

# TP1 : oXygen

TEI@Oxford

Jan 2011

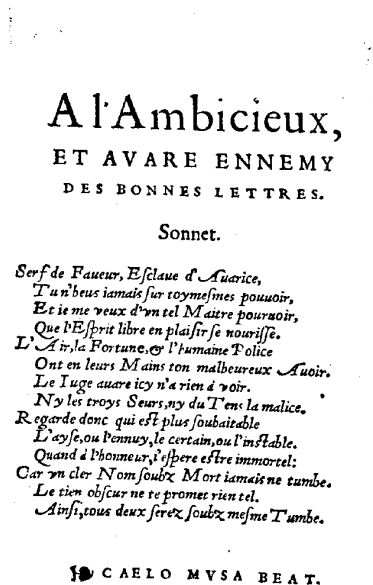
## 1 À la découverte d' Oxygen - 2

Dans ce deuxième exercice, nous allons approfondir votre expérience avec Oxygen pour :

- l'encodage XML d'un document existant
- la représentation des structures aperçues dans un texte littéraire
- l'affichage en mode « auteur »

## 2 Sample text

Dans cet exercice, nous allons baliser ce sonnet de Du Bellay, un écrivain du XVI<sup>e</sup> siècle :

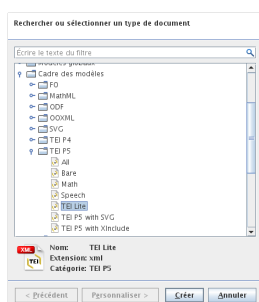


Pour vous simplifier la vie, on a déjà préparé une transcription prête à l'emploi... ne la regardez pas tout de suite !

## 3 Création du document

Ouvrez Oxygen. Cliquez sur l'icône Nouveau en haut à gauche, (ou sélectionnez Nouveau du menu Fichier, ou tapez CTRL-N.) pour ouvrir la boîte de dialogue Nouveau

- Cette fois, choisissez Cadre des modèles, ensuite TEI P5, et ensuite TEI-Lite, pour sélectionner un schéma TEI plus complet.



## 6 STRUCTURATION DU DOCUMENT

- Comme auparavant, Oxygen vous propose un squelette du document ; vous observerez que le schéma TEI lite décrit un document TEI plus élaboré avec un choix de balises plus important.
- Complétez l'entête TEI avec des indications de titre ('Un sonnet du Du Bellay numérisé'), distribution, et source ('encodé à partir d'un facsimilé de la première édition produit au CESR, Tours')
- Enlevez toute la partie proposée par défaut pour le `<body>` (mais retenez les balises `<body>` et `</body>`).

## 4 Ajout de texte au document

Rien ne vous empêche de taper à la main tout le poème directement. Mais pour gagner du temps on vous propose la démarche suivante :

- Assurez-vous que le curseur soit toujours entre les deux balises `<body>` et `</body>` de votre document vide
- Dans le menu Document, sélectionnez Fichier, et ensuite Insérer un fichier
- Naviguez jusqu'au fichier duBellay.txt dans le dossier Travaux qui devrait se trouver sur votre Bureau, et insérez-le. Ce fichier est aussi disponible sur le web, à cet URL )

## 5 Attention, Will Robinson !

Votre document est rempli de taches rouges!

```
1 <?xml version="1.0" encoding="UTF-8"?>
2 <doc type="text">
3 A FORTIBUS,
4 ET AVARE SIBIBUS
5 DES BONNES LETTRES
6 Sonnet.
7
8
9 Serf de Faveur, Esclave d'Avare,
10 Tu n'heus jamais sur toi-même pouvoir,
11 Et je me vœux d'un tel Maître pourvoir,
12 Que l'Esprit libre en plaisir se nourisse.
13
14
15 L'AR, la Fortune, & l'humaine Police
16 Ont en leurs mains ton malheureux Avoir
17 Le Jug' avare n'y a rien à voir
18 Ny les trois Seurs, ny du Tens la malice.
19
20
21 Regarde donc qui est plus souhaitable
22 L'ave, ou l'emvy, le certain, ou l'instable
23 Quand à l'honneur, j'espere estre immortel.
24
25
26 Car un cher Nôm souz Mort jamais ne tombe
27 Le tien abour ne te promet rien tel.
28 Mais, tous deux s'ens' souz meisme Tumba,
29
30
31 CAELO MYSA BEAT.
32
33 </doc>
34
```

**[Error]** The entity name must immediately follow the '@' in the entity reference.  
Grille Auteur

On va régler cela petit à petit. Commençons avec la partie la plus sérieuse : où le texte lui-même est affiché en rouge.

- Oxygen vous indique que la partie en rouge n'est pas bien formée
- Il y a un message en bas de l'écran qui vous aidera (un peu) à identifier le problème. Sauriez vous comment le résoudre ?
- Conseil : souvenez-vous du fait que l'esperluette est un caractère magique en XML, qui doit donc être représentée de manière indirecte.

## 6 Structuration du document

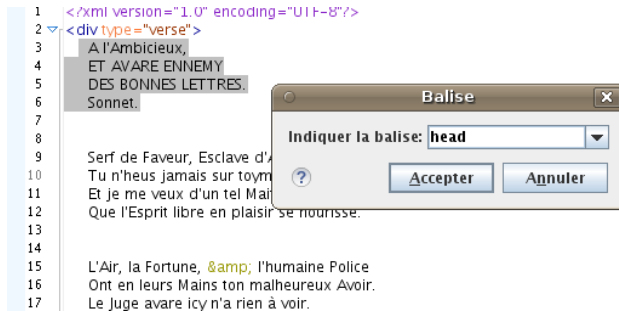
Il est possible d'identifier dans ce document plusieurs sous-parties. En particulier, il contient :

- un titre ... à baliser avec la balise `<head>`
- quatre strophes ... à baliser avec la balise `<lg>`
- des vers ... à baliser avec la balise `<l>`

- une citation d'Horace ... à baliser avec la balise `<quote>`

Allons-y !

- Avec la souris, sélectionnez les quatre premières lignes du document 'À l'ambitieux et avare ... sonnet'; ceux-ci constituent son titre.
- Tapez CTRL-E (ou sélectionnez XML Refactoring et ensuite Entourer des balises dans le menu Document
- Oxygen vous propose toutes les balises disponibles à cet emplacement.
- Sélectionnez `<head>` et cliquez Accepter



- Ce sonnet est bicéphale : coupez en deux parties le `<head>` existant.. La première partie est le titre. La seconde partie est le mot sonnet. Chacune de ces deux parties devrait être balisée avec un `<head>` au début et à la fin, en vous plaçant avant ou après le mot 'Sonnet' et en vous servant de CTRL-ALT-D ou d'une autre manière comme lors du premier exercice.

## 7 Balisage du poème

Enfin, il faut travailler un peu sur l'étiquetage du poème lui-même. Notre but serait de baliser chaque vers, se servant de la balise TEI `<l>` ('line') et en amont de baliser chaque ensemble de vers constituant une strophe avec la balise TEI `<lg>`.

Bien sur, il y a plusieurs manières de faire. On va vous décrire une méthode assez rapide, mais si vous préfériez une autre démarche, on n'insiste pas !

- Sélectionnez avec la souris tout le texte du poème, jusqu'à la citation d'Horace 'caelo musa beat'.
- Entourez toute le poème d'une balise `<l>`.
- Entourez la citation d'Horace d'une balise `<quote>`
- Le petit carré vert apparaît ! Avons-nous terminé ? Hélas non : un document peut être valide, tout en contenant des mensonges ! Voyez-vous des mensonges ? Cliquez sur le bouton Indentation (ou tapez CTRL-SHIFT-P).

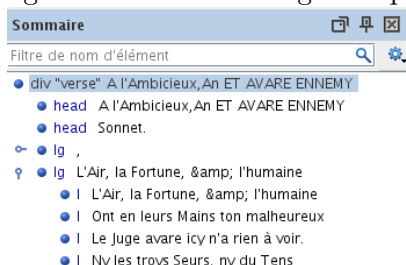
```
<body>
<head>A l'Ambitieux, ET AVARE ENNEMY DES BONNES LETTRES. Sonnet.</head>
<l>Serf de Faveur, Esclave d'Avarice, Tu n'heus jamais sur toymesmes pouvoir, Et je me veux
d'un tel Maître pourvoir, Que l'Esprit libre en plaisir se nourisse. L'Air, la Fortune,
&amp; l'humaine Police Ont en leurs Mains ton malheureux Avoir. Le Juge avare icy n'a rien à
voir. Ny les troys Seurs, ny du Tens la malice. Regarde donc qui est plus souhaitable
L'ayse, ou l'ennuy, le certain, ou l'instable. Quand à l'honneur, j'espere estre immortel:
Car un cler Non souzb Mort jamais ne tombe. Le tien obscur ne te promet rien tel. Ainsi,
tous deux serez souzb mesme Tumble. </l>
<quote> CAELO MUSAE BEAT.</quote>
</body>
```

Parce qu'elles ne sont pas explicités par le balisage, la distinction entre chaque vers et entre chaque strophe ne sont plus affichées. Il faut donc les baliser. (Si vous ne voyez pas bien où indiquer ces séparations, vous pouvez ré-établir l'affichage originel en tapant CTRL-Z)

- Mettez le curseur au début de chaque vers, et distinguez-le du précédent, en vous servant du bouton `division` ou en tapant ALT-SHIFT-D
- Si cela vous semble plus rapide, vous pouvez aussi ‘tricher’ : copiez la séquence `</l><l>`, et collez-la au commencement de chaque vers.
- Quand vous aurez fini, cliquez encore sur le bouton `Indentation`. Ou sont les strophes ?
- Entourez tous les vers (tous les `<l>`) avec une balise `<lg>`.
- Divisez cette strophe unique en quatre strophes distinctes. Attention : il faut effectuer cette distinction au bon endroit — entre la fin d’un vers et le début du suivant, (i.e. après une balise fermante et avant une balise ouvrante).

### 8 Affichage du texte balisé

Est-ce que la disparition graduelle de notre beau sonnet au dessus d’un nuage de balises commence à vous inquiéter un peu ? Cela nous permet au moins de visualiser la structure émergente du sonnet – regardez par exemple dans la fenêtre ‘Sommaire’ à gauche :



Pour contrôler l’affichage d’un text balisé, il faut préciser comment nous désirons visualiser chaque balise. Ces précisions se font avec ce qu’on appelle une *feuille de style*.

- Sélectionnez `XML Document` et ensuite `Associer une feuille de style` à partir du menu `Document`.
- Naviguez jusqu’au fichier `duBellay.css` dans le dossier `Travaux` qui devrait se trouver sur votre Bureau, et associez-le (ce fichier est aussi disponible à l’URL sur le web ).
- Notez qu’une ligne contenant un ”processing instruction” est ajouté au début de votre document.
- Contrôlez l’effet sur l’affichage de votre document au mode `Auteur`.

On peut changer cet affichage. Ouvrez le fichier `duBellay.css` avec `Oxygen` et modifiez-le, si cela vous intéresse. Mais il y a des limites :

- CSS n’est pas un langage XML, et ne permet que de modifier l’affichage du contenu d’un document XML.
- On peut changer la taille de police, l’étendu des marges, les couleurs etc.
- Changez la couleur du titre de **red** en **purple**.

Des transformations beaucoup plus intéressantes sont possibles avec un autre langage de feuille de style : le XSLT. On verra cela prochainement...