

Appendix 3: World Maps

Chris Bentz

14 June, 2021

Session Info

Give the session info (reduced).

```
## [1] "R version 3.6.3 (2020-02-29)"  
## [1] "x86_64-pc-linux-gnu"
```

Load Libraries

If the libraries are not installed yet, you need to install them using, for example, the command: `install.packages("ggplot2")`.

```
library(readr)  
library(ggmap)  
library(maps)  
library(gridExtra)  
library(ggrepel)
```

Give the package versions.

```
##      ggrepel gridExtra      maps      ggmap      ggplot2      readr  
##    "0.8.2"      "2.3"    "3.3.0"    "3.0.0"    "3.3.2"    "1.4.0"
```

Load the Data

Load Glottolog (Version 4.1) language information combined with information on the language sample of the IWMLC.

```
languages <- as.data.frame(read_csv("https://raw.githubusercontent.com/IWMLC/language-complexity-metrics/master/data/IWMLC_languages.csv"))
```

Simple Stats

```
length(unique(languages$isocodes)) # number of languages according to iso: 79
```

```
## [1] 80
```

```
length(unique(languages$glottocode)) # number of languages according to glottolog: 79
```

```
## [1] 80
```

```
length(unique(languages$family_id)) # 34

## [1] 34

unique(languages$macroarea) # macroareas: 6

## [1] "Africa"          "Papunesia"        "Eurasia"          "South America"
## [5] "North America" "Australia"        NA
```

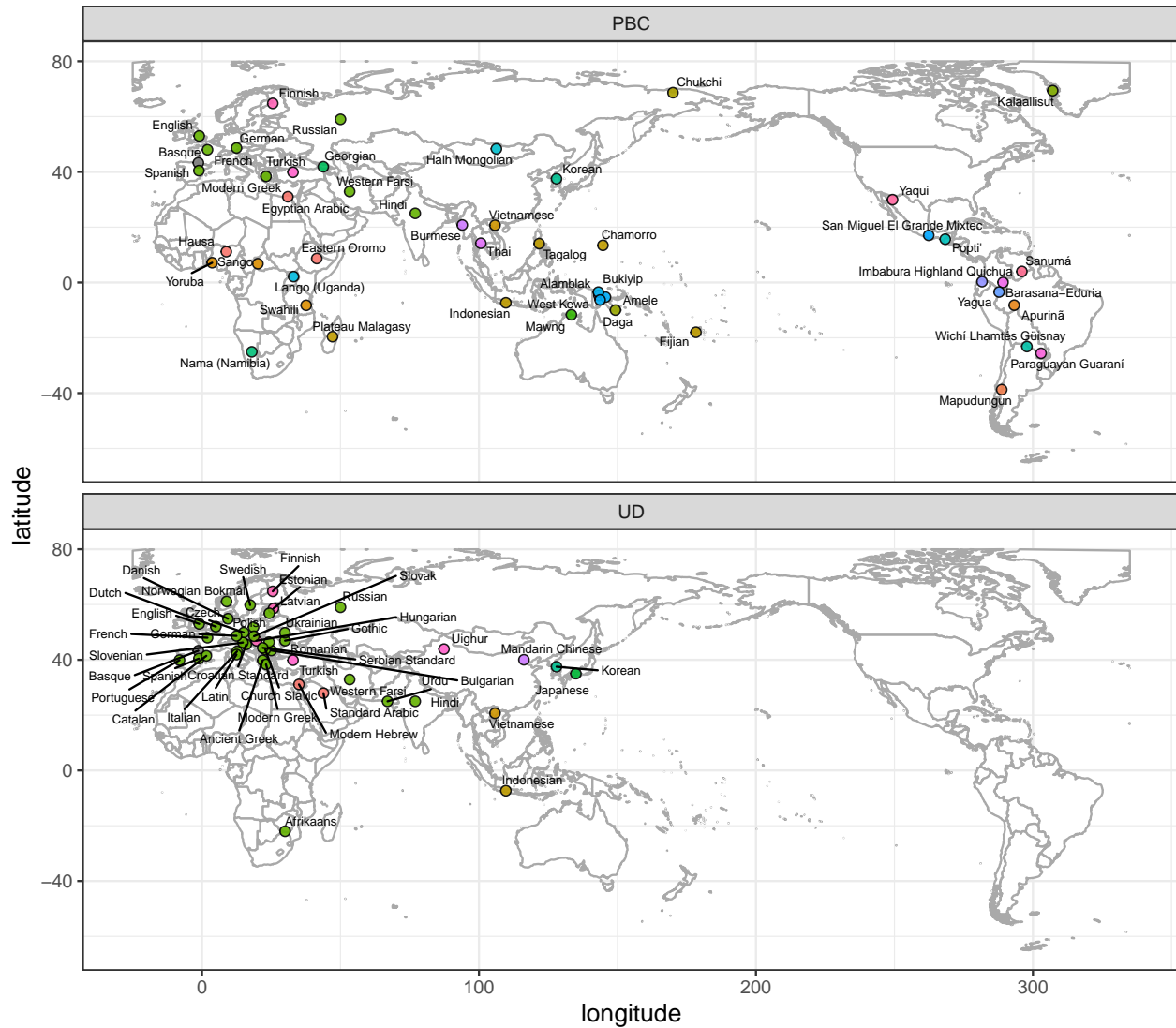
Pre-Processing

```
# remove Norwegian Nynorsk (nno) since this has NAs in glottolog
languages <- languages[languages$isocodes != "nno", ]
# add 360 to longitudes for languages with longitudes < -25
# (this is necessary to create a Pacific centered map)
languages$longitude[languages$longitude < -25] <-
  languages$longitude[languages$longitude < -25] + 360
# select non-European languages
```

World Map

World maps with macroarea information from Glottolog.

```
# create world map
world <- map_data("world", wrap = c(-25, 335))
family.map <- ggplot() +
  geom_polygon(data = world, aes(x = long, y = lat, group = group),
    fill = "white", colour = "darkgrey") +
  geom_point(data = languages, aes(x = longitude, y = latitude,
    fill = family_id),
    alpha = 0.9, size = 2, pch = 21) +
  geom_text_repel(data = languages, aes(x = longitude, y = latitude,
    label = name), size = 2,
    box.padding = unit(0.2, 'lines'), force = 1) +
  scale_y_continuous(limits = c(-65, 80)) +
  #scale_x_continuous(breaks = c(-180, -90, 0, 90, 180)) +
  labs(x = "longitude", y = "latitude", fill = "Language Family") +
  theme_bw() +
  facet_wrap(~ corpus, nrow = 2) +
  theme(axis.title.x = element_text(size = 12),
    axis.title.y = element_text(size = 12),
    title = element_text(size = 12),
    legend.title = element_text(size = 10),
    legend.text = element_text(size = 10),
    legend.position = "bottom")
family.map
```



- Language Family
- afro1255 ● daga1274 ● kore1284 ● otom1299 ● tupi1275
 - arau1255 ● eski1264 ● mata1289 ● peba1241 ● turk1311
 - araw1281 ● indo1319 ● maya1287 ● quec1387 ● ural1272
 - atla1278 ● iwai1246 ● mong1349 ● sepi1257 ● utoa1244
 - aust1305 ● japo1237 ● nilo1247 ● sino1245 ● yano1268
 - aust1307 ● kart1248 ● nucl1708 ● taik1256 ● NA
 - chuk1271 ● khoe1240 ● nucl1709 ● tuca1253

Save to file.

```
ggsave("Figures/WorldMap/worldMaps.pdf", family.map,
       dpi = 300, scale = 1, width = 8, height = 10, device = cairo_pdf)
```