Worldwide Geographical Clines of Complexity

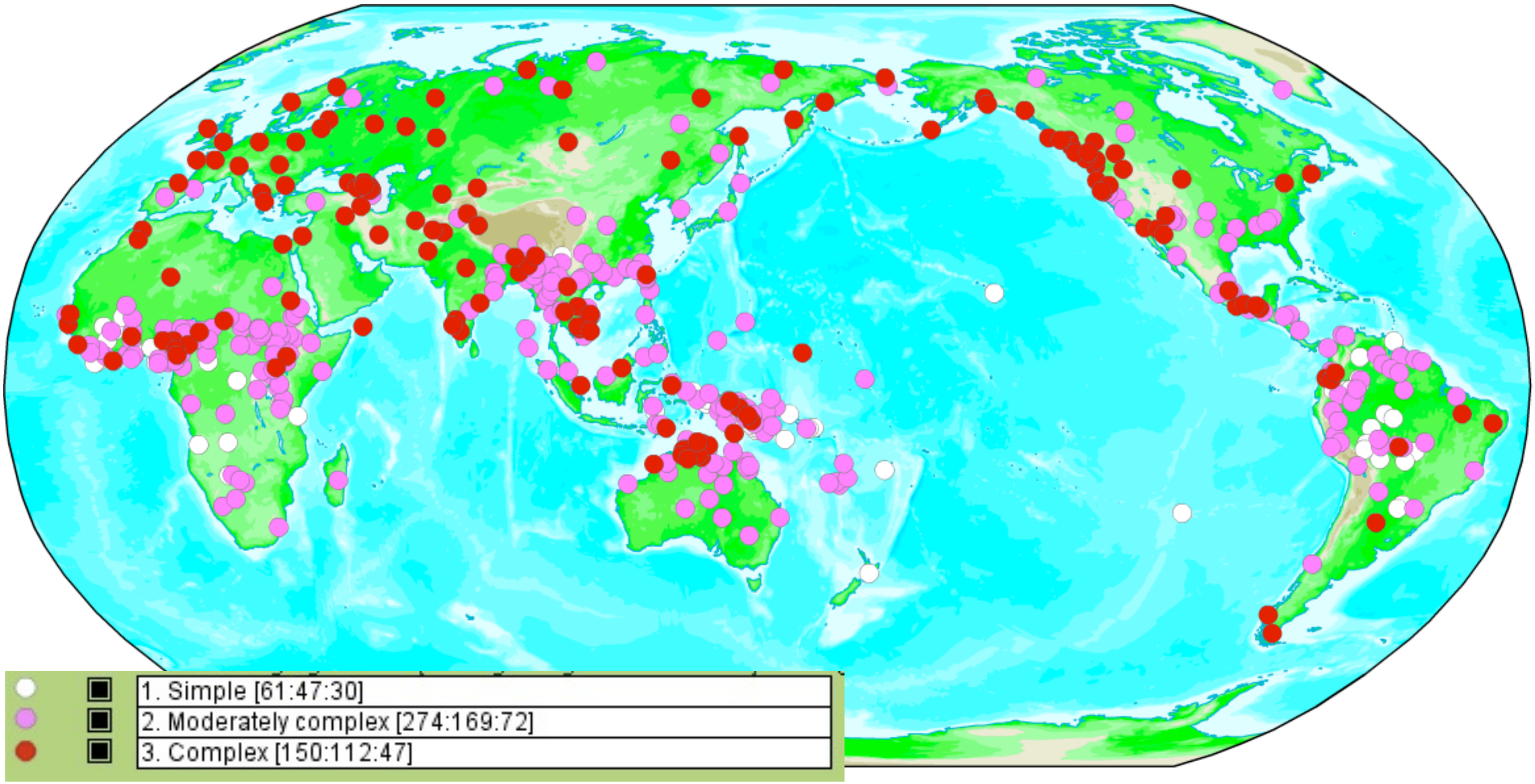
Michael Cysouw
Max Planck Institute for Evolutionary Anthropology

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

Haspelmath, Dryer, Gil, Comrie (eds.) 2005. The World Atlas of Language Structures. Oxford: OUP.

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.


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.
Search
powered by Google
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 ...
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 OSX, Mac OS9.2 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.

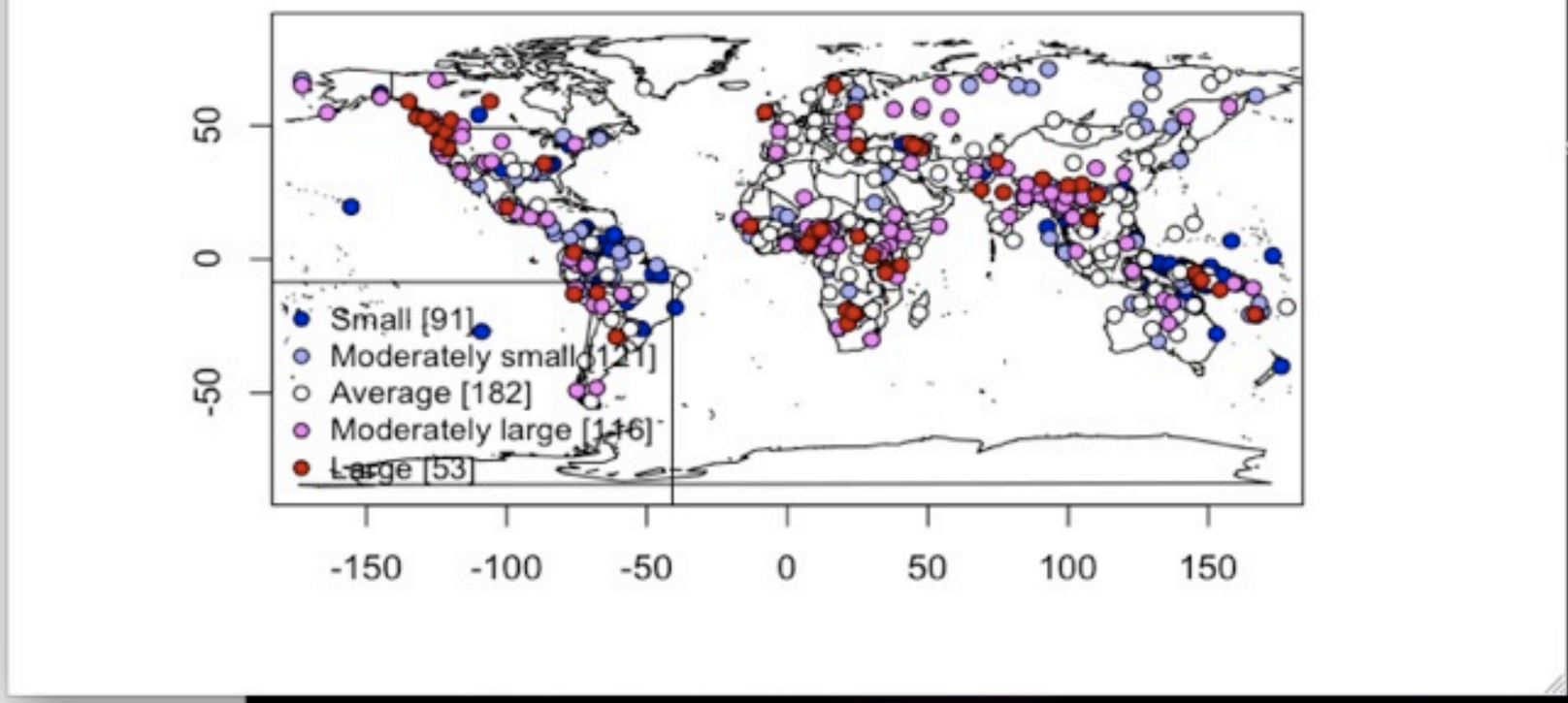
wals.info

R Console

~/Documents/Linguistics/Scientific Production/5. Talks/

```
R version 2.9.0 (2009-04-17)
Copyright (C) 2009 The R Foundation for Statistical Computing
ISBN 3-900051-07-0

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

  Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

[R.app GUI 1.28 (5395) i386-apple-darwin8.11.1]

> source("/Applications/WALS_R/wals.R",chdir=T)
> showMap(1,cex=1)
> table(mat[,1],mat[,2])

    1  2  3
  1 11 51 29
  2 19 71 31
  3 28 80 73
  4 22 53 40
  5 11 31 10
> chisq.test(table(mat[,1],mat[,2]))

    Pearson's Chi-squared test

data:  table(mat[, 1], mat[, 2])
X-squared = 15.3455, df = 8, p-value = 0.05277

> round(chisq.test(table(mat[,1],mat[,2]))$residuals,digits=1)

     1    2    3
  1 -1.0  0.7 -0.1
  2 -0.1  1.2 -1.4
  3 -0.3 -1.3  1.8
  4  0.8 -0.7  0.4
  5  0.9  0.9 -1.7
>
```

Quartz 2 [*]



Complexity

- Many different possible definitions
- Here: a simple and easily applicable approach
- **More oppositions** and **strict structure** is interpreted as more complex

Feature/Chapter 1: Consonant Inventories

by Ian Maddieson

This chapter consists of the map data (show map) and the chapter text.

Map Configuration

You may configure the appearance of the map using the tools below and then hit show map

Select icons for the values [help]

- Small (91 languages)
- Moderately small (121 languages)
- Average (182 languages)
- Moderately large (116 languages)
- Large (53 languages)

total: 563

Feature/Chapter 42: Pronominal and Adnominal Demonstratives

by Holger Diessel

This chapter consists of the map data (show map) and the chapter text.

Map Configuration

You may configure the appearance of the map using the tools below and then hit show map

Select icons for the values [help]

- Identical (143 languages)
- Different stem (37 languages)
- Different inflection (21 languages)

total: 201

The World Atlas of Language Structures Online

Home Features

Feature/Chapter 83: Order of Object and Verb

by Matthew S. Dryer

This chapter consists of the map data (show map) and the chapter text.

Map Configuration

You may configure the appearance of the map using the tools below and then hit show map

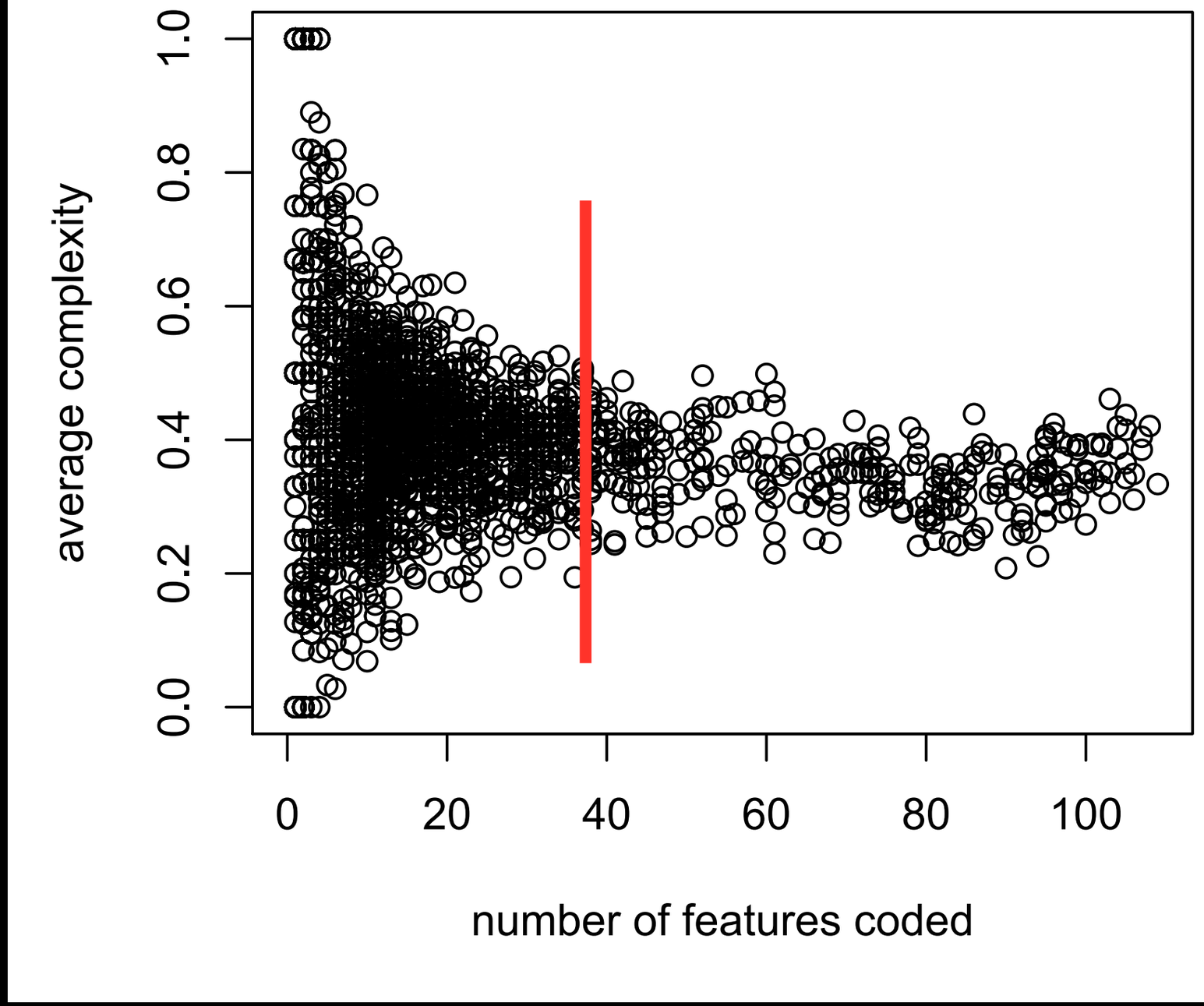
Select icons for the values [help]

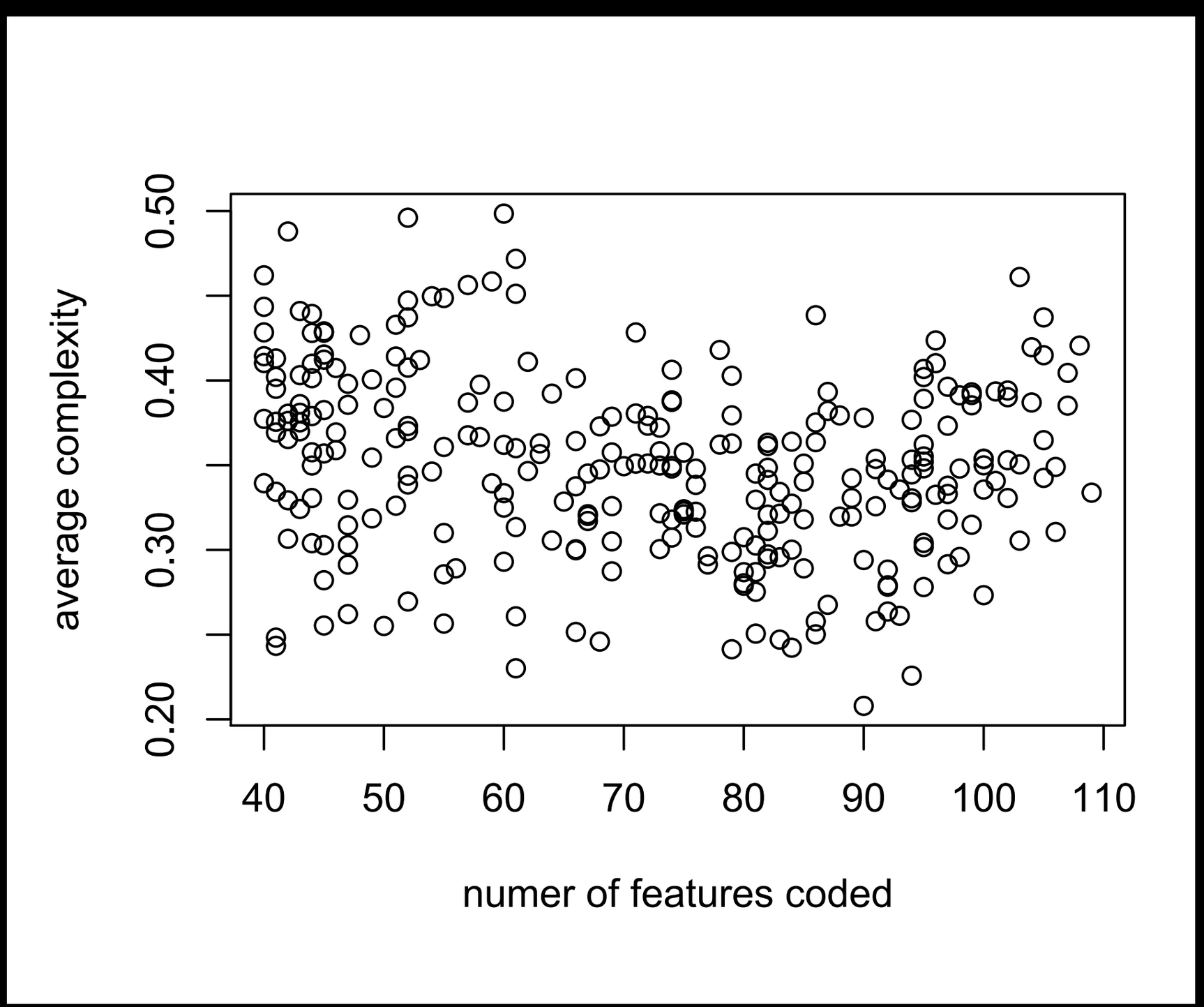
- OV (640 languages)
- VO (640 languages)
- No dominant order (90 languages)

total: 1370

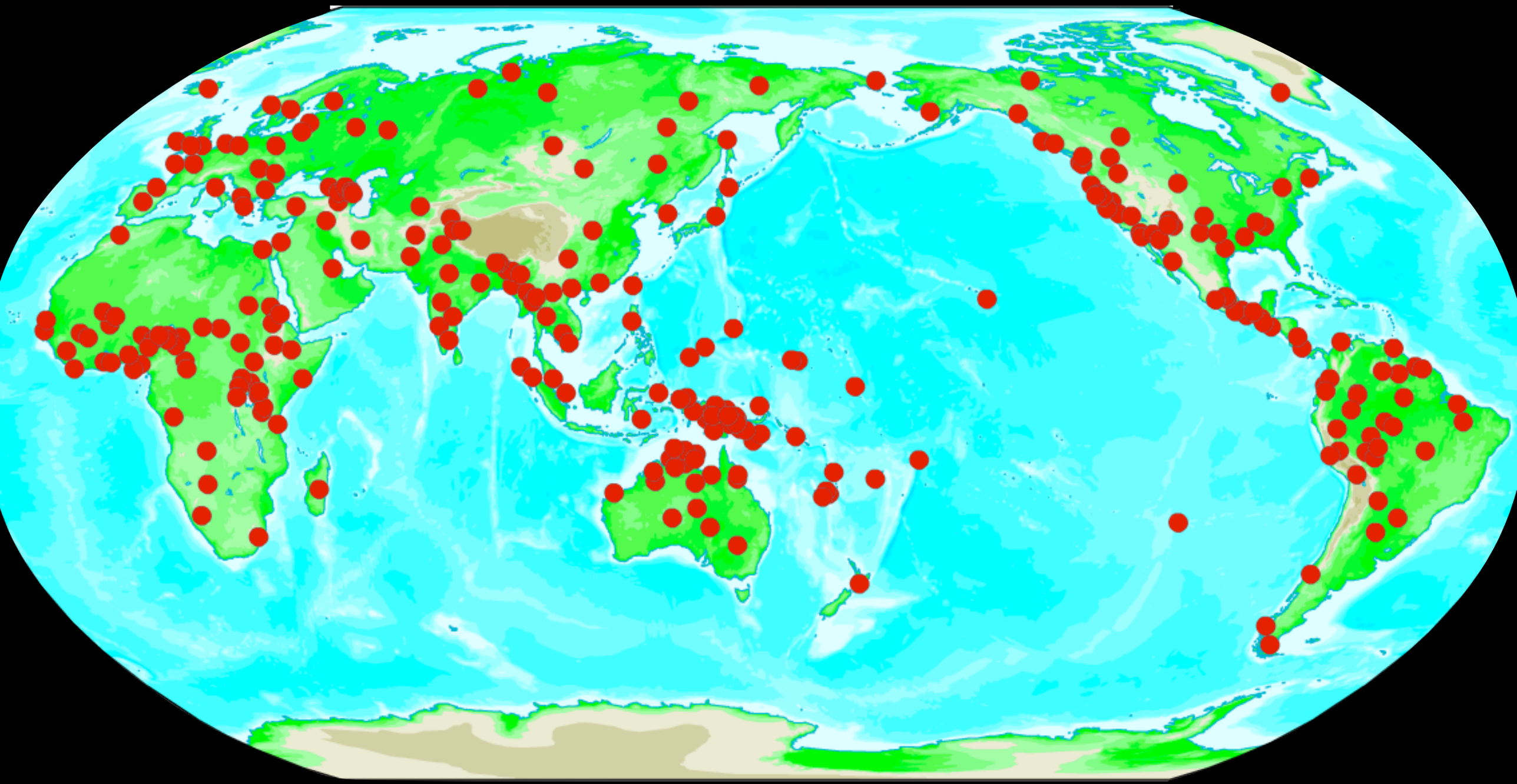
Average Complexity

- Average complexity is easily computed for each language in WALS
- However: WALS is notoriously incomplete
- An ad-hoc selection of languages is necessary


average complexity
0.0
0.2
0.4
0.6
0.8
1.0
0
20
40
60
80
100
number of features coded


average complexity
0.20
0.30
0.40
0.50
40
50
60
70
80
90
100
110
numer of features coded

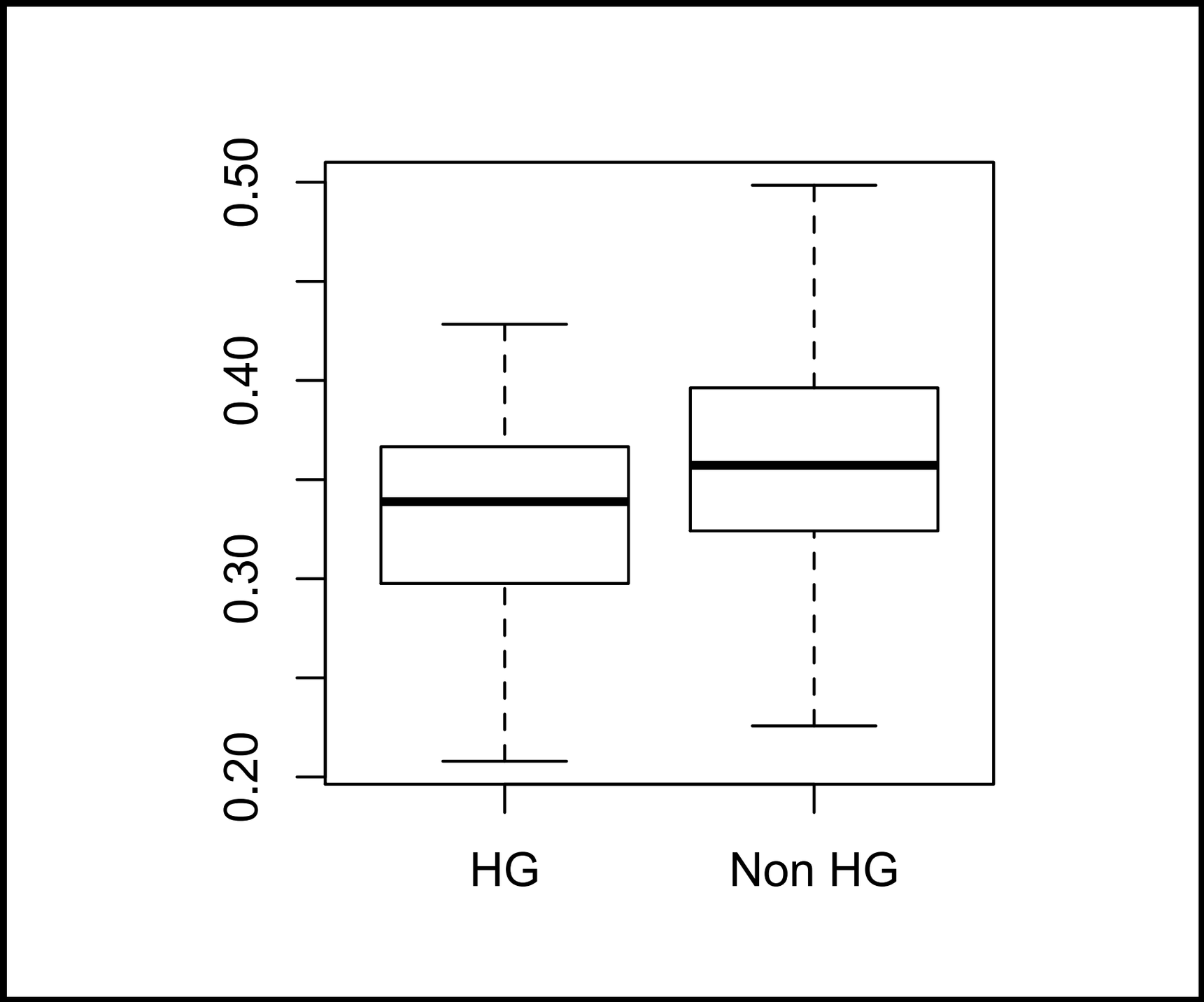
280 language sample

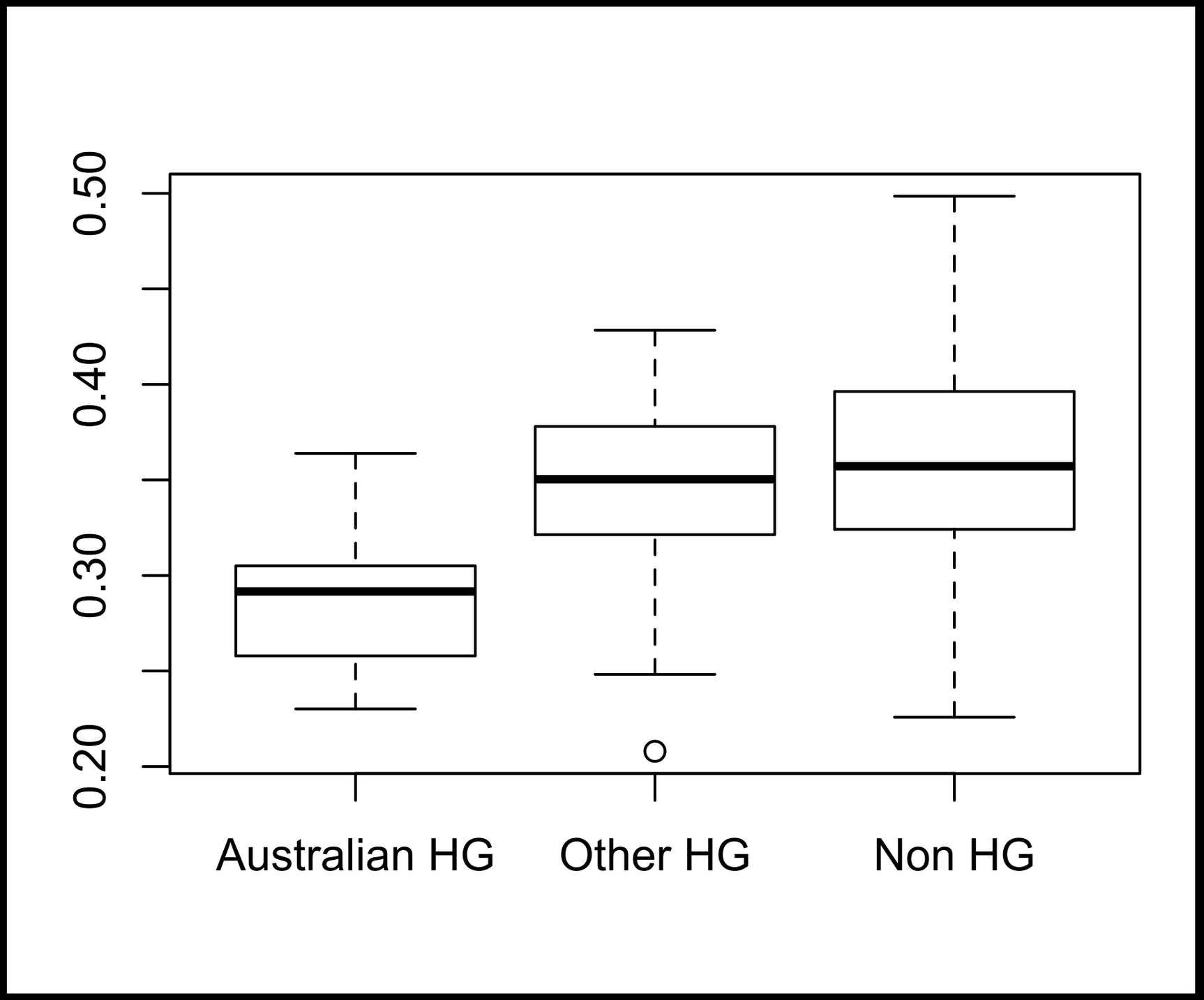


Subsistence

- Is there any relation between the complexity of a language and cultural factors?
- Specifically here: is there a correlation with hunter-gatherer lifestyle?
- List of languages spoken by hunter-gatherer communities as prepared by Güldemann et al. (in press)

Güldemann, Tom, Patrick McConvell and Richard Rhodes (eds.). In press. Hunter-gatherers and Linguistic History: A Global Perspective. Cambridge University Press.


0.50
0.40
0.30
0.20
HG
Non HG


0.50
0.40
0.30
0.20
Australian HG
Other HG
Non HG

Preferences of hunter-gatherer languages

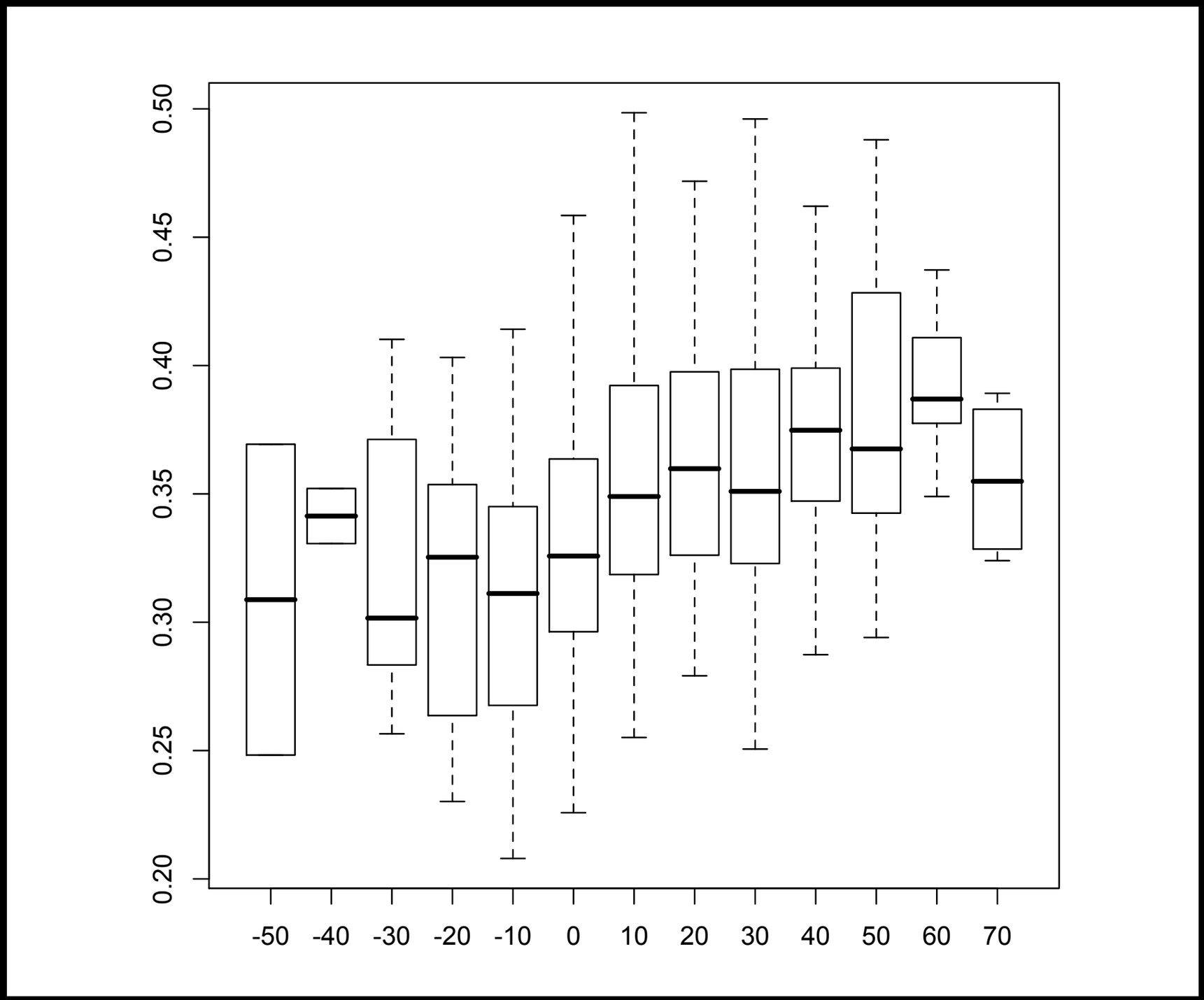
- No dominant constituent order
- No adpositions
- Subject clitics on variable hosts
- Initial interrogatives

Cysouw, Michael & Bernard Comrie. In press. Some observations on typological features of hunter-gatherer languages.

Geography

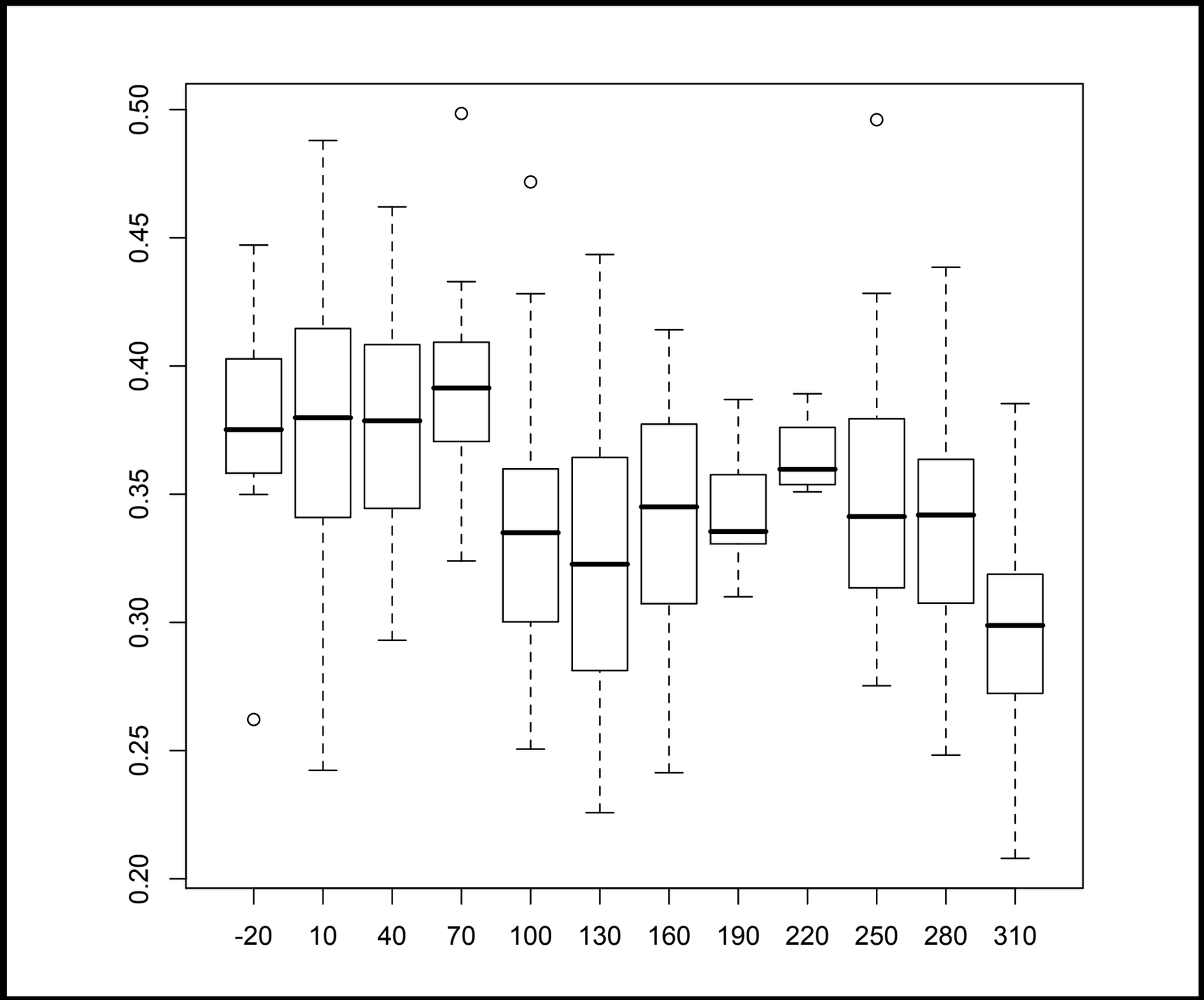
- Is there any relation between the complexity of a language and geography?
- The study of geographical distribution of linguistic variation is still in its infancy
- The following notes are thus somewhat preliminary

Latitude



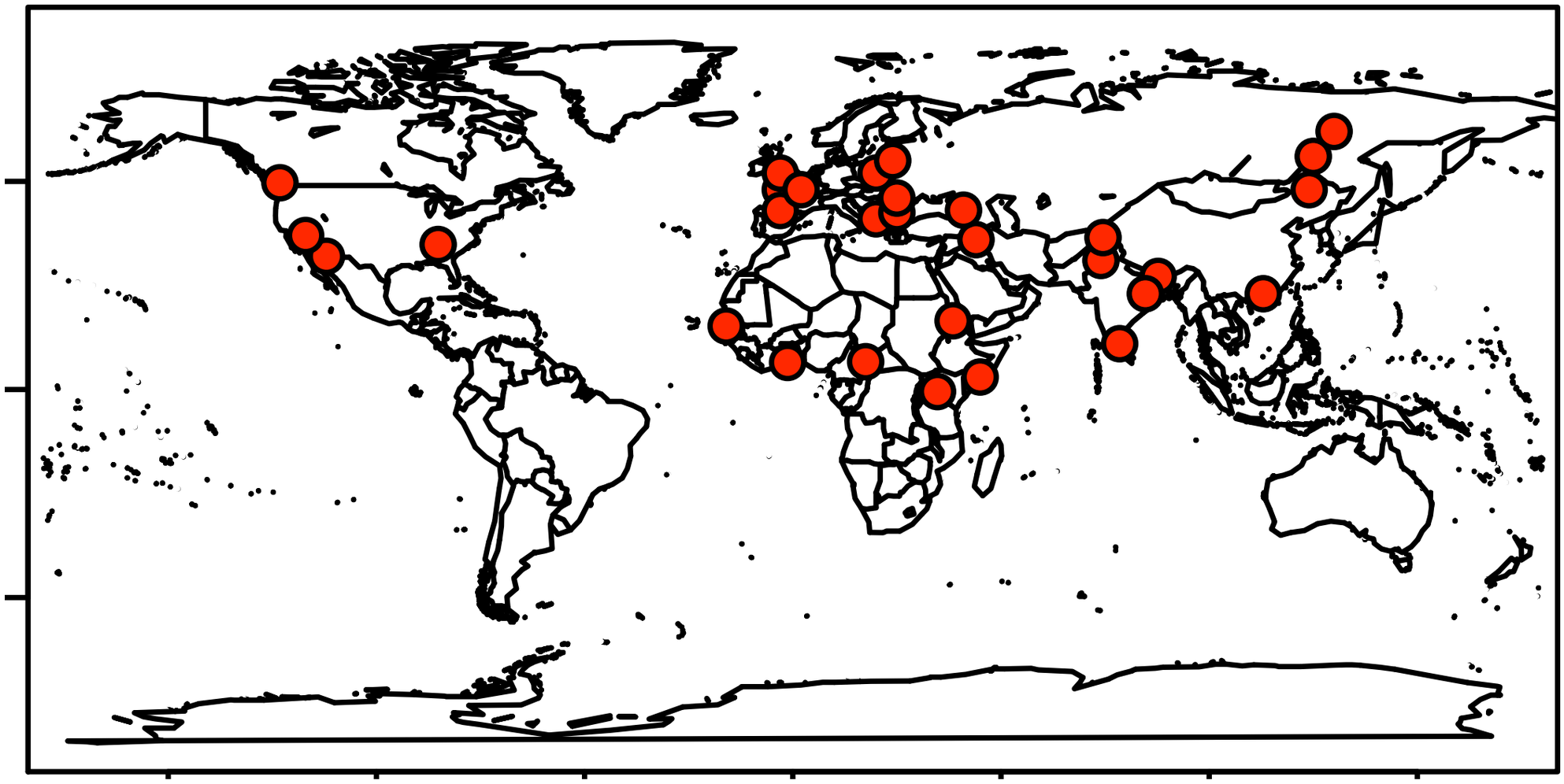
South ← → North

Longitude

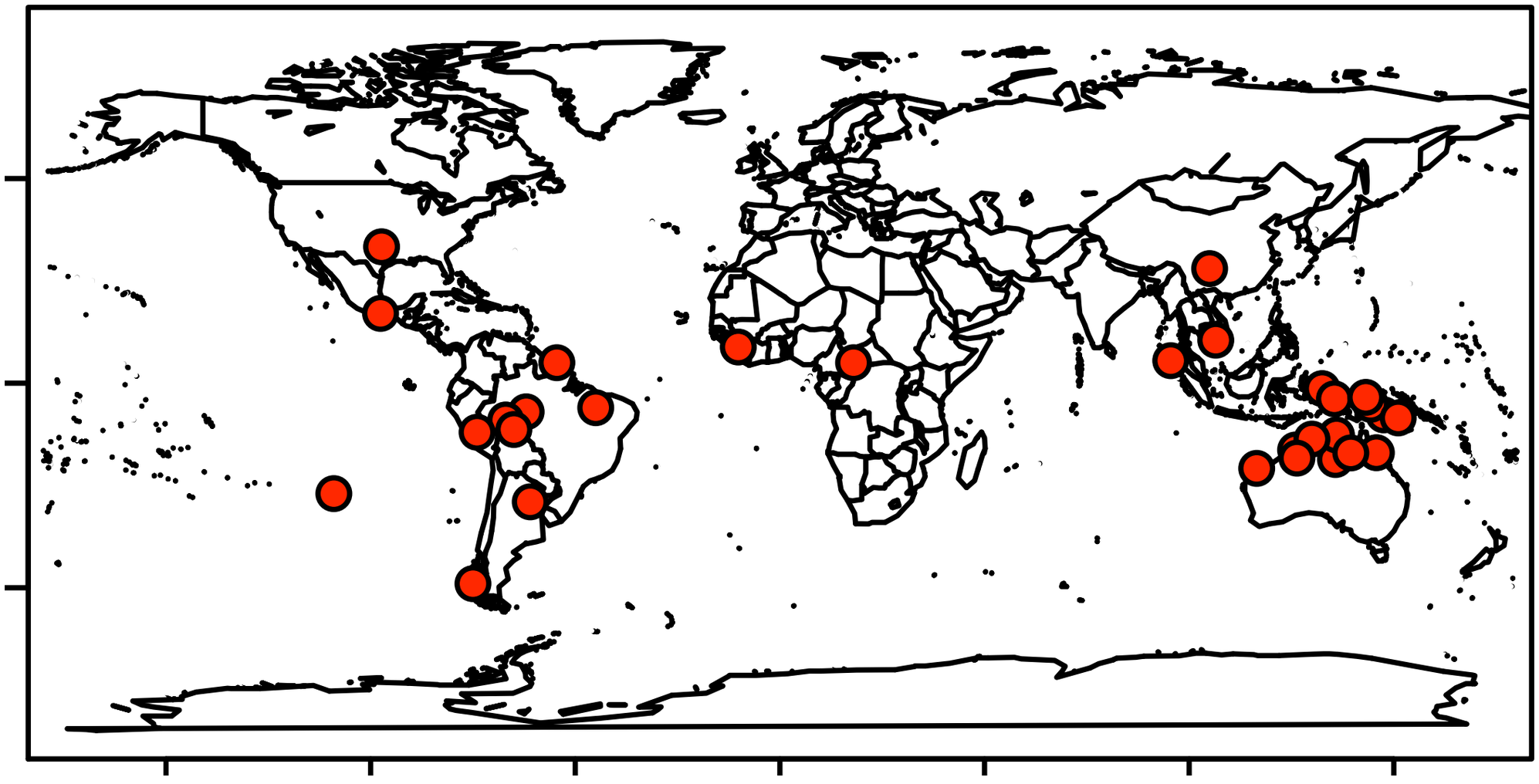


West ← → East

Most Complex

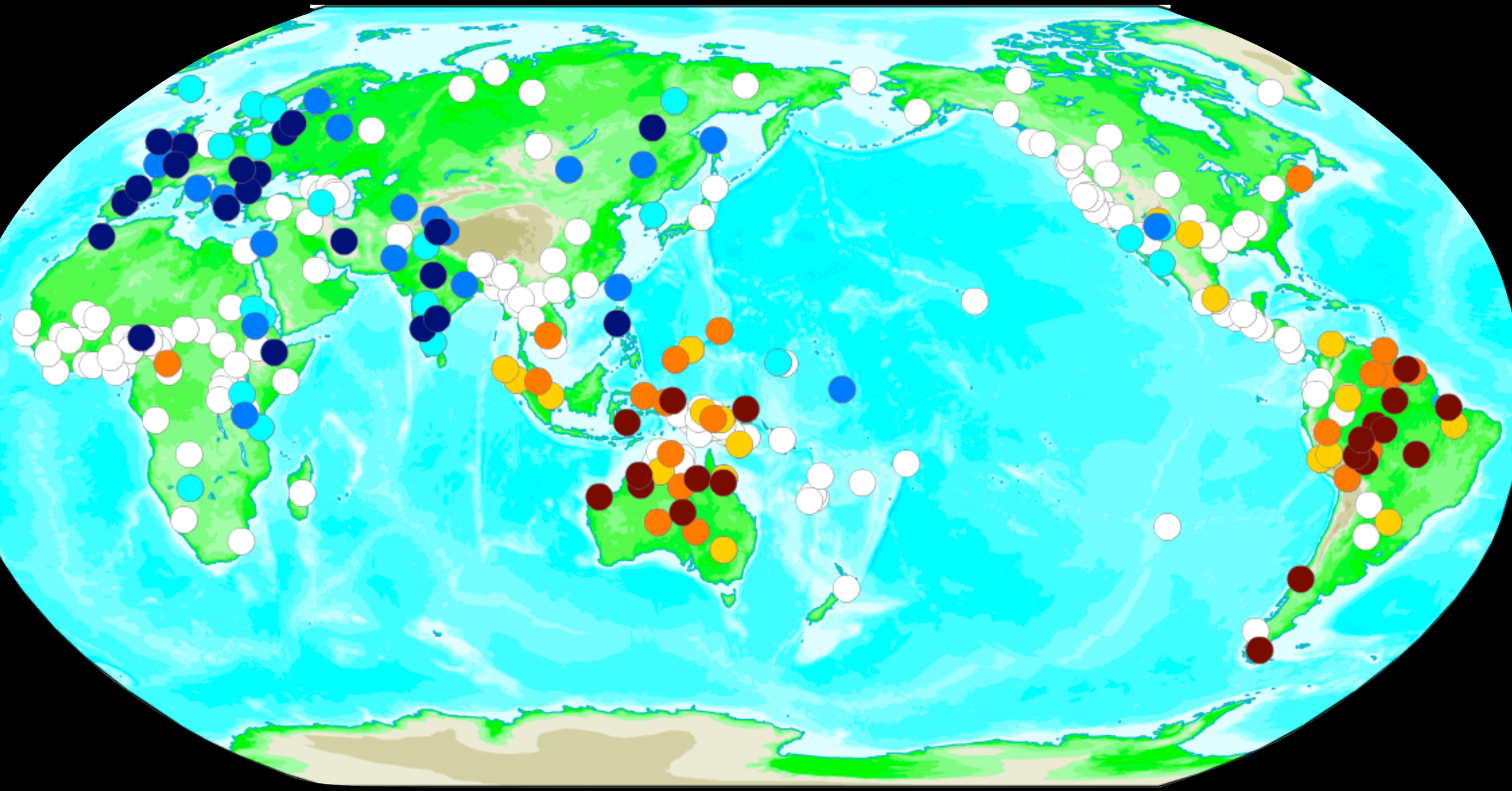


Least Complex



Smoothing Geographical Distributions

- Language variation is geographically rather haphazardly distributed
- To show general trends, take for each language the average of the closest languages
- here: average of the language itself and it's two closest neighbours in the sample



What does this mean?

- Language with low complexity are found at the fringes of human settlement on the globe: this might be a sign on old language structure
- Languages with low complexity are the those farthest away from Europe: this indicates the eurocentricity of the feature selection in WALS (and in linguistics in general)