

Nooj's Chinese module

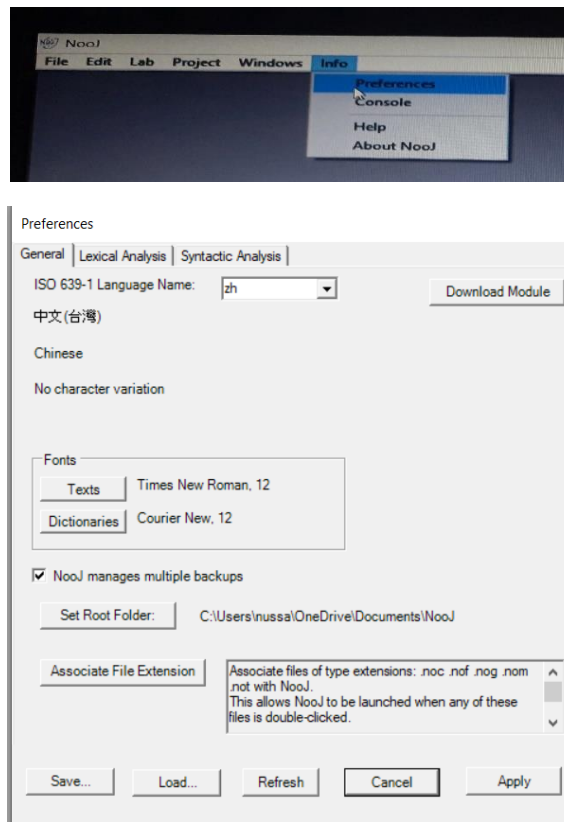
To install the module, unzip the file zh.zip in your Documents/NooJ folder so that it contains the three folders: “Lexical Analysis”, “Projects” and “Syntactic Analysis” as in:

Documents/NooJ/zh/Lexical Analysis

Documents/NooJ/zh/Projects

Documents/NooJ/zh/Syntactic Analysis

To process Chinese texts, select the language “zh” in **Info** → **Preferences** as shown in the following screenshots:



This module contains three folders:

1. In the "Lexical Analysis" folder:

- the electronic dictionary “dictionary.dic” contains 63,567 entries. This dictionary is associated with morphological grammar “derivation.nof”.
- Spelling Mistakes.dic contains 421 frequent mistakes`
- Transport ZH-FR.dic is an example of a ZH to FR bilingual dictionary

When compiled, this dictionary can be selected in **Info → Preferences → Lexical Analysis** as "dictionary.nod".

2. In the "Project" folder:

- Corpus.nod contains 481,871 characters.
- Disamb.nod is a project for the grammar disambiguation which is applied to the corpus.
- ENPN.nod is a demonstration of the ENPN grammar that recognizes named entities in the text.

3. In the "Syntactic Analysis" folder:

- Classifiers.nog recognizes all numbers below 1 billion as well as sequences Verb-Nombre-Classifieur
- Discontinuous expressions.nog recognizes expressions that accept insertions
- Economy.nog recognizes technical terms in Economy
- ENPN.nog recognizes named entities of persons
- Health.nog recognizes names of disease
- Legal.nog recognizes compounds words and expressions of law and justice.
- Reduplication.nog is an example of grammar that recognizes reduplication in Chinese
- Spelling Mistakes.nog recognizes the spelling mistakes in the corresponding dictionary and produce their correction
- Transport ZH-FR.nog recognizes technical terms in Transport and produces their French translation

For any information, please contact: zhencai1122@hotmail.com