

Using Massively Parallel Text for Language Typology

Michael Cysouw

Deutscher Sprachatlas, Philipps Universität Marburg

Quantitative Language Comparison

Dictionaries

aatyétsíí *pron.dem.* aquellos dos (hombres o animales machos más distantes).

átyoj *voc.* cuñado. (ver **tónuúbe**)

átyónuúbe, átyónujte *s.* mi cuñado. (ver **tónuúbe**)

átyónulle *s.* mi cuñada. (ver **tónulle**) (sinón. **búhjíí**)

áátyumi *s.col.* especie de pajarito muy lindo cuyas plumas de color celeste se usaban para hacer adornos.

áátyúmiího un pajarito.

átyúvájkeu *s.* especie de bejuco que se usa para amarrar manguarés y balsas.

átsaacúnu *abs.* efecto de...

[átsáácunu] *vi.* malograrse la carne. (sinón. **mévaacúnu**)

aatsári *s.col.* especie de roedor nocturno parecido al ratón.

aatsáriu un roedor.

átsii *conj.pron.* ese lugar (mencionado en la oración anterior).

átsihdyu de ese lugar.

átsihvu a ese lugar.

átsímujtéba *s.* inundación grande, crecienta grande.

átsímujtóhi *s.* restinga.

áátsíime *s.* mis hijos. (ver **tsíime**)

aau *conj.pron.* ése (bola, pelota, etc. mencionado en la oración anterior).

ava (ver **áa**)

avyába *s.pos.* cogollo.

coomíí **avyába** cogollo de ungurahui.

avye, ávyécunu *abs.* dolencia, dolor. ¡Áyúú, íñéhaca ihdyu avye! ¡Ay, qué dolor! || estado de...

[avye, ávyécunu] *ve.* 1. doler; tener pena, sentir pena. *Mítyane táhbááu oke avyécunúhi.* Me duele mucho el

estómago. 2. ser picante.

[ávyé(h)/avyéwu] *adj.* 1. doloroso. *Oke naníyó wátsihcyú avyéwu.* Los azotes que me dio mi tío fueron muy dolorosos. 2. picante.

ávyejpácyo *s.* licor; aguardiente; cualquier bebida alcohólica.

ávyejpiítte *abs.* efecto de...

[ávyejpiítte] *ve.* estar dolorido, tener dolor. *Mítyane ó ávyejpiítte o í-cúné boóne.* Estoy muy dolorido después de jugar.

avyéjtso *abs.* agravamiento. || acción de...

[ávyejtso] *vt.* agravar; intensificar, hacer más fuerte. *Cuuve ó ávyéjtsoó wácyujpa amomé íbóhítuki.* Por la tarde haré más fuerte el veneno para pescar para que los peces no revivan.

ávyēju *abs.* 1. reino. *Naavéné avyéjú íñe ímítyunéhjí.* El reino de Satanás consiste en cosas malas. 2. gloria. *Píí-vyéébé avyéjú máájtyumí íévéhóówari.* La gloria de Dios se ve en el cielo.

ávyéjuúbe, ávyéjújte, ávyéjújte *s.* jefe, curaca, cacique; patrón, señor; rey, soberano; presidente.

ávyéjújkimyéi *abs.* superación, engrandecimiento. || acción de...

[ávyéjújkimyéi] *vr.* superarse, engrandecerse, volverse importante. *Ídátsohrée néhíjcyaaabe tsúúca ávyéjújkimyéi náhjíheho ímyéénúneri.* El que era considerado humilde se volvió importante por la bodega que tiene. (sinón. **íhnaóóve, íhstutúúve**)

ávyéjújtsaméi *abs.* acción y efecto de...

[ávyéjújtsaméi] *vr.* exaltarse, ser exaltado, ser engrandecido.

ávyéjújtso *abs.* exaltación, glorificación.

Grammars

(1)a. Segmentos Consonantais

Oclusivas

p t k

Africadas

ts tʃ

Fricativas

β s ʃ ʒ h

Tepe

r

Nasais

m n

Aproximantes

w y

b. Segmentos Vocálicos

i í u

a

(Parallel) Texts

MATÉO

Matéó Jetsóodíttyú cáátúnuháámi

Jetsocríjtó íhdé múnááúvu
(Rc. 3.23-38)

1 ¹Dabíi, Aavaráa, íjcyámútsiúvúj tsiiménémúhaabéhjáa Jetsocríjtó íjcyáábé íhdé múnááúvú ílluúme: ²Diiyéhjáa Aavaráa tsímavá Itsáake. Aabéhjáa Itsáa tsímavá Jacóóbóke. Aabéhjáa Jacóóbó tsímavá Jodáake tsíjtyehjí íñahbémuma. ³Aabéhjáa Jodáa tsímavá Tamáadívú Páareé, Tsáaraá, íjcyámútsikye. Aabéhjáa Páare tsímavá Ejróoke. Aabéhjáa Ejróo tsímavá Aráake. ⁴Aabéhjáa Aráa tsímavá Aminadáake. Aabéhjáa Aminadáa tsímavá Naatsóoke. Aabéhjáa Naatsóo tsímavá Tsamóoke. ⁵Aabéhjáa Tsamóo tsímavá Bóoke Raadívú. Aabéhjáa Bóo tsímavá Obéeke Róótídivu. Aabéhjáa Obéé tsímavá Itsaíikye. ⁶Aabéhjáa Itsaíí tsímavá ávyéjuube Dabíikye. Aabéhjáa Dabíí tsímavá Tsaromóoke Oriá táabadívu.

⁷Aabéhjáa Tsaromóo tsímavá Roboáake. Aabéhjáa Roboáa tsímavá Abíake. Aabéhjáa Abíá tsímavá Áátsáake. ⁸Aabéhjáa Áátsa tsímavá Jotsapáake. Aabéhjáa Jotsapáa tsímavá Joráake. Aabéhjáa Joráa tsímavá Otsiáke. ⁹Aabéhjáa Otsiá tsímavá Joatáake. Aabéhjáa Joatáa tsímavá Acáake. Aabéhjáa Acáa tsímavá Etsékíake. ¹⁰Aabéhjáa Etsékíá tsímavá Manatséake. Aabéhjáa Manatséé tsímavá

Amóoke. Aabéhjáa Amóo tsímavá Jotsiáke. ¹¹Aabéhjáa Jotsiá tsímavá Jeconíake tsíjtyehjí íñahbémuma. Téjyúhjáa Babiróónía múnáa hójtsiri méihdeéjte íjraéemu.

¹²Aaméhjáa diityé hójtsí pañétú íjchívyéné boone Jeconíá tsímavá Tsaratiéeke. Aabéhjáa Tsaratiéé tsímavá Tsorobabéke. ¹³Aabéhjáa Tsorobabéé tsímavá Abiíoóke. Aabéhjáa Abiíoó tsímavá Eriakiikye. Aabéhjáa Eriakíí tsímavá Atsóoke. ¹⁴Aabéhjáa Atsóo tsímavá Tsadóoke. Aabéhjáa Tsadóo tsímavá Akíikye. Aabéhjáa Akíí tsímavá Eriáoóke. ¹⁵Aabéhjáa Eriáoó tsímavá Ereatsáake. Aabéhjáa Ereatsáa tsímavá Matáake. Aabéhjáa Matáa tsímavá Jacóóbóke. ¹⁶Aabéhjáa Jacóóbó tsímavá Jotséeke Maaríá tajíí íjcyáábeke. Aanéhjáa diille Maaríá tsímavá Jetsóoke, 'Pájtyetétsoóbe' néeyíoné 'Críjtoó' némeíbyeke.

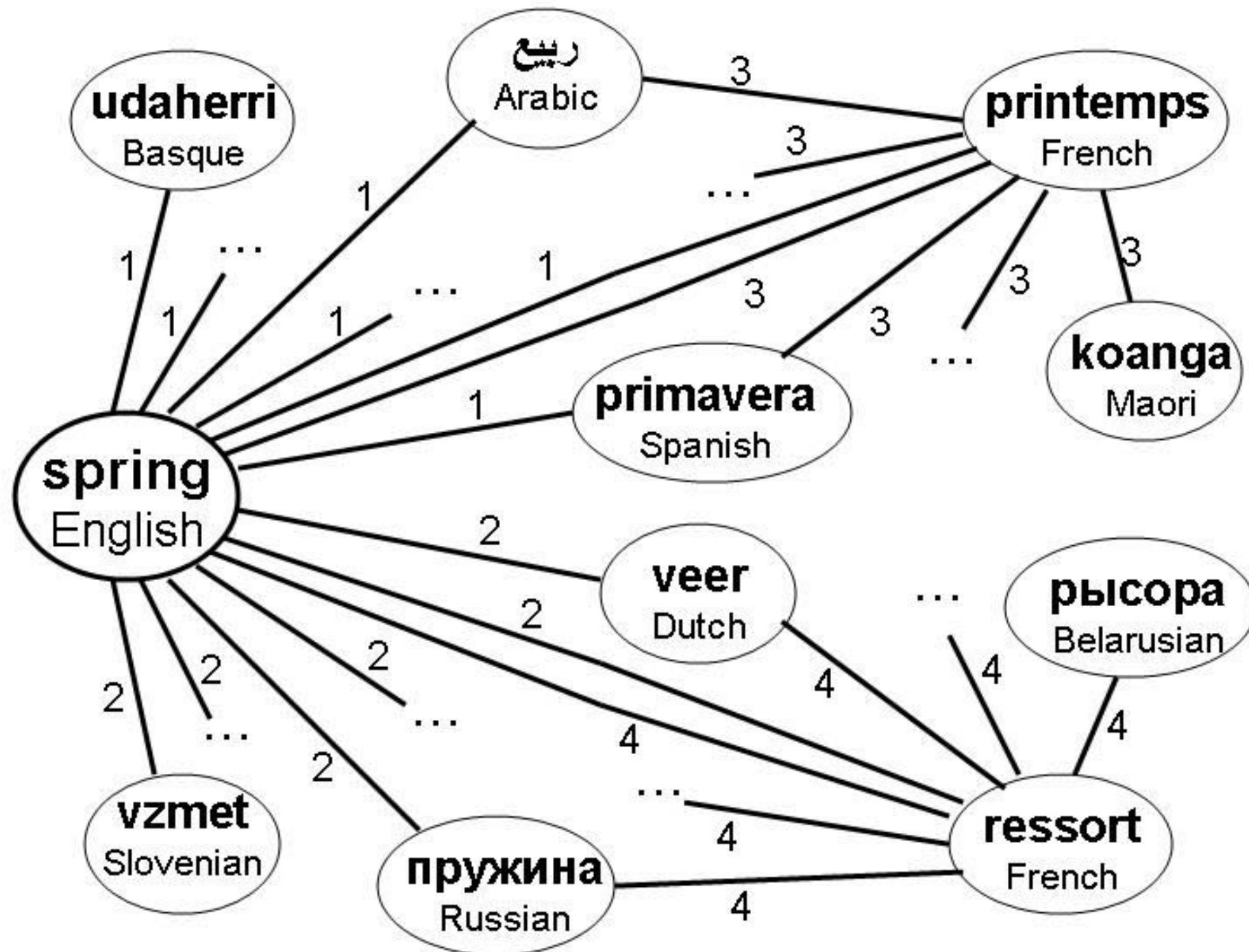
¹⁷Ehduméhjáa Jetsóo íhdé múnááúvu. Diiyéhjáa muurá Aavaráaúvú tsímavá tsáapíikye. Aabéhjáa dsíjívénetu dííbyé hajchíyējuco cámaavéné éhniínevu. Aabéhjáa idyé dsíjívénetu dííbyé hajchíyējuco cámaavéne. Ehdúhjáa icyóhodu 'camácámá' líiyaatéhíjcyámé 14-ihjyúvú Dabíiúvuke ditye tsímavanévújuco. Áabeúvudíttyúhjáa idyé líiyaatéhíjcyámé téhdunéré 14-íjyúva. Áijyúhjáa diityéké Babiróónía

Goals

- Explain language similarity/dissimilarity
- Language universals
 - ▶ i.e. regularities that hold independently of historical processes
- Phylogeny reconstruction
 - ▶ i.e. 'language trees', but language history has a very strong influence from horizontal process

Interesting Problems

- Use non-parallel resources
 - ▶ e.g. make inferences on a translation graph
- Extract parallel resources
 - ▶ Dictionaries → Wordlist
 - ▶ Grammars → Typological comparison
 - ▶ Texts → Parallel Sentences
- Chunking within parallel resources
- Alignment between parallel resources

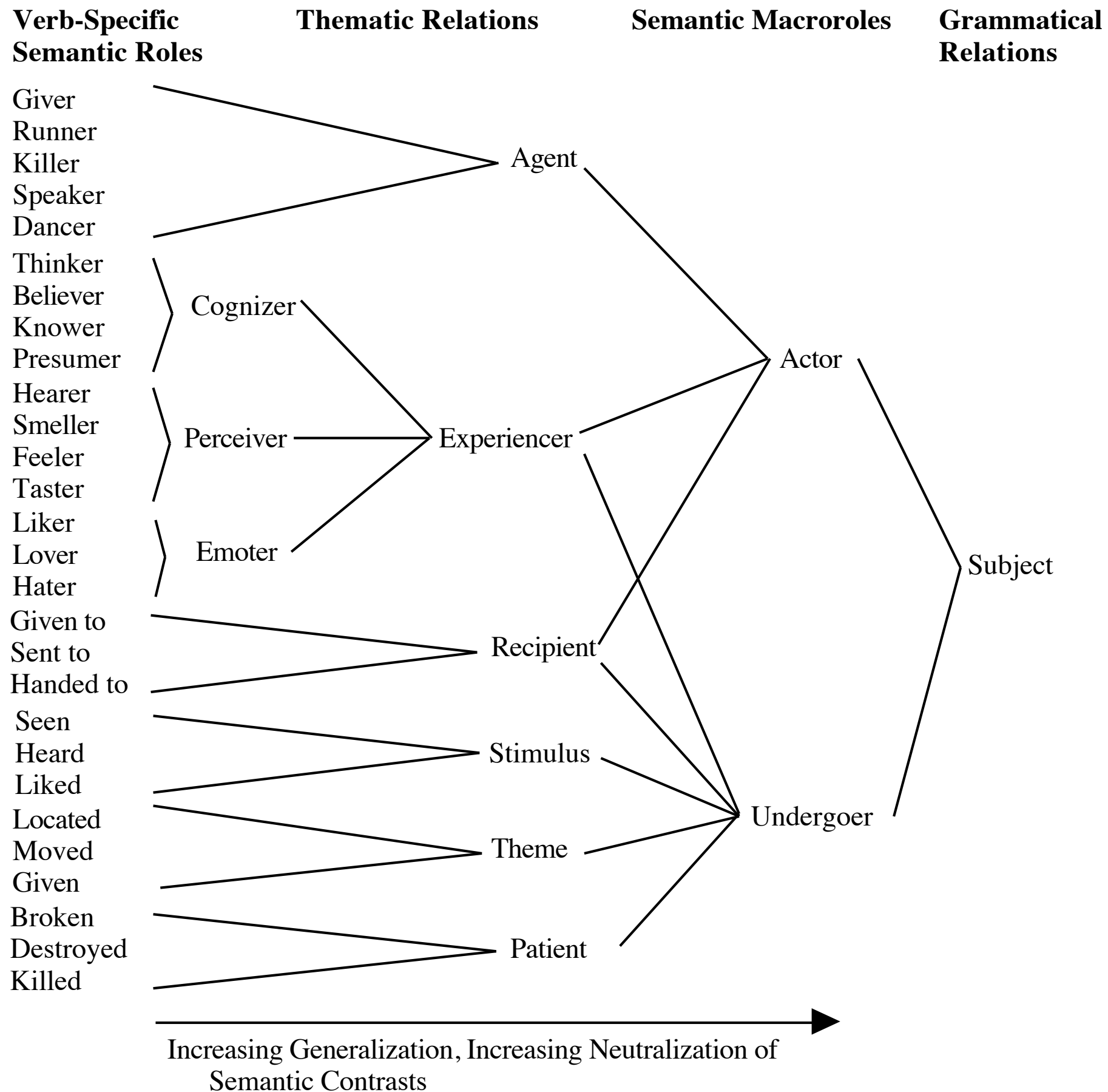


Alignment

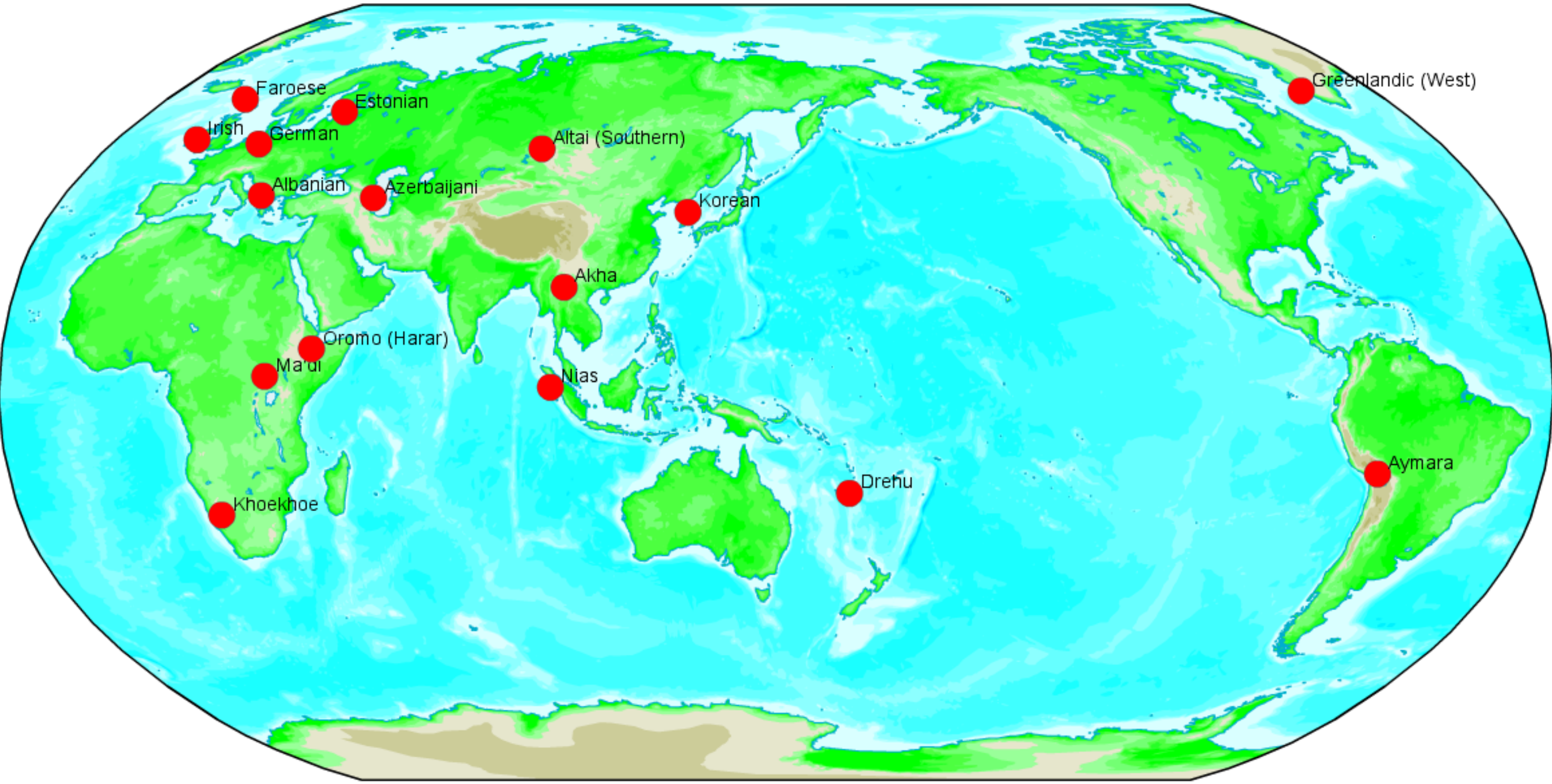
- **Sound-alignment between words**
 - ▶ find regular sound correspondences, and use them to infer cognate words across languages
- **Word-alignment between sentences**
 - ▶ find functional equivalences across languages, and use them to infer typological parameters

| Language | IDS | meaning | alignment |
|---------------|--------------|-----------------|------------------------|
| Nivaclé | 5.124 | unripe | n i y i ? y a |
| <i>Pilagá</i> | 5.125 | rotten | n i č i ? y a |
| Wichí | 10.240 | drip | n i t o n |
| Wichí | 4.591 | dribble | n i t u n |
| <i>Pilagá</i> | 5.130 | drink | n i y o m |
| Wichí | 5.370/3.655 | spoon/shell | l a n e k |
| <i>Toba</i> | 5.370 | spoon | l e m e k |
| <i>Pilagá</i> | 5.370 | spoon | l e m e k |
| Nivaclé | 15.830 | wet | w a ? a i |
| <i>Pilagá</i> | 1.329/1.320 | ocean/sea | w a ʁ a i |
| Wichí | 7.330/10.710 | chimney/road | n o y i h |
| <i>Mocoví</i> | 10.710 | road | n a ? i k |
| Nivaclé | 14.332 | for a long time | k a x u ? |
| <i>Mocoví</i> | 14.332 | for a long time | ḵ a w a ? |
| Wichí | 1.520/14.530 | sun/clock | h ^w a l a ? |
| <i>Toba</i> | 1.520 | sun | n a l a ? |
| Wichí | 9.220 | cut | y i s e t |
| <i>Pilagá</i> | 9.110 | do/make | y i ? e t |
| Wichí | 9.210/4.760 | hit/kill | i l o n |
| <i>Toba</i> | 18.420 | call by name | i l o n |
| Wichí | 17.172 | imitate | i t e n |
| <i>Mocoví</i> | 16.510 | dare | i t e n |
| Wichí | 4.858 | scar | l a h ɲ i |
| <i>Pilagá</i> | 4.374 | footprint | l i i ɲ i |
| Maca | 3.585 | hawk | m i y o |
| <i>Toba</i> | 3.950 | frog | m i y o |
| Nivaclé | 19.590 | prevent | f a ? m a t a n |
| <i>Mocoví</i> | 16.670 | tell lies | n a ? m a h a n |
| Chorote | 5.123 | ripe | y o w e ? |
| <i>Toba</i> | 5.123 | ripe | y a m o ḵ |

Inducing Semantic Roles



- 1 What important information is contained in the Bible?
- 2 Who is the Bible's author?
- 3 Why should you study the Bible?
- 4 The Bible is a precious gift from God.
- 5 The Bible alone tells us what we must do to please God.
- 6 The Bible was written by some 40 different men over a period of 1,600 years, beginning in 1513 B.C.E.
- 7 So God in heaven, not any human on earth, is the Author of the Bible.
- 8 God made sure that the Bible was accurately copied and preserved.
- 9 More Bibles have been printed than any other book.
- 10 Not everyone will be happy to see you studying the Bible, but do not let that stop you.
- 11 But the Bible tells us that there is only one TRUE God.
- 12 But when the Bible was written, the name Jehovah appeared in it some 7,000 times
- 13 God is a Spirit, says the Bible.
- 14 The Bible reveals Jehovah's personality to us.
- 15 The Bible tells us that he is also merciful, kind, forgiving, generous, and patient.
- 16 We learn about God from creation and from the Bible.
- 17 Another way we can learn about God is by studying the Bible.
- 18 By disobeying God's command, the first man, Adam, committed what the Bible calls sin.
- 19 This is what the Bible refers to as the ransom.
- 20 Some of your loved ones may become very angry because you are studying the Bible.
- 21 What is the Bible's view of separation and of divorce?
- 22 The Bible says that a husband is the head of his family.
- 23 Parents need to spend time with their children and study the Bible with them,
- 24 When marriage mates have problems getting along together, they should try to apply Bible counsel.
- 25 The Bible urges us to show love and to be forgiving.
- 26 But God does not approve of them if they come from false religion or are against Bible teachings.
- 27 The only two birthday celebrations spoken of in the Bible were held by persons who did not worship Jehovah.
- 28 The Bible teaches that only a few people are on the narrow road to life.
- 29 The Bible foretold that after the death of the apostles, ...
- 30 True Christians love one another, respect the Bible, and preach about God's Kingdom.
- 31 Another mark of true religion is that its members have a deep respect for the Bible.
- 32 They try to live by the Bible in their everyday life.
- 33 The Bible is the basis for what is taught.
- 34 By now you have learned many good things from the Bible.



Faroese

Estonian

Irish

German

Albanian

Azerbaijani

Altai (Southern)

Korean

Akha

Nias

Khoekhoe

Maui

Oromo (Harar)

Drehu

Greenlandic (West)

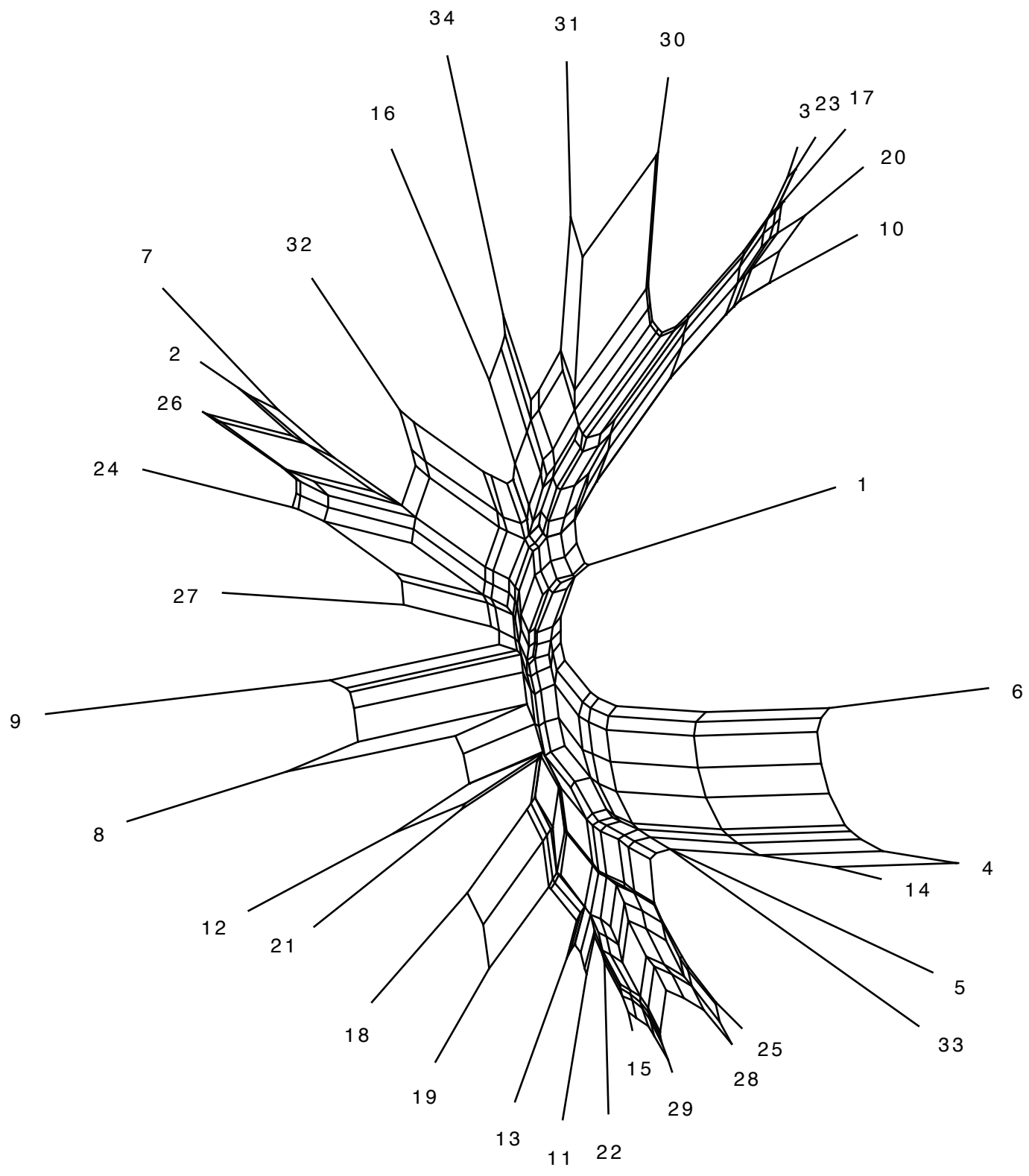
Aymara

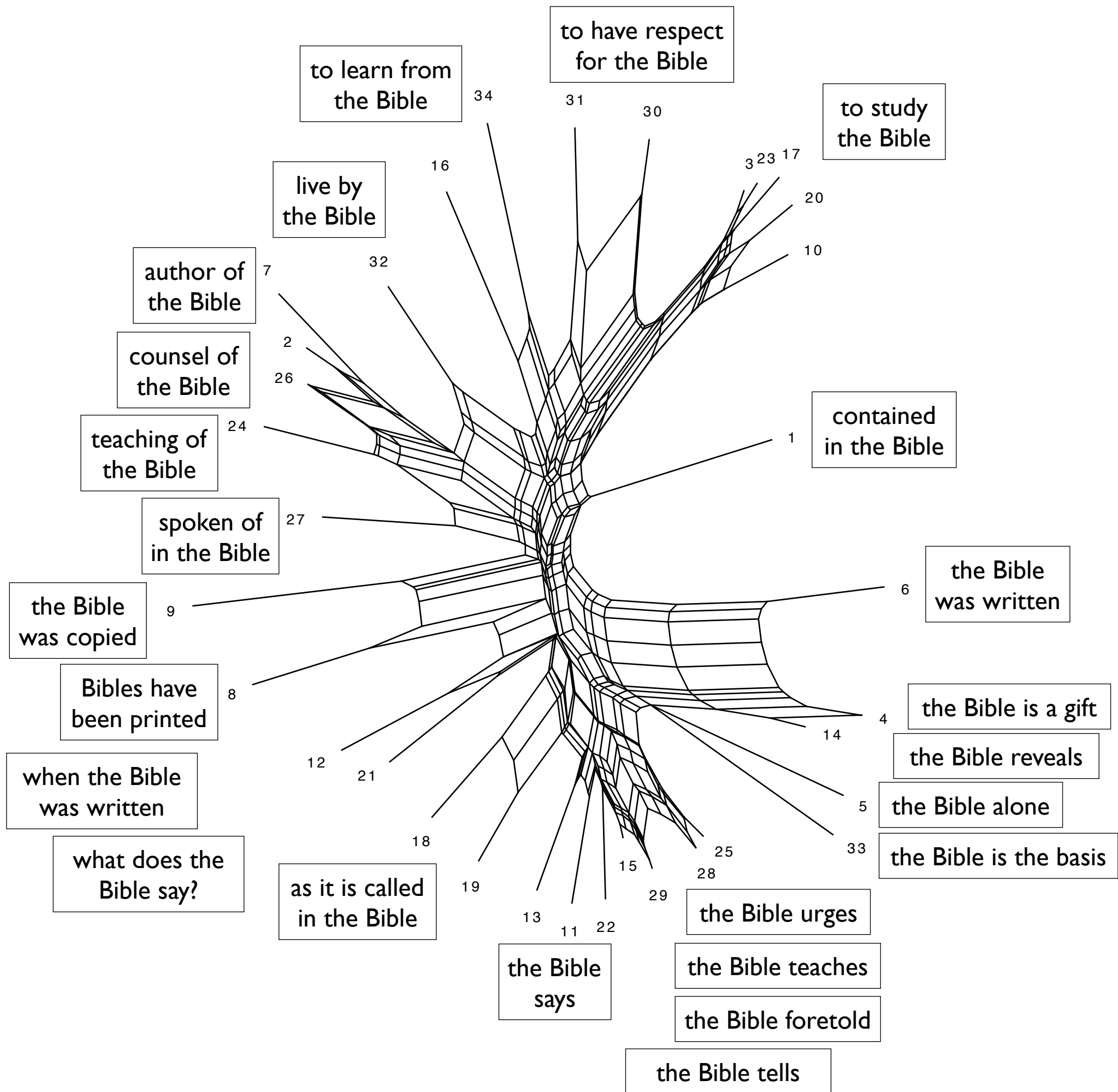
| Language | Genus | Family | Macroarea | Grammatical “Alignment” |
|-------------|------------------|---------------|-----------|-------------------------|
| Oromo | Eastern Cushitic | Afro-Asiatic | Africa | Marked nominative |
| Khoekhoe | Central Khoisan | Khoisan | Africa | Nominative-accusative |
| Ma'di | Moru-Ma'di | Nilo-Saharan | Africa | No case marking (?) |
| Albanian | Albanian | Indo-European | Europe | Nominative-accusative |
| Irish | Celtic | Indo-European | Europe | No case marking (?) |
| Faroese | Germanic | Indo-European | Europe | Nominative-accusative |
| Estonian | Finnic | Uralic | Europe | Nominative-accusative |
| Altai | Turkic | Altaic | Asia | Nominative-accusative |
| Azerbaijani | Turkic | Altaic | Asia | Nominative-accusative |
| Korean | Korean | Korean | Asia | Nominative-accusative |
| Akha | Burmese-Lolo | Sino-Tibetan | Asia | Ergative-absolutive (?) |
| Drehu | Oceanic | Austronesian | Pacific | Active-inactive |
| Nias | Sundic | Austronesian | Pacific | Marked absolutive |
| Greenlandic | Eskimo-Aleut | Eskimo-Aleut | America | Ergative-absolutive |
| Aymara | Aymaran | Aymaran | America | Marked Nominative (?) |

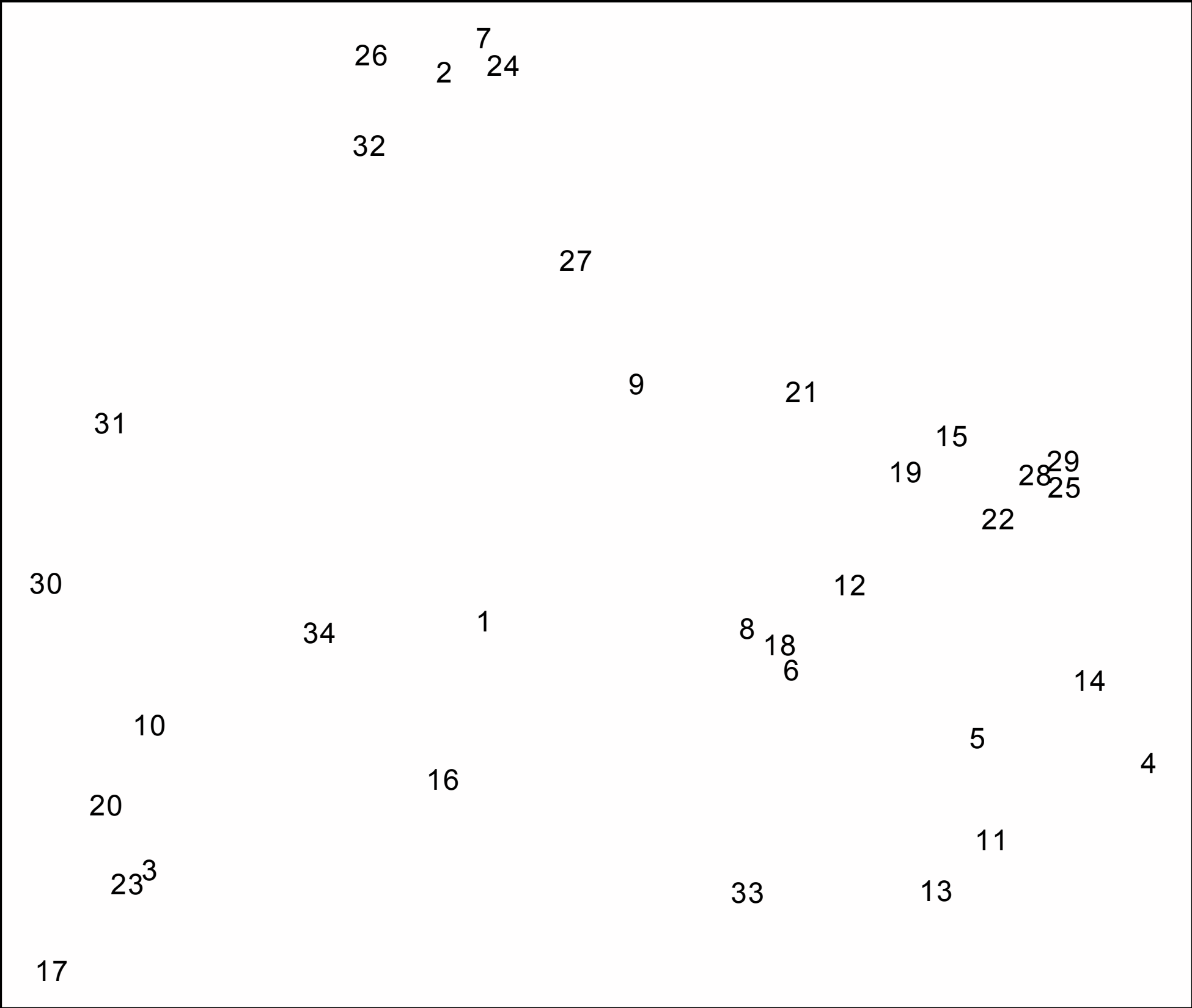
| Albanian | Faroese | Estonian | Greenlandic |
|----------------------------------|-------------------------------|-----------------------------|----------------------------------|
| <i>bibla</i> Nominative | <i>biblian</i> Nominative | <i>piibel</i> Nominative | <i>biibilip</i> Ergative |
| <i>biblën</i> Accusative | <i>bibliuna</i> Accusative | <i>piiblit</i> Partitive | <i>biibli</i> Absolutive |
| <i>biblës</i> Genitive/Dative | <i>bibliunnar</i> Genitive | <i>piibli</i> Genitive | <i>biibilmik</i> Instrumental |
| ... | <i>bibliuni</i> Dative | <i>piiblis</i> Inessive | <i>biibilmi</i> Locative |
| | ... | <i>piiblist</i> Elative | ... |
| | | ... | |

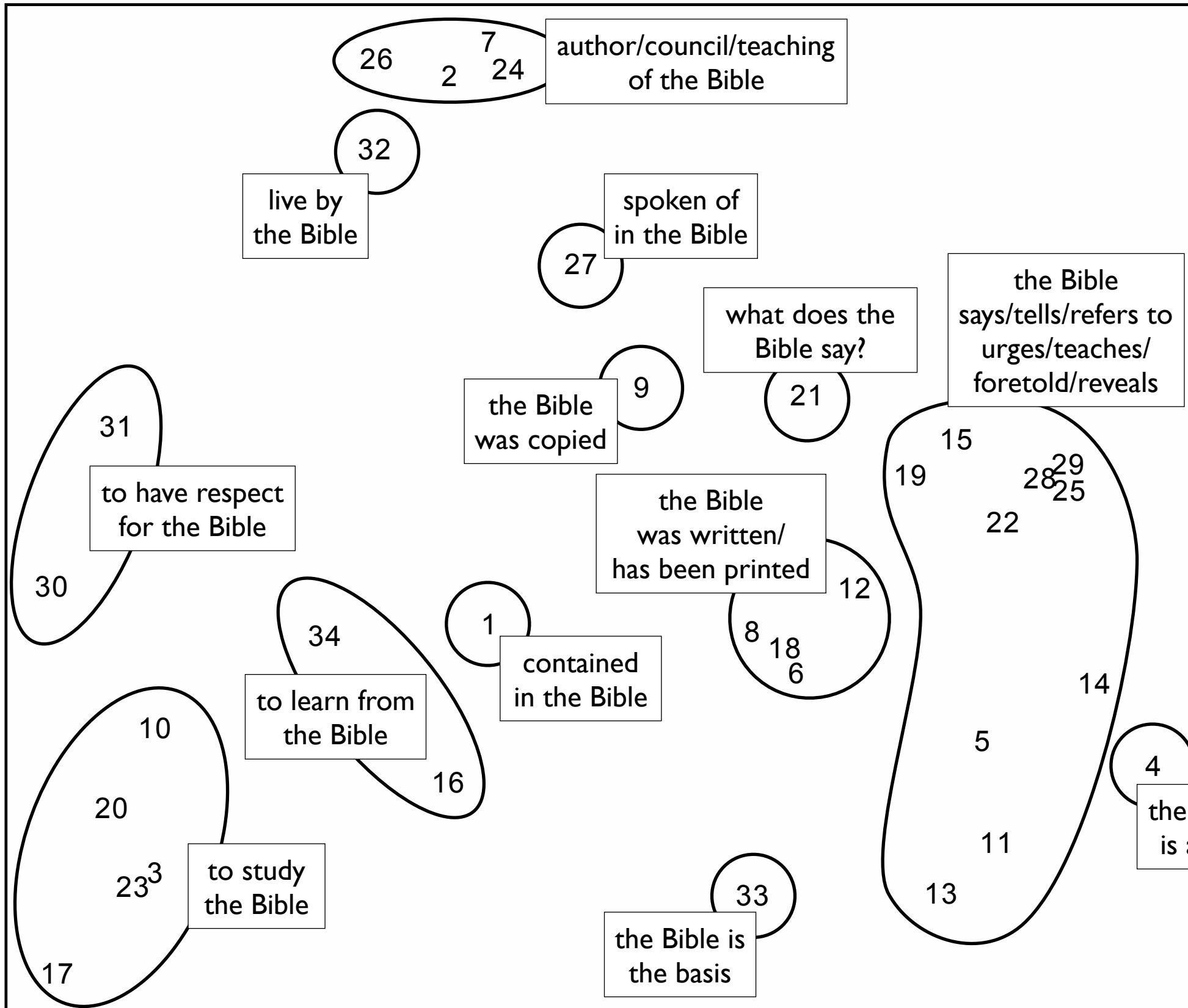
| Context | Albanian | Faroese | Estonian | Greenlandic |
|---------|----------|------------|----------|------------------|
| 1 | bibla | bíbliuni | piibel | biibili |
| 2 | biblës | bíbliunnar | piibli | biibilimik |
| 3 | biblën | bíbliuna | piiblit | biibili |
| 4 | bibla | bíblían | piibel | biibili |
| 5 | bibla | bíblían | piibel | biibilip |
| 6 | bibla | bíbliuna | piibli | biibili |
| 7 | biblës | bíbliunnar | piibli | biibilimut |
| 8 | bibla | bíblían | piiblit | biibilip |
| 9 | bibla | NA | piiblit | biibili |
| 10 | biblën | bíbliuna | piiblit | biibilimik |
| 11 | bibla | bíblían | piibel | biibilimili |
| 12 | bibla | bíblían | piibel | biibilili |
| 13 | bibla | bíblían | piibel | biibilimi |
| 14 | bibla | bíblían | piibel | biibilimi |
| 15 | bibla | bíblían | piibel | biibilimi |
| 16 | bibla | bíbliuni | piibli | biibililu |
| 17 | biblën | bíbliuna | piiblit | biibilimik |
| 18 | bibla | bíblían | piiblis | biibilip |
| 19 | bibla | bíblían | piiblis | biibilimi |
| 20 | biblën | bíbliuna | piiblit | biibilimik |
| 21 | NA | bíblían | piibel | biibilimi |
| 22 | bibla | bíbliuni | piibel | biibili |
| 23 | biblën | bíbliuna | piiblit | biibilimillu |
| 24 | biblike | bíblían | piibli | biibilimi |
| 25 | bibla | bíblían | piibel | biibilimi |
| 26 | biblës | bíbliunnar | piibli | biibilimi |
| 27 | bibla | bíblían | piiblis | biibilimi |
| 28 | bibla | bíblían | piibel | biibilimi |
| 29 | bibla | bíblían | piibel | biibilimi |
| 30 | biblën | bíbliuna | piiblist | biibilimik |
| 31 | biblën | bíbliuni | piibli | biibilimik |
| 32 | biblës | bíbliuni | piibli | biibili |
| 33 | bibla | bíbliuna | piibel | biibilimik |
| 34 | bibla | bíbliuni | piiblist | biibilimeersunik |

| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|
| 0.00 | 0.67 | 0.50 | 0.75 | 0.67 | 0.75 | 0.69 | 0.75 | 0.73 | 0.56 | 0.44 | 0.63 | 0.53 | 0.75 | 0.53 | 0.56 | 0.63 | 0.50 | 0.63 | 0.63 | 0.60 | 0.50 | 0.56 | 0.69 | 0.75 | 0.57 | 0.44 | 0.69 | 0.56 | 0.63 | 0.63 | 0.50 | 0.75 | 0.63 |
| 0.67 | 0.00 | 0.73 | 0.80 | 0.73 | 0.67 | 0.20 | 0.73 | 0.64 | 0.60 | 0.80 | 0.67 | 0.93 | 0.73 | 0.57 | 0.80 | 0.73 | 0.73 | 0.73 | 0.67 | 0.50 | 0.67 | 0.73 | 0.40 | 0.60 | 0.15 | 0.53 | 0.67 | 0.67 | 0.60 | 0.60 | 0.38 | 0.73 | 0.80 |
| 0.50 | 0.73 | 0.00 | 0.88 | 0.80 | 0.69 | 0.81 | 0.75 | 0.80 | 0.25 | 0.75 | 0.75 | 0.73 | 0.88 | 0.80 | 0.69 | 0.13 | 0.69 | 0.81 | 0.19 | 0.73 | 0.81 | 0.06 | 0.75 | 0.88 | 0.64 | 0.69 | 0.88 | 0.88 | 0.44 | 0.63 | 0.64 | 0.75 | 0.69 |
| 0.75 | 0.80 | 0.88 | 0.00 | 0.40 | 0.31 | 0.81 | 0.56 | 0.53 | 0.88 | 0.44 | 0.44 | 0.40 | 0.13 | 0.47 | 0.75 | 0.94 | 0.63 | 0.63 | 0.94 | 0.53 | 0.44 | 0.94 | 0.88 | 0.25 | 0.93 | 0.81 | 0.31 | 0.44 | 1.00 | 1.00 | 0.79 | 0.50 | 0.94 |
| 0.67 | 0.73 | 0.80 | 0.40 | 0.00 | 0.60 | 0.80 | 0.60 | 0.79 | 0.80 | 0.47 | 0.47 | 0.57 | 0.47 | 0.50 | 0.73 | 0.87 | 0.53 | 0.60 | 0.87 | 0.57 | 0.53 | 0.80 | 0.80 | 0.47 | 0.77 | 0.67 | 0.47 | 0.47 | 0.93 | 1.00 | 0.85 | 0.53 | 0.87 |
| 0.75 | 0.67 | 0.69 | 0.31 | 0.60 | 0.00 | 0.69 | 0.63 | 0.53 | 0.75 | 0.63 | 0.63 | 0.67 | 0.44 | 0.67 | 0.81 | 0.75 | 0.75 | 0.75 | 0.88 | 0.60 | 0.56 | 0.75 | 0.81 | 0.56 | 0.79 | 0.81 | 0.63 | 0.63 | 0.81 | 0.81 | 0.71 | 0.63 | 0.94 |
| 0.69 | 0.20 | 0.81 | 0.81 | 0.80 | 0.69 | 0.00 | 0.63 | 0.60 | 0.75 | 0.81 | 0.75 | 0.93 | 0.75 | 0.67 | 0.81 | 0.81 | 0.81 | 0.81 | 0.81 | 0.60 | 0.75 | 0.81 | 0.50 | 0.75 | 0.29 | 0.63 | 0.69 | 0.69 | 0.69 | 0.69 | 0.50 | 0.81 | 0.81 |
| 0.75 | 0.73 | 0.75 | 0.56 | 0.60 | 0.63 | 0.63 | 0.00 | 0.47 | 0.88 | 0.56 | 0.50 | 0.67 | 0.63 | 0.53 | 0.75 | 0.75 | 0.63 | 0.69 | 0.81 | 0.80 | 0.69 | 0.75 | 0.75 | 0.63 | 0.79 | 0.75 | 0.56 | 0.56 | 0.75 | 0.75 | 0.86 | 0.81 | 0.75 |
| 0.73 | 0.64 | 0.80 | 0.53 | 0.79 | 0.53 | 0.60 | 0.47 | 0.00 | 0.73 | 0.80 | 0.73 | 0.86 | 0.67 | 0.64 | 0.80 | 0.87 | 0.87 | 0.87 | 0.80 | 0.86 | 0.67 | 0.87 | 0.73 | 0.67 | 0.69 | 0.73 | 0.67 | 0.73 | 0.80 | 0.80 | 0.62 | 0.80 | 0.80 |
| 0.56 | 0.60 | 0.25 | 0.88 | 0.80 | 0.75 | 0.75 | 0.88 | 0.73 | 0.00 | 0.81 | 0.75 | 0.80 | 0.81 | 0.80 | 0.69 | 0.25 | 0.69 | 0.81 | 0.19 | 0.67 | 0.88 | 0.25 | 0.69 | 0.81 | 0.57 | 0.63 | 0.81 | 0.88 | 0.44 | 0.63 | 0.57 | 0.63 | 0.81 |
| 0.44 | 0.80 | 0.75 | 0.44 | 0.47 | 0.63 | 0.81 | 0.56 | 0.80 | 0.81 | 0.00 | 0.56 | 0.27 | 0.38 | 0.27 | 0.69 | 0.81 | 0.44 | 0.44 | 0.81 | 0.53 | 0.31 | 0.75 | 0.81 | 0.38 | 0.86 | 0.63 | 0.31 | 0.19 | 0.88 | 0.88 | 0.86 | 0.56 | 0.81 |
| 0.63 | 0.67 | 0.75 | 0.44 | 0.47 | 0.63 | 0.75 | 0.50 | 0.73 | 0.75 | 0.56 | 0.00 | 0.53 | 0.50 | 0.47 | 0.69 | 0.81 | 0.44 | 0.56 | 0.81 | 0.33 | 0.63 | 0.75 | 0.63 | 0.50 | 0.64 | 0.63 | 0.56 | 0.56 | 0.88 | 0.88 | 0.71 | 0.69 | 0.81 |
| 0.53 | 0.93 | 0.73 | 0.40 | 0.57 | 0.67 | 0.93 | 0.67 | 0.86 | 0.80 | 0.27 | 0.53 | 0.00 | 0.27 | 0.36 | 0.60 | 0.73 | 0.40 | 0.33 | 0.80 | 0.43 | 0.40 | 0.73 | 0.73 | 0.33 | 0.77 | 0.60 | 0.33 | 0.33 | 0.80 | 0.80 | 0.85 | 0.60 | 0.73 |
| 0.75 | 0.73 | 0.88 | 0.13 | 0.47 | 0.44 | 0.75 | 0.63 | 0.67 | 0.81 | 0.38 | 0.50 | 0.27 | 0.00 | 0.33 | 0.69 | 0.88 | 0.56 | 0.50 | 0.88 | 0.40 | 0.44 | 0.88 | 0.75 | 0.13 | 0.79 | 0.69 | 0.19 | 0.31 | 0.94 | 0.94 | 0.79 | 0.44 | 0.88 |
| 0.53 | 0.57 | 0.80 | 0.47 | 0.50 | 0.67 | 0.67 | 0.53 | 0.64 | 0.80 | 0.27 | 0.47 | 0.36 | 0.33 | 0.00 | 0.73 | 0.87 | 0.33 | 0.27 | 0.80 | 0.50 | 0.20 | 0.80 | 0.53 | 0.20 | 0.57 | 0.33 | 0.20 | 0.07 | 0.80 | 0.80 | 0.64 | 0.60 | 0.73 |
| 0.56 | 0.80 | 0.69 | 0.75 | 0.73 | 0.81 | 0.81 | 0.75 | 0.80 | 0.69 | 0.69 | 0.69 | 0.60 | 0.69 | 0.73 | 0.00 | 0.63 | 0.63 | 0.75 | 0.69 | 0.80 | 0.75 | 0.69 | 0.75 | 0.75 | 0.71 | 0.75 | 0.75 | 0.81 | 0.69 | 0.56 | 0.57 | 0.63 | 0.44 |
| 0.63 | 0.73 | 0.13 | 0.94 | 0.87 | 0.75 | 0.81 | 0.75 | 0.87 | 0.25 | 0.81 | 0.81 | 0.73 | 0.88 | 0.87 | 0.63 | 0.00 | 0.75 | 0.88 | 0.19 | 0.80 | 0.94 | 0.13 | 0.81 | 0.94 | 0.71 | 0.75 | 0.94 | 0.94 | 0.31 | 0.50 | 0.79 | 0.63 | 0.63 |
| 0.50 | 0.73 | 0.69 | 0.63 | 0.53 | 0.75 | 0.81 | 0.63 | 0.87 | 0.69 | 0.44 | 0.44 | 0.40 | 0.56 | 0.33 | 0.63 | 0.75 | 0.00 | 0.25 | 0.75 | 0.53 | 0.44 | 0.69 | 0.63 | 0.56 | 0.64 | 0.38 | 0.56 | 0.44 | 0.81 | 0.81 | 0.64 | 0.69 | 0.69 |
| 0.63 | 0.73 | 0.81 | 0.63 | 0.60 | 0.75 | 0.81 | 0.69 | 0.87 | 0.81 | 0.44 | 0.56 | 0.33 | 0.50 | 0.27 | 0.75 | 0.88 | 0.25 | 0.00 | 0.81 | 0.47 | 0.31 | 0.81 | 0.56 | 0.38 | 0.57 | 0.31 | 0.38 | 0.25 | 0.81 | 0.81 | 0.64 | 0.69 | 0.75 |
| 0.63 | 0.67 | 0.19 | 0.94 | 0.87 | 0.88 | 0.81 | 0.81 | 0.80 | 0.19 | 0.81 | 0.81 | 0.80 | 0.88 | 0.80 | 0.69 | 0.19 | 0.75 | 0.81 | 0.00 | 0.80 | 0.88 | 0.19 | 0.75 | 0.81 | 0.64 | 0.69 | 0.81 | 0.88 | 0.44 | 0.63 | 0.64 | 0.69 | 0.69 |
| 0.60 | 0.50 | 0.73 | 0.53 | 0.57 | 0.60 | 0.60 | 0.80 | 0.86 | 0.67 | 0.53 | 0.33 | 0.43 | 0.40 | 0.50 | 0.80 | 0.80 | 0.53 | 0.47 | 0.80 | 0.00 | 0.60 | 0.73 | 0.40 | 0.40 | 0.46 | 0.53 | 0.47 | 0.47 | 0.80 | 0.80 | 0.62 | 0.67 | 0.87 |
| 0.50 | 0.67 | 0.81 | 0.44 | 0.53 | 0.56 | 0.75 | 0.69 | 0.67 | 0.88 | 0.31 | 0.63 | 0.40 | 0.44 | 0.20 | 0.75 | 0.94 | 0.44 | 0.31 | 0.88 | 0.60 | 0.00 | 0.88 | 0.75 | 0.31 | 0.71 | 0.56 | 0.31 | 0.19 | 0.88 | 0.81 | 0.57 | 0.56 | 0.75 |
| 0.56 | 0.73 | 0.06 | 0.94 | 0.80 | 0.75 | 0.81 | 0.75 | 0.87 | 0.25 | 0.75 | 0.75 | 0.73 | 0.88 | 0.80 | 0.69 | 0.13 | 0.69 | 0.81 | 0.19 | 0.73 | 0.88 | 0.00 | 0.75 | 0.88 | 0.64 | 0.69 | 0.88 | 0.88 | 0.44 | 0.63 | 0.71 | 0.75 | 0.69 |
| 0.69 | 0.40 | 0.75 | 0.88 | 0.80 | 0.81 | 0.50 | 0.75 | 0.73 | 0.69 | 0.81 | 0.63 | 0.73 | 0.75 | 0.53 | 0.75 | 0.81 | 0.63 | 0.56 | 0.75 | 0.40 | 0.75 | 0.75 | 0.00 | 0.63 | 0.21 | 0.38 | 0.63 | 0.63 | 0.69 | 0.63 | 0.43 | 0.81 | 0.75 |
| 0.75 | 0.60 | 0.88 | 0.25 | 0.47 | 0.56 | 0.75 | 0.63 | 0.67 | 0.81 | 0.38 | 0.50 | 0.33 | 0.13 | 0.20 | 0.75 | 0.94 | 0.56 | 0.38 | 0.81 | 0.40 | 0.31 | 0.88 | 0.63 | 0.00 | 0.64 | 0.56 | 0.06 | 0.19 | 0.88 | 0.88 | 0.64 | 0.50 | 0.81 |
| 0.57 | 0.15 | 0.64 | 0.93 | 0.77 | 0.79 | 0.29 | 0.79 | 0.69 | 0.57 | 0.86 | 0.64 | 0.77 | 0.79 | 0.57 | 0.71 | 0.71 | 0.64 | 0.57 | 0.64 | 0.46 | 0.71 | 0.64 | 0.21 | 0.64 | 0.00 | 0.36 | 0.64 | 0.64 | 0.57 | 0.57 | 0.31 | 0.86 | 0.71 |
| 0.44 | 0.53 | 0.69 | 0.81 | 0.67 | 0.81 | 0.63 | 0.75 | 0.73 | 0.63 | 0.63 | 0.63 | 0.60 | 0.69 | 0.33 | 0.75 | 0.75 | 0.38 | 0.31 | 0.69 | 0.53 | 0.56 | 0.69 | 0.38 | 0.56 | 0.36 | 0.00 | 0.56 | 0.44 | 0.63 | 0.63 | 0.43 | 0.81 | 0.69 |
| 0.69 | 0.67 | 0.88 | 0.31 | 0.47 | 0.63 | 0.69 | 0.56 | 0.67 | 0.81 | 0.31 | 0.56 | 0.33 | 0.19 | 0.20 | 0.75 | 0.94 | 0.56 | 0.38 | 0.81 | 0.47 | 0.31 | 0.88 | 0.63 | 0.06 | 0.64 | 0.56 | 0.00 | 0.13 | 0.88 | 0.88 | 0.64 | 0.50 | 0.81 |
| 0.56 | 0.67 | 0.88 | 0.44 | 0.47 | 0.63 | 0.69 | 0.56 | 0.73 | 0.88 | 0.19 | 0.56 | 0.33 | 0.31 | 0.07 | 0.81 | 0.94 | 0.44 | 0.25 | 0.88 | 0.47 | 0.19 | 0.88 | 0.63 | 0.19 | 0.64 | 0.44 | 0.13 | 0.00 | 0.88 | 0.88 | 0.71 | 0.56 | 0.81 |
| 0.63 | 0.60 | 0.44 | 1.00 | 0.93 | 0.81 | 0.69 | 0.75 | 0.80 | 0.44 | 0.88 | 0.88 | 0.80 | 0.94 | 0.80 | 0.69 | 0.31 | 0.81 | 0.81 | 0.44 | 0.80 | 0.88 | 0.44 | 0.69 | 0.88 | 0.57 | 0.63 | 0.88 | 0.88 | 0.00 | 0.31 | 0.57 | 0.75 | 0.56 |
| 0.63 | 0.60 | 0.63 | 1.00 | 1.00 | 0.81 | 0.69 | 0.75 | 0.80 | 0.63 | 0.88 | 0.88 | 0.80 | 0.94 | 0.80 | 0.56 | 0.50 | 0.81 | 0.81 | 0.63 | 0.80 | 0.81 | 0.63 | 0.63 | 0.88 | 0.57 | 0.63 | 0.88 | 0.88 | 0.31 | 0.00 | 0.36 | 0.81 | 0.56 |
| 0.50 | 0.38 | 0.64 | 0.79 | 0.85 | 0.71 | 0.50 | 0.86 | 0.62 | 0.57 | 0.86 | 0.71 | 0.85 | 0.79 | 0.64 | 0.57 | 0.79 | 0.64 | 0.64 | 0.64 | 0.62 | 0.57 | 0.71 | 0.43 | 0.64 | 0.31 | 0.43 | 0.64 | 0.71 | 0.57 | 0.36 | 0.00 | 0.79 | 0.64 |
| 0.75 | 0.73 | 0.75 | 0.50 | 0.53 | 0.63 | 0.81 | 0.81 | 0.80 | 0.63 | 0.56 | 0.69 | 0.60 | 0.44 | 0.60 | 0.63 | 0.63 | 0.69 | 0.69 | 0.69 | 0.67 | 0.56 | 0.75 | 0.81 | 0.50 | 0.86 | 0.81 | 0.50 | 0.56 | 0.75 | 0.81 | 0.79 | 0.00 | 0.81 |
| 0.63 | 0.80 | 0.69 | 0.94 | 0.87 | 0.94 | 0.81 | 0.75 | 0.80 | 0.81 | 0.81 | 0.81 | 0.73 | 0.88 | 0.73 | 0.44 | 0.63 | 0.69 | 0.75 | 0.69 | 0.87 | 0.75 | 0.69 | 0.75 | 0.81 | 0.71 | 0.69 | 0.81 | 0.81 | 0.56 | 0.56 | 0.64 | 0.81 | 0.00 |

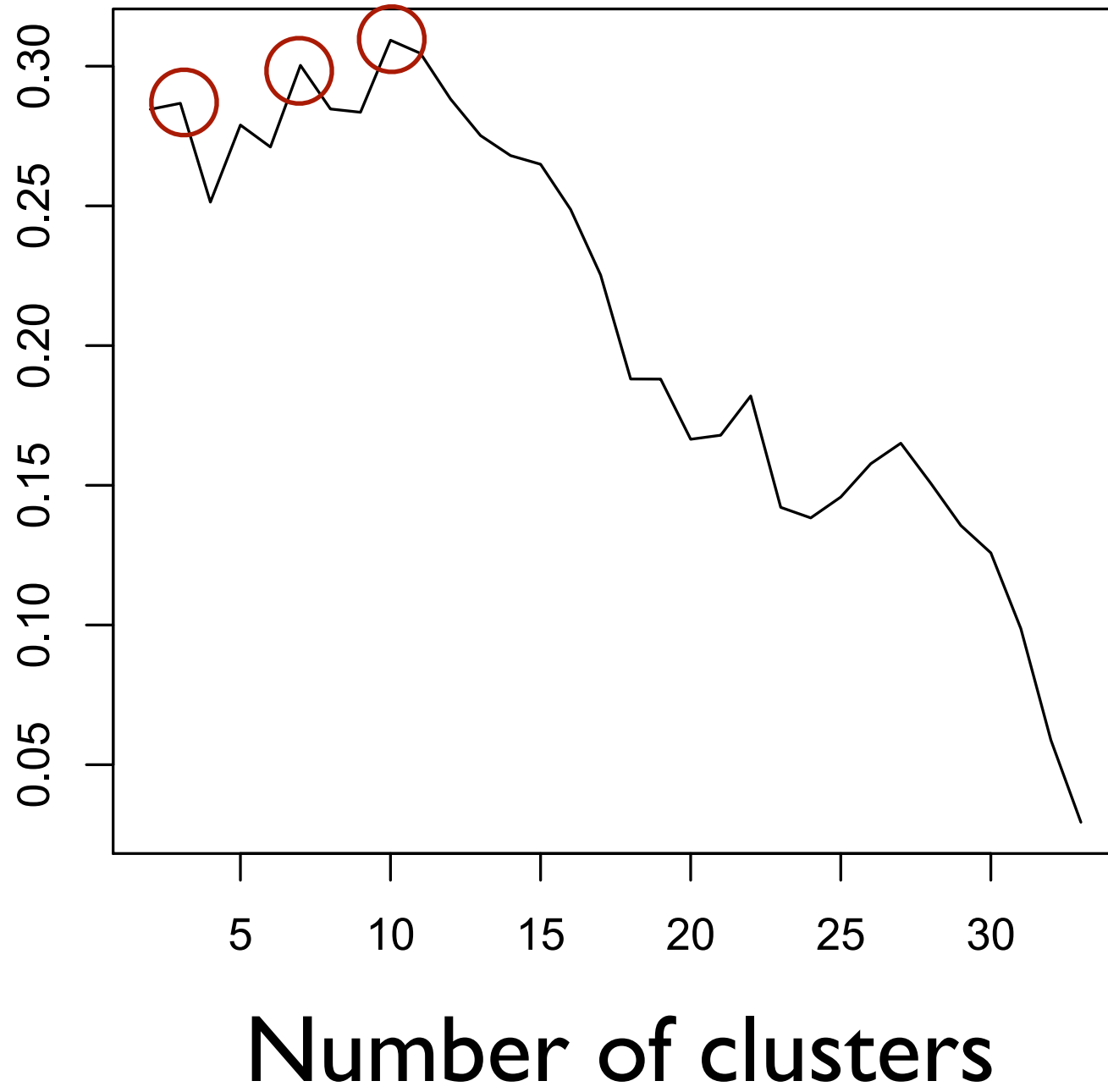


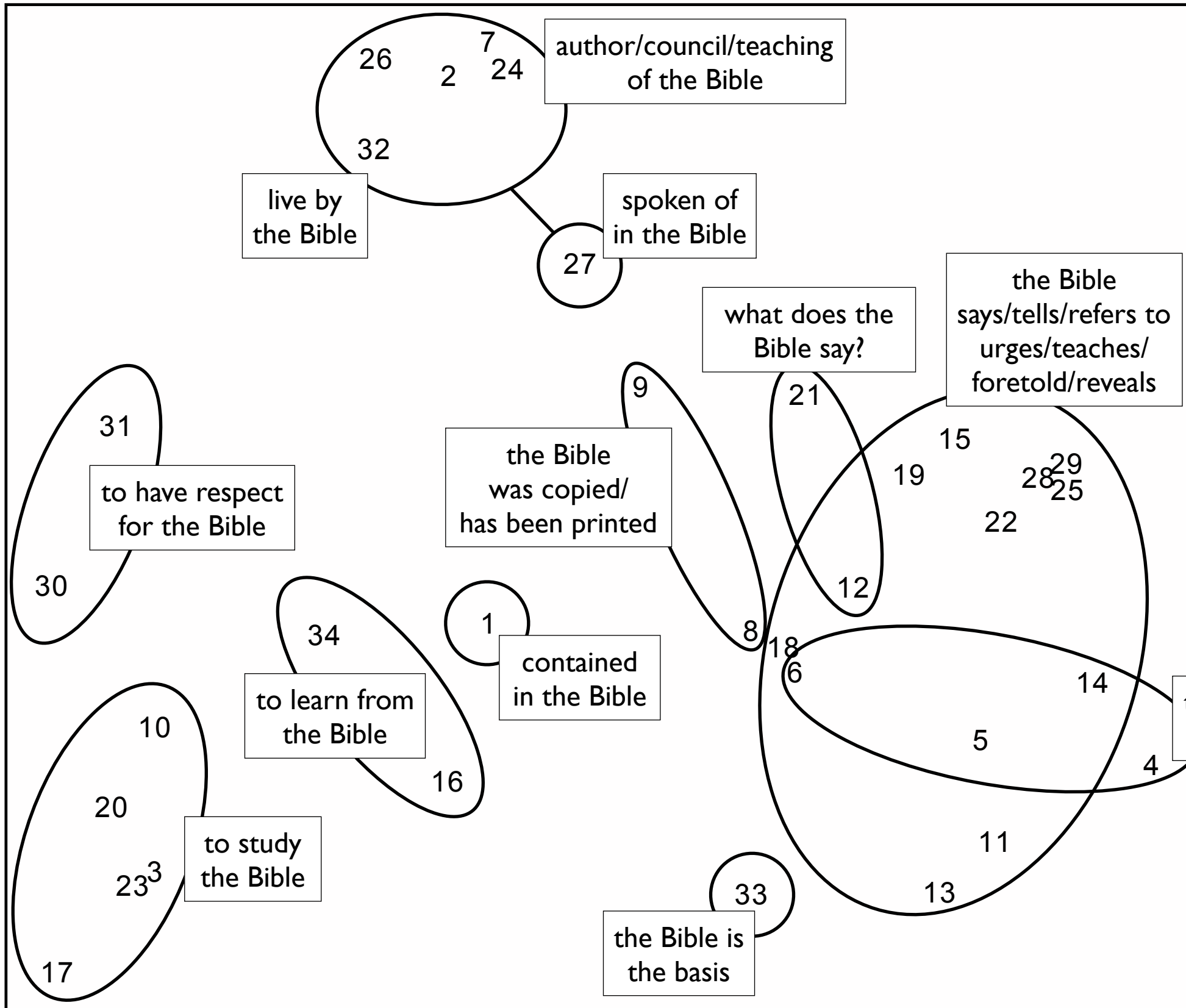


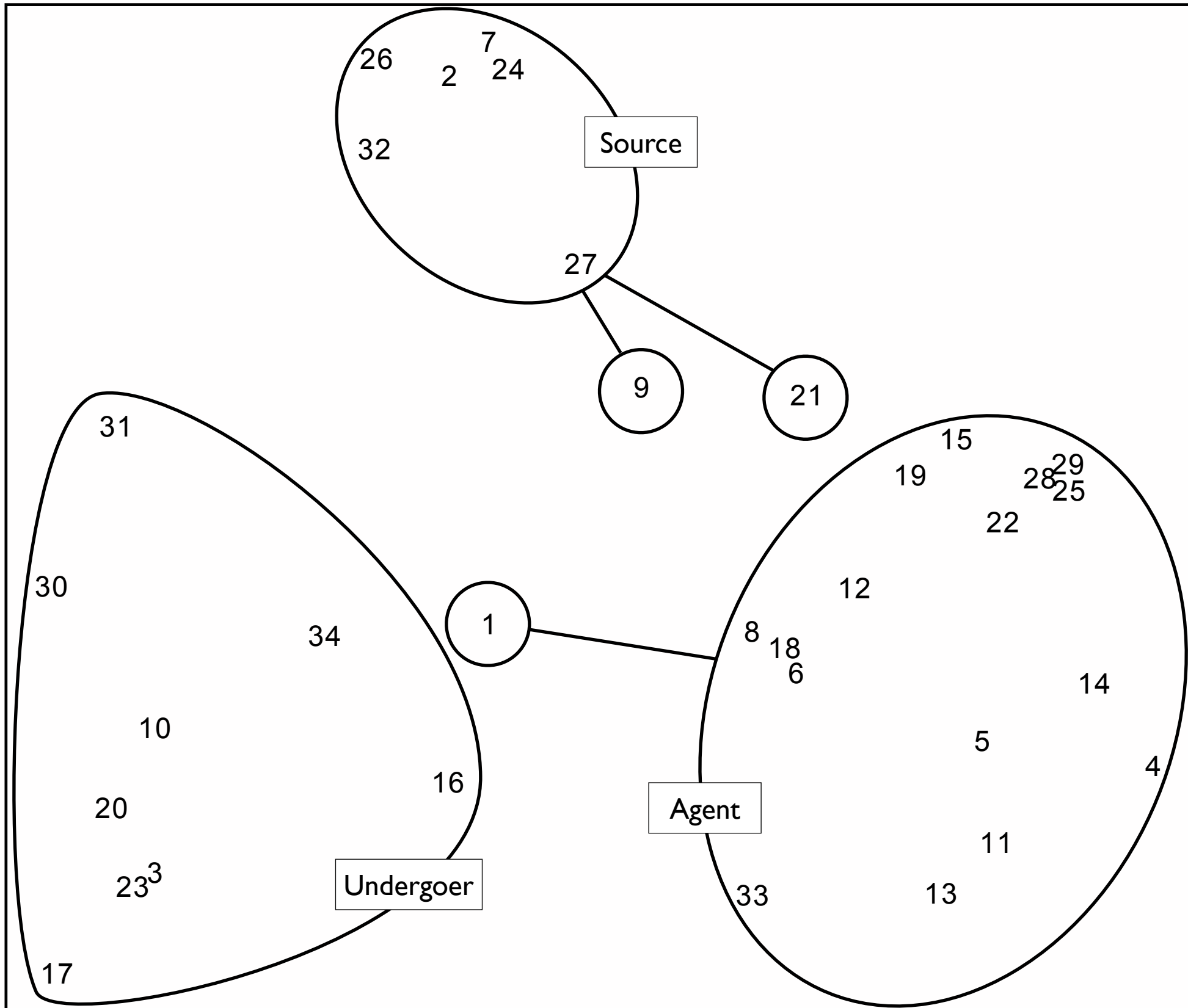


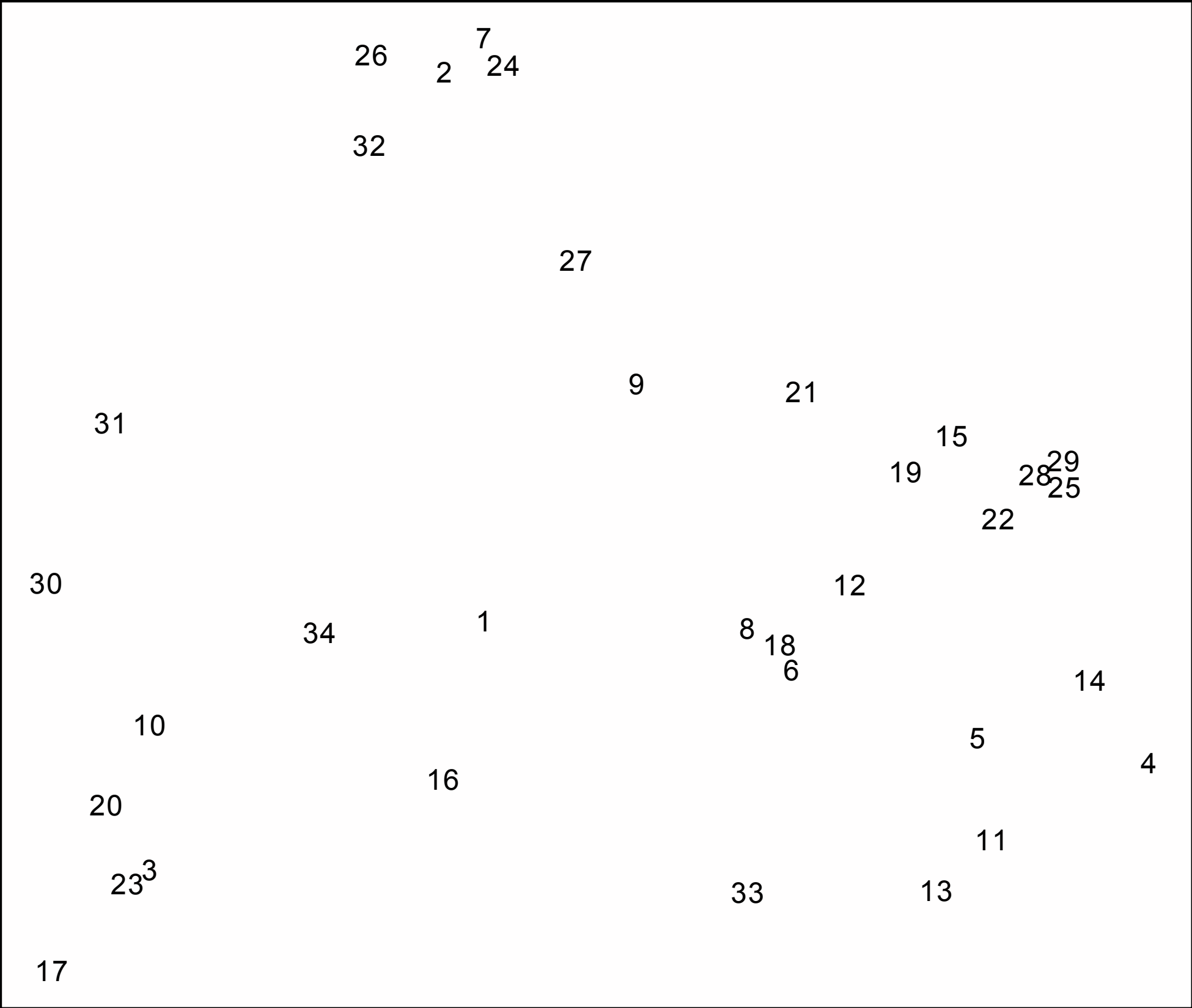


“Fit” of clustering

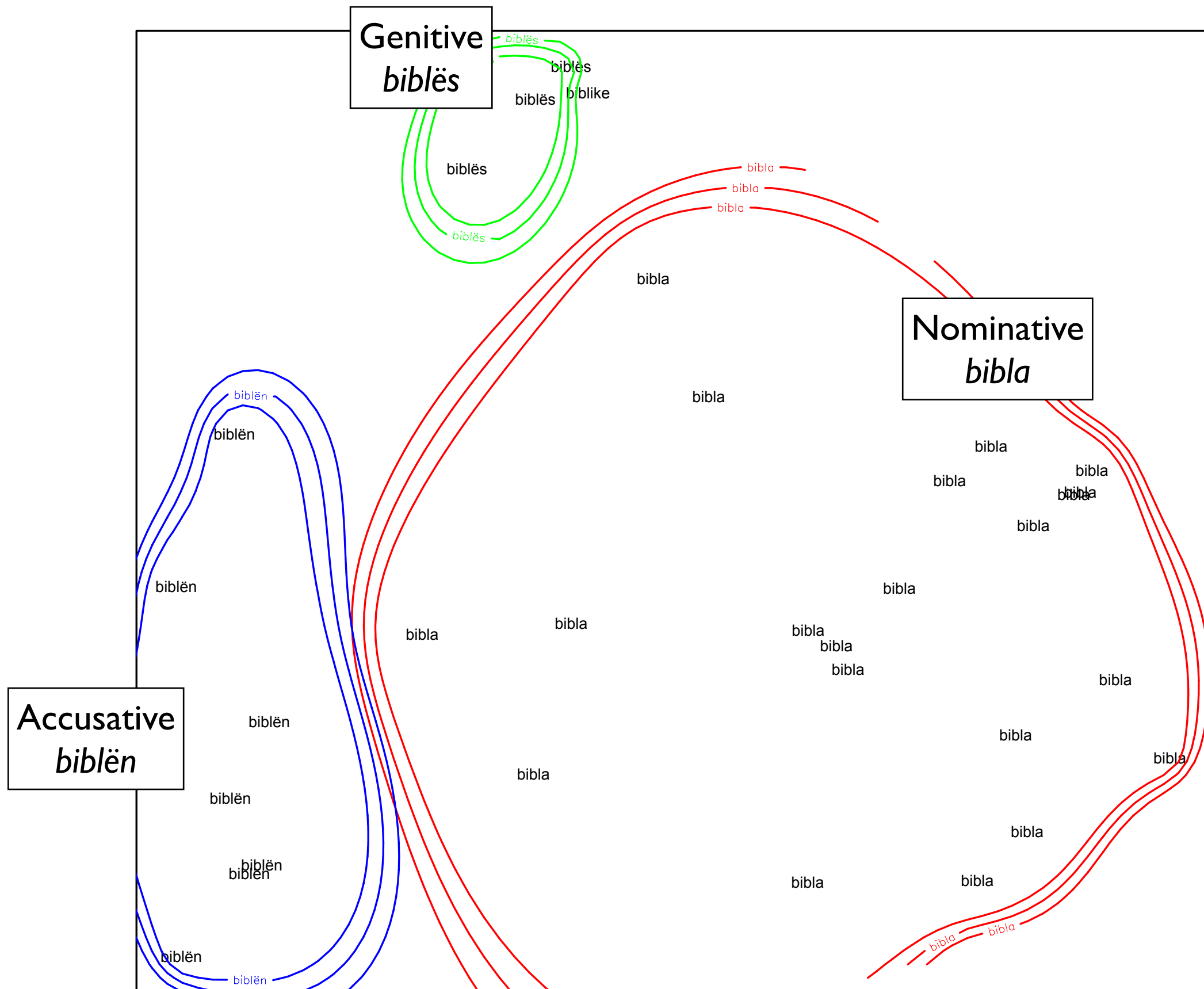








albanian



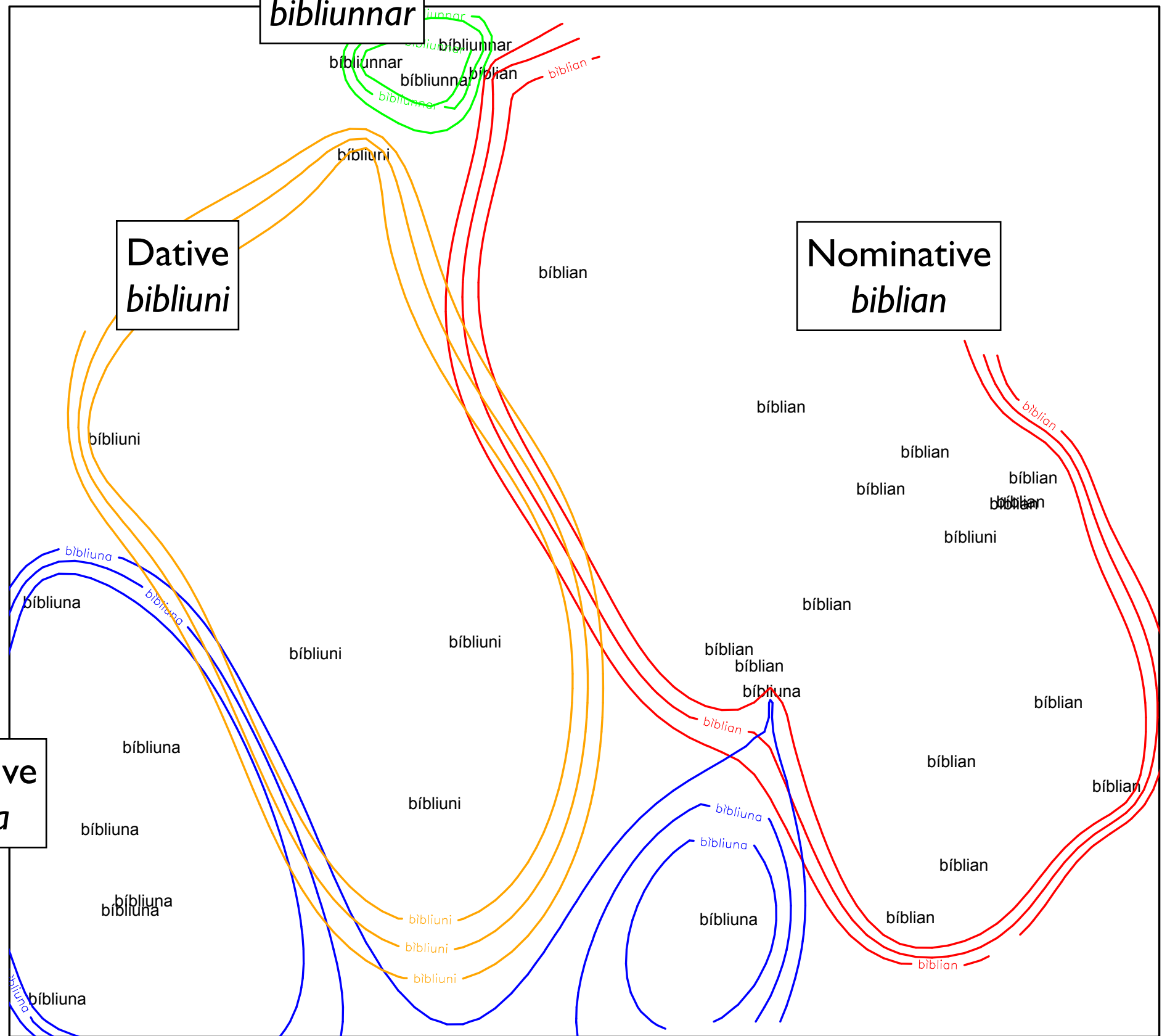
faroese

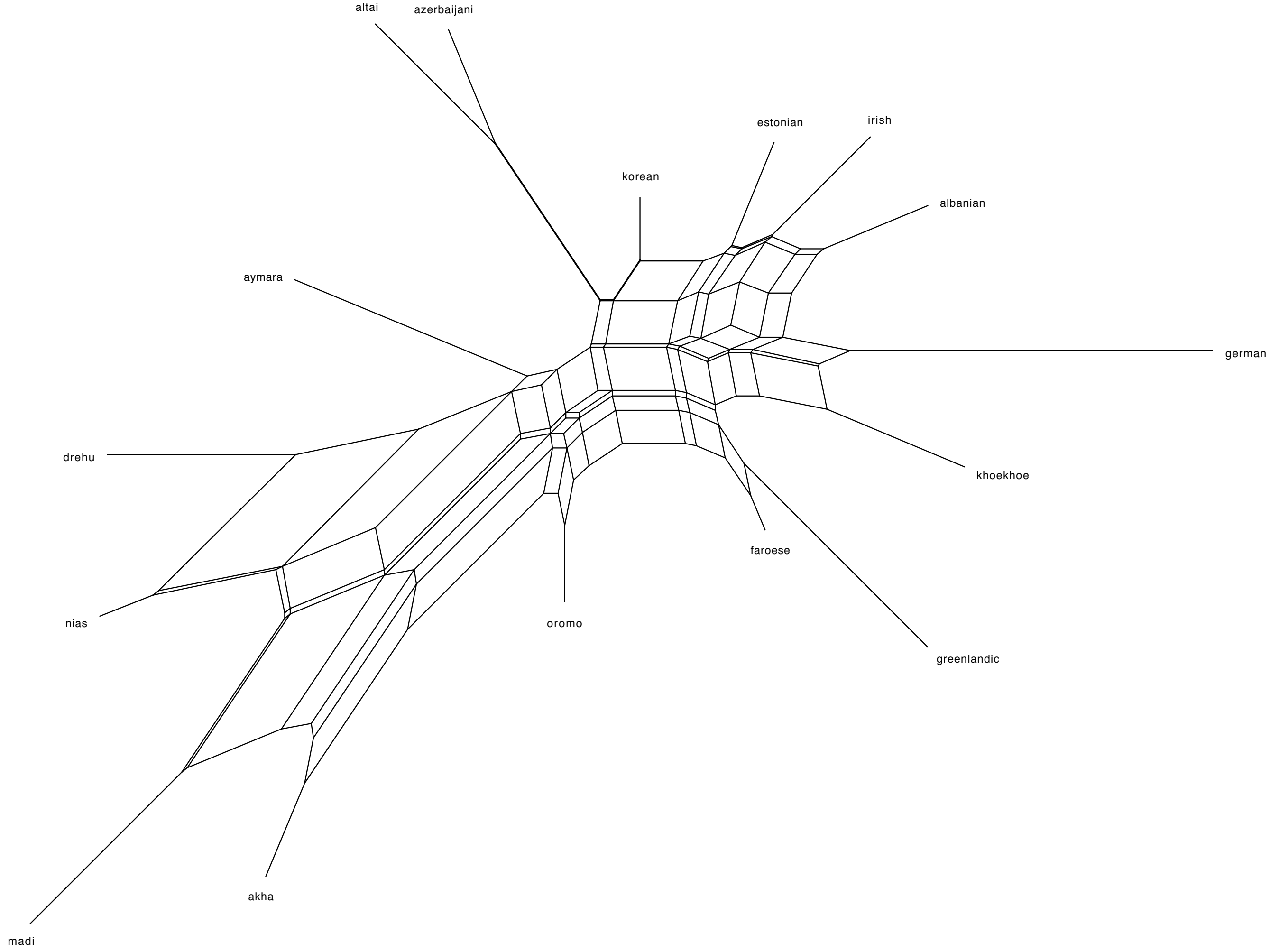
Genitive
bibliunnar

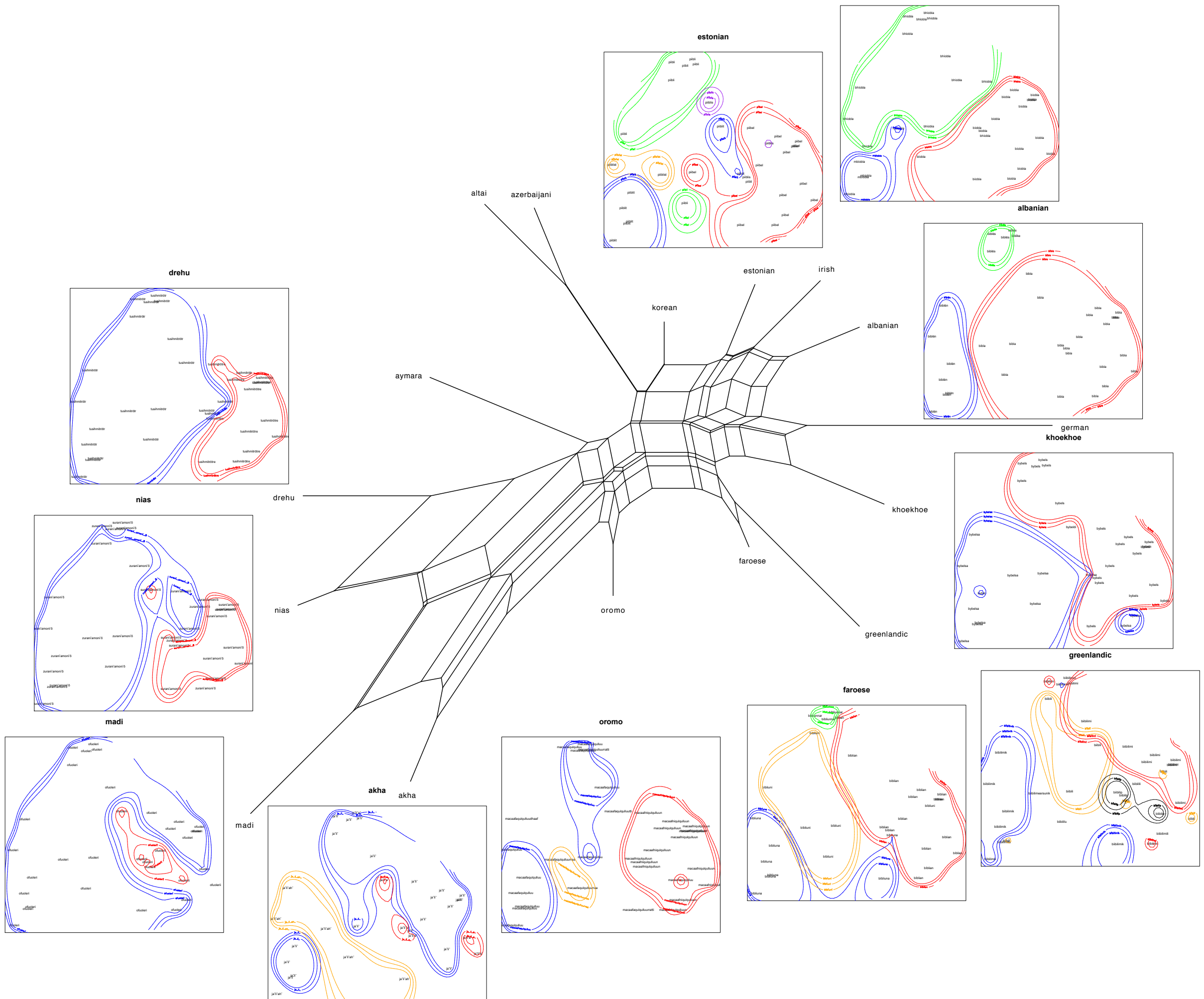
Dative
bibliuni

Nominative
biblian

Accusative
bibliuna







altai
azerbaijani

estonian

irish

albanian

drehu

estonian

irish

albanian

korean

german
khoekhoe

aymara

khoekhoe

nias

drehu

faroese

oromo

greenlandic

nias

greenlandic

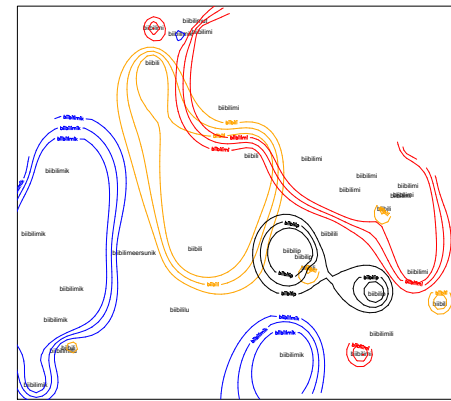
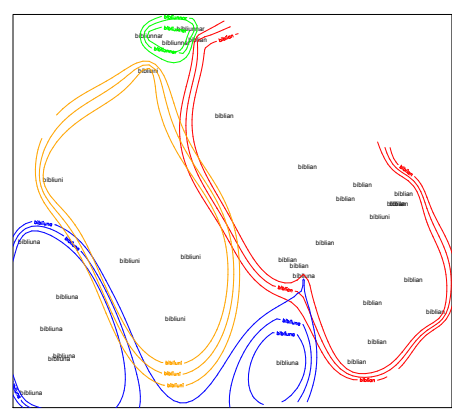
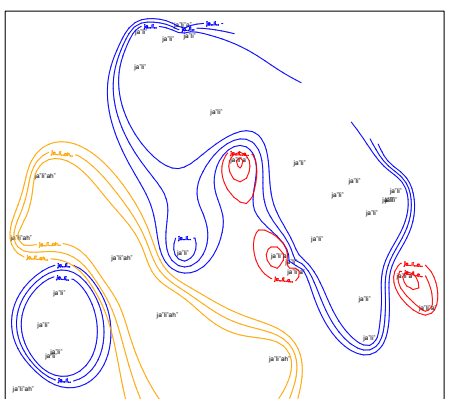
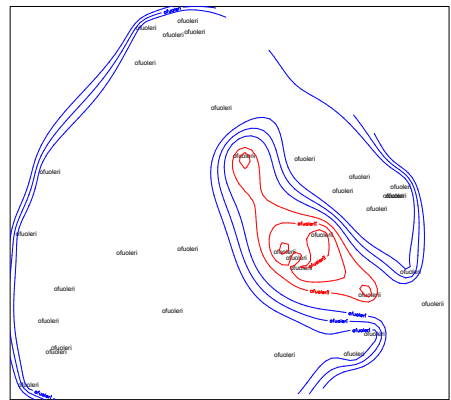
madi

faroese

oromo

akha
akha

madi



Conclusions

- Cross-linguistic diversity in marking gives us a window into grammatical structure
- Grammatical relations can be induced from parallel texts
- Typology can be induced from the language-specific clustering in texts