

Extra EXERCISE Normalization - INSTRUCTIONS:

TOOL (Creating a new TEI-XML File).

- Follow the steps as in exercise 1 or reuse the TEI-XML structure you already have created (for now, just the structure, not the text).

TRANSCRIPTION

Transcribe everything as it appears on the old print page, also the special character, the long S (f).

- Transcribe from the scan “Extra.3-Lope-sonnet-1707.pdf” the sonnet in French or use the transcription provided “Extra.2-Lope-sonnet-1707-Transcript.txt”. (We don’t transcribe the hyphenation, gue-re. Have a look into the transcript).
- To add the special character f in oXygen: Edit -> Insert from Character Map...; Search for ‘long s’.
- The ampersand (&) needs to be escaped, because it is used precisely to mark special characters. So, instead of using &, into oXygen you should transcribe `&` (as an editor you could also decide to normalize it with “et” before the encoding).

ENCODING

We are encoding a normalized version of the transcribed version. Take into consideration some of the steps from exercise 1 for the sonnet structure and/or reuse the TEI-XML file you already have created.

- Start just tagging the difference f/s (fait/sait).
- You need to add these elements:

`<choice>`: groups a number of alternative encodings. `<orig>`: original form. `<reg>`: regularization.

- It should look like:

```
<choice>
  <orig>fait</orig>
  <reg>sait</reg>
</choice>
```

- Feeling strong? Keep tagging the normalized version: e.g. defelpere/désespère, pouvois/pouvais, hazard/hasard, desja/déjà, joye/joie, etc.
- In order to know which word to tag, compare the original with the normalized version provided “Extra.2-Lope-sonnet-1707-TranscriptNormalized.txt”.

Already done?

The first letter in the original is highlighted, **D**oris. Do we encode it?

- Search in the TEI Guidelines 3.3.2.2 Emphatic Words and Phrases for a solution looking for ‘hi’ `<hi>` and the attribute `@rend`.

The poem has also a title.

- We need the element `<head>` with the attribute `@title`

VISUALIZATION (Web browser and oXygen editor mode).

- Follow the steps as in exercise 1.
- The CSS stylesheet will show the regularized version. In order to visualize the original version: open the file “visualize-tei-xml.css” in oXygen; place the cursor on line 105 `/*reg {display: none}*/`; go to Document -> Edit -> Toggle Comment; place the cursor on line 101 `orig {display: none;}`; go to Document -> Edit -> Toggle Comment. Save the changes and refresh the page in the browser.

- Tip! You can just remove/add the characters (opening) `/*` and `*/` (closing) to toggle the code. Don't forget to save the changes.

EXERCISE SOLUTION

- I provided you with a TEI-XML file already encoded “Extra.4-Lope-sonnet-1707-done.xml”. Feel free to have a look inside or ask me for help.