Resources for Further Learning

This is basically a compiled list of all the links provided in the slides (plus a few bonus links). Don't get overwhelmed! Think of these as options to dive further into if you are interested, or starting points for when you get stumped with a coding/stats problem. This list is by no means exhaustive nor really even comprehensive, but can hopefully give an idea of how much support there is for R users and encourage you to become a more confident, competent one yourself ☺

General / Getting Started:

- Download R and R Studio: https://posit.co/download/rstudio-desktop/
- Working Directories: https://support.posit.co/hc/en-us/articles/200711843-Working-Directories-and-Workspaces-in-the-RStudio-IDE
- Cheatsheets: https://rstudio.com/resources/cheatsheets/
- Introverse. Provides interactive tutorials of many key R packages: https://sjspielman.github.io/introverse/articles/introverse_online.html
- Intro to R as a programming language: http://r-guide.steveharoz.com/index.html

Wrangling:

- The tidyverse: https://tidyverse.tidyverse.org/
- Video tutorial of data manipulation: https://www.youtube.com/watch?v=Zc_ufg4uW4U
- Joining: https://www.garrickadenbuie.com/project/tidyexplain/
- Tidying (general): https://tidydatatutor.com/

Plotting:

- Free ggplot book: https://ggplot2-book.org/
- Video workshop by Thomas Lin Pedersen:
    - Part 1: https://www.youtube.com/watch?v=h29g21z0a68
    - Part 2: https://www.youtube.com/watch?v=0m4yywqNPVY
- R Graph Gallery: https://r-graph-gallery.com/

Finding help!

- R Studio Community: https://community.rstudio.com/
- Stack Overflow (my / many people's go-to): https://stackoverflow.com/questions/tagged/r

Linguistics-specific packages:

1. **quanteda:** quantitative analysis of textual data (corpus management, tokenization, stemming, text analysis)
2. **tm:** text mining,; create and manipulate text documents, preprocessing, clustering , topic modeling.
3. **stringr:** strings and regular expressions. Good for pattern matching, string manipulation, and text cleaning
4. **udpipe:** Parse raw text; enables part-of-speech tagging, dependency parsing, and morphological analysis of text in multiple languages.

5. **koRpus:** quantitative text analysis (specifically for German. Handles tokenization, stemming, readability analysis, and other linguistic measures.
6. **phonR:** phonological analysis: phoneme transcription, syllabification, phonetic feature analysis, and phonological rule application.
7. **lme4:** Fitting linear mixed-effects models. For analyzing data with hierarchical or nested structures (e.g., experimentsal or language acquisition studies)
8. **text2vec:** text vectorization and feature extraction. Convert textual data into numerical representations, such as bag-of-words, TF-IDF, and word embeddings
9. **WordNetR:** Accessing and working with WordNet, a lexical database that provides semantic relationships between words. I
10. **tm.plugin.webmining**: Extention of 'tm' package. Scrape and process text from websites