Digital Research Toolkit for Linguists

Week 11: Typesetting linguistic documents with LaTeX

Anna Prysłopska

June 17, 2024

Psycholinguistics and Cognitive Modeling Lab

Homework

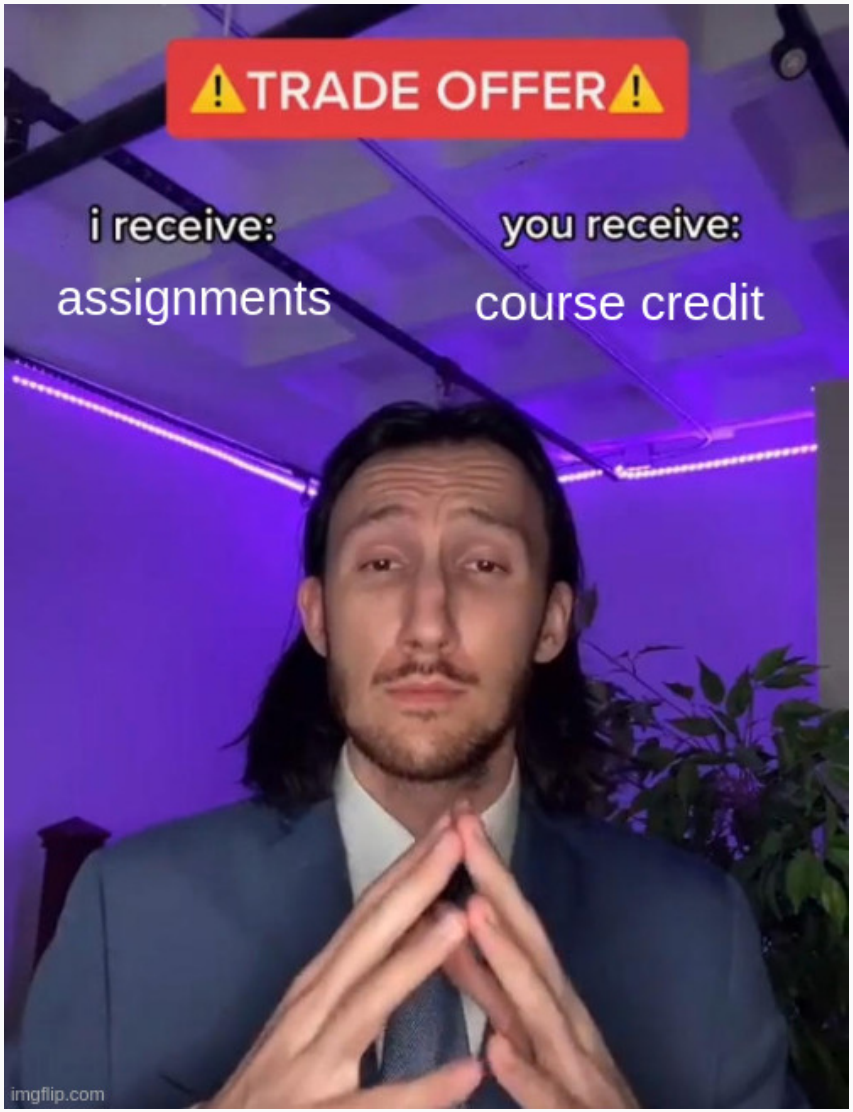
"Create a basic XeLaTeX document for the Noisy channel experiment (as on page 29 in the handout and the provided LaTeX files). Upload the resulting files to ILIAS as one compressed file (ZIP or otherwise). Use the scientific document structure."

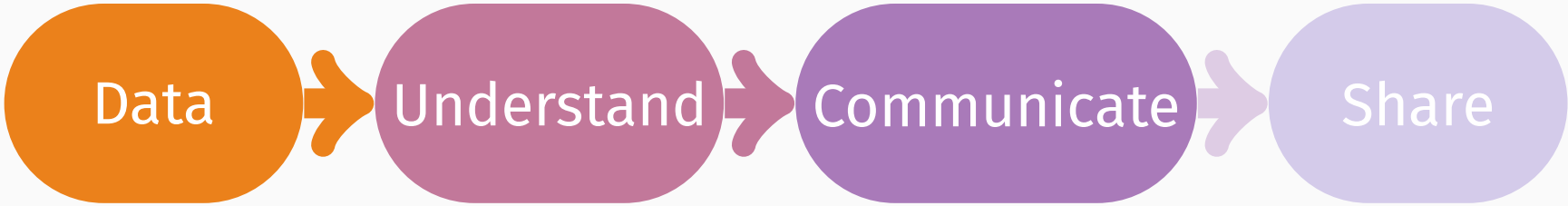
- ✖ No full document (e.g. missing discussion and analysis)
- ✖ Used AI without attribution → plagiarism
- ✖ No author
- ✖ One file, not **files**
- ? The experiment had 20, 23, 25, 27, 30 participants?

Questions?

Course credit

✔ complete $n - 2$ assignments ✖ complete $< n - 2$ assignments





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

LaTeX recap

General information

Name: ’latɛç or ˈlɑːtɛx or ˈleɪtɛx but NOT ˈleɪtɛks.

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

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

Our version: XeTeX & XeLaTeX

Input: Text file **TEX** + commands

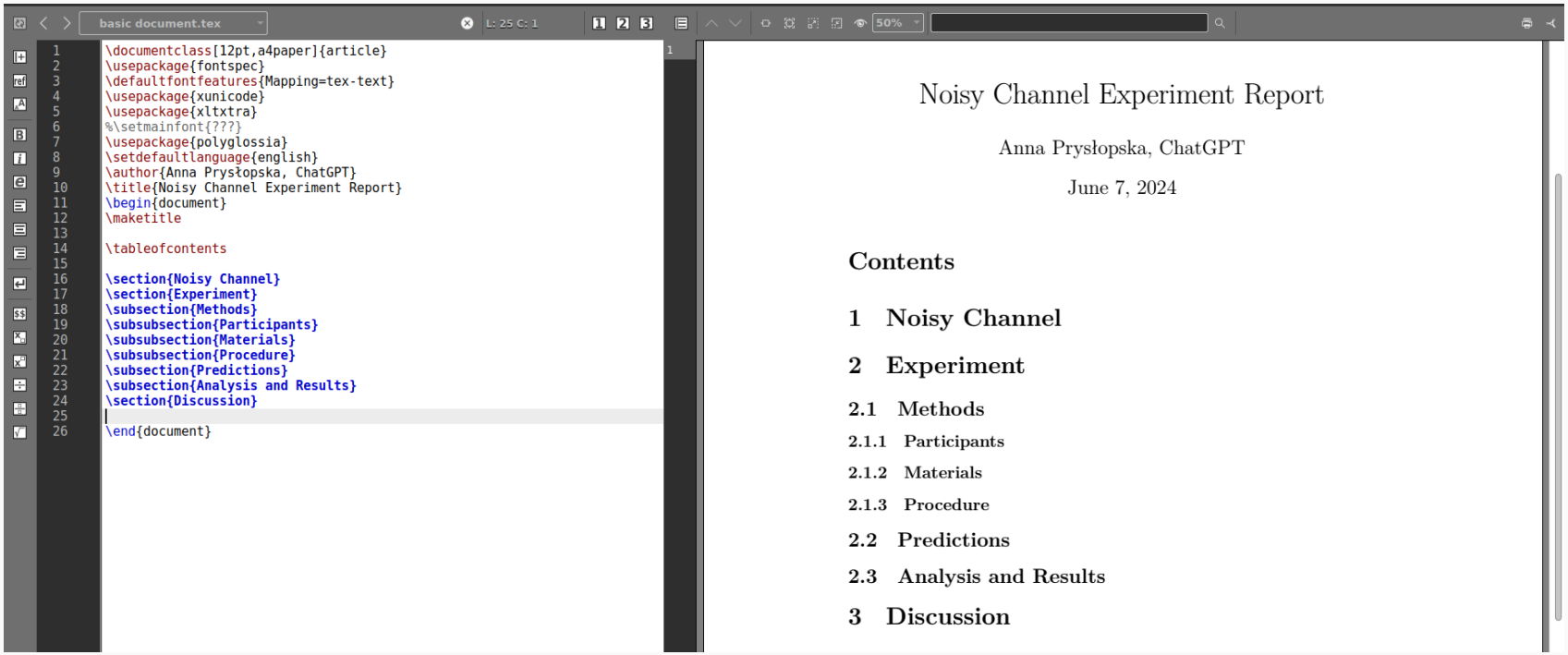
Output: **PDF** + **LOG** + **AUX** + ...

Compilation: XeLaTeX → XeLaTeX (→ biliography → XeLaTeX)

Document structure: (1) Document class. (2) Preamble. (3) Document.

First document

`\maketitle`	Creates title
`\section{…}`	Creates section
`\subsection{…}`	Creates subsection
`\subsection{…}`	Creates subsubsection



Editing text

Text structure

Numering

```
\part{…}
\chapter{…}
\section{…}
\subsection{…}
\subsubsection{…}
\paragraph{…}
\subparagraph{…}
```

→ will be included in TOC

No numering

```
\part*{…}
\chapter*{…}
\section*{…}
\subsection*{…}
\subsubsection*{…}
\paragraph*{…}
\subparagraph*{…}
```

→ will NOT be included in TOC

`\tableofcontents` = TOC

`\addcontentsline{WHERE}{LEVEL}{WHAT}`

WHERE = toc (table of contens)

LEVEL = part, chapter, section, …

WHAT = the exact text, e.g. “Abstract”

Commands

LaTeX commands typeset text elements:

`\chapter{Introduction}` = Make “Introduction” a chapter

The commands are case sensitive:

✔ `\chapter{…}` ✖ `\Chapter{…}` ✖ `\CHAPTER{…}`

Commands: General format

`\` start with a backslash

`a-z` letters only

✱ sometimes in “plain” and “stared” variants

⚟ some can take arguments and options:
`\command[optional parameter]{argument}`

⊗ no arguments → LaTeX ignores any whitespace after the command:

```
\section{Introduction}
\LaTeX is neat
\LaTeX{} is idiosyncratic
```

1. Introduction
LaTeXis neat
LaTeX is idiosyncratic

Make some space

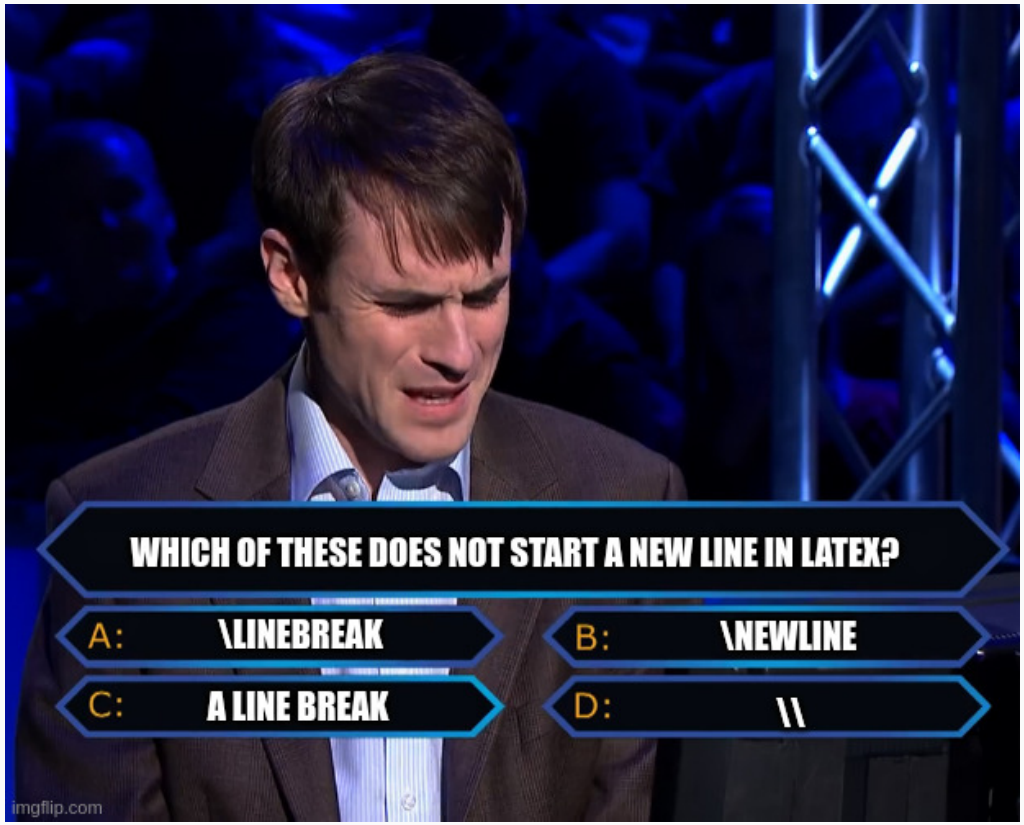
Whitespace

space, tab, ‘invisible’ characters (vertical and horizontal)

LaTeX assumes that **only one consecutive whitespace character** makes sense and will correct you if you add more.

Starting new lines with a space is a typographic sin and will be ignored.

Make some space



A **single line break** is considered one whitespace.
Two line breaks are considered a single paragraph break, as are **3+ line breaks**.

Gaps and breaks

`\hspace{…}`	Give me …px/cm/em/etc. horizontal space…
`\hspace*{…}`	…I MEAN IT
`\hfill`	Fill the page with whitespace horizontally
`\vspace{…}`	Give me …px/cm/em/etc. vertical space…
`\vspace*{…}`	…I MEAN IT
`\vfill`	Fill the page with whitespace vertically
`\newline` or `\linebreak`	Start a new line
`\\`	Start a new paragraph…
`\\*`	…but don't start a new page
`\newpage`	Start a new page, fill the last one with whitespace
`\pagebreak`	Start a new page, stretch the content on the last one
`\phantom{…}`	Pretend there is an invisible … here

Formatting text

Making text *feel* `very` special

Italics	`\textit{}` or `\emph{}`
Bold	`\textbf{}`
Underline	`\underline{}`
`Typewriter`	`\texttt{}`
SMALL CAPS	`\textsc{}`
Subscript	`$_{Subscript}$`
Superscript	`$^{\textrm{Superscript}}$`
Color	`\textcolor{red}{Color}`

Colors thanks to the package **`xcolor`**. You can define your own colors in the preamble or use preexisting ones:
`https://en.wikibooks.org/wiki/LaTeX/Colors`

Avoid ALL CAPS and SMALL CAPS

all words have the same shape and are difficult to distinguish

I HATE LATEX. AND WHEN I SAY THAT, I REALLY MEAN HATE. IT MAKES ME RANT AND RAVE AND FOAM AT THE MOUTH IN ANGER. SO HERE'S THE FIRST THING. I INSTALL LATEX. ALREADY IT'S OVER 1GB TO INSTALL. WHY? NO IDEA.

I hate latex. And when i say that, i really mean hate. It makes me rant and rave and foam at the mouth in anger. So here's the first thing. I install latex. Already it's over 1GB to install. Why? No idea.

Accessibility in Typography

Use **bold** and *italics* sparingly

it's easy to overemphasize and lose readers, while assistive technology doesn't emphasize yet

Avoid underline

underlined text is a link, like the source of that LaTeX rant

Use typewriter for URLs or emails no mistaking I for l or b-d

Accessibility in Typography

Avoid ~~strikethrough~~, superscript , and *subscript*
this text is read as usual by screen readers

Fonts and their sizes

Relative to option in `\documentclass[]{}`

Set globally `\large` or locally `{\large }`

`\tiny`	tiny
`\scriptsize`	scriptsize
`\footnotesize`	footnotesize
`\small`	small
`\normalsize`	normalsize
`\large`	large
`\Large`	Large
`\LARGE`	LARGE
`\huge`	huge
`\Huge`	Huge

Cross References: hyperref

`\url{URL}`	clickable URL
`\href{URL}{TEXT}`	clickable URL hyperlink with custom text
`\label{KEY}`	given name
`\ref{KEY}`	reference
`\pageref{KEY}`	page number[1]

[1] This is a footnote. `\footnote{This is a footnote.}`

Cross References: Convention

Conventional key structure (but you can use any **unique** name)

Key	Usage	Example
`ex:foo`	example	sentence in 2a
`fig:foo`	figure	Figure 1
`sec:foo`	section	Section 6
`tab:foo`	table	Table 1
`li:foo`	list	List 9
	missing reference	**??**

Environments

Lists

Format blocks of text (e.g. lists, text alignment, tables, figures, poetry).
Start with `\begin{…}` and end with `\end{…}`.
Must be contained and not empty.

```
\begin{itemize}
\item First
\item[+] Second
\item[--] Third
\end{itemize}
```

- First
+ Second
– Third

```
\begin{enumerate}
\item First
\item[ii.] Second
\item[(3)] Third
\end{enumerate}
```

1. First
ii. Second
(3) Third

Alignment

```
\begin{flushleft}
'Twas brillig, and the slithy toves
Did gyre and gimble in the wabe:
All mimsy were the borogoves,
And the mome raths outgrabe.
\end{flushleft}
```

'Twas brillig, and the slithy toves Did gyre and gimble in the wabe: All mimsy were the borogoves, And the mome raths outgrabe.

```
\begin{center}
“Beware the Jabberwock, my son!
The jaws that bite, the claws that catch!
Beware the Jubjub bird, and shun
The frumious Bandersnatch!”
\end{center}
```

“Beware the Jabberwock, my son! The jaws that bite, the claws that catch! Beware the Jubjub bird, and shun The frumious Bandersnatch!”

```
\begin{flushright}
He took his vorpal sword in hand;
Long time the manxome foe he sought---
So rested he by the Tumtum tree
And stood awhile in thought.
\end{flushright}
```

He took his vorpal sword in hand; Long time the manxome foe he sought— So rested he by the Tumtum tree And stood awhile in thought.

Tabular and tables

```
\begin{tabular}{|l|c|r|}
\hline
A table & With & No caption \\
\hline
A & a & i \\
B & b & ii \\
\hline
\end{tabular}
```

A table	With	No caption
A	a	i
B	b	ii

```
\begin{table}
\begin{tabular}{lcr}
A table & With & A caption \\
\hline
A & a & i \\
B & b & ii \\
\end{tabular}
\caption{Neat}\label{tab:neat}
\end{table}
```

A table	With	A caption
A	a	i
B	b	ii

Table 1: Neat

Generators in editor or e.g. `https://www.tablesgenerator.com/`.

Images and figures: graphix

```
\includegraphics[scale=1]{cat.jpg}
```

Figure 1: Dog

```
\begin{figure}
\includegraphics
[width=1\textwidth]{dog.jpg}
\caption{Dog}\label{fig:dog}
\end{figure}
```

Language support

`\usepackage[utf8]{inputenc}`
input encoding
`\usepackage[english]{babel}`
document language
`\usepackage{csquotes}`
quotation marks

”Funny” joke	`"Funny" joke`
“Funny” joke	``` ``Funny'' joke ```
„Funny” joke	`,,Funny'' joke`
“Funny” joke	`\enquote{Funny} joke`
‘Funny’ joke	`\enquote*{Funny} joke`

Hyphenation is automatic, but can malfunction for technical terms. You can specify hyphenation of a word in the preamble

```
\hyphenation{keepittogether make-big-chunks
su-per-ca-li-fra-gi-lis-tic-ex-pi-a-li-do-cious}
```

Special characters

Some characters have special uses

- `&` in tables,
- `\` in commands,
- `{ }` in commands,
- `_` for subscript,
- `^` for superscript,
- `$` in math,
- `%` to leave comments,
- `#` for parameters and alignment,
- `~` for spacing

LaTeX will complain if you use them willy-nilly.

Special characters

Special characters must be prefixed/escaped with `\` to render them.

```
  & \&
  \ \textbackslash
{ } \{ \}
  _ \_
  ^ \^
  $ \$
  % \%
  # \#
  ~ \~
```

There are more special characters: “-” hyphen, “–” en-dash, “—” em-dash, “−”minus etc. You can copy & paste most in XeLaTeX or look them up:
`http://tug.ctan.org/info/symbols/comprehensive/symbols-a4.pdf`

Glosses

Numbered examples: gb4e

Make this the last package you load because it can conflict with other packages.

```
\begin{exe}
\ex
\gll Holla die Waldfee\\
     holla the wood.fairy\\
\glt Well, I never!
\end{exe}
```

(1) Holla die Waldfee
holla the wood.fairy
Well, I never!

Numbered examples: gb4e

```
\begin{exe}
\ex \begin{xlist}
\ex \gll Alter Schwede\\
         old Swede\\
\end{xlist}
\end{exe}
```

(2) a. Alter Schwede
old Swede

Numbered examples: gb4e

```
\begin{exe}
\ex
\gll \# die Familienvater\\
{} the$_{fem}$ family.father\\
\glt the family father
\end{exe}
```

(3) # die Familienvater
the$_{fem}$ family.father
the family father ‘

IPA symbols

Accents and IPA with `tipa`

XeLaTeX àcçêpts möst ðiñgs (as long as your font has them!) but you might want to typeset more complex linguistic stuff.

```
\textipa{f@"nEtIks}
```

fə'nɛtɪks

Symbol name	*Macro name*	*Symbol*
Turned A	`\textturna`	ɐ
Glottal stop	`\textglotstop`	ʔ
Right-tail D	`\textrtaild`	ɖ
Small capital G	`\textscg`	ɢ
Hooktop B	`\texthtb`	ɓ
Curly-tail C	`\textctc`	ɕ
Crossed H	`\textcrh`	ħ
Old L-Yogh ligature	`\textOlyoghlig`	ɮ
Beta	`\textbeta`	β

Full documentation:
`http://www.l.u-tokyo.ac.jp/~fkr/tipa/tipaman.pdf`

Semantic formulae

Math, symbols, and semantics: amsmath, amssymb, stmaryrd

Math mode is an easy way of typesetting mathematic formulae. Switching from (default) text mode to math mode (and back) is done by using: `$ ... $`. Everything in between is parsed as math (cf. p. 10).

$$\frac{a^2 + b^2 \neq \delta}{\bigstar}$$

```
$\dfrac{a^2 + b^2 \neq \delta}{\bigstar}$
```

Semantic brackets require the package `stmaryrd`.

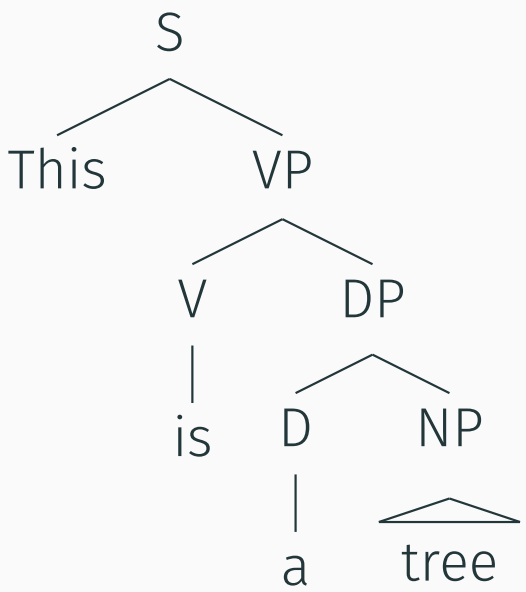
$$[\![be_3]\!] \equiv \lambda P^i \lambda x^s \exists x^k [\forall x^o [P^i(x^o) \Leftrightarrow R'(x^o, x^k)] \wedge R(x^s, x^k)]$$

```
$\llbracket be_3 \rrbracket \equiv
\lambda P^i \lambda x^s \exists x^k [\forall x^o
[P^i (x^o) \Leftrightarrow R'(x^o,x^k)]
\land R(x^s,x^k)]$
```

Syntactic trees

Growing trees with `qtree`

Many other packages out there (`xyling`, `xy`, `tikz-qtree`, `forest`).
You can add arrows with `tree-dvips`.




```
\Tree [.S This
    [.VP [.V is ]
    [.DP [.D a ]
        \qroof{tree}.NP ] ] ]
```

Questions?

Summary

- ✔ basic LaTeX commands
- ✔ scientific document structure
- ✔ typography in LaTeX
- ✔ creating tables
- ✔ including plots
- ✔ cross-referencing and hyperlinking
- ✔ LaTeX for linguists
- ⏩ bibliography management and large projects

Homework assignment

Homework assignment due June 21 at 15:30

- Complete assignment 9 ($\rightarrow$ ILIAS)
- Read “Bibliography management with biblatex”: `https://www.overleaf.com/learn/latex/Bibliography_management_with_biblatex`