Working Title

Dr Massimiliano Canzi

May 5th 2022

```
library(ggpubr)
```

```
## Loading required package: ggplot2
```

```
library(lme4)
```

```
## Loading required package: Matrix
```

```
## Registered S3 methods overwritten by 'lme4':
##   method                          from
##   cooks.distance.influence.merMod car
##   influence.merMod                car
##   dfbeta.influence.merMod         car
##   dfbetas.influence.merMod        car
```

```
library(lmerTest)
```

```
##
## Attaching package: 'lmerTest'
```

Section 1

Introduction

Who am I

Who am I

- Lab Manager / Data Scientist @ Uni Konstanz

Who am I

- Lab Manager / Data Scientist @ Uni Konstanz
- PhD in Linguistics @ Uni of Manchester

Who am I

- Lab Manager / Data Scientist @ Uni Konstanz
- PhD in Linguistics @ Uni of Manchester
- MSc Forensic Speech Science @ Uni of York

Who am I

- Lab Manager / Data Scientist @ Uni Konstanz
- PhD in Linguistics @ Uni of Manchester
- MSc Forensic Speech Science @ Uni of York
- MA Linguistics @ Uni of Manchester

Who am I

Interested in R, experimental design, reproducibility, open science

Once Upon a Time... in Hollywood

Once Upon a Time… in Hollywood

Planning a (linguistics) experiment and filming a Hollywood movie are not as far apart as you would think:

- Write a script

Once Upon a Time… in Hollywood

Planning a (linguistics) experiment and filming a Hollywood movie are not as far apart as you would think:

- Write a script
- Set a production budget

Once Upon a Time… in Hollywood

Planning a (linguistics) experiment and filming a Hollywood movie are not as far apart as you would think:

- Write a script
- Set a production budget
- Casting, locations, props, storyboards

Once Upon a Time… in Hollywood

Planning a (linguistics) experiment and filming a Hollywood movie are not as far apart as you would think:

- Write a script
- Set a production budget
- Casting, locations, props, storyboards
- Shoot the film

Once Upon a Time… in Hollywood

Planning a (linguistics) experiment and filming a Hollywood movie are not as far apart as you would think:

- Write a script
- Set a production budget
- Casting, locations, props, storyboards
- Shoot the film
- Editing, colour grading, VFX

Once Upon a Time… in Hollywood

Planning a (linguistics) experiment and filming a Hollywood movie are not as far apart as you would think:

- Write a script
- Set a production budget
- Casting, locations, props, storyboards
- Shoot the film
- Editing, colour grading, VFX
- Press, festivals

Once Upon a Time... in Hollywood

Once Upon a Time… in Hollywood

Planning a (linguistics) experiment and filming a Hollywood movie are not as far apart as you would think:

- Formulate a research question

Once Upon a Time… in Hollywood

Planning a (linguistics) experiment and filming a Hollywood movie are not as far apart as you would think:

- Formulate a research question
- Budget, weight methodologies, availability of participants and tools

Once Upon a Time… in Hollywood

Planning a (linguistics) experiment and filming a Hollywood movie are not as far apart as you would think:

- Formulate a research question
- Budget, weight methodologies, availability of participants and tools
- Recruitment, equipment setup, hardware and software

Once Upon a Time… in Hollywood

Planning a (linguistics) experiment and filming a Hollywood movie are not as far apart as you would think:

- Formulate a research question
- Budget, weight methodologies, availability of participants and tools
- Recruitment, equipment setup, hardware and software
- Data collection

Once Upon a Time… in Hollywood

Planning a (linguistics) experiment and filming a Hollywood movie are not as far apart as you would think:

- Formulate a research question
- Budget, weight methodologies, availability of participants and tools
- Recruitment, equipment setup, hardware and software
- Data collection
- Data wrangling, analaysis and visualisation

Once Upon a Time… in Hollywood

Planning a (linguistics) experiment and filming a Hollywood movie are not as far apart as you would think:

- Formulate a research question
- Budget, weight methodologies, availability of participants and tools
- Recruitment, equipment setup, hardware and software
- Data collection
- Data wrangling, analaysis and visualisation
- Paper, conferences

Once Upon a Time… in Hollywood

Just like when shooting a film. Plan ahead.

Once Upon a Time… in Hollywood

Just like when shooting a film. Plan ahead.

The more things you can accurately predict and plan, the more solid your design and experiment will be. That's pretty much all there is to it.

Thank you!

Questions?

Just kidding

Just kidding

Important: Starting now, most of the advice in this presentation is based on personal experience. As often is the case, there are many ways one can reach a destination. I am simply presenting some of the tips that helped me along the way.

Section 2

Pre-Production

Research Questions

Research Questions

Section 3

Production

Section 4

Post-Production

Pre-processing and Analysis

Let's have a look at some R code

Pre-processing and Analysis

Pre-processing and Analysis

Welcome back!

Pre-processing and Analysis

Pre-processing and Analysis

Sometimes, it's not as straightforward i.e. the story of event-related potentials (ERP)

Reporting results

Reporting results

- Standardisation

Reporting results

- Standardisation
- Supporting material

Reporting results

- Standardisation
- Supporting material
- package `report`

Reporting results

We fitted a linear mixed model (estimated using REML and nloptwrap optimizer) to predict rt with condition, fricative and participant_device_type (formula: rt ~ condition * fricative + participant_device_type). The model included condition, participant_private_id and item as random effects (formula: list(~condition | participant_private_id, ~1 | item)). The model's total explanatory power is substantial (conditional R2 = 0.41) and the part related to the fixed effects alone (marginal R2) is of 0.08. The model's intercept, corresponding to condition = NM, fricative = FF and participant_device_type = computer, is at 6.89 (95% CI [6.81, 6.97], t(11044) = 171.01, p < .001). Within this model:

- The effect of condition [WM] is statistically non-significant and positive (beta = 4.10e-03, 95% CI [-0.04, 0.05], t(11044) = 0.17, p = 0.867; Std. beta = 7.57e-03, 95% CI [-0.08, 0.10])
- The effect of fricative [SH] is statistically significant and negative (beta = -0.22, 95% CI [-0.30, -0.15], t(11044) = -5.52, p < .001; Std. beta = -0.41, 95% CI [-0.56, -0.27]) …

Reporting results

Analyses were conducted using the R Statistical language (version 4.0.3; R Core Team, 2020) on macOS Big Sur 10.16, using the packages ggpubr (version 0.4.0; Alboukadel Kassambara, 2020), Matrix (version 1.3.2; Douglas Bates and Martin Maechler, 2021), lme4 (version 1.1.26; Douglas Bates et al., 2015), ggplot2 (version 3.3.5; Wickham. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York, 2016.), stringr (version 1.4.0; Hadley Wickham, 2019), tidyr (version 1.1.2; Hadley Wickham, 2020), forcats (version 0.5.1; Hadley Wickham, 2021), readr (version 1.4.0; Hadley Wickham and Jim Hester, 2020), dplyr (version 1.0.4; Hadley Wickham et al., 2021), tibble (version 3.1.5; Kirill Müller and Hadley Wickham, 2021), lmerTest (version 3.1.3; Kuznetsova A et al., 2017), purrr (version 0.3.4; Lionel Henry and Hadley Wickham, 2020), sjPlot (version 2.8.9; Lüdecke D, 2021), viridis (version 0.5.1; Simon Garnier, 2018), viridisLite (version 0.4.0; Simon Garnier et al., 2021) and tidyverse (version 1.3.0; Wickham et al., 2019).

References

- Alboukadel Kassambara (2020). ggpubr: 'ggplot2' Based Publication Ready Plots. R package version 0.4.0. https://CRAN.R-project.org/package=ggpubr
- Douglas Bates and Martin Maechler (2021). Matrix: Sparse and Dense Matrix Classes and Methods. R package version 1.3-2. https://CRAN.R-project.org/package=Matrix
- Douglas Bates, Martin Maechler, Ben Bolker, Steve Walker (2015). Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software, 67(1), 1-48. doi:10.18637/jss.v067.i01.
- H. Wickham. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York, 2016.
- Hadley Wickham (2019). stringr: Simple, Consistent Wrappers for Common String Operations. R package version 1.4.0. https://CRAN.R-project.org/package=stringr …