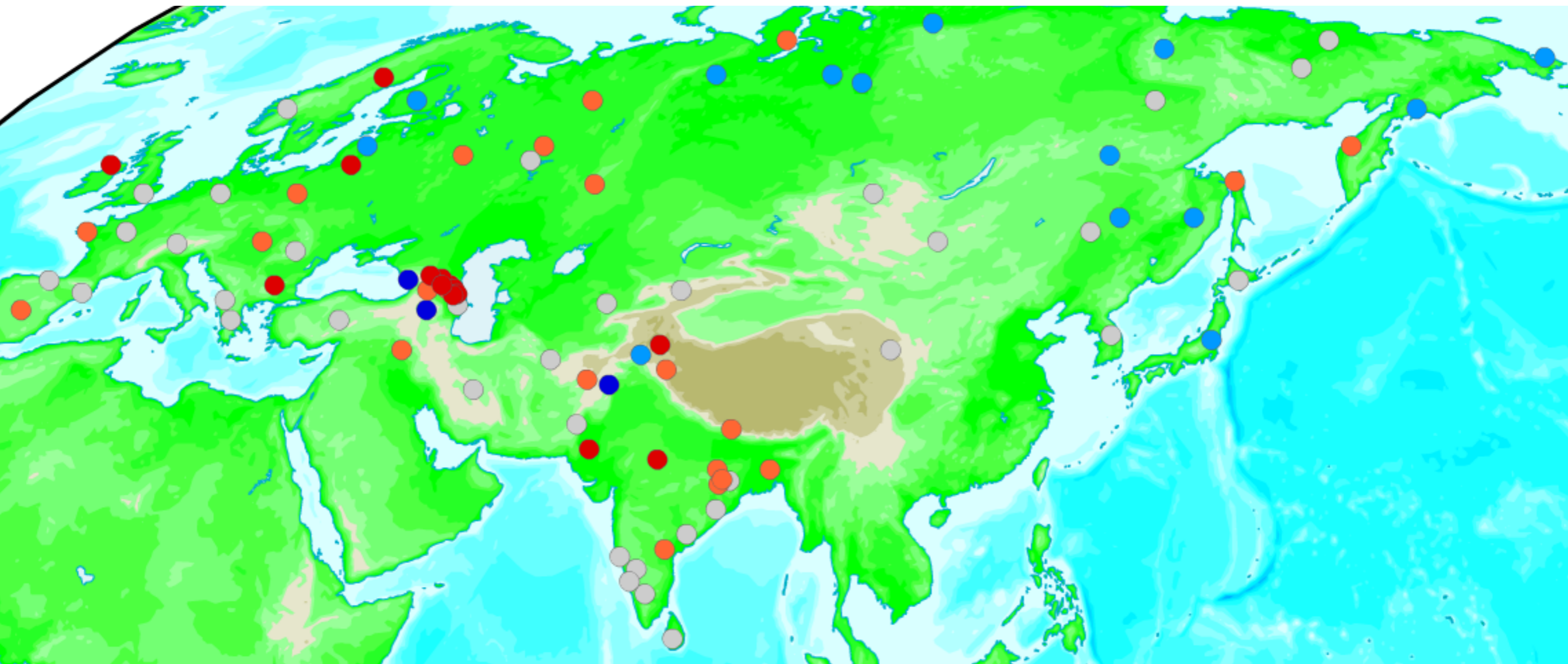
Average Complexity



high complexity

low complexity

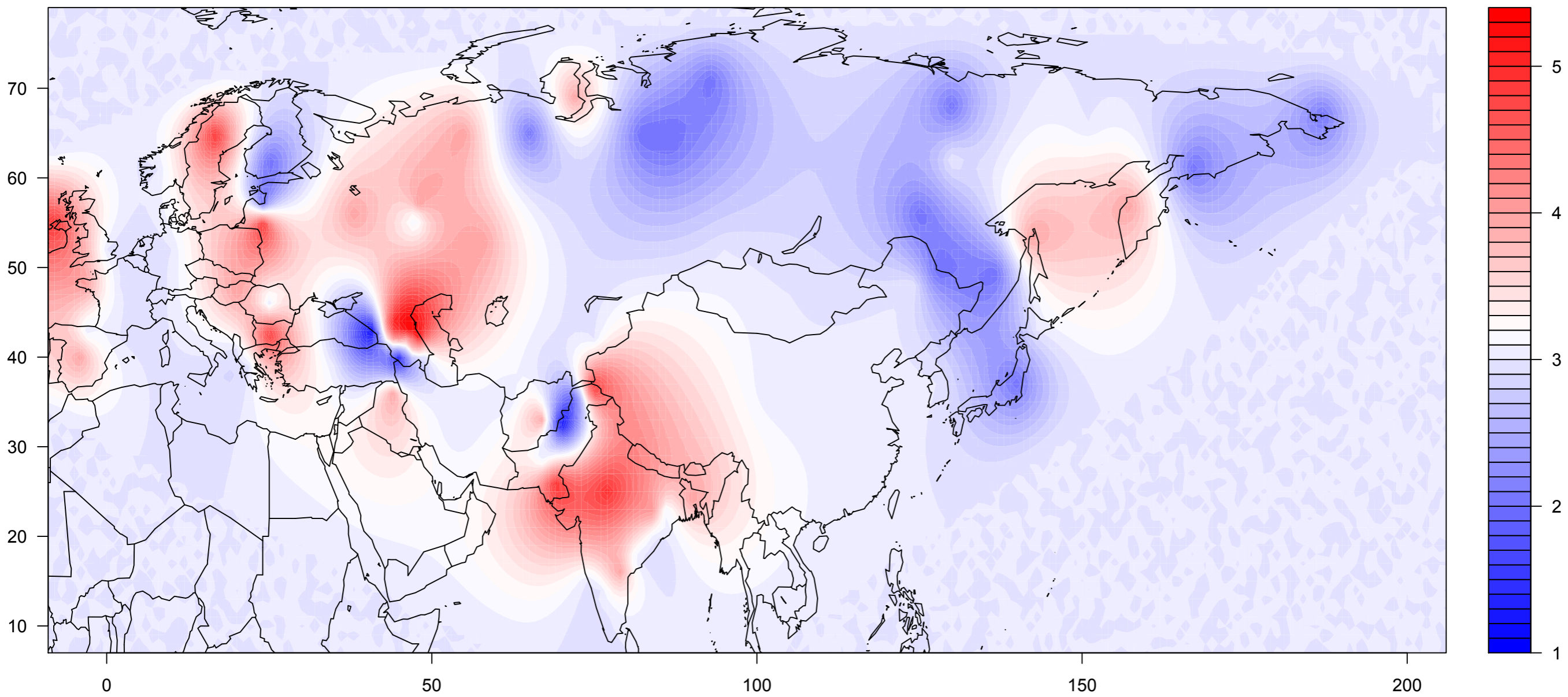
Consonant Inventories

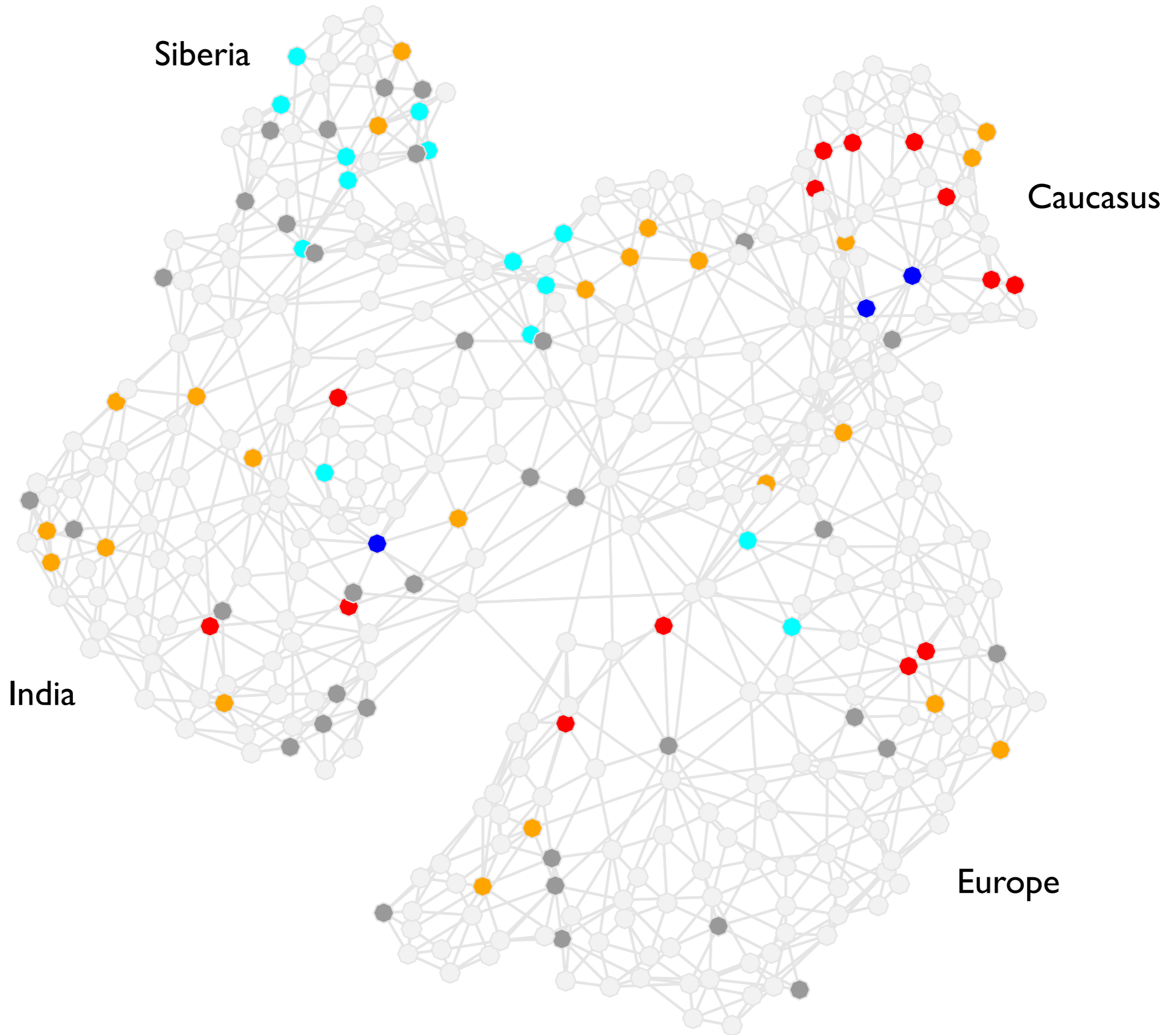


few consonants

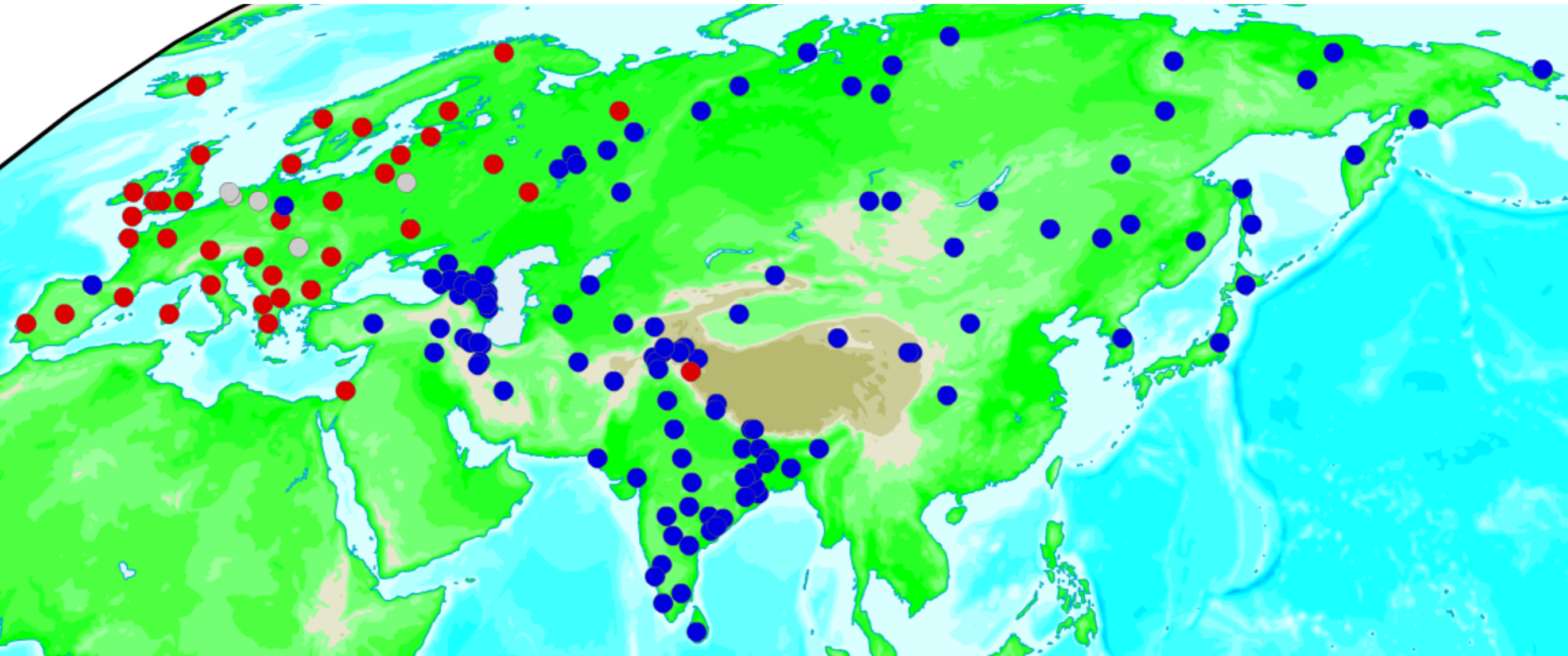
many consonants

Maddieson, Ian (2005) 'Consonant Inventories' in: Martin Haspelmath, Matthew S. Dryer, David Gil, & Bernard Comrie (eds.) *World Atlas of Language Structures*. Oxford: Oxford University Press.





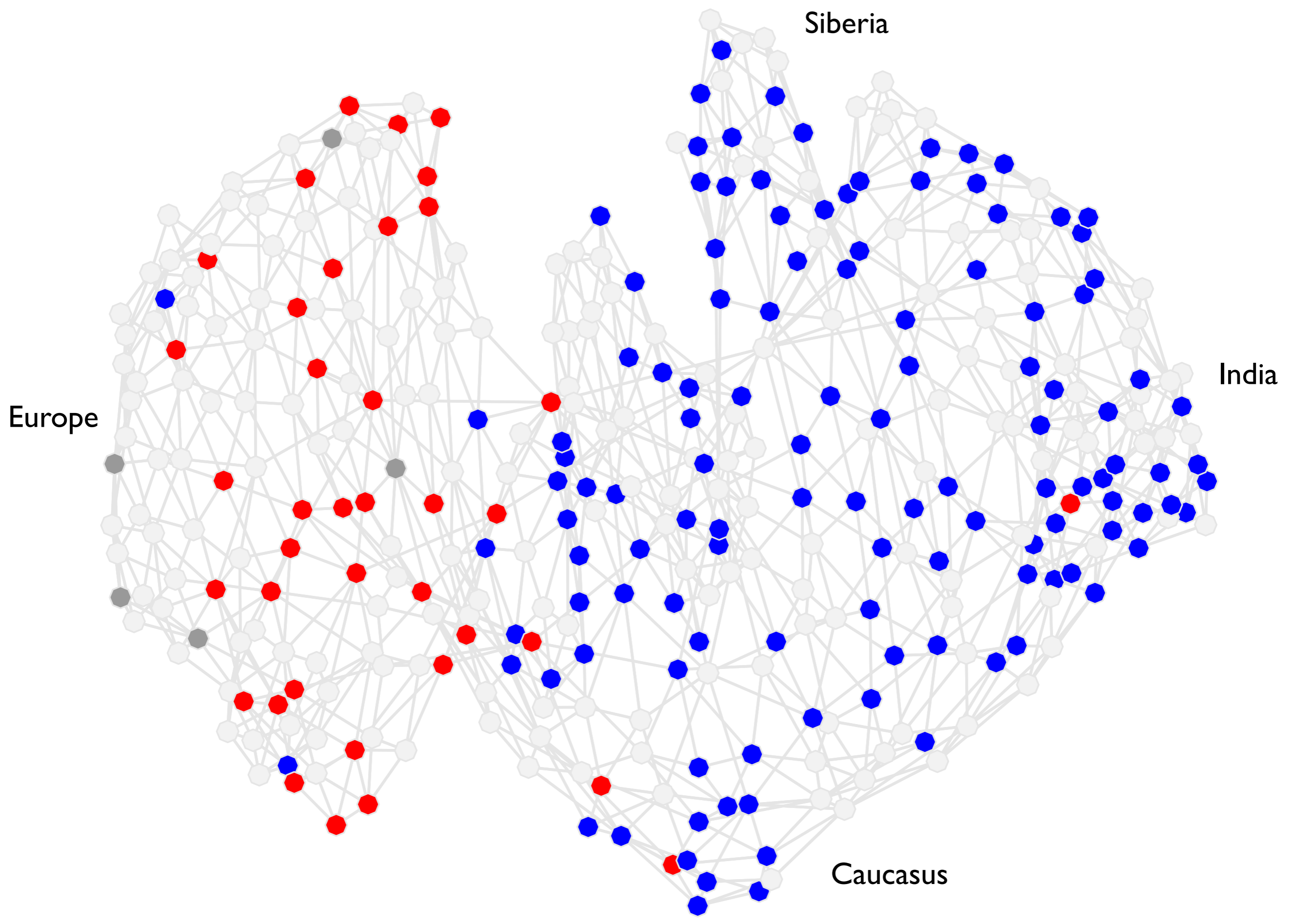
Order of Object and Verb



he sees the house
(Verb-Object)

he the house sees
(Object-Verb)

Dryer, Matthew S. (2005) 'Order of Object and Verb' in: Martin Haspelmath, Matthew S. Dryer, David Gil, & Bernard Comrie (eds.) *World Atlas of Language Structures*. Oxford: Oxford University Press.

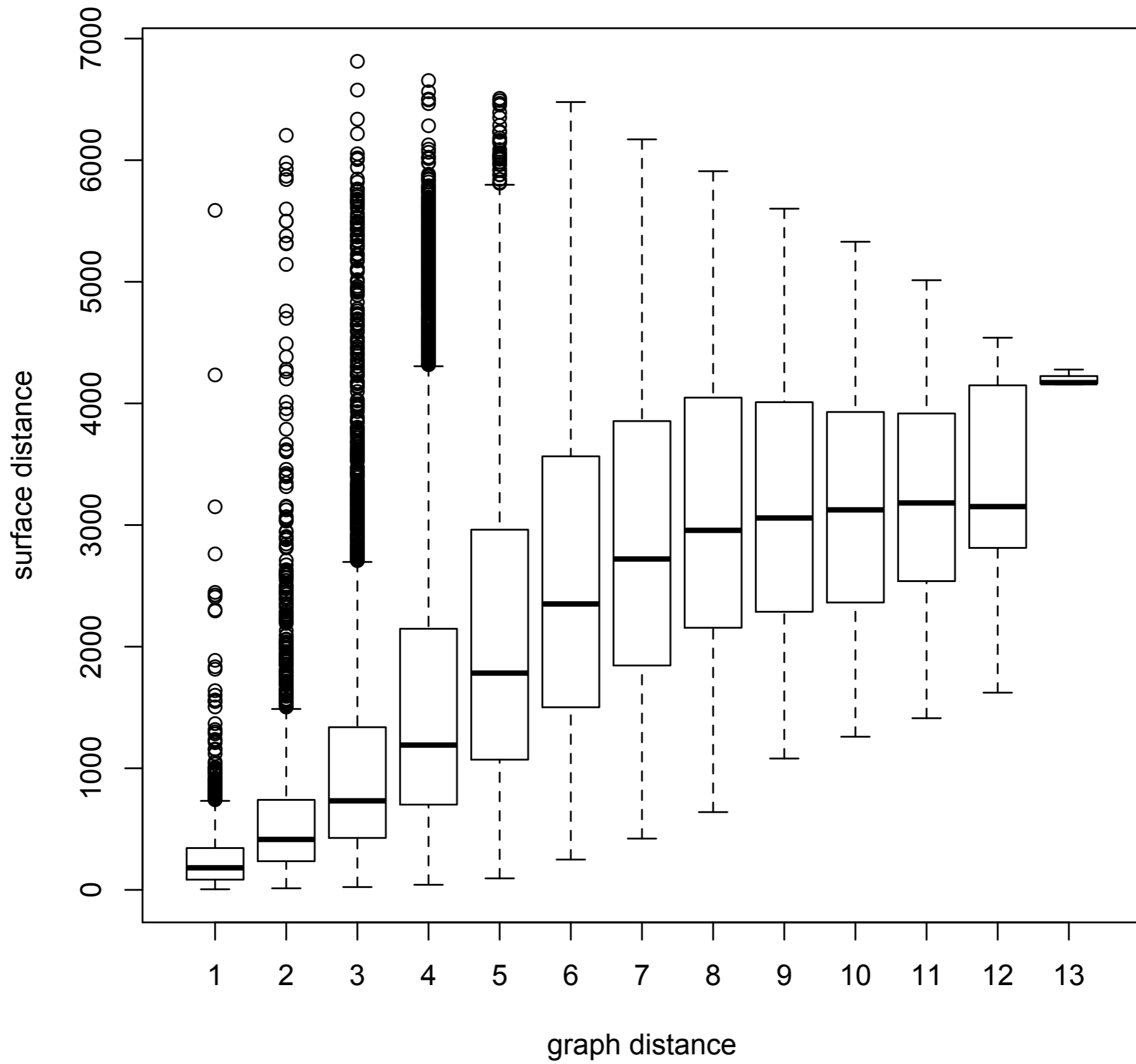


Europe

Siberia

India

Caucasus



Conclusions

- Geographical distance can be used as a proxy to the probability of social contact
- Geographical distance seems to be correlated with language similarity
- Geographical display of linguistic diversity can (and should) be improved
- Linguists need some help with all of this!