

# **Historical Reconstruction through Comparison of Meaning**

Michael Cysouw  
LMU Munich

**Die Bedeutung eines Wortes ist  
sein Gebrauch in der Sprache.**

Ludwig Wittgenstein,  
Philosophische Untersuchungen, no. 43.

# Proposal

- Define the meaning of a linguistic form as the set of all contexts in which it occurs
- Under this strongly extensionalistic definition of meaning, variation in meaning becomes readily measurable
- The differences in meaning between cognate forms reveal historical processes

# Proposal

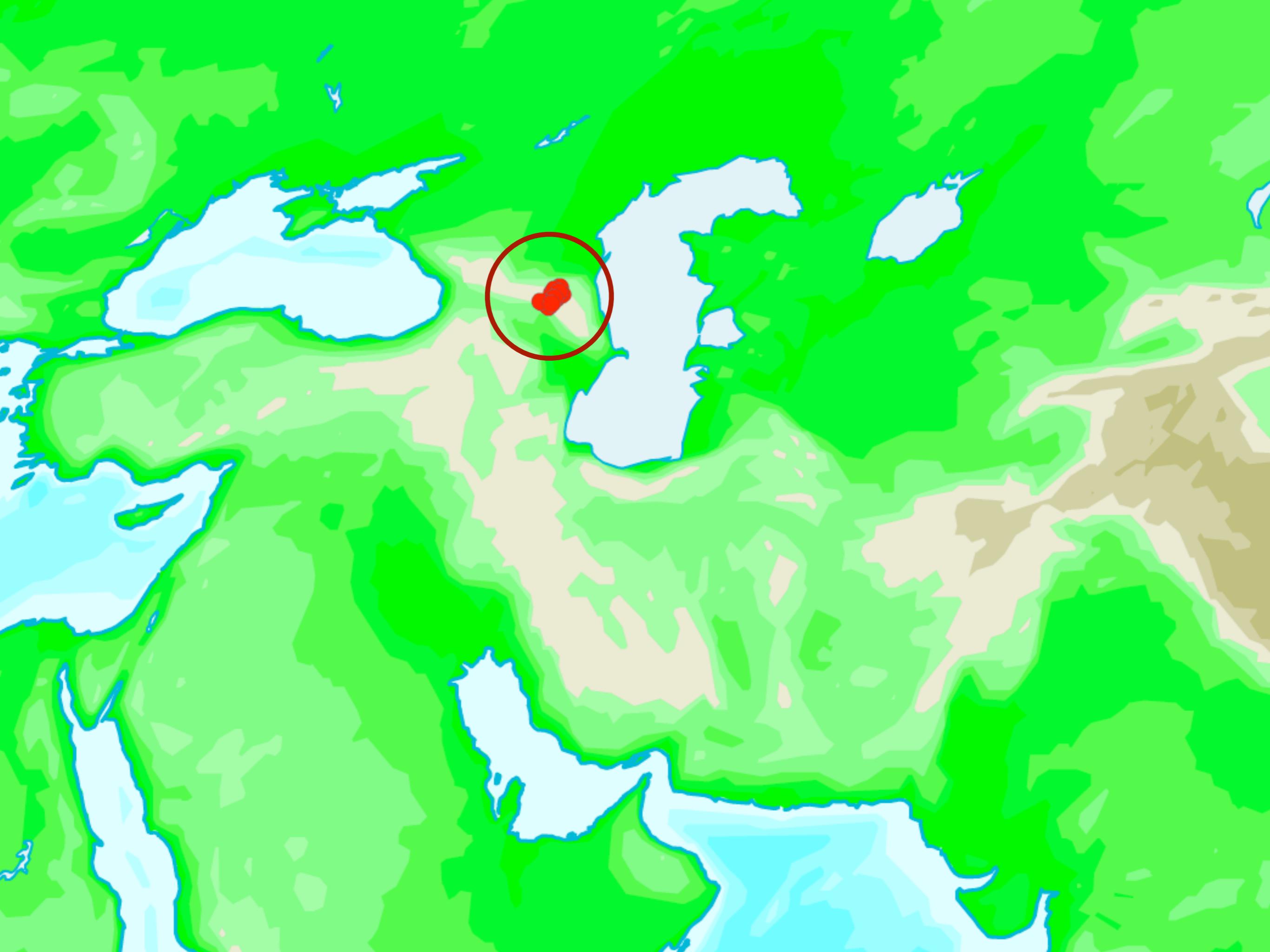
- Define the **meaning** of a linguistic form as the set of all contexts in which it occurs
- Under this strongly extensionalistic definition of **meaning**, variation in **meaning** becomes readily measurable
- The differences in **meaning** between cognate forms reveal historical processes

# Proposal

- Define the **property M** of a linguistic form as the set of all contexts in which it occurs
- Under this strongly extensionalistic definition of **property M**, variation in **property M** becomes readily measurable
- The differences in **property M** between cognate forms reveal historical processes



**“go” periphrasis in Tsezic**



**Tsezic**

**West Tsezic**

**East Tsezic**

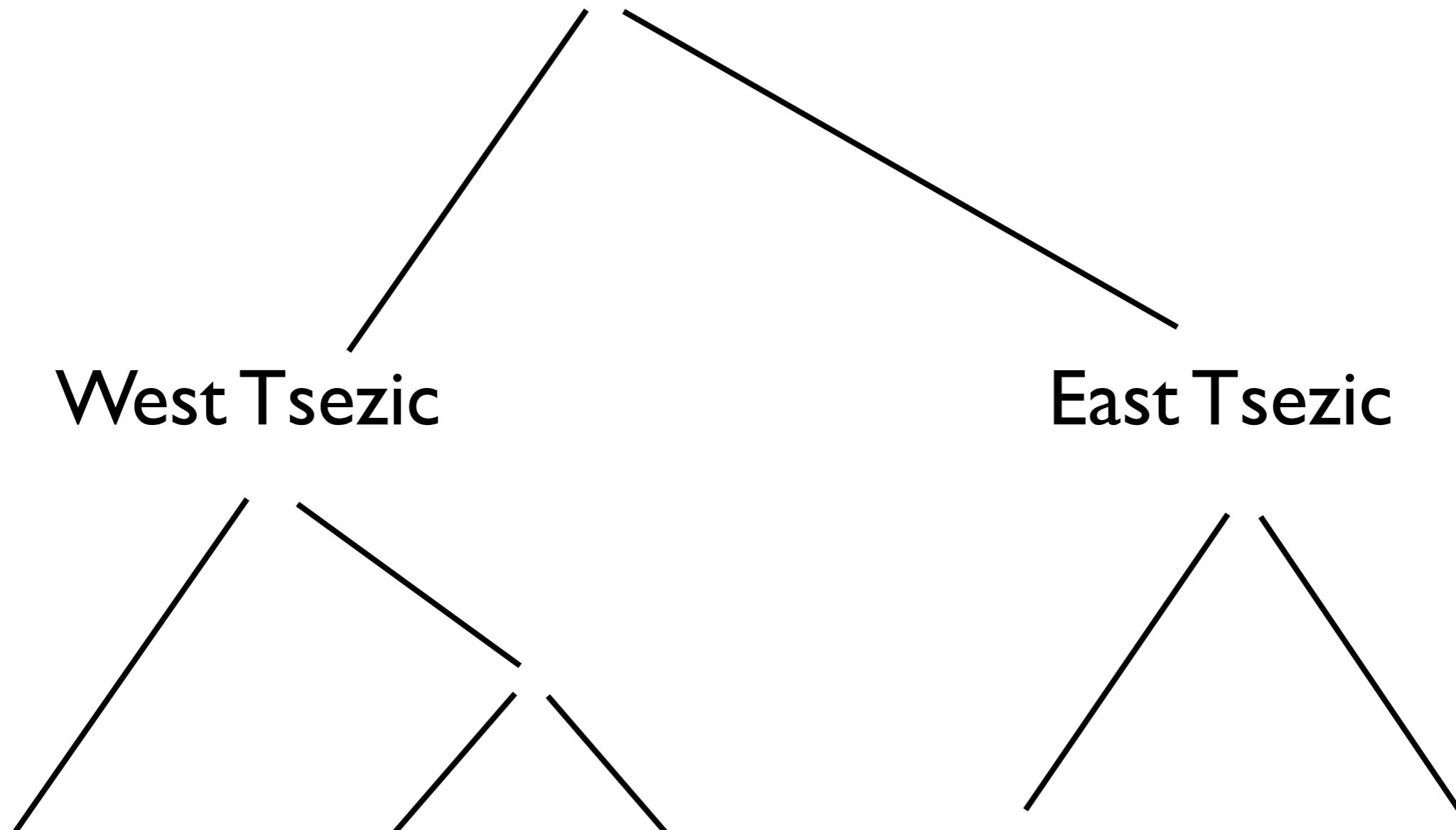
**Khwarshi**

**Tsez**

**Hinuq**

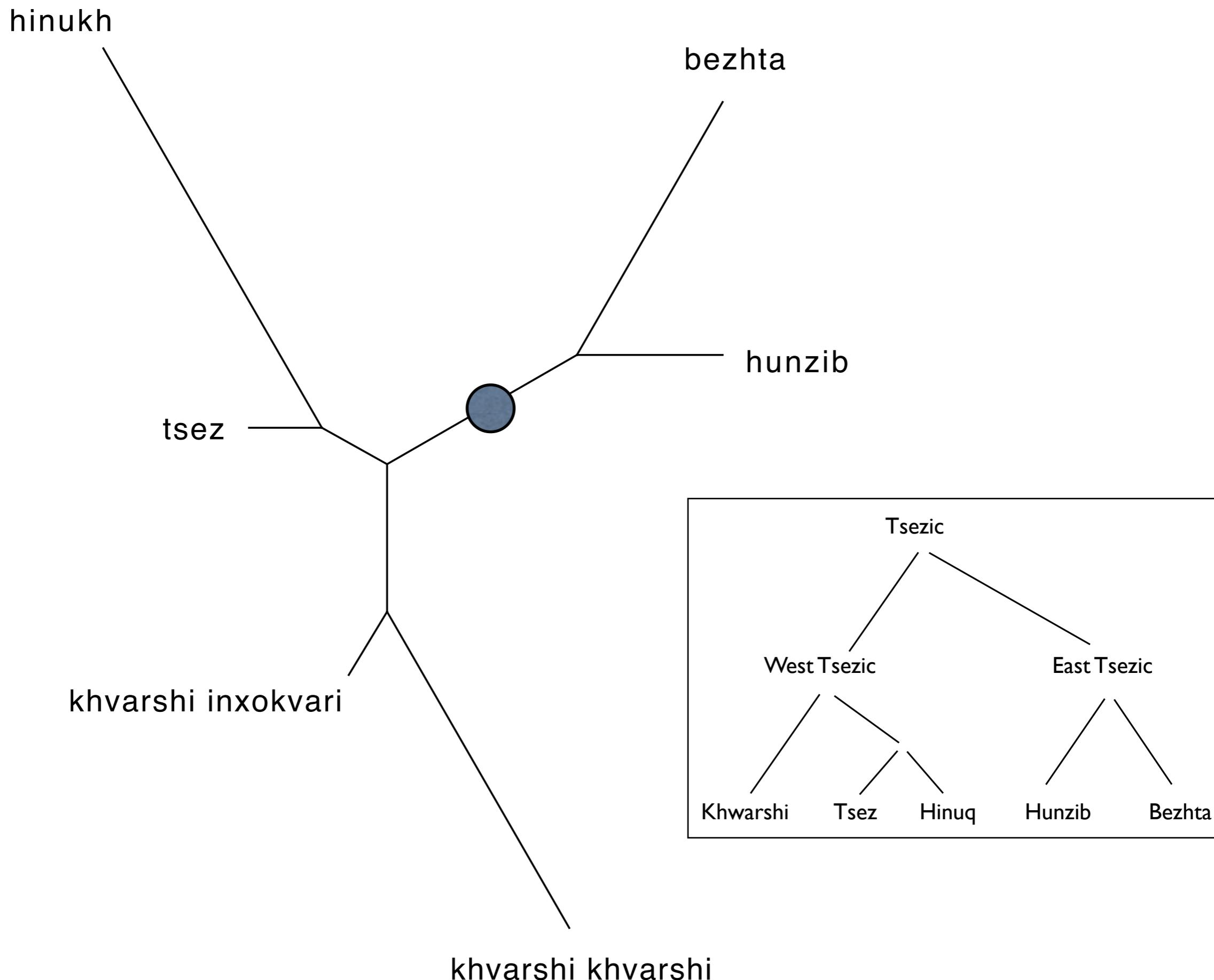
**Hunzib**

**Bezhta**

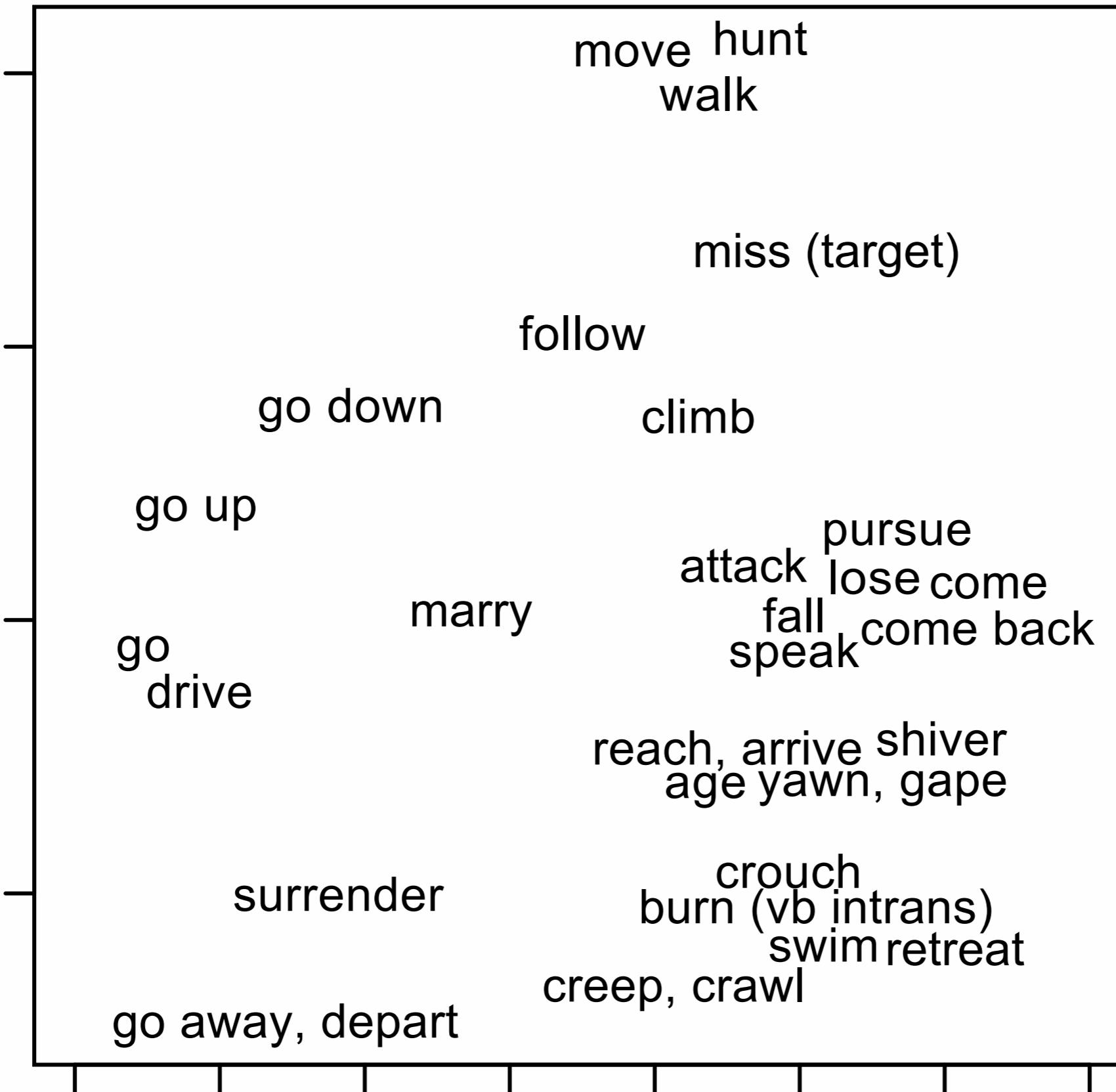


IDS number	English	Hinukh
2.330	marry	xoddoho yi᷑'a
10.470	go	i᷑'a
10.471	climb	šid i᷑'a
10.472	go up	toššid i᷑'a
10.473	go down	᷑'iyo i᷑'a
10.520	follow	hezzo i᷑'a
10.530	pursue	xalbun hezzo i᷑'a
10.650	drive	mašina᷑'o i᷑'a
10.660	ride	gulu᷑'o i᷑'a
18.210	speak	a᷑'i i᷑'a
20.430	attack	k <sup>w</sup> erer bi᷑'a
20.460	surrender	k <sup>w</sup> ar bi᷑'a

	bezhta	hinukh	hunzib	khvarshi inxokvari	khvarshi khvarshi	tsez
marry			0	0	0	
yawn	0	0	0	0		0
shiver	0	0	0	0		0
move		0		0	0	0
fall	0		0	0	0	0
creep	0	0	0	0	0	
walk		0		0	0	0
go						
go up					0	
climb	0			0	0	0
go down				0	0	
go away	0					
follow			0	0	0	
pursue	0		0	0	0	0
arrive	0	0	0		0	0
drive						
ride						
lose	0		0	0	0	0
age	0	0	0	0		0
speak	0		0	0	0	0
attack	0		0	0	0	0
retreat	0	0	0	0	0	0
surrender	0				0	
hunt		0		0	0	0
miss		0	0	0	0	0



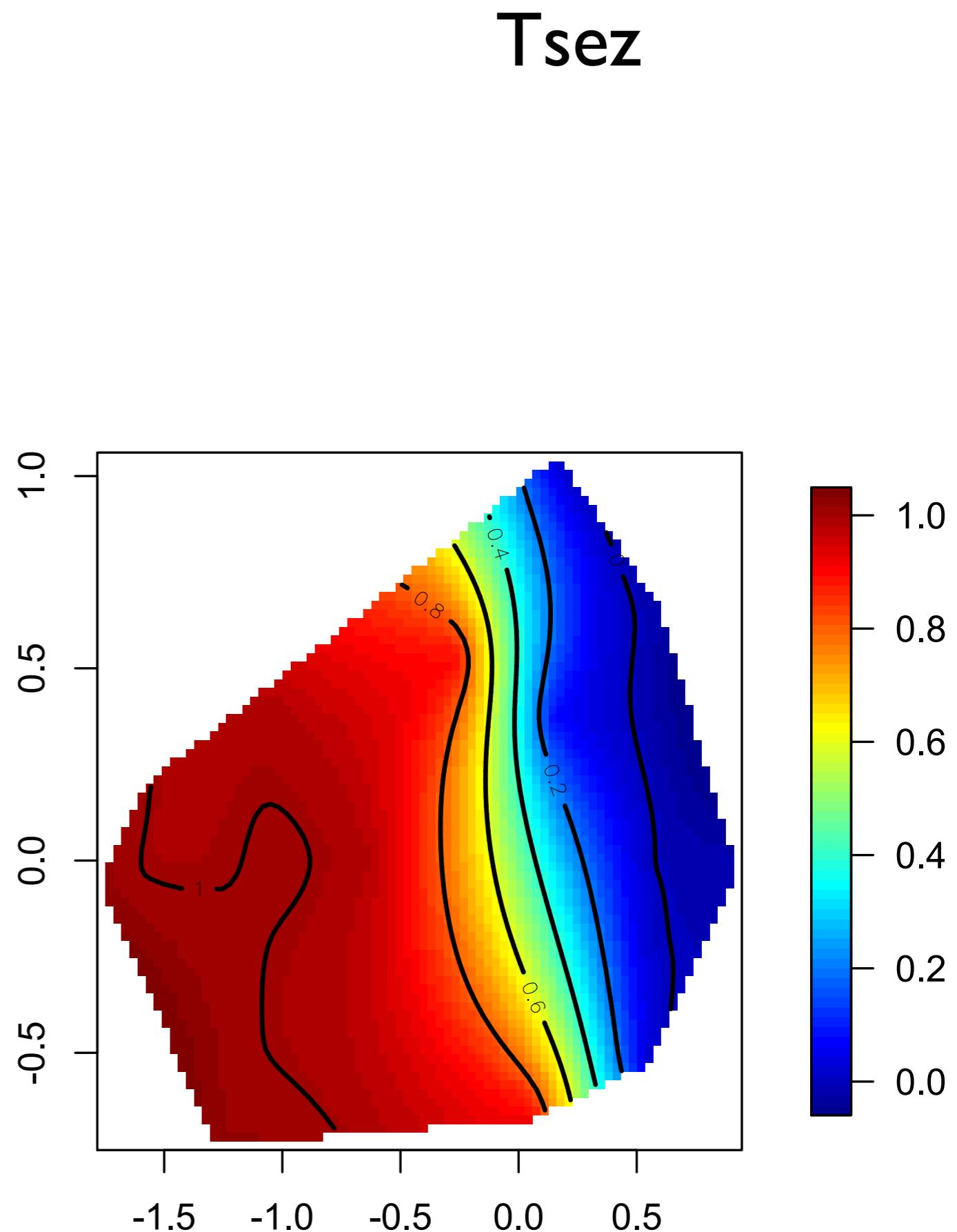
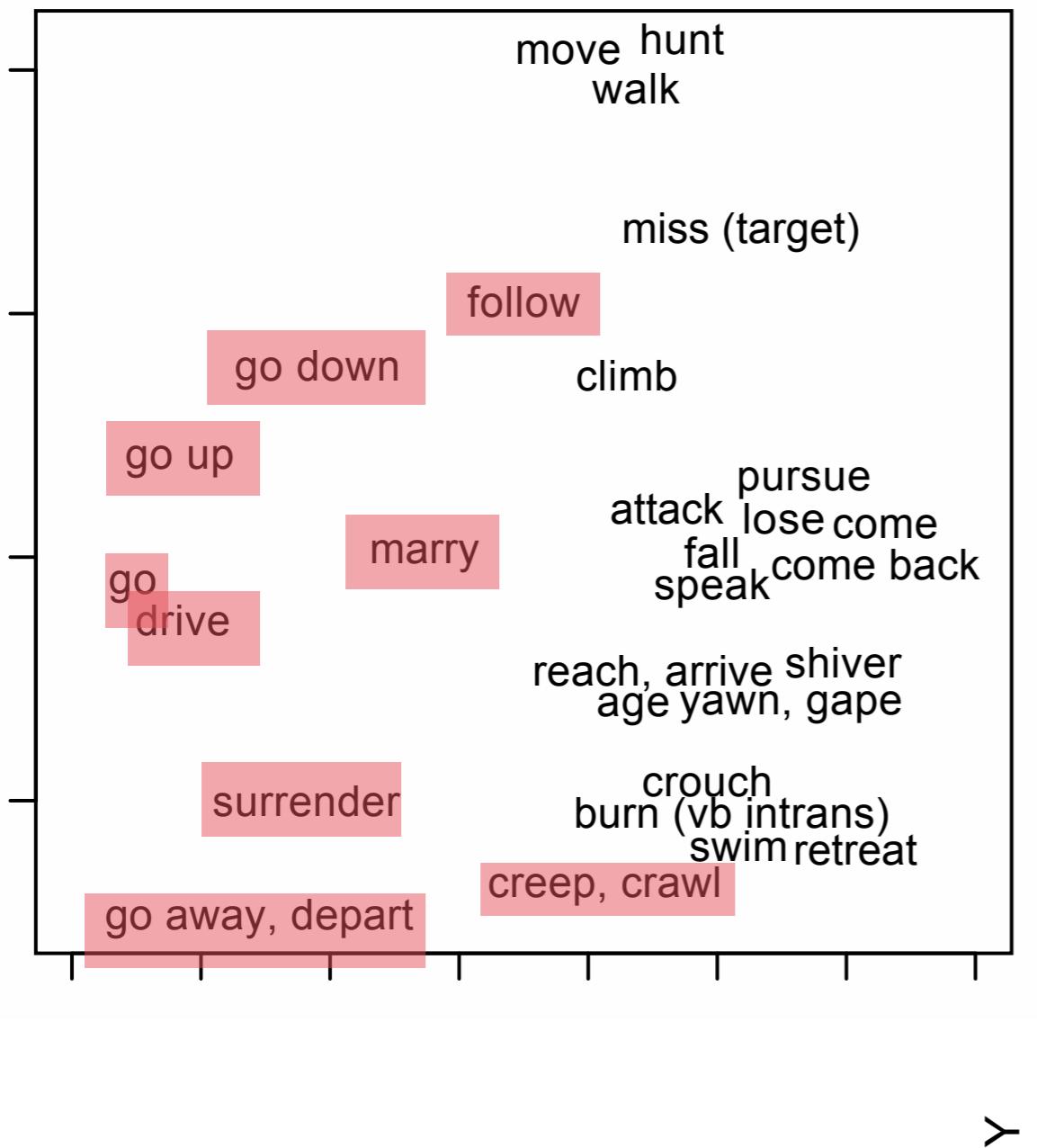
	bezhta	hinukh	hunzib	khvarshi inxokvari	khvarshi khvarshi	tsez
marry	I	I	0	0	0	I
yawn	0	0	0	0	I	0
shiver	0	0	0	0	I	0
move	I	0	I	0	0	0
fall	0	I	0	0	0	0
creep	0	0	0	0	0	I
walk	I	0	I	0	0	0
go	I	I	I	I	I	I
<b>go up</b>	I	I	I	I	0	I
climb	0	I	I	0	0	0
<b>go down</b>	I	I	I	0	0	I
go away	0	I	I	I	I	I
follow	I	I	0	0	0	I
pursue	0	I	0	0	0	0
arrive	0	0	0	I	0	0
drive	I	I	I	I	I	I
ride	I	I	I	I	I	I
lose	0	I	0	0	0	0
age	0	0	0	0	I	0
speak	0	I	0	0	0	0
attack	0	I	0	0	0	0
retreat	0	0	0	0	0	0
surrender	0	I	I	I	0	I
hunt	I	0	I	0	0	0
miss	I	0	0	0	0	0

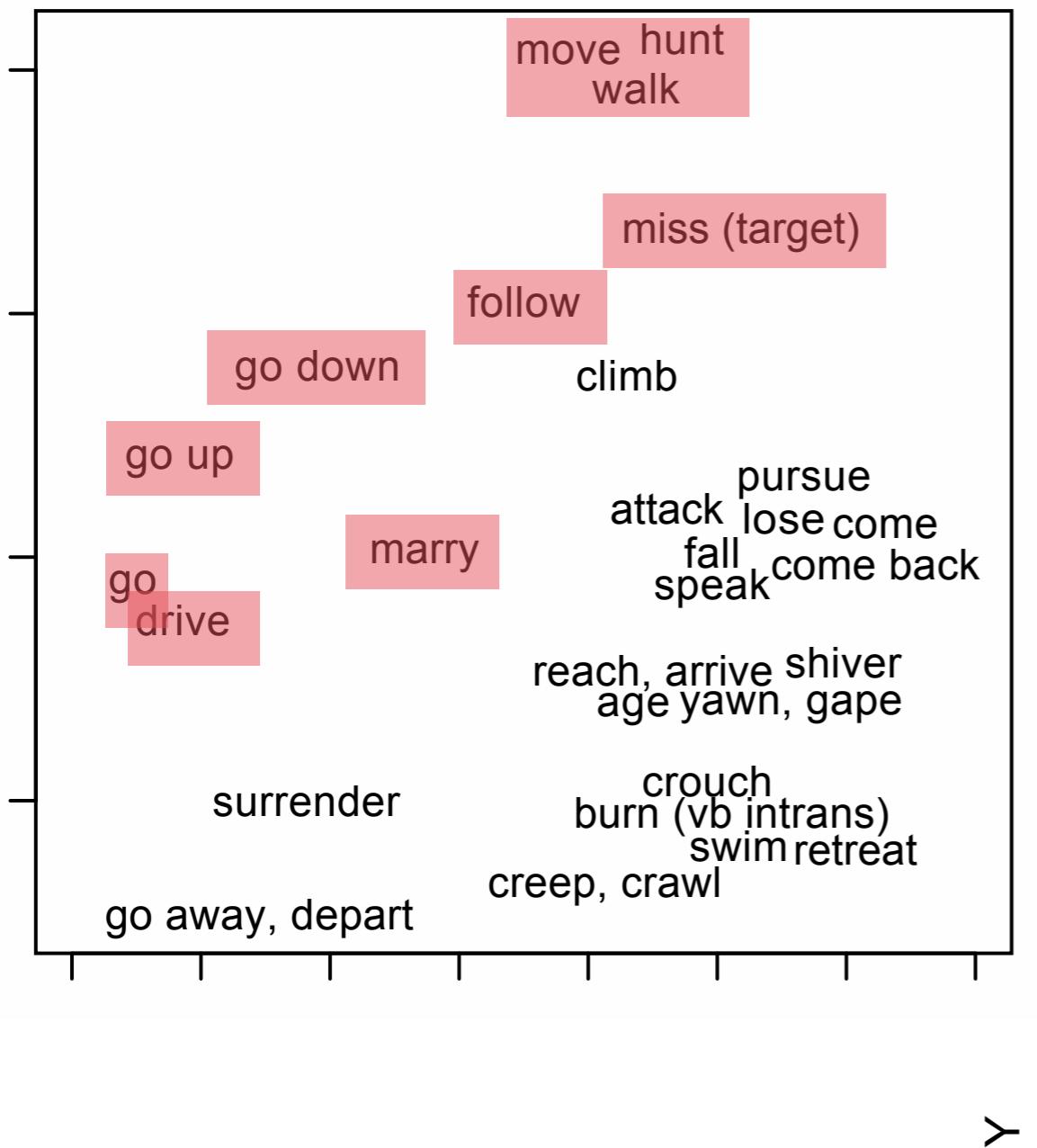


A scatter plot showing the distribution of verb meanings across a 2D space defined by two axes. The vertical axis has tick marks at the top and bottom. The horizontal axis has tick marks at the left and right. The plot area is bounded by a black rectangular frame.

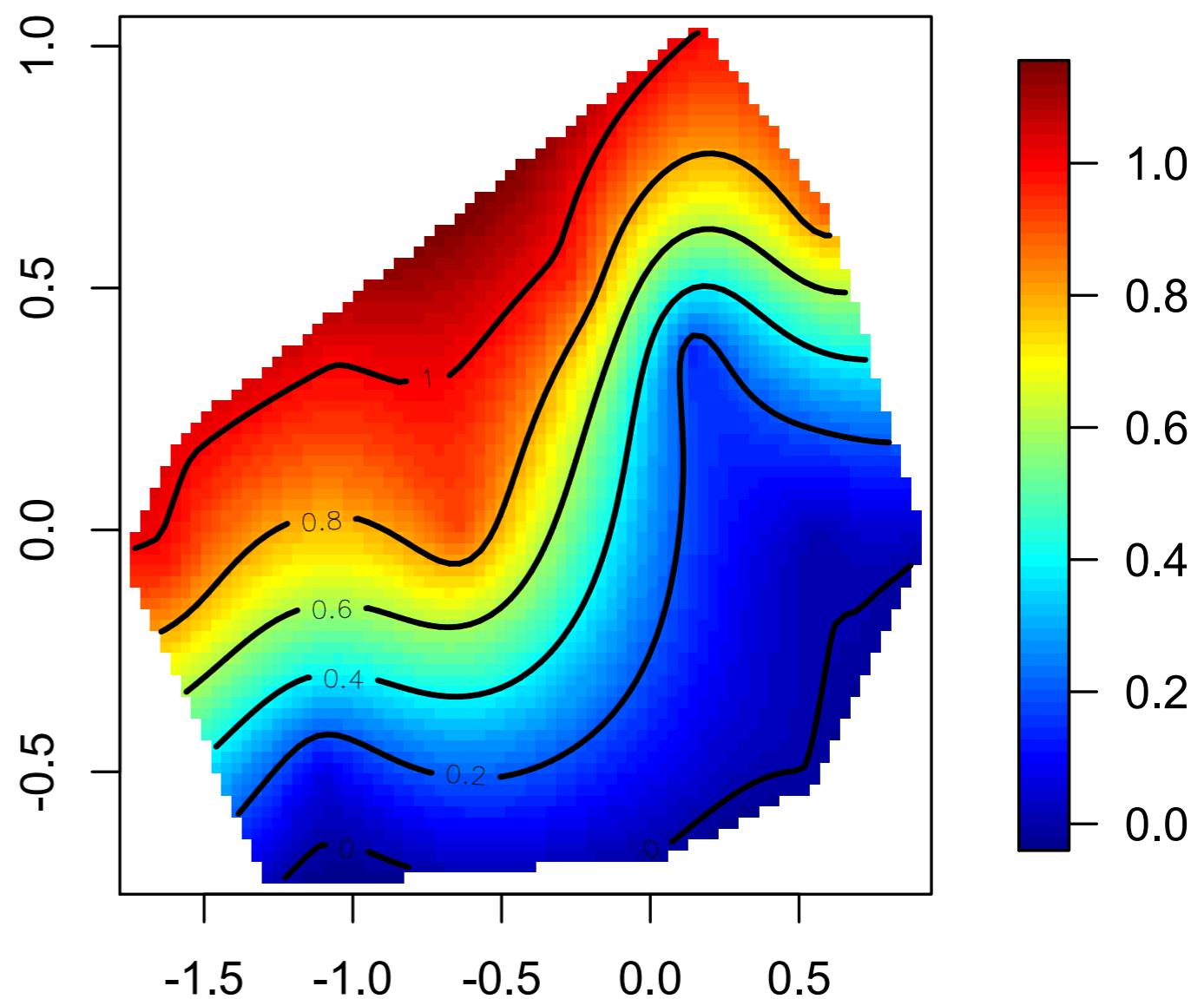
The verbs are plotted as follows:

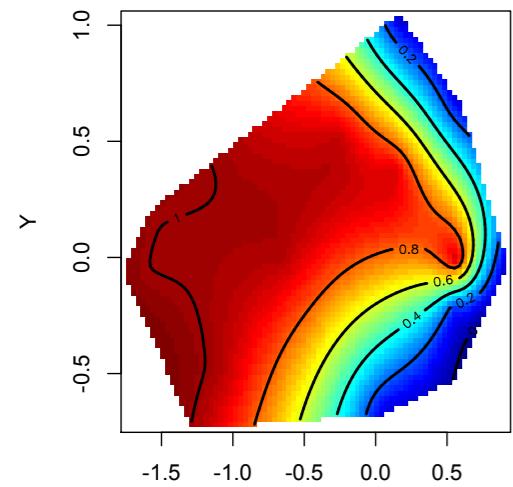
- Top row (vertical axis tick mark level):
  - move
  - hunt
  - walk
- Second row (vertical axis tick mark level):
  - miss (target)
- Third row (vertical axis tick mark level):
  - follow
- Fourth row (vertical axis tick mark level):
  - go down
  - climb
- Fifth row (vertical axis tick mark level):
  - go up
  - attack
  - pursue
  - lose
  - come
- Sixth row (vertical axis tick mark level):
  - go
  - marry
  - fall
  - speak
  - come back
- Seventh row (vertical axis tick mark level):
  - drive
  - reach, arrive
  - shiver
  - age
  - yawn, gape
- Eighth row (vertical axis tick mark level):
  - surrender
  - crouch
  - burn (vb intrans)
  - swim
  - retreat
- Ninth row (vertical axis tick mark level):
  - go away, depart
  - creep, crawl



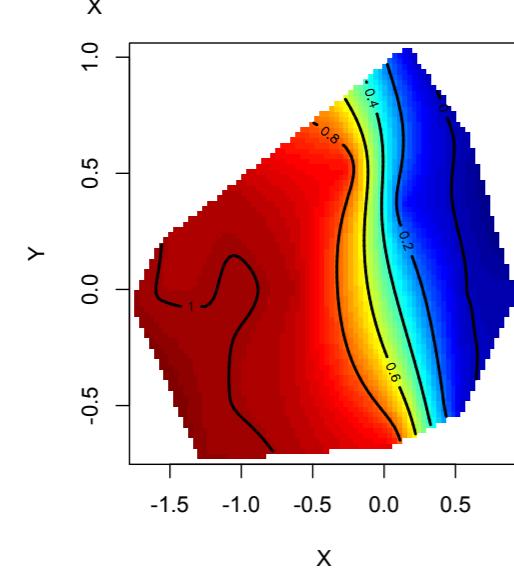


# Bezhta

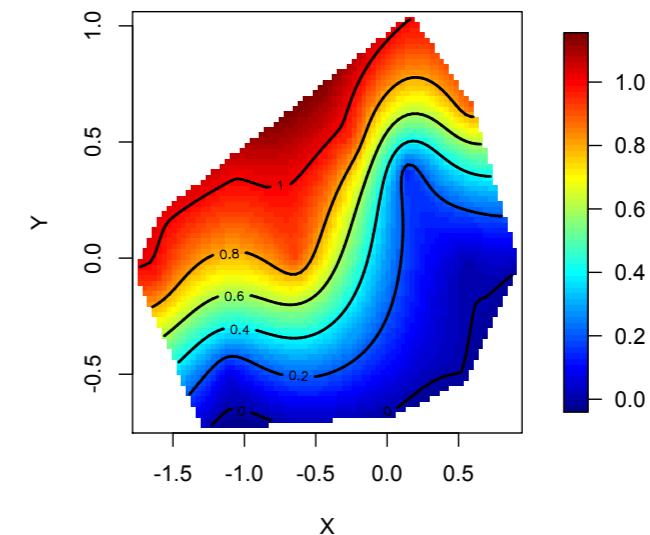




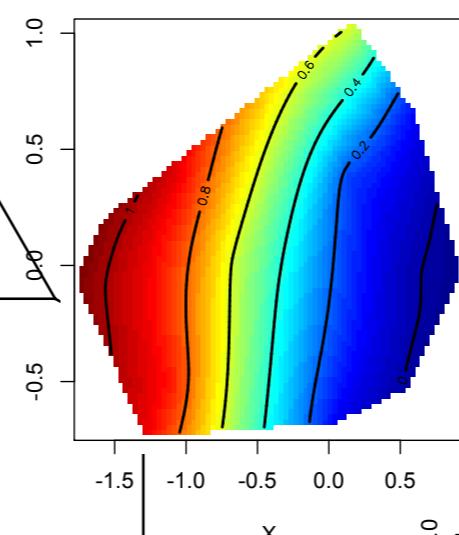
hinukh



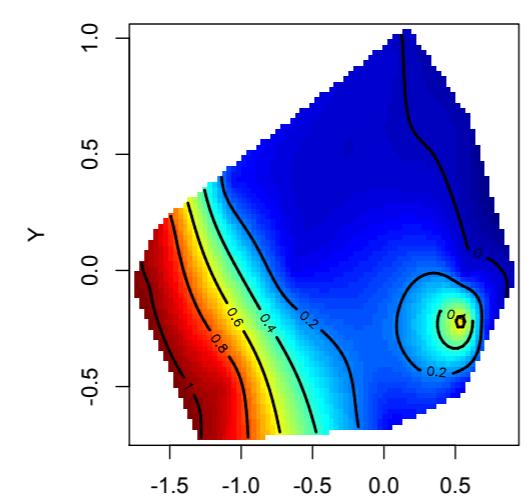
tsez



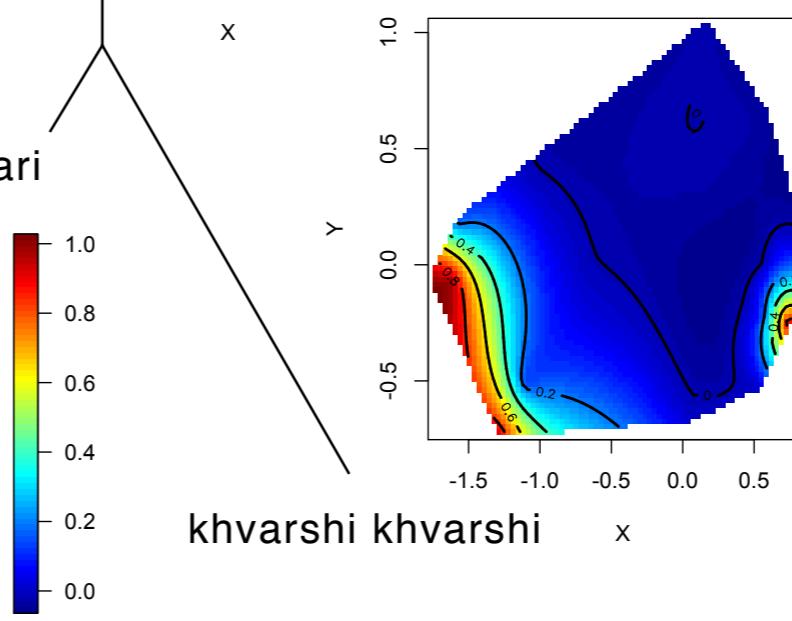
bezhta



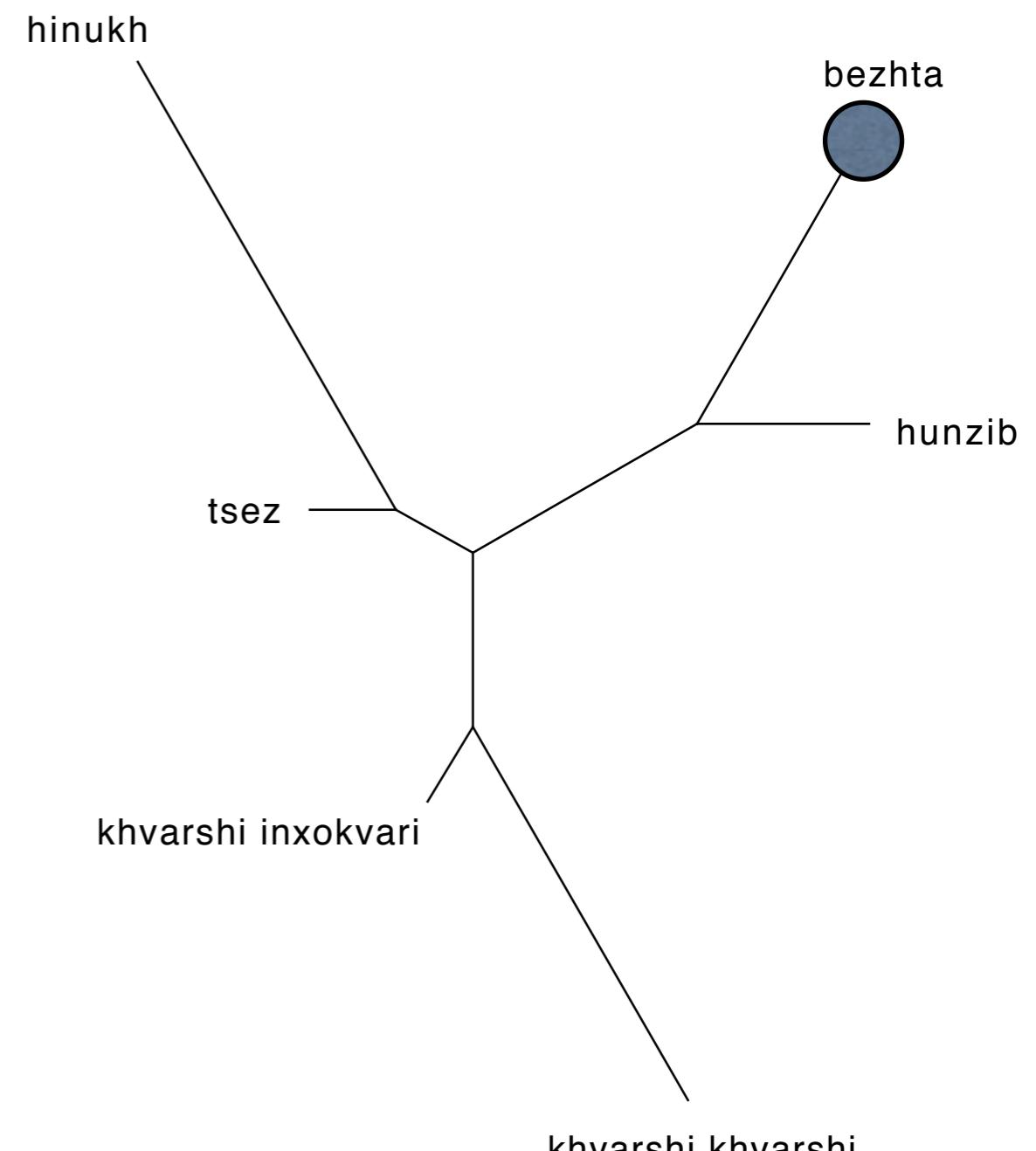
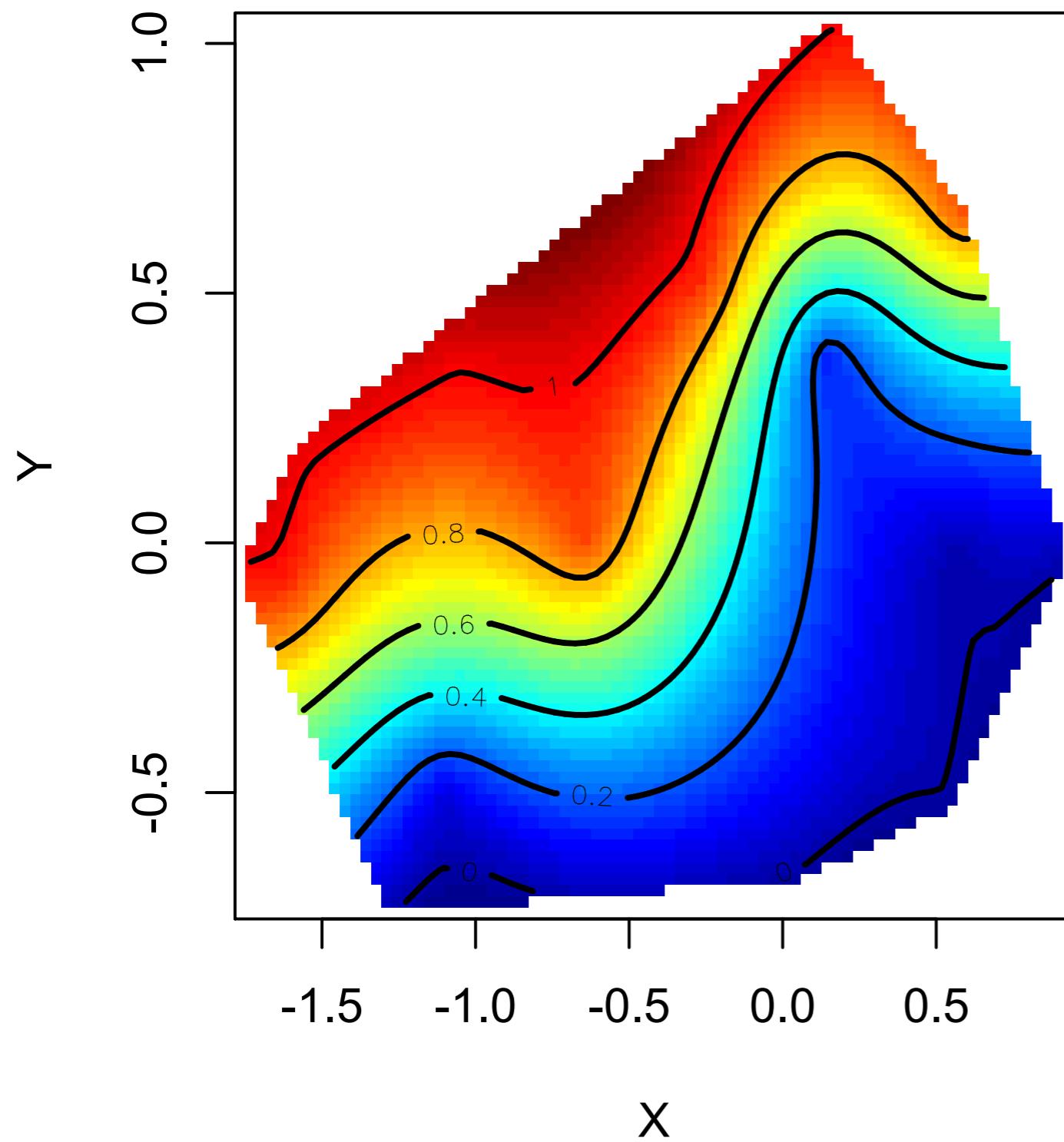
hunzib

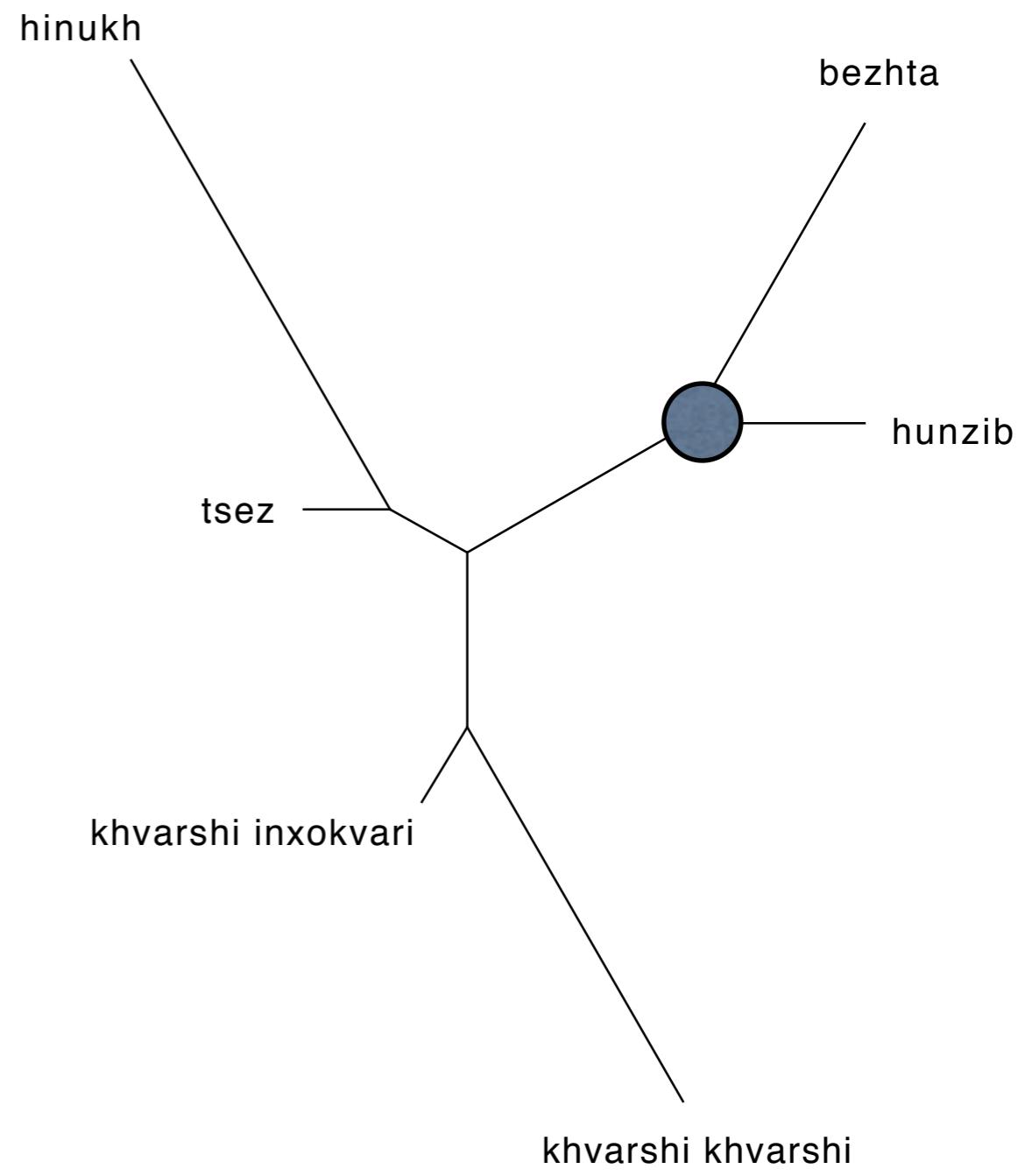
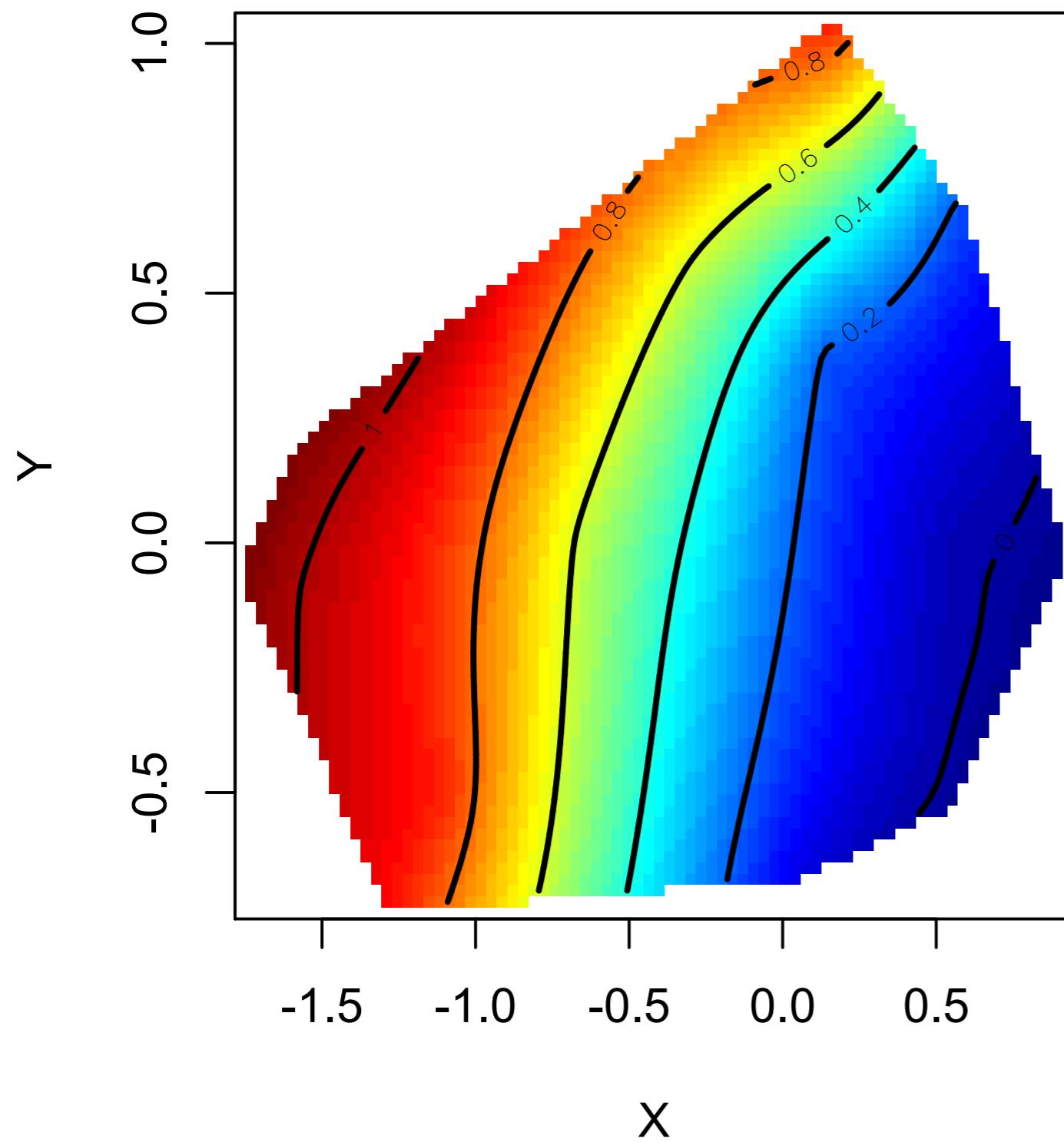


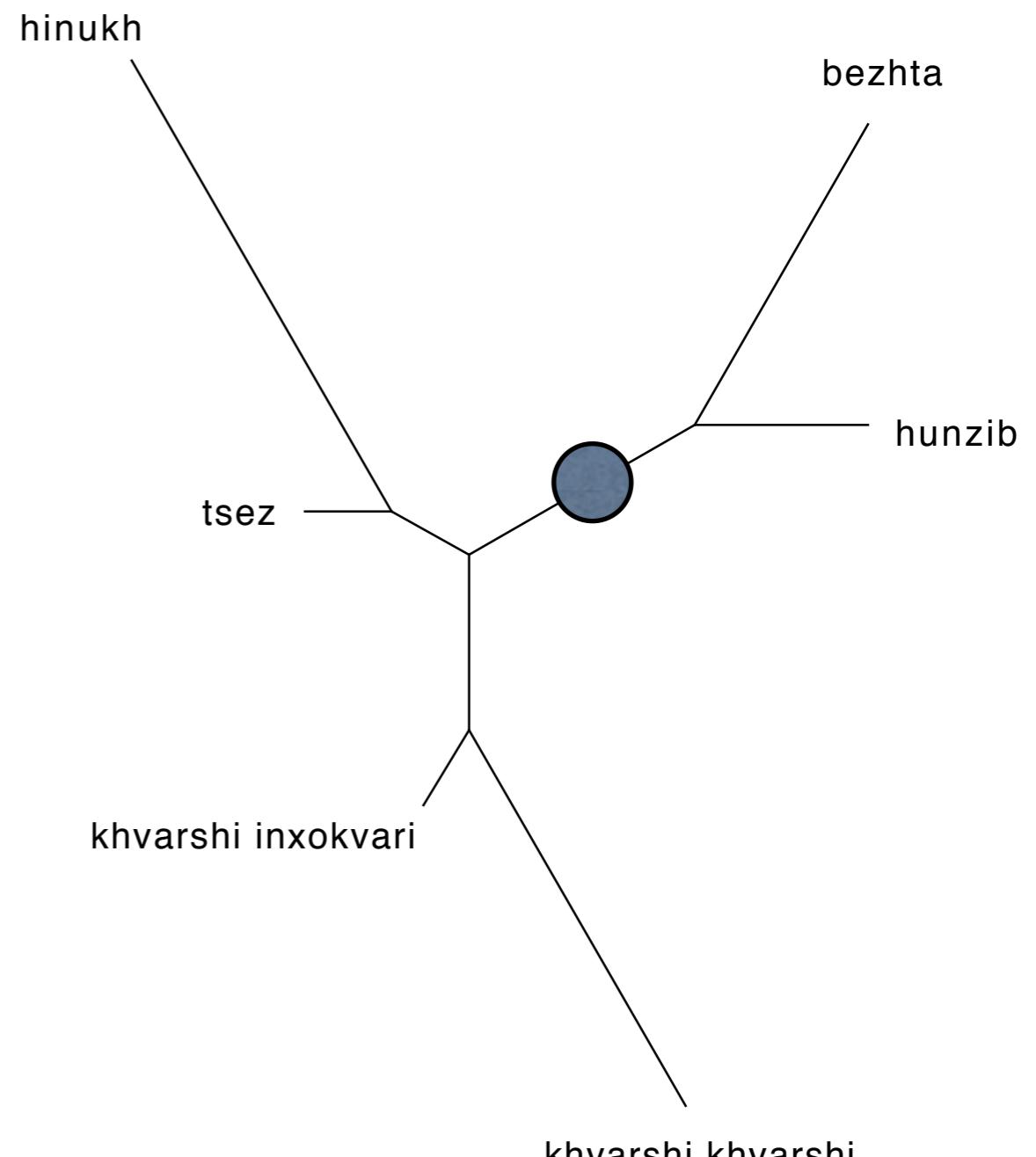
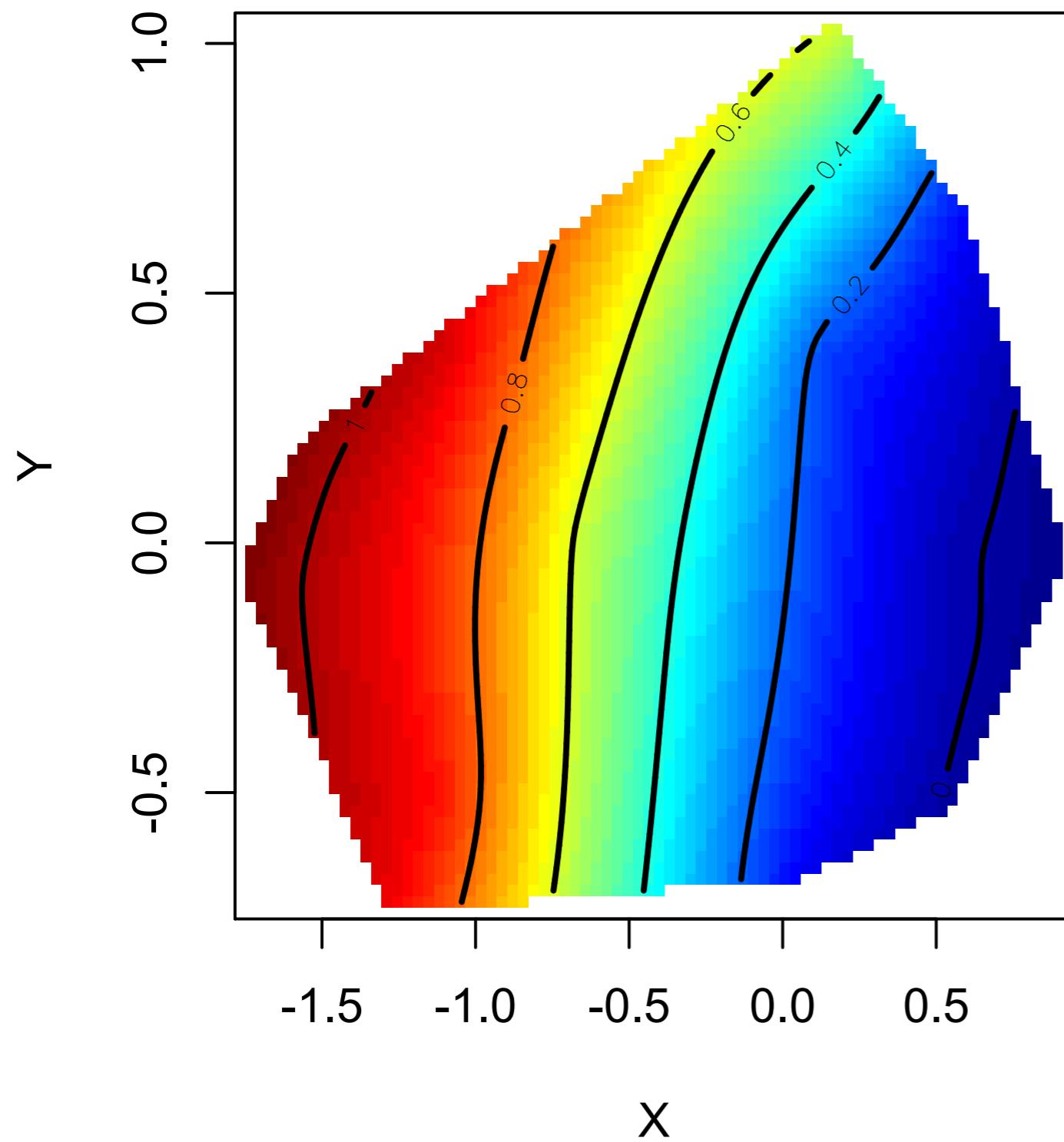
khvarshi inxokvari

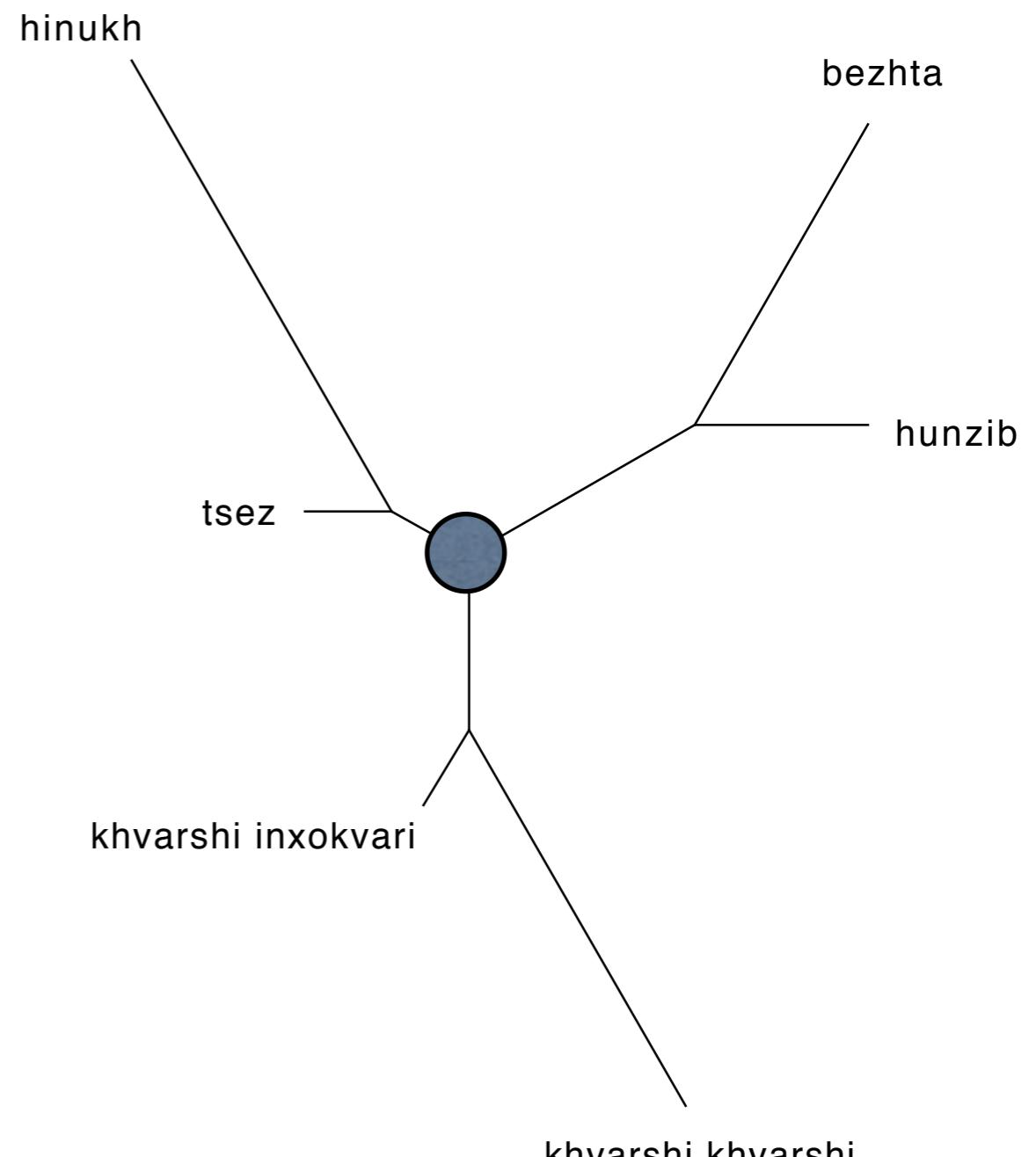
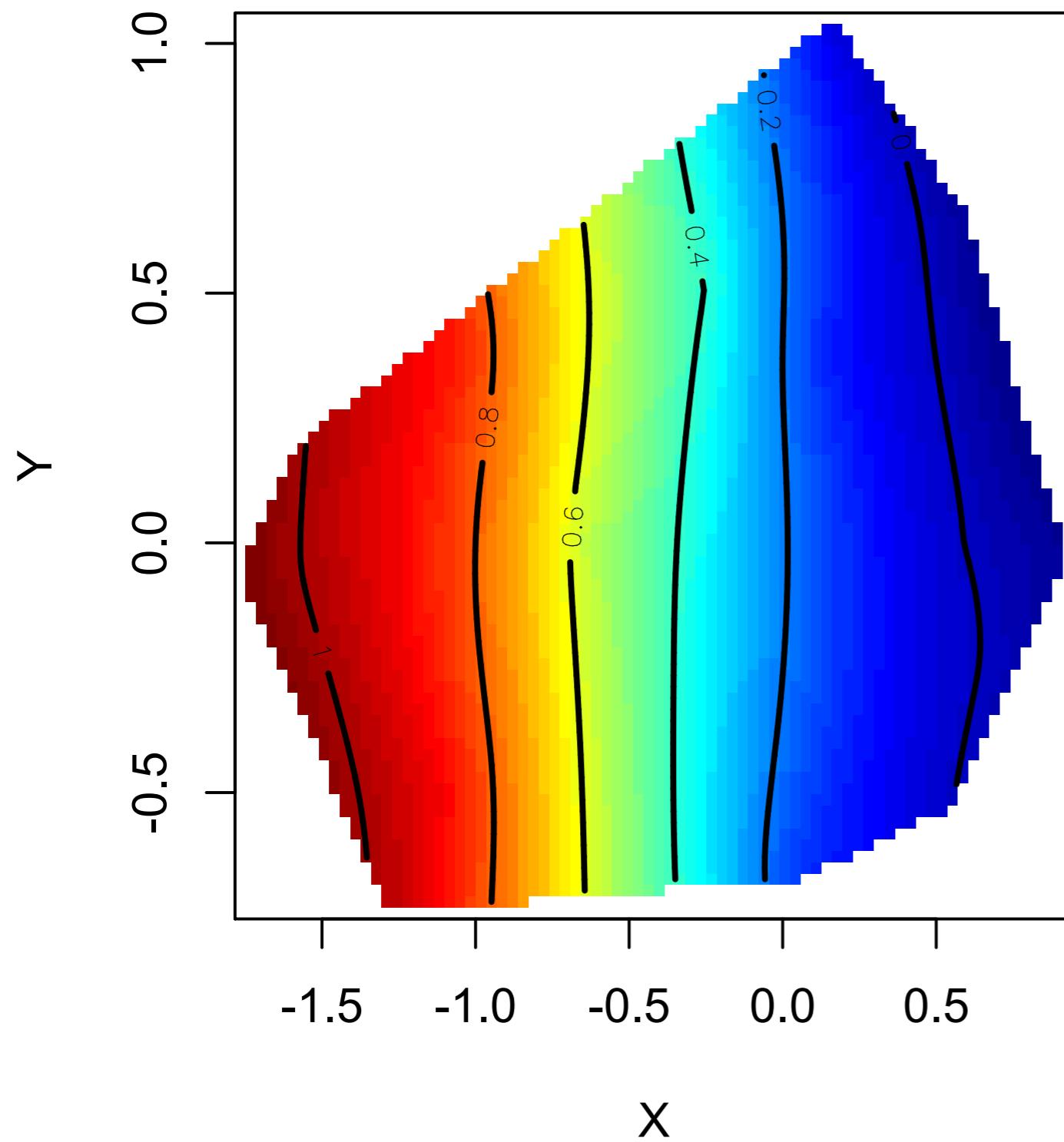


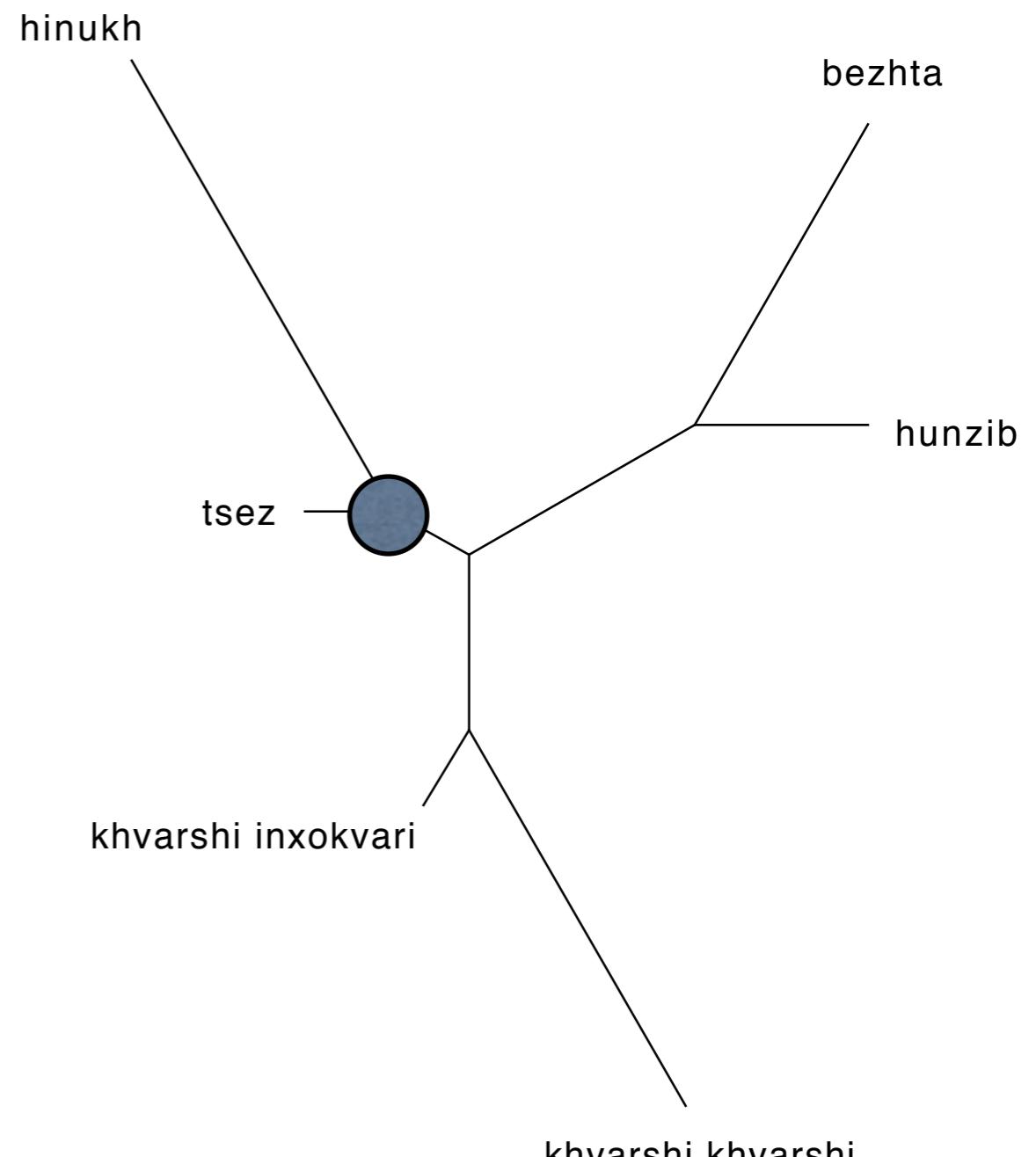
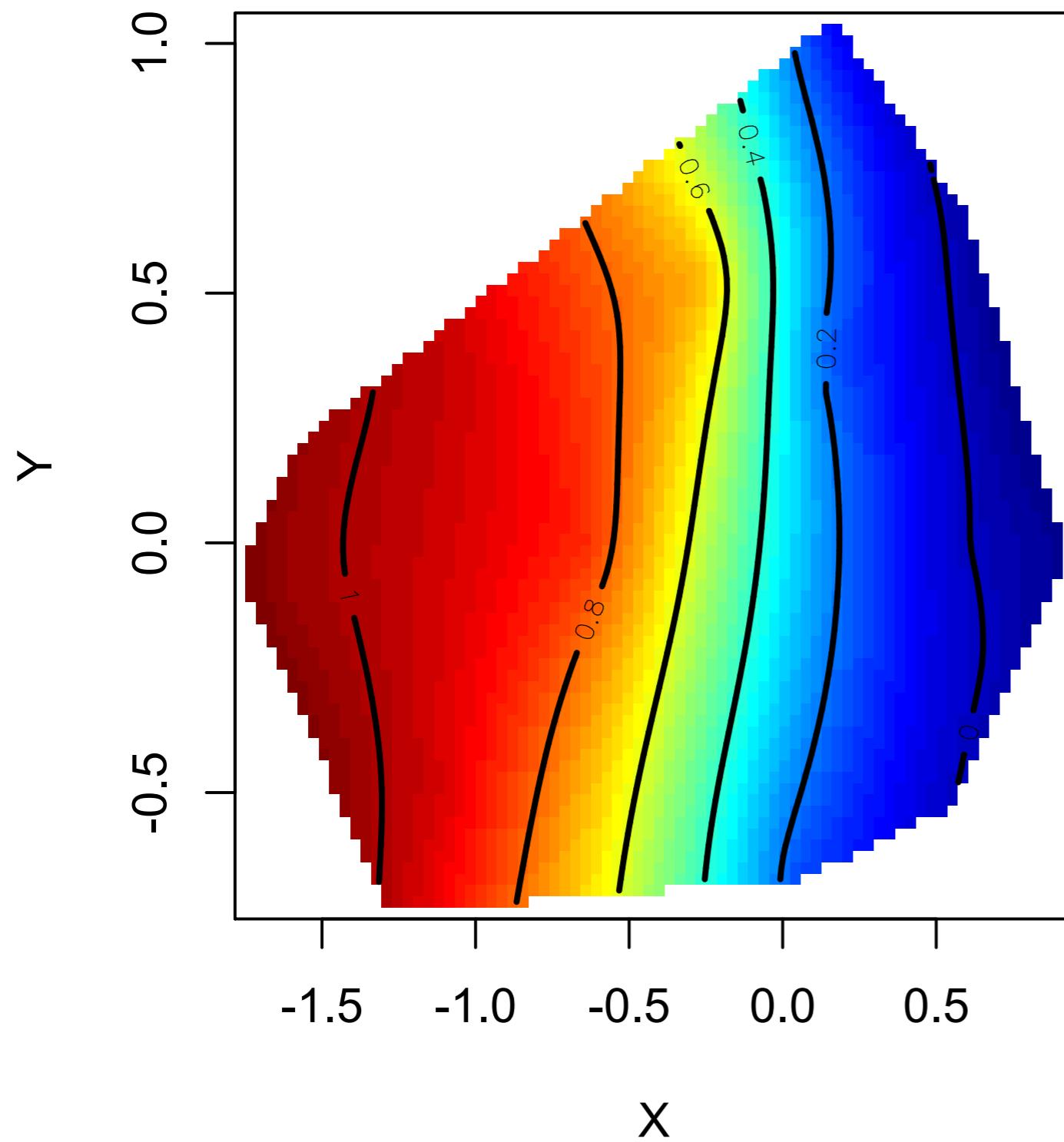
khvarshi khvarshi

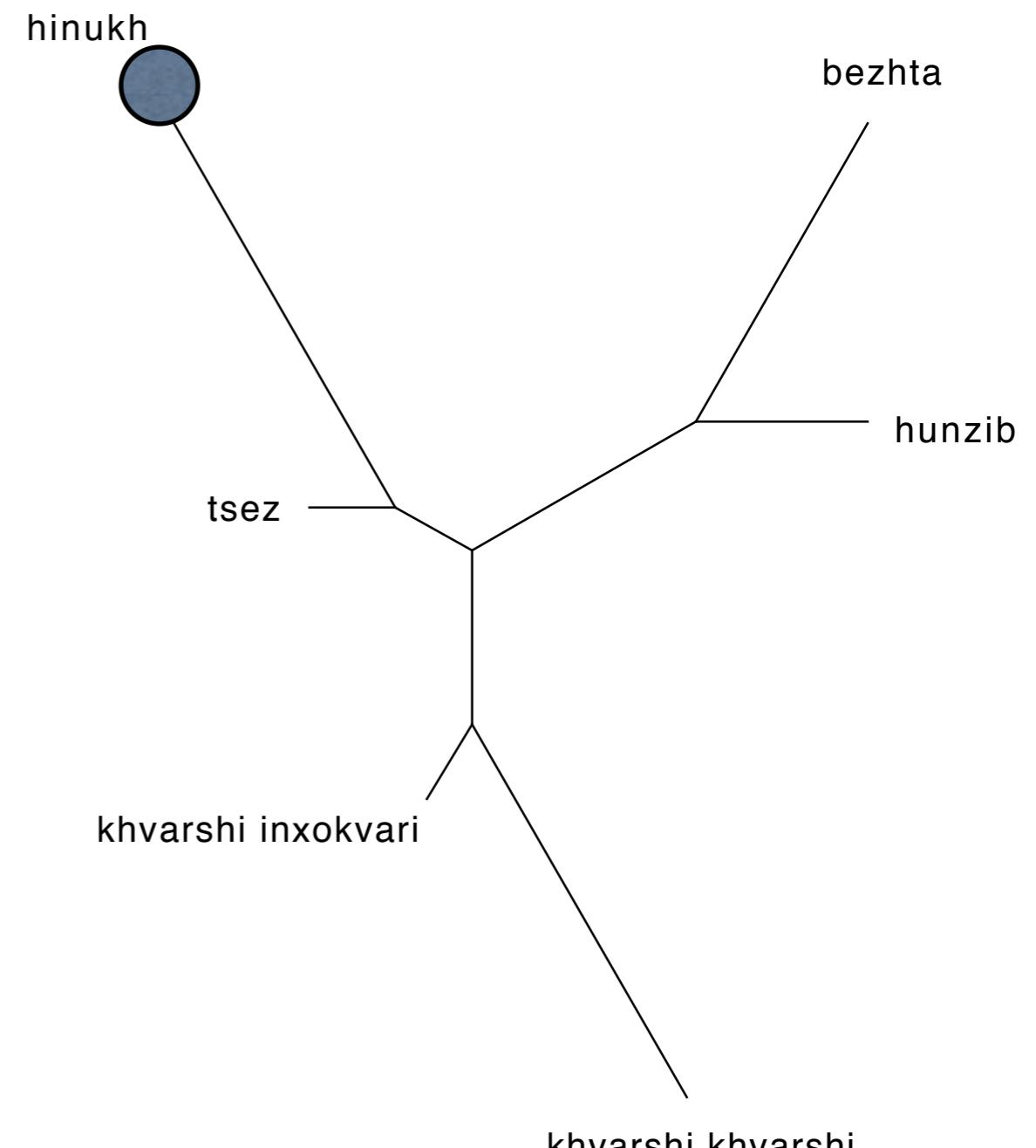
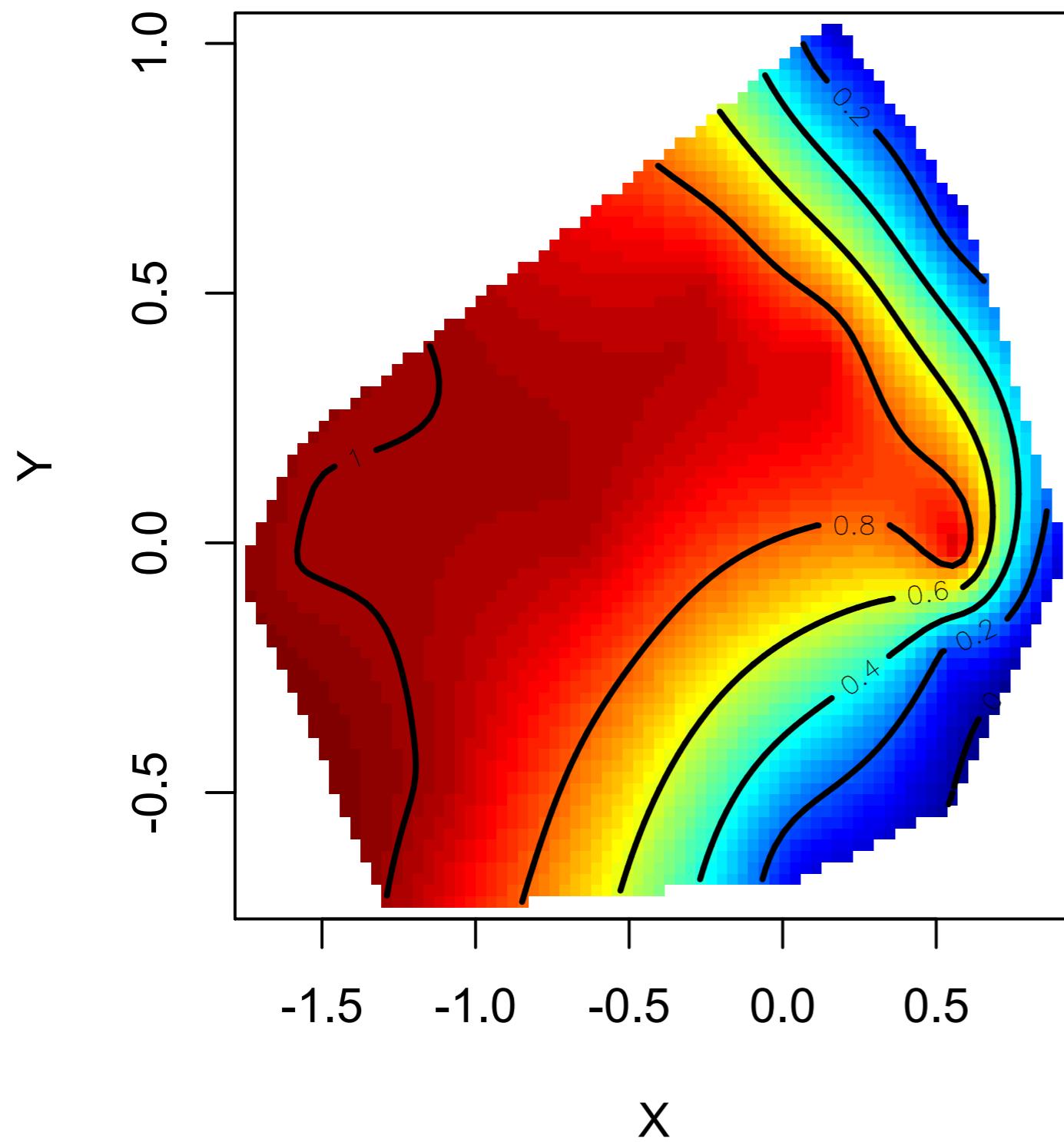


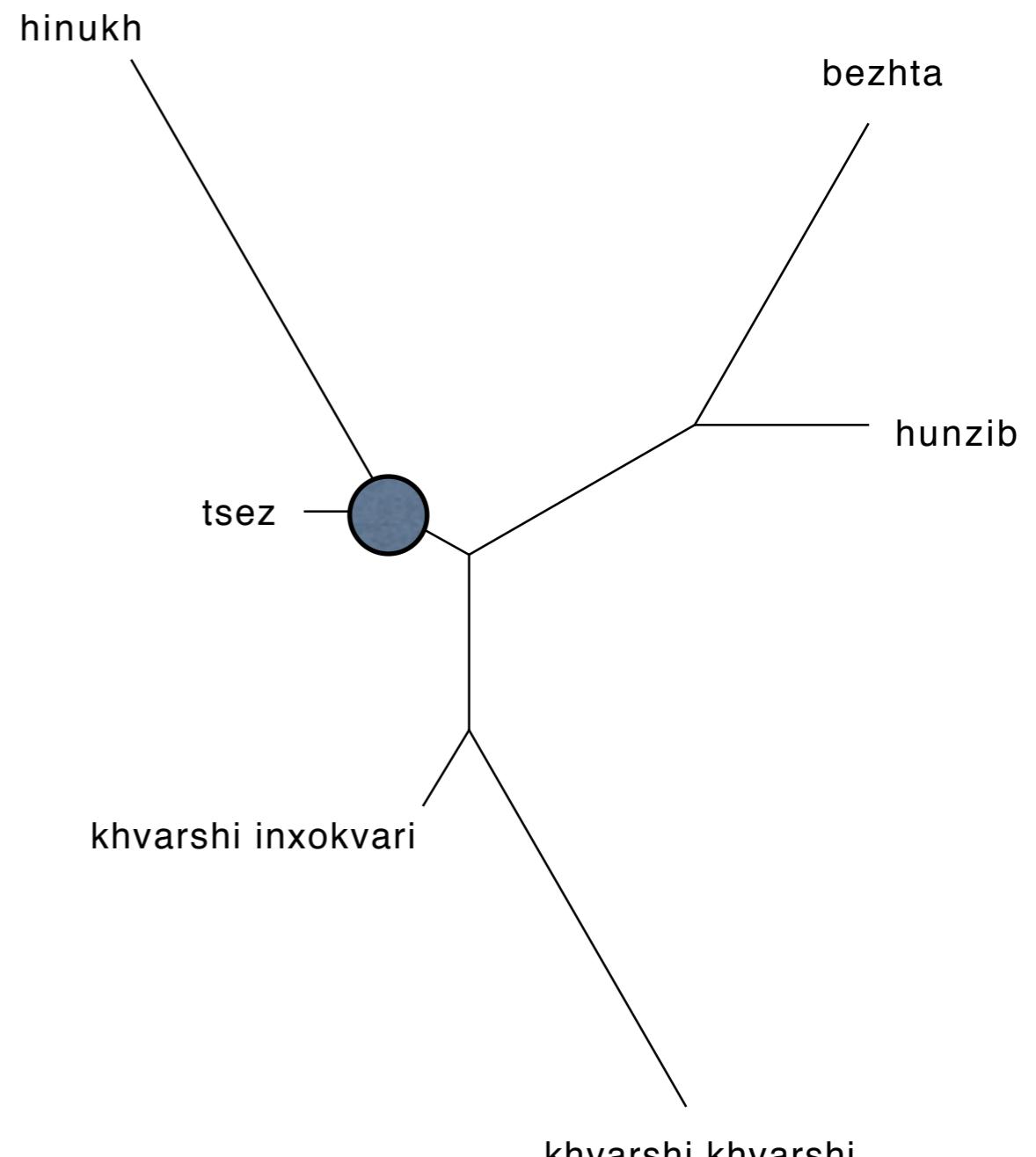
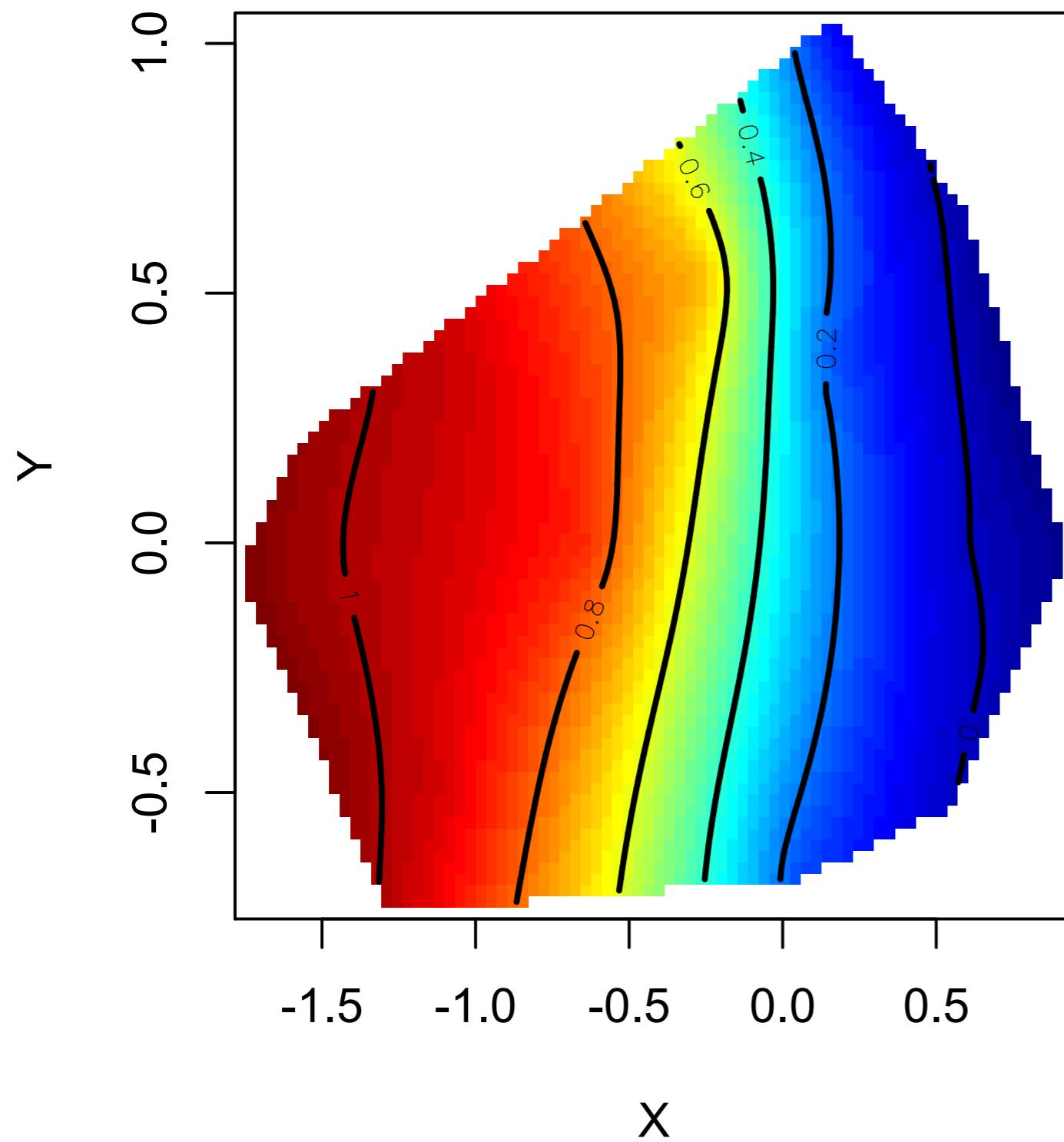


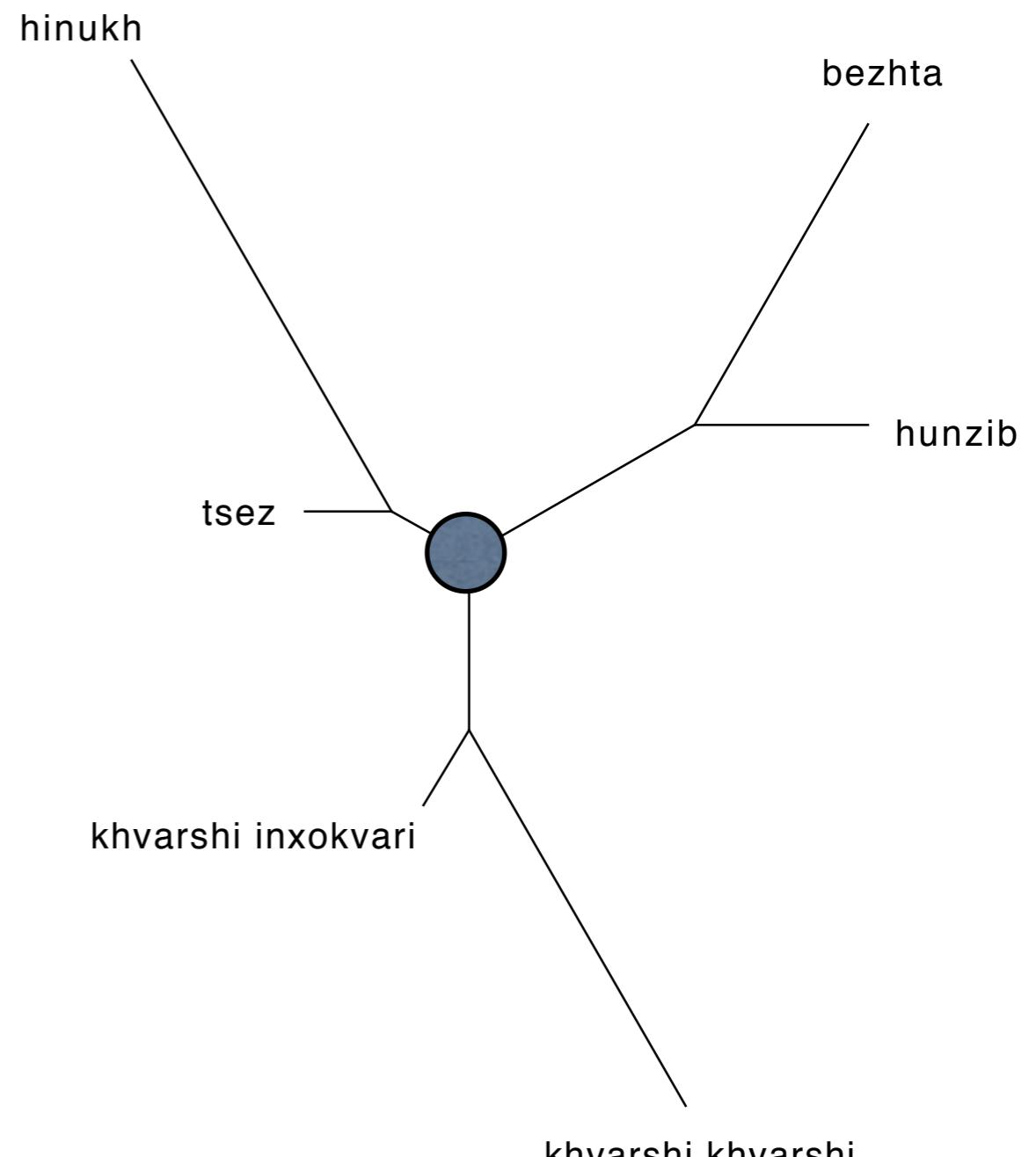
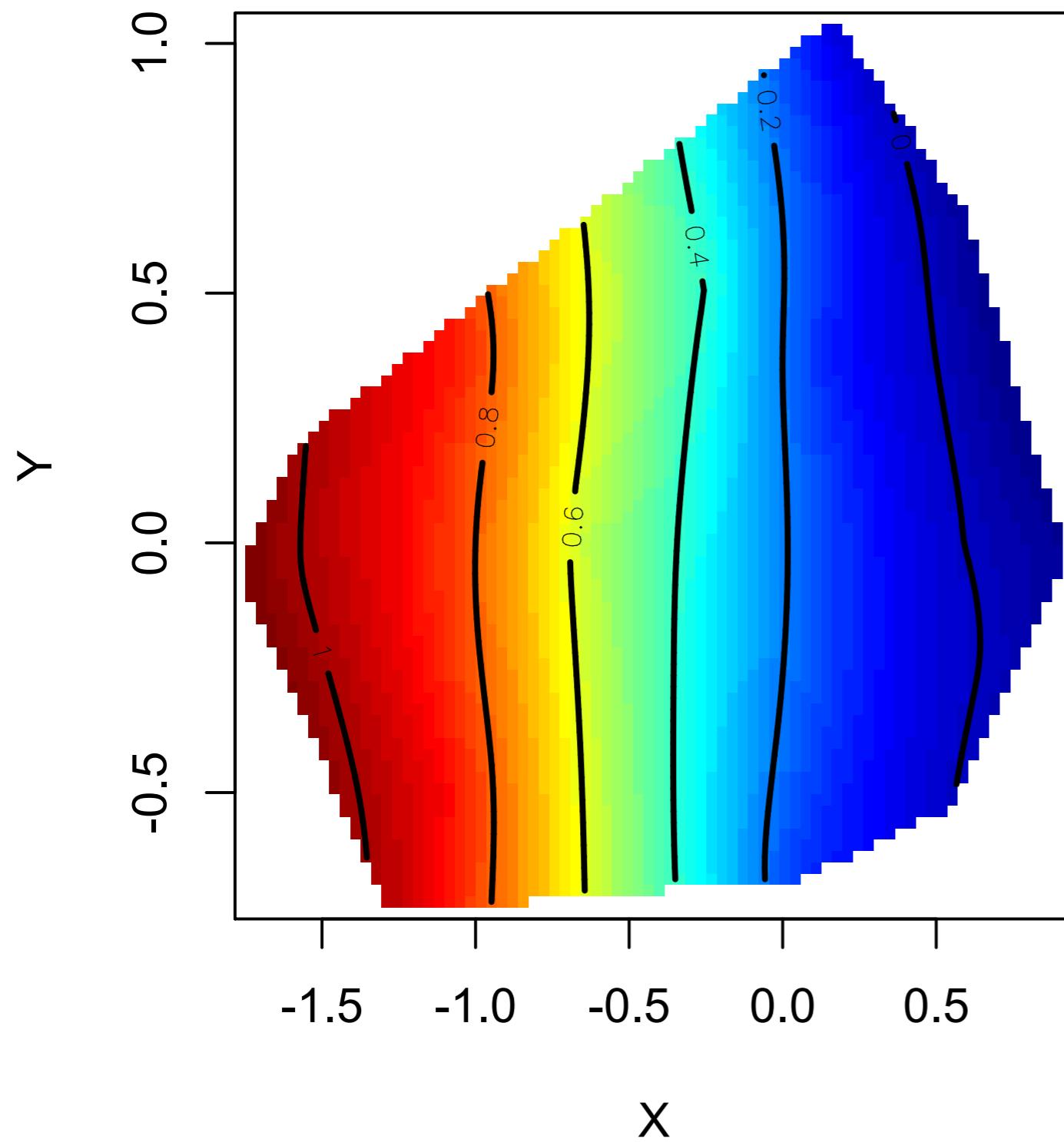


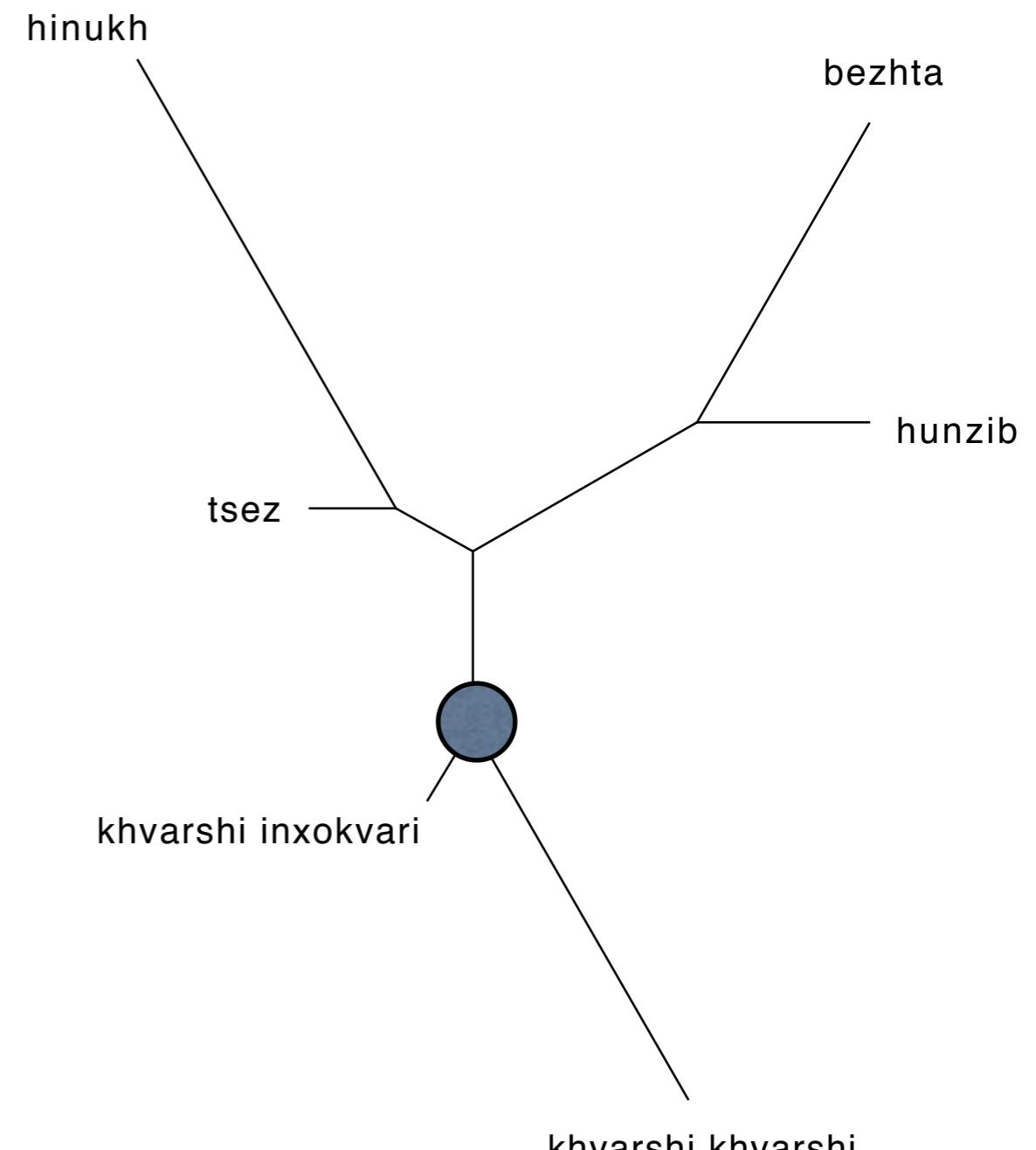
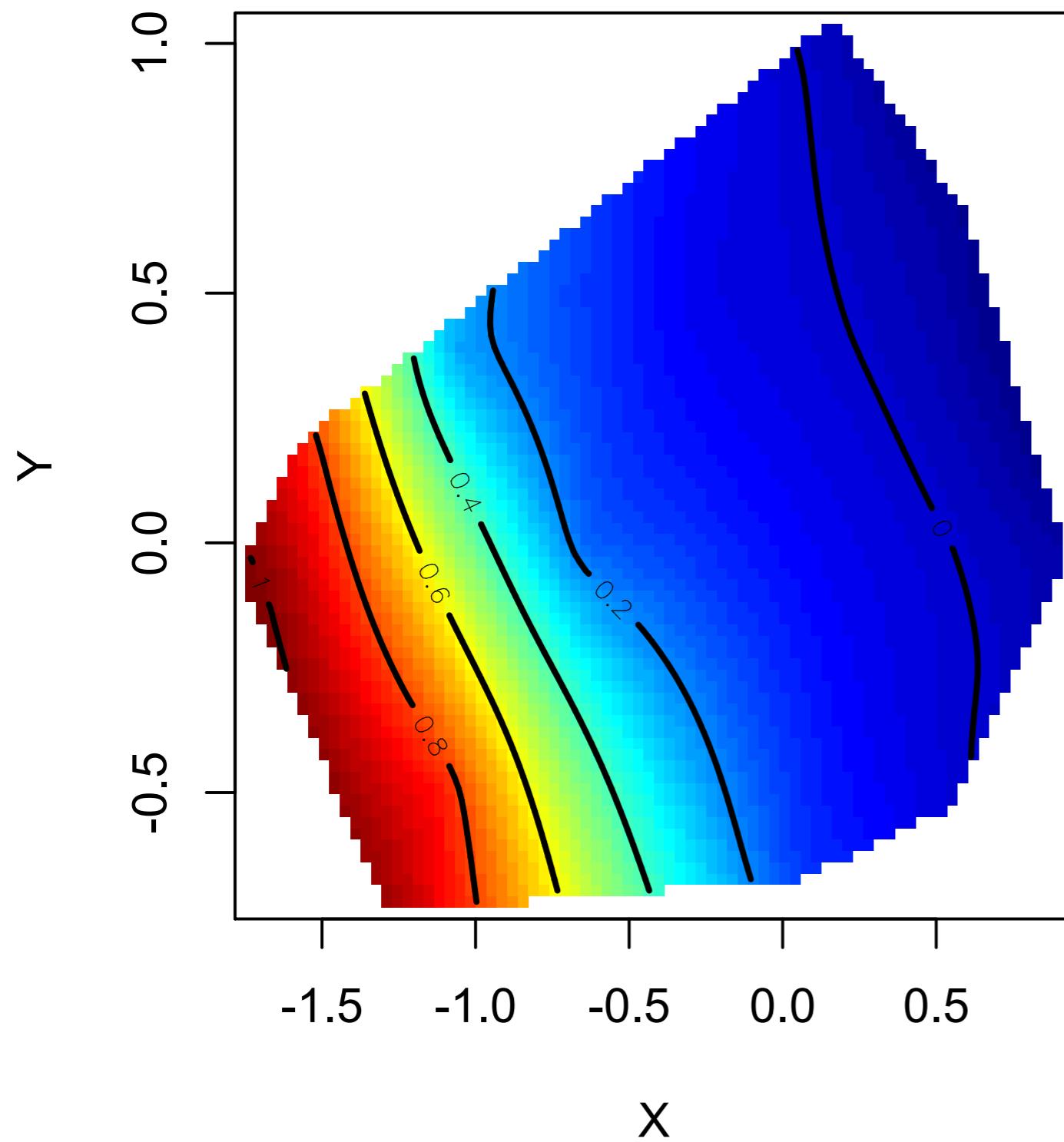


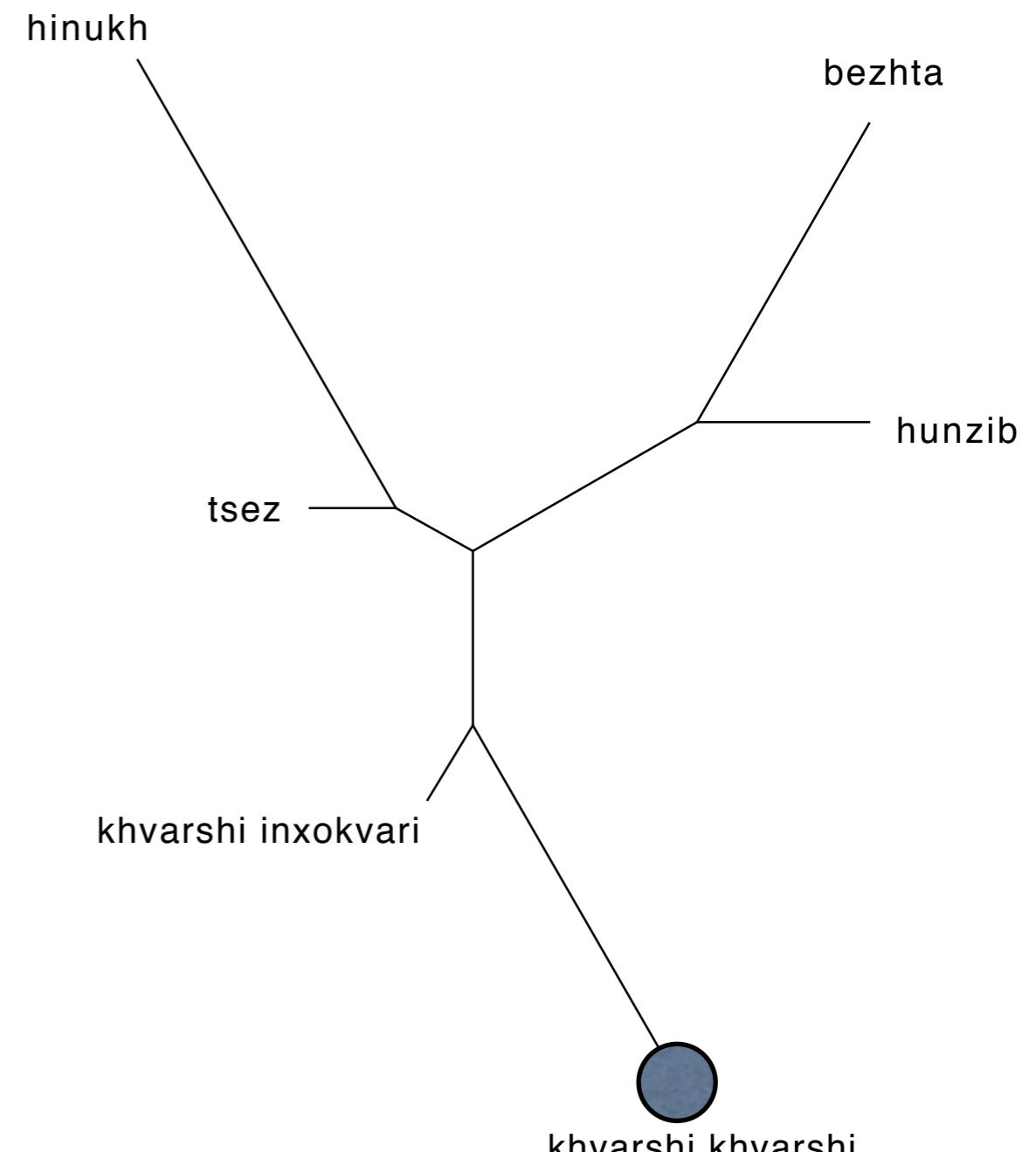
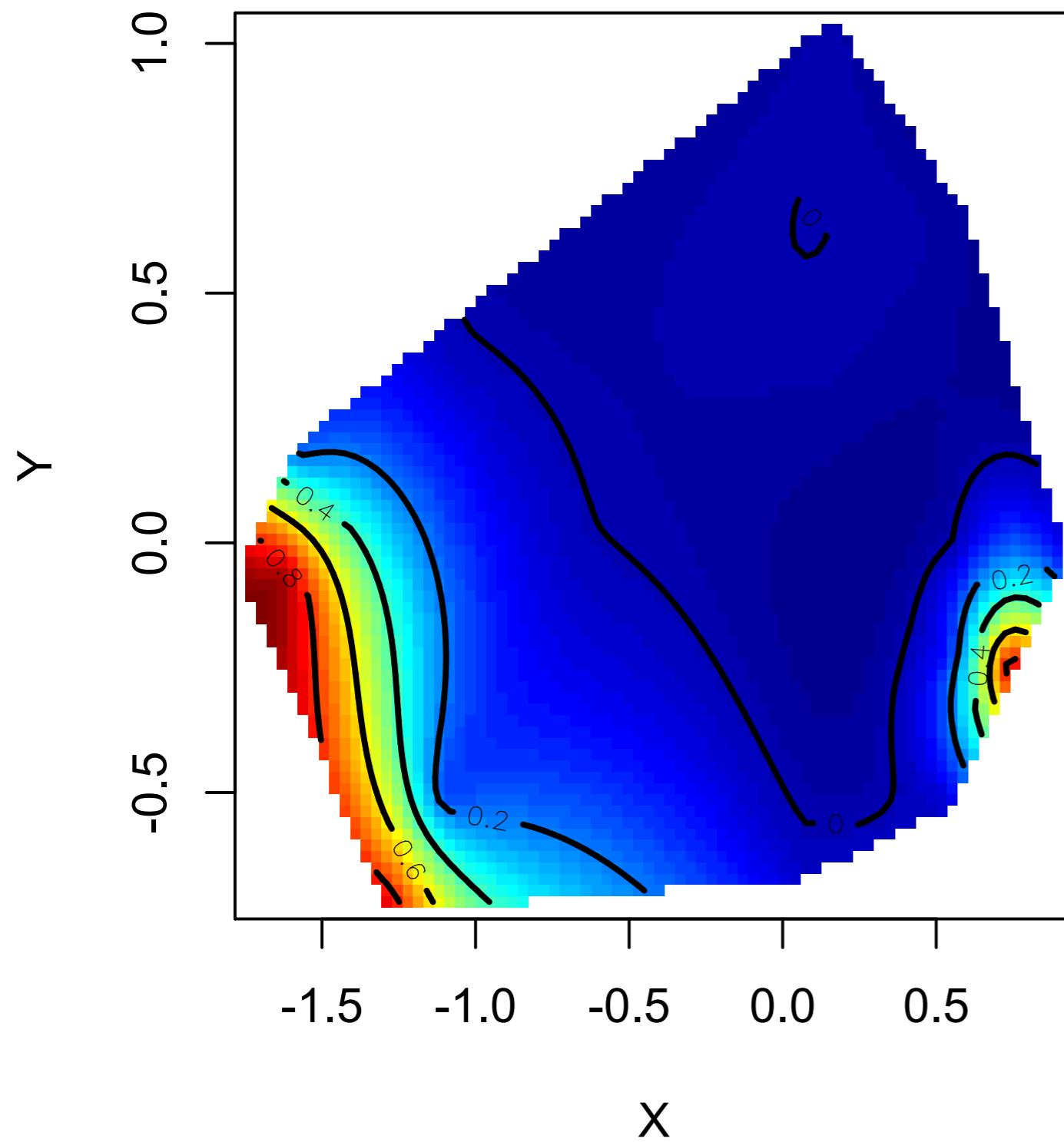


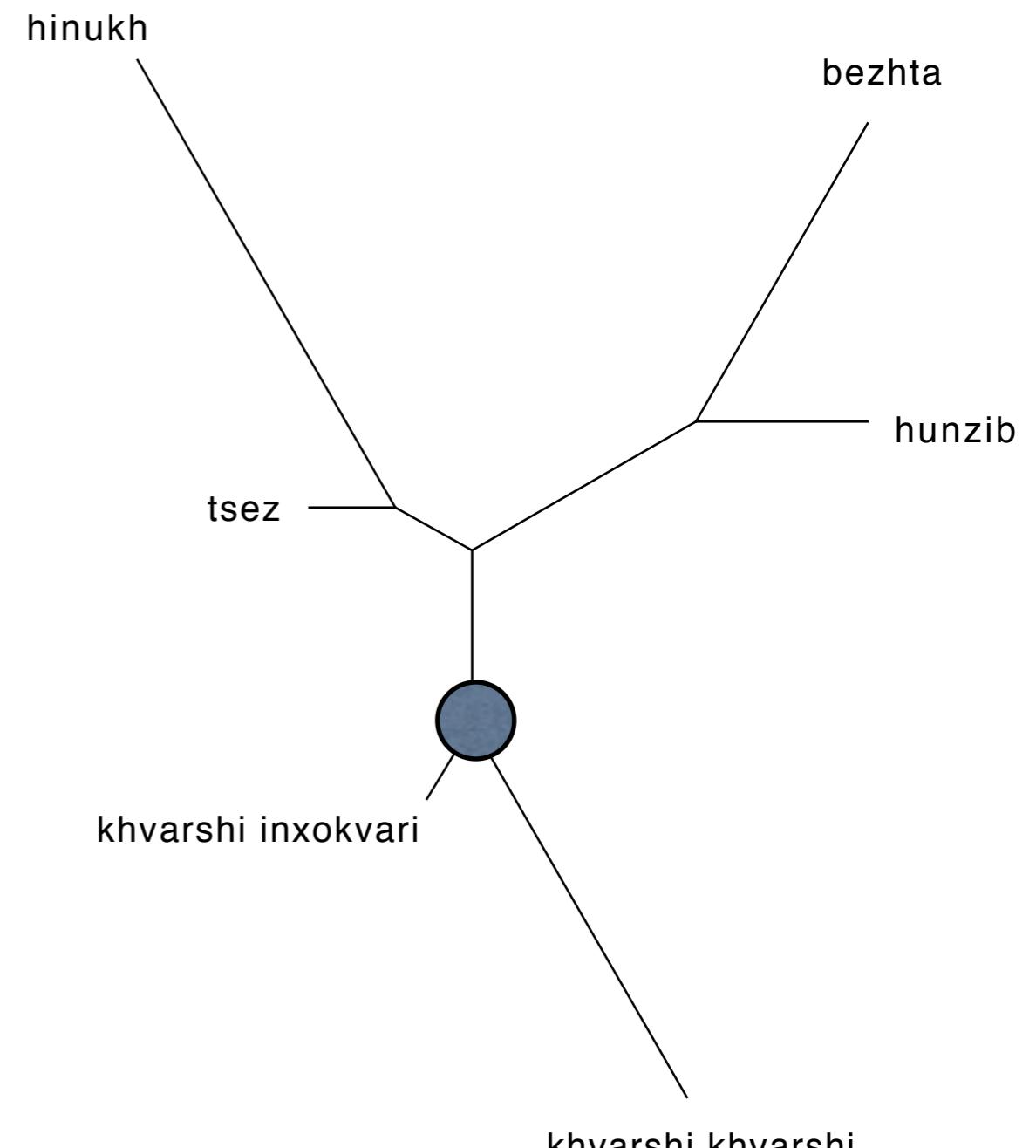
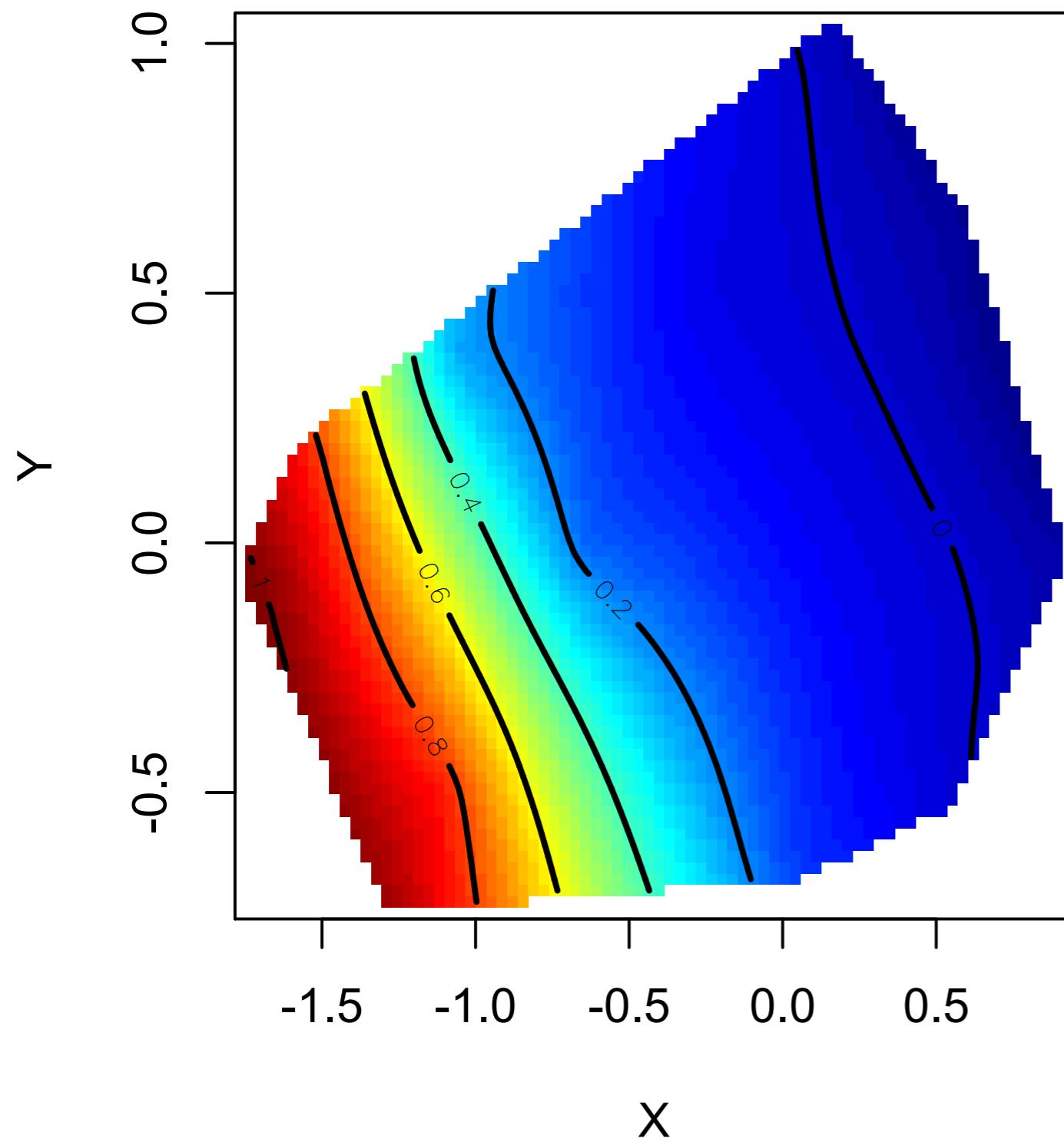


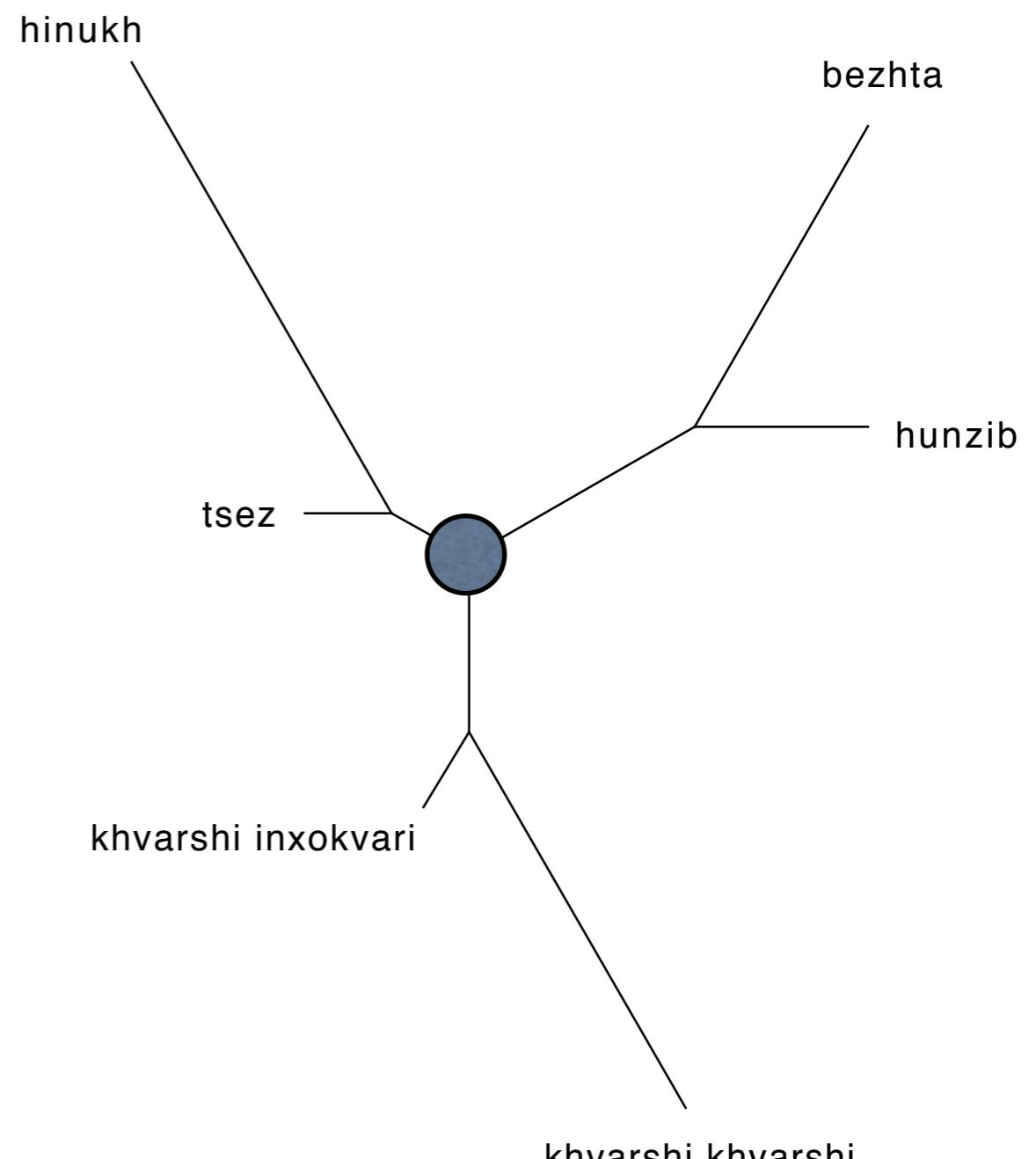
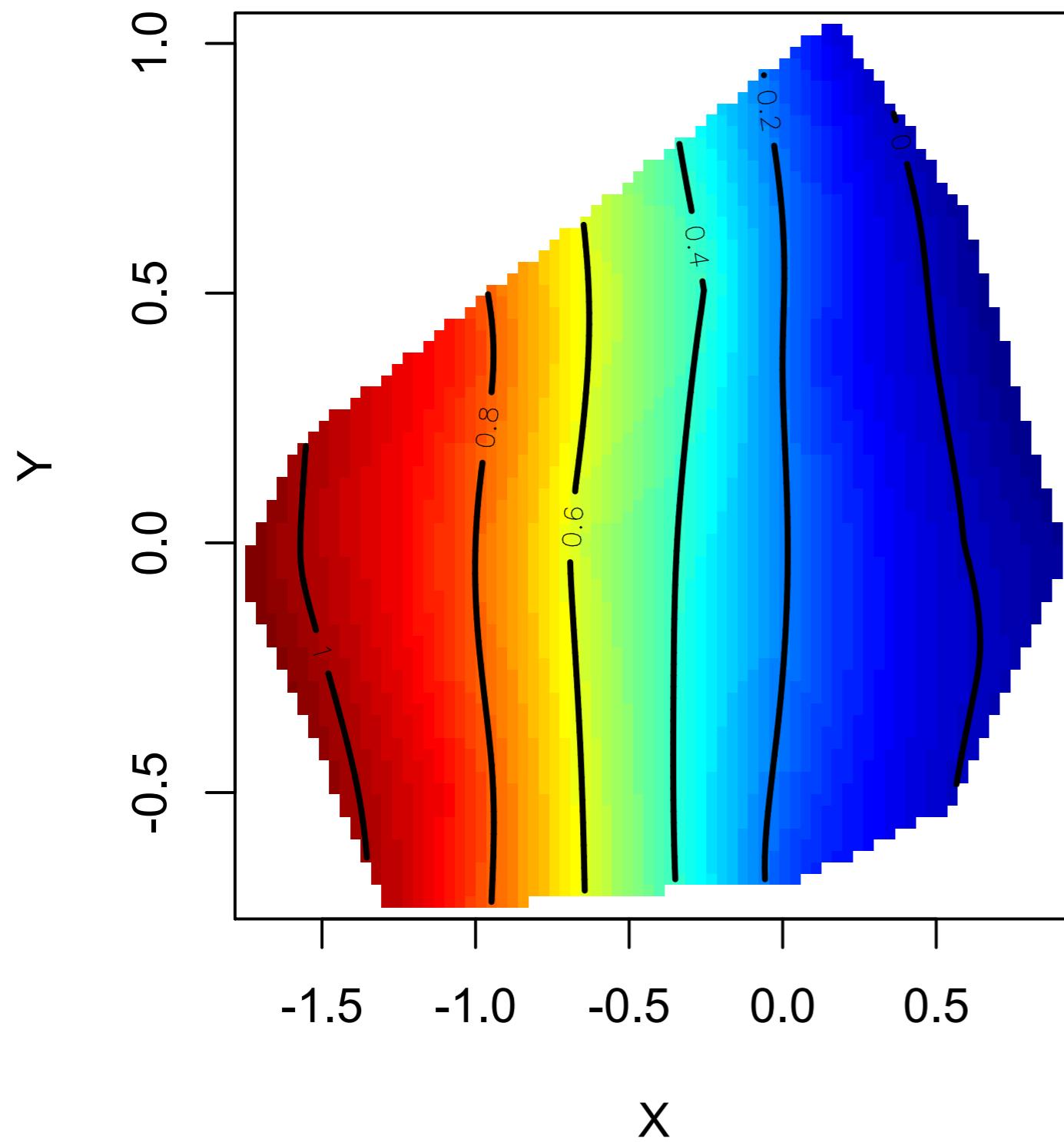


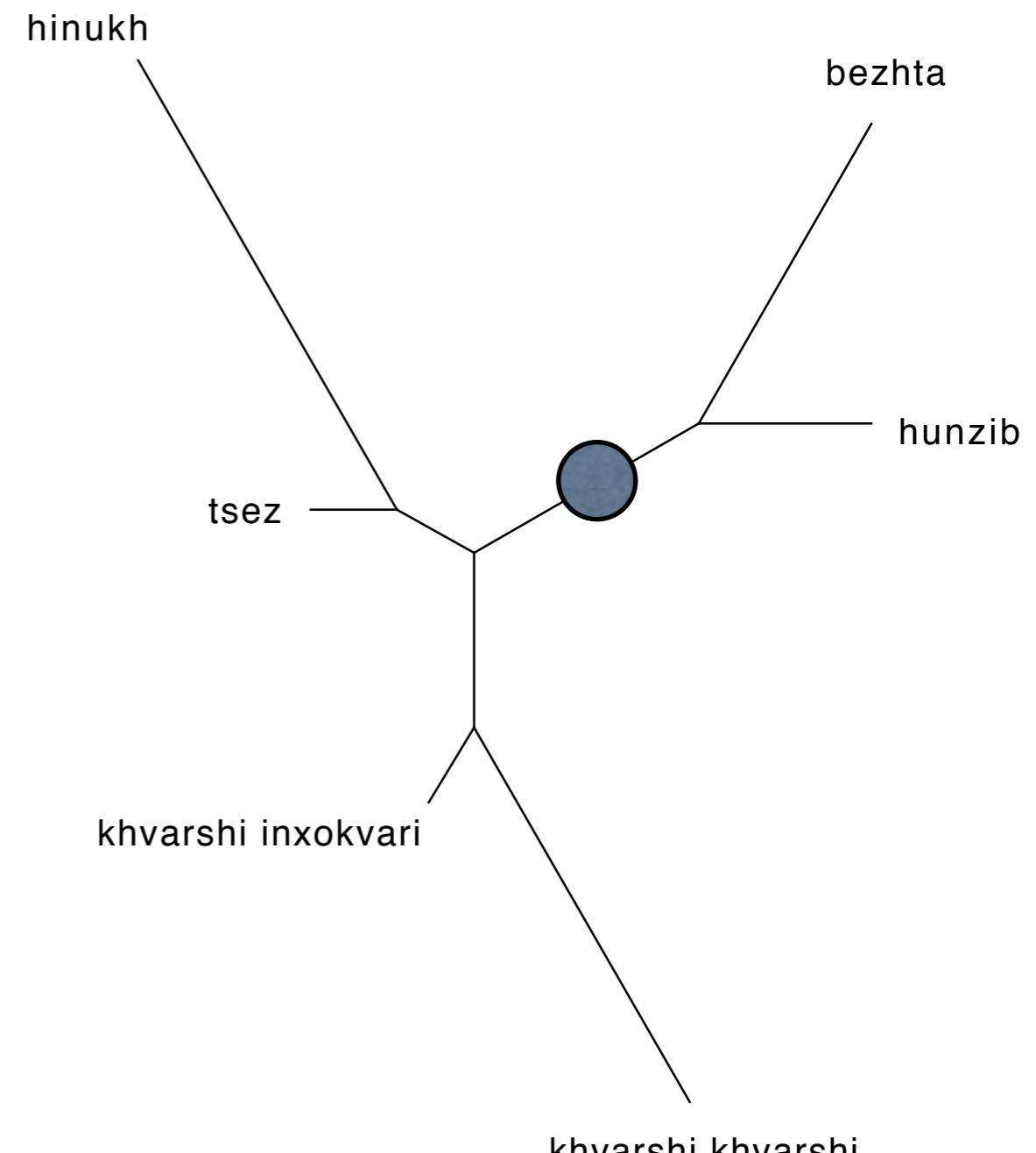
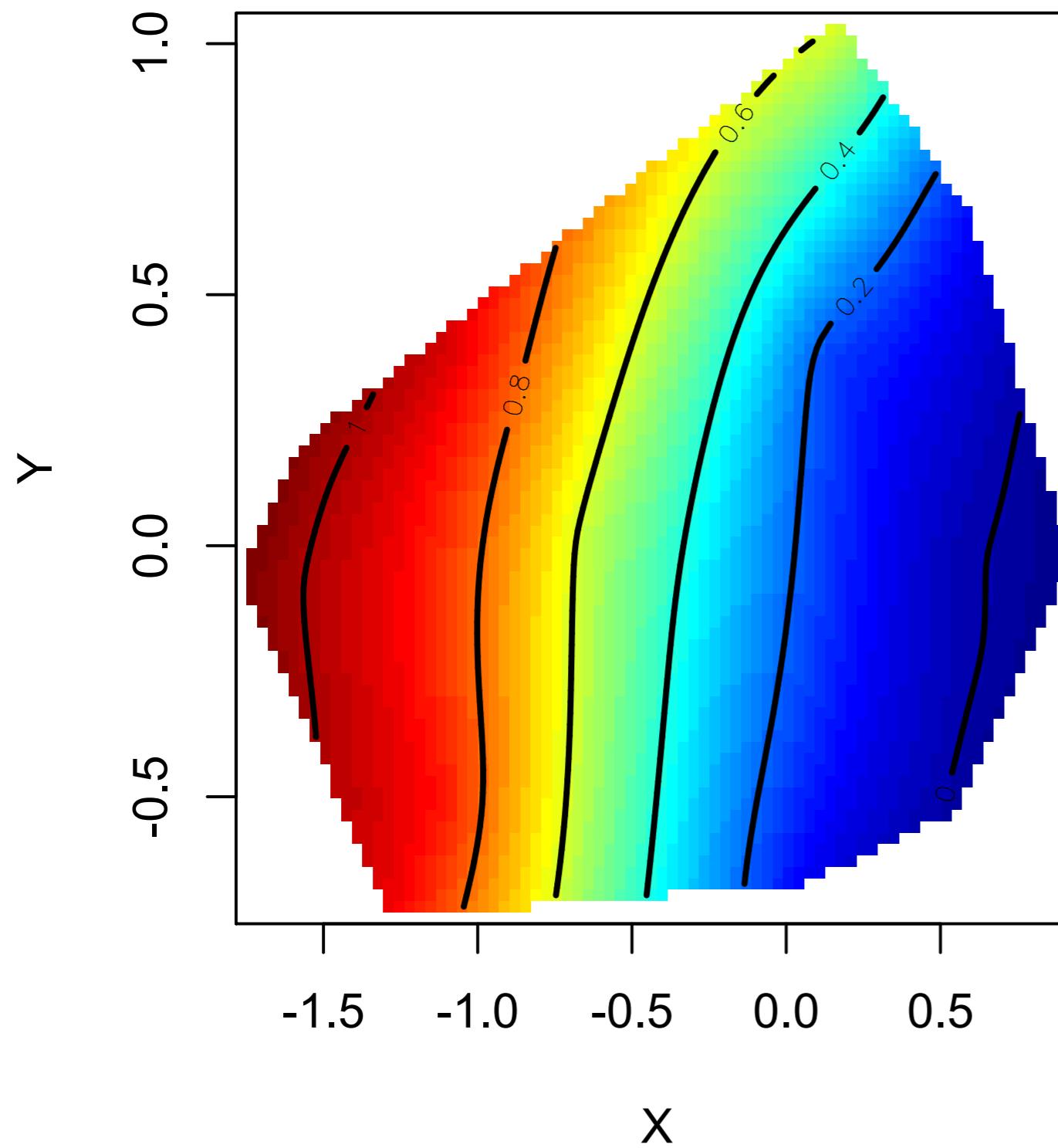


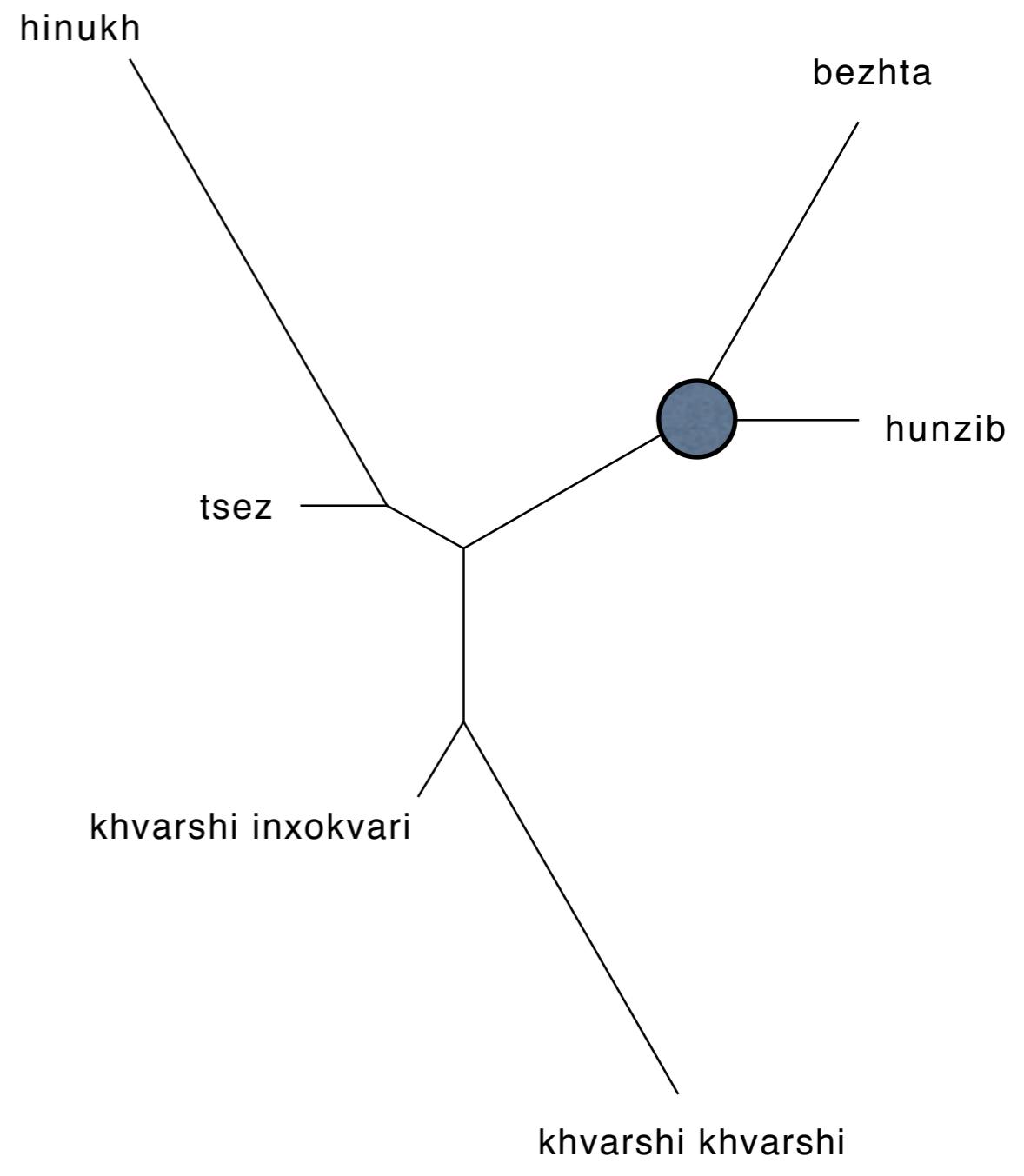
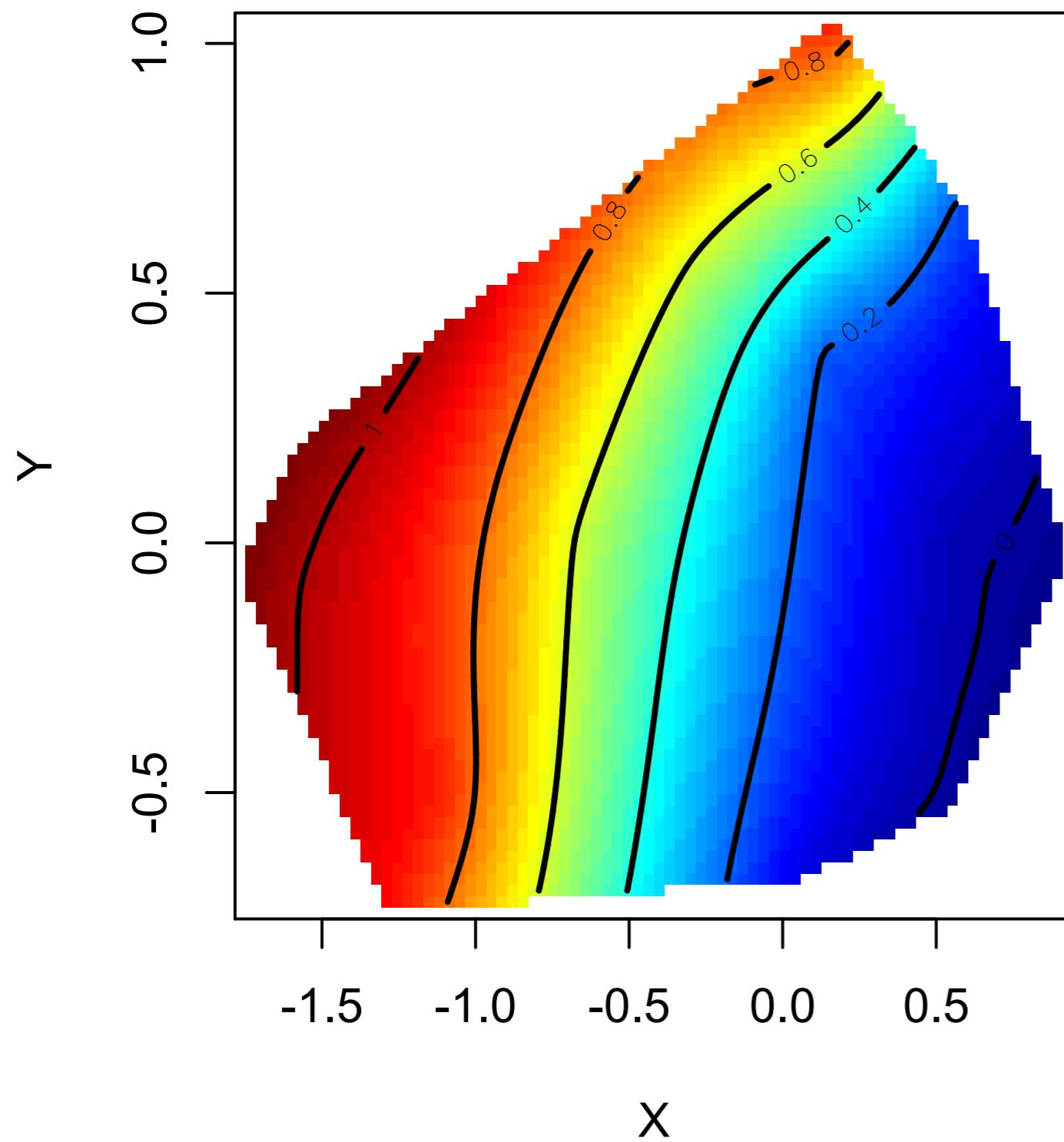


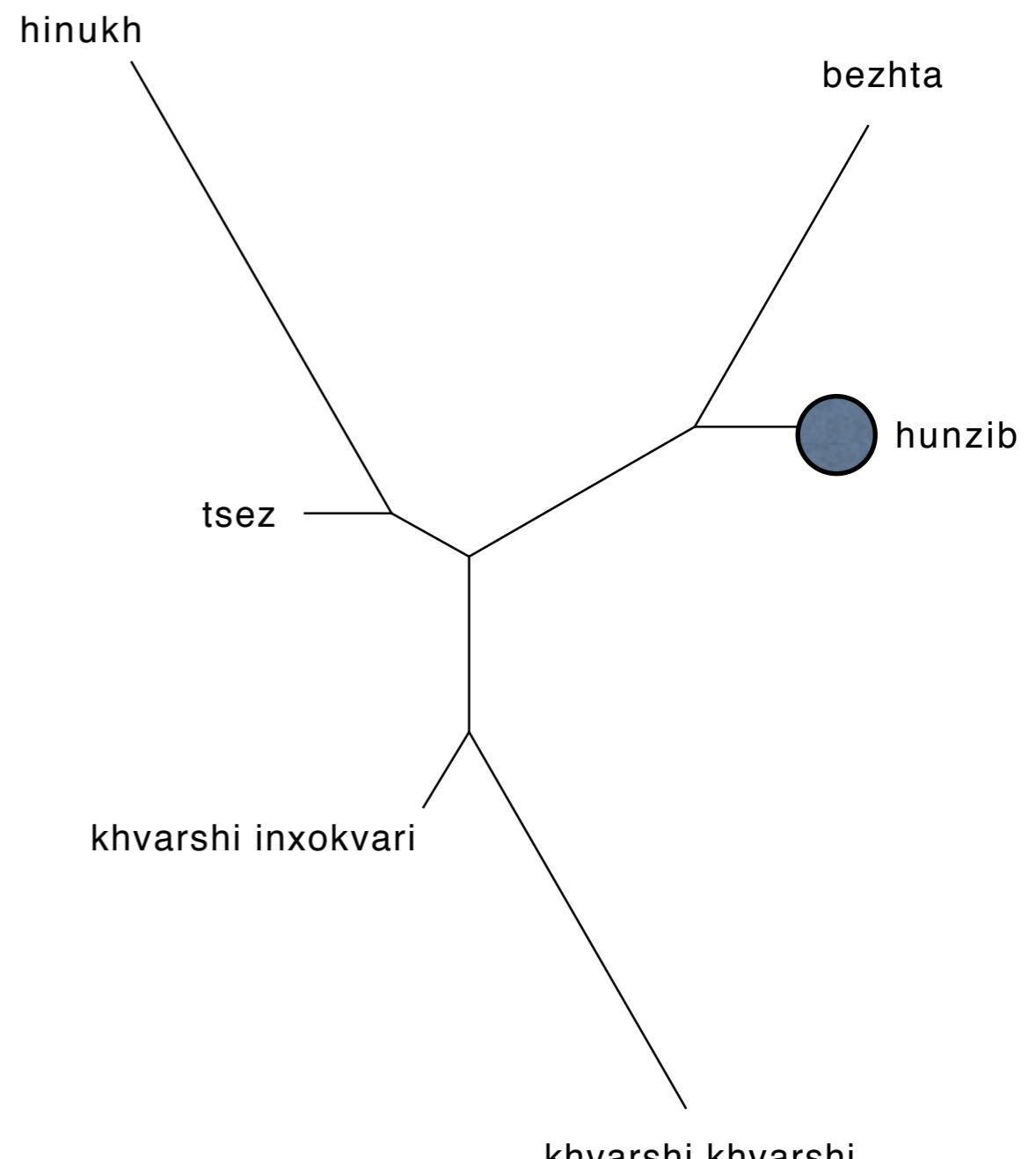
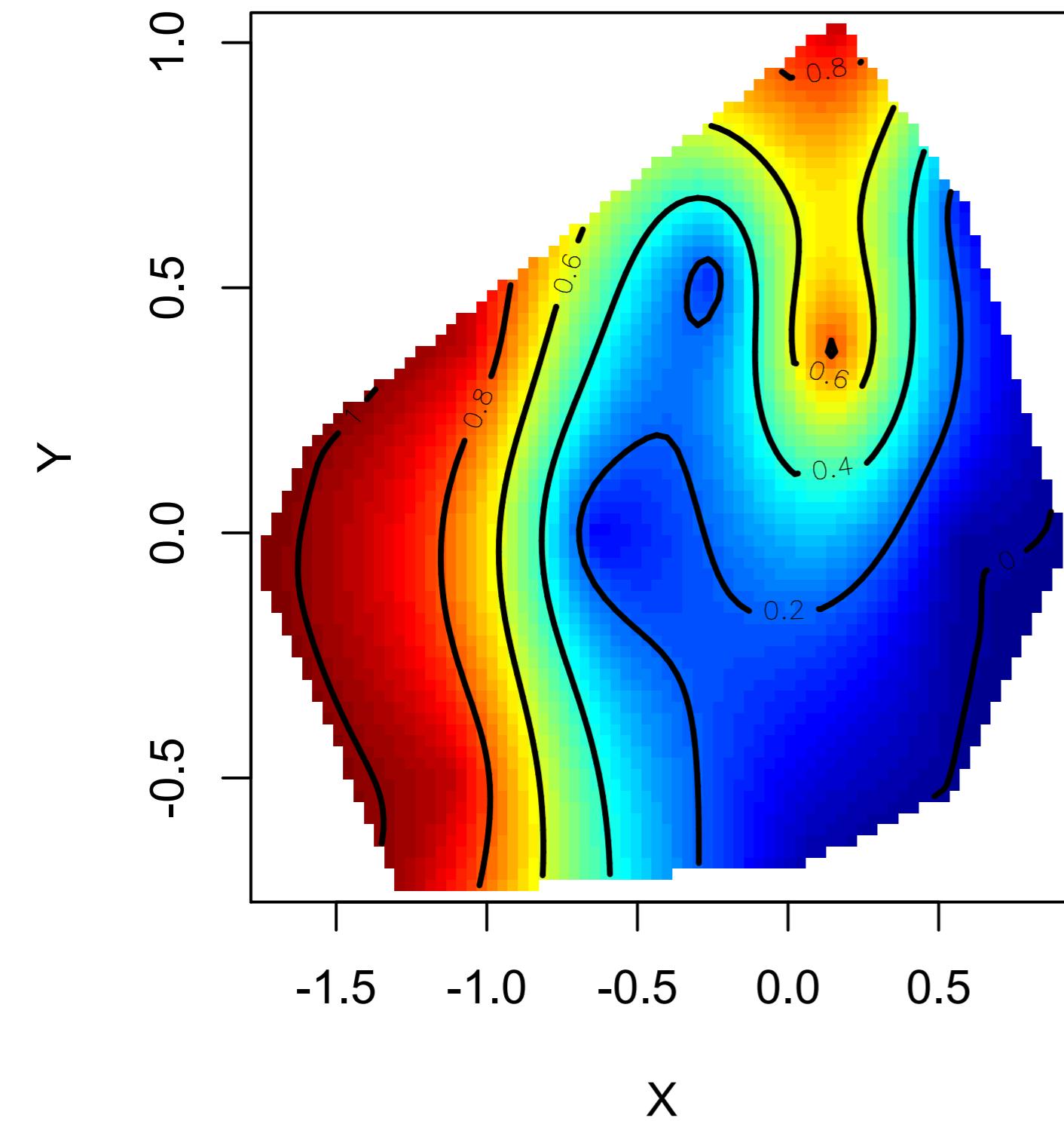












# **Universal Declaration of Human Rights**

<i>in, im</i>	German
<i>in</i>	Dutch
<i>in</i>	English
<i>yn</i>	Frisian
<i>in</i>	Afrikaans
<i>in</i>	Scots
אין	Yiddish
í	Icelandic
<i>i</i>	Swedish
<i>i</i>	Nynorsk
<i>i</i>	Bokmål
í	Faroese
<i>i</i>	Danish

# Article 10

**English:** “everyone is entitled in full equality to a fair and public hearing by an independent and impartial tribunal , in the determination of his rights and obligations and of any criminal charge against him .”

**Afrikaans:** “elkeen het , in volle gelykheid , die reg tot ’n regverdige en openbare verhoor deur ’n onafhanklike en objektiewe tribunaal , in die bepaling van sy regte en verpligte en die ondersoek van enige kriminele saak teen hom .”

**German:** “jeder hat bei der feststellung seiner rechte und pflichten sowie bei einer gegen ihn erhobenen strafrechtlichen beschuldigung in voller gleichheit anspruch auf ein gerechtes und öffentliches verfahren vor einem unabhängigen und unparteiischen gericht .”

**Danish:** “enhver har under fuld ligeberettigelse krav på en retfærdig og offentlig behandling ved en uafhængig og upartisk domstol , når der skal træffes en afgørelse med hensyn til hans rettigheder og forpligtelser og med hensyn til en hvilken som helst mod ham rettet strafferetslig anklage .”

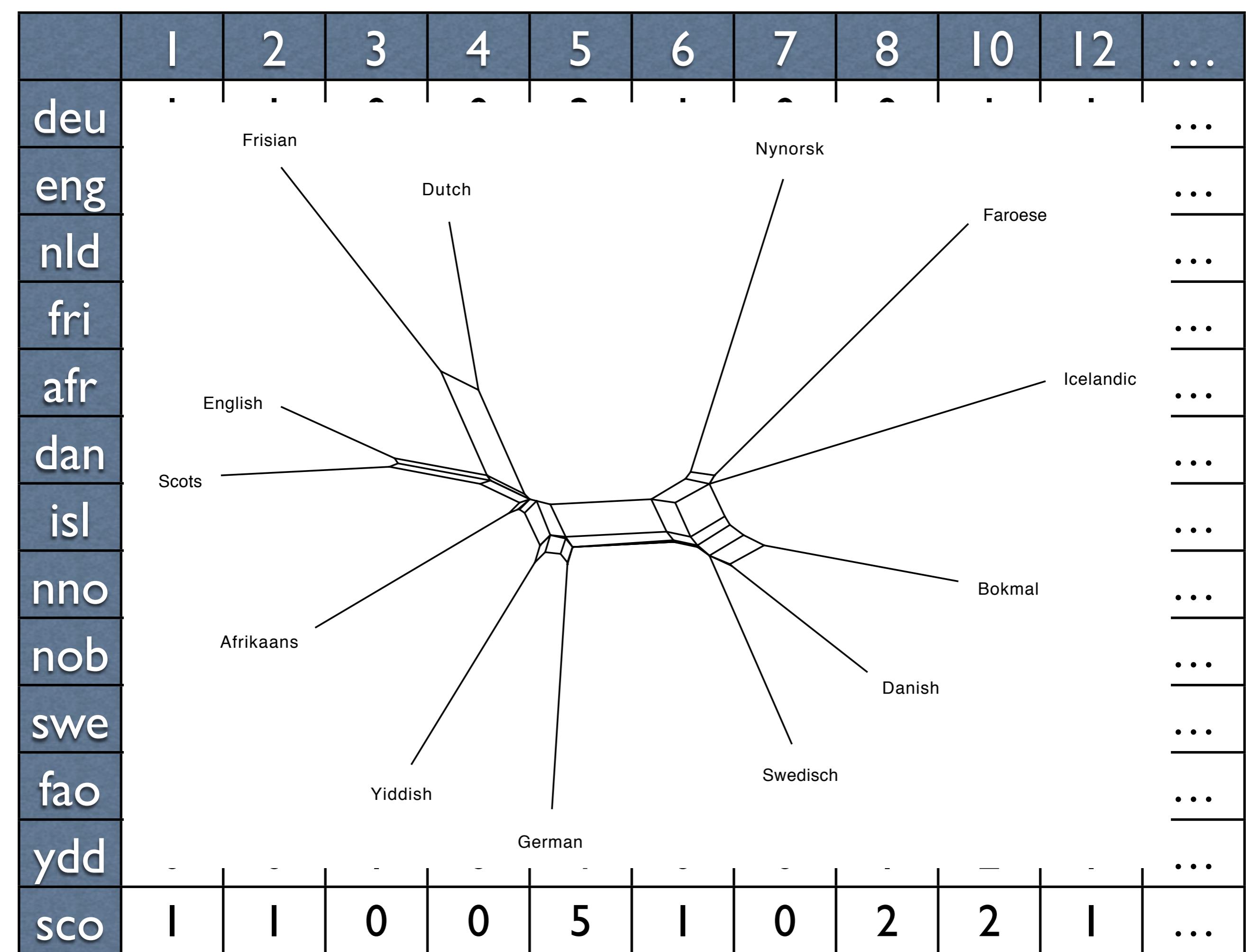
# Article 10

**English:** “everyone is entitled **in** full equality to a fair and public hearing by an independent and impartial tribunal , **in** the determination of his rights and obligations and of any criminal charge against him .”

**Afrikaans:** “elkeen het , **in** volle gelykheid , die reg tot 'n regverdige en openbare verhoor deur 'n onafhanklike en objektiewe tribunaal , **in** die bepaling van sy regte en verpligte en die ondersoek van enige kriminele saak teen hom .”

**German:** “jeder hat **bei** der feststellung seiner rechte und pflichten sowie bei einer gegen ihn erhobenen strafrechtlichen beschuldigung **in** voller gleichheit anspruch auf ein gerechtes und öffentliches verfahren vor einem unabhängigen und unparteiischen gericht .”

**Danish:** “enhver har **under** fuld ligeberettigelse krav på en retfærdig og offentlig behandling ved en uafhængig og upartisk domstol , **når** der skal træffes en afgørelse med hensyn til hans rettigheder og forpligtelser og med hensyn til en hvilken som helst mod ham rettet strafferetslig anklage .”



# Article 10

**English:** “everyone is entitled **in full equality** to a fair and public hearing by an independent and impartial tribunal , in the determination of his rights and obligations and of any criminal charge against him .”

**Afrikaans:** “elkeen het , **in volle gelykheid** , die reg tot 'n regverdige en openbare verhoor deur 'n onafhanklike en objektiewe tribunaal , in die bepaling van sy regte en verpligte en die ondersoek van enige kriminele saak teen hom .”

**German:** “jeder hat bei der Feststellung seiner Rechte und Pflichten sowie bei einer gegen ihn erhobenen strafrechtlichen Beschuldigung **in voller Gleichheit** Anspruch auf ein gerechtes und öffentliches Verfahren vor einem unabhängigen und unparteiischen Gericht .”

**Danish:** “enhver har **under fuld ligeberettigelse** krav på en retfærdig og offentlig behandling ved en uafhængig og upartisk domstol , når der skal træffes en afgørelse med hensyn til hans rettigheder og forpligtelser og med hensyn til en hvilken som helst mod ham rettet strafferetslig anklage .”

# Article 10

**English:** “everyone is entitled in full equality to a fair and public hearing by an independent and impartial tribunal , **in the determination of his rights** and obligations and of any criminal charge against him .”

**Afrikaans:** “elkeen het , in volle gelykheid , die reg tot ’n regverdige en openbare verhoor deur ’n onafhanklike en objektiewe tribunaal , **in die bepaling van sy regte** en verpligte en die ondersoek van enige kriminele saak teen hom .”

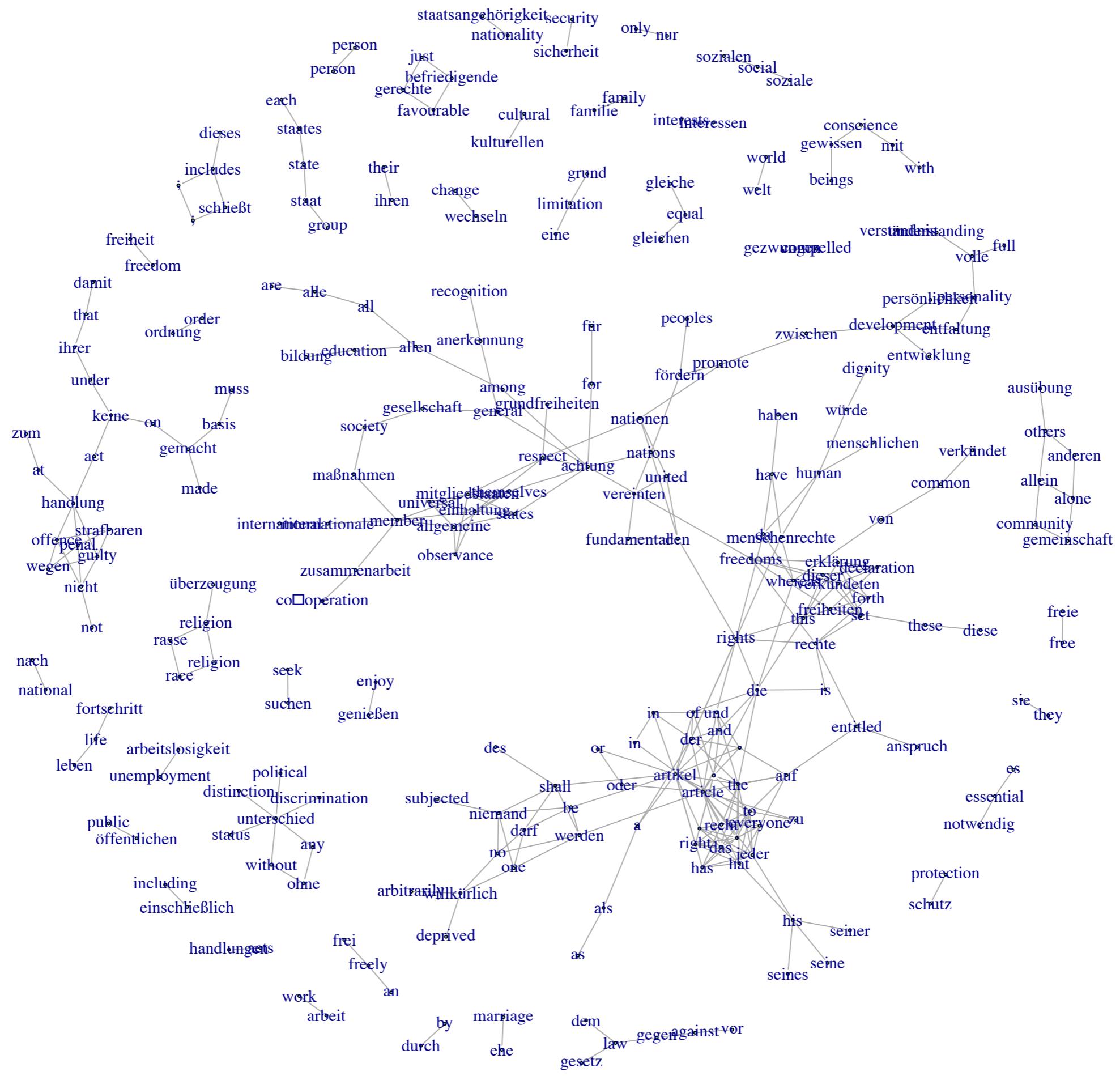
**German:** “jeder hat **bei der Feststellung seiner Rechte** und Pflichten sowie bei einer gegen ihn erhobenen strafrechtlichen Beschuldigung in voller Gleichheit Anspruch auf ein gerechtes und öffentliches Verfahren vor einem unabhängigen und unparteiischen Gericht .”

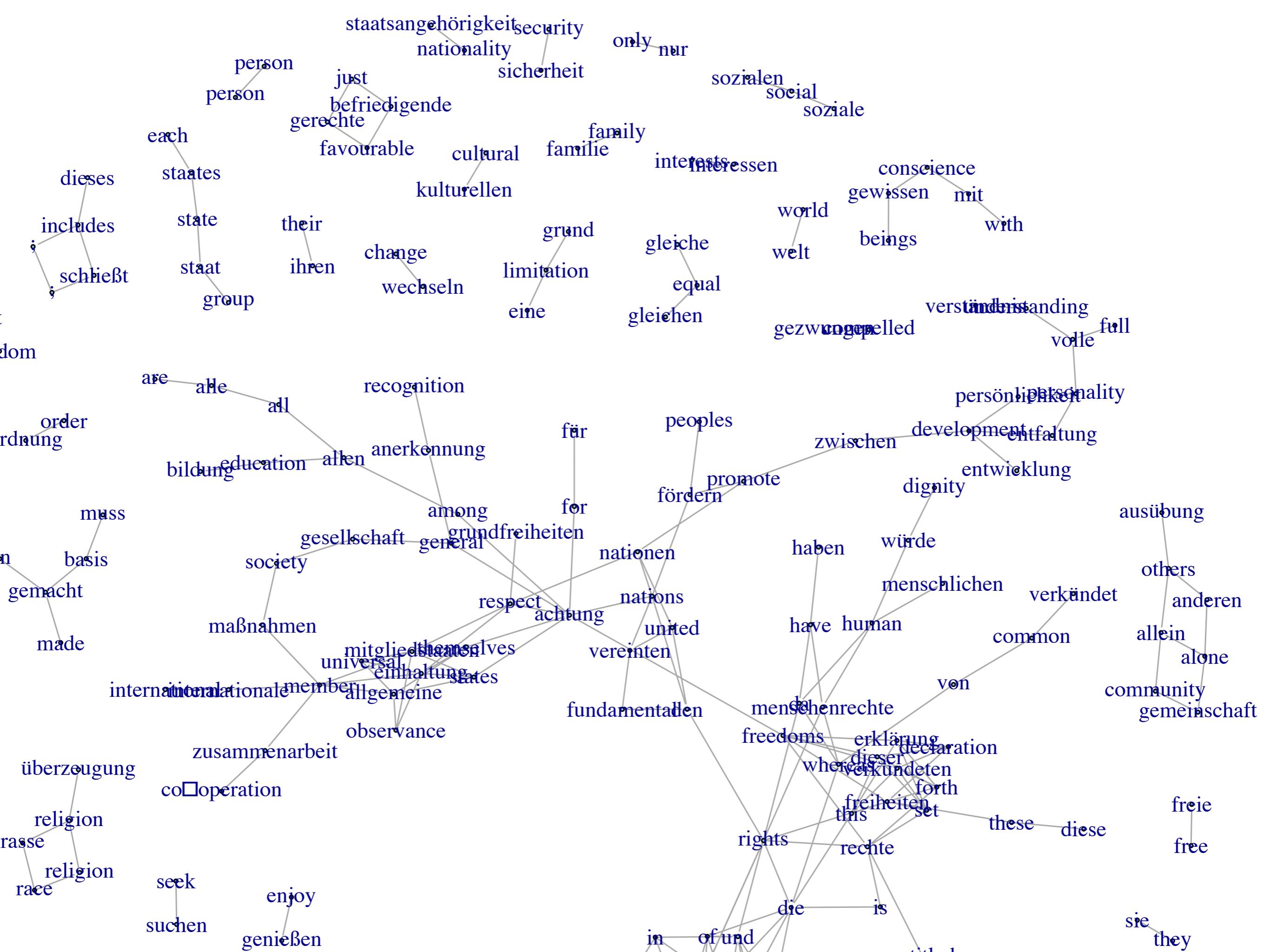
**Danish:** “enhver har under fuld ligeberettigelse krav på en retfærdig og offentlig behandling ved en uafhængig og upartisk domstol , **når der skal træffes en afgørelse med hensyn til hans rettigheder** og forpligtelser og med hensyn til en hvilken som helst mod ham rettet strafferetslig anklage .”

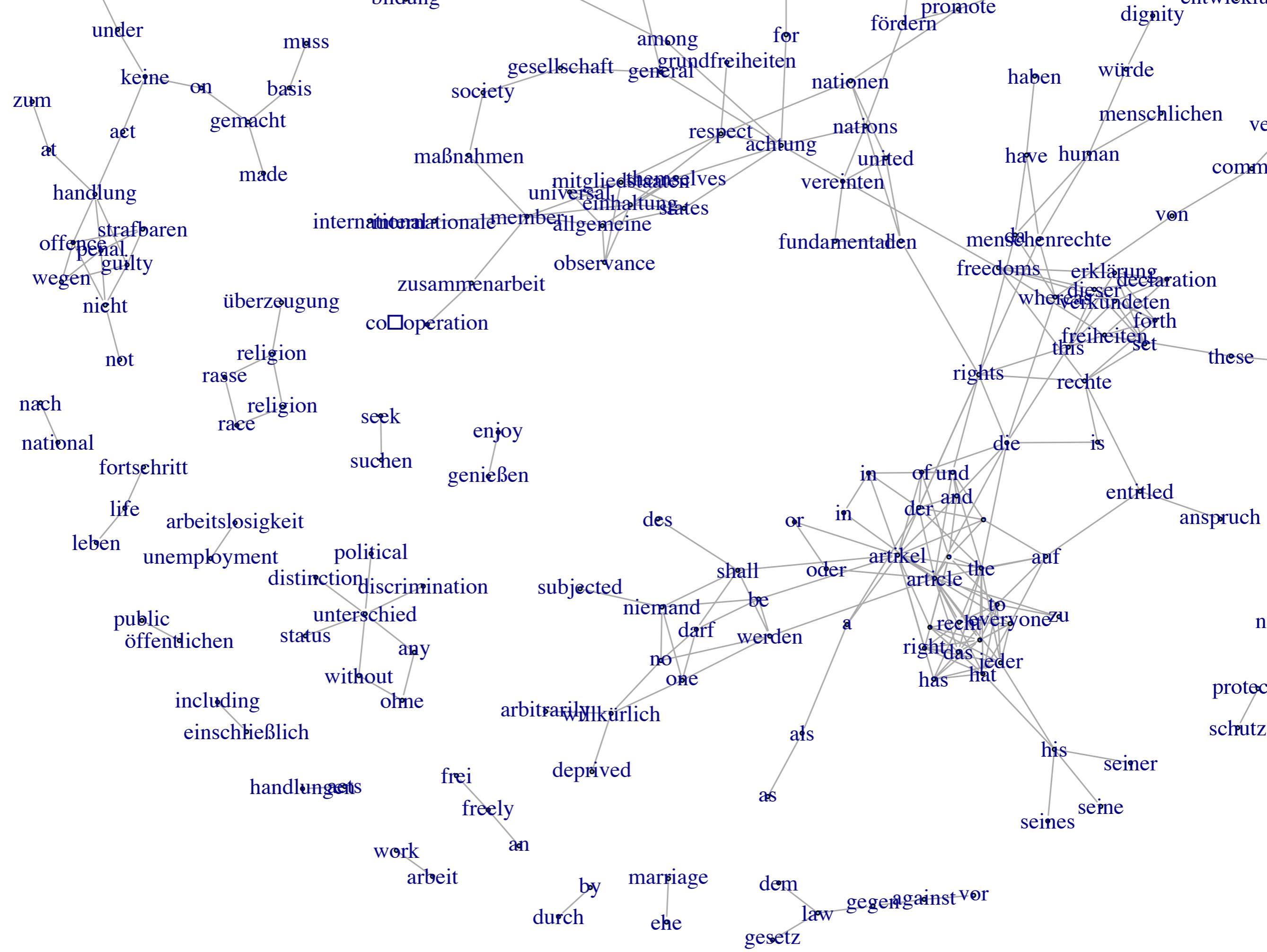
- For each word from each language, compute how often it co-occurs with each word in each other language
- Only for the 58 paragraphs in the declaration
- significance of co-occurrence of word  $w$  and word  $v$  based on Poisson distribution:

$$sig(w, v) = \lambda^k \cdot e^{-\lambda} \cdot k!$$

$\lambda = (\text{frequency of word } w) \cdot (\text{frequency of word } v) \div (\text{number of contexts})$   
 $k = \text{frequency of co-occurrence of words } w \text{ and } v$







# Preamble, sentence 5

## English:

“whereas the peoples of the united nations have **in** the charter reaffirmed their faith **in** fundamental human rights , **in** the dignity and worth of the human person and **in** the equal rights of men and women and have determined to promote social progress and better standards of life **in** larger freedom ,”

## German:

“da die völker der vereinten nationen **in** der charta ihren glauben an die grundlegenden menschenrechte , an die würde und den wert der menschlichen person und an die gleichberechtigung von mann und frau erneut bekräftigt und beschlossen haben , den sozialen fortschritt und bessere lebensbedingungen **in** größerer freiheit zu fördern ,”

# Preamble, sentence 5

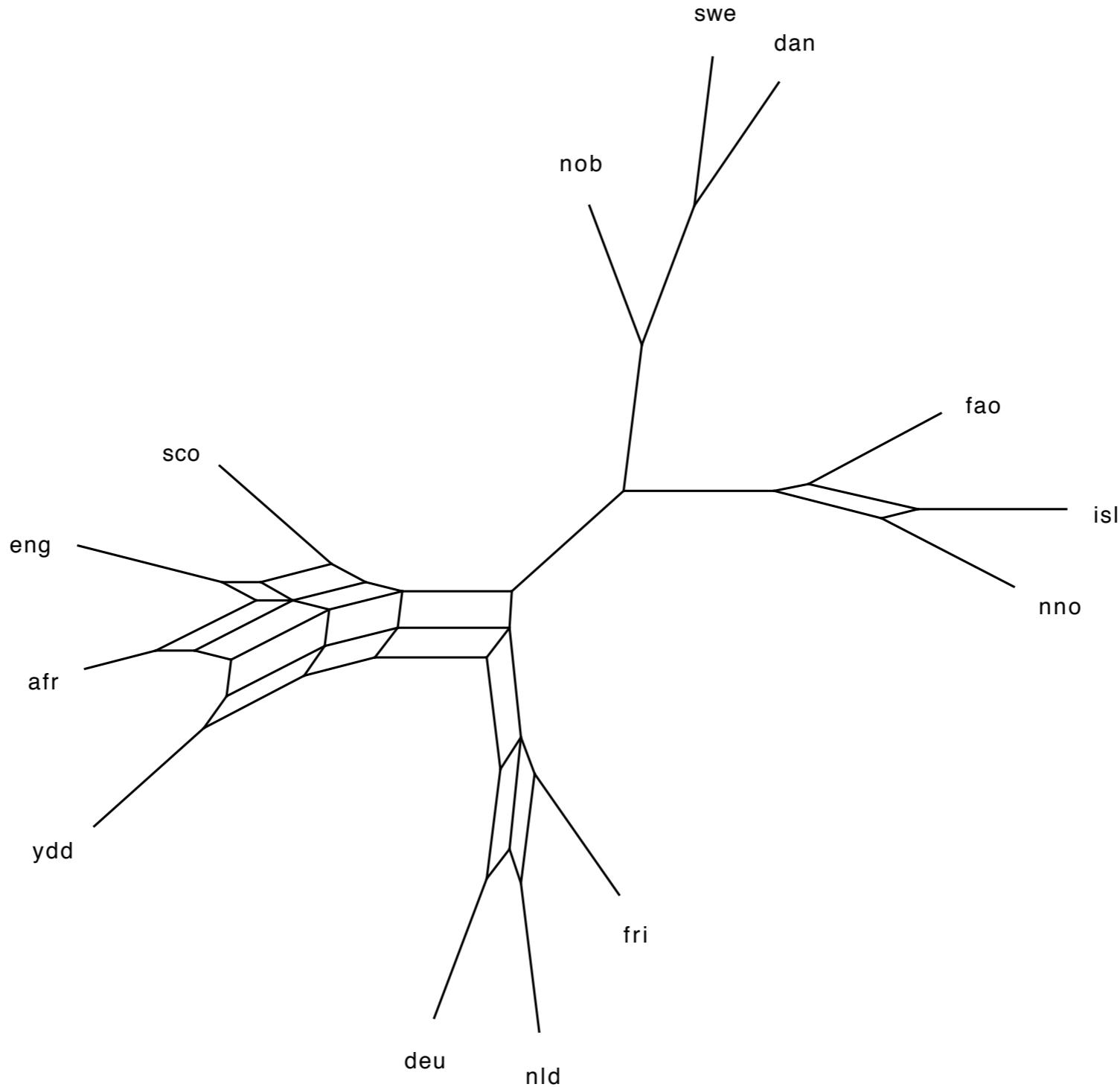
## English:

“whereas the peoples of the **united nations** have **in** the **charter** reaffirmed their faith **in** fundamental human rights , **in** the dignity and worth of the human person and **in** the equal rights of men and women and have determined to promote social progress and better standards of life **in larger freedom** ,”

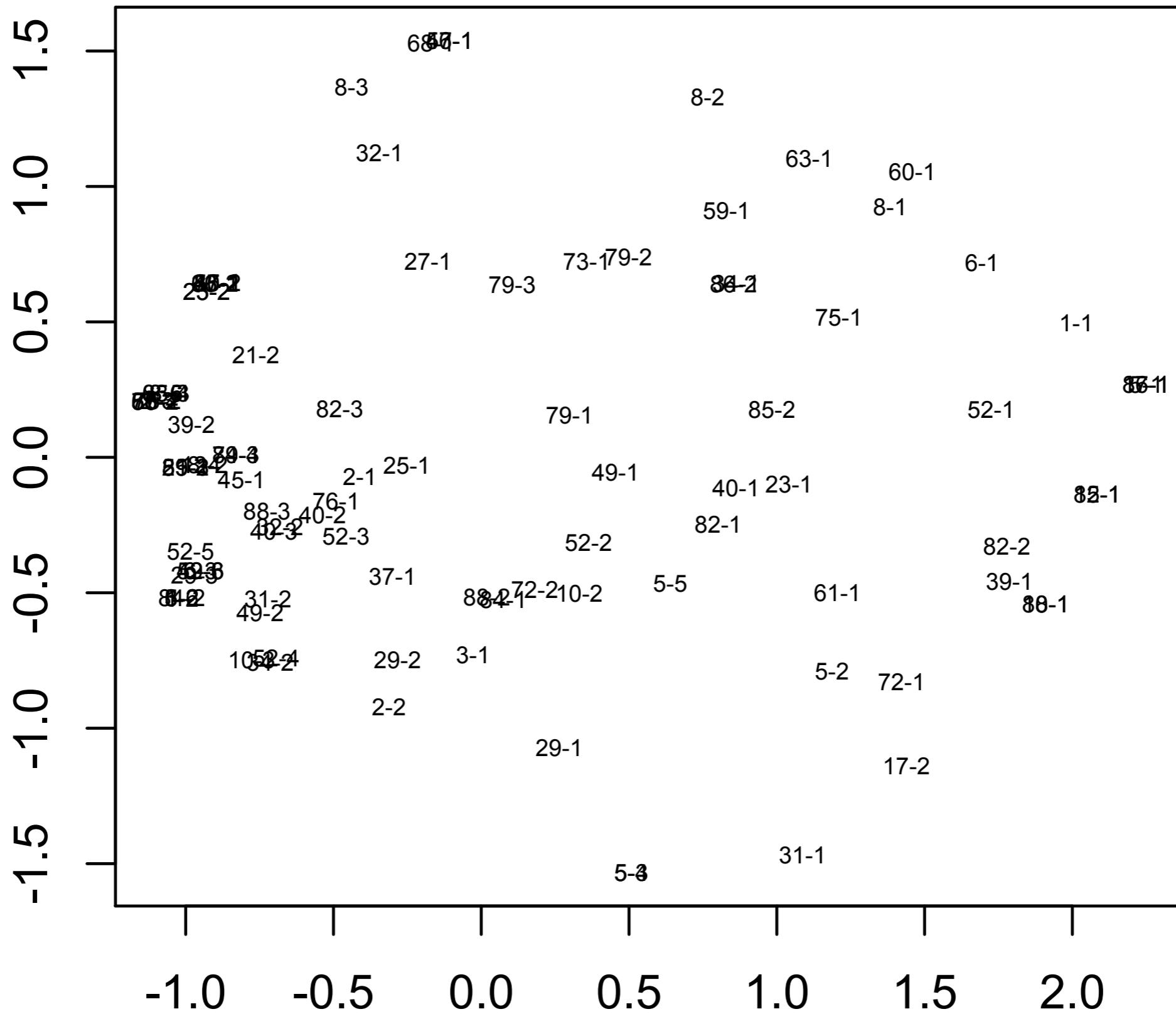
## German:

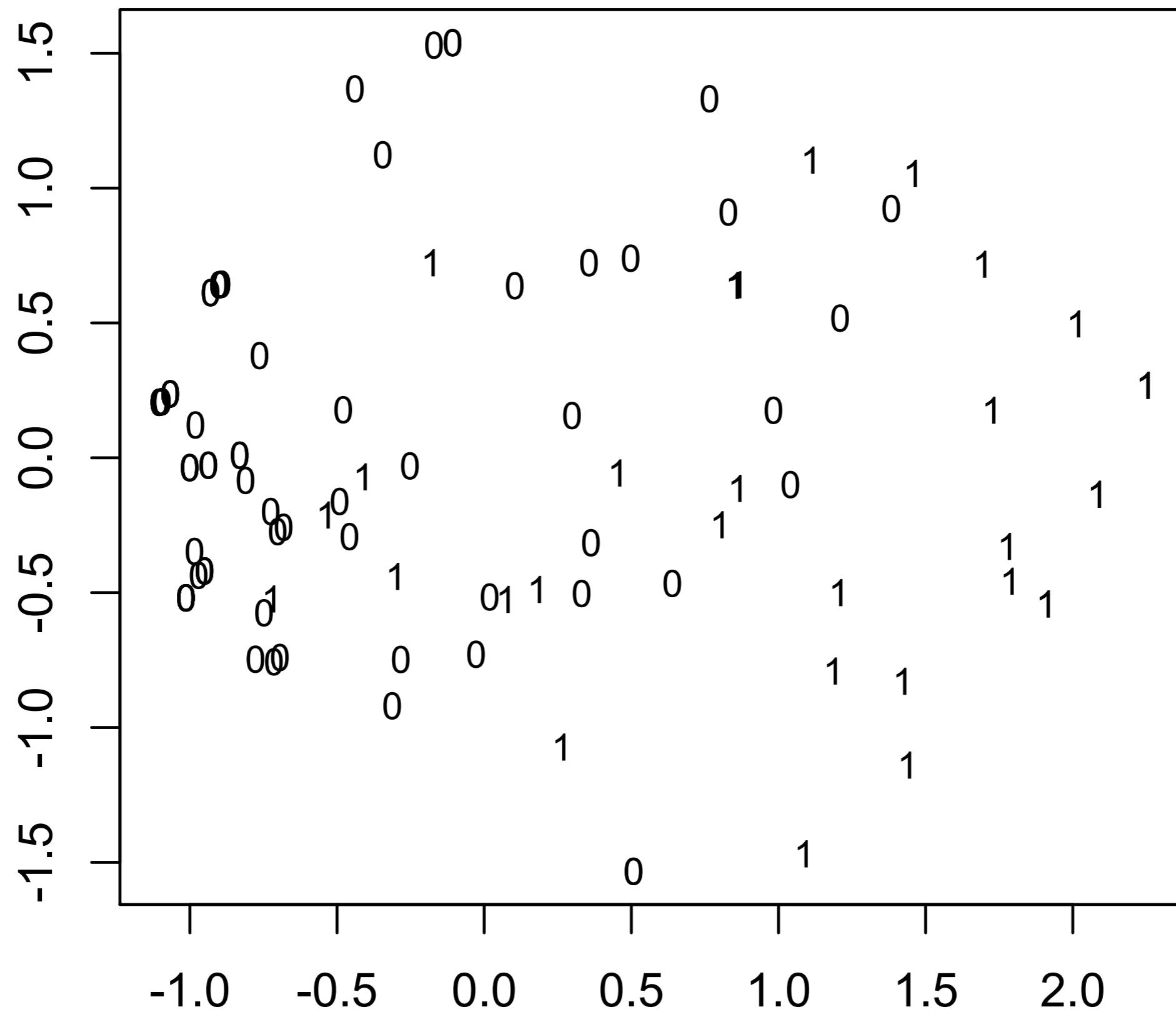
“da die völker der **vereinten nationen** **in** der **charta** ihren glauben an die grundlegenden menschenrechte , an die würde und den wert der menschlichen person und an die gleichberechtigung von mann und frau erneut bekräftigt und beschlossen haben , den sozialen fortschritt und bessere lebensbedingungen **in größerer freiheit** zu fördern ,”

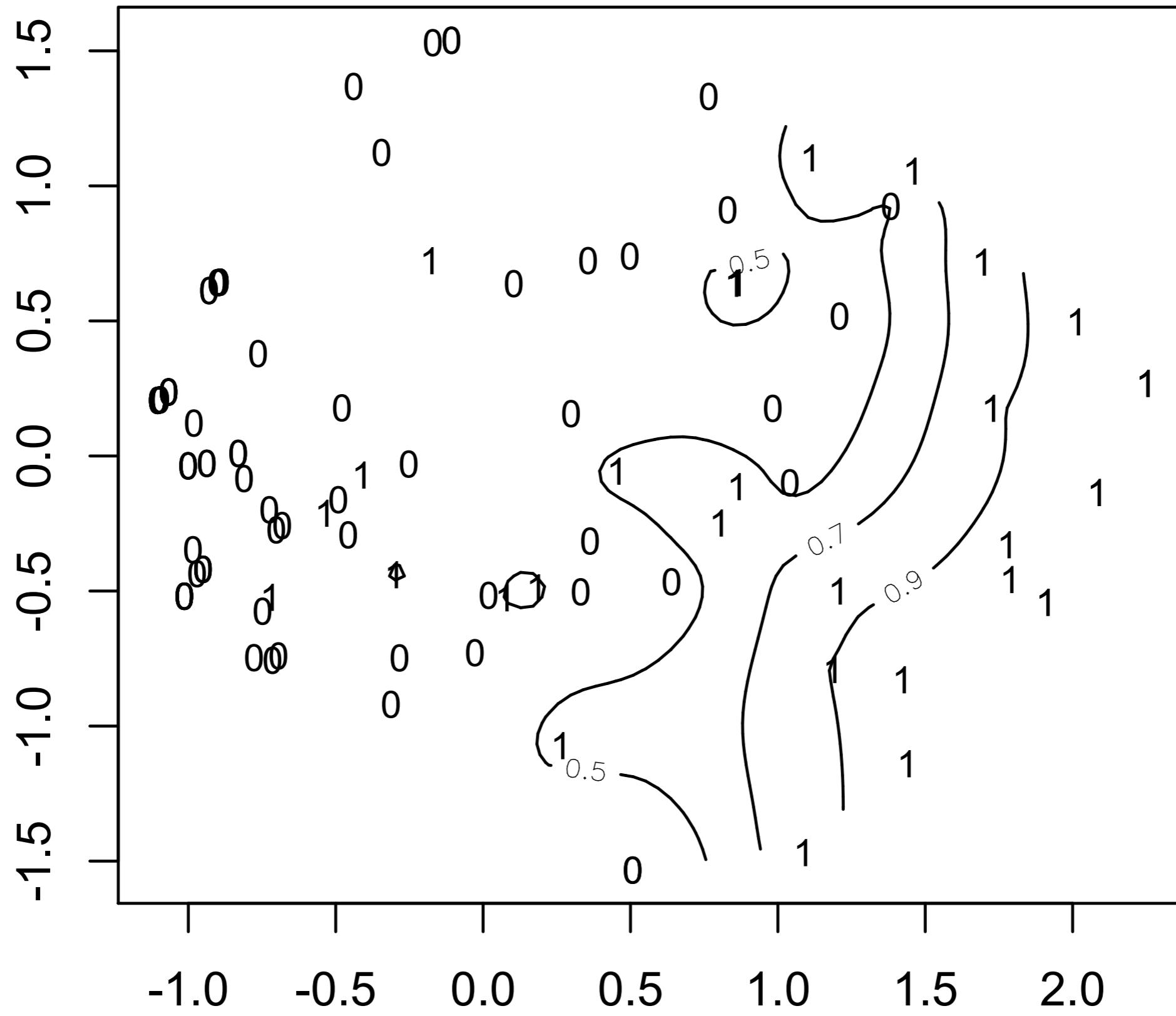
deu	2							
eng	5							
nld	5							
fri	6							
afr	5							
ydd	4							
sco	5							
nno	3							
nob								
swe								
fao	3							
dan								
isl								

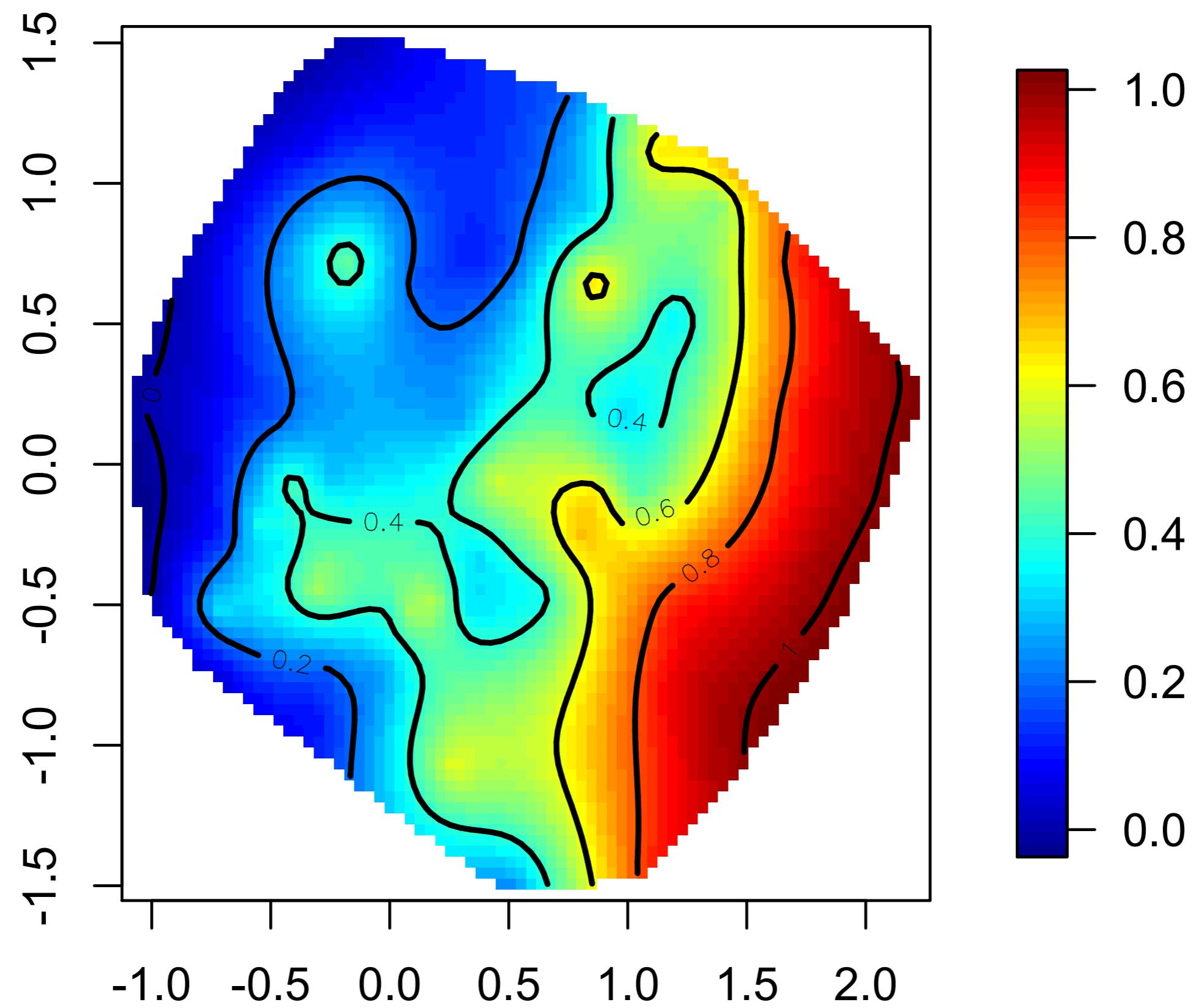


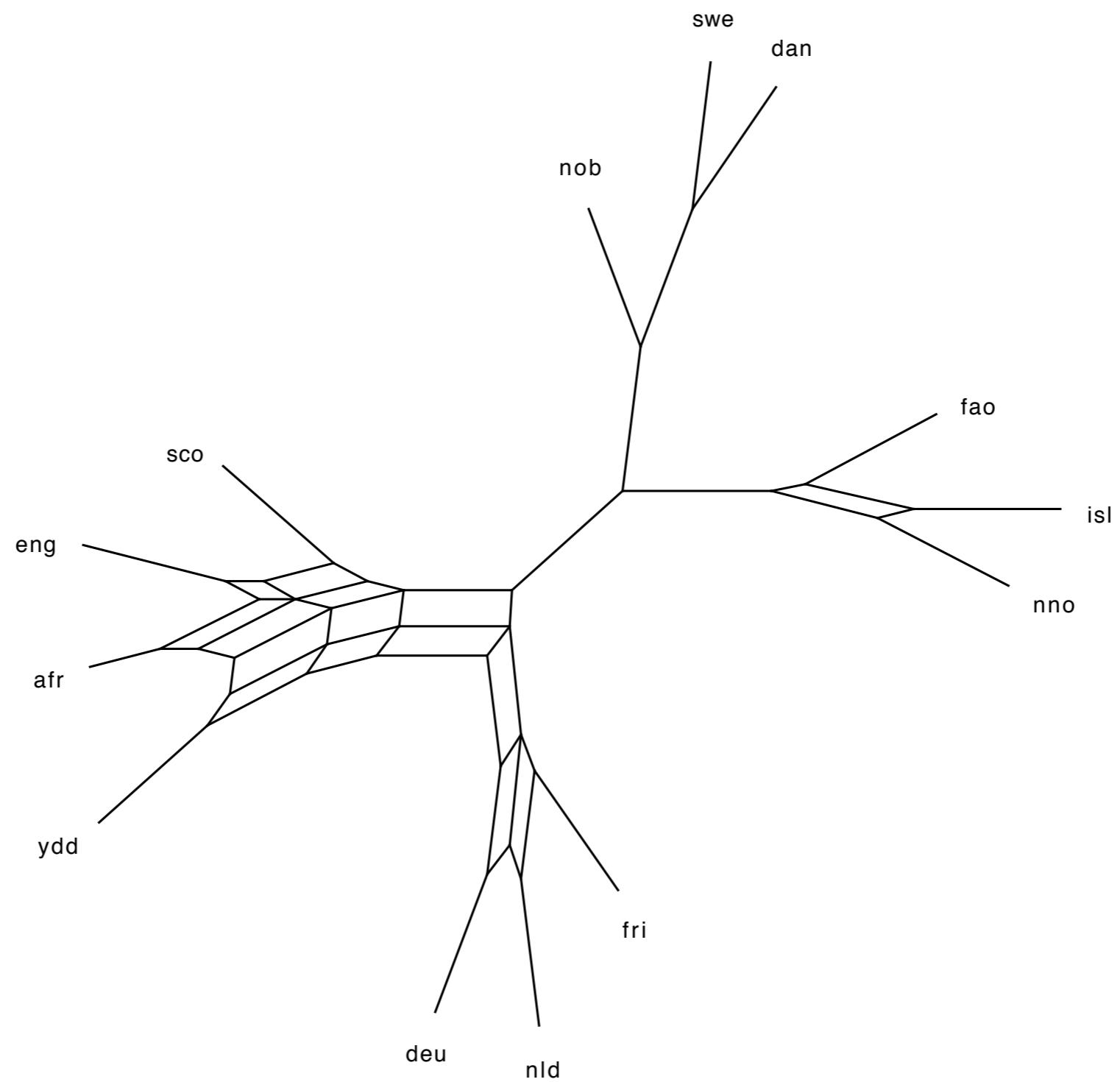
**Consensus network of 4 optimal trees  
according to ‘dollo’ Maximum Parsimony**

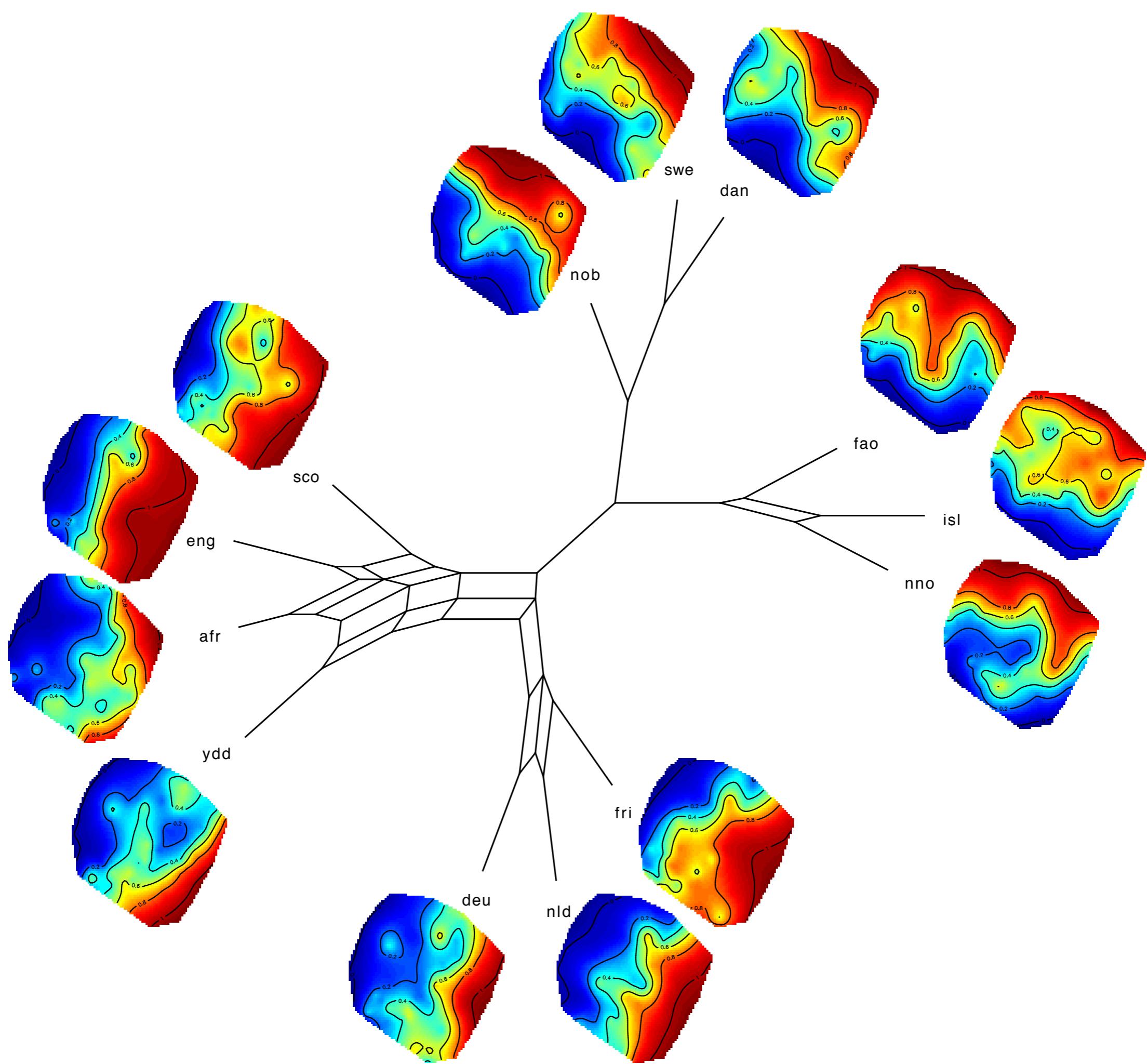












# Conclusion

- Traditional comparative-historical linguistics is a great method, but it only deals with one aspect: changes in the phoneme system
- Changes in meaning are normally only considered a nuisance
- But: using details of difference in meaning, more independent evidence for historical relationships can be uncovered
- Text-mining statistics are incredibly helpful