



ir pam

Manuel d'encodage : Poésie latine biblique : manuscrits, inscriptions et commentaires

Manuel d'encodage : Poésie latine biblique : manuscrits, inscriptions et commentaires
© 2023 by Inès Burri and Bruno Bureau is licensed under [CC BY-NC-SA 4.0](https://creativecommons.org/licenses/by-nc-sa/4.0/)

Table des matières

1- Encoder les textes	3
1. Avant toute chose : qu'est-ce que le XML ?.....	3
a. Les balises	3
b. Les attributs	3
2. Préparer les textes des poètes.....	4
a. Présentation de l'encodage général	4
b. Écrire une référence	4
3. Préparer les textes des commentaires	5
a. Présentation de l'encodage général	5
b. Normaliser les références bibliques	5
c. Écrire un renvoi biblique avec citation	6
d. Écrire un renvoi biblique sans citation	7
e. Mentionner une erreur de référence	8
2- Transformer les textes pour les visualiser	8
1. Le dossier de travail	8
2. Appliquer le scénario qui permet de comparer les textes à partir de la Bible	8
3. Appliquer le scénario qui permet de comparer les poètes à la Bible à partir de chaque poète. ...	10
4. Appliquer le scénario pour visualiser les références et citations dans les commentaires	10

1- Encoder les textes

1. Avant toute chose : qu'est-ce que le XML ?

a. Les balises

L'ensemble du corpus est encodé en XML ou eXtensible Markup Language, qui est un langage informatique de balisage générique. Les contenus textuels sont compris dans des balises qui les structurent, les organisent et les caractérisent.

Nous avons ici l'exemple d'un texte « Francesco Stella ». Les balises `<persName>` l'encadrent et nous indiquent qu'il s'agit d'un nom de personne. Une **balise ouvrante** (composée de deux chevrons simple `<>`) indique le début de la mention et une **balise fermante** (avec pour le premier chevron l'ajout d'un slash `</>`) la fin.

```
<persName>Francesco Stella</persName>
```

Figure 1 Encodage d'un anthroponyme

Nous pouvons enrichir l'information en ajoutant une balise `<forename>` pour le prénom, et une seconde `<surname>` pour le nom. Vous remarquerez que ces autres balises sont comprises dans la première puisqu'elles dépendent d'elle. Une balise peut en contenir autant que possible tant que celles-ci ne se chevauchent pas entre elles.

```
<persName>  
  <forename>Francesco </forename>  
  <surname>Stella</surname>  
</persName>
```

Figure 2 Encodage enrichi d'un anthroponyme

Par exemple, l'encodage ci-dessous ne peut pas fonctionner car la balise `<surname>` s'ouvre dans la balise `<persName>` mais se ferme en dehors.

```
<persName>  
  <forename>Francesco </forename>  
  <surname>Stella</persName>  
</surname>
```

Figure 3 Exemple d'un "overlap"

Il existe également des balises, dites auto-fermantes (avec un chevron final fermant `</>`), qui ajoutent une information dans l'encodage en lui-même et qui ne possèdent pas de contenu textuel.

b. Les attributs

L'ajout d'un ou de plusieurs attributs à une balise viennent affiner le balisage. Voici, par exemple, l'encodage de cette courte biographie de M. Stella :

« Born in 1962, Francesco Stella is full professor of Medieval Latin Literature at the University of Siena. »

Notre texte est compris dans une balise `<p>` (pour paragraphe) et contient un attribut `@xml:lang` « en » qui précise qu'il est écrit en anglais. L'attribut `@type` « birth_date » de la balise `<date>` nous permet d'identifier le chiffre comme une **date de naissance**. Puis, on suit la même logique pour le statut professionnel (`<roleName>`) et le lieu de rattachement (`<orgName>`).

Par convention, un attribut, lorsqu'on en parle est précédé du signe @. Mais, lorsqu'on l'utilise dans l'encodage on ne met jamais l'@. Comme dans l'exemple suivant.

```

<p xml:lang="en">Born in
  <date type="birth_date">1962</date>,
  <persName>
    <forename>Francesco</forename>
    <surname>Stella</surname>
  </persName>
  is <roleName type="profession">full professor</roleName>
  of Medieval Latin Literature at the
  <orgName>University of
    <placeName>
      <settlement type="city">Siena</settlement>
    </placeName>
  </orgName>
.</p>

```

Figure 4 Encodage d'une biographie

2. Préparer les textes des poètes

a. Présentation de l'encodage général

Les livres des poèmes ont été balisés dans des `<div1>` auquel on a ajouté un `@type` « Liber », un `@n` (= number) et un `@xml:id` (= un identifiant unique). Une balise `<p>` (= paragraphes) regroupe des balises `<seg>` qui servent, ici, à encoder les vers d'une même phrase. Elles possèdent un `@n` et un `@xml:id`. On y trouve des balises autofermantes `<lb>` pour marquer les débuts de vers¹, qu'on complète avec un `@type` « versus » et un `@n`.

```

<div1 type="Liber" n="2" xml:id="Hild_2">
<p>
  <seg n="1" xml:id="Hild_2_1">
    <lb type="versus" n="1"/>Fert Amalechites quoniam Saul ipse peremit,
    <lb type="versus" n="2"/>Protinus in Siceleg David eum perimit.
  </seg>
  <seg n="2" xml:id="Hild_2_2">
    <lb type="versus" n="3"/>Ungitur a Iuda rex David versus in Ebron,
    <lb type="versus" n="4"/>Eius et imperio dux Ioab est equitum.
  </seg>
ETC...

```

Figure 5 Encodage général des poèmes

b. Écrire une référence

Ainsi, l'arborescence des textes étant déjà encodés, il reste à baliser les renvois aux passages bibliques cités. On part du principe qu'un segment renvoie au minimum à un passage biblique voire à plusieurs. Si un segment n'a aucun correspondant dans la Bible, le fait de ne pas réaliser l'encodage, décrit ci-dessous, suffit à faire comprendre qu'il s'agit d'un segment original sans source dans le texte biblique.

Les informations concernant les renvois doivent être balisées dans les `<seg>`. À l'intérieur de celles-ci et à leur commencement, on utilisera une balise `<listRef>`, qui contiendra une ou plusieurs balises `<ref/>` autofermantes. C'est ici qu'il faudra mentionner la/les référence(s) à l'aide d'un attribut `@cRef`.

¹ Nous avons choisi d'utiliser des balises autofermantes pour éviter les possibles *overlap* avec les `<seg>`.

Pour l'@cRef on veillera particulièrement à lui donner la valeur telle qu'exprimée par l'@xml:id du fichier correspondant. Dans le cas suivant, il s'agit de renvois à 1 Samuel (désigné par l'@xml:id S1) auquel nous ajoutons le numéro du chapitre et du verset, séparés par un *underscore*. Attention, l'utilisation d'une espace ou d'un tiret dans la valeur d'un attribut le rendrait inutilisable et bloquerait la transformation.

```
<seg n="1" xml:id="Hild_1_1">
  <listRef>
    <ref cRef="S1_1_1"/>
    <ref cRef="S1_1_2"/>
    <ref cRef="S1_1_3"/>
  </listRef>
  <lb type="versus" n="1"/>Helcana de Ramatha Phenennae sponsus et Annae
  <lb type="versus" n="2"/>Venerat in Silo sacra Deo facere.
</seg>

ETC...
```

Figure 6 Encodage des renvois bibliques dans les poèmes

3. Préparer les textes des commentaires

a. Présentation de l'encodage général

Les commentaires suivent le même principe d'arborescence à la différence qu'il a fallu ajouter une sous-partie pour les chapitres, encodés à l'aide de balises <div2> et accompagnées d'un @type « Caput » et d'un @n. Suivant le même principe que pour les poèmes, nous retrouvons des <p> pour les paragraphes, et des <seg> pour les phrases.

```
<div1 type="Proemium" n="1"> (...)
<div1 type="Liber" n="1"> (...)
<div1 type="Liber" n="2">
  <div2 type="Caput" n="1">
    <p>seg> Nunc autem parienti canticum dicere quid est, nisi conditorem,
    quem ineffabiliter diligit, veris praeconiis praedicare?</seg><seg> Dicit
    ergo: (I Reg. II, 1.) Exsultavit cor meum in Domino.</seg>
  </p>
```

Figure 7 Encodage général des commentaires

b. Normaliser les références bibliques

Toutes les références bibliques sont normalisées pour généraliser les renvois. Nous utiliserons celles qui sont fournis sur la page [Wikipédia : « Abréviations livres de la Bible »](#). Pour compléter la référence, on ajoutera un *underscore*, le numéro du chapitre, un autre *underscore* et le numéro du verset. Ex : Jb_1_1 → Livre 1 de job, verset 1

Pour les autres références, nous utiliserons celles du [Thesaurus Linguae Latinae](#).

Les références correspondant au *Liber Regum* sont :

- Pour le *Liber primus Samuelis* : S1
- Pour le *Liber secundus Samuelis* : S2
- Pour le *Liber tertius Regum* : R3
- Pour le *Liber quartus Regum* : R4

c. Écrire un renvoi biblique avec citation

Les balises `<cit>` servent à encoder les renvois bibliques à l'aide d'un `@source` contenant la valeur normalisée du renvoi. Une balise `<quote>` vient encadrer dans un dernier temps la citation. Il faut ensuite supprimer la `référence` dans le texte puisque l'information, désormais présente dans l'encodage, est devenue inutile. Même chose pour les guillemets qui sont supplantés par la balise `<quote>`. Attention à bien supprimer également les `espaces` avant ou après les parenthèses.

```
<cit source="S1_1_2">
<quote>(Vers. 2.) Qui habebat duas uxores, nomen uni
Anna, nomen secundae Phenenna: fueruntque
Phenennae filii, Annae vero non erant liberi.</quote>
</cit>
```

Figure 8 Encodage d'un renvoi biblique, avec une citation de la Bible de Samuel

Que faire si le renvoi contient plusieurs citations d'un même livre ? Dans ce cas de figure, toutes les indications communes seront comprises dans l'`@source` de la balise `<cit>` et les différences seront renseignées dans de nouvelles balises. Ainsi, au commencement du `<quote>`, il faudra créer une balise `<listRef>` qui contiendra autant de balises `<ref>` que de renvois, et ces dernières posséderont un `@cRef` dont la valeur correspondra au renvoi normalisé de la référence. De nouveau, il faudra penser à supprimer les `références` et les `espaces` inutiles dans le texte.

```
<cit source="Jn_10">
  <quote>
    <listRef>
      <ref cRef="Jn_10_17"/>
      <ref cRef="Jn_10_18"/>
    </listRef>
    <seg>Propterea me diligit Pater, quia ego pono animam, et
iterum sumam eam: nemo tollit eam a me, sed ego pono eam
a meipso.</seg>
    <seg>Potestatem habeo ponendi eam, et iterum sumendi
eam (Jn. X, 17, 18)</seg>
  </quote>
</cit>
```

Figure 9 Encodage d'un renvoi des références multiples d'un même auteur

Si le renvoi contient une citation qui se réfère à différents livres de la Bible, on choisira la première référence citée pour l'`@source` et le reste pour les `@cRef`. Ne pas oublier de supprimer les `références` et les `espaces` inutiles dans le texte.

```
<cit source="Mt_3_17">
  <quote>
    <listRef>
      <ref cRef="Mt_17_5"/>
      <ref cRef="1P_2_17"/>
    </listRef>
    Hic est filius meus dilectus, in quo mihi bene complacui. (Mt. III, 17; XVII, 5; 1P. II, 17)
  </quote>
</cit>
```

Figure 10 Encodage d'un renvoi à une citation commune à plusieurs bibles

Lorsqu'une citation est coupée par un verbe de déclaration, on encodera le début de la citation avec un `<cit>` sans `@source` avec un `<quote>`, et la deuxième partie avec un `<cit>` `@source` suivi d'un `<quote>` (voir l'exemple ci-dessous).

```
<cit><quote> Praecipe</quote></cit>, ait Dominus,  
<cit source="Ex_27"><quote> filiis Israel ut afferant tibi oleum de arboribus olivarum  
purissimum, piloque contusum, ut ardeat lucerna semper in tabernaculo Domini, et  
collocabunt eam Aaron et filii ejus, ut usque mane luceat coram Domino </quote></cit>
```

Figure 11 Encodage d'une citation en deux parties

d. Écrire un renvoi biblique sans citation

Les renvois ne contiennent pas toujours de citation.

Dans le cas où il n'y a qu'une seule référence, nous utiliserons une balise `<cit>` avec un `@type` « `reference` », puis une balise `<ref>` autofermante et un `@cRef` contenant la référence normalisée. Pour les mêmes raisons que précédemment, il faudra supprimer la `référence` dans le texte.

```
<cit type="reference">  
<ref cRef="Ac_4_4"/>(Ac. IV, 4)  
</cit>
```

Figure 12 Encodage d'un renvoi biblique sans citation

S'il y en a plusieurs, nous utiliserons une balise `<cit>` avec un `@type` « `reference` », puis une balise `<ref>`, puis une balise `<listRef>` enfermant des balises `<ref>` autofermantes avec, pour chacune, un `@cRef` contenant une des références normalisées. De la même manière, il faudra supprimer la `référence` dans le texte.

D'autre part, et comme précédemment, si les renvois se réfèrent à un même auteur et ou même livre, il faudra ajouter un `@cRef` à la première balise `<ref>` contenant les informations similaires. Attention, si les renvois sont différents, nous laisseront la première `<ref>` vide (voir figure 12).

```
<cit type="reference">  
  <ref cRef="HIER">  
    <listRef>  
      <ref cRef="HIER_adv_Pelag_1_8"/>  
      <ref cRef="HIER_adv_Iovin_1_10"/>  
    </listRef>  
  </ref>(AdVers. Pelag., lib. I, c. 8 et contra Jovin. lib. I, c. 10)  
</cit>
```

Figure 14 Encodage de renvois bibliques multiples mais communes sans citation

```
<cit type="reference">  
  <ref>  
    <listRef>  
      <ref cRef="Mt_17_3"/>  
      <ref cRef="Lc_9_30"/>  
    </listRef>  
  </ref>  
</cit>
```

Figure 13 Encodage de renvois bibliques différents sans citation

e. Mentionner une erreur de référence

Il est possible de rencontrer des erreurs de référence dans le texte. Il faudra donc les corriger et conserver l'erreur constatée. Pour cela nous utiliserons une balise `<choice>` qui renseignera l'erreur dans l'`@source` d'une balise `<sic>`, et une balise `<corr>` contenant la citation encodée comme précédemment, c'est-à-dire avec un `<cit>` `@source` et d'un `<quote>`. La balise `<sic>` devra se positionner après la balise `</corr>`.

```
<choice>
  <corr>
    <cit source="Es_5"><quote>ne pluant super eam imbrem.</quote></cit>
  </corr>
  <sic source="Es_6"/>
</choice>
```

Figure 15 Encodage d'une citation qui possède une erreur de référence dans le texte

2- Transformer les textes pour les visualiser

Une fois les XMLs dûment enrichis, il faut désormais exporter toutes ces données pour les rendre intelligibles. C'est le rôle des fichiers de transformations XSL(s) qui auront pour fonction de créer de nouveaux fichiers à partir des XML.



Figure 16 Schéma de transformation d'XMLs vers HTML

Cette opération passe par un scénario de transformation où l'on indique quels seront le(s) XML(s) à transformer, par quel XSL, et à quel endroit seront enregistrés les documents. Il est primordial de respecter scrupuleusement chacune de ces étapes pour obtenir une visualisation des données transformées. Pour nous, il s'agira de fichiers HTMLs car notre visualisation est un site web.

1. Le dossier de travail

L'ensemble de nos données sont enregistrées dans un dossier « PoesieLatineBiblique »². Vous y trouverez :

- un dossier « XML » contenant les fichiers XML
- un dossier « XSLT » contenant les fichiers XSL
- un dossier « XHTML » où sont enregistrés les fichiers HTML qui contiennent toutes les informations textuelles du site
- un dossier « ZIP » contenant les fichiers XML compressés pour permettre leur téléchargement sur les pages de lecture des textes.
- un dossier « Manuels » contenant les fichiers PDF des manuels de présentation et d'encodage du site.

Les deux autres dossiers, « CSS » et « images », servent à la mise en page du site.

2. Appliquer le scénario qui permet de comparer les textes à partir de la Bible

La construction des fichiers de visualisation pour la comparaison des versets des textes de la Bible aux poètes et aux commentaires, se fabrique à partir du fichier *Master.xml* et du fichier *Extraction_Bible.xsl*.

Sur *Oxygen*, cliquez sur le bouton qui sert à paramétrer un scénario de transformation.

² Ces données seront disponibles au téléchargement dans un entrepôt Nakala à la fin du financement ANR-DFG PoBLAM en mars 2022.

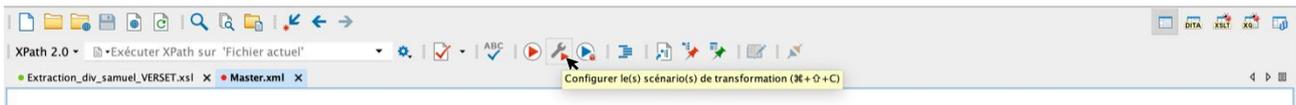


Figure 17 Icône de configuration d'un scénario de transformation

Un onglet s'ouvre. Cliquez sur « Nouveau », et choisissez « XML transformation with XSLT » (voir la figure : étape n°1). Dans cette nouvelle fenêtre, il faut y renseigner le fichier XML et le fichier XSL, en cliquant sur les dossiers à gauche de la barre correspondante et en retrouvant le chemin qui mènent jusqu'aux fichiers. (voir la figure : étape n°2). Dans cette même fenêtre sur l'onglet « Sortie », choisissez le chemin de sortie : c'est-à-dire le dossier XHTML du projet et validez. Attention, vous ne devez pas changer le nom du document et laisser « Untitled ». Si le logiciel vous dit que le fichier existe déjà et qu'il risque d'être écrasé cliquez sur « Remplacer ». De retour sur la fenêtre, cochez la case « XHTML », et enfin cliquez sur « OK ». (voir la figure : étape n°3).

Vérifiez que votre scénario est bien coché et cliquez sur « Appliquez le/les scénario(s) associé(s) » en bas à droite. Vous pouvez désormais actualiser le site pour voir vos modifications.

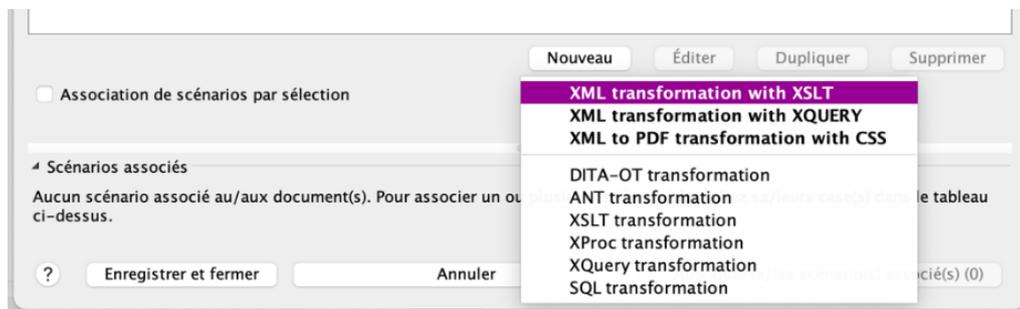


Figure 18 Configuration d'un nouveau scénario de transformation étape n°1

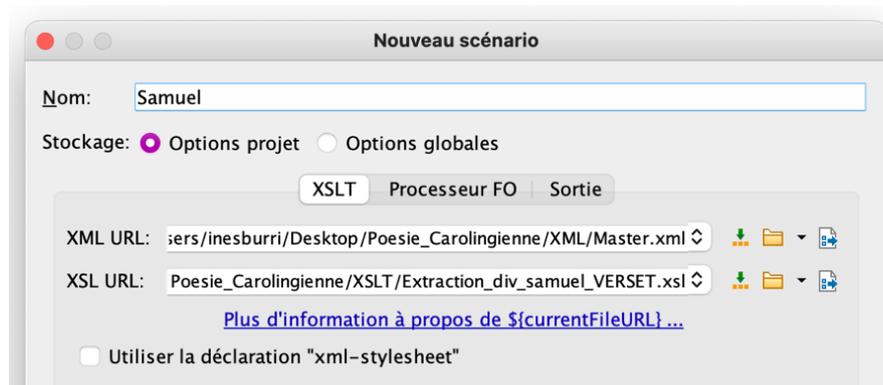


Figure 19 Configuration d'un nouveau scénario de transformation étape n°2

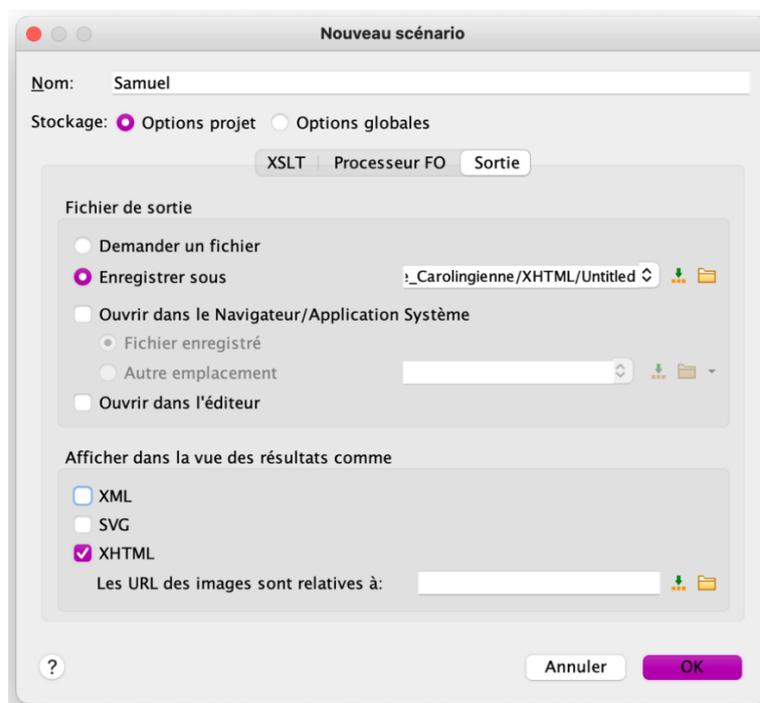


Figure 20 Configuration d'un nouveau scénario de transformation étape n°3

3. Appliquer le scénario qui permet de comparer les poètes à la Bible à partir de chaque poète.

Pour les comparaisons des poètes à la Bible, il s'agira des mêmes étapes et un chemin de sortie rigoureusement identique. La seule différence réside dans le choix du fichier XSL : il faudra choisir le fichier *Extraction_seg_poetes.xml* et l'appliquer au fichier *Master.xml*.

4. Appliquer le scénario pour visualiser les références et citations dans les commentaires

Pour visualiser les textes des commentaires avec les liens et citations, nous aurons de nouveau les mêmes étapes et le même chemin de sortie. La seule différence concerne le choix de l'XSL : *ext_tei_lectseule_commentaires.xml* (que vous trouverez dans le dossier « lectures seules »).

5. Le contenu des fichiers XSLT

Les feuilles de transformation XSLT sont entièrement commentées pour expliquer étape par étape comment les utiliser, et comprennent des titres pour structurer l'ensemble. Par commodité, ils commencent par un commentaire en majuscules décrivant leur objectif et indiquant le fichier XML sur lequel l'appliquer.