

The Importance of Rara for Linguistics

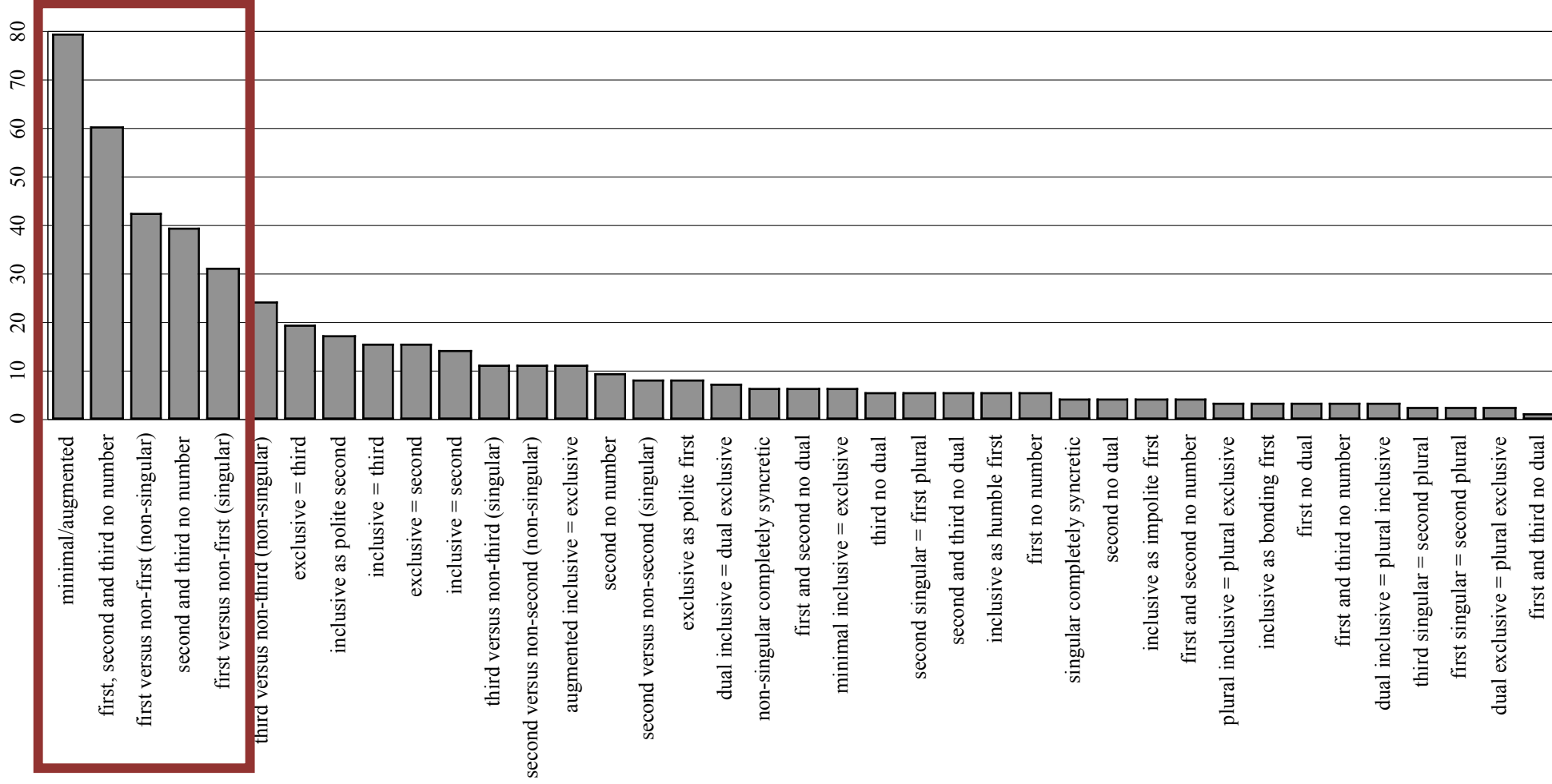
Michael Cysouw

Max Planck Institute for Evolutionary Anthropology

Argumentation

- Why should be care about rara ?
- The existence of rara is no exception
- Rara are not rare !
- Linguistic theories that consider rara to be exceptional have a problem

Rare person marking



**Intermezzo:
'deconstructing' typology**

When is something a rarum ?

- Something is a rarum because it is different from other, more widespread possibilities
- However, every structure in every language is different in a sense
- Real question: how to establish groups of 'similar' languages

	L ₁	L ₂	L ₃	L ₄	L ₅	L ₆	L ₇	L ₈	...
L ₁	1	1	1	0	0	0	0	0	
L ₂	1	1	1	0	0	0	0	0	
L ₃	1	1	1	0	0	0	0	0	
L ₄	0	0	0	1	1	0	0	0	
L ₅	0	0	0	1	1	0	0	0	
L ₆	0	0	0	0	0	1	1	1	
L ₇	0	0	0	0	0	1	1	1	
L ₈	0	0	0	0	0	1	1	1	
...									

Undifferentiated Typology

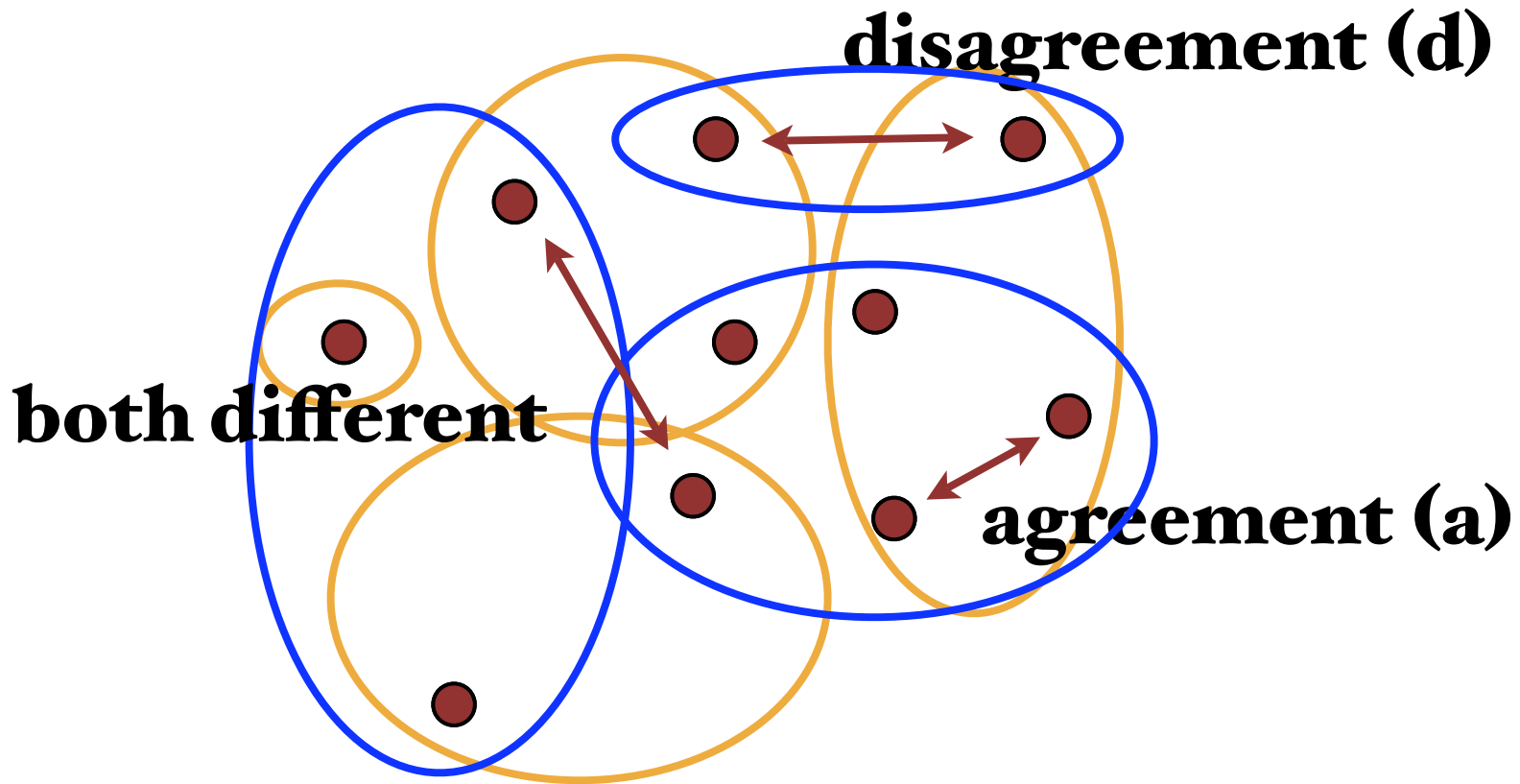
	L ₁	L ₂	L ₃	L ₄	L ₅	L ₆	L ₇	L ₈	...
L ₁	1	0.55	0.72	0.31	0.70	0.61	0.50	0.58	
L ₂	0.55	1	0.55	0.31	0.40	0.44	0.31	0.48	
L ₃	0.72	0.55	1	0.29	0.53	0.51	0.48	0.60	
L ₄	0.31	0.31	0.29	1	0.38	0.36	0.26	0.27	
L ₅	0.70	0.40	0.53	0.38	1	0.64	0.51	0.46	
L ₆	0.61	0.44	0.51	0.36	0.64	1	0.57	0.43	
L ₇	0.50	0.31	0.48	0.26	0.51	0.57	1	0.47	
L ₈	0.58	0.48	0.60	0.27	0.46	0.43	0.47	1	
...									

'Deconstructed' Typology

B. Wälchli's data on motion events

- 72 languages
- 335 clauses for each language from Bible
- clauses describing motion events
- here, only the lexical verb used is included

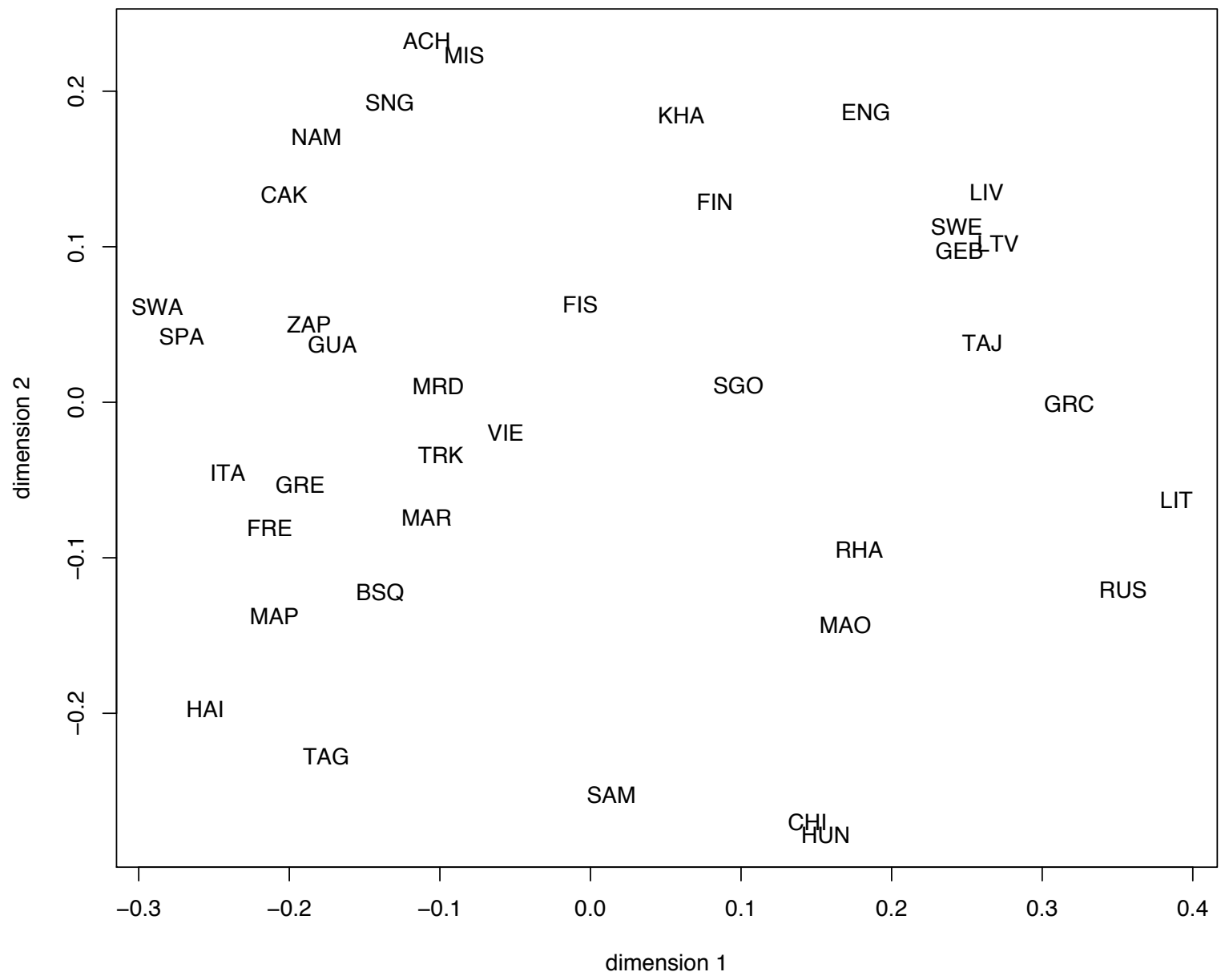
	MRD	LIT	ENG	FRE
1050	sams	eiti	go	aller
1070	sams	eiti	come	venir
1090	sams	eiti	come	venir
1104	lisems	kopti	come	sortir
1105	valgoms	zengti	descend	descendre
1114	–	–	come	se faire entendre
1120	vetjams	varyti	drive	pousser
1140	sams	eiti	come	se rendre
1160	jutams	eiti	walk	marcher

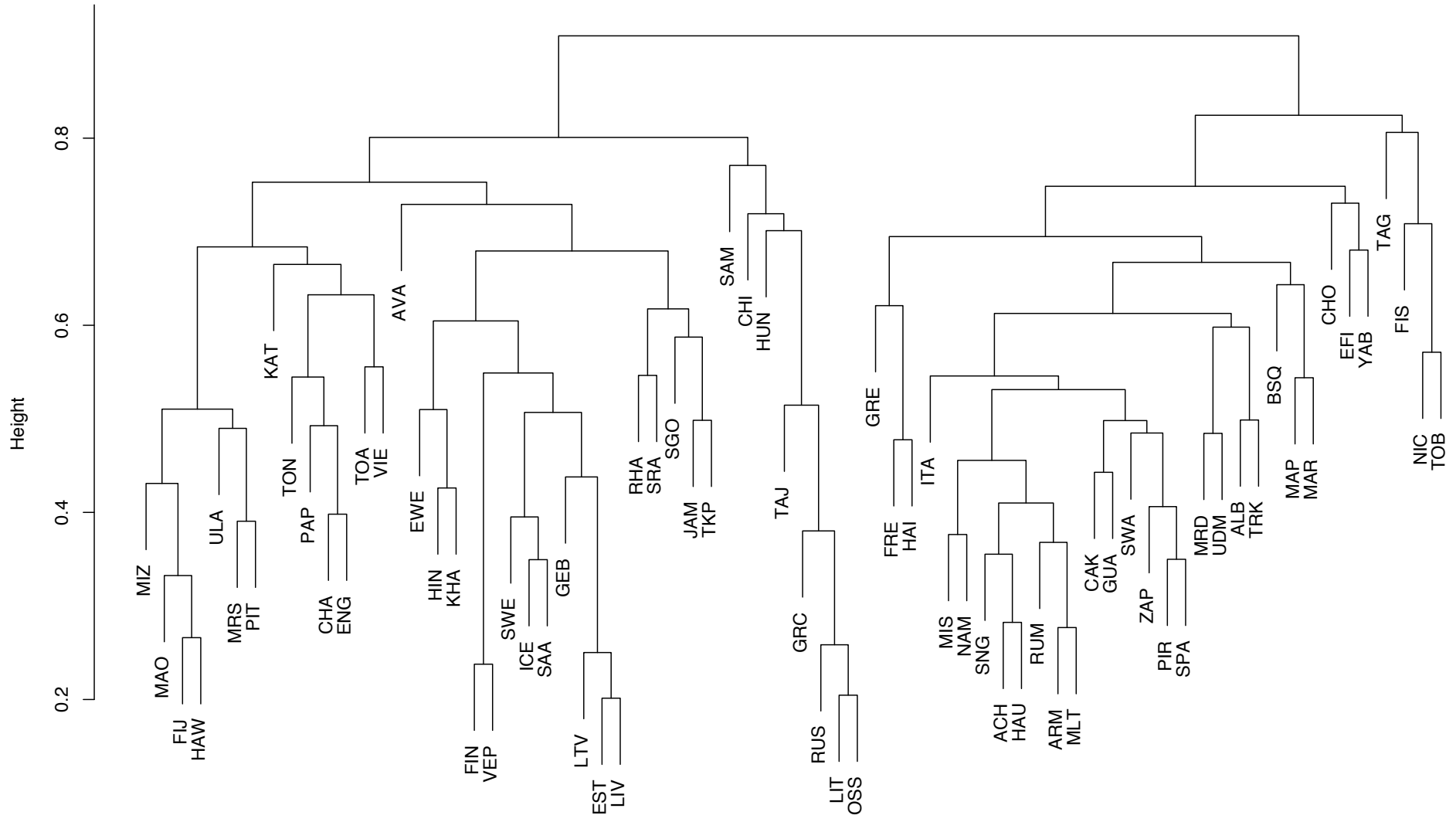


Jaccard similarity:

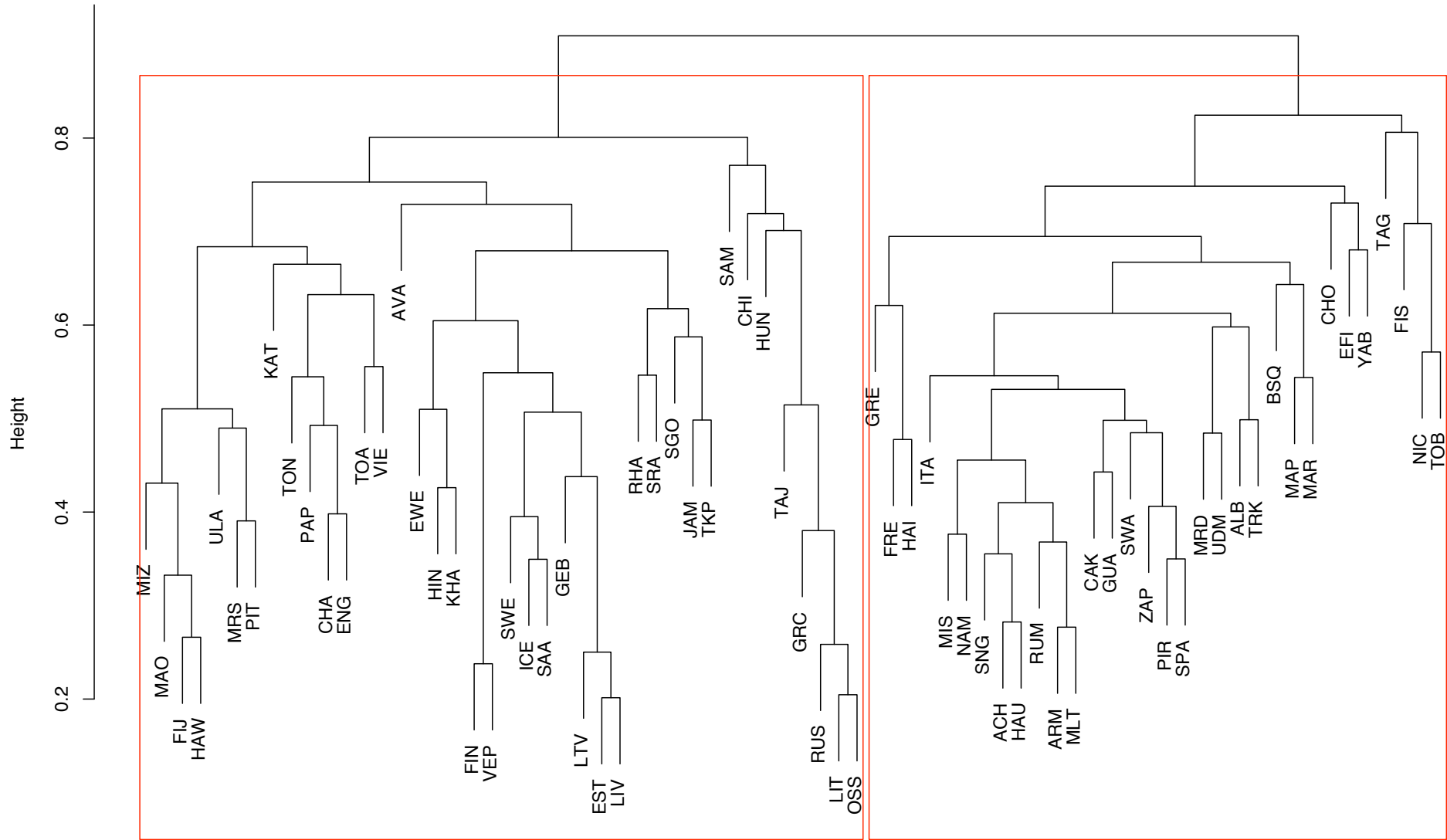
$$\frac{a}{a+d}$$

	ACH	ALB	ARM	AVA	BSQ	CAK	CHA	CHI	CHO	EFI	ENG	EST	EWE	FIJ	FIN	FIS	FRE	GEB	GRC	GRE	GUA	HAI	HAU	HAW	HIN	HUN	ICE	ITA	JAM	KAT	KHA	LIV	LTV	MAO	MAP	MA		
ACH	1	0.43	0.62	0.31	0.41	0.57	0.6	0.22	0.41	0.51	0.55	0.46	0.54	0.35	0.46	0.35	0.45	0.44	0.4	0.39	0.52	0.41	0.72	0.41	0.55	0.23	0.48	0.47	0.32	0.36	0.55	0.37	0.47	0.45	0.36	0.43	0.4	
ALB	0.43	1	0.53	0.26	0.41	0.45	0.41	0.2	0.32	0.34	0.38	0.3	0.33	0.21	0.29	0.4	0.33	0.25	0.41	0.53	0.35	0.49	0.29	0.34	0.28	0.3	0.5	0.3	0.28	0.37	0.26	0.28	0.33	0.28	0.4	0.4		
ARM	0.62	0.53	1	0.27	0.4	0.63	0.61	0.26	0.37	0.4	0.54	0.46	0.47	0.34	0.42	0.4	0.41	0.44	0.4	0.39	0.57	0.33	0.67	0.36	0.53	0.24	0.46	0.49	0.32	0.31	0.49	0.34	0.46	0.49	0.34	0.37	0.3	
AVA	0.31	0.26	0.27	1	0.27	0.24	0.27	0.25	0.18	0.22	0.32	0.36	0.27	0.31	0.33	0.21	0.26	0.32	0.24	0.25	0.31	0.23	0.26	0.28	0.32	0.25	0.34	0.26	0.28	0.25	0.32	0.29	0.33	0.34	0.3	0.27	0.3	
BSQ	0.41	0.41	0.4	0.27	1	0.41	0.34	0.22	0.25	0.28	0.29	0.28	0.32	0.29	0.32	0.21	0.36	0.26	0.23	0.37	0.44	0.33	0.38	0.3	0.32	0.26	0.28	0.42	0.25	0.32	0.34	0.25	0.27	0.3	0.34	0.36	0.3	
CAK	0.57	0.45	0.63	0.24	0.41	1	0.55	0.21	0.39	0.4	0.4	0.38	0.42	0.3	0.36	0.34	0.44	0.31	0.32	0.37	0.56	0.35	0.61	0.31	0.43	0.21	0.37	0.48	0.27	0.26	0.41	0.29	0.38	0.39	0.29	0.41	0.3	
CHA	0.6	0.41	0.61	0.27	0.34	0.55	1	0.24	0.41	0.41	0.4	0.6	0.41	0.46	0.33	0.43	0.39	0.43	0.43	0.39	0.47	0.33	0.59	0.39	0.45	0.24	0.5	0.46	0.3	0.35	0.57	0.33	0.43	0.46	0.33	0.37	0.3	
CHI	0.22	0.2	0.26	0.25	0.22	0.21	0.24	1	0.16	0.2	0.27	0.37	0.22	0.28	0.26	0.19	0.19	0.31	0.28	0.18	0.25	0.18	0.23	0.23	0.25	0.29	0.3	0.2	0.33	0.25	0.3	0.29	0.33	0.33	0.25	0.25	0.2	
CHO	0.41	0.32	0.37	0.18	0.25	0.39	0.41	0.16	1	0.29	0.28	0.24	0.31	0.21	0.26	0.22	0.27	0.24	0.22	0.27	0.33	0.27	0.45	0.23	0.27	0.17	0.26	0.32	0.21	0.26	0.31	0.21	0.24	0.27	0.2	0.29	0.3	
EFI	0.51	0.34	0.4	0.22	0.28	0.4	0.41	0.2	0.29	1	0.32	0.27	0.36	0.23	0.31	0.27	0.32	0.26	0.23	0.32	0.33	0.32	0.51	0.26	0.35	0.18	0.29	0.37	0.22	0.27	0.41	0.2	0.29	0.27	0.25	0.32	0.3	
ENG	0.55	0.38	0.54	0.32	0.29	0.4	0.6	0.27	0.28	0.32	1	0.58	0.47	0.47	0.48	0.37	0.32	0.52	0.5	0.34	0.39	0.25	0.5	0.46	0.54	0.33	0.6	0.37	0.44	0.35	0.58	0.5	0.56	0.6	0.41	0.31	0.3	
EST	0.46	0.3	0.46	0.36	0.28	0.38	0.41	0.37	0.24	0.27	0.58	1	0.46	0.47	0.49	0.48	0.3	0.28	0.58	0.49	0.25	0.39	0.21	0.41	0.38	0.48	0.39	0.63	0.29	0.46	0.31	0.58	0.6	0.8	0.79	0.42	0.31	0.3
EWE	0.54	0.33	0.47	0.27	0.32	0.42	0.46	0.22	0.31	0.36	0.47	0.46	1	0.29	0.42	0.29	0.35	0.4	0.33	0.3	0.38	0.31	0.52	0.31	0.52	0.25	0.46	0.32	0.32	0.31	0.49	0.35	0.43	0.49	0.32	0.33	0.3	
FIJ	0.35	0.26	0.34	0.31	0.29	0.3	0.33	0.28	0.21	0.23	0.47	0.49	0.29	1	0.32	0.19	0.22	0.36	0.47	0.22	0.34	0.21	0.36	0.73	0.34	0.36	0.43	0.28	0.35	0.32	0.38	0.66	0.45	0.48	0.67	0.33	0	
FIN	0.46	0.31	0.42	0.33	0.32	0.36	0.43	0.26	0.26	0.31	0.48	0.48	0.42	0.32	1	0.42	0.21	0.45	0.36	0.33	0.38	0.24	0.45	0.3	0.5	0.28	0.48	0.36	0.37	0.28	0.48	0.38	0.46	0.48	0.3	0.27	0.3	
FIS	0.35	0.29	0.4	0.21	0.21	0.34	0.39	0.19	0.22	0.27	0.37	0.3	0.29	0.19	0.42	1	0.32	0.36	0.25	0.29	0.35	0.21	0.39	0.22	0.41	0.21	0.34	0.32	0.26	0.19	0.34	0.25	0.32	0.33	0.2	0.19	0.2	
FRE	0.45	0.4	0.41	0.26	0.36	0.44	0.43	0.19	0.27	0.32	0.32	0.28	0.35	0.22	0.31	0.32	1	0.3	0.22	0.43	0.43	0.52	0.39	0.24	0.36	0.21	0.35	0.51	0.25	0.26	0.34	0.22	0.26	0.28	0.23	0.35	0.3	
GEB	0.44	0.33	0.44	0.32	0.26	0.31	0.43	0.31	0.24	0.26	0.52	0.58	0.4	0.36	0.45	0.36	0.3	1	0.41	0.26	0.36	0.2	0.41	0.36	0.49	0.31	0.57	0.3	0.39	0.25	0.46	0.48	0.56	0.59	0.33	0.22	0	
GRC	0.4	0.25	0.4	0.24	0.23	0.32	0.39	0.28	0.22	0.23	0.5	0.49	0.33	0.47	0.36	0.25	0.22	0.41	1	0.21	0.29	0.2	0.38	0.45	0.37	0.34	0.44	0.25	0.37	0.25	0.42	0.71	0.52	0.51	0.4	0.22	0.2	
GRE	0.39	0.41	0.39	0.25	0.37	0.37	0.43	0.18	0.27	0.32	0.34	0.25	0.3	0.22	0.33	0.29	0.43	0.26	0.21	1	0.38	0.38	0.42	0.24	0.31	0.23	0.28	0.42	0.24	0.27	0.39	0.2	0.24	0.28	0.24	0.35	0.3	
GUA	0.52	0.53	0.57	0.31	0.44	0.56	0.47	0.25	0.33	0.33	0.39	0.39	0.38	0.34	0.38	0.35	0.43	0.36	0.29	0.38	1	0.38	0.51	0.33	0.48	0.28	0.38	0.47	0.31	0.3	0.46	0.3	0.37	0.42	0.35	0.49	0.4	
HAI	0.41	0.35	0.33	0.23	0.33	0.35	0.33	0.18	0.27	0.32	0.25	0.21	0.31	0.21	0.24	0.21	0.52	0.2	0.2	0.38	0.38	1	0.39	0.23	0.25	0.21	0.22	0.44	0.21	0.25	0.28	0.18	0.19	0.23	0.21	0.42	0.3	
HAU	0.72	0.49	0.67	0.26	0.38	0.61	0.59	0.23	0.45	0.51	0.5	0.41	0.52	0.36	0.45	0.39	0.39	0.41	0.38	0.42	0.51	0.39	1	0.38	0.51	0.24	0.43	0.47	0.33	0.31	0.51	0.33	0.43	0.44	0.34	0.44	0.4	
HAW	0.41	0.29	0.36	0.28	0.3	0.31	0.39	0.23	0.23	0.26	0.46	0.38	0.31	0.73	0.3	0.22	0.24	0.36	0.45	0.24	0.33	0.23	0.38	1	0.37	0.35	0.37	0.31	0.35	0.34	0.38	0.59	0.38	0.41	0.67	0.32	0.3	
HIN	0.55	0.34	0.53	0.32	0.32	0.43	0.45	0.25	0.27	0.35	0.54	0.48	0.52	0.34	0.5	0.41	0.36	0.49	0.37	0.31	0.48	0.25	0.51	0.37	1	0.31	0.51	0.35	0.35	0.33	0.57	0.41	0.53	0.53	0.37	0.29	0.3	
HUN	0.23	0.28	0.24	0.25	0.26	0.21	0.24	0.29	0.17	0.18	0.33	0.39	0.25	0.36	0.28	0.21	0.21	0.31	0.34	0.23	0.28	0.21	0.24	0.35	0.31	1	0.3	0.26	0.39	0.28	0.3	0.39	0.35	0.42	0.35	0.28	0.2	
ICE	0.48	0.3	0.46	0.34	0.28	0.37	0.5	0.3	0.26	0.29	0.6	0.63	0.46	0.43	0.48	0.34	0.35	0.57	0.44	0.28	0.38	0.22	0.43	0.37	0.51	0.3	1	0.34	0.39	0.28	0.55	0.48	0.58	0.61	0.35	0.26	0.3	
ITA	0.47	0.5	0.49	0.26	0.42	0.48	0.46	0.2	0.32	0.37	0.37	0.29	0.32	0.28	0.36	0.32	0.51	0.3	0.25	0.42	0.47	0.44	0.47	0.31	0.35	0.26	0.34	1	0.24	0.27	0.36	0.26	0.27	0.33	0.28	0.44	0.3	
JAM	0.32	0.3	0.32	0.28	0.25	0.27	0.3	0.33	0.21	0.22	0.44	0.46	0.32	0.35	0.37	0.26	0.25	0.39	0.37	0.24	0.31	0.21	0.33	0.35	0.35	0.39	0.39	0.24	1	0.32	0.41	0.42	0.4	0.38	0.34	0.32	0.2	
KAT	0.36	0.28	0.31	0.25	0.32	0.26	0.35	0.25	0.26	0.27	0.35	0.31	0.31	0.32	0.28	0.19	0.26	0.25	0.25	0.27	0.3	0.25	0.31	0.34	0.33	0.28	0.28	0.27	0.32	1	0.38	0.27	0.3	0.33	0.35	0.34	0.3	
KHA	0.55	0.37	0.49	0.32	0.34	0.41	0.57	0.3	0.31	0.41	0.58	0.58	0.49	0.38	0.48	0.34	0.34	0.46	0.42	0.39	0.46	0.28	0.51	0.38	0.57	0.3	0.55	0.36	0.41	0.38	1	0.37	0.56	0.53	0.41	0.37	0.3	
LIT	0.37	0.26	0.34	0.29	0.25	0.29	0.33	0.29	0.21	0.2	0.5	0.6	0.35	0.66	0.38	0.25	0.22	0.48	0.71	0.2	0.3	0.18	0.33	0.59	0.41	0.39	0.48	0.26	0.42	0.27	0.37	1	0.56	0.68	0.55	0.25	0.2	
LIV	0.47	0.28	0.46	0.33	0.27	0.38	0.43	0.33	0.24	0.29	0.56	0.8	0.43	0.45	0.46	0.32	0.26	0.56	0.52	0.24	0.37	0.19	0.43	0.38	0.53	0.35	0.58	0.27	0.4	0.3	0.56	0.56	1	0.75	0.39	0.26	0.3	
LTV	0.45	0.33	0.49	0.34	0.3	0.39	0.46	0.33	0.27	0.27	0.6	0.79	0.49	0.48	0.48	0.33	0.28	0.59	0.51	0.28	0.42	0.23	0.44	0.41	0.53	0.42	0.61	0.33	0.38	0.33	0.53	0.68	0.75	1	0.43	0.31	0.3	
MAO	0.36	0.28	0.34	0.3	0.34	0.29	0.33	0.25	0.2	0.25	0.41	0.42	0.32	0.67	0.3	0.2	0.23	0.33	0.4	0.24	0.35	0.21	0.34	0.67	0.37	0.35	0.35	0.28	0.34	0.35	0.41	0.55	0.39	0.43	1	0.3	0.3	
MAP	0.43	0.4	0.37	0.27	0.36	0.41	0.37	0.25	0.29	0.32	0.31	0.31	0.33	0.33	0.27	0.19	0.35	0.22	0.22	0.35	0.49	0.42	0.44	0.32	0.29	0.28	0.26	0.44	0.32	0.34								

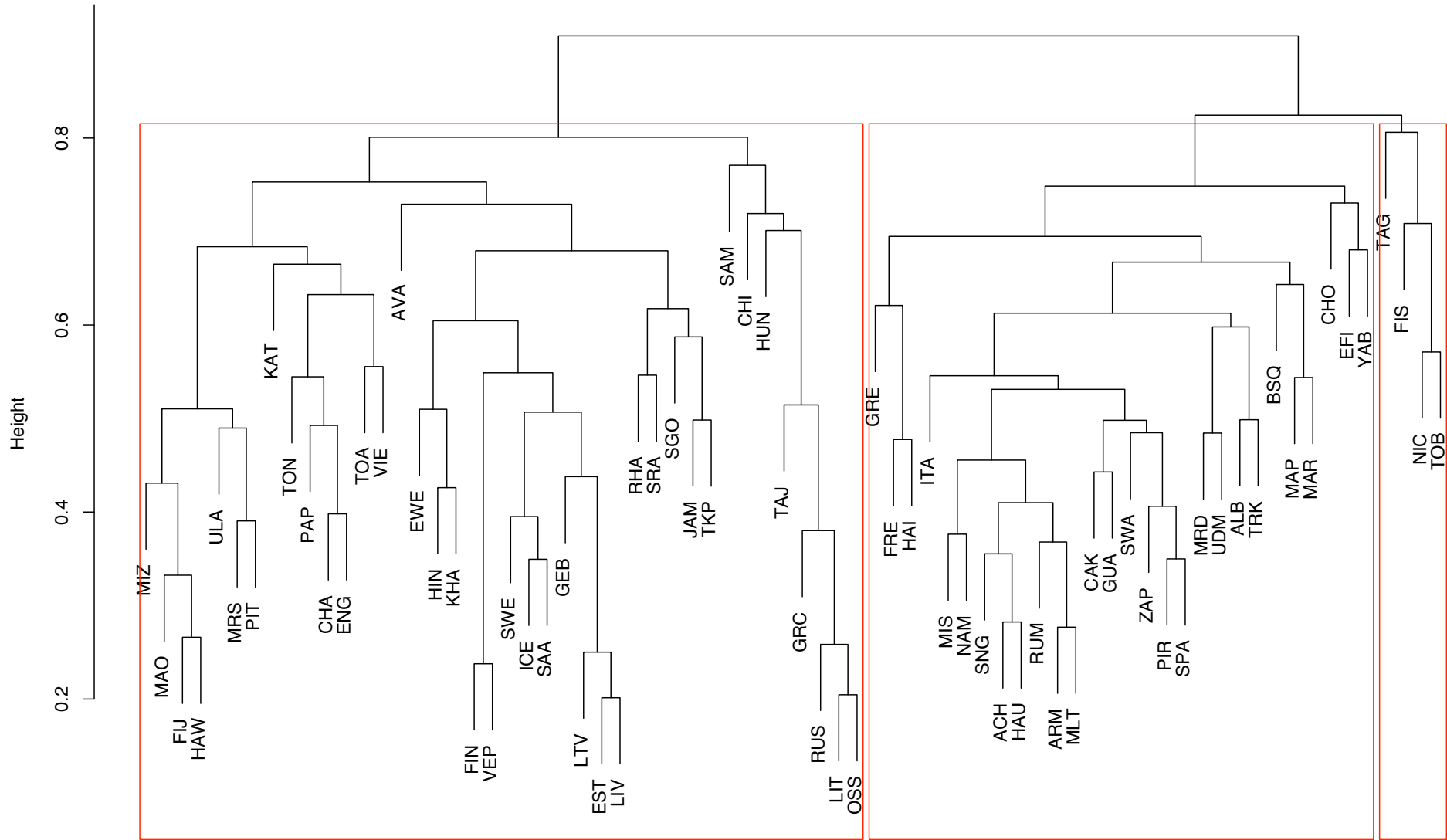




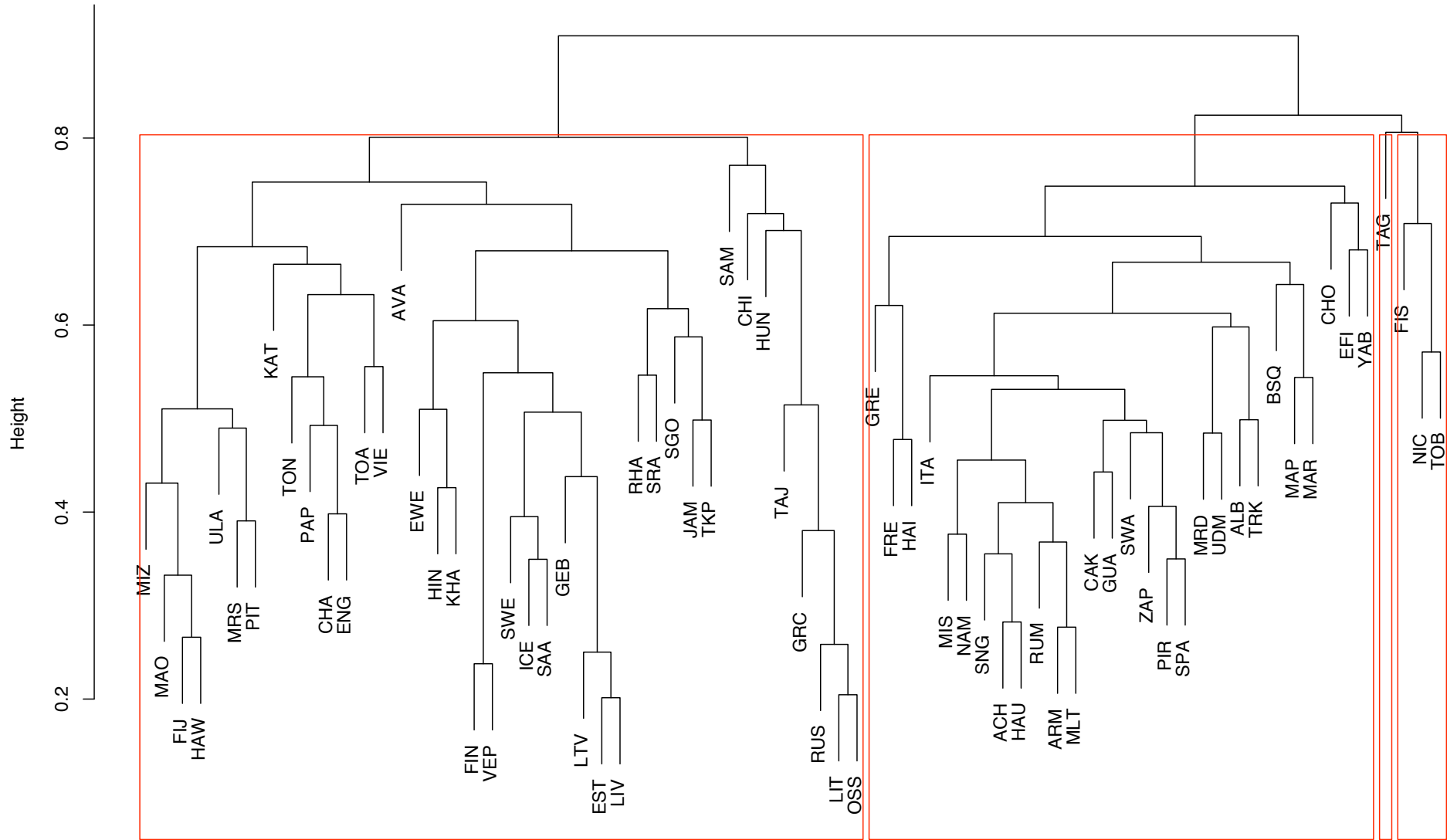
hclust (*, "complete")



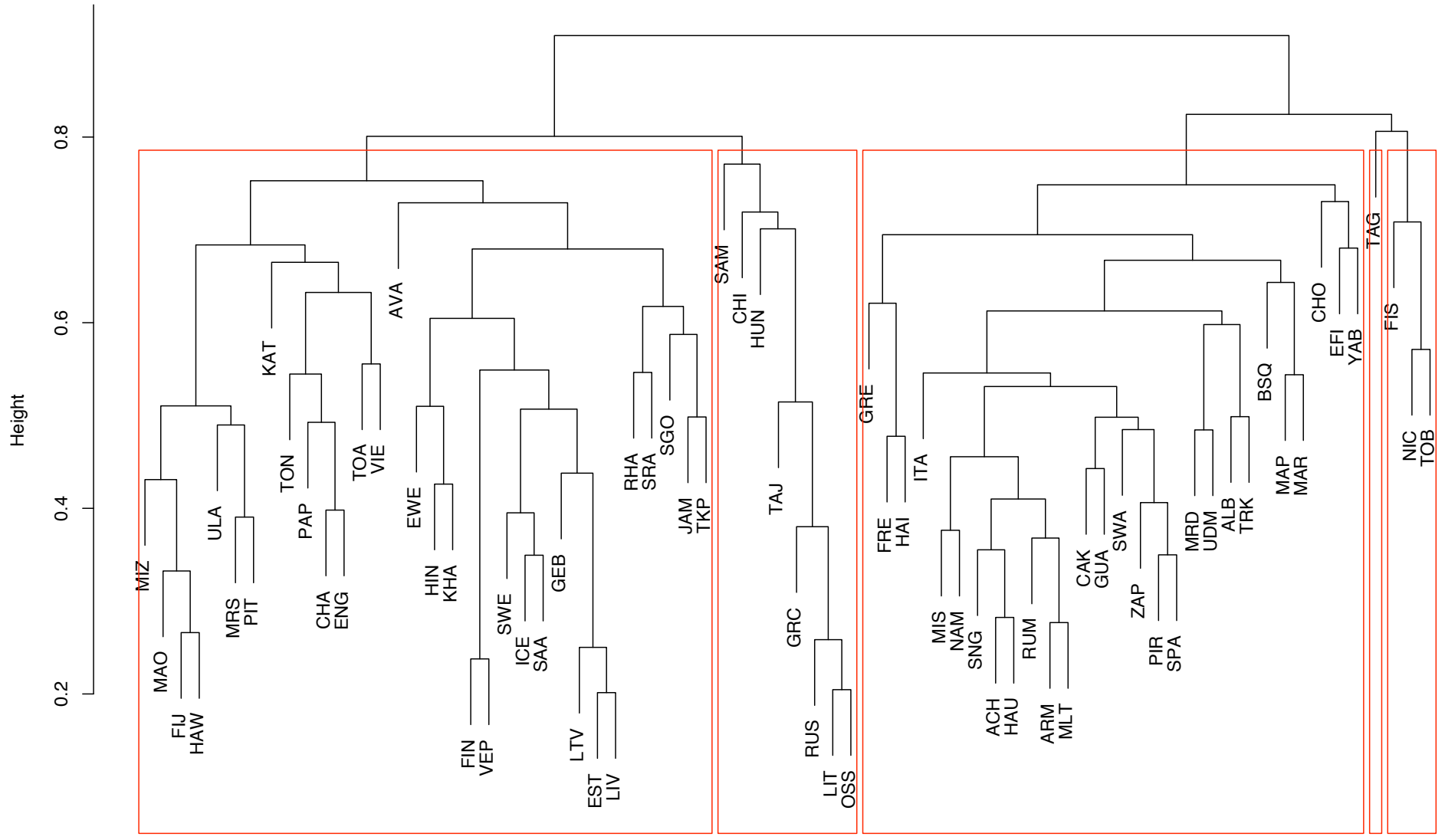
hclust (*, "complete")



hclust (*, "complete")



hclust (*, "complete")

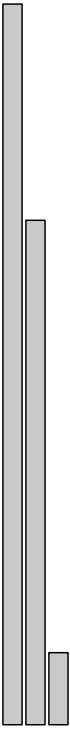


hclust (*, "complete")

40
30
20
10
0



40
30
20
10
0





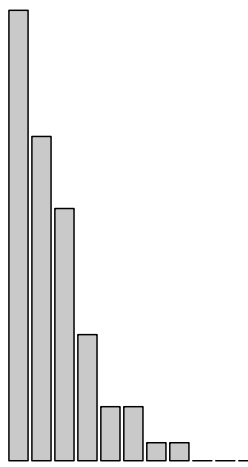




40
30
20
10
0



40
30
20
10
0

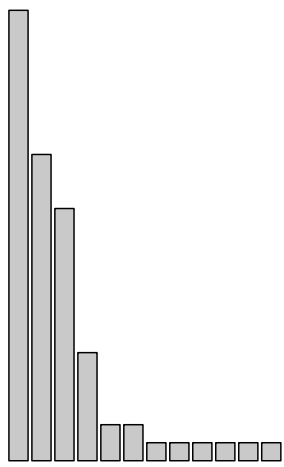




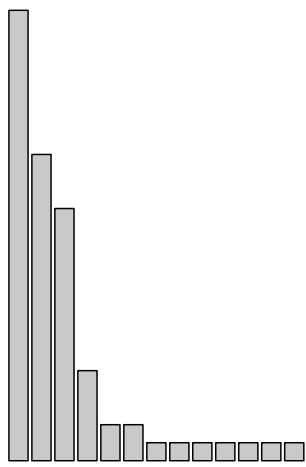




40
30
20
10
0



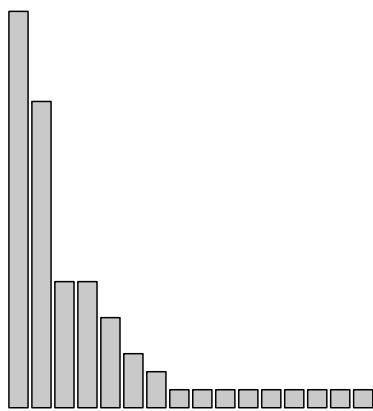
40
30
20
10
0







40
30
20
10
0











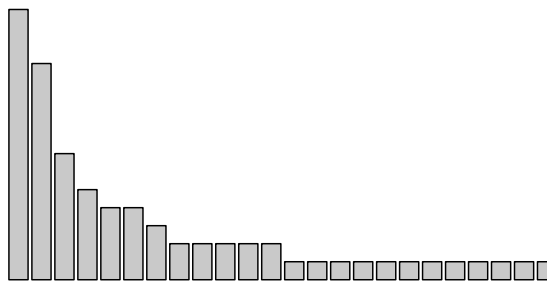




40
30
20
10
0



40
30
20
10
0

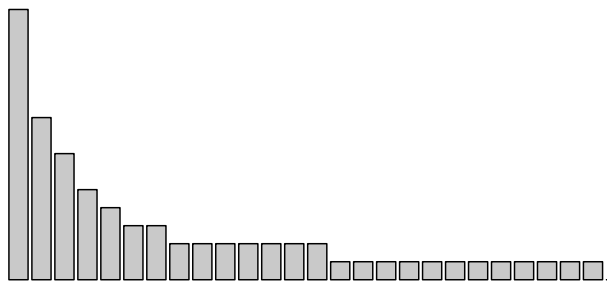


40
30
20
10
0





40
30
20
10
0







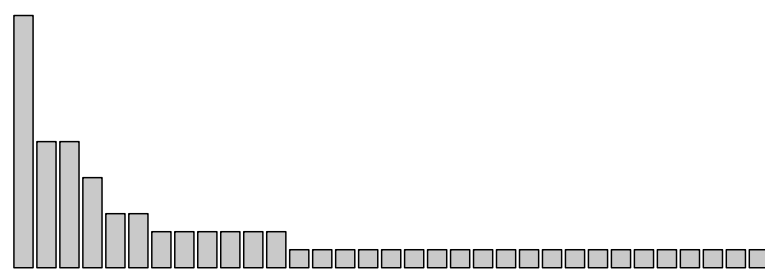




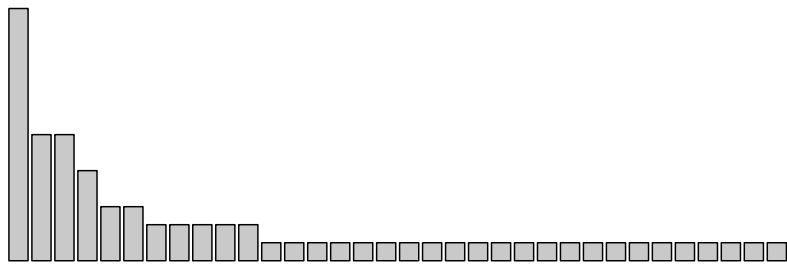




40
30
20
10
0



40
30
20
10
0



Interpretation

- Being faithful to a particular grouping criterium leads to many rara
- Linguists (grammarians) do not seem to like such distributions



WALS



The World Atlas of Language Structures

edited by M. Haspelmath, M. S. Dryer, D. Gil, B. Comrie

The Interactive Reference Tool

developed by Hans-Jörg Bibiko

OXFORD
UNIVERSITY PRESS

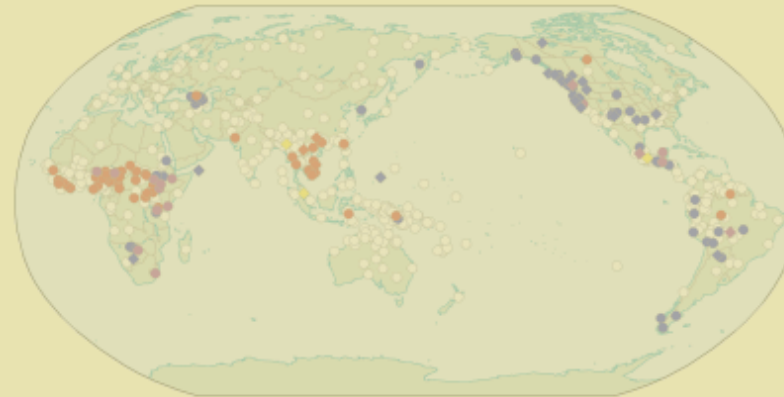
FEATURE VIEWER

LANGUAGE VIEWER

COMPOSER

ABOUT THIS PROGRAM

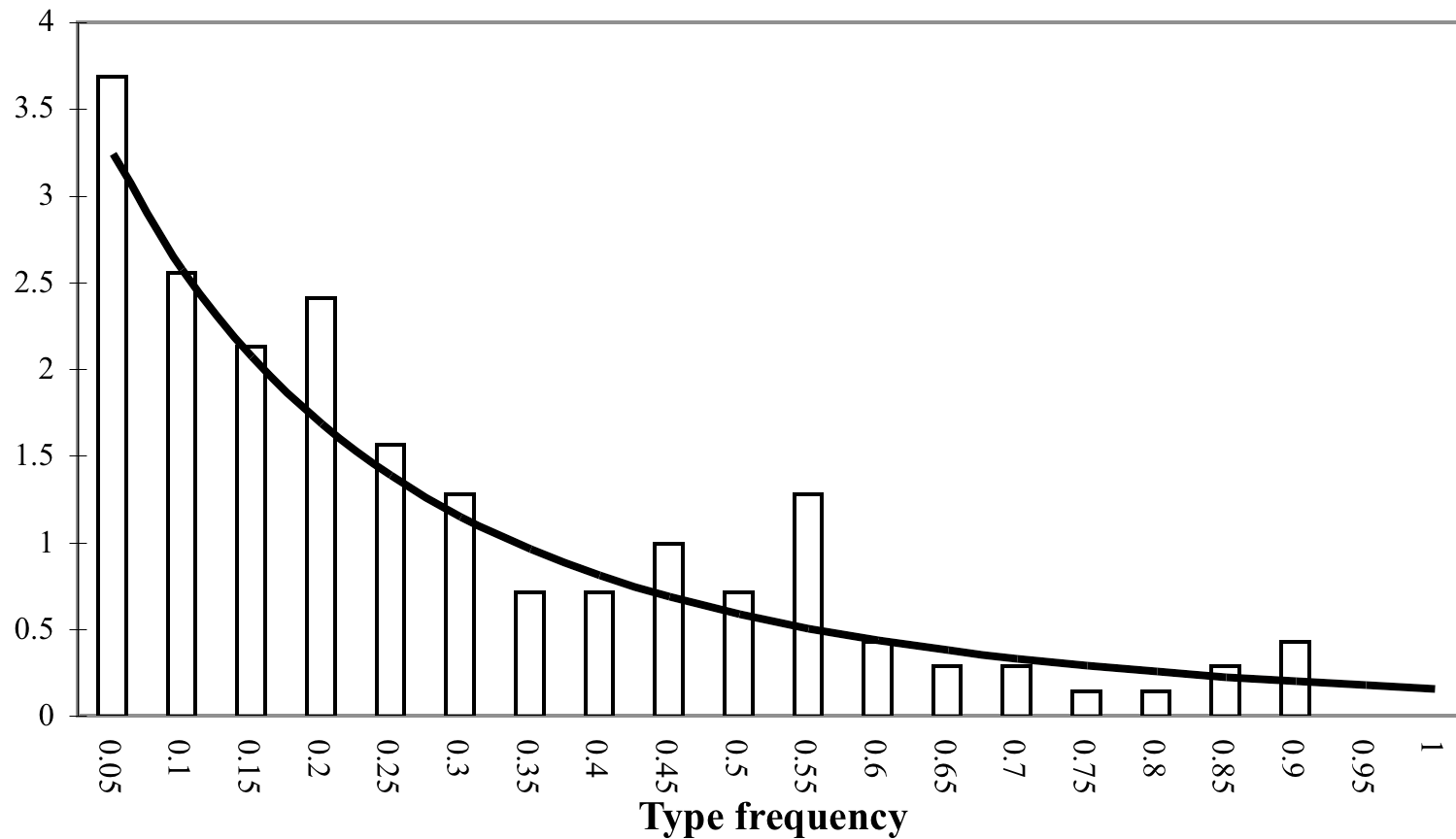
GUIDED TOUR



World Atlas of Language Structures

- 2,600 languages
- 140 characteristics, 670 types
- 60,000 datapoints
- 50 different authors
- survey of what typologists have been interested in over the last decades

random selection of values from WALS



Summary

- Rara are not rare !
- The existence of rara is a solid fact of every domain of linguistic structure
- How would a linguistic theory look like that incorporates this finding ?

A personal opinion

- Everything is possible in the structure of human language
- If needed, the human brain can manage quite odd behaviour
- However, as long as there is no social 'need', quirky structures do not occur



MAX-PLANCK-GESellschaft

**Max Planck Institute
for Evolutionary Anthropology**