Three of a kind?

Multi-alignment of amino-acids, sounds and words

Michael Cysouw
Philipps-Universität Marburg

Philipps Universität Marburg

Multi-alignment

A. Biological Phylogenetics
   (alignment of nucleotides/amino-acids)

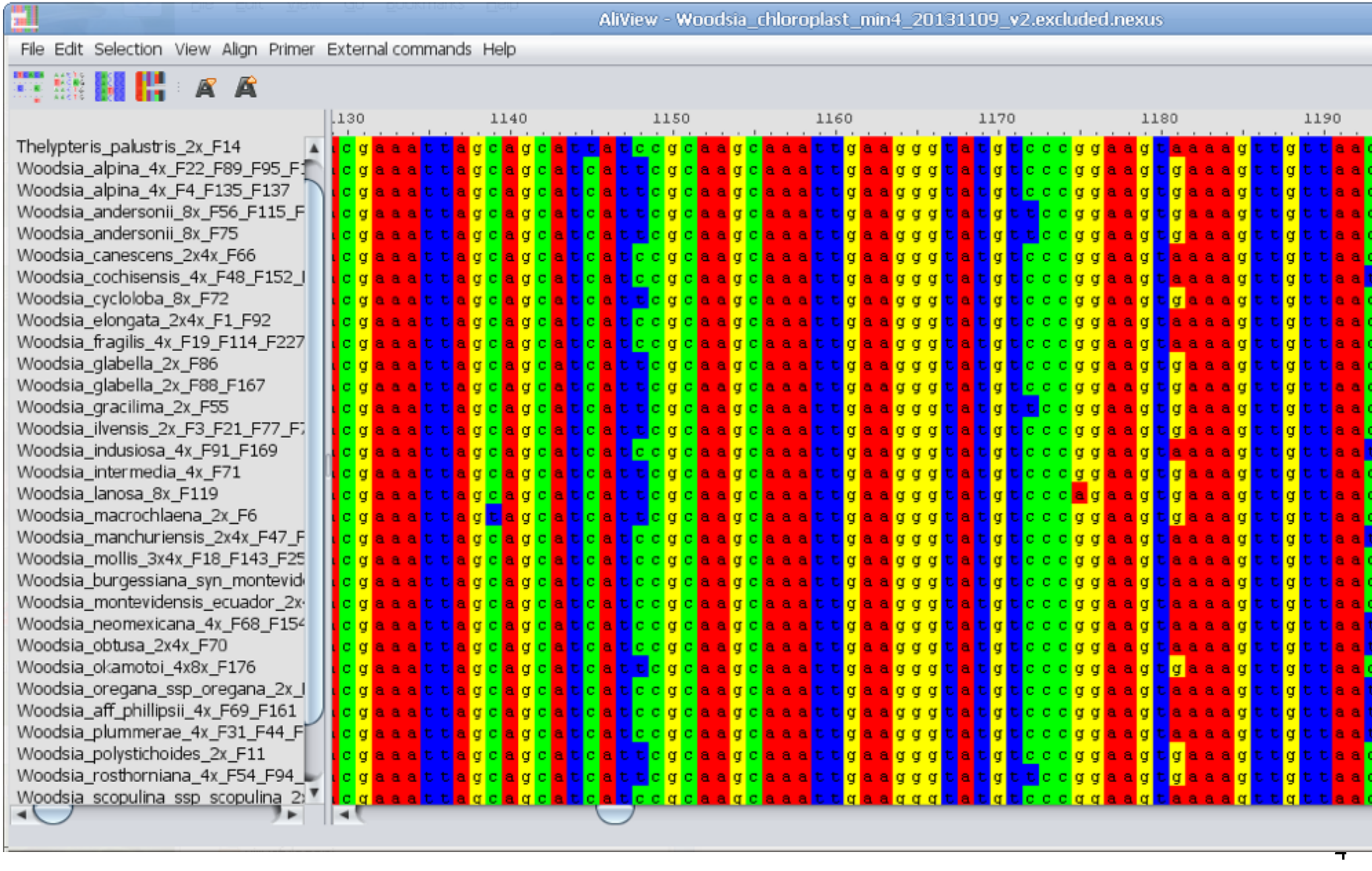
B. Historical Linguistics
   (alignment of sounds in words)

C. Functional Language Comparison
   (alignment of words in sentences)

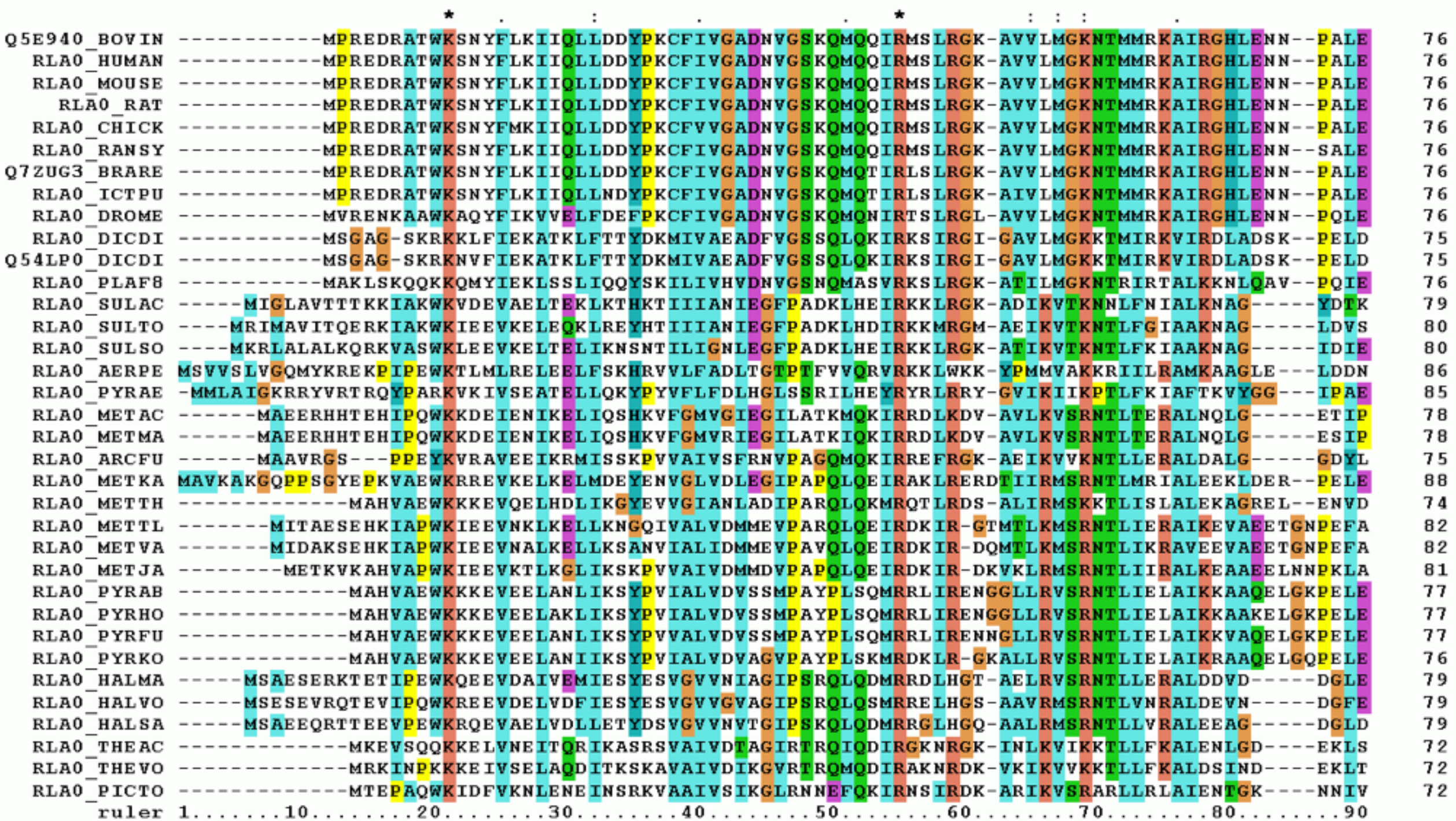
A: Biological Phylogenetics

- Multi-alignment methods actively developed since the 1980s
- Start with pairwise alignment, then build up multiple alignment starting from most similar pairs (using "guide tree")

Multi-alignment of nucleotides (4-letter alphabet)



Multi-alignment of amino-acids (20-letter alphabet)



B: Historical Linguistics

- Multi-alignment is just a fancy name for sound correspondences
- Each sound correspondence is "aligned" in a column, possibly adding empty cells
- It is a useful and consistent way to represent comparative data (both between languages or dialects)
- First automatic multi-alignments by Kondrak (2009) and Prokić (2009). Currently furthest developed by Mattis List (LingPy)

Language	IDS	meaning	alignment					
Pilagá	15.810/15.820	heavy/light	d	e	s	a	l	i
Toba	15.810	heavy	d	e	s	a	l^y	i
Mocoví	15.810/15.820	heavy/light	r	e	s	a	l^y	i

Pilagá	9.440	build	n	ʔ	o	ʁ	o	–	s	e	g	e	m
Toba	9.440	build	n	ʔ	o	ʁ	o	o	š	i	g	e	m
Mocoví	9.440	build	n	o	ʔ	ʁ	o	n	š	i	g	i	m

Steiner, Lydia, Peter F Stadler, and Michael Cysouw. "A Pipeline for Computational Historical Linguistics." *Language Dynamics and Change* 1, no. 1 (2011): 89-127 doi:10.1163/221058211X570358.

LOCATION	WORD
Aachen	a̟ːph
Adorf	a̟ːbʱə
Ahrbergen	o̝→o̟phə̝
Albersloh	ɑːpʱə̟̈
Allna	ɑɸh
Altenberg	ʌfɛ̟̈
Altentrüdin	a̟f
Altlandsberg	ɑʾfəʾ
Altwarp	oːph
Astfeld	ɒˤːpʰə
Atzendorf	afɛ̟
Ballhausen	ʌʾfə
Bardenfleth	ɔ̟ːp͡ɸ
Barssel	ɒ̟ːpʱə
Bempflingen	afː
Bennin	o̝pʰ
Billingsbach	a̟f
Bockelwitz	ʌ̟və̟̈
Bonn	ɑːpʾ
Borstendorf	ɤ̝fː
Breddin	ɒːph
Brelingen	ɑ̟f͡βə
Bremscheid	ɒˤːpḫə̟
…	…

A	FF	E
a̟ː	ph	-
ɑ̟ː	bʱ	ə
o̝→o̟	ph	ə̟
ɑː	pʱ̈	ə̟̈
ɑ	ɸh	-
ʌ	f	ɛ̟̈
a̟	f	-
ɑʾ	f	əʾ
oː	ph	-
ɒˤː	pʰ	ə
a	f	ɛ̟
ʌʾ	f	ə
ɔ̟ː	p͡ɸ	-
ɒ̟ː	pʱ	ə
a	fː	-
o̝	pʰ	-
a̟	f	-
ʌ̟	v	ə̟̈
ɑː	pʾ	-
ɤ̝	fː	-
ɒː	ph	-
ɑ̟	f͡β̞	ə
ɒˤː	pḫ	ə̟
…	…	…

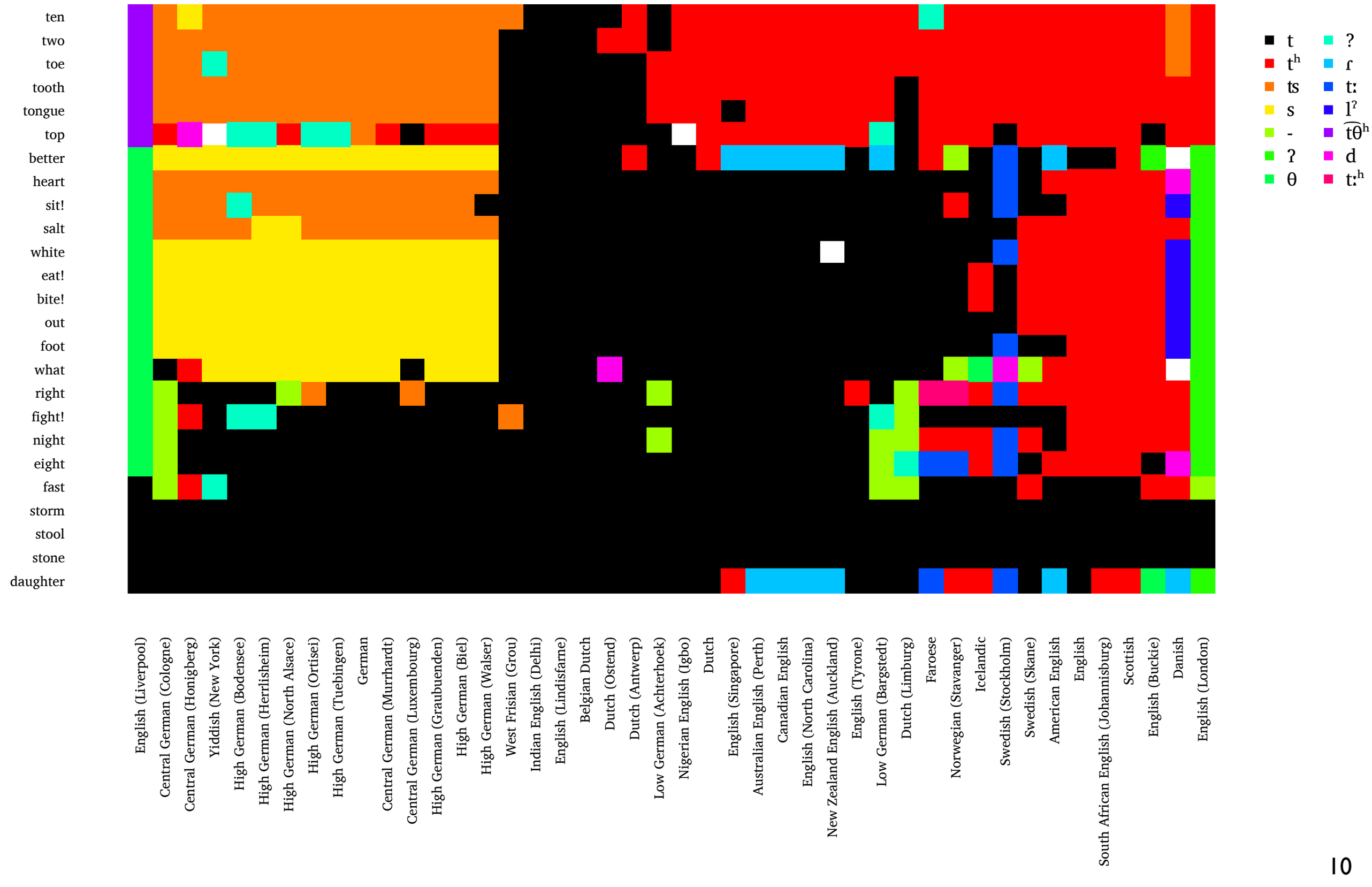
Affe (German Dialect data)

tongue

American English	tʰ	ʌ	ŋ	-	-
Canadian English	tʰ	ʌ	ŋ	-	-
Central German (Cologne)	ʦ	ʊ	ŋ	-	-
Central German (Honigberg)	ʦ	aː	-	-	-
Central German (Luxembourg)	ʦ	ɔ	ŋ	-	-
Central German (Murrhardt)	ʦ	ʊ	ŋ	-	-
Danish	tʰ	ʊ	ŋ	-	ə
Dutch (Antwerp)	t	ʌ	ŋ	-	-
Belgian Dutch	t	ʊ	ŋ	-	-
Dutch (Limburg)	t	ʊ	ŋ	-	-
Dutch (Ostend)	t	ʊ	ŋ	-	ə
Dutch	tʰ	ʌ	ŋ	-	-
New Zealand English (Auckland)	tʰ	ɐ	ŋ	-	-
English (Buckie)	tʰ	ʌ	ŋ	-	-
Indian English (Delhi)	t	ɐ	ŋ	-	-
Nigerian English (Igbo)	tʰ	ʌ	ŋ	ɡ	-
South African English (Johannisburg)	tʰ	ɜ	ŋ	-	-
English (Lindisfarne)	t	ɔ	ŋ	-	-
English (Liverpool)	t͡θʰ	ʊ	ŋ	ɡ	-
English (London)	tʰ	ɐ	ŋ	-	-
English (North Carolina)	tʰ	ɜː	ŋ	-	-
Australian English (Perth)	tʰ	ɐ	ŋ	-	-
English (Singapore)	t	ɑ	ŋ	-	-
English	tʰ	ɐ	ŋ	-	-
English (Tyrone)	tʰ	ɔ	ŋ	-	-
Faroese	tʰ	ʊ	ŋ	k	a
German	ʦ	ʊ	ŋ	-	ə
High German (North Alsace)	ʦ	ʊ	ŋ	-	-
High German (Biel)	ʦ	ʊ	ŋ	-	ə
High German (Bodensee)	ʦ	ʊ	ŋ	-	ə
High German (Graubuenden)	ʦ	ʊ	ŋ	ɡ	ɐ
High German (Herrlisheim)	ʦ	ʊ	ŋ	-	-
High German (Ortisei)	ʦ	ʊ	ŋ	ɡ	ɛ
High German (Tuebingen)	ʦ	u	ŋ	ɡ	-
High German (Walser)	ʦ	ʊ	ŋ	ɡ	ə
Icelandic	tʰ	ʊ	ŋ	k	a
Low German (Achterhoek)	tʰ	ʊ	ŋ	-	ə
Low German (Bargstedt)	tʰ	ʊ	ŋ	-	-
Norwegian (Stavanger)	tʰ	ʊ	ŋ	-	ə
Scottish	tʰ	ʌ	ŋ	-	-
Swedish (Skane)	tʰ	Øʏ	ŋ	j	ɐ
Swedish (Stockholm)	tʰ	ʊ	ŋ	-	ə
West Frisian (Grou)	t	ɔ	ŋ	-	ə
Yiddish (New York)	ʦ	u	ŋ	ɡ	-

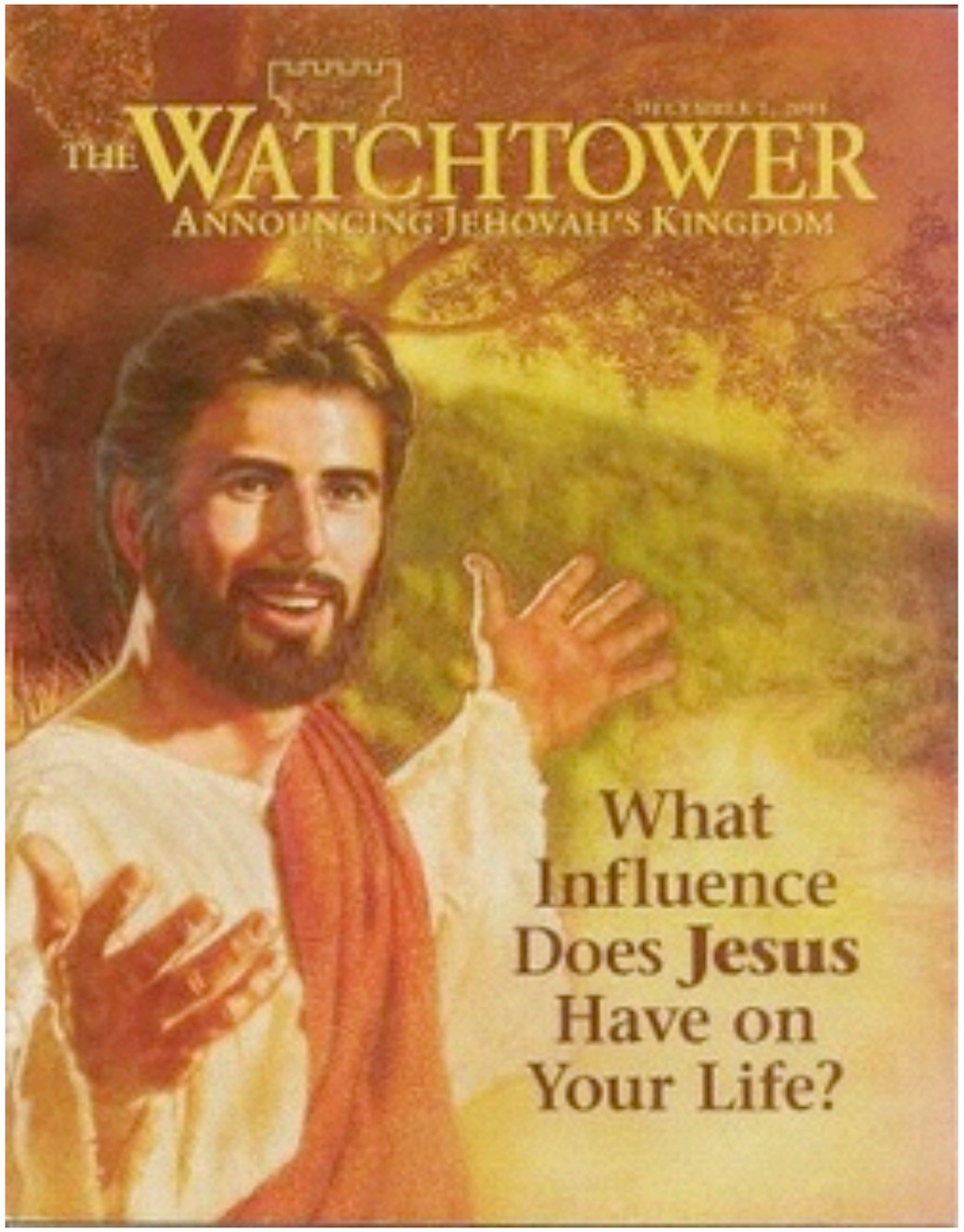
(Data from Paul Haggerty, aligned by the LingPy library maintained by Mattis List)

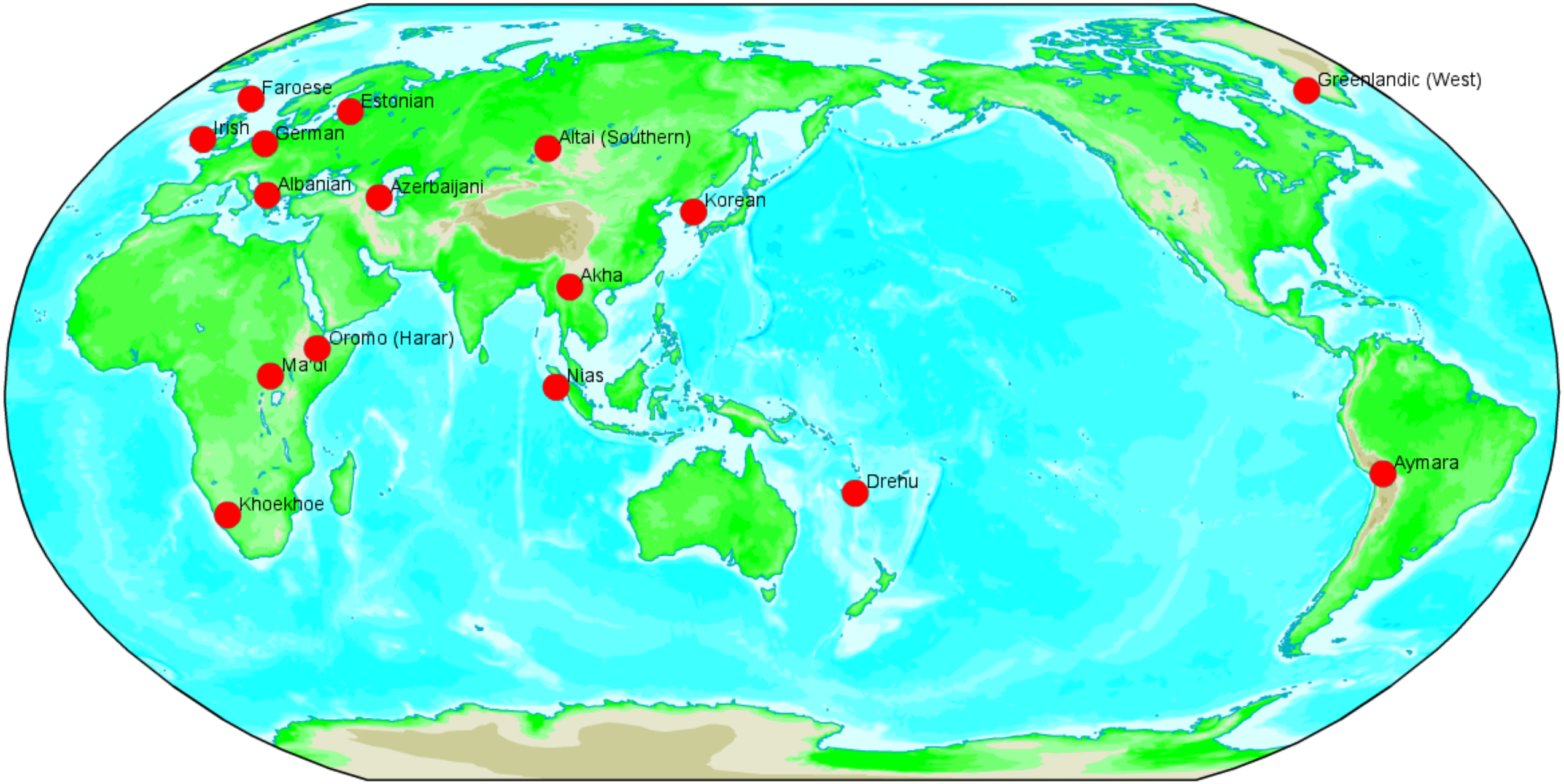
Correspondences with high frequency of [t]



C: Functional Language Comparison

- Comparing morphosyntactic patterns in languages all over the world
- "Language Typology"
- Traditional problem how to compare disparate languages.
- Solution: Multi-alignment


THE WATCHTOWER
ANNOUNCING JEHOVAH'S KINGDOM
What Influence Does Jesus Have on Your Life?


Faroese
Estonian
Irish
German
Albanian
Azerbaijani
Altai (Southern)
Korean
Akha
Oromo (Harar)
Ma'di
Nias
Khoekhoe
Drehu
Greenlandic (West)
Aymara

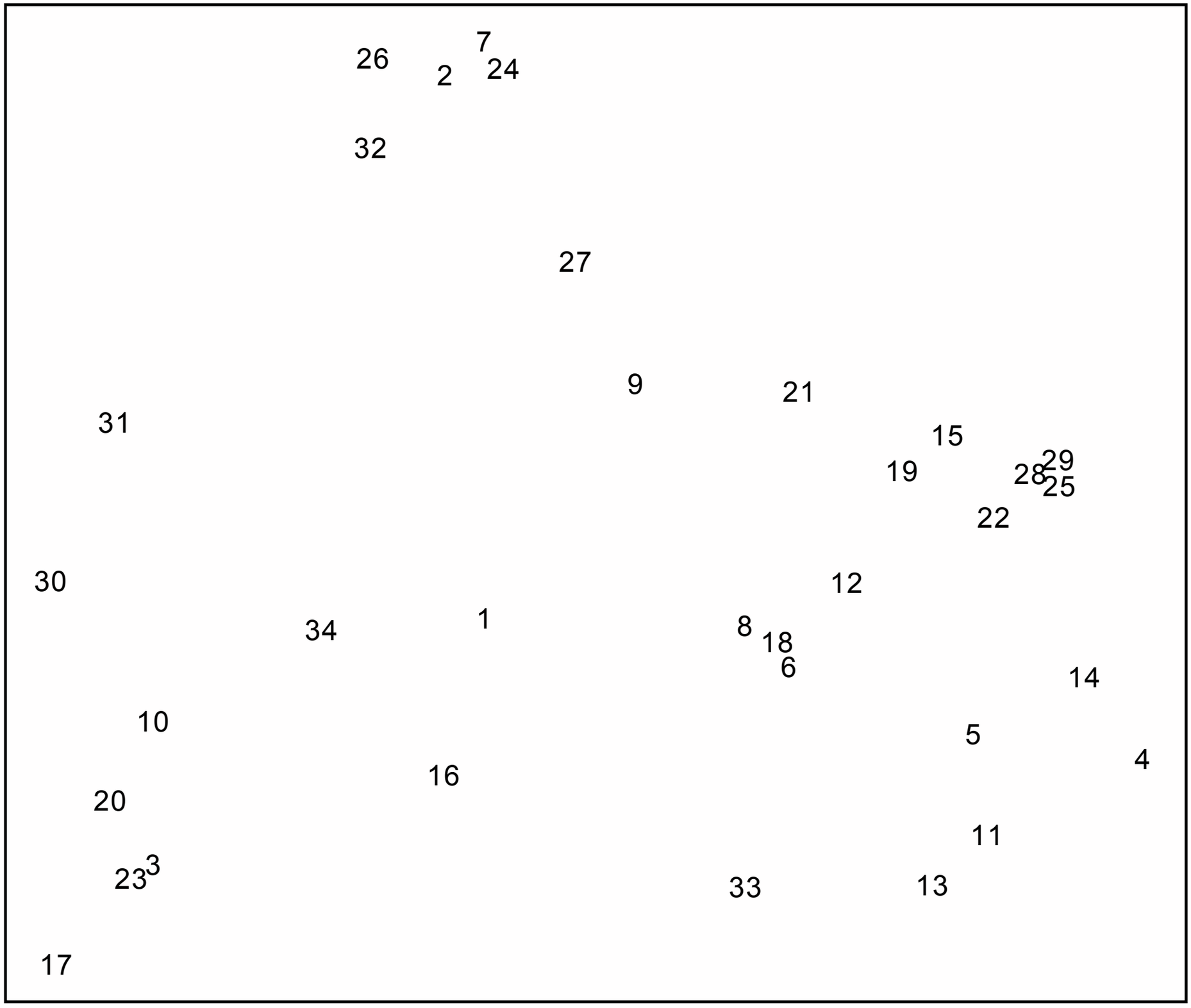
1 What important information is contained in the Bible?
2 Who is the Bible's author?
3 Why should you study the Bible?
4 The Bible is a precious gift from God.
5 The Bible alone tells us what we must do to please God.
6 The Bible was written by some 40 different men over a period of 1,600 years, beginning 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.

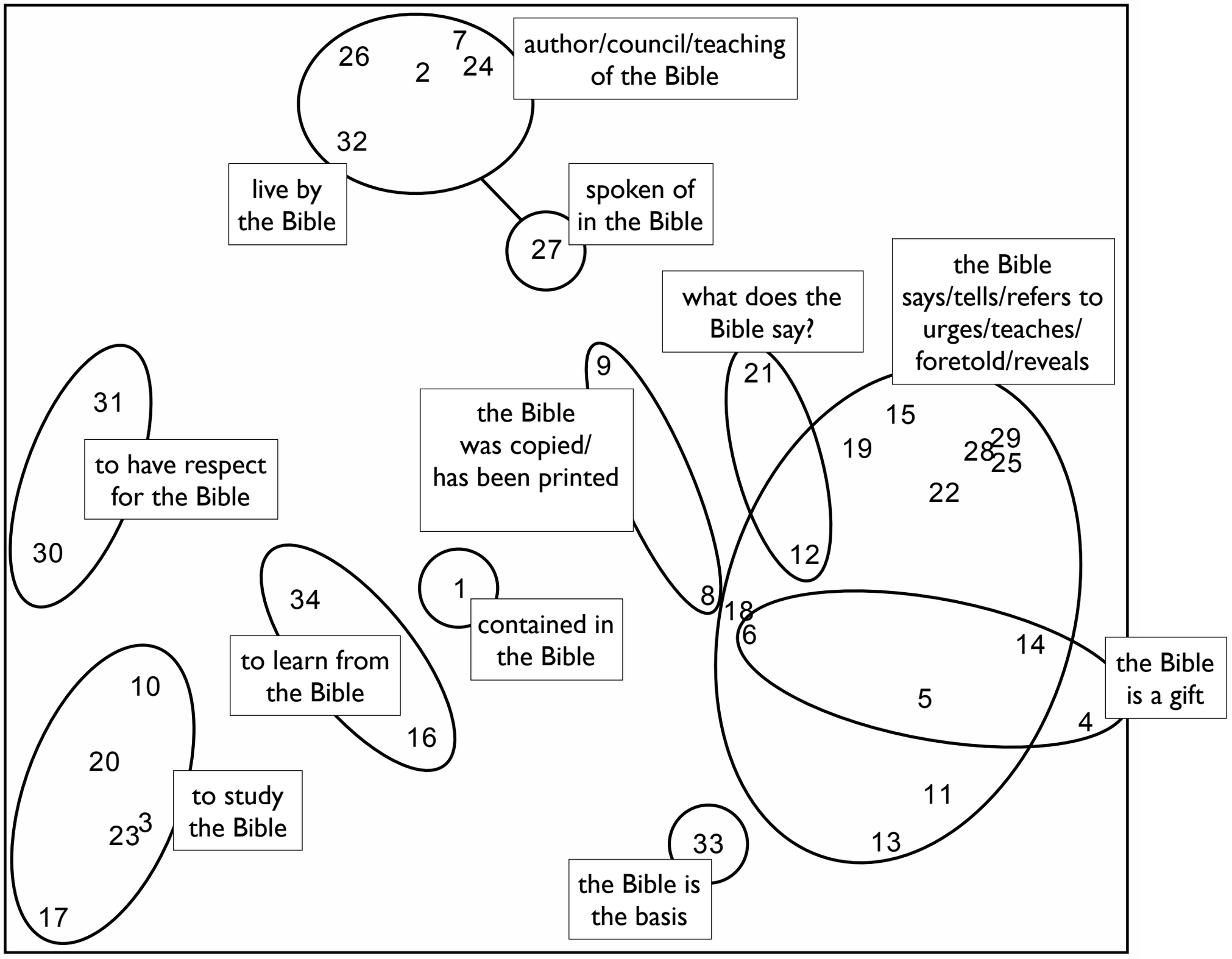
Cysouw, Michael. 2014. Inducing semantic roles. In Silvia Luraghi & Heiko Narrog (eds.), *Perspectives on Semantic Roles*, Amsterdam: Benjamins.

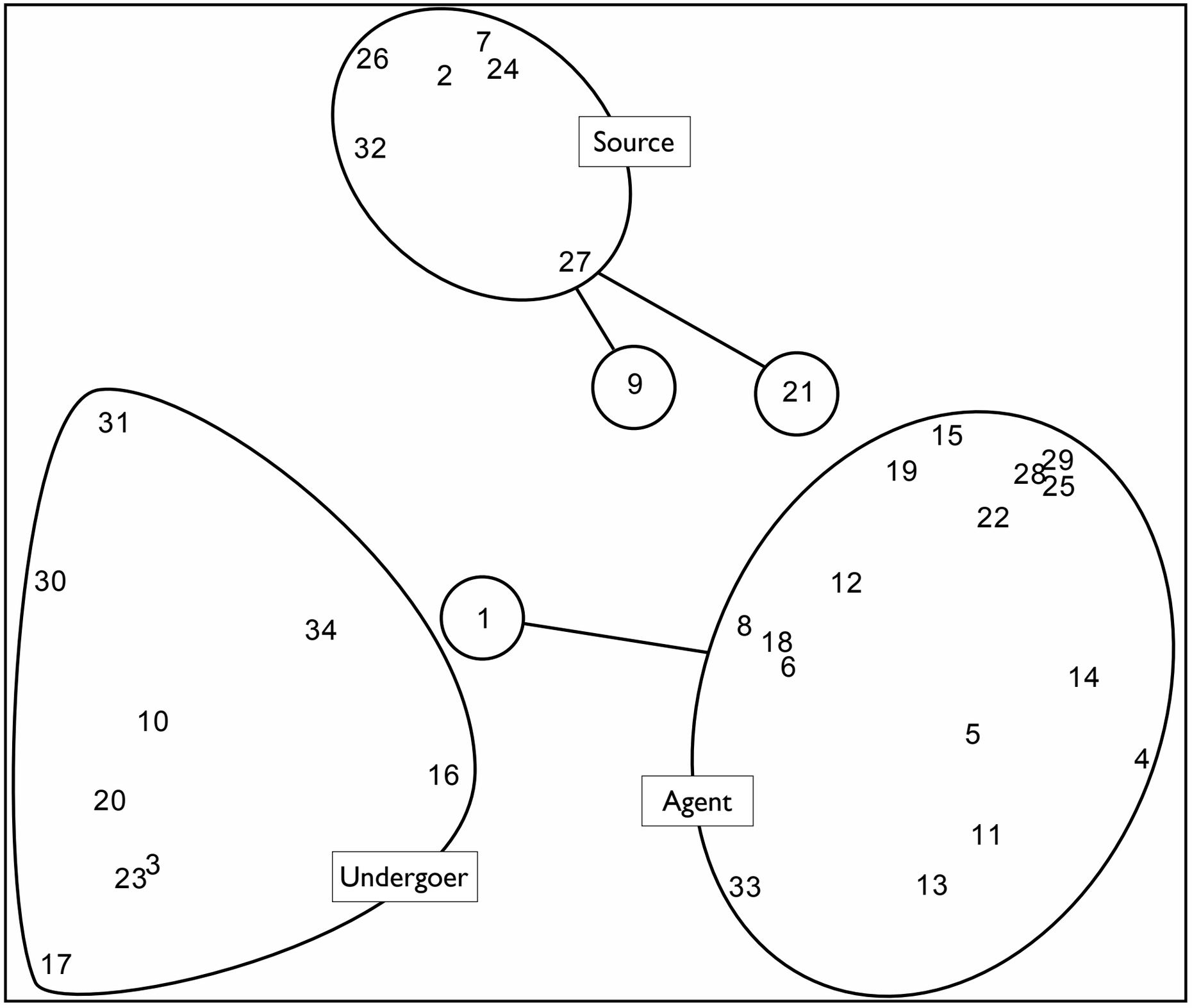
Albanian	Faroese	Estonian	Greenlandic
bibla Nominative	*biblian* Nominative	*þiibel* Nominative	*biibiliþ* Ergative
biblën Accusative	*bibliuna* Accusative	*þiiblit* Partitive	*biibli* Absolutive
biblës Genitive/Dative	*bibliunnar* Genitive	*þiibli* Genitive	*biibilmik* Instrumental
…	*bibliuni* Dative	*þiiblis* Inessive	*biibilmi* Locative
	…	*þiiblist* Elative	…
		…	

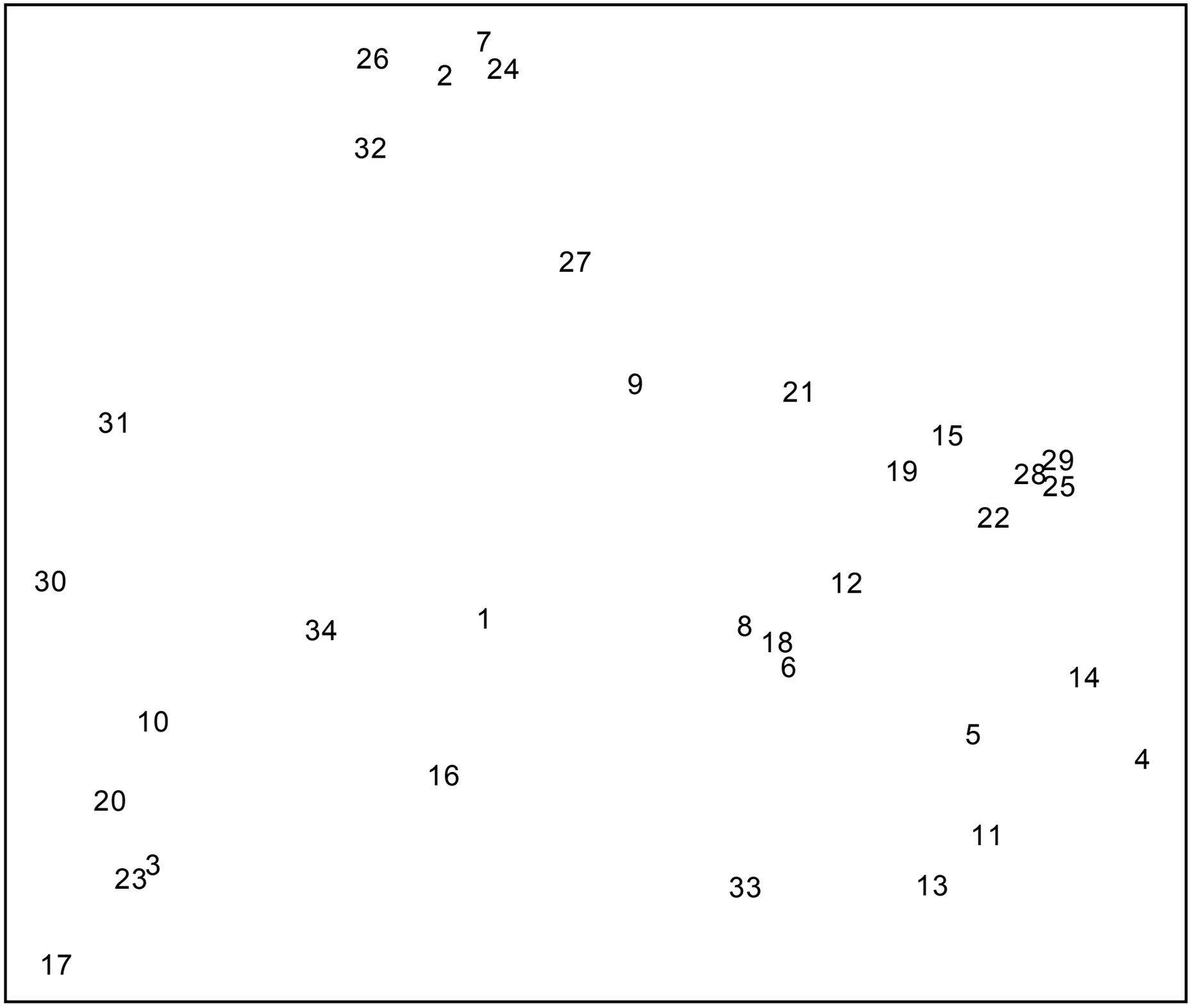
Context	Albanian	Faroese	Estonian	Greenlandic
1	bibla	bíbliuni	piibel	biibili
2	biblës	bíbliunnar	piibli	biibilimik
3	biblën	bíbliuna	piiblit	biibili
4	bibla	bíblian	piibel	biibili
5	bibla	bíblian	piibel	biibilip
6	bibla	bíbliuna	piibli	biibili
7	biblës	bíbliunnar	piibli	biibilimut
8	bibla	bíblian	piiblit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

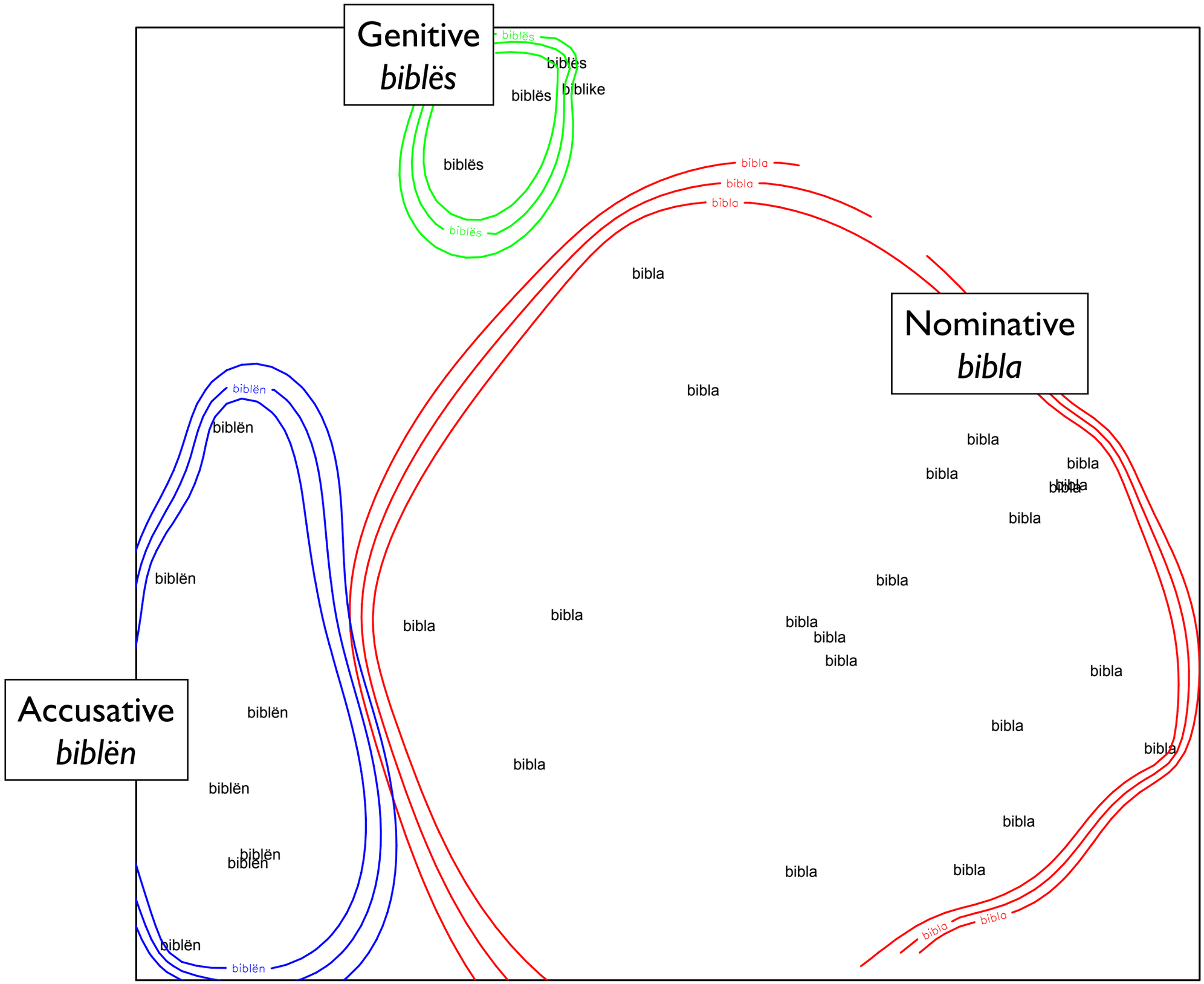

26
7
2
24
32
27
9
21
31
15
19
29
28
25
22
30
12
1
34
8
18
6
14
10
5
4
16
20
11
3
23
33
13
17


author/council/teaching of the Bible
26
7
24
2
32
live by the Bible
spoken of in the Bible
27
what does the Bible say?
the Bible says/tells/refers to urges/teaches/ foretold/reveals
9
21
31
15
19
29
28
25
22
to have respect for the Bible
the Bible was copied/ has been printed
30
12
1
34
8
18
6
14
contained in the Bible
to learn from the Bible
the Bible is a gift
10
5
16
4
20
11
to study the Bible
23
3
13
33
the Bible is the basis
17

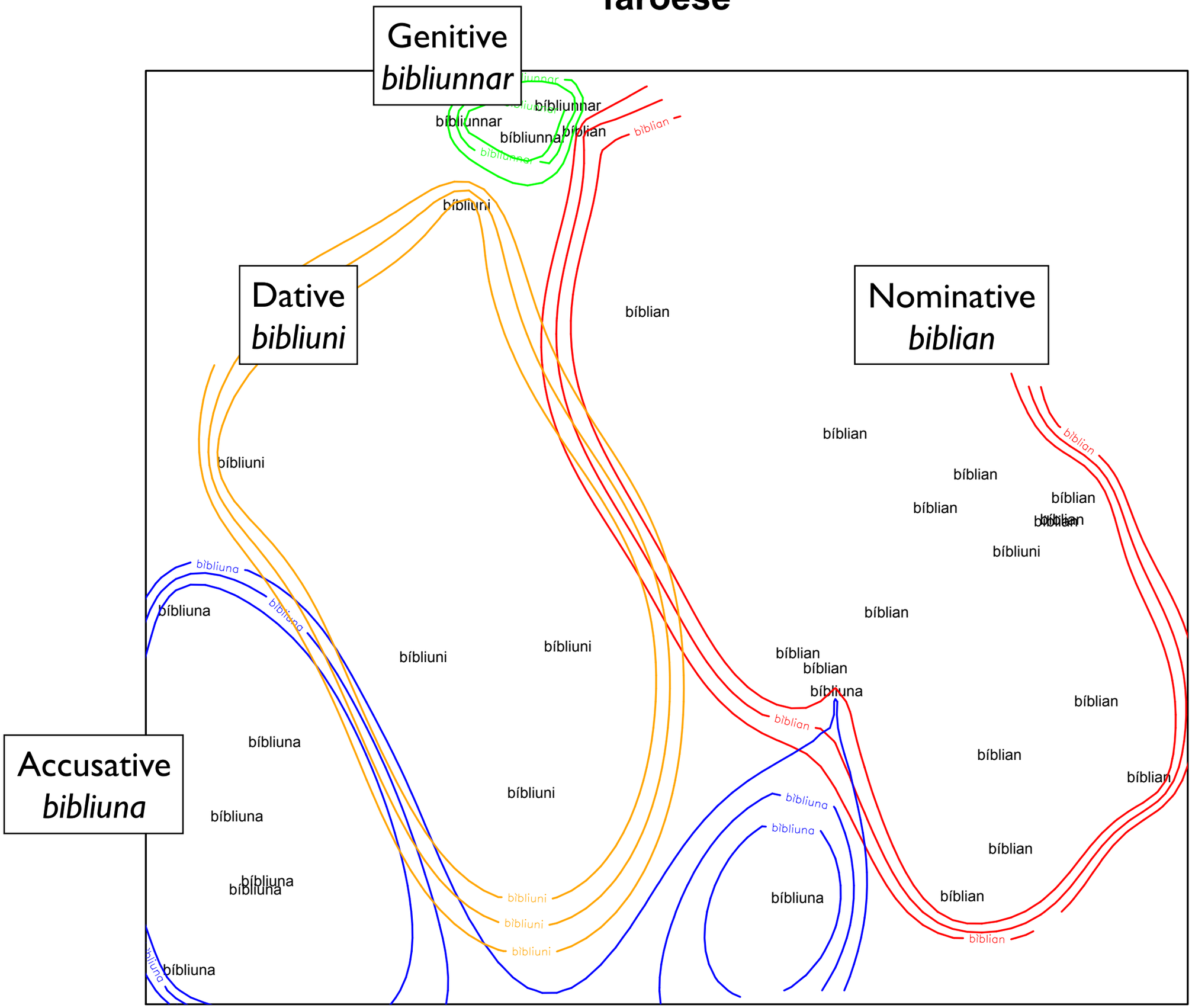

7
26
2
24
32
Source
27
9
21
31
15
19
29
28
25
22
30
12
1
34
8
18
6
14
10
5
4
16
20
Agent
11
23
3
Undergoer
33
13
17


26
7
2
24
32
27
9
21
31
15
19
29
28
25
22
30
12
1
34
8
18
6
14
10
5
4
16
20
11
3
23
33
13
17

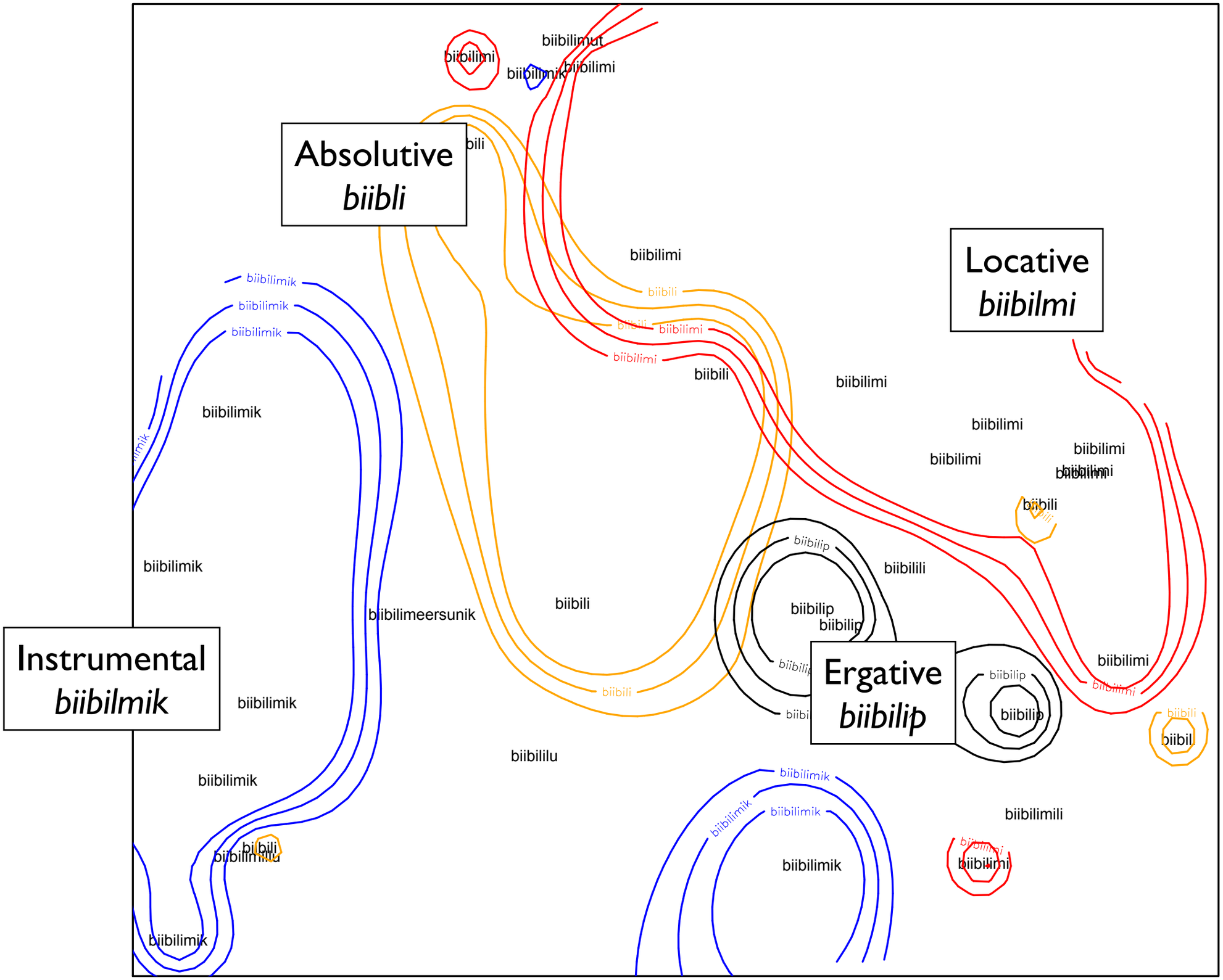
albanian


Genitive
biblës
biblës
biblës
biblike
biblës
biblës
biblës
bibla
bibla
bibla
bibla
Nominative
bibla
bibla
bibla
bibla
bibla
bibla
bibla
bibla
bibla
bibla
bibla
bibla
bibla
bibla
bibla
bibla
bibla
bibla
bibla
bibla
bibla
bibla
bibla
biblën
biblën
biblën
Accusative
biblën
biblën
biblën
biblën
biblën
biblën
biblën

faroese


Genitive
bibliunnar
Dative
bibliuni
Nominative
biblian
Accusative
bibliuna
bíbliunnar
bíbliunnar
bíbliunnar
bíblian
bìblian
bìbliunnar
bìbliunnar
bíbliuni
bíblian
bíblian
bíbliuni
bíblian
bíblian
bíblian
bíblian
bíblian
bìblian
bíbliuni
bìbliuna
bìbliuna
bíbliuna
bíblian
bíbliuni
bíbliuni
bíblian
bíblian
bíbliuna
bìblian
bíblian
bíbliuna
bíblian
bíblian
bìbliuna
bìbliuna
bíbliuni
bíbliuna
bíblian
bíbliuna
bíbliuna
bíblian
bíbliuna
bìbliuni
bìbliuni
bìbliuni
bìblian
bìbliuna
bíbliuna

greenlandic



Multi-alignment

A. Biological Phylogenetics
   (alignment of nucleotides/amino-acids)

B. Historical Linguistics
   (alignment of sounds in words)

C. Functional Language Comparison
   (alignment of words in sentences)

Three of a kind?

- Basically the same problem:
  - substrings within strings have to be aligned
- However, there are differences
  - the information of the substrings differs strongly
  - the lengths of the strings differs strongly

Parts/Length Ratio
("informativity of the parts")

- DNA: order 10^{-3} (viz. ~ 4 / 1000s)
  (nucleotides vs. length of strings)
- Protein: order 10^{-1} (viz. ~ 20 / 100s)
  (amino acids vs. length of protein)
- Words: order 10^{+1} (viz. ~ 40 / 6)
  (sounds vs. length of word)
- Sentences: order 10^{+3} (viz. ~ 1000s / 10)
  (number of words vs. length of sentence

A different view of multi-alignment

- Multi-alignment is mostly seen as the problem of adding gaps in the right places
- Crossing alignments (“metathesis”) is seen as a nuisance / special case
- Now: words in sentences have massive crossing alignments.

A different view of multi alignment

- Consider multi-alignment as constrained partitioning ("flat clustering")
- Partitioning
  - take all elements in all species/languages, and put them into groups ("mulit-alignments")
- Constraints, e.g.
  - prefer groups across species/languages
  - ordering can be more or less strongly obeyed