



## 2. Installing NooJ for Java

on LINUX, macOS, UNIX and Windows computers

### A step-by-step guide

This series of tutorials is based upon work from COST Action  
**Multi3Generation CA18231**, supported by COST  
(European Cooperation in Science and Technology).

COST (European Cooperation in Science and Technology) is a funding agency for research and innovation networks. Our Actions help connect research initiatives across Europe and enable scientists to grow their ideas by sharing them with their peers. This boosts their research, career and innovation, cf. [www.cost.eu](http://www.cost.eu)

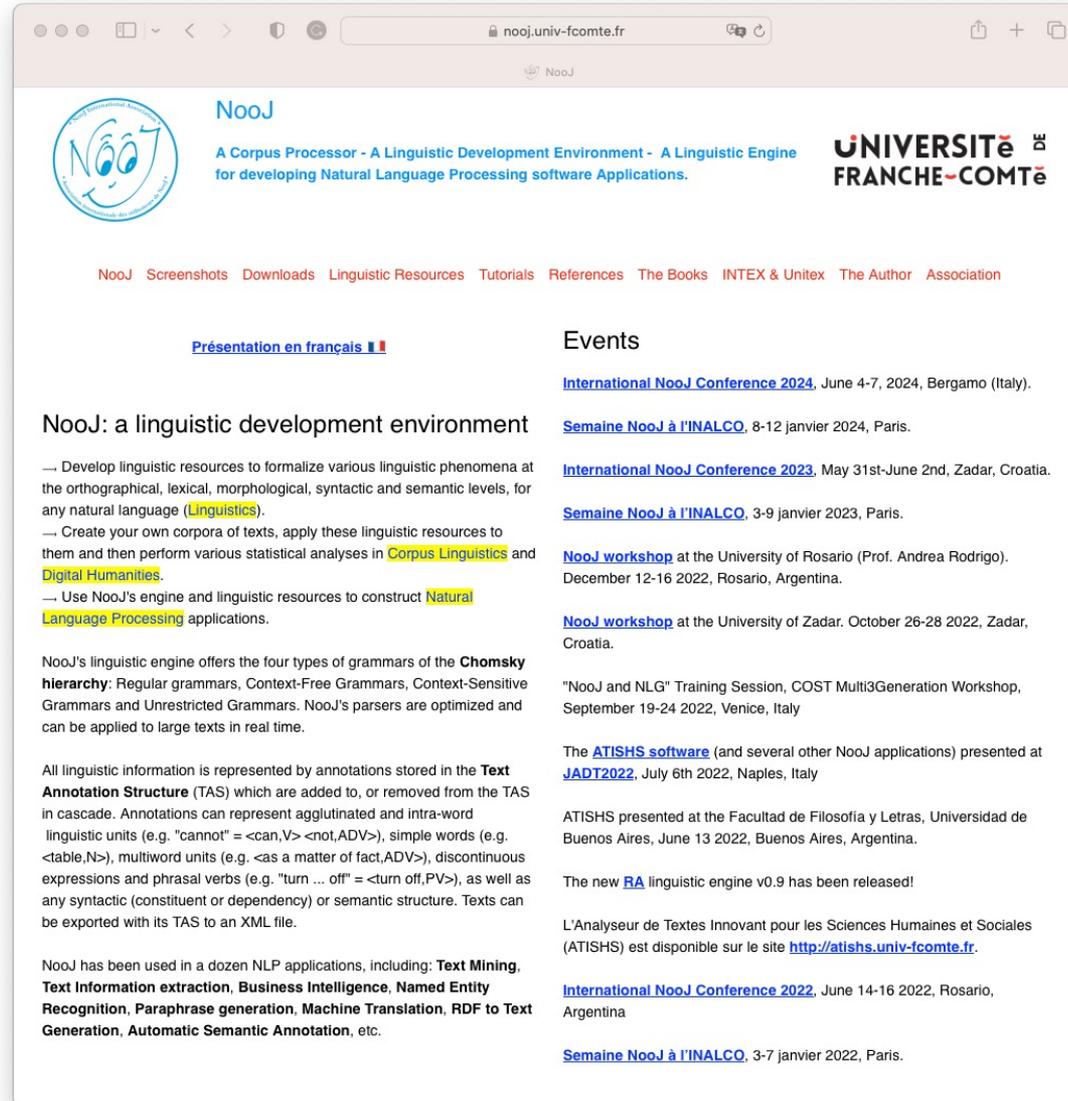
# Installing NooJ

## The four variants

1. NooJ for Windows: the complete toolbox. Runs on Windows (.NET technology).
2. NooJ for Java: a simplified version of NooJ for demos and students of Software Engineering: well-documented open-source (JAVA technology)
3. ATISHS for macOS: a French application in the Digital Humanities, various statistical tools (macOS technology)
4. RA for Linux, macOS, and Windows: the new engine, used in the WEB application <https://webnooj.univ-fcomte.fr> (Swift, Vapor technology)

# NooJ for Java: Step 1

Go to <https://nooj.univ-fcomte.fr>



The screenshot shows the homepage of the NooJ website. At the top left is the NooJ logo, a blue circle with a stylized face and the text 'NooJ'. To its right is the text 'NooJ' and a description: 'A Corpus Processor - A Linguistic Development Environment - A Linguistic Engine for developing Natural Language Processing software Applications.' On the top right is the logo of 'UNIVERSITÉ FRANCHE-COMTÉ'. Below the header is a navigation menu with links: 'NooJ', 'Screenshots', 'Downloads', 'Linguistic Resources', 'Tutorials', 'References', 'The Books', 'INTEX & Unitex', 'The Author', and 'Association'. The main content area is divided into two columns. The left column has a link 'Présentation en français' and a section 'NooJ: a linguistic development environment' with three bullet points describing its capabilities. The right column has an 'Events' section with several links to conferences and workshops, including 'International NooJ Conference 2024', 'Semaine NooJ à l'INALCO', and 'NooJ workshop' at various universities.

**NooJ**  
A Corpus Processor - A Linguistic Development Environment - A Linguistic Engine  
for developing Natural Language Processing software Applications.

UNIVERSITÉ FRANCHE-COMTÉ

[NooJ](#) [Screenshots](#) [Downloads](#) [Linguistic Resources](#) [Tutorials](#) [References](#) [The Books](#) [INTEX & Unitex](#) [The Author](#) [Association](#)

[Présentation en français](#)

## NooJ: a linguistic development environment

- Develop linguistic resources to formalize various linguistic phenomena at the orthographical, lexical, morphological, syntactic and semantic levels, for any natural language ([Linguistics](#)).
- Create your own corpora of texts, apply these linguistic resources to them and then perform various statistical analyses in [Corpus Linguistics](#) and [Digital Humanities](#).
- Use NooJ's engine and linguistic resources to construct [Natural Language Processing](#) applications.

NooJ's linguistic engine offers the four types of grammars of the **Chomsky hierarchy**: Regular grammars, Context-Free Grammars, Context-Sensitive Grammars and Unrestricted Grammars. NooJ's parsers are optimized and can be applied to large texts in real time.

All linguistic information is represented by annotations stored in the **Text Annotation Structure (TAS)** which are added to, or removed from the TAS in cascade. Annotations can represent agglutinated and intra-word linguistic units (e.g. "cannot" = <can,V> <not,ADV>), simple words (e.g. <table,N>), multiword units (e.g. <as a matter of fact,ADV>), discontinuous expressions and phrasal verbs (e.g. "turn ... off" = <turn off,PV>), as well as any syntactic (constituent or dependency) or semantic structure. Texts can be exported with its TAS to an XML file.

NooJ has been used in a dozen NLP applications, including: **Text Mining**, **Text Information extraction**, **Business Intelligence**, **Named Entity Recognition**, **Paraphrase generation**, **Machine Translation**, **RDF to Text Generation**, **Automatic Semantic Annotation**, etc.

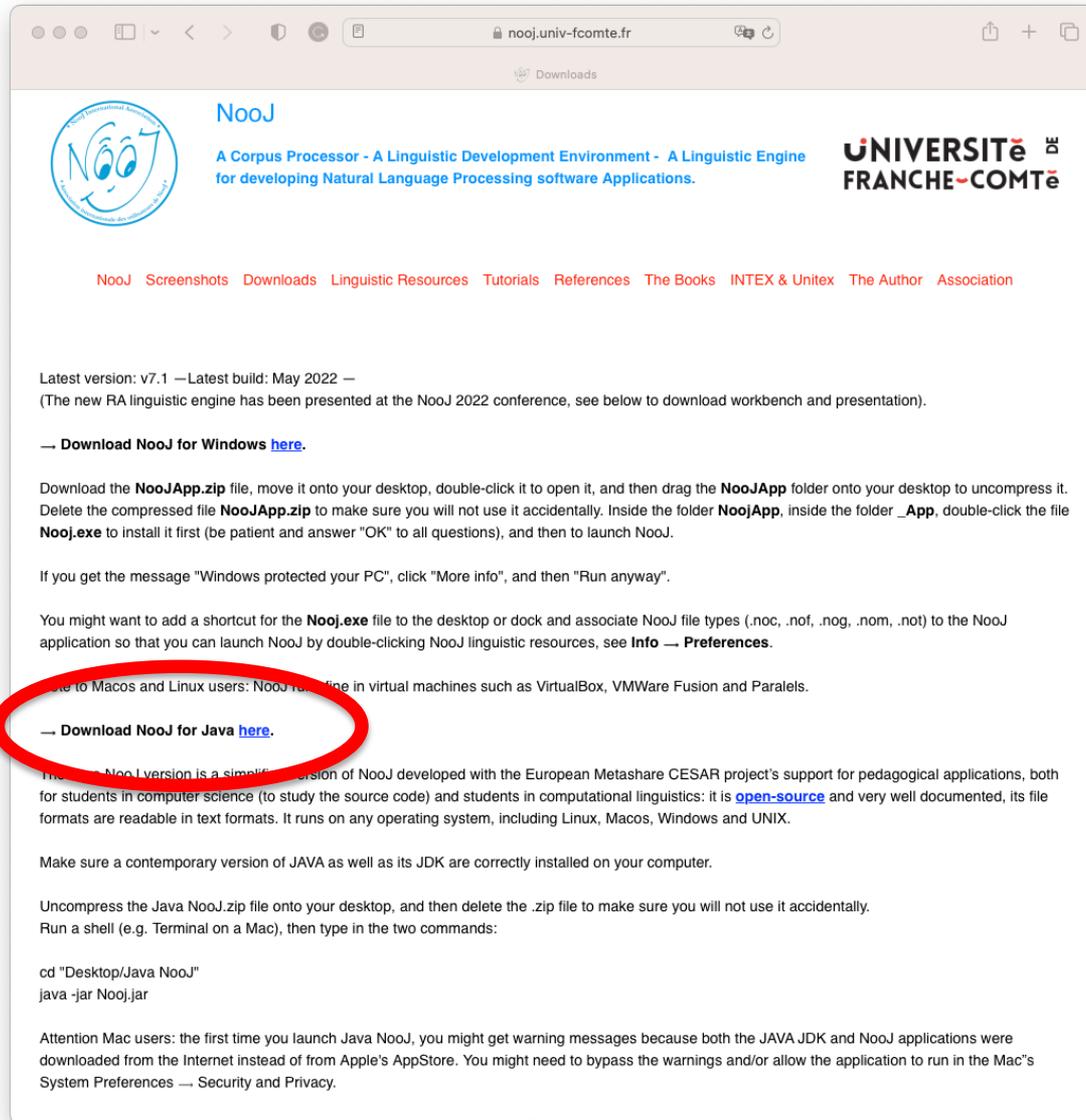
## Events

- [International NooJ Conference 2024](#), June 4-7, 2024, Bergamo (Italy).
- [Semaine NooJ à l'INALCO](#), 8-12 janvier 2024, Paris.
- [International NooJ Conference 2023](#), May 31st-June 2nd, Zadar, Croatia.
- [Semaine NooJ à l'INALCO](#), 3-9 janvier 2023, Paris.
- [NooJ workshop](#) at the University of Rosario (Prof. Andrea Rodrigo). December 12-16 2022, Rosario, Argentina.
- [NooJ workshop](#) at the University of Zadar. October 26-28 2022, Zadar, Croatia.
- "NooJ and NLG" Training Session, COST Multi3Generation Workshop, September 19-24 2022, Venice, Italy
- The [ATISHS software](#) (and several other NooJ applications) presented at [JADT2022](#), July 6th 2022, Naples, Italy
- ATISHS presented at the Facultad de Filosofía y Letras, Universidad de Buenos Aires, June 13 2022, Buenos Aires, Argentina.
- The new [RA](#) linguistic engine v0.9 has been released!
- L'Analyseur de Textes Innovant pour les Sciences Humaines et Sociales (ATISHS) est disponible sur le site <http://atishs.univ-fcomte.fr>.
- [International NooJ Conference 2022](#), June 14-16 2022, Rosario, Argentina
- [Semaine NooJ à l'INALCO](#), 3-7 janvier 2022, Paris.

# NooJ for Java: Step 2

## Go to “Downloads” page

Click here  
to download  
file:  
**Java NooJ.zip**



nooj.univ-fcomte.fr

Downloads

 **NooJ**  
A Corpus Processor - A Linguistic Development Environment - A Linguistic Engine  
for developing Natural Language Processing software Applications.

UNIVERSITÉ DE  
FRANCHE-COMTÉ

[NooJ](#) [Screenshots](#) [Downloads](#) [Linguistic Resources](#) [Tutorials](#) [References](#) [The Books](#) [INTEX & Unitex](#) [The Author](#) [Association](#)

Latest version: v7.1 — Latest build: May 2022 —  
(The new RA linguistic engine has been presented at the NooJ 2022 conference, see below to download workbench and presentation).

— [Download NooJ for Windows here.](#)

Download the **NooJApp.zip** file, move it onto your desktop, double-click it to open it, and then drag the **NooJApp** folder onto your desktop to uncompress it. Delete the compressed file **NooJApp.zip** to make sure you will not use it accidentally. Inside the folder **NooJApp**, inside the folder **\_App**, double-click the file **NooJ.exe** to install it first (be patient and answer “OK” to all questions), and then to launch NooJ.

If you get the message “Windows protected your PC”, click “More info”, and then “Run anyway”.

You might want to add a shortcut for the **NooJ.exe** file to the desktop or dock and associate NooJ file types (.noc, .nof, .nog, .nom, .not) to the NooJ application so that you can launch NooJ by double-clicking NooJ linguistic resources, see [Info](#) — [Preferences](#).

For MacOS and Linux users: NooJ can be run in virtual machines such as VirtualBox, VMWare Fusion and Paralels.

— [Download NooJ for Java here.](#)

The Java NooJ version is a simplified version of NooJ developed with the European Metashare CESAR project’s support for pedagogical applications, both for students in computer science (to study the source code) and students in computational linguistics: it is [open-source](#) and very well documented, its file formats are readable in text formats. It runs on any operating system, including Linux, MacOS, Windows and UNIX.

Make sure a contemporary version of JAVA as well as its JDK are correctly installed on your computer.

Uncompress the Java NooJ.zip file onto your desktop, and then delete the .zip file to make sure you will not use it accidentally.  
Run a shell (e.g. Terminal on a Mac), then type in the two commands:

```
cd "Desktop/Java NooJ"  
java -jar Nooj.jar
```

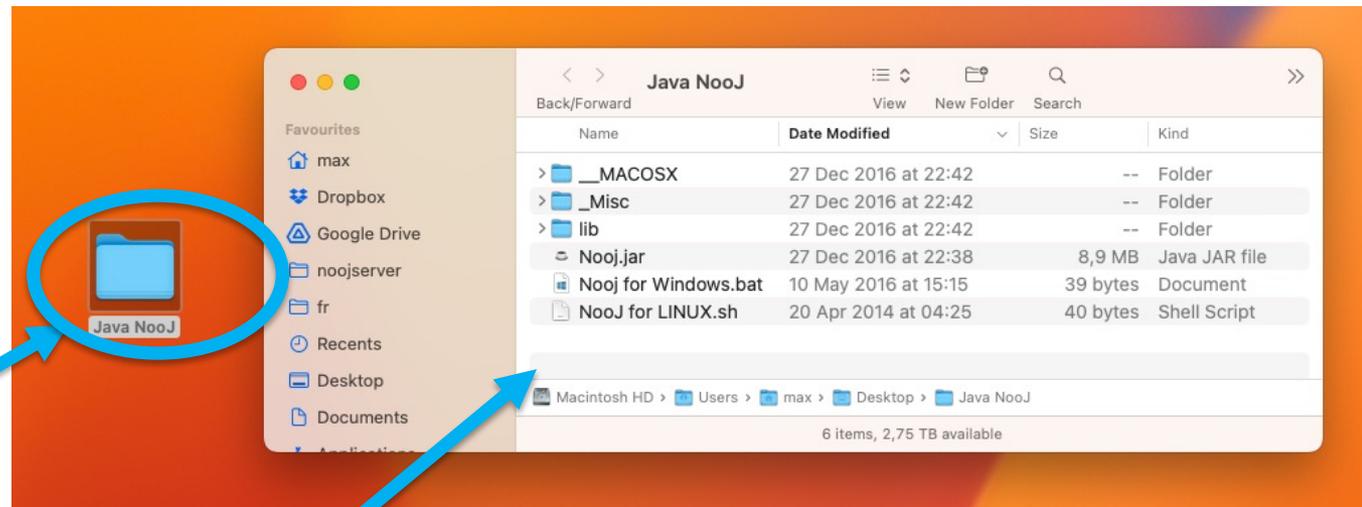
Attention Mac users: the first time you launch Java NooJ, you might get warning messages because both the JAVA JDK and NooJ applications were downloaded from the Internet instead of from Apple’s AppStore. You might need to bypass the warnings and/or allow the application to run in the Mac’s System Preferences — Security and Privacy.

RA: A new linguistic engine for NooJ. Download workbench [here](#). Check out RA sources [here](#).

# NooJ for Java: Step 3

Uncompress the **Java NooJ.zip** package

- Decompress (Unzip)  
**Java NooJ.zip**  
to get folder **JavaNooJ**  
Drag folder to Desktop



- Double click folder JavaNooJ to open it.  
Check that the file **Nooj.jar** is present.

# NooJ for Java: Step 4

## run Nooj.jar

- run the **Terminal** App
- Enter the command **cd "Desktop/Java NooJ"**
- Enter the command **java -jar Nooj.jar**



```
~/Desktop/Java NooJ — 74x13
~/Desktop/Java NooJ — zsh
[~ % cd "Desktop/Java NooJ"
Desktop/Java NooJ % java -jar Nooj.jar
```

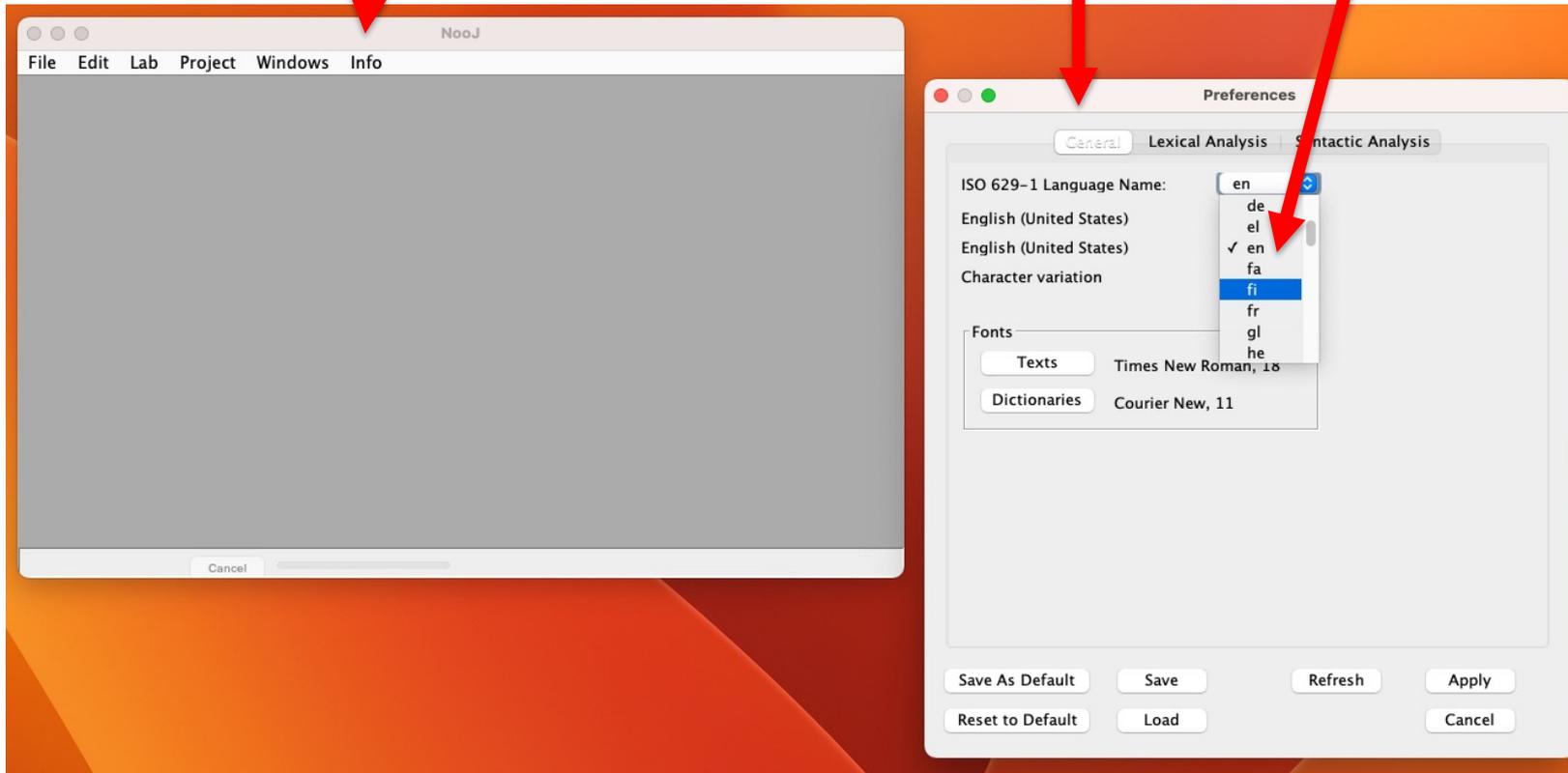
# NooJ for Java: Step 5

## Set Default Language

- To set the default language, click **Info > Preferences**

- Go to the **General** tab

- Select Language, e.g. "en" for English



# NooJ for Java: Step 6

## Test NooJ

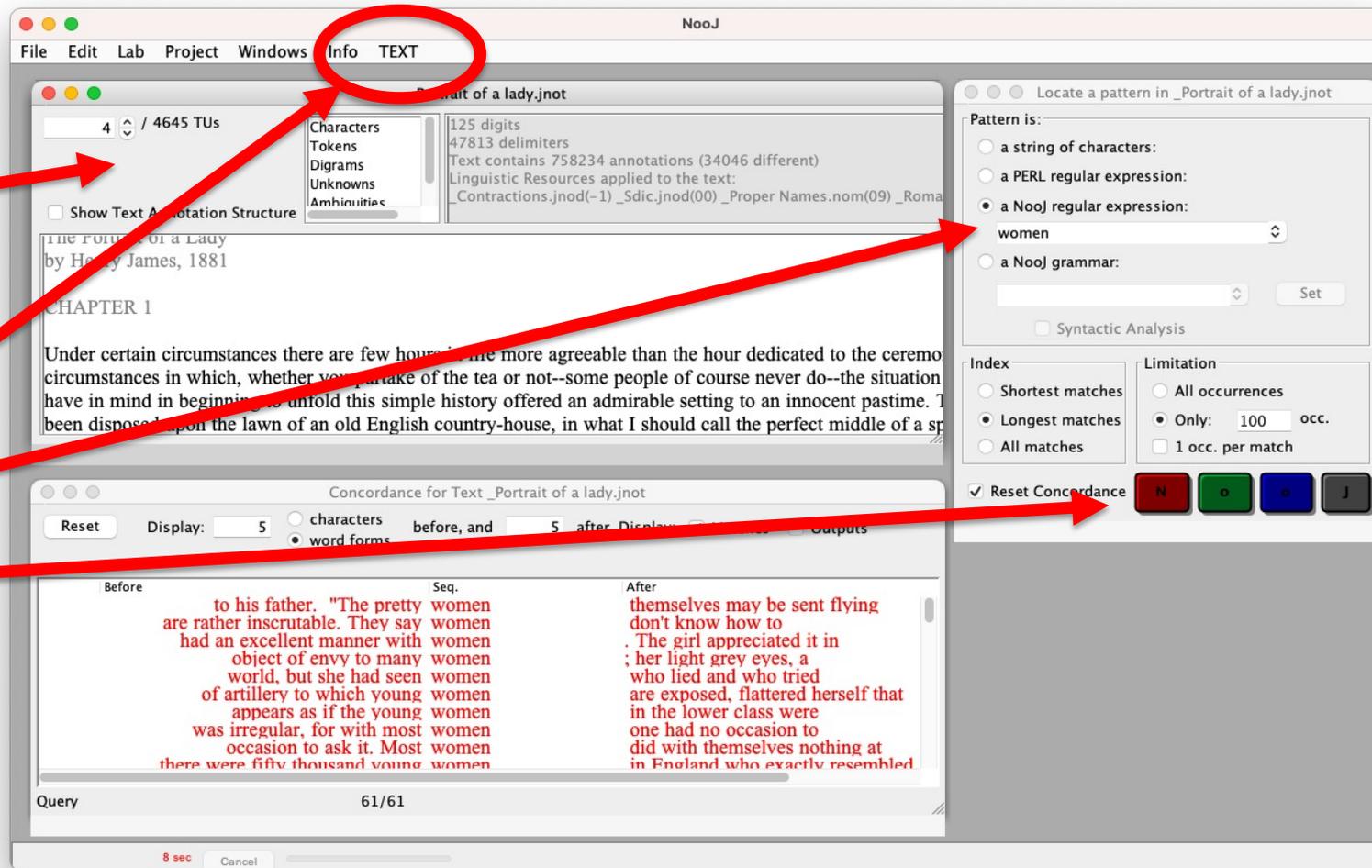
If the “en” language was selected:

- click **File > Open > Text**, then select “\_The portrait of a lady” to open text.

- click **TEXT > Locate**

- Enter **women**

- click the red button to compute the concordance for wordform **women**



# Useful Information

- You might want to move the **Java NooJ** folder from the desktop to another folder, e.g., **/Applications**. Make sure you know how to manage read and execute rights, so that all users can still run NooJ, and (of course) modify the first command (**cd "Java NooJ"**) accordingly.
- You might want to create a bash, csh or zsh shell script that runs the two commands, and create a shortcut to this shell, so that you do not need to navigate to **Java NooJ** everytime to run NooJ.
- Do not mess with the content of the **Java NooJ** folder: never move, edit, rename or even touch any file that are the **Java NooJ** folder. Always work in your personal **Documents/ONooJ** folder: all your files should be stored in your personal folder.
- Some file names start with the « **\_** » character. Never modify these files, as they might get modified, updated or even deleted during a subsequent update of the NooJ software. If you need to modify their content, just save the resulting files with another name that does not contain the « **\_** » prefix.

# Problems?

- macOS does not like software that do not come from the Apple store. The first time you run NooJ, it will produce a warning.
  1. Click the Apple icon (top left of screen), then select **Settings > Privacy & Security**
  2. Select the software, then click **Open**
- If the java command does not work, Java is not installed on your computer.
  - Go to: <https://www.oracle.com/java/technologies/downloads/>
  - Select Operating System (Linux, macOS or Windows)
  - Select your computer hardware (ARM64 or x64) and installer
- An anti-virus does not like to download software from the internet.
  - temporarily unload antivirus.
- Other problem?
  - Contact [max.silberztein@gmail.com](mailto:max.silberztein@gmail.com)



# CONGRATULATIONS



You have successfully installed

**NooJ for Java**

(Now you have to read the documentation)

