

Denis Creissels and Fernando Zúñiga, with George Moroz

### **3 Languages examined or referred to in the present book**

#### **1 The maps**

Maps 1–8 were created by George Moroz with the help of the R (R Core Team 2023) package *lingtypology* (Moroz 2017).

#### **2 The classification**

Languages are listed alphabetically within each of the following five areas: North America, South America, Africa, Pacific, and Eurasia.

The genetic affiliation of the languages listed in this index is provided along the lines of the classification of languages adopted in WALS Online (Dryer and Haspelmath 2013), with only some minor modifications that are commented in footnotes. For languages that are not explicitly listed in WALS, the classification copies that of closely related languages that feature in WALS and for which it can be taken for granted that they belong to the same genus.

Apart from pidgins and creoles, for which no genetic affiliation is proposed in WALS, the general principle of this classification is that, as a rule, it does not mention the intermediate groupings that appear in other classifications, restricting to only two levels: genus and family. FAMILY refers to the highest level widely accepted by specialists, while GENUS is commented on in WALS Online (<https://wals.info/languoid/genealogy>) as follows:

The notion genus is explained in Dryer (1989). It is intended as a level of classification which is comparable across the world, so that a genus in one family is intended to be comparable in time depth to genera in other parts of the world. The choice of term is intended to match the general idea of genus in biological classification, where a genus is a set of species that are clearly closely related to each other (and where words in everyday language often correspond to genera rather than species). In the genealogical classification of languages, a genus is a group of languages whose relatedness is fairly obvious without systematic comparative analysis, and which even the most conservative “splitter” would accept. Genealogical groups deeper than a genus are often less obvious and in the absence of detailed comparative work are often not universally accepted. If there is evidence of time depth of groups, the genus would not have a time depth greater than 3500 or 4000 years. A genus may have a time depth much less than this, but if the time of the split of one group of languages from other languages in the family appears to be greater than 4000 years, then this constitutes a reason to say that this group of languages is a separate genus.

Here, as in WALS, levels of classification lower than that of genus are not taken into account. For example, the classification of Somali refers to the genus “Lowland East

Cushitic”, ignoring the lower-level groupings “Nuclear Lowland East Cushitic” and “Omo-Tana”. Similarly, Northern Paiute is classified as a member of the Northern Uto-Aztecan genus, ignoring the lower-level grouping “Numic”.

An intermediate level between genus and family, that of subfamily, is only provided when the genetic affiliation of the language in question is more commonly characterized with reference to the subfamily than to the genus to which it belongs. For example, Baule is classified here as “Baule, Tano, Kwa, Niger-Congo”. The reason is that, in the literature, whatever the validity of the Kwa subfamily within the Niger-Congo family may be, “Kwa”, rather than “Tano”, is the label commonly used to characterize the position of Baule within the Niger-Congo family. More generally, the subfamilies of Niger-Congo (Atlantic, Kwa, Benue-Congo), Afroasiatic (Cushitic) and Eastern Sudanic (Nilotic) that do not meet the definition of genera (and whose very validity as genetic groupings is sometimes questionable) but are commonly mentioned in the classification of Niger-Congo, Afroasiatic, and Eastern Sudanic languages have been systematically added.

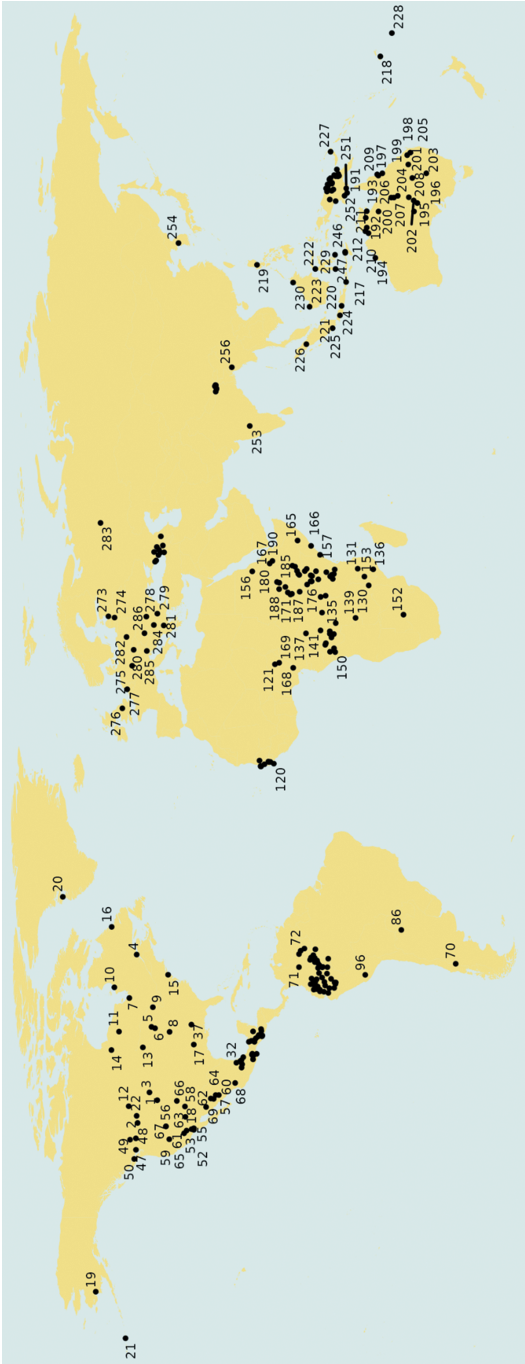
This said, some particular cases must be considered. First, some groups of lects commonly treated as constituting a single (macro-)language, like the dialectal varieties of Wolof, are commonly regarded as isolates within a genetic grouping—in the case of Wolof, North Atlantic = Atlantic = Niger-Congo—whose time depth exceeds that admitted for genera. In such cases, the name of the language is repeated as that of the genus that coincides with the (macro-)language in question. For example, Wolof is classified here as “Wolof (language), Wolof (genus), Atlantic (subfamily), Niger-Congo (family)”.

A second particular case is that of languages belonging to a genus that is not commonly regarded as included into a higher-level family (and consequently, also has the status of family in the sense given here to this term). In such cases, the name of the language is followed by a single label referring to the genus. Mayan languages are a case in point. For example, K’iche’ is classified here as “K’iche’ (language), Mayan (genus and family)”.

The third particular case concerns language with no known relatives, such as Cofán (Colombia). In such cases, the name of the language or language variety is followed by a single label that simply reproduces the name of the language, for example “Cofán (language), Cofán (genus and family)”.

## References

- Dryer, Matthew S. 1989. Large linguistic areas and language sampling. *Studies in Language* 13. 257–292.
- Dryer, Matthew S. & Martin Haspelmath (eds.). 2013. *WALS Online* (v2020.3) [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.7385533> (Available online at <https://wals.info>, Accessed on 2023–03–08.)
- Moroz, George. 2017. *lingtypology*: easy mapping for Linguistic Typology. <<https://CRAN.R-project.org/package=lingtypology>>.
- R Core Team. 2023. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.



Map 1: World.

NORTH AMERICA		
language	genetic affiliation	number
Acjachemem	Northern Uto-Aztecan, Uto-Aztecan	52
Arapaho	Algonquian, Algic	1
Blackfoot	Algonquian, Algic	2
Caddo	Caddo, Caddoan	17
Cahuilla	Northern Uto-Aztecan, Uto-Aztecan	53
Central Alaskan Yup'ik	Eskimo, Eskimo-Aleut	19
Chalcatongo Mixtec	Mixtec, Oto-Manguean	43
Cheyenne	Algonquian, Algic	3
Chickasaw	Muskogean	37
Chontal	Mayan	24
Classical Nahuatl	Aztecan, Uto-Aztecan	54
Cupeño	Northern Uto-Aztecan, Uto-Aztecan	55
Ch'ol	Mayan	23
Ch'orti'	Mayan	25
Eastern Highlands Otomi	Otomian, Oto-Manguean	38
Gosiute Shoshoni	Northern Uto-Aztecan, Uto-Aztecan	56
Guarijío	Tarahumaran, Uto-Aztecan	57
Hopi	Northern Uto-Aztecan, Uto-Aztecan	58
Hul'q'umi'num	Central Salish, Saolishan	47
Innu	Algonquian, Algic	4
Itzaj	Mayan	26
Ixil	Mayan	27
Kalaallisut	Eskimo, Eskimo-Aleut	20
Kaqchikel	Mayan	28
K'iche'	Mayan	29
Ktunaxa	Kutenai	22
Menominee	Algonquian, Algic	5
Meskwaki	Algonquian, Algic	6
Moose Cree	Algonquian, Algic	7
Myaamia	Algonquian, Algic	8
Nishnaabemwin	Algonquian, Algic	9
Northern East Cree	Algonquian, Algic	10 <sup>1</sup>
Northern Paiute	Northern Uto-Aztecan, Uto-Aztecan	59
Northern Tepehuan	Tepiman, Uto-Aztecan	60
Northern Zapotec	Zapotecan, Oto-Manguean	39
Ojicree	Algonquian, Algic	11
Okanagan	Interior Salish, Salishan	48
Otomi	Otomian, Oto-Manguean	40
Pahka'anil	Northern Uto-Aztecan, Uto-Aztecan	61
Pima Bajo	Tepiman, Uto-Aztecan	62
Plains Cree	Algonquian, Algic	12
Poqomam	Mayan	30
Q'anjob'al	Mayan	31
Quiegolani Zapotec	Zapotecan, Oto-Manguean	44
San Felipe Otomi	Otomian, Oto-Manguean	41

<sup>1</sup> Southern East Cree (mentioned in Chapter 18) does not appear on the map but is spoken south of Northern East Cree.



(continued)

NORTH AMERICA		
language	genetic affiliation	number
San Lucas Quiaviní Zapotec	Zapotecan, Oto-Manguean	42
Serrano	Northern Uto-Aztecan, Uto-Aztecan	63
Shuswap	Interior Salish, Salishan	49
Sliammon	Central Salish, Salishan	50
Southeastern Huastec	Mayan	32
Southwestern Ojibwe	Algonquian, Algic	13
Swampy Cree	Algonquian, Algic	14
Tamazulápam Mixe	Mixe-Zoque	45
Tarahumara	Tarahumaran, Uto-Aztecan	64
Teotitlán Zapotec	Zapotecan, Oto-Manguean	46
Tojolabal	Mayan	33
Tselal	Mayan	34
Tsotsil	Mayan	35
Tümpisa Shoshoni	Northern Uto-Aztecan, Uto-Aztecan	65
Tz'utujil	Mayan	36
Unami	Algonquian, Algic	15
Unangan	Aleut, Eskimo-Aleut	21
Upper Necaxa Totonac	Totonacan	51
Ute	Northern Uto-Aztecan, Uto-Aztecan	66
Walapai	Yuman, Hokan	18
Western Naskapi	Algonquian, Algic	16
Western Shoshoni	Northern Uto-Aztecan, Uto-Aztecan	67
Wixárika / Huichol	Corachol, Uto-Aztecan	68
Yaqui	Cahita, Uto-Aztecan	69



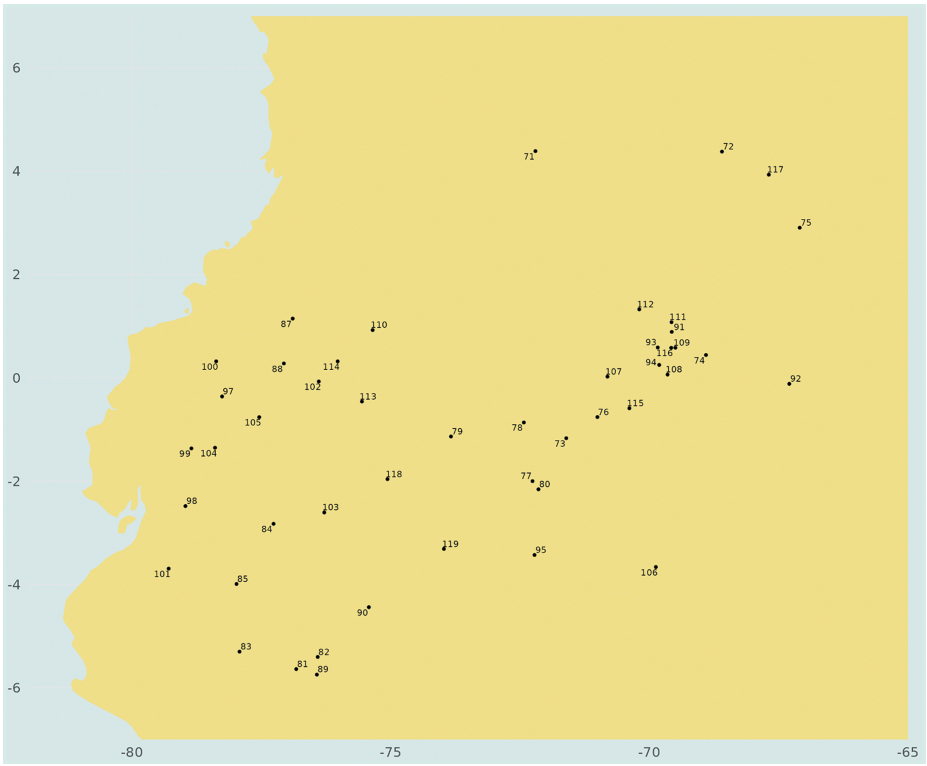
Map 2: Southern North America.

SOUTH AMERICA		
language	genetic affiliation	number
Achagua	Japura-Colombia, Arawakan	71
Aguaruna	Jivaroan	83
Arabela	Zaparoan	116
Ayacucho Quechua	Quechuan	96
Barasano	Tucanoan	107
Calderón Highland Quichua	Quechuan	97
Camsá	Camsá	87
Cañar Highland Quichua	Quechuan	98
Carapana	Tucanoan	108
Chimborazo Highland Quichua	Quechuan	99
Cofán	Cofán	88
Dâw	Nadahup	92
Desano	Tucanoan	109
Hup	Nadahup	93
Imbabura Highland Quichua	Quechuan	100
Iquito	Zaparoan	119
Kakua	Cacua-Nukak	91
Koreguaje	Tucanoan	110
Kotiria	Tucanoan	111
Kubeo	Tucanoan	112
Loja Highland Quichua	Quechuan	101
Mapudungun	Araucanian	70
Miraña	Boran	77
Muinane	Witoto, Witotoan	78
Muniche	Muniche	89
Murui	Witoto, Witotoan	79
Napo Lowland Quechua	Quechuan	102
Ocaina	Witoto, Witotoan	80
Pastaza Quechua	Quechuan	103
Piapoco	Japura-Colombia, Arawakan	72
Puinave	Puinave	117
Resígaro	Japura-Colombia, Arawakan	73
Salasaca Highland Quichua	Quechuan	104
Secoya	Tucanoan	113
Shawi	Cahuapanan	81 <sup>2</sup>
Shiwiar	Jivaroan	84
Shiwilu	Cahuapanan	82
Siona	Tucanoan	114
Tanimuka	Tucanoan	115
Tariana	Japura-Colombian, Arawakan	74
Tena Lowland Quichua	Quechuan	105
Tikuna	Tikuna	106
Toba / Qom	Qom, Guaycuruan	86

<sup>2</sup> San Martín Quechua (mentioned in Chapter 11) does not appear on the map but is spoken south of Shawi.

(continued)

SOUTH AMERICA		
language	genetic affiliation	number
Tukano	Tucanoan	116
Urarina	Urarina	80
Wampis	Jivaroan	85
Warekena	Alto Orinoco, Arawakan	75
Yagua	Peba-Yaguan	95
Yuhup	Nadahup	94
Yukuna	Japura-Colombian, Arawakan	76



AFRICA		
language	genetic affiliation	number
Agar Dinka	Western Nilotic, Nilotic, Eastern Sudanic	171
Akie	Southern Nilotic, Nilotic, Eastern Sudanic	172
Alagwa	Southern Cushitic, Cushitic, Afroasiatic	154
Amharic	Semitic, Afroasiatic	191
Arbore	Lowland East Cushitic, Cushitic, Afroasiatic	155
Asimjeeg Datooga	Southern Nilotic, Nilotic, Eastern Sudanic	173
Ateso	Eastern Nilotic, Nilotic, Eastern Sudanic	174
Barbayiiga Datooga	Southern Nilotic, Nilotic, Eastern Sudanic	175
Bari	Eastern Nilotic, Nilotic, Eastern Sudanic	176
Baule	Tano, Kwa, Niger-Congo	169
Beja	Beja, Cushitic, Afroasiatic	156
Bijogo	Bijogo, Atlantic, Niger-Congo	120
Boni	Lowland East Cushitic, Cushitic, Afroasiatic	157
Burunge	Southern Cushitic, Cushitic, Afroasiatic	158
Cherang'any	Southern Nilotic, Nilotic, Eastern Sudanic	177
Chewa	Bantu, Benue-Congo, Niger-Congo	130
Chingoni	Bantu, Benue-Congo, Niger-Congo	131
Dhaasanac	Lowland East Cushitic, Cushitic, Afroasiatic	159
Dholuo	Western Nilotic, Nilotic, Eastern Sudanic	178
Ding	Bantu, Benue-Congo, Niger-Congo	132
Elmolo	Lowland East Cushitic, Cushitic, Afroasiatic	160
Gedee	Highland East Cushitic, Cushitic, Afroasiatic	161
Gisamjanga Datooga	Southern Nilotic, Nilotic, Eastern Sudanic	179
Gombe Fula	Fula-Serer, Atlantic, Niger-Congo <sup>3</sup>	121
Iraqw	Southern Cushitic, Cushitic, Afroasiatic	162
Jóola Fóoñi	Joola, Atlantic, Niger-Congo	122
Jumjum	Western Nilotic, Nilotic, Eastern Sudanic	180
Kikuyu	Bantu, Benue-Congo, Niger-Congo	133
Kongo ya Leta	Bantu, Benue-Congo, Niger-Congo	134
Laalaa	Cangin, Atlantic, Niger-Congo	123
Lango	Western Nilotic, Nilotic, Eastern Sudanic	181
Lengola	Bantu, Benue-Congo, Niger-Congo	135
Lomwe	Bantu, Benue-Congo, Niger-Congo	136
Londo	Bantu, Benue-Congo, Niger-Congo	137
Luba-Kasai	Bantu, Benue-Congo, Niger-Congo	138
Lunda	Bantu, Benue-Congo, Niger-Congo	139
Maasai	Eastern Nilotic, Nilotic, Eastern Sudanic	182
Mabaan	Western Nilotic, Nilotic, Eastern Sudanic	183
Mandinka	Central Mande, Mande <sup>4</sup>	170
Mankanya	Manjaku-Mankanya, Atlantic, Niger-Congo <sup>5</sup>	124

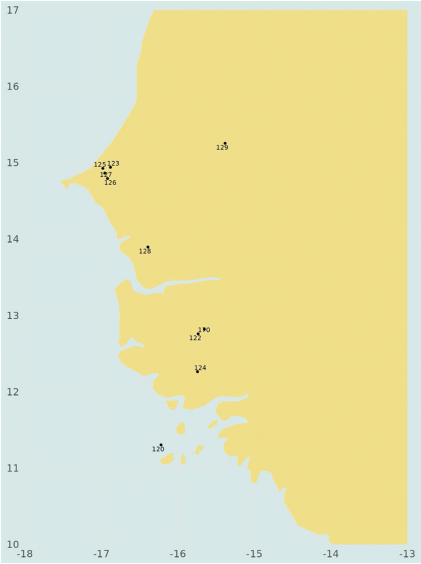
<sup>3</sup> *Peul* is the French name of the language called *Fula* in English.

<sup>4</sup> In the Mande family, West Mande does not meet the definition of a genus, and is best regarded as a subfamily.

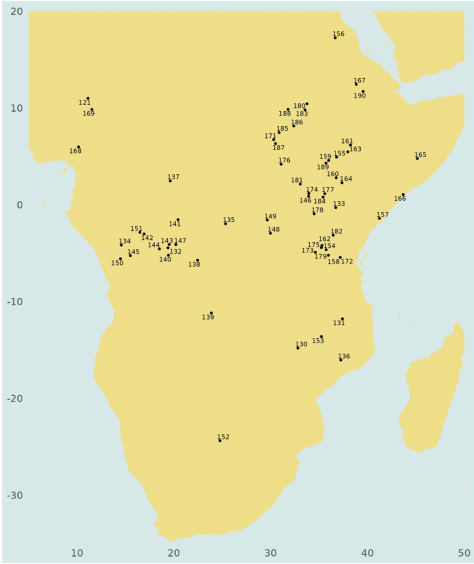
<sup>5</sup> Manjaku and Mankanya belong to the same genus, but there is no reason to select “Manjaku” rather than “Mankanya” as a label for the genus that includes Manjaku and Mankanya.

(continued)

AFRICA		
language	genetic affiliation	number
Mankon	Wide Grassfields, Benue-Congo, Niger-Congo	168
Mbuun	Bantu, Benue-Congo, Niger-Congo	140
Mongo-Nkundo	Bantu, Benue-Congo, Niger-Congo	141
Nandi	Eastern Nilotic, Nilotic, Eastern Sudanic	184
North Boma	Bantu, Benue-Congo, Niger-Congo	142
Ndut	Cangin, Atlantic, Niger-Congo	125
Ngwi	Bantu, Benue-Congo, Niger-Congo	143
Noon	Cangin, Atlantic, Niger-Congo	126
Northern Luo	Western Nilotic, Nilotic, Eastern Sudanic	185
Nsong	Bantu, Benue-Congo, Niger-Congo	144
Ntandu	Bantu, Benue-Congo, Niger-Congo	145
Nuer	Western Nilotic, Nilotic, Eastern Sudanic	186
Nyole	Bantu, Benue-Congo, Niger-Congo	146
Nzadi	Bantu, Benue-Congo, Niger-Congo	147
Oromo	Lowland East Cushitic, Cushitic, Afroasiatic	163
Palor	Cangin, Atlantic, Niger-Congo	127
Reel	Western Nilotic, Nilotic, Eastern Sudanic	187
Rendille	Lowland East Cushitic, Cushitic, Afroasiatic	164
Rundi	Bantu, Benue-Congo, Niger-Congo	148
Rwanda	Bantu, Benue-Congo, Niger-Congo	149
Seereer	Fula-Serer, Atlantic, Niger-Congo	128
Shilluk	Western Nilotic, Nilotic, Eastern Sudanic	188
Sikongo	Bantu, Benue-Congo, Niger-Congo	150
Somali	Lowland East Cushitic, Cushitic, Afroasiatic	165
Tiene	Bantu, Benue-Congo, Niger-Congo	151
Turkana	Eastern Nilotic, Nilotic, Eastern Sudanic	189
Tswana	Bantu, Benue-Congo, Niger-Congo	152
Tunni	Lowland East Cushitic, Cushitic, Afroasiatic	166
Wolof	Wolof, Atlantic, Niger-Congo	129
Xamtanga	Central Cushitic, Cushitic, Afroasiatic	167
Yao	Bantu, Benue-Congo, Niger-Congo	153



Map 4: Western Africa.

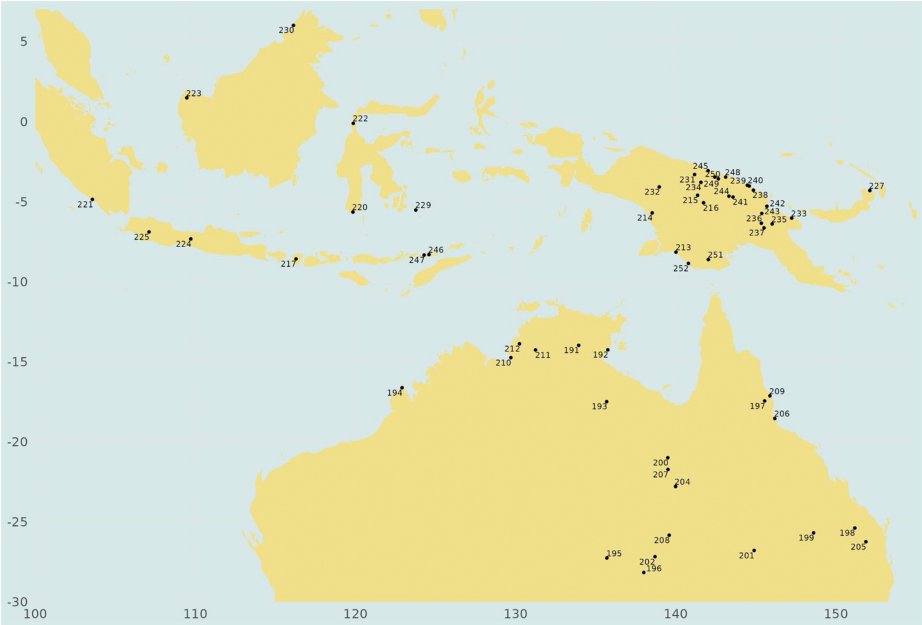


Map 5: Eastern and southern Africa.

PACIFIC		
language	genetic affiliation	number
Abui	Alor-Pantar, Greater West Bomberai	246
Alamblak	Sepik Hill, Sepik	244
Amele	Mabuso, Trans-New Guinea	242
Ampenan Sasak	Malayo-Sumbawan, Austronesian	217
Arabana-Wangkanguru	Central Pama-Nyungan, Pama-Nyungan	195
Aruamu / Mikarew	Ruboni, Ramu-Lower Sepik	238
Bardi	Nyulnyulan	194
Barupu	Warupu, Skou	245
Boumaa Fijian	Oceanic, Austronesian	218
Coastal Marind	Marind-Yaqai, Trans-New Guinea	213
Central Asmat	Asmat-Kamrau Bay	214
Dalabon	Marne, Gunwinyguan	191
Diyari	Central Pama-Nyungan, Pama-Nyungan	196
Dyirbal	Northern Pama-Nyungan, Pama-Nyungan	197
Fore	Fore-Gimi, Trans-New Guinea	237
Goreng-Goreng	Southeastern Pama-Nyungan, Pama-Nyungan	198
Gungabula	Northern Pama-Nyungan, Pama-Nyungan	199
Hua	Siane-Yagaria, Trans-New Guinea	236
Ilocano	Northern Luzon, Austronesian	219
Imonda	Border	231
Kalkatungu	Northern Pama-Nyungan, Pama-Nyungan	200
Kopar	Lower Sepik, Ramu-Lower Sepik	239
Lower Grand Valley Dani	Dani, Trans-New Guinea	232
Makasar	South Sulawesi, Austronesian	220

(continued)

PACIFIC		
language	genetic affiliation	number
Margany	Northern Pama-Nyungan, Pama-Nyungan	201
Marrithiyel	Bringen, Western Daly	212
Mian	Ok, Trans-New Guinea	215
Mountain Arapesh	Kombio-Arapesh, Torricelli	248
Murrinhpatha	Murrinhpatha, Southern Daly	210
Nasal	Nasal, Austronesian	221
Nen	Nambu, Yam	251
Ngamini	Central Pama-Nyungan, Pama-Nyungan	202
Ngan'gityemerri	Ngankikurungkurr, Southern Daly	211
Ngiyambaa	Southern Pama-Nyungan, Pama-Nyungan	203
Ngkolmpu	Kanum, Yam	252
Pendau	Celebic, Austronesian	222
Pitta-Pitta	Central Pama-Nyungan, Pama-Nyungan	204
Salako	Land Dayak, Austronesian	223
Selepet	Huon, Trans-New Guinea	233
Standard Indonesian	Malayo-Sumbawan, Austronesian	224
Sundanese	Malayo-Sumbawan, Austronesian	225
Tairora	Tairora, Trans-New Guinea	235
Tauya	Rai Coast, Trans-New Guinea	243
Teiwa	Alor-Pantar, Greater West Bomberai	247
Telefol	Ok, Trans-New Guinea	216
Toba Batak	Northwest Sumatra-Barrier Islands, Austronesian	226
Tolai	Oceanic, Austronesian	227
Tongan	Oceanic, Austronesian	228
Tukang Besi	Celebic, Austronesian	229
Urim	Urim, Torricelli	249
Waka-Waka	Southern Pama-Nyungan, Pama-Nyungan	205
Wambaya	Wambayan, Mirndi	193
Warrgamay	Northern Pama-Nyungan, Pama-Nyungan	206
Watam	Lower Ramu, Ramu-Lower Sepik	240
West Coast Bajau	Sama-Bajaw, Austronesian	230
Wubuy	Nunggubuyu, Gunwinyguan	192
Yalarnnga	Northern Pama-Nyungan, Pama-Nyungan	207
Yale	Mek, Trans-New Guinea	234
Yaluyandi	Central Pama-Nyungan, Pama-Nyungan	208
Yeri	Wapei, Torricelli	250
Yidiny	Northern Pama-Nyungan, Pama-Nyungan	209
Yimas	Lower Sepik, Ramu-Lower Sepik	241



**Map 6:** Southwestern Pacific and northern Australia.

EURASIA		
language	genetic affiliation	number
Abaza	Northwest Caucasian	268
Abkhaz	Northwest Caucasian	269
Akhvakh	Avar-Andic-Tsezic, Nakh-Daghestanian	267
Bantawa	Kiranti, Sino-Tibetan <sup>6</sup>	255
Bulgarian	Slavic, Indo-European	279
Czech	Slavic, Indo-European	280
Dutch	Germanic, Indo-European	275
English	Germanic, Indo-European	276
Georgian	Kartvelian	263 <sup>7</sup>
German	Germanic, Indo-European	277
Hakha Lai	Kuki-Chin, Sino-Tibetan	256
Hayu	Kiranti, Sino-Tibetan	257
Hungarian	Ugric, Uralic	287
Kabardian	Northwest Caucasian	270
Khaling	Kiranti, Sino-Tibetan	258

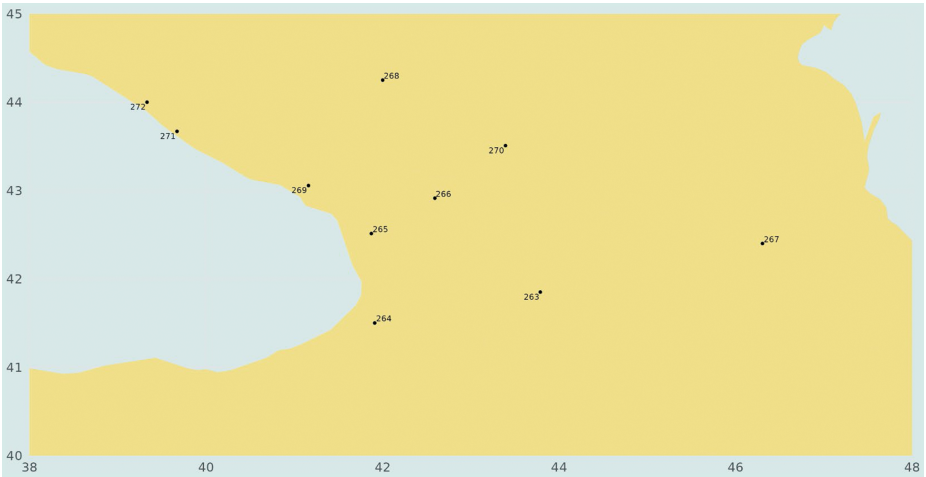
<sup>6</sup> The genus labeled “Himalayish” in WALS is more commonly labeled “Kiranti”. Moreover, “Himalayish” is ambiguous, since not all authors use it with the same extension.

<sup>7</sup> Number 263 also applies to Old Georgian, which was spoken on roughly the same territory as Modern Georgian.



(continued)

EURASIA		
language	genetic affiliation	number
Korean	Koreanic <sup>8</sup>	254
Latvian	Baltic, Indo-European	273
Laz	Kartvelian	264
Limbu	Kiranti, Sino-Tibetan	259
Lithuanian	Baltic, Indo-European	274
Macedonian	Slavic, Indo-European	281
Marathi	Indic, Indo-European	253
Mingrelian	Kartvelian	265
Polish	Slavic, Indo-European	282
Puma	Kiranti, Sino-Tibetan	260
Romanian	Italic, Indo-European <sup>9</sup>	278
Russian	Slavic, Indo-European	283
Serbian	Slavic, Indo-European	284
Slovene	Slavic, Indo-European	285
Svan	Kartvelian	266
Thulung	Kiranti, Sino-Tibetan	261
Ubykh	Northwest Caucasian	271
West Circassian	Northwest Caucasian	272
Yakkha	Kiranti, Sino-Tibetan	262



**Map 7:** Caucasus.

<sup>8</sup> The classification of Korean depends on whether Jeju is considered a Korean dialect, or a closely related language constituting alongside Korean a genus that can be labeled “Koreanic”.

<sup>9</sup> A strict application of Dryer’s (1989) definitions leads to the conclusion that the Romance languages do not constitute a genus, and that the genus to which they belong is rather the Italic branch of Indo-European.