



4. Installing RA Workbench

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

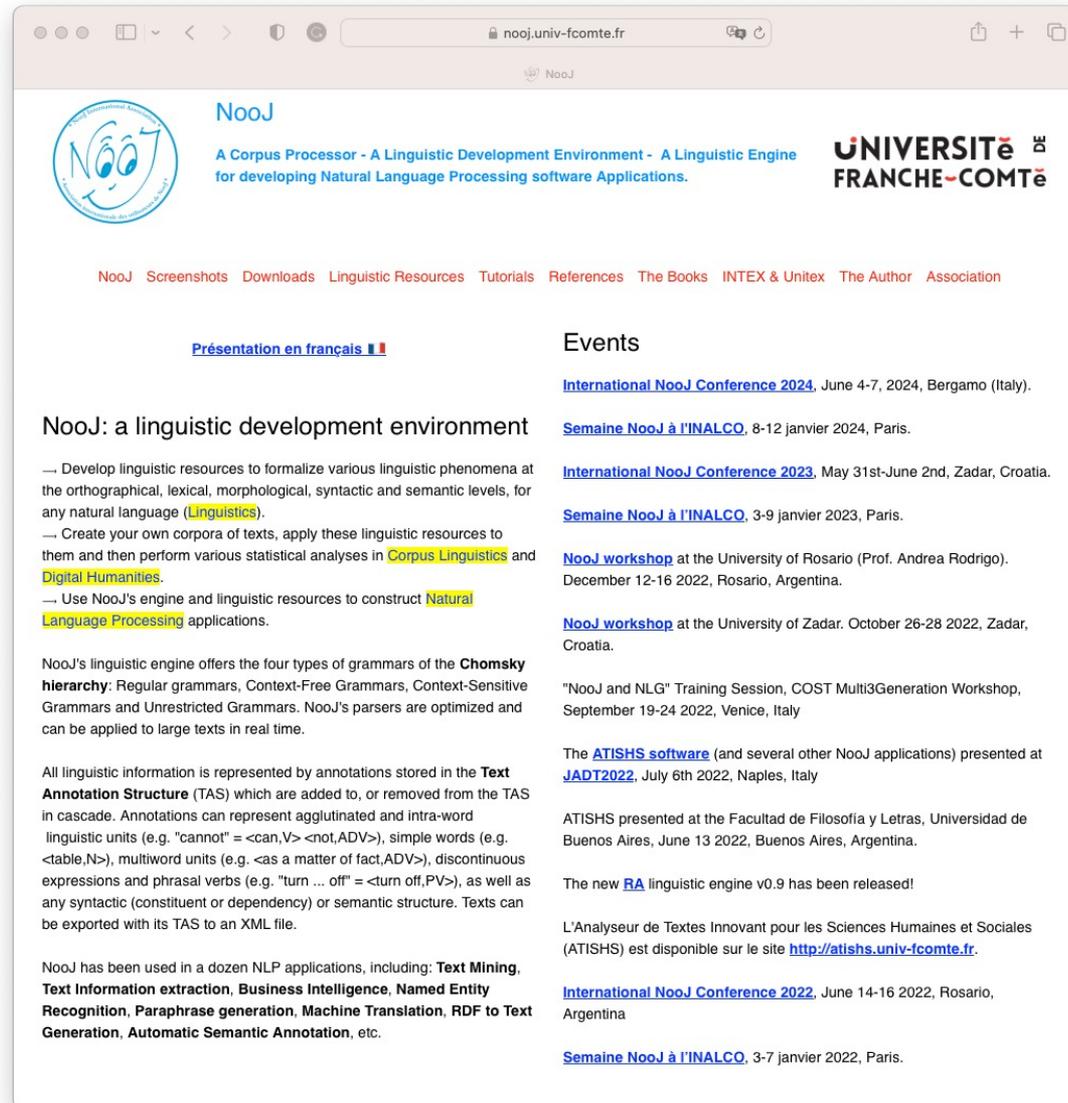
Installing NooJ

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)

RA Workbench: Step 1

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



The screenshot shows the homepage of the NooJ website. At the top, there 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 features a section titled "Présentation en français" with a sub-header "NooJ: a linguistic development environment". Below this, there are three bullet points describing the environment's capabilities: developing linguistic resources, creating corpora, and using the engine for natural language processing applications. A paragraph explains the Chomsky hierarchy supported by NooJ. Another paragraph details the Text Annotation Structure (TAS) and its use in representing linguistic information. A final paragraph lists various NLP applications where NooJ has been used. The right column is titled "Events" and lists several conferences and workshops, including the International NooJ Conference 2024, Semaine NooJ à l'INALCO, and NooJ workshop at the University of Rosario and Zadar. At the bottom, there is a link to the ATISHS software and a note about the release of the RA linguistic engine v0.9.

 **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>.

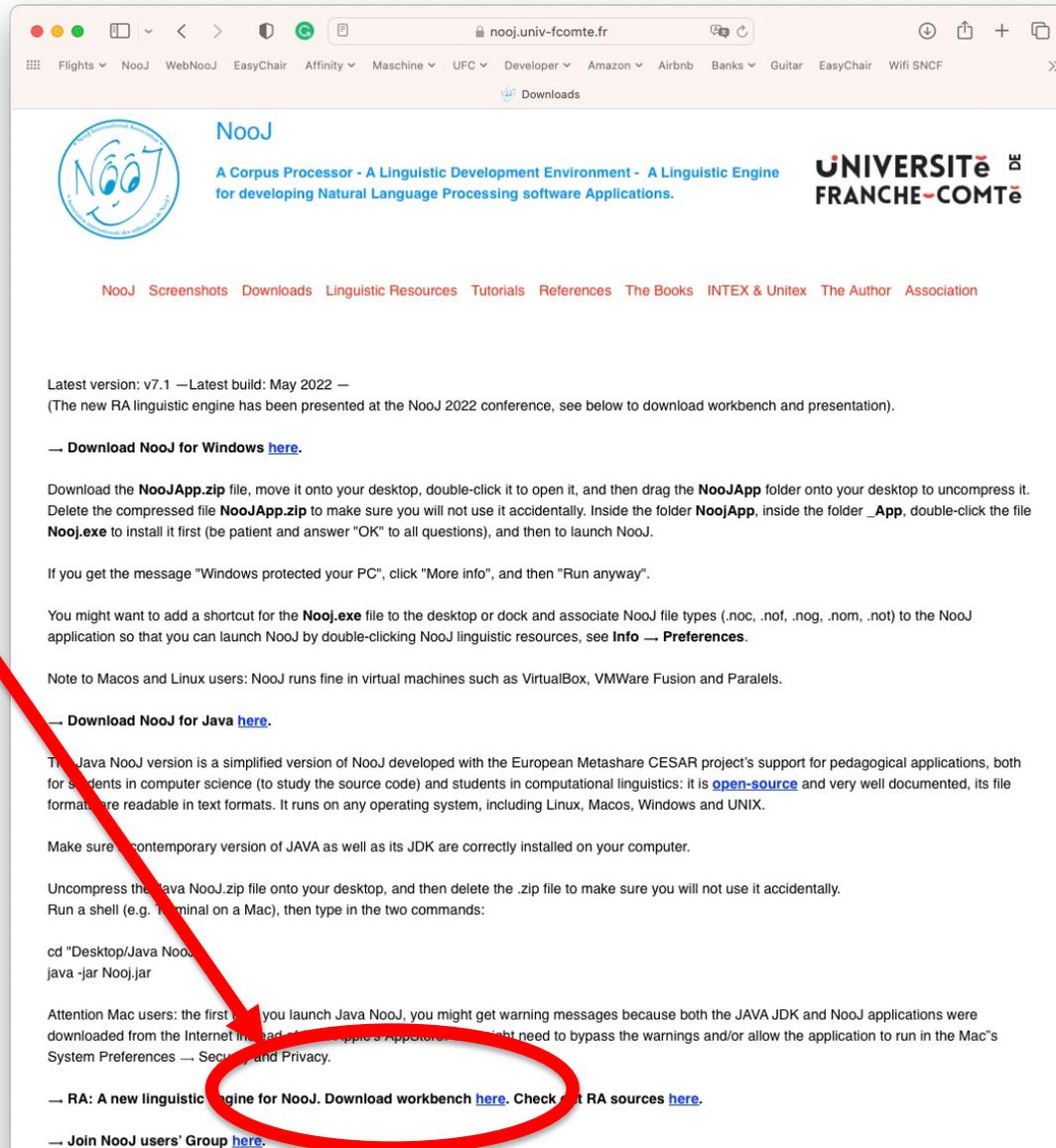
[International NooJ Conference 2022](#), June 14-16 2022, Rosario, Argentina

[Semaine NooJ à l'INALCO](#), 3-7 janvier 2022, Paris.

RA Workbench: Step 2

Go to “Downloads” page

Click here
to download
file:
RA-Workbench.zip



nooj.univ-fcomte.fr

Flights NooJ WebNooJ EasyChair Affinity Maschine UFC Developer Amazon Airbnb Banks Guitar EasyChair Wifi SNCF

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 & Unisex](#) [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](#).

Note to MacOS and Linux users: NooJ runs fine 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
```

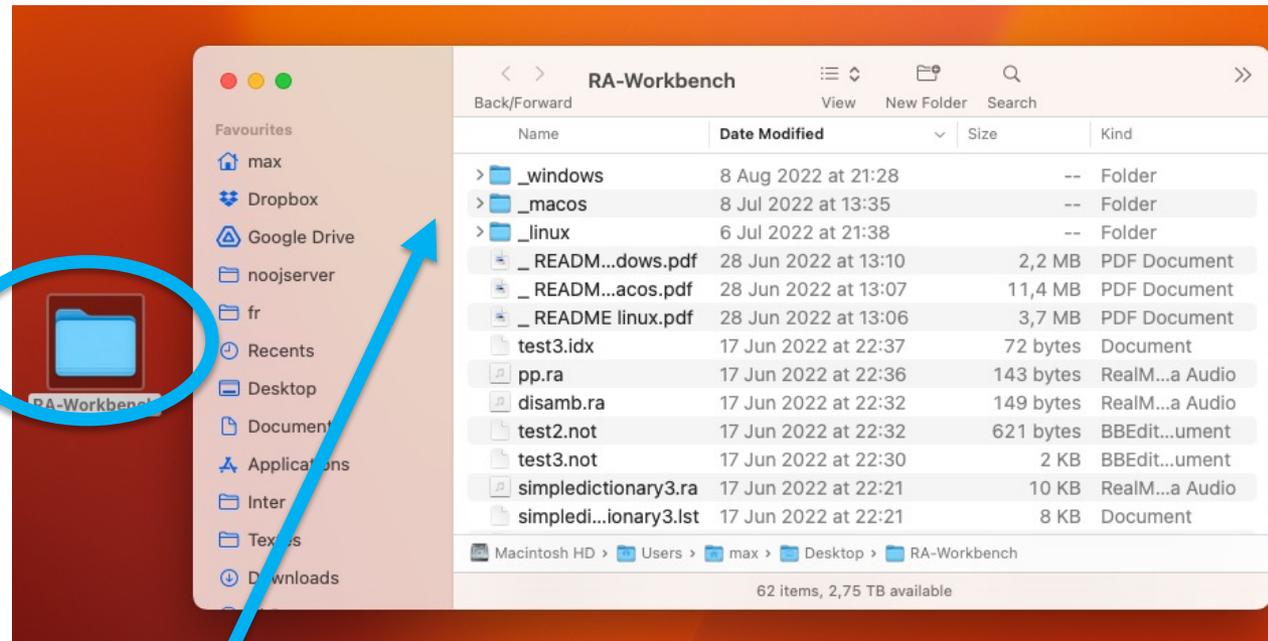
Attention Mac users: the first time you launch Java NooJ, you might get warning messages because both the JAVA JDK and NooJ applications were downloaded from the Internet instead of the App Store. You might need to bypass the warnings and/or allow the application to run in the Mac's System Preferences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

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

— [Join NooJ users' Group here](#).

RA Workbench: Step 3

Uncompress the **RA-Workbench.zip** package



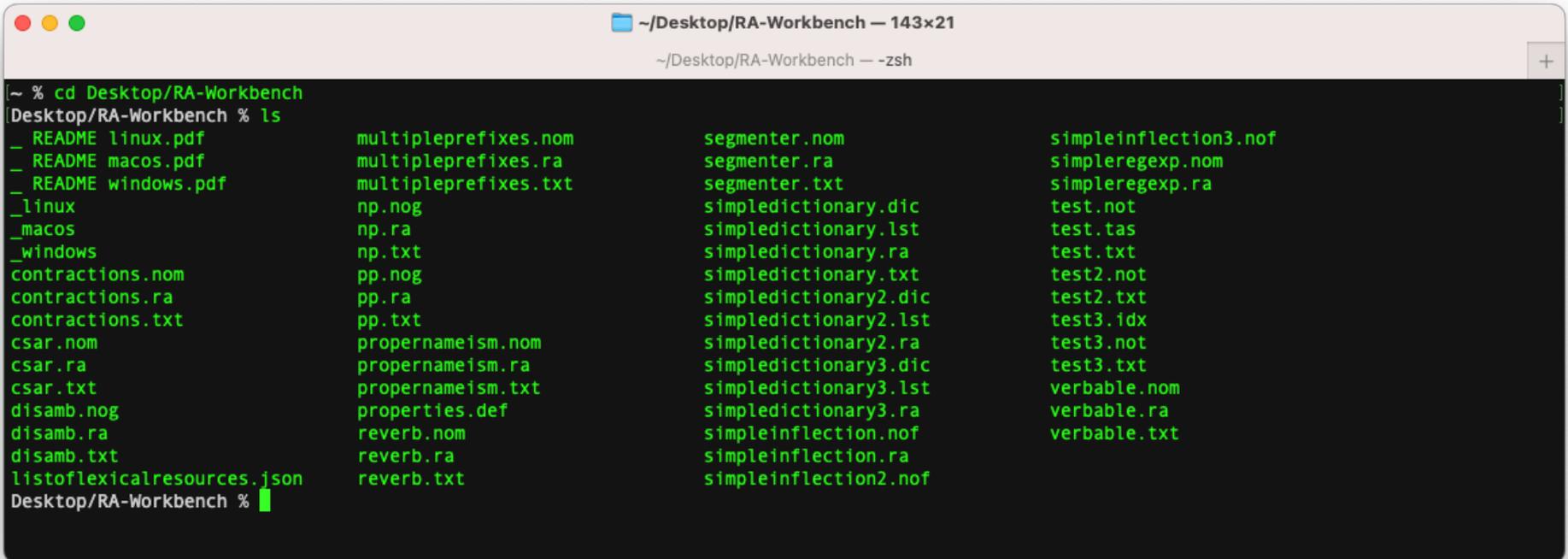
- Decompress (Unzip)
RA-Workbench.zip
to get folder **RA-Workbench**
Drag folder to Desktop

- Double click folder **RA-Workbench** to open it.
Check that the three folders `_windows`, `_macos`, `_linux` are present.

RA Workbench: Step 4

Check **RA** content

- run the **Terminal** App
- Enter the command **cd "Desktop/RA-Workbench"**
- Enter the command **ls** to display the content of the workbench



```
~/Desktop/RA-Workbench — 143x21
~/Desktop/RA-Workbench — -zsh
~ % cd Desktop/RA-Workbench
Desktop/RA-Workbench % ls
_ README linux.pdf      multipleprefixes.nom  segmenter.nom         simpleinflection3.nof
_ README macos.pdf      multipleprefixes.ra   segmenter.ra          simpleregexp.nom
_ README windows.pdf   multipleprefixes.txt  segmenter.txt         simpleregexp.ra
_ linux                 np.nog                simpledictionary.dic  test.not
_ macos                 np.ra                  simpledictionary.lst  test.tas
_ windows               np.txt                 simpledictionary.ra   test.txt
contractions.nom       pp.nog                 simpledictionary.txt  test2.not
contractions.ra        pp.ra                  simpledictionary2.dic test2.txt
contractions.txt       pp.txt                 simpledictionary2.lst test3.idx
csar.nom               propernameism.nom     simpledictionary2.ra  test3.not
csar.ra                propernameism.ra     simpledictionary3.dic test3.txt
csar.txt               propernameism.txt    simpledictionary3.lst verbable.nom
disamb.nog             properties.def         simpledictionary3.ra  verbable.ra
disamb.ra              reverb.nom            simpleinflection.nof  verbable.txt
disamb.txt             reverb.ra             simpleinflection.ra
listoflexicalresources.json reverb.txt             simpleinflection2.nof
```

RA Workbench: Step 5

Test RA

A few commands:

\$ **cat test3.txt** to display the sample text **test3.txt**

\$ **_macos/segment test.txt segmenter.ra** to segment text into sentences and create **test3.not**

\$ **_macos/lexicalanalysis test.not simplifiedictionary3.ra properties.def** to lookup dictionary and add annotations to **test3.not**

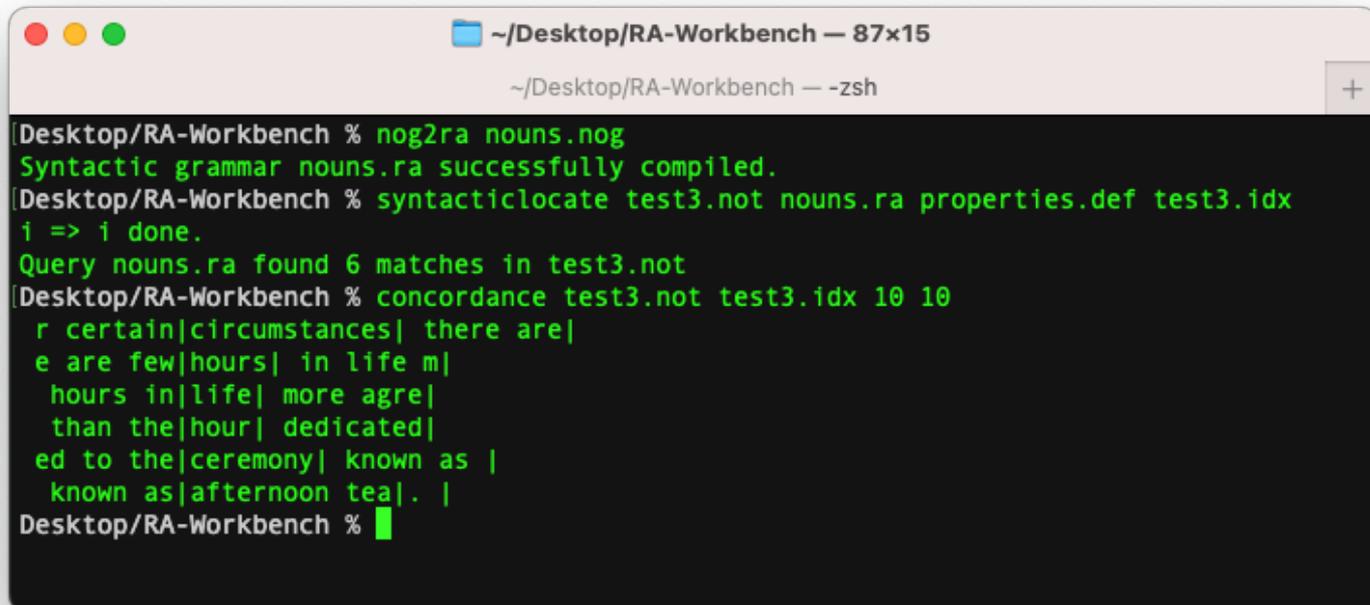
\$ **_macos/printtas test3.not** to display the Text Annotation Structure of text **test3.not**

\$ **echo "Main = <N> ; " > nouns.nog** to create the syntactic grammar **nouns.nog**

\$ **_macos/nog2ra nouns.nog** to compile the syntactic grammar into finite-state machine **nouns.ra**

\$ **_macos/syntacticlocate test3.not nouns.ra properties.def test3.idx** to construct the index of all nouns and save it in **test3.idx**

\$ **_macos/concordance test3.not test3.idx 10 10** to display the concordance of all nouns



```
~/Desktop/RA-Workbench — 87x15
~/Desktop/RA-Workbench — zsh
Desktop/RA-Workbench % nog2ra nouns.nog
Syntactic grammar nouns.ra successfully compiled.
Desktop/RA-Workbench % syntacticlocate test3.not nouns.ra properties.def test3.idx
1 => 1 done.
Query nouns.ra found 6 matches in test3.not
Desktop/RA-Workbench % concordance test3.not test3.idx 10 10
r certain|circumstances| there are|
e are few|hours| in life m|
hours in|life| more agre|
than the|hour| dedicated|
ed to the|ceremony| known as |
known as|afternoon tea|. |
Desktop/RA-Workbench % █
```

Problems?

- If the system does not find the RA commands, add the corresponding directory to the \$PATH variable, *e.g.*, in **.profile** or **.zshrc** file in your home directory, add the line:
 - for macOS: **PATH="~/Desktop/RA-Workbench/_macos:\$PATH"**
 - for LINUX: **PATH="~/Desktop/RA-Workbench/_linux:\$PATH"**
- During the unzip process, some operating systems reset files read and/or execute permissions.
 1. Get to the corresponding directory, *e.g.*, **cd _macos**
 2. Change permissions, *e.g.*, **chmod u+rx ***
- Other problem?
 - Contact max.silberztein@gmail.com



CONGRATULATIONS



You have successfully installed
RA Workbench

(Now you have to read the documentation)

