Breaking Bad Habits in Experimental and Quantitative Research in Linguistics

Dr Massimiliano Canzi

mcanzi/2022 daw | May 5th 2022

Section 1

Introduction

Massimiliano "Max" Canzi

• Lab Manager / Data Scientist @ Uni Konstanz

Introduction

000000000

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

Introduction

000000000

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

Introduction

000000000

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

Interested in: \mathbf{R} , experimental designs, $\mathbf{reproducibility}$, open science, improving existing methodologies, data visualisation, honest scientific communication...

Post-Production

About the LingLabs

Consortium of linguistic laboratories at the University of Konstanz spanning a variety of research topics and fields.

Consortium of linguistic laboratories at the University of Konstanz spanning a variety of research topics and fields.

• Digital Signal Processing in R by Dr Cristopher Carignan (UCL). 8th and 9th of June, 2-4 pm CEST

Introduction

0000000000

Consortium of linguistic laboratories at the University of Konstanz spanning a variety of research topics and fields.

- Digital Signal Processing in R by Dr Cristopher Carignan (UCL). 8th and 9th of June, 2-4 pm CEST
- Collecting and exploring social media data in R by Dr Lisa Donlan (UoManchester). \(\beta 22nd \) and \(23rd \) of \(Sept. \) 2-4 \(pm \) \(CEST \)

Introduction

0000000000

Consortium of linguistic laboratories at the University of Konstanz spanning a variety of research topics and fields.

- Digital Signal Processing in R by Dr Cristopher Carignan (UCL). 8th and 9th of June, 2-4 pm CEST
- Collecting and exploring social media data in R by Dr Lisa Donlan (UoManchester). \(\beta 22nd \) and \(23rd \) of \(Sept. \) 2-4 \(pm \) \(CEST \)
- ...

Introduction

0000000000

Consortium of linguistic laboratories at the University of Konstanz spanning a variety of research topics and fields.

- Digital Signal Processing in R by Dr Cristopher Carignan (UCL). 8th and 9th of June, 2-4 pm CEST
- Collecting and exploring social media data in R by Dr Lisa Donlan (UoManchester). \(\beta 22nd \) and \(23rd \) of \(Sept, \) 2-4 \(pm \) \(CEST \)
- ...

Introduction

0000000000

Consortium of linguistic laboratories at the University of Konstanz spanning a variety of research topics and fields.

- Digital Signal Processing in R by Dr Cristopher Carignan (UCL). 8th and 9th of June, 2-4 pm CEST
- Collecting and exploring social media data in R by Dr Lisa Donlan (UoManchester). \(\beta 22nd \) and \(23rd \) of \(Sept, \) 2-4 \(pm \) \(CEST \)

...

https://www.ling.uni-konstanz.de/forschung/workshops/

Planning a (linguistics) experiment and a movie production are not as far apart as you would think...

Planning a (linguistics) experiment and a movie production are not as far apart as you would think...

• Write a script

Introduction

Planning a (linguistics) experiment and a movie production are not as far apart as you would think...

• Write a script

Introduction

0000000000

Set a production budget

Planning a (linguistics) experiment and a movie production are not as far apart as you would think...

Write a script

Introduction

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

Planning a (linguistics) experiment and a movie production are not as far apart as you would think...

 \bullet Write a script

Introduction

- Set a production budget
- Casting, locations, props, storyboards
- \bullet Shoot the film

Planning a (linguistics) experiment and a movie production are not as far apart as you would think...

• Write a script

Introduction

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

Planning a (linguistics) experiment and a movie production are not as far apart as you would think...

• Write a script

Introduction

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



Figure 1: Severance (2022)

Planning a (linguistics) experiment and a movie production are not as far apart as you would think...

Planning a (linguistics) experiment and a movie production are not as far apart as you would think...

 \bullet Formulate a research question

Introduction

Planning a (linguistics) experiment and a movie production are not as far apart as you would think...

• Formulate a research question

Introduction

0000000000

• Budget, weigh methodologies, availability of participants and tools

Planning a (linguistics) experiment and a movie production are not as far apart as you would think...

• Formulate a research question

Introduction

- Budget, weigh methodologies, availability of participants and tools
- Recruitment, equipment setup, hardware and software

Planning a (linguistics) experiment and a movie production are not as far apart as you would think...

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

Introduction

Planning a (linguistics) experiment and a movie production are not as far apart as you would think...

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

Introduction

0000000000

• Data wrangling, analaysis and visualisation

Planning a (linguistics) experiment and a movie production are not as far apart as you would think...

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

Introduction

- Data wrangling, analaysis and visualisation
- Papers, conferences

Just like when shooting a film, **plan ahead**.

Post-Production

Motion Pictures / Scientific Experiments

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?

Thank you!

Thank you!

 $\mathbf{Just} \ \mathbf{kidding...} \ \operatorname{let's} \ \operatorname{get} \ \operatorname{started}.$

Section 2

Pre-Production

Find a direct, achievable link between the ${f theory}$ and its ${f implementation}.$

If your research question is too complex $\mathbf{scale}\ \mathbf{down},\ \mathbf{zoom}\ \mathbf{in}.$

Find a direct, achievable link between the ${f theory}$ and its ${f implementation}.$

If your research question is too complex $\mathbf{scale}\ \mathbf{down},\ \mathbf{zoom}\ \mathbf{in}.$

Develop clear, ${f testable}$ hypotheses.

Find a direct, achievable link between the **theory** and its **implementation**.

• What **method** is best to answer the question?

- What **method** is best to answer the question?
- How many experiments?

- What **method** is best to answer the question?
- How many experiments?
- What experimental **design**?

- What **method** is best to answer the question?
- How many experiments?
- What experimental **design**?
- How (and what) are variables going to be coded?

Choose a method that fits your research question, but keep ease of implementation in mind.

Choose a method that fits your research question, but keep ease of implementation in mind.

Know your method well! Methods comes with baggage, meaning that often the chosen method determines conventions in experimental design, data analysis, presentation of results, etc.

Choose a method that fits your research question, but keep ease of implementation in mind.

Know your method well! Methods comes with baggage, meaning that often the chosen method determines conventions in experimental design, data analysis, presentation of results, etc.

Don't be afraid to innovate.

Film or Mini Series?

Film or Mini Series?

Multiple experiments $\underline{\text{might}}$ allow for a less complex experimental design, follow-up studies, ...

Film or Mini Series?

Multiple experiments $\underline{\text{might}}$ allow for a less complex experimental design, follow-up studies, ...

However, more experiments also mean more time spent on participant recruitment, data collection, \dots

If you were to take one thing home from today, make it this one:

If you were to take one thing home from today, make it this one:

 \gg less is more \ll

If you were to take one thing home from today, make it this one:

» less is more «

Every time you add a categorical predictor with two levels, for example, you should double your data. Can you afford it?

If you were to take one thing home from today, make it this one:

» less is more «

Every time you add a categorical predictor with two levels, for example, you should double your data. Can you afford it?

Use tools like simr to help you determine sample size (Green and MacLeod 2016).

If you were to take one thing home from today, make it this one:

» less is more «

Every time you add a categorical predictor with two levels, for example, you should double your data. Can you afford it?

Use tools like simr to help you determine sample size (Green and MacLeod 2016).

For the now visibly upset Bayesians...

Choosing an experimental method and design is equivalent to choosing the **resolution** of your study.

Choosing an experimental method and design is equivalent to choosing the **resolution** of your study.

Choosing an experimental method and design is equivalent to choosing the **resolution** of your study.

It is, again, really important that your design allows you to confidently test your hypotheses as you intend to.

• Know your **limits**

Choosing an experimental method and design is equivalent to choosing the **resolution** of your study.

- Know your **limits**
- Know your **goals**

Choosing an experimental method and design is equivalent to choosing the **resolution** of your study.

- Know your **limits**
- Know your **goals**
- Choose a balanced design

Choosing an experimental method and design is equivalent to choosing the **resolution** of your study.

- Know your **limits**
- Know your **goals**
- Choose a balanced design
- Register your report?

Registered Reports

- Mertzen, D., Lago, S., & Vasishth, S. (2021). The benefits of preregistration for hypothesis-driven bilingualism research. Bilingualism: Language and Cognition, 24(5), 807-812.
- Roettger, T. B. (2021). Preregistration in experimental linguistics: Applications, challenges, and limitations. Linguistics, 59(5), 1227-1249.

One mistake I see often is related to variable coding, for example..

• Binary instead of continuous..

- Binary instead of continuous..
- $\bullet\,$ 5-point instead of 100-point scales

- Binary instead of continuous...
- 5-point instead of 100-point scales
- 100-point scales instead of 5-point scales

- Binary instead of continuous..
- $\bullet\,$ 5-point instead of 100-point scales
- $\bullet~$ 100-point scales instead of 5-point scales
- ...

Section 3

Production

Things to consider:

 \bullet Location of the experiment (e.g. lab, online)

- \bullet Location of the experiment (e.g. lab, online)
- $\bullet\,$ Number of items

- \bullet Location of the experiment (e.g. lab, online)
- Number of items
- Number of participants

- Location of the experiment (e.g. lab, online)
- Number of items
- Number of participants
- What sample?

- Location of the experiment (e.g. lab, online)
- Number of items
- Number of participants
- What sample?

Don't underestimate the power of the internet.

Don't underestimate the power of the internet.

Social medias, esp. $\mathbf{Twitter}$, are a great place to advertise for your online studies.

Don't underestimate the power of the internet.

Social medias, esp. $\mathbf{Twitter}$, are a great place to advertise for your online studies.

Online platforms are getting better and better, allowing for more complex designs and data collection types (e.g. reaction times, Stoet 2017).

A balance between number of items and participants is necessary, and often dictated by the methodology

A balance between number of items and participants is necessary, and often dictated by the methodology e.g. ERP experiments.

A balance between number of items and participants is necessary, and often dictated by the methodology e.g. ERP experiments.

Ideally, you would have an infinite number of both.

A balance between number of items and participants is necessary, and often dictated by the methodology e.g. ERP experiments.

Ideally, you would have an infinite number of both.

Always ask your participants a little more, especially if it's a lab-based experiment.

Casting

Casting

Are university students in your university representative enough of the sample you're looking for?

• Write everything down (hardware, software, locations, etc.)

- Write everything down (hardware, software, locations, etc.)
- Get all the forms you need (participant consent, data protection)

- Write everything down (hardware, software, locations, etc.)
- Get all the forms you need (participant consent, data protection)
- Plan a storage solution for data and anonymise whenever possible

- Write everything down (hardware, software, locations, etc.)
- Get all the forms you need (participant consent, data protection)
- Plan a storage solution for data and anonymise whenever possible
- Hire a research assistant

Section 4

Post-Production

Know your data!

Let's have a look at some R code

model_one <- rt_log %>% lmer(formula = rt ~ condition * fricative + partic

Welcome back!

Welcome back!

Some references:

- Baayen, R. H., & Milin, P. (2010). Analyzing reaction times. International Journal of Psychological Research, 3(2), 12-28.
- Leys, C., Ley, C., Klein, O., Bernard, P., & Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. Journal of experimental social psychology, 49(4), 764-766.

When modelling:

• Your most complex model should be clear to you before you even start collecting data.

When modelling:

- Your most complex model should be clear to you before you even start collecting data.
- What is a fixed effect and what is a random effect in your experiment?

When modelling:

- Your most complex model should be clear to you before you even start collecting data.
- What is a fixed effect and what is a random effect in your experiment?
- Slopes galore

Sometimes, it's not as straightforward

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

Reporting your results:

Reporting your results:

• Standardise (e.g. APA)

Reporting your results:

- Standardise (e.g. APA)
- \bullet Provide supporting material

Reporting your results:

- Standardise (e.g. APA)
- \bullet Provide supporting material
- \bullet package report

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]) ...

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).



Figure 2: Severance (2022)

When modelling, I like to provide plots showing model estimates, rather than plotting raw data.

When modelling, I like to provide plots showing model estimates, rather than plotting raw data.

For example, you can use package sjPlot, or do it yourself with packages ggeffects and ggplot2.

Production

000000

Color Grading

Again, » less is more «. A few tips:

• Use package viridis whenever possible – https://cran.rproject.org/web/packages/viridis/vignettes/intro-to-viridis.html

Production

Color Grading

- Use package viridis whenever possible https://cran.rproject.org/web/packages/viridis/vignettes/intro-to-viridis.html
- Differentiate with shape, not just with colour

- Use package viridis whenever possible https://cran.rproject.org/web/packages/viridis/vignettes/intro-to-viridis.html
- Differentiate with shape, not just with colour
- Label everything clearly

- Use package viridis whenever possible https://cran.rproject.org/web/packages/viridis/vignettes/intro-to-viridis.html
- Differentiate with shape, not just with colour
- Label everything clearly
- Don't cheat!

- Use package viridis whenever possible https://cran.rproject.org/web/packages/viridis/vignettes/intro-to-viridis.html
- Differentiate with shape, not just with colour
- Label everything clearly
- Don't cheat!
- Check with journals

Experiment done, data analysed, results reported, paper written. Time to **get it** out into the world.

Experiment done, data analysed, results reported, paper written. Time to **get it** out into the world.

Consider **Open Access** whenever possible.

Experiment done, data analysed, results reported, paper written. Time to **get it** out into the world.

Consider **Open Access** whenever possible.

Share your data and your code with your paper.

Ending Credits

- Baayen, R. H., & Milin, P. (2010). Analyzing reaction times. International Journal of Psychological Research, 3(2), 12-28.
- Green, P., & MacLeod, C. J. (2016). SIMR: an R package for power analysis of generalized linear mixed models by simulation. Methods in Ecology and Evolution, 7(4), 493-498.
- Leys, C., Ley, C., Klein, O., Bernard, P., & Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. Journal of experimental social psychology, 49(4), 764-766.
- Stoet, G. (2017). PsyToolkit: A novel web-based method for running online questionnaires and reaction-time experiments. Teaching of Psychology, 44(1), 24-31.