Digital Research Toolkit for Linguists

Week 10: ETFX basics

Anna Prysłopska June 10, 2024

Psycholinguistics and Cognitive Modeling Lab

Submit QMD and HTML

Keep **all the code** you need for analyzing and visualizing the data.

1

Table, list, plot

python



assignmentqmd.sec



renamed to assignment.qmd
that's why I asked for both QMD and HTML

A report should contain your name and title of the report. Also **some** (rudimentary) description of the experiment.

title: "Reporting and Documenting with Quarto"

author: "Anna Prysłopska"

format: html
editor: visual

•••

format output type (HTML, PDF, DOC, ...), not quarto

 $\textbf{editor} \hspace{1cm} \textbf{type of text editor (visual or not} \rightarrow \textbf{leave out)}$

```
66 Reference the table, list, and plot in the report text by
hyperlinking/cross-referencing. 55
: Table caption. {#tbl-id}
                                              Table ID & caption
[Table reference](#tbl-id)
                                               Manual reference
ลtbl-id
                                Automatically formated reference
```{r}
| label: tbl-id
#| fig-cap: "Table caption."
print(table)
Now I can cross-reference atbl-id in the text.
```

Code type goes in the curly brackets {}. Code itself goes in between the ```:

```
```{r}
sessionInfo()
```
```

**66** Include at least one **plot** of the data. **99** 

plot != image

If you want to include an image, include it in the upload.



# Data Understand Communicate Share

R & RStudio, packages, data types, formats, encoding import from workspace, assign values, operations, clean, filter, arrange, select, merge, group, summarize, export, visualize document, create clean and beautiful reports connect, collaborate, backup

# Table of contents

- 1. LETEX
- 2. Typesetting documents
- 3. Basic document structure
- 4. Scientific document structure
- 5. Wrap-up
- 6. Homework assignment

**ETEX** 

# Beautiful Typesetting with LaTeX

Overfull \hbox (9.895pt too wide)



# What is ETEX



Show Menu

### An introduction to LaTeX

LaTeX, which is pronounced «Lah-tech» or «Lay-tech» (to rhyme with «blech» or «Bertolt Brecht»), is a

**LaTeX** (/ˈlɑːtɛk/ *LAH-tek* or /ˈleɪtɛk/ *LAY-tek*, $^{[2]}$  [Note 1] often stylized as  $\text{LAT}_{FX}$ ) is a software

"English words like 'technology' stem from a Greek root beginning with the letters τεχ...; and this same Greek word means art as well as technology. Hence the name TeX, which is an uppercase form of τεχ.

Insiders pronounce the  $\chi$  of TeX as a Greek chi, not as an ' $\chi$ ', so that TeX rhymes with the word blecchhh. It's the 'ch' sound in Scottish words like loch or German words like ach, it's a Spanish ' $\eta$ ' and a Russian 'kh'. When you say it correctly to your computer, the terminal may become slightly moist."

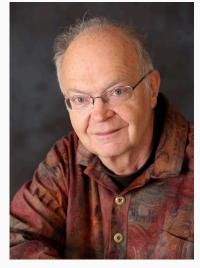
-- Donald Knuth

### Modern Greek [edit]

In Modern Greek, it has two distinct pronunciations: In front of high or front vowels (Ie/ or Ii/) it is pronounced as a voiceless palatal fricative [ç], as in German Ich or Iike some pronunciations of "h" in English words like hew and human. In front of low or back vowels (Ia/, Io/ or Iu/) and consonants, it is pronounced as a voiceless velar fricative ([x]), as in German ach or Spanish J. This distinction corresponds to the ich-Laut and ach-Laut of German.

Name: 'latɛç, 'lɑ:tɛx or 'leɪtɛx but you can do whatever as long as it's not 'leɪtɛks

# What is **ETEX**



Donald Knuth © Hector Garcia-Molina

**Originally:** T<sub>E</sub>X was a computer program for typesetting text and mathematic formulae.

Now it comes in different flavors: <code>MT\_EX</code>, <code>pdfT\_EX</code>, <code>pdfMT\_EX</code>, <code>X\_JMT\_EX</code>, <code>LuaMT\_EX</code>, <code>LuaMT\_EX</code>, ...

Aim: typeset and print at the highest typographical quality using a predefined, professional layout.

Installation: Base and packages (https://ctan.org/)



# ETEX vs. Word

### Reporting and Documenting with Quarto

Anna Prysłopska, ,

<A-Tean@army.org>

### Introduction

This a Quarto document. It shows a lot of features. You can use the source or visual editors to write your report and the reader the file to the format of your choosing. Quarto primarily works with UTP-8 encoded text.

### Structure

Give your file structure. Use sections to organize the content into logical bits:

### # Header 1

## Header 2 ### Header 3 #### You get the idea

If you're using the visual editor, simply click on the option you want from the drop down menu, as in Figure 11.

You can also include horizontal rules if you want to be funcy. Just use 3+ asterisks \*\*\*\*\*\* or dashes -----

### Basic text formatting

You can make your text stand out by making it stalic \* italic \* \_italic\_ or bold \*\*bold\*\* \_\_\_bbid\_\_. Spacing matters: \* 'this is broken\*\*. In the visual editor, a let of the options are realible as the usual keyboard shortests, but you can also format text by clicking [Figure 2].

### Reporting and Documenting with Quarto

Anna Prysłopska, 孫子, Стамен Гигов Григоров «А-Теанбагну, огд»

### Introduction

This a Quarto document. It shows a lot of features. You can use the source or visual editors to write your report and the render the file to the format of your choosing. Quarto primarily works with ULFS encoded text.

### Structure

Give your file structure. Use sections to organize the content into logical bits:

### # Header 1 ## Header 2 ### Header 3

#### You get the idea

If you're using the visual editor, simply click on the option you want from the drop down menu, as in Figure 1).



### Figure 1: Adding sections to Quarto document.

You can also include horizontal rules if you want to be fancy, Just use 3+ asterisks \*\*\*\*\*\* or dashes · · · · · ·

### Basic text formatting

You can make your text stand out by making it infile "italie" \_italie\_or bold 
"bold" \_bold . Spacing matters: "filis is broken". In the visual editor, a lot of the 
options are available as the usual keyboard shortcuts, but you can also format text by 
dicking (Figure 2).

# 町EX vs. Word

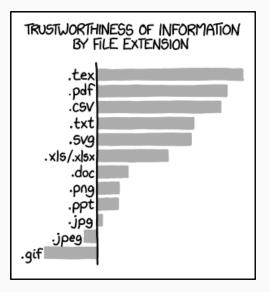
|             | ETEX                     | Word                     |
|-------------|--------------------------|--------------------------|
| Typesetting | program                  | user                     |
| Design      | program                  | user                     |
| Changes     | appear later (after com- | updated dynamically      |
|             | pilation)                | (interactively)          |
| Output      | abstract thinking        | what you see is what you |
|             |                          | get                      |
| Focus       | writing                  | writing, layout, design  |
| Rights      | free and OS              | proprietary              |
| Documents   | literally anything       | limited to most common   |
|             |                          | formats                  |

# Advantages and disadvantages: **MEX**

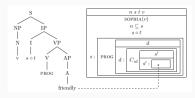
| PROS                                           | CONS                                                     |
|------------------------------------------------|----------------------------------------------------------|
| Free and portable                              | Less common                                              |
| Professionally crafted layouts                 | Made by opinionated nerds                                |
| Beautiful documents with minimal effort        | Changing the design requires patience & arcane knowledge |
| Pictures/tables appear where they should       | Naming files matters                                     |
| Typesetting mathematical & semantic formu-     | People might get jealous                                 |
| lae, symbols, syntactic trees, automata, DRSs, |                                                          |
| IPA etc. is supported & convenient             |                                                          |
| Footnotes, TOC, references, bibliographies are |                                                          |
| easy, dynamic, & automatic                     |                                                          |
| Free packages for all typographical needs      | Some packages hate each other                            |
| Just text                                      | Debugging is annoying                                    |
| Fast once learned                              | Steep learning curve                                     |

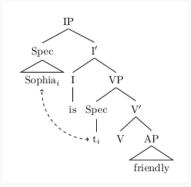
# Advantages and disadvantages: Word

| PROS                      | CONS                                         |
|---------------------------|----------------------------------------------|
| Minimal learning curve    | Creating professional & beautiful documents  |
|                           | takes time and effort                        |
| WYSIWYG                   | Math is difficult, slow, & often ugly        |
| Known & ridiculously pop- | Glosses, DRSs, syntactic trees are a pain to |
| ular                      | make and place                               |
| Easy-to-use reviewing     | Bibliography is clunky                       |
| tools                     |                                              |
|                           | (Cross)references are buggy & slow           |
|                           | Repositioning elements creates chaos         |
|                           |                                              |

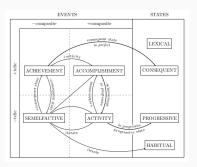


Munroe (2013)

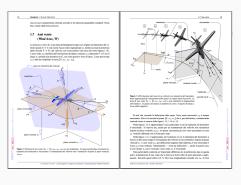




- (52) Sophia is being friendly.
  - a. [Sophia] ≡ SOPHIA<sub>+animate</sub>
  - b.  $[be_3] \equiv \lambda P[P_{+active}]$
  - c.  $[be_3 \text{ friendly}] \equiv \lambda P[P_{+active}](\text{friendly})$  $\equiv \lambda x[\text{friendly}_{+active}(x)]$
  - $$\begin{split} \text{d.} & \quad \text{[Sophia is being friendly]} \\ & \equiv \lambda x [\text{Friendly}_{+active}(x)] (\text{Sophia}_{+animate}) \\ & \quad \equiv \text{Friendly}_{+active}(\text{Sophia}_{+animate}) \end{split}$$







Tilburg Science Hub (2024)

Credinciosii: Doamne miluieste. Preotul: Pe Preasfânta, curata, preabinecuvântata, mărita stăpâna noastră, de Dumnezeu Născâtoarea și pururea Fecioara Maria, cu toți

sfintii sd o pomenim.

People: Lord, have mercy. Priest: Let us remember our most holy, pure, blessed, and elorious Lady, the Mother of God and ever virgin Mary, with all the saints.

Credinciosii: Preasfântă Născătoare de Dumnezeu, miluieste ne nei noi. Preoful: Pe noi însine si unii ne altii si toată

Prea Sfân tă Născă-toa-re de Dum ne-xeu, milu-ie-ste-ne pe noi. People: Most holy Mother of God, have mercy Priest: Let us commit ourselves and one another

viața noastră, lui Hristos, Dumneneu să o dâm. and all our life to Christ our God.

Credinciosii: Tie Doamne Preotul: Gi Tie se cuvine toată slava, cinstea si închinăciunea, Tatălui și Fiului și Sfântului Duh, acum si pururea si în vecii vecilor.

People: To You, O Lord.

Priest: For to You belong all glory, honor, and worship to the Father and the Son and the Holy Spirit, now and forever and to the ages of ages.

Credinciosii: Amin

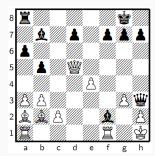
People: Amen.

### антірондь їнтаї? THE PIRST ANTIPHON2

Credinciosii: Mărire Tatălui și Fiului și Sfântului Duh, si acum si pururea si în vecii vecilor. Amin. Binecuvintează suflete al meu ne Domnul. Si toate cele dinläuntrul meu numele

cel Sfânt al lui. Binecuyântat esti Doamne

Bi-ne-cu-ván - tat..... People: Glory to the Bather, and to the Son. and to the Holy Spirit, now and ever, and to the ages of ages. Amen. Bless the Lord, O my soul, And all that is within me, bless His holy Name, Blessed are You O Lord



<sup>&</sup>lt;sup>2</sup>Cuvântul «antifon» vine din limba greacă și înseamnă -treaptā-. Antiforul Întâi cuprinde începtul Promuhii 102 si -step-. The First Antiphon encompasses the beginning of este o clintare de slávire a lui Dumnezeu.

<sup>&</sup>lt;sup>2</sup>The word «antiphon» comes from Greek, meaning Paules 102, and it is a song of peake to God.



### Le premier liure de Moyfe,

Dift Genefe.

Jacquari processos d'irrigue i associet qu'ex cotas qu'aprile i irrigue d'i provincia sa susta coque. D'i som-mément des Perus américo, qui est glid sant decumi quispen le delage, d'i en effend d'11 5 V 5. C 18 R 1 5 T defen-da d'istras frien de char.

CHAPITRE I. Consider the old the first many B, vo., the first many aging of some price, two (is in instance agil, so the Arthonous, in coloque) and of adjusted, 1,12, all these fixed town for manner, is gold a accomplished, by these



vuide.& les tenebres effoyent für les abylines : & l'Elprit de Dieu deftoit espandu par deffus les eaux. 3 Adone Dieu dit, 'Qu'll y ait lumie-

re, 'Et la lumiere fut. + Et Dieu vid & la lumiere effoit bonne : & fepara la lumiere des tenebres. s Et Dieu appela la lumiere tour, & les renebres muid. Lors fur faid le foir & le marin du premier jour-6 Puis Dieu dit, 'Qu'il y ait vne fef-

tendue entre les eaux, & quelle fepare les beaux dhuec les eaux. 7 Dieu donc fit leftendue, & diuffa

les eaux, qui efforent fous lithendue. dauec celles, qui eftovent fur leftendue. Et fut ainfi faict. 8 Et Dieu appela littendue, Gel. Lors

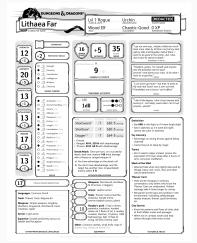
fur faict le foir & le marin du fecond mence - 9 ¶Puis Dieu dit, + 'Que les eaux, qui font fous le ciel, fovent affemblees en

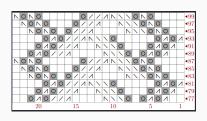
vn lieu, & que le fec apparoiffe. Et fut ainti faich. 10 Et Dieu appelale fec, Terre, & Baffern blee des eaux, mers. Et Dieu vid que celà eftoit bon.

11 Et Dieu dlt, Que la terre produife verdure, herbe produifant femence, & arbre fruidtier, faifant fruid felon fon efocce, lectuel ait fa feméce en fov-mefme fur la terre. Et fut sinfi fsiét. 12. La terre déc produifit verdure, her-

be produifant femêre felon fon efnece. & arbre fans fruiêt, lequel auoit fa femence en foymelme felon fon efpeor. Fr Dieu vid oue orlà effoit bon. 12 Lors fur faich le foir & le marin du

troifieme lour. 14 ¶ Apres Dieu dît, 1 kQu'îl yait lumi naires en lestendue du ciel, pour separer la nuich du jour : & foret en fignes,







### Mathematics

In mathematics he was greater Than Tycho Brahe, or Erra Pater: For he, by geometric scale, Could take the size of pots of ale; Resolve, by sines

and tangents straight, If bread or butter wanted weight; And wisely tell what hour o the day The clock does strike, by Algebra.

Samuel Butler (1612–1680)

### Arranging plots

### patchwork

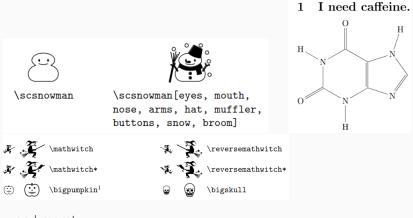
makes it ridiculously simple to combine separate ggplots into the same graphic.

### cowplot

provides various features to make plots beautiful, including aligning and arranging plots.



\*\*



... and more!

# The T<sub>E</sub>X family tree

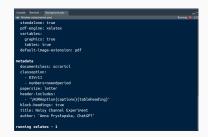
TFX The Original<sup>TM</sup> but basic and outputs DVI T<sub>F</sub>X extended and includes packages ETFX pdfT<sub>F</sub>X T<sub>F</sub>X improved and outputs PDF pdflET<sub>F</sub>X **ETFX** improved and outputs PDF T<sub>F</sub>X expanded beyond English + fonts  $X_{\exists}T_{F}X$ Xalta file format for XaTeX engine LuaT<sub>F</sub>X T<sub>F</sub>X + Lua (good for macros) LuaET<sub>E</sub>X file format for LuaT<sub>E</sub>X engine ConT<sub>F</sub>Xt T<sub>F</sub>X + interface for advanced typography features AMST<sub>E</sub>X † T<sub>F</sub>X extension by American Mathematical Society

... and more!

Typesetting documents

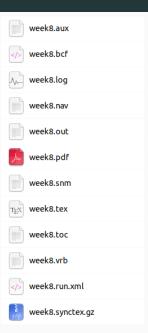
# $R o allestrick{ET_EX}$





# Output





| AUX BBL BCF BIB BLG LOF LOG LOT NAV OUT | mostly reference information prepared bibliography data citations for PDF your list of citations bibliography log list of figures compilation log list of tables navgation bookmarks |
|-----------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| PDF                                     | PDF file                                                                                                                                                                             |
| RUN.XML                                 | bibliography <b>XML</b> file                                                                                                                                                         |
| SNM                                     | information used by <b>pgfimage</b> for slides                                                                                                                                       |
| SYNTEX.GZ                               | correspondence between PDF and TEX                                                                                                                                                   |
| TEX                                     | ₽T <sub>E</sub> X file                                                                                                                                                               |
| TOC                                     | table of contents                                                                                                                                                                    |
| VRB                                     | verbatim material for fragile slides                                                                                                                                                 |

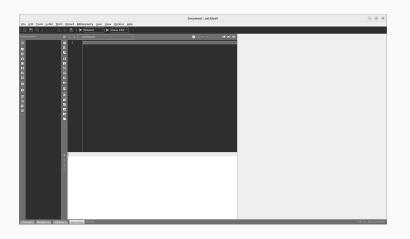
# Output



# Workflow

- 1. Input: text file = text + commands that tell MEX how to typeset the text
- 3. Output: PDF + LOG + AUX + ...

## Texmaker



https://www.xm1math.net/texmaker/

You can also use Overleaf: https://www.overleaf.com

Basic document structure

### Global document structure

- 1. Document class
- 2. Preamble
  - Packages
  - · Commands
  - · Default settings
  - · Title
  - Author
  - Date
  - ...
- 3. Start of document
- 4. Document
- 5. End of document

\documentclass[options]{...}

```
\usepackage[options]{...}
\newcommand{\xmark}{\ding{55}}}
```

\graphicspath{{./images/}}

\title{...}

\author{...}

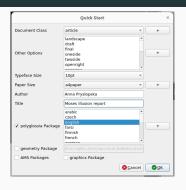
\date{...}

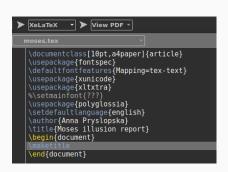
\begin{document}

In a hole in the ground there lived a hobbit.

\end{document}

# First document





- 1. Create new file
- 2. Wizard  $\rightarrow$  Quick X3ETEX document
- 3. Save file
- 4. Run X3ETEX
- 5. Run View PDF
- 6. Profit!

#### First document

\maketitle

Creates title

\section{...}

Creates section

\subsection{...}

Creates subsection

\subsection{...}

Creates subsubsection



#### **Document classes**

#### \document{...}

- article articles in scientific journals, presentations, short reports, program documentation, etc.
- minimal minimal document, mainly used for debugging (see also standalone)
- report longer reports with several chapters, small books, thesis, etc.
- book books
- memoir similar to book and report but supports more design-related options; for poetry, fiction, non-fiction, and mathematical works etc.
- · letter letters
- beamer presentations

(Almost) complete list: https://ctan.org/topic/class

## **Document options**

#### \document[...]{report}

- 10pt, 11pt, 12pt main font size (default is 10pt)
- · a4paper, letterpaper, b5paper, ... paper size
- titlepage, notitlepage put the title on a separate page or not
- twocolumn typeset the document in 2 columns
- twoside, oneside typeset as double or single sided output (affects e.g. margins, not the printer)
- · landscape typeset in landscape mode
- openright, openany chapters begin either only on right hand pages or on the next page available
- draft highlights hyphenation and justification problems, shows placeholders instead of included images

### **Packages**

Declared **only** in the preamble

Sometimes the order matters

Not all packages are **compatible** with each other or your T<sub>E</sub>X distribution

#### \usepackage{...}

- inputenc accept different input encodings
- · babel multilingual support (see also polyglossia)
- csquotes simplified quotation marks
- amsmath mathematic equations
- fontawesome neat symbols
- xcolor defining colors
- hyperref crossreferencing and links
- graphicx including pictures
- biblatex bibliography
- tikz making fancy/complicated plots and pictures

## Packages for linguistics

#### \usepackage{...}

- gb4e or Covington glosses and example numbering
- · tipa IPA symbols (in 对距X you can input them directly)
- OTtablx OT Tableaux
- qtree & tree-dvips syntactic trees (many other alternatives)
- stmaryrd semantic brackets
- drs Discourse Representation Structures
- phonrule phonological rules

See also Dickinson and Herring (2008) and LaTeX/Linguistics - Wikibooks (2024).

## Package options

```
\usepackage[...]{ }
\usepackage[utf8]{inputenc}
\usepackage[english]{babel}
\usepackage[backend=biber, sorting=nyt, sortcites=true, indexing=cite, useprefix=false, maxcitenames=2, style=authoryear-comp]{biblatex}
```

# Scientific document structure

## Typical scientific document

| 1. Title                                                                              | witty or not, but informative                               |
|---------------------------------------------------------------------------------------|-------------------------------------------------------------|
| 2. Abstract                                                                           | tl;dr                                                       |
| 3. Introduction/background topic, background, why is this important, hypothesis, aim, |                                                             |
| 4. Methods                                                                            | description of how the research was conducted               |
| 5. Analysis                                                                           | data analysis info                                          |
| 6. Results                                                                            | study findings                                              |
| 7. Discussion                                                                         | interpretation of the results                               |
| 8. References                                                                         | works cited                                                 |
| 9. Appendices calculations,                                                           | supplementary material (items, detailed extra figures etc.) |



Wrap-up

#### Summary

- ✓ Knitting to PDF from R
- ✓ MEX document structure
- ✓ Scientific document structure
- ► More typesetting with 上X

Homework assignment

## Homework assignment due June 14th 15:30

- ② Complete assignment 7 (→ ILIAS)

  Create a basic X¬M¬EX document for the Noisy channel experiment
  (as on page 29). Upload the resulting files to ILIAS.
- **?** Read chapter 2 of "The Not So Short Introduction to  $\text{ET}_{\text{E}} X 2_{\varepsilon}$ " (Oetiker et al. 2023).

#### References

- Dickinson, Markus and Josh Herring (2008). ETeXfor Linguists.

  Accessed: 2024-06-07. URL:
- https://cl.indiana.edu/~md7/08/latex/slides.pdf. LaTeX/Linguistics - Wikibooks (2024). Accessed: 2024-06-07. URL:
- https://en.wikibooks.org/wiki/LaTeX/Linguistics.

  Munroe, Randall (2013). xkcd: File Transfer. Accessed: 2024-06-07.
- URL: https://xkcd.com/1301/.

  Oetiker, Tobias et al. (2023). The Not So Short Introduction to ETEX.

  Accessed: 2024-06-07, URL:
  - https://tobi.oetiker.ch/lshort/lshort.pdf.
- Tilburg Science Hub (2024). Accessed: 2024-06-07. URL: https://tilburgsciencehub.com.