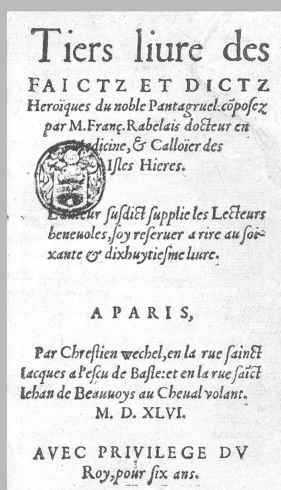


Manuel d'encodage XML-TEI

Renaissance et temps modernes
Imprimés - manuscrits



Recommandations pour une application TEI
Projet BVH
Bibliothèques Virtuelles Humanistes :
<http://www.bvh.univ-tours.fr>



Nicole Dufournaud

Avec la collaboration de Marie-Luce Demonet, Toshinori Uetani et Thierry Vincent

Avec les suggestions de Lou Burnard et Jean-Daniel Fekete

Révision: Vanessa Le Rolle, Laetitia Bontemps, Sandrine Breuil, Rémi Jimenes, Alice Nué

Version 2 (1er juillet 2009)



Table des matières

Introduction.....	4
I Notions.....	6
1. Principes de balisage.....	9
1) Les éléments TEI.....	10
2) Les attributs.....	11
2. Structure du texte TEI : quelques notions.....	11
1) L'en-tête TEI : <teiHeader>.....	12
2) Le texte de l'œuvre.....	14
II Balisages.....	16
1. Balisages formels et sémantiques.....	16
1) Balisage formel.....	16
(1) Le balisage structurel.....	16
(2) La mise en forme typographique.....	19
(3) Les interventions éditoriales.....	28
2) Balisage sémantique.....	35
(1) Les citations et références bibliographiques.....	35
(2) Les langues.....	37
(3) L'onomastique : noms de personnes, dénomination sociale et toponymes.....	38
(4) La datation calendaire : <date>.....	40
2. Images typographiques.....	40
III En-tête TEI <teiHeader>.....	43
1. Description bibliographique du fichier : <fileDesc> file description.....	43
2. Description de l'encodage : <encodingDesc> encoding description.....	47
3. Description non bibliographique : <profileDesc> text-profile description.....	48
4. Journal de modifications : <revisionDesc> revision description.....	49
IV Texte de l'œuvre.....	50
1. Structure complexe : les Euvres de Louise Labé.....	50
2. Page de titre et pièces liminaires.....	52
Remarque importante.....	55
Annexe 1 : Liste des livres encodés (juin 2009).....	56
Annexe 2 : Liste des langues et dialectes utilisés.....	58
Annexe 3 : liste des éléments et des attributs.....	59
Index des illustrations.....	66

Introduction

TEI (Text Encoding Initiative) est un projet universitaire pluridisciplinaire visant à uniformiser autant que possible le codage de documents en vue de leur échange et de leur publication en ligne ou hors ligne. Il s'agit d'un format de codage de documents, dit «structuré» : il a besoin d'un langage, XML, pour aider à la saisie d'un texte en lui donnant une structure compatible à la fois avec les exigences des différentes communautés qui s'en servent (bibliothécaires, historiens, philosophes, spécialistes de littérature, linguistes, amateurs) et avec les possibilités des outils de consultation. Les recommandations du consortium TEI (<http://www.tei-c.org/>) dans sa version P5 (décembre 2007 puis novembre 2008) ont été suivies, en s'inspirant du manuel TEI des *Cahiers Gutenberg* de 1996 pour élaborer le présent manuel¹.

Cet ouvrage n'a pas pour objectif de remplacer la lecture attentive des directives de TEI (<http://www.tei-c.org/P5/Guidelines/>), mais d'en décrire une application et de guider son utilisation pour les textes allant du début de l'imprimerie à la période de normalisation typographique au XIX^e siècle. Il devrait être complété par une section sur la plate-forme de travail élaborée à l'INRIA² : un tel outil permet une utilisation plus simple du document en cours d'encodage grâce à son environnement convivial. Il autorise une navigation entre les fichiers XML et l'affichage HTML, et permet la visualisation d'index et de tables des matières générés automatiquement. Il facilite aussi le travail en équipe.

Nous nous appuyons sur des exemples pratiques. Trois ouvrages du XVI^e siècle français nous ont servi à illustrer concrètement les principaux usages d'encodage de TEI :

1. Louise Labé, *Euvres*, Lyon, Jean de Tournes, 1555 (d'après un exemplaire original disponible sur le site « Gallica » de la Bibliothèque de France) ;
2. [Elie VINET ?], *Discours non plus mélancoliques que divers*, Poitiers, Enguilbert de Marnef, 1556, d'après un exemplaire reproduit sur Gallica, et un autre, daté de 1557, de la Médiathèque François-Mitterrand de Poitiers en ligne sur BVH ;
3. François Rabelais, *Le Tiers Livre*, Paris, Chrétien Wechel, 1546 (BVH), et Paris, Michel Fezandat, 1552³.

D'autres œuvres encodées ultérieurement complètent cette liste⁴.

Leur structure complexe et les problèmes divers rencontrés au cours de leur encodage ont permis de vérifier l'adéquation des éléments existant dans TEI —voire d'en proposer d'autres lorsque cela s'est révélé nécessaire—, et d'élaborer un protocole qui guide l'application à d'autres textes publiés pendant trois siècles. Conçu en premier lieu à l'usage des BVH (« Bibliothèques Virtuelles Humanistes ») du Centre d'Études Supérieures de la Renaissance (CESR) pour des textes français de la

1 *Cahiers Gutenberg*, « Numéro spécial : TEI Text Encoding Initiative », dir. Jacques ANDRE, Rennes, n° 24, juin 1996.

2 INRIA : Institut National de Recherche en Informatique et en Automatique, équipe-projet AVIZ, INRIA Saclay – Île-de-France, dirigée par Jean-Daniel Fekete, que nous remercions pour sa précieuse contribution.

3 Nous avons travaillé en parallèle sur deux exemplaires de cette édition : celui de la Bibliothèque nationale de France (Rés. Y2 2162) et la reproduction en ligne de l'exemplaire de la collection Gordon (Université de Virginie, R 258) ; les variantes constatées (et encodées) nous ont permis de découvrir que ce dernier représentait un état du texte inconnu jusqu'à présent.

4 Voir en annexe I la liste des ouvrages encodés.

Renaissance, ce manuel pourra servir de base à d'autres projets portant sur d'autres documents manuscrits ou imprimés.

Encodage et transcription

Si les exemples donnés sont français, la TEI « Renaissance et temps modernes » ne réalise pas un encodage spécifique des textes écrits en cette langue, car TEI offre la possibilité d'usages multilingues de la description structurelle. L'universalité des éléments permet la compatibilité des analyses, et les sources numérisées dans le cadre des BVH relèvent de tous les domaines encyclopédiques, dans plusieurs langues anciennes et modernes.

Cette application montre comment on peut adapter les éléments à la description des textes anciens, en respectant leur présentation originale : pagination, lignage, disposition des titres, insertion des éléments illustrés, caractéristiques typographiques. Elle se fonde sur une transcription « patrimoniale » du texte, soucieuse de donner le plus grand nombre possible d'informations tout en autorisant les traitements automatiques pour les recherches sur la langue, les données historiques, les différentes versions du document et les variantes dont le niveau de précision peut être adapté aux recherches désirées.

Cette représentation informatique du document est le plus petit dénominateur commun aux exigences des différentes communautés des sciences humaines, en leur laissant le soin d'y ajouter des annotations relevant de leurs spécialités. Les recommandations de ce manuel sont autant que possible d'ordre générique, autorisant des options qui permettent de les adapter à d'autres bases textuelles.

Par rapport aux usages habituels de TEI, les difficultés rencontrées lors de son application aux textes de cette période viennent toutes du désir de fidélité à l'exemplaire transcrit. Nous avons tenté de rendre compatibles l'exigence de restitution et la possibilité d'effectuer des requêtes impliquant des comptages, et donc une certaine régularisation, notamment des graphies multiples. L'un des principaux intérêts du langage XML est de permettre de résoudre par l'encodage la très grande variation typographique et textuelle des documents : cette souplesse ne doit pas cependant être un frein à l'acquisition et à l'encodage rapides des corpus textuels. Nous avons utilisé cette procédure en même temps que nous développons, avec des laboratoires d'informatique, des routines d'automatisation fondées sur l'analyse de structure, ce qui constitue l'un des axes de recherche interdisciplinaire du projet « BVH »⁵.

Au fur et à mesure que l'équipe a progressé dans la mise au point de cette application, il est apparu qu'une telle structure pouvait aussi servir à modéliser les ouvrages qui ne seraient publiés qu'en mode image, mais dont les divisions seraient naturellement de même nature : titres, chapitres, paragraphes, éléments illustrés, etc., de façon à préparer et à faciliter l'alignement des deux modes de présentation du texte, un idéal que les usagers des bibliothèques numériques attendent avec impatience et auquel le consortium TEI travaille.

Les œuvres encodées peuvent être des manuscrits ou des imprimés, mais les sources textuelles qui ont servi à l'élaboration du manuel d'encodage sont des imprimés. La nature même du texte influe sur la structure du document : un poème ne sera pas encodé comme un roman car sa structure n'est pas la même (strophes pour le premier et paragraphes dans des chapitres pour le second). Moyennant quelques modifications, ce manuel pourra aussi s'appliquer à l'encodage des manuscrits⁶.

5 Le Laboratoire LI-RFAI (dir. Jean-Yves Ramel) et le consortium Navidomass (projet ANR 2007-2009).

I Notions

Commençons par deux pages du *Tiers Livre* de Rabelais (1546) :

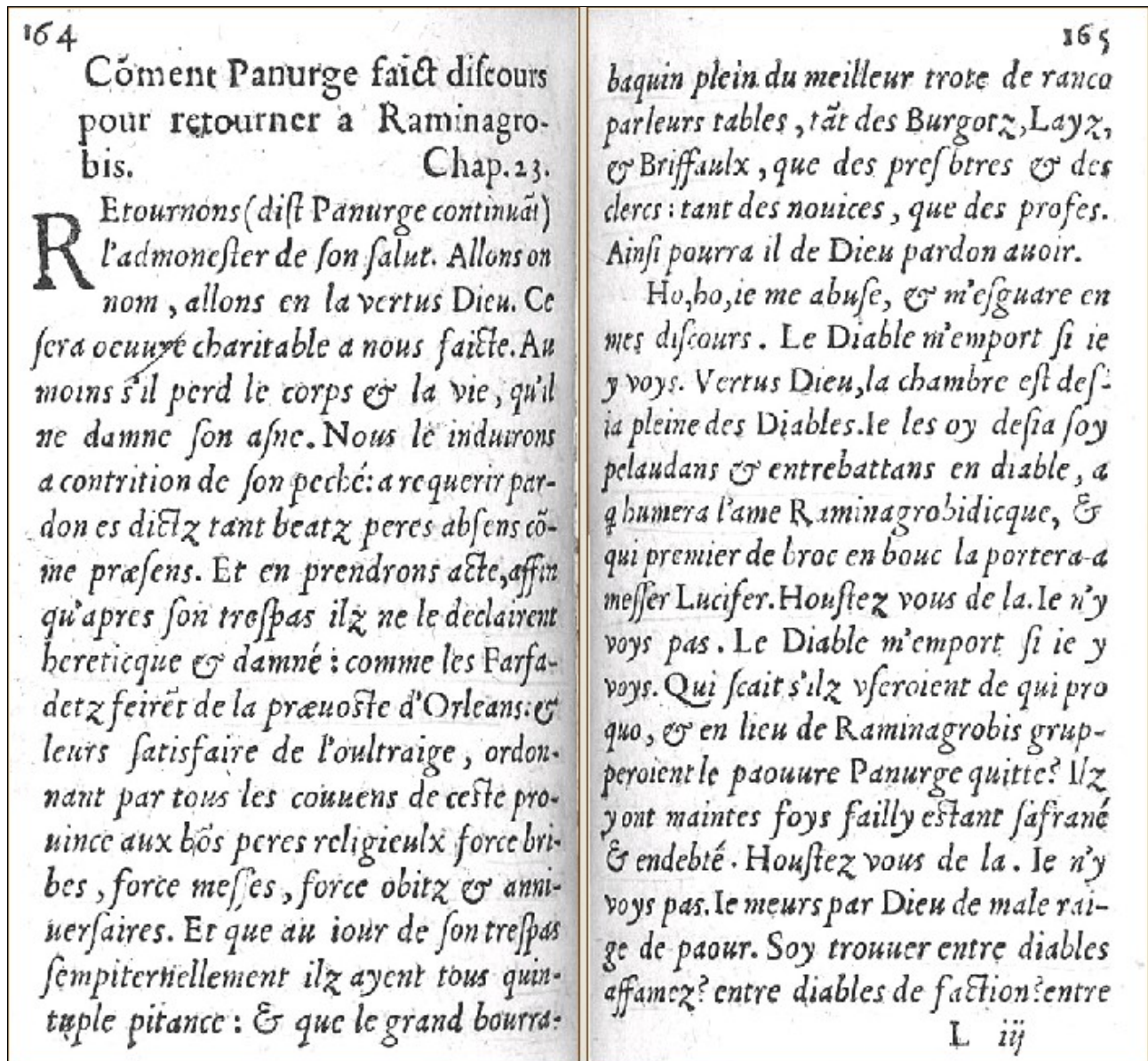


Illustration 1: Chapitre 23 du Tiers Livre de Rabelais, édition 1546

Une transcription effectuée avec un traitement de texte donne le résultat suivant :

6 Voir le protocole TEI et plus particulièrement les deux chapitres « Manuscripts Description » et « Representation of Primary Sources » <http://www.tei-c.org/P5/Guidelines/>

164

Comment Panurge faict discours
pour retourner a Raminagro-
bis. Chap. 23.

REtournons (*dist Panurge continuant*)
l'admonester de son salut. Allons on
nom, allons en la vertus Dieu. Ce
sera oeuvre charitable a nous faicte. Au
moins s'il perd le corps & la vie, qu'il
ne damne son asne. Nous le induirons
a contrition de son peché: a requérir par-
dou es dictz tant beatz peres absens com-
me praesens. Et en prendrons acte, affin
qu'apres son trespas ilz ne le declairent
hereticque & damné: comme les Farfa-
detz feirent de la praevoste d'Orleans: &
leurs satisfaire de l'oultraige, ordon-
nant par tous les couvens de ceste pro-
vince aux bons peres religieux force bri-
bes, force messes, force obitz & anni-
versaires. Et que au jour de son trespas
sempiternellement ilz ayent tous quin-
tuple pitance: & que le grand bourra-

165

baquin plein du meilleur trote de ranco
par leurs tables, tant des Burgotz, Layz,
& Briffaulx, que des presbtres & des
clercs: tant des novices, que des profes.
Ainsi pourra il de Dieu pardon avoir.

Ho, ho, je me abuse, & m'esguare en
mes discours. Le Diable m'emporte si je
y voys. Vertus Dieu, la chambre est des-
ja pleine des Diables. Je les oy desja soy
pelaudans & entrebattans en diable, a
qui humera l'ame Raminagrobicque, &
qui premier de broc en bouc la portera a
messer Lucifer. Houstez vous de la. Je n'y
voys pas. Le Diable m'emport si je y
voys. Qui scait s'ilz useroient de qui pro
quo, & en lieu de Raminagrobis grup-
peroit le paouvre Panurge quitte? Ilz
y ont maintes foyz failly estant safrané
& endebté. Houstez vous de la. Je n'y
voys pas. Je meurs par Dieu de male rai-
ge de paour. Soy trouver entre diables
affamez? entre diables de faction? entre

L iij

Illustration 2: Transcription du chapitre 23 du Tiers Livre

Les limites de cette transcription établie avec un traitement de texte apparaissent ainsi :

- les numéros de page, la signature (L iij) et le titre du chapitre ne sont pas délimités et font partie du texte ;
- les paragraphes se différencient grâce à des fins de ligne forcées et un retrait d'alinéa ;
- les choix éditoriaux ne sont pas explicites : les abréviations sont développées ; les césures de mots en fin de ligne sont respectées ;
- les noms de personnes et de lieux ne sont pas repérables.

Nous présentons le même exemple sous une troisième forme conforme aux recommandations de XML/TEI qui rend explicite les règles de codage :

- les numéros de page et la signature font l'objet d'une délimitation spécifique et le saut de page est distinct ;
- le texte est explicitement structuré : titre du chapitre et paragraphes ;

- les noms de personnes et de lieux sont marqués ;
- les rendus typographiques sont indiqués ;
- il a été ajouté un « sic » pour marquer le mot « asne » là où on attendait « ame ».

Enfin, les abréviations ne sont pas encodées, ni les césures de fin de ligne ; le caractère & est remplacé par &7.

```
<pb/>
<fw place="top-left" type="pageNum">164</fw>
<div type="chapter" n="23">
  <head>
    <lb/>Comment <name>Panurge</name> faict discours
    <lb/>pour retourner a <name>Raminagro-
    <lb/>bis</name>. <hi rend="right">Chap.23.</hi>
  </head>
  <p>
    <lb/>RE<hi rend="it">tournons (dist <name>Panurge</name> continuant)
    <lb/>l'admonester de son salut. Allons on
    <lb/>nom, allons en la vertus Dieu. Ce
    <lb/>sera oeuvre charitable a nous faicte. Au
    <lb/>moins s'il perd le corps &amp; la vie, qu'il
    <lb/>ne damne son <sic>asne</sic>. Nous le induirons
    <lb/>a contrition de son peché: a requerir par-
    <lb/>don es dictz tant beatz peres absens com-
    <lb/>me praesens. Et en prendrons acte, affin
    <lb/>qu'apres son trespas ilz ne le declairent
    <lb/>hereticque &amp; damné comme les Farfa-
    <lb/>detz feirent de la praeuostie d'<placeName>Orleans</placeName>:
    &amp;
    <lb/>leurs satisfaire de l'oultraige, ordon-
    <lb/>nant par tous les couuens de ceste pro-
    <lb/>vince aux bons peres religieulx force bri-
    <lb/>bes, force messes, force obitz &amp; anni-
    <lb/>versaires. Et que au jour de son trespas
    <lb/>sempiternellement ilz ayent tous quin-
    <lb/>tuple pitance: &amp; que le grand bourra-
    <pb/>
    <fw place="top-right" type="pageNum">165</fw>
    <lb/>baquin plein du meilleur trote de ranco
    <lb/>par leurs tables, tant des Burgotz, Layz,
    <lb/>&amp; Briffaulx, que des presbtres &amp; des
    <lb/>clercs: tant des novices, que des profes.
    <lb/>Ainsi pourra il de Dieu pardon avoir.</hi>
  </p>
  <p>
    <lb/><hi rend="it">Ho, ho, je me abuse, &amp; m'esquare en
    <lb/>mes discours. Le Diable emport si je
    <lb/>y voys. Vertus Dieu, la chambre est des-
    <lb/>ja pleine des Diables. Je les oy desja soy
    <lb/>pelaudans &amp; entrebattans en diable, a
    <lb/>qui humera l'ame Raminagrobidicque, &amp;
    <lb/>qui premier de broc en bouc la portera a
    <lb/>messer <name>Lucifer</name>. Houstez vous de la. Je n'y
    <lb/>voys pas. Le Diable m'emport si je y
    <lb/>voys. Qui scait s'ilz useroient de qui pro
```

7 Voir « Remarque importante » en fin de manuel.


```

<lb/>quo, &amp; en lieu de <name>Raminagrobis</name> grup-
<lb/>peroient le paouvre <name>Panurge</name> quitte? Ilz
<lb/>y ont maintes foyz failly estant safrané
<lb/>&amp; endebt . Houstez vous de la. Je n'y
<lb/>voys pas. Je meurs par Dieu de male rai-
<lb/>ge de paour. Soy trouver entre diables
<lb/>affamez? entre diables de faction? entre
<fw place="bot-center" type="sig">L iij</fw>

```

Le texte a une structure physique, logique et s mantique. L'encodage montre les divisions de type chapitre, compos es de paragraphes dans lesquels se juxtaposent des lignes qu'il est possible de num roter.

1. Principes de balisage

« Les  l ments, qui sont les structures de base des documents XML, sont d limit s par des balises de d but et des balises de fin qui portent le contenu » est une d finition parmi beaucoup d'autres⁸.

TEI est une norme de codage de textes qui repose sur XML. Des marques sont utilis es pour identifier le contenu du texte : elles sont appel es « balises » (en anglais *tags*). Dans XML, le marquage est constitu  des chevrons « < » et « > », par exemple :

```
<title>Discours non plus m lancoliques...
```

Dans cet exemple, la balise <...> entoure le mot `title` mis entre les chevrons.

En g n ral, la s quence textuelle est encadr e par une balise ouvrante et une balise fermante :

```
<p>Le Tiers Livre de Fran ois Rabelais.</p>
```

Le `p` est marqu  deux fois : le premier est dans la balise ouvrante <...>, soit la balise de d but ; le second est dans la balise fermante </...>, soit la balise de fin qui est compos e du chevron fermant mais  galement de la barre « / » qui ici signifie la fin du marquage des  l ments du texte.

Nous pouvons reprendre notre premier exemple, celui du titre :

```
<title>Discours non plus m lancoliques ...</title>
```

Le mot `title` est mis une seconde fois   la fin du titre.

Il y a deux r gles essentielles dans XML : la syntaxe des balises telle que nous venons de la d crire et la r gle d'imbrication stricte des  l ments. Ainsi, on peut imbriquer un  l ment dans un autre :

```
<p>Le Tiers Livre de <name>Fran ois Rabelais</name>.</p>
```

Il est interdit de faire chevaucher les  l ments comme le montre l'exemple suivant :

```
<p>Le Tiers Livre de <name>Fran ois Rabelais.</p></name>
```

⁸ Simon St. Laurent et Michael Fitzgerald, *XML Pr cis et concis*, traduction de Philippe Ensarguet et Jean-No l Gadreau, Paris, Ed. O'Reilly, 2005.

Ces règles simples sont fondamentales : si elles ne sont pas suivies scrupuleusement, le document XML produit est invalide.

TEI repose sur XML pour sa syntaxe. Tout document TEI doit être conforme aux règles édictées précédemment, avec une liste d'éléments valides et des règles d'emboîtements autorisées.

1) Les éléments TEI

On appelle « élément » la partie du texte qui se trouve dans les balises après le chevron « < » de la balise ouvrante, et répétée dans la balise fermante. Dans l'exemple précédent, l'élément est indiqué par la lettre *p* qui signifie *paragraph*⁹. Dorénavant, nous dirons « élément » pour signifier l'élément de la balise ; par exemple, nous écrirons l'élément *paragraph* ainsi : <p>.

Deux cas sont possibles :

1. les éléments ont un contenu, constitué d'une partie d'un texte ; il se place entre deux balises, la première ouvrante <...> et la seconde fermante </...> caractérisée par / qui se positionne **à gauche** de la balise. Par exemple, le titre du livre de Rabelais est encodé ainsi :

```
<title>Le Tiers Livre</title>
```

Le texte encodé est « Le Tiers Livre » et l'élément TEI est « *title* ».

2. les éléments n'ont pas de contenu : les balises se raccourcissent alors en une seule grâce au caractère / qui se positionne **à droite** de la balise ; par exemple, le saut de page ou *page break* :

```
<pb/>
```

Les éléments peuvent eux-mêmes être définis et classés dans deux catégories¹⁰ :

1. Dans la première catégorie, la structure du document est définie : il est possible de dresser l'organisation logique des œuvres comme la page de titre, les pièces liminaires, etc. La transcription physique du document signale les principaux traits éditoriaux ainsi que les informations supplémentaires comme les déchirures, les corrections, les ajouts, les suppressions, les abréviations. Dans le cas d'une source imprimée, les éléments éditoriaux diffèrent et concernent la typographie : titre courant, pied de page, signatures. Nous appelons cette catégorie : balisage formel et typographique.
2. La seconde est composée d'éléments analytiques succincts comme l'indication des noms propres, des toponymes, des dates ou de tout élément d'étude pour lequel il est possible de dresser des typologies. Nous appelons cette catégorie : balisage sémantique.

⁹ Tous les termes anglais seront en italique à l'exception des exemples.

¹⁰ Selon Jacques André, « le principe de base est "à chacun son métier..." : aux auteurs de travailler sur le "contenu" de leurs œuvres et aux typographes ou maquettistes de régler les problèmes de forme... » Voir *Cahiers Gutenberg*, article cité, p. 16.

Il existe une hiérarchie dans les éléments. Une certaine catégorie d'éléments est dite imbriquée dans une autre catégorie d'éléments. Par exemple, l'adresse est décomposée selon le modèle actuel, soit « rue » et « ville » ; l'élément `<address>` comprend toujours l'élément `<addrLine>`. Il n'est pas utilisé seul :

```
<address>
  <addrLine>au mont S. Hilaire, </addrLine>
  <addrLine>a l'hostel d'Albret.</addrLine>
</address>
```

Dans certaines configurations, certains éléments indépendants peuvent être imbriqués dans d'autres : par exemple, l'élément `<date>` s'utilise seul ou parfois imbriqué dans d'autres éléments comme dans celui qui est dédié à la bibliographie `<bibl>`.

Souvent, les éléments ne suffisent pas. Pour les caractériser, on a besoin d'y associer un paramètre spécifique : ce sont les « attributs ».

2) Les attributs

On appelle « attribut » un paramètre associé à un élément. Il peut être simple comme `n` pour *number* ou *name*, ou plus élaboré comme `resp="#ND"` pour pointer vers un autre endroit du fichier : le caractère `#` informe que la responsable de l'encodage Nicole Dufournaud a pour identifiant `ND` et se trouve référencée dans le fichier¹¹.

Un attribut est lui-même associé à une valeur. Par exemple, une date sera encodée ainsi :

```
<date when="1515-04-28">28 avril 1515</date>
```

Dorénavant, nous écrirons les attributs sous la forme suivante :

```
attribut="valeur"
```

ou encore précédé de l'arobase :

```
@attribut
```

Il existe des attributs globaux – c'est-à-dire utilisés par tous les éléments – mais également des attributs spécifiques à quelques éléments. Par exemple, l'attribut `@n` pour *number* est un attribut global alors que `@when` est un attribut réservé à l'élément `<date>`.

2. Structure du texte TEI : quelques notions

L'encodage global d'un document standard se présente ainsi :

¹¹ Les identifiants peuvent être notés dans le header.

```

<tei>
  <teiHeader>
    ...
  </teiHeader>
  <text>
    <front>...</front>
    <body>
      <div>
        ...
      </div>
    </body>
    <back>...</back>
  </text>
</tei>

```

Dans la suite de ce document, nous présentons et expliquons les principes élémentaires d'encodage pour baliser la structure d'un texte.

La structure de base d'un texte TEI repose sur un document¹². Dans le projet BVH, le document est en général un livre imprimé, mais ce peut être également un manuscrit. La structure d'un texte TEI décrit le contenu d'un document manuscrit ou imprimé qui comprend une partie textuelle prépondérante.

Comme la structure d'un document touche à l'aspect physique – il est composé de feuillets constitués de pages – on dit que la structure est physique. Une description logique simple du document peut être la suivante : la page de titre, les pages liminaires, le corps du texte et les annexes. La page de titre – s'il y en a une – est composée d'un gros titre ou de plusieurs titres, d'un auteur, etc. La page d'un imprimé peut être décrite ainsi : un numéro de page, un titre courant, un pied de page, etc.

À la structure physique du document se combine une structure logique : l'organisation du contenu et son découpage en parties, chapitres et sections. Si la structure physique s'adresse à l'œil, on entend par structure logique l'organisation du document en entités comme titre, chapitre, section, paragraphe, note, citation, tableau, liste, etc.

La notion de structure est fondamentale pour éditer, échanger et analyser des textes grâce aux recommandations TEI. S'y ajoute une autre notion : celle de l'en-tête spécifique à TEI que l'on nomme *header*. Tout texte de format TEI est composé d'un en-tête TEI placé avant le corps du texte lui-même.

1) L'en-tête TEI : <teiHeader>

L'en-tête contient des informations analogues à celles que l'on trouve sur la page de titre d'un texte imprimé complété par l'aspect numérique de la source : nous parlons de page de titre électronique. Tous les éléments de catalogage d'un ouvrage et de la transcription sont précisés ici.

¹² Sur le sujet, nous conseillons les travaux de Jacques André qui nous ont fortement inspirés, particulièrement « Numérisation et codage de livres anciens », *Documents numériques*, spécial « Numérisation et patrimoine », vol. 7, n° 3-4, 2003. Voir le site personnel de Jacques André : <http://jacques-andre.fr/>

Certains éléments de l'en-tête contiennent du texte libre, d'autres sont groupés :

- les éléments dont les noms se terminent par *Stmt* (*Statement*) englobent un groupe d'éléments divers ;
- les éléments dont les noms se terminent par *Desc* (*Description*) contiennent une description.

L'en-tête comprend quatre parties :

1. `<fileDesc>` une description bibliographique du fichier électronique regroupant le titre électronique, les responsables de l'édition et de la publication, ainsi que la description bibliographique de l'œuvre originale ;
2. `<encodingDesc>` le rapport entre le texte électronique et la source originale ;
3. `<profileDesc>` une description des aspects non bibliographiques ;
4. `<revisionDesc>` un historique des révisions du fichier électronique.

L'en-tête doit prendre en compte l'aspect juridique lié à la propriété intellectuelle et au droit d'auteur.

Il peut être réduit à quelques lignes : le titre électronique, l'éditeur électronique et une bibliographie sommaire de la source originale. Nous proposons un en-tête minimal qui doit avoir la structure suivante :

```
<teiHeader>
  <fileDesc>
    <titleStmt>
      <title>...</title>
    </titleStmt>
    <publicationStmt>
      <authority>...</authority>
    </publicationStmt>
    <sourceDesc>
      <p>...</p>
    </sourceDesc>
  </fileDesc>
</teiHeader>
```

Remarquons l'indentation des lignes dans notre présentation : ces retraits permettent une bonne appréciation visuelle de la structure du document. L'indentation n'est pas obligatoire, mais elle est fortement conseillée.

Dans l'exemple des *Discours*, l'en-tête peut être réduit selon l'encodage suivant :

```

<teiHeader>
  <fileDesc>
    <titleStmt>
      <title>Discours non plus mélancoliques que divers,
        attribué à Elie Vinet, Poitiers, 1556: édition
        électronique format XML/TEI P5.</title>
    </titleStmt>
    <publicationStmt>
      <authority>Université François-Rabelais</authority>
    </publicationStmt>
    <sourceDesc>
      <p>...</p>
    </sourceDesc>
  </fileDesc>
</teiHeader>

```

Cependant, il est recommandé de le compléter en apportant des informations sur la source originale, ou en donnant les modalités de l'édition électronique. L'expérience a montré que plus riche est l'en-tête, plus le fichier est réutilisé. L'en-tête TEI est développée en fin de manuel.

2) *Le texte de l'œuvre*

Le texte est structuré au maximum en cinq composantes ; deux sont obligatoires et trois optionnelles :

1. <text> élément obligatoire, il peut être unitaire ou composite comme dans un recueil;
2. <front> élément optionnel, il regroupe les pièces situées avant le début du texte lui-même comme la page de titre de l'ouvrage et les pièces liminaires ;
3. <group> élément optionnel, il regroupe plusieurs textes ;
4. <body> élément obligatoire, il regroupe le corps entier d'un texte unitaire ;
5. <back> élément optionnel, il regroupe les annexes.

La structure du texte est décrite avec plus de détails à la fin du manuel. Nous présentons ici une structure simple.

L'élément <text> représente le début du texte de la source originale. Il suit immédiatement la balise indiquant la fin de l'en-tête TEI </teiHeader> et précède les éléments <front> et <body>. Par exemple :

```

<text>
  <front>
    ...
  </front>
  <body>
    ...
  </body>
</text>

```

L'élément `<front>` contient tous les éléments qui précèdent le corps du texte au sens strict. Imbriquée dans l'élément `<front>`, la page de titre nécessite l'élément `<titlePage>` ; elle peut être suivie de pièces liminaires comme un privilège, une dédicace, une préface, des textes d'autres auteurs, etc. Tous ces éléments sont compris dans l'élément `<front>`.

```
<text>
  <front>
    <titlePage>
      <docTitle>
        <titlePart> ... </titlePart>
      </docTitle>
      <docImprint>
        ...
      </docImprint>
    </titlePage>
  </front>
</text>
```

La page de titre reprend tous les éléments de la page de titre de l'œuvre originale comme le titre, l'auteur, l'adresse typographique, la date, etc. Nous reprendrons en détail ces éléments dans la quatrième partie du manuel.

L'élément `<body>` est utilisé pour marquer le contenu de l'œuvre dans sa globalité.

Enfin, l'élément `<back>` contient toutes sortes d'annexes qui suivent le corps du texte. Par exemple, il est utilisé pour les corrections ajoutées par l'imprimeur ou encore pour le colophon quand il est imprimé en fin d'ouvrage.

II Balisages

L'encodage se distingue de l'affichage en HTML qui n'est pas pris en charge par XML/TEI, mais par une **feuille de style**.

Le projet BVH repose sur le respect patrimonial des sources, qu'elles soient littéraires ou historiques. Il peut être important de préserver les changements de caractères et les alternances typographiques, et de spécifier l'emplacement des illustrations. De même, le marquage a un sens qui favorise l'analyse historique et l'interprétation des sources. Le respect de la typographie donne par ailleurs une meilleure concordance avec les images numériques des livres imprimés (les facsimilés, selon la terminologie TEI) et améliore la lisibilité du texte. Le projet BVH ne prévoit pas la restitution diplomatique intégrale de la source.

1. Balisages formels et sémantiques

Les balisages formels et typographiques observent la restitution patrimoniale de la source et les balisages sémantiques donnent une première analyse du texte concernant principalement les noms et les dates.

1) *Balisage formel*

Une œuvre est structurée en un ensemble de parties qui peuvent être divisées et subdivisées en chapitres, sections, paragraphes, etc. Nous distinguons plusieurs marquages : le premier varie selon le type de texte à encoder ; le second concerne la mise en forme ; le troisième précise les interventions éditoriales sur le texte.

(1) *Le balisage structurel*

Le corps du texte balisé avec l'élément `<body>` est structuré selon le genre de texte : un recueil de poèmes n'a pas la même structure qu'une pièce de théâtre, un roman et encore moins un dictionnaire. Cependant tous ces textes ont en commun d'être souvent répartis en divisions, d'avoir des retours à la ligne, des sauts de lignes et des sauts de pages.

1. Division `<div>` associé à l'élément `<head>`

L'élément `<div>` est une subdivision du corps du texte. Ce n'est pas un élément obligatoire. Quand il est utilisé, il peut être accompagné d'un élément optionnel `<head>` qui identifie la division.

Les éléments `<div>` peuvent s'emboîter les uns dans les autres. Ils sont dotés d'un attribut de typologie pour les caractériser, ainsi que de l'attribut `@n` :

```
<div n="1" type="chapter">
```

ce qui signifie que la division est le chapitre numéro 1.

L'élément `<head>` contient toute forme de titre de la division, par exemple :

```
<head>ARGUMENT.</head>
```

ce qui signifie que le texte réel est « ARGUMENT. » Cet élément est également utilisé en dehors de l'élément `<div>` comme nous le verrons par la suite.

2. Paragraphe : `<p>`

L'élément `<p>` indique un paragraphe, segmentation très fréquente pour les textes en prose. Parfois, il n'est pas nécessaire de découper le texte en divisions lorsqu'il n'y a pas de sous-parties. Dans ce cas, après l'élément `<body>`, nous mettons directement l'élément `<p>`¹³.

Par exemple, la préface des *Euvres* de Louise Labé ne nécessite pas un découpage en divisions, les paragraphes suffisent :

```
<text type= "preface">
  <front>
    ...
  </front>
  <body>
    <p>Estant le temps venu...</p>
  </body>
</text>
```

C'est la **feuille de style** qui décide de l'affichage du paragraphe : elle peut montrer uniquement un retour à la ligne ou un paragraphe suivi d'un saut de ligne. Si un retrait de la première ligne est exigé, il devra être notifié manuellement grâce à l'attribut `@rend` et la valeur « indent » :

```
<p rend="indent">
```

Attention : certains imprimés ont des paragraphes sans retrait et même sans alinéa. Ces blancs qui aèrent la page peuvent être balisés avec l'élément `<space>` combiné avec `<p>`.

3. Texte théâtral : `<sp>` pour *speech* et `<speaker>`

Les *Euvres* de Louise Labé comprennent des discours sous forme de répliques comme dans un texte théâtral. Nous avons utilisé les éléments `<sp>` pour *speech* et `<speaker>`, puis nous les avons combinés avec l'élément `<p>` :

¹³ Mais cela peut être indispensable pour le traitement de bases de données textuelles.

```
<sp who="#amour">
  <speaker>AMOUR</speaker>
  <p>Qui est cette fole qui me...</p>
</sp>
<sp who="#folie">
  <speaker>FOLIE</speaker>
  <p>Tu ne m'usses pù empescher,...</p>
</sp>
```

L'élément `<sp>` utilise l'attribut `@who` pour différencier les locuteurs.

Les indications de mise en scène (didascalies) peuvent être encodées grâce à l'élément `<stage>` :

```
<stage>Folie se fait invisible, tellement, qu'Amour ne la
peut assener.</stage>
```

4. Prose et vers : `<lg>` pour *linegroup* et `<l>` pour *line*

Pour coder la structure des textes poétiques, deux éléments ont été utilisés. Dans les élégies et les sonnets, l'élément `<lg>` pour *linegroup* contient un groupe de lignes de vers, la strophe, et `<l>` ou *line* une seule ligne. L'élément `<p>` est alors inutile.

```
<lg n="2">
  <l>PAR TOY, AMI, TANT VESQUI ENFLAMMEE,</l>
  <l>QU'EN LANGUISSANT PAR FEU SUIS CONSUMEE,</l>
  <l>QUI COUVE ENCOR SOUS MA CENDRE EMBRAZEE,</l>
  <l>SI NE LA RENS DE TES PLEURS APAIZEE.</l>
</lg>
```

Ici l'élément `<lg>` utilise l'attribut `@n` pour numéroté les strophes ; l'élément `<l>` pourrait également utiliser cet attribut pour numéroté les lignes.

5. Tête et fin de division : `<head>` et `<trailer>`

Nous avons vu l'utilisation de l'élément `<head>` qui suit l'élément `<div>` pour indiquer un sous-titre. Les exemples étant nombreux et répartis tout au long du manuel, nous ne revenons pas sur la spécificité de cet élément souvent imbriqué dans un autre élément.

À la fin d'une division, il apparaît parfois une mention explicite de fin. L'élément `<trailer>` en contient la formule :

```
<trailer>FIN DES EUVRES DE LOUÏZE LABE' LIONNOIZE.</trailer>
```

(2) La mise en forme typographique

Les principes adoptés par le projet BVH (à la suite de ce qui a déjà été pratiqué dans la base textuelle Epistemon) stipulent que le repérage des éléments textuels s'effectue par rapport à la disposition de la page originale : respect du lignage et de la pagination. Des éléments servent à marquer cette segmentation fondamentale.

1. Retour à la ligne <lb> pour *linebreak*

Cet élément marque le début d'une nouvelle ligne. Comme le contenu de l'élément est vide, il est raccourci ainsi :

<lb/>

Si l'on veut numéroté les lignes, on ajoute l'attribut @n.

2. Coupure de mots en fin de ligne

Pour indiquer une coupure de mots en fin de ligne et afin de répondre au besoin des requêtes sur des mots entiers, un élément manque dans le protocole TEI. En effet, ces coupures doivent être encodées selon que le tiret de continuation en fin de ligne est présent ou non :

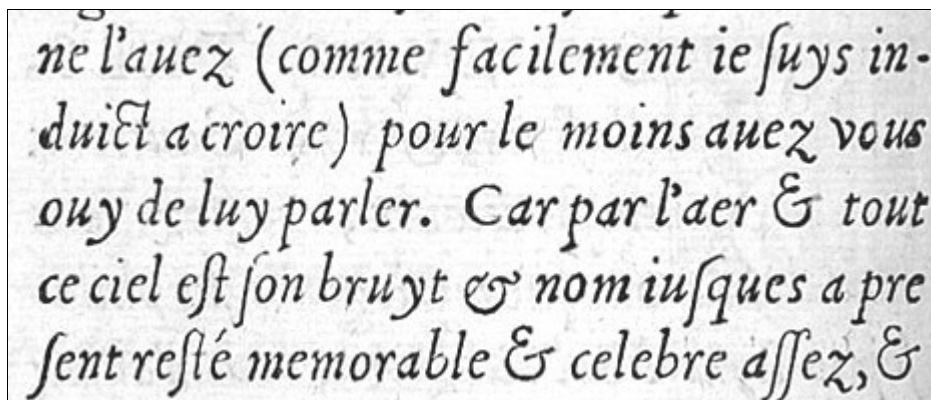


Illustration 3: Fins de ligne dans le Tiers Livre de Rabelais 1546, prologue (p. 2)

En attendant une solution globale pour ce problème, le balisage s'effectue comme suit :

- Avec tiret :

<p>... (comme facilement je suys in-<lb/>duict a croire)...</p>

- Sans tiret :

<p>...jusques a pre<lb/>sent...</p>

Le même problème se pose pour les mots coupés en fin de page.

3. Saut de page <pb> pour *pagebreak*

La limite entre une page et la suivante est encodée grâce à l'élément <pb>. Tout comme le *line-break*, il peut comprendre l'attribut *n* afin de numéroté les pages. Par exemple :

```
<pb n="10"/>
```

Cet élément est essentiel pour le lien avec l'image des pages du livre numérisé.

Remarques :

- Ne pas oublier que la première page doit être aussi numérotée ;
- Le numéro *n* n'est pas obligatoirement identique à celui qui figure sur la page du livre, encodé grâce à l'élément <fw> (voir paragraphe suivant) ;
- Si la page du texte pointe vers son image en fac-similé, l'élément <pb> est indispensable.

4. En-tête et pied de page : <fw> pour *Forme Work*

Il est utile de signaler les titres courants qui sont sur la même ligne que les numéros de page, et les pieds de page. L'élément <fw> pour *forme work* sert dans les deux cas.

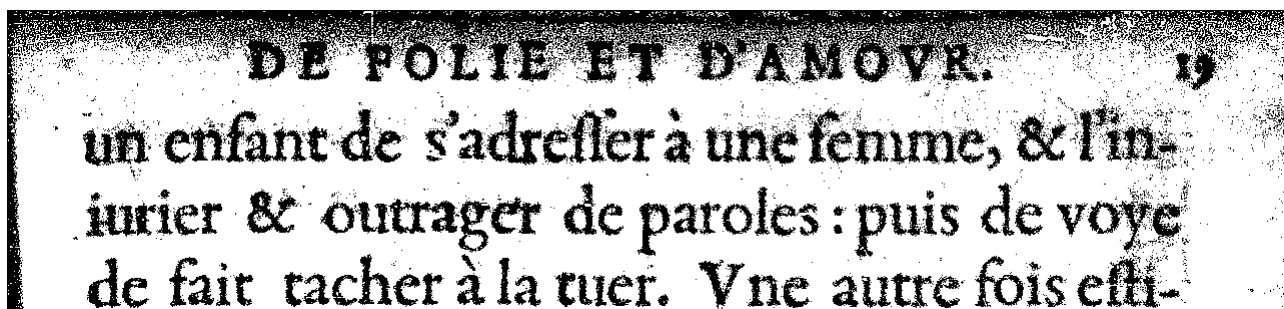


Illustration 4: Titre courant des Euvres de Louise Labé (p. 19)

Pour le titre courant, cet élément précise le rendu typographique, sa place dans la page et son type grâce aux attributs @place et @type. Dans l'exemple, le titre courant et le numéro de page seront encodés ainsi :

```
<fw place="top-center" type="head">DE FOLIE ET D'AMOUR</fw>
<fw place="top-right" type="pageNum">19</fw>
```

Les valeurs des attributs sont :

- *top* pour le haut de page et *bot* (pour *bottom*) pour le bas de page,
- *center*, *left* ou *right* pour le centre, la gauche ou la droite de la page,
- *head* pour le titre courant,

- *pageNum* pour le numéro de la page ou du feuillet.

Pour le pied de page, l'élément précise la place des « signatures » et des « réclames » (au sens typographique). Prenons la page 3 du même ouvrage :

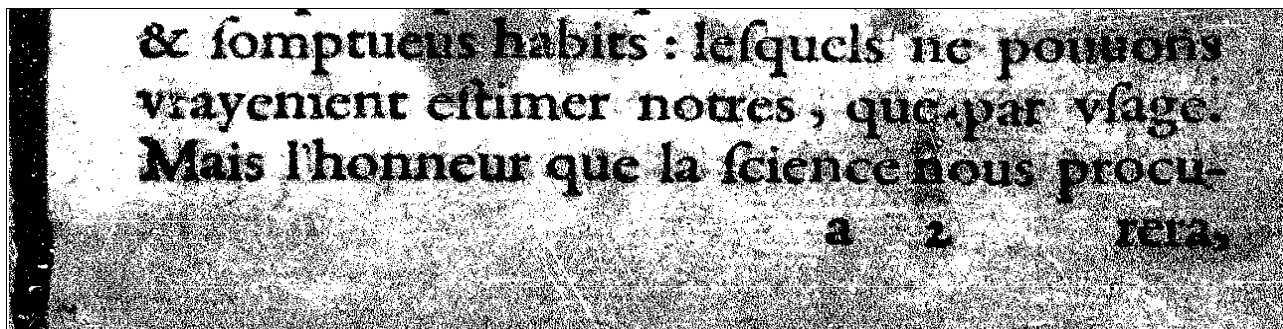


Illustration 5: Pied de page des Euvres de Louise Labé (p. 3)

La signature et la réclame seront encodées ainsi :

```
<fw place="bot-center" type="sig">a 2</fw>
<fw place="bot-right" type="catch">rera,</fw>
```

Les attributs utilisés sont les mêmes, seule leur valeur change :

- *sig* pour *signature*, c'est-à-dire la signature,
- *catch* pour *catchword*, c'est-à-dire la réclame.

Cet élément permet une restitution fidèle autant que possible de la structure physique de la page originale, car ici la signature n'est pas au centre.

5. Mise en valeur : `<hi>` pour *highlighted*

L'élément `<hi>` marque un mot ou une série de mots graphiquement distincts du texte environnant. Il utilise l'attribut `@rend` pour *rendition* qui rend l'aspect physique de la chaîne de caractères.

```
<hi rend="it">
```

L'attribut reflète l'aspect visuel et ne donne aucune interprétation. Ici, l'attribut `@rend` signale la présentation physique de la police qui peut être : *italic* ou *it*, *sc* pour *small caps*, *bold* (gras), mais également *center*. Toutes ces valeurs sont déclarées dans la feuille de style.

6. Annotations imprimées : `<note>`

L'élément `<note>` est utilisé pour indiquer une note située dans la marge qu'elle soit ou non signalée par un appel de note. Dans les *Euvres* de Louise Labé, une note imprimée figure en marge :

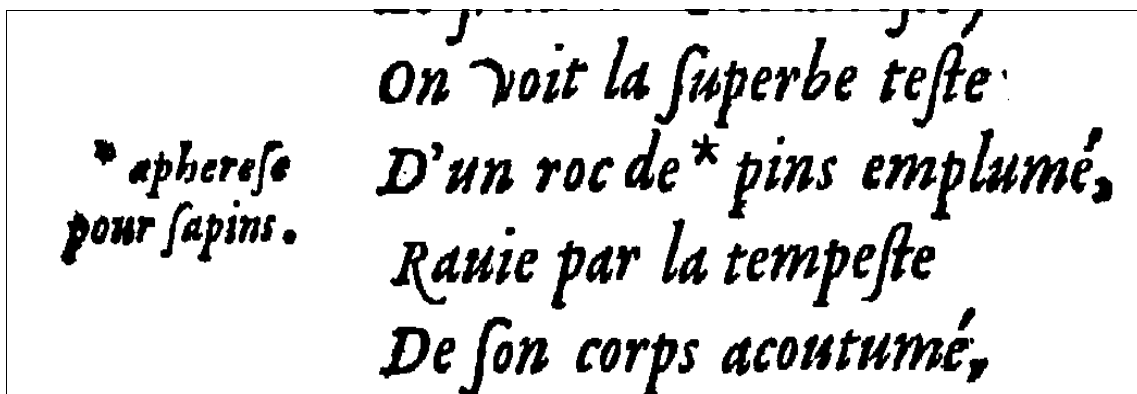


Illustration 6: Appel de note dans les *Euvres de Louise Labé* (p. 154)

Nous proposons l'encodage suivant :

```
<lg>
...
<l>On voit la superbe teste</l>
<l>D'un roc de*<note place="margin-left">* apherese<lb/>pour
sapins.</note> pins emplumé,</l>
<l>Ravie par la tempeste</l>
<l>De son corps acoutumé,</l>
...
</lg>
```

Il est possible de dresser une typologie de ces annotations qui sont souvent des explications, des sous-titres, des entrées d'index, grâce à l'attribut @type.

7. Listes et tableaux : <list> et <table>

- L'élément <list> est utilisé pour marquer toutes sortes de listes simples. L'élément <p> le précède pour indiquer un paragraphe, s'il y a lieu. Les éléments <item> séparent toutes les séquences de la liste. L'élément <head> contient l'en-tête de la liste.

Nous donnons ici deux exemples : une liste simple et une liste à deux colonnes.

```
<p>
  <list>
    <head>FAUTES A CORRIGER<lb/>en l'impression.</head>
    <item>Page 33. lig. 7. lis plait, pour pleint</item>
    <item>Page 67. lig. 15. lis fait, pour fay</item>
    <item>Page 102. lig. 16. lis Tous ses voisins</item>
    <item>Page 141. lig. 5. lis Il nous faut eores aquiter</item>
  </list>
</p>
```

Les listes peuvent être également marquées à l'aide de lettres ou de chiffres, ou bien avec des puces. Il s'agit alors de mettre l'attribut @type et la valeur correspondante : soit « ordered » pour ordonner, soit « bulleted » pour les listes marquées par des puces,

soit « simple » pour des listes non marquées ou enfin « runon » pour indiquer le style en continu de la liste sans retour à la ligne.

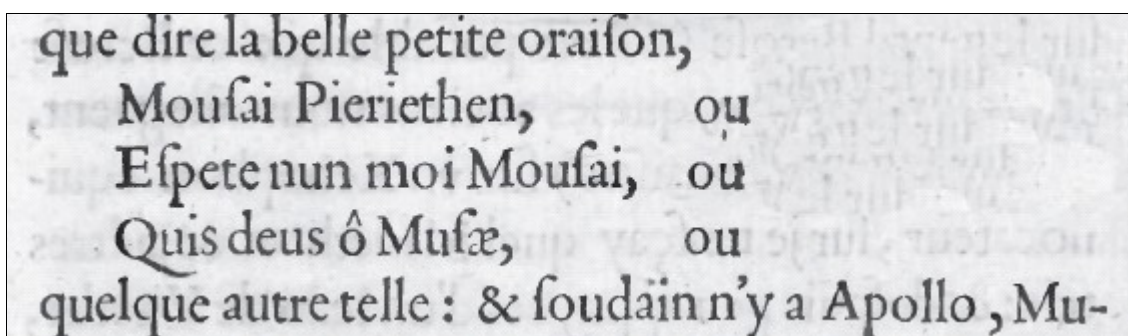
Un tableau simple à deux colonnes peut être traité comme une liste de type glossaire (ou notices de dictionnaire) ; nous utilisons alors l'élément `<label>` à la place du premier `<item>` qui contient alors soit le terme défini dans un glossaire, soit l'étiquette associée à un élément dans une liste.

- Quand le tableau est plus complexe, l'élément `<table>` est utilisé avec d'autres éléments imbriqués pour constituer un tableau : l'élément `<row>` (rangée) constitue une ligne de la table et l'élément `<cell>` (cellule) chaque colonne ; on peut y ajouter un élément `<head>` pour donner un titre au tableau. Les attributs `@rows` ou `@cols` (colonnes) indiquent le nombre de lignes ou de colonnes du tableau.

Il s'agit d'effectuer un choix éditorial entre les deux modes : liste ou tableau. Lorsqu'on veut respecter la présentation d'une œuvre comme dans le projet BVH, le choix sera dicté par la représentation la plus fidèle à l'original.

■ *Discours non plus mélancoliques que divers :*

Prenons un exemple tiré du chapitre 1 des *Discours*, qui contient un tableau simple à deux colonnes :



que dire la belle petite oraison,	
Moufai Pieriethen,	ou
Espete nun moi Moufai,	ou
Quis deus ô Musæ,	ou
quelque autre telle : & soudain n'y a Apollo, Mu-	

Illustration 7: Tableau simple à deux colonnes tiré des *Discours* (p. 1)

La présentation en tableau à deux colonnes a été préférée à une liste de type « glossaire », car la première colonne ne correspond ni à une étiquette ni à un label mais à du texte, ce qui donne le résultat suivant :

```
<table>
  <row cols="2">
    <cell>Mousai Pieriethen,</cell><cell>ou</cell>
  </row>
  <row>
    <cell>Espete nun moi Mousai,</cell><cell>ou</cell>
  </row>
  <row>
    <cell>Quis deus ô Musæ,</cell><cell>ou</cell>
  </row>
</table>
```

Le résultat à l'écran est approchant de l'original, à savoir :

que dire la belle petite oraison,
 Mousai Pieriethen, ou
 Espete nun moi Mousai, ou
 Quis deus ô Musae, ou
 quelque autre telle: & soudain n'y a Apollo, Mu-

Illustration 8: Affichage d'un tableau simple à deux colonnes tiré des Discours (p. 1)

- Dépliant inséré page 9 des *Discours non plus mélancoliques que divers*

Un tableau imprimé se trouve sur le dépliant inséré page 9 du même livre ; il comporte huit colonnes. Nous en présentons les trois premières :

Les Iuifs les nom- moient jadis en ceste forte, en leur langage,	Les Romains Chretiens jadis,	Les Portugalois aujourd'hui,
Prima Sabbati, Secunda Sabbati, Tertia Sabbati, Quarta Sabbati, Quinta Sabbati, Sexta Sabbati, Sabbatum.	Dominica dies, Secunda Feria, Tertia Feria, Quarta Feria, Quinta Feria, Sexta Feria, Sabbatum.	Domingo, Secunda Feira, Terça Feira, Quarta Feira, Quinta Feira, Sesta Feira, Sábado.

Illustration 9: Dépliant des Discours (p. 9)

Dans cet exemple plus complexe, nous avons disposé verticalement les éléments `<cell>` par commodité. L'attribut `@role` a été ajouté dans l'élément `<row>` : la valeur « label » est suggérée pour des étiquettes – ici les peuples – et la valeur « data » pour les données proprement dites :

```
<table rows="8" cols="9">
  <row role="label">
    <cell>Les Juifs les nom-
    <lb/>moient jadis en ceste
    <lb/>sorte, en leur langage,</cell>
    <cell>Les Romains<lb/>chretiens jadis,</cell>
```

```

        <cell>Les Portugalois<lb/>aujourd'hui,</cell>
        <cell>Les Castil-<lb/>lans aujour-<lb/>hui,</cell>
        <cell>Les Gaulois<lb/> & amp; François</cell>
        <cell>Italiens,</cell>
        <cell>Alemans,</cell>
        <cell>Bretons,</cell>
        <cell>Basques,</cell>
    </row>
    <row role="data">
        <cell>Prima Sabbati,</cell>
        <cell>Dominica dies,</cell>
        <cell>Domingo,</cell>
        <cell>Domingo,</cell>
        <cell>Dimanche,</cell>
        <cell>Domenica,</cell>
        <cell>Sondach,</cell>
        <cell>Dei Suul,</cell>
        <cell>Igándeia,</cell>
    </row>
    <row role="data">
        <cell>Secunda Sabbati,</cell>
        <cell>Secunda Feria,</cell>
        <cell>Secunda Feira,</cell>
        <cell>Lunes</cell>
        <cell>Lundi,</cell>
        <cell>Lunedì,</cell>
        <cell>Maendach,</cell>
        <cell>Dei luun,</cell>
        <cell>Astelechéna,</cell>
    </row>
    ...
</table>

```

Il est bien entendu possible d'ajouter toutes sortes d'attributs dans les éléments des tableaux comme l'attribut @rend.

Le résultat est proche de la page originale :

Les Juifs les nom-	Les Romains	Les Portugalois
moient jadis en ceste	chretiens jadis,	aujourd'hui,
sorte, en leur langage,		
Prima Sabbati,	Dominica dies,	Domingo,
Secunda Sabbati,	Secunda Feria,	Secunda Feira,
Tertia Sabbati,	Tertia Feria,	Terça Feira,
Quarta Sabbati,	Quarta Feria,	Quarta Feira,
Quinta Sabbati,	Quinta Feria,	Quinta Feira,
Sexta Sabbati,	Sexta Feria,	Sesta Feira,
Sabbatum.	Sabbatum.	Sábado.

*Illustration 10: Affichage du dépliant des Discours
(p. 9)*

Table des matières

Une table des matières (*toc* ou *table of contents*) a une spécificité : l'alignement à droite du numéro de page, comme dans les *Discours*.

Table des choses contenues en ce liure.	
De noz Historiens qui cherchent l'origine de noz Gaulois & François, Chapitre premier,	page 1.
Des noms des iours de la semaine, Chap. 2.	4.
Que c'est que More, Chap. 3.	13.
Histoire d'Hercules, Pyrene, Bebrix, Brettan, Celtine, Celte, Gaule Celtique, Chap. 4.	13.
Des Grammeriens François, Chap. 5.	18.
Du nom de la riuiere d'Arar, qui s'appelle aujourduy la Saone, & de la ville de Longdown, qui est maintenant Lion sur le Rosne, Chap. 6.	21.

Illustration 11: Table des matières des *Discours*

Voici l'encodage qui en résulte :

```
<div type="toc">
  <table>
    <head>Table des choses contenues en ce livre.</head>
    <row>
      <cell>De noz Historiens qui cherchent l'origine de noz Gau-
      <lb/>lois & François, Chapitre premier,</cell>
      <cell>page 1.</cell>
    </row>
    <row>
      <cell>Des noms des jours de la semaine, Chap. 2.</cell>
      <cell>4.</cell>
    </row>
    <row>
      <cell>Que c'est que More, Chap. 3.</cell>
      <cell>13.</cell>
    </row>
    <row>
      <cell>Histoire d'Hercules, Pyrene, Bebrix, Brettan,
      Celti-
      <lb/>ne, Celte, Gaule Celtique, Chap. 4.</cell>
      <cell>13.</cell>
    </row>
    <row>
      <cell>Des Grammeriens François, Chap. 5.</cell>
      <cell>18.</cell>
    </row>
    <row>
      <cell>Du nom de la riuiere d'Arar, qui s'appelle aujourduy
      la Saone, & de la ville de Longdown, qui est mainte-
      nant Lion sur le Rosne, Chap. 6.</cell>
      <cell>21.</cell>
    </row>
  </table>
</div>
```

```

        <cell>Des Grammeriens François, Chap, 5.</cell>
        <cell>18.</cell>
    </row>
    <row>
        <cell>Du nom de la riviere d'Arar, qui s'appelle
        aujourd'uy
        <lb/>la Saone, & de la ville de Lougdoun, qui est
        mainte-
        <lb/>nant Lion sur le Rosne, Chap. 6.</cell>
        <cell>21.</cell>
    </row>
    ...
</table>
</div>

```

Et voici le résultat qui s'affiche :

Table des choses contenues en ce livre.	
De noz Historiens qui cherchent l'origine de noz Gau- lois & François, Chapitre premier, page	I.
Des noms des jours de la semaine, Chap. 2.	4.
Que c'est que More, Chap. 3.	13.
Histoire d'Hercules, Pyrene, Bebrix, Brettan, Celti- ne, Celte, Gaule Celtique, Chap. 4.	13.
Des Grammeriens François, Chap. 5.	18.
Du nom de la riviere d'Arar, qui s'appelle aujourd'uy la Saone, & de la ville de Lougdoun, qui est mainte- nant Lion sur le Rosne, Chap. 6.	21.

Illustration 12: Affichage d'un extrait de la table des matières des Discours

Il reste à centrer le titre et à mettre le texte en italiques pour que le résultat soit approchant de l'original.

Dans le *Tiers Livre* de Rabelais, la table des matières est différente, comme le montre cet extrait :

P	ROLOGVE du docteur Rabelais au- theur du liure. feul. 4. Comment Pantagruel transporta vne colo- nie de Vtopiens en Dipsodie. f. 11.
	Cōment Panurge feut faict Chastellain de Salmiguondin en Dipsodie, & man- geoit son bled en herbe. f. 15.
	Comment Panurge loue les Debiteurs & Emprunteurs. f. 18.
	Continuation du discours de Panurge, a la louange des Presteurs & Debiteurs. feucillet 22.

Illustration 13: Extrait de la table des matières du Tiers
Livre de Rabelais 1552

Tout en écartant à ce stade la lettrine P, on peut choisir d'encoder la table des matières sous forme de liste, à savoir :

```
<div type="toc">
  <list type="contents">
    <item>PROLOGUE du
    <lb/>docteur Rabelais au-
    <lb/>theur du livre.<ref><abbr>feul.</abbr>4.</ref></item>
    ...
  </list>
</div>
```

Dans une liste de type « contents », on peut prévoir d'aligner à droite son contenu automatiquement. L'élément `<ref>` permet de pointer vers la page correspondante grâce à l'attribut `@target`. Il doit être déclaré dans la feuille de style.

(3) Les interventions éditoriales

Les interventions éditoriales sont de deux types : celles que nous signalons, parce qu'elles figurent dans le texte imprimé, comme les abréviations ; celles que nous ajoutons pour une meilleure compréhension du texte ou de la transcription, qui sont des annotations.

1. Abréviations : `<abbr>`

Les abréviations du texte sont de deux types : d'une part, sous la forme d'une suppression de lettres (« F. » pour « François ») et, d'autre part, sous la forme d'un « brévigraphe » qui symbolise la réduction d'un mot comme dans « vous » abrégé en **vo⁹**.

Dans le projet BVH, les mots abrégés de type *brevigraph*, très fréquents à la Renaissance, sont développés. Le choix éditorial a été de ne pas les marquer avec l'attribut « brevigraph » afin d'alléger l'encodage. Une transcription diplomatique les indiquerait.

L'autre type d'abréviation concerne les suppressions de lettres comme dans les *Euvres* de Louise Labé où les noms des locuteurs « Amour » et « Folie » sont abrégés en « AM. » et « FOL. » :

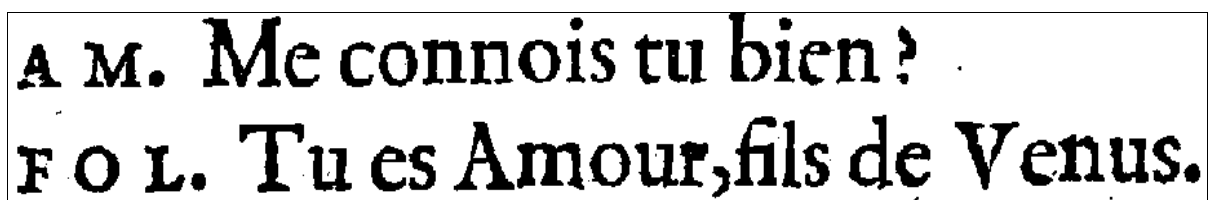


Illustration 14: *Euvres de Louise Labé* (p. 11)

Ce second type d'abréviation peut faire l'objet d'un développement grâce à l'élément `<expn>` :

```
<abbr>AM.</abbr>
<expn>AMOUR</expn>
```

Seule l'abréviation « AM. » s'affichera en HTML, la forme développée étant réservée aux requêtes des moteurs de recherche.

2. État de l'œuvre : `<unclear>`, ``, `<gap>`, `<supplied>` et `<add>`

Des fragments de texte peuvent avoir été ôtés, effacés ou ajoutés par accident ou volontairement par des mains différentes : imprimeur, lecteur, etc. Cinq éléments peuvent être combinés pour signaler ces problèmes de lecture et d'intervention.

- L'élément `<unclear>` est rarement utilisé dans les livres imprimés. Cependant, quelques taches et des caractères usés peuvent être marqués ainsi. L'attribut `@reason` en indique la raison. Prenons un exemple dans le *Tiers Livre* de 1546 :

```
<p>...la tierce partie du monde, <unclear
reason="damaged">bestes</unclear><lb/>...</p>
```

- Le second élément est `` pour *deletion*. Dans les *Discours*, une lettre a été barrée :

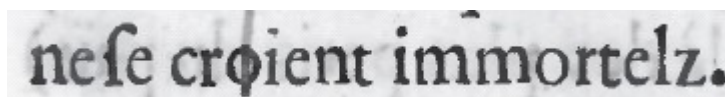


Illustration 15: *Rature dans les Discours* (p. 66)

Cette intervention est encodée ainsi :


```
<p>...ne se cr<del>o</del>ient immortelz...</p>
```

Quand le nom de responsable est connu, l'attribut @hand peut être utilisé pour l'identifier. Dans le cas des *Discours*, il s'agit probablement de l'éditeur Enguilbert de Marnef (selon Marie-Luce Demonet).

```
<p>...ne se cr<del hand="#marnef" resp="#MLD">o</del>ient
    immortelz...</p>
```

- L'élément <gap> concerne des passages qui peuvent être absents pour des raisons matérielles ou éditoriales et que l'on renonce à restituer. Il doit être accompagné de l'élément <desc> pour *description* qui autorise l'insertion d'autres éléments.

```
<gap reason="grec">
  <desc>
    <bibl><title>ODE GRECQUE</title>, attribuée à <author>Jacques
    Peletier du Mans</author>, <author>Jean-Antoine de
    Baïf</author>, <author>Antoine du Moulin</author> ou
    <author>Henri Estienne</author></bibl>
  </desc>
</gap>
```

Quand il est vide de contenu, il peut être raccourci ainsi :

```
<gap reason="illegible"/>
```

Cet élément peut également être utilisé à la place de l'élément <unclear> quand une partie du texte à transcrire est complètement illisible ; l'attribut @reason y est spécifié également.

- Un quatrième élément, <supplied>, est utilisé pour compléter <gap> ou <unclear>. Il marque le texte restitué et sa source grâce à l'attribut @source qui indique l'édition ou l'exemplaire qui a permis la restitution, référence préalablement définie dans le Header. C'est ainsi que pour l'encodage du *Courtisan* de Baldassare Castiglione de 1538, nous avons consulté une édition de 1537 :

```
<teiHeader>
...
  <normalization>
    <bibl xml:id="ed1537"
xml:base="http://gallica2.bnf.fr/ark:/12148/bpt6k49708n.r=le+courtisan.-
langFR#"> <title>Les Quatre livres du Courtisan, </title> [trad. Jacques
Colin], <pubPlace>[Lyon]</pubPlace>, <editor>[Denys de Harsy]</editor>,
<date>[1537], </date>
    </bibl>
  </normalization>
...
</teiHeader>
<text>
...
  <supplied source="#ed1537">labhorrist</supplied>
...
</text>
```

- L'élément `<add>` permet d'ajouter un caractère ou un texte non imprimé dans l'ouvrage ; parmi les attributs possibles, retenons `@hand` et `@place`. Dans le privilège du roi des *Discours non plus mélancoliques*, un exemple nous est donné avec cette correction d'époque en troisième ligne :

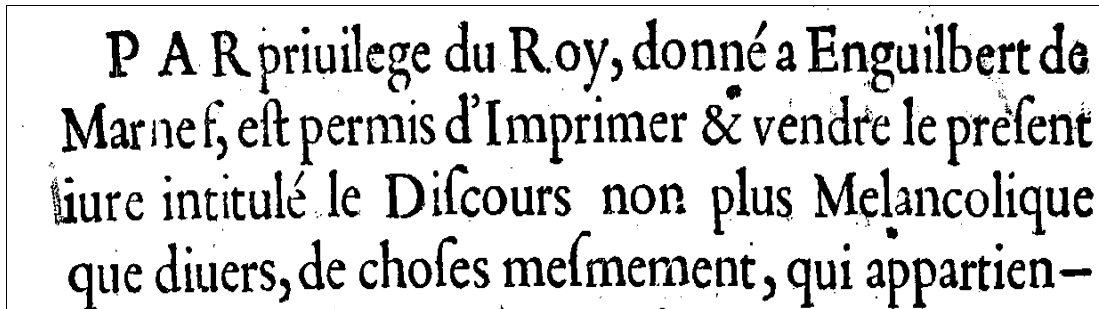


Illustration 16: Ajout dans le privilège du roi des *Discours*

Nous proposons l'encodage suivant pour représenter le « l » de « livre » ajouté à la main :

```
<p>
<lb/>PAR privilege du Roy, donné a Enguilbert de
<lb/>Marnef, est permis d'Imprimer &amp; vendre le present
<lb/><add hand="editor">l</add>ivre intitulé le <title>Discours non plus
Melancolique
<lb/>que divers, de choses mesmement, qui appartiennent
<lb/>a nostre France : &amp; a la fin La maniere de bien
<lb/>&amp; justement entoucher les Lucs &amp; Guitermes.</title>...
</p>
```

Dans cet exemple, l'ajout semble avoir été effectué par l'éditeur, ici Enguibert de Marnef.

Quand il s'agit d'ajouts de mains anonymes, nous les encodons sous la forme suivante :

```
<add place="margin-left" hand="unknown">...</add>
```

La feuille de style définit l'emplacement de l'appel de note, qu'il soit à gauche ou à droite.

3. Erreurs et régularisations : `<sic>` et `<corr>`, `<orig>` et `<reg>`

D'autres interventions éditoriales concernent les corrections et les régularisations, base de tout travail d'édition. Elles utilisent les quatre éléments liés deux à deux, à savoir : `<sic>` et `<corr>` pour *correction* , `<orig>` pour *original form* et `<reg>` pour *regularization*.

- L'élément `<sic>` contient le texte original bien qu'il apparaisse incorrect ou inexact. Les erreurs manifestes peuvent être une erreur de pagination ou une inversion de lettres. Dans la *La Deffence, et illustration de la Langue francoyse* de Du Bellay, on lit :

... que la Nature a <sic>créés</sic>...

Ici, la correction n'est pas nécessaire. En revanche, les erreurs manifestes peuvent concerner la pagination ou une inversion de lettres. Pour les corriger, nous pouvons utiliser l'élément `<corr>`.

L'élément `<sic>` peut s'employer indépendamment de l'élément `<corr>`. Si ce dernier est utilisé, le choix éditorial est de privilégier la forme correcte du mot et de ne pas afficher la forme originale erronée.

- L'élément `<corr>` contient la forme correcte d'un mot rectifié par nous. Dans *La Deffence* encore (p. 5 recto), « Et tout ainsi qu'entres les Auteurs » est corrigé ainsi :

```
<p>...
    Et tout ainsi qu'
    <corr>entre</corr>
    les Auteurs...
</p>
```

Si la correction est apportée grâce à une autre édition, l'attribut `@source` peut être ajouté. C'est le cas dans le *Tiers Livre* de 1546 :

```
<corr source="#1552">ne</corr>
```

Enfin, certaines interventions éditoriales concernent les régularisations. Un mot (ou une expression) peut être régularisé avec les éléments `<orig>` et `<reg>` :

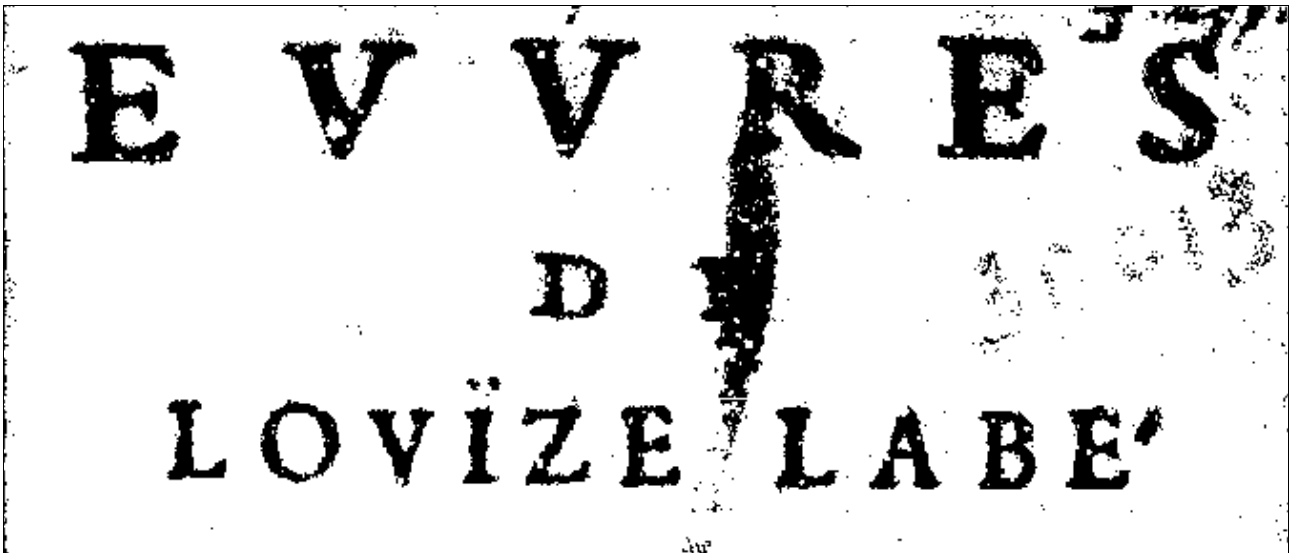


Illustration 17: Extrait de la page de titre des *Euvres de Louise Labé*

- L'élément `<orig>` contient la forme originale de la source, ici le titre de l'ouvrage et le nom de l'auteur :

```

<front>
<titlePage>
  <docTitle>
    <titlePart>
      <orig>EVVRES</orig>
    </titlePart>
  </docTitle>
  <docAuthor>
    <orig>Lovize Labe'</orig>
  </docAuthor>
</titlePage>
</front>

```

- L'élément <reg> contient la forme régularisée¹⁴.

```

<front>
<titlePage>
  <docTitle>
    <titlePart>
      <reg>OEUVRES</reg>
    </titlePart>
  </docTitle>
  <docAuthor>Louise Labé</docAuthor>
</titlePage>
</front>

```

La feuille de style applique les consignes d'affichage. Pour afficher le texte, des choix éditoriaux sont nécessaires pour les éléments concernant les erreurs, les régularisations mais également les abréviations : nous utilisons alors l'élément <choice>. Reprenons les exemples précédents auxquels nous appliquons ce nouvel élément :

- Les éléments <orig> et <reg> avec <choice>

```

<front>
<titlePage>
  <docTitle>
    <titlePart>
      <choice>
        <orig>EVVRES</orig>
        <reg>OEUVRES</reg>
      </choice>
    </titlePart>
  </docTitle>
  <docAuthor>Lovize Labe'</docAuthor>

```

¹⁴ La normalisation des noms de personnes et de lieux peut s'effectuer grâce aux index régularisés sur une plate-forme de travail adéquate.

```
</titlePage>
</front>
```

Selon le choix éditorial, la feuille de style peut afficher soit la version originale, « EVVRES », soit la version régularisée, « OEUVRES ».

- Les éléments `<sic>` et `<corr>` avec `<choice>`

```
<p>...
  Et tout ainsi qu'
  <choice>
    <sic>entres</sic>
    <corr>entre</corr>
  </choice>
  les Auteurs...
</p>
```

Dans ce second cas, la feuille de style affichera la version corrigée.

- Les éléments `<abbr>` et `<expan>` avec `<choice>`

```
<choice>
  <abbr>AM.</abbr>
  <expan>AMOUR</expan>
</choice>
```

Dans ce dernier cas, notre feuille de style affiche la version abrégée.

Enfin, ces éléments peuvent se combiner entre eux. Par exemple, dans le *Tiers Livre* de Rabelais, une main anonyme a rectifié l'erreur évidente de numérotation du chapitre 38 :

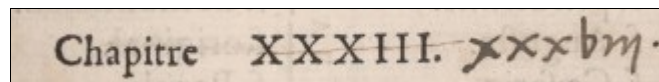


Illustration 18: Ajout et rature dans le Tiers Livre de Rabelais

Nous proposons l'encodage suivant :

```
Chapitre <sic><del>XXXIII.</del></sic><add hand="lecteur"
  place="online">xxxviij.</add>
```

L'attribut `@place` est ajouté pour spécifier l'emplacement de l'ajout dans le texte.

2) **Balisage sémantique**

Ce type de balisage permet de mettre en évidence les noms propres et des éléments temporels comme les dates. Des expressions ou des suites de mots peuvent également faire l'objet d'un marquage permettant de dresser une typologie en vue d'analyses ultérieures.

(1) *Les citations et références bibliographiques*

Dans *le Tiers Livre*, on dénombre une multitude de citations qui ne sont pas forcément délimitées par des guillemets ni facilement identifiables :

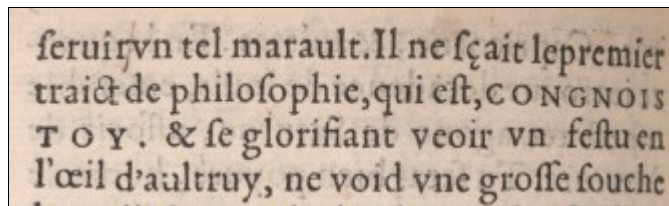


Illustration 19: Expression tirée du Tiers Livre de Rabelais (p. 83 verso)

A titre exceptionnel, nous envisageons de les encoder : c'est un choix éditorial. Dans cet exemple, nous encodons l'expression « Congnois toy » grâce à l'élément `<quote>` qui marque un passage ou une expression attribués à un agent extérieur au texte :

```
<lb/>servir un tel marault. Il ne sçait le premier
<lb/>traict de philosophie, qui est, <quote>CONGNOIS
<lb/>TOY</quote>. & se glorifiant veoir un festu en...
```

Quand l'expression devient une citation littéraire, nous utilisons l'élément `<quote>` imbriqué dans l'élément `<cit>` pour *citation*. À la page 37 des *Discours*, une citation pourrait être encodée avec trois éléments imbriqués : `<cit>`, `<quote>` et `<l>` pour chaque ligne de vers.

```
<cit>
  <quote>
    <l>Ainsi s'en va chatoulheus de la gorge</l>
    <l>Ledit valet, monté comme un saint George.</l>
  </quote>
</cit>
```

Par ailleurs, dans certains textes, apparaissent des références bibliographiques plus ou moins précises :

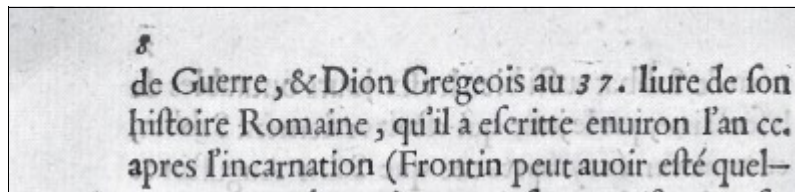


Illustration 20: Référence bibliographique dans les Discours
(p. 8)

Nous utilisons les trois éléments suivants : `<bibl>` pour *bibliographic citation*, `<author>` et `<title>`, les deux derniers étant imbriqués dans le premier :

```
<lb/>de Guerre, & <bibl><author>Dion Gregeois</author> au <hi  
rend="italic">37.</hi>livre de son  
<lb/><title>histoire Romaine</title></bibl>, qu'il a escritte environ  
l'<date when="200">an cc</date>.  
<lb/>apres l'incarnation (Frontin peut avoir esté quel-
```

L'élément `<bibl>` marque les références bibliographiques ici peu structurées. Dans notre exemple, seuls l'auteur et le titre sont cités¹⁵.

Nous combinons les citations avec les références bibliographiques quand une mention d'un autre ouvrage est donnée avec ses références :

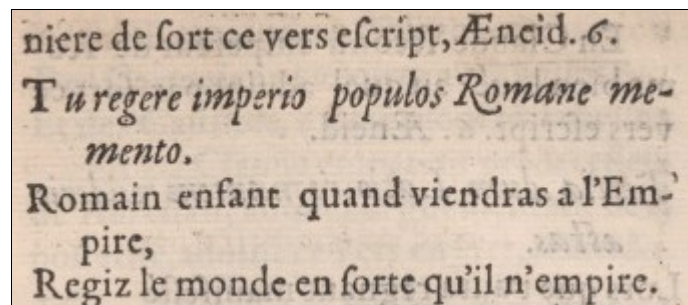


Illustration 21: Citation avec une référence biblio-
graphique (Tiers Livre de 1552 p. 38)

La citation qui apparaît littéralement est marquée grâce aux deux éléments complémentaires `<quote>` et `<bibl>`.

```
<lb/>niere de sort ce vers escript,  
<cit>  
<bibl><title>Aeneid</title>. 6.</bibl>  
<lb/>  
<quote>  
<l><hi rend="it">Tu regere imperio populos Romane me-  
<lb/>mento</hi>.</l>  
</quote>  
<quote>  
<l>Romain enfant quand viendras a l'Em-
```

15 Dans le cas de références complètes, nous conseillons d'autres éléments imbriqués comme `<date>`, `<editor>`, `<imprint>`, `<publisher>`, `<pubPlace>`, etc.


```
<lb/>pire,</l>
<l>Regiz le monde en sorte qu'il n'empire.</l>
</quote>
</cit>
```

Dans tous les cas, l'attribut qui est privilégié pour les éléments `<cit>` et `<quote>` est `@type`.

Dans certains passages, des paroles ou des pensées peuvent être mises en évidence. L'élément `<said>` est alors retenu. Dans les *Discours*, la page 38 en est un exemple :

```
<p>...<said>He, maistre<lb/> mon ami,</said> dit il,...</p>
```

L'attribut `@who` identifie éventuellement le locuteur.

(2) Les langues

Les langues utilisées dans les œuvres sont déclarées dans l'en-tête TEI¹⁶. Dans le corps du texte, l'élément `<foreign>` marque le mot ou l'expression écrit dans une autre langue que celle de l'œuvre ; on y associe l'attribut `@xml:lang` et, en valeur, le code de la langue. La nomenclature des codes standards correspond à la norme ISO 639¹⁷, comme ici pour le latin.

```
<foreign xml:lang="la">restare</foreign>
```

La norme ISO 639-2 a défini le français de la Renaissance (1400-1600) par le code "frm". Nous proposons de mettre cette valeur grâce à l'attribut `@xml:lang` dans l'élément `<text>` :

```
<text xml:lang="frm">
```

Par défaut le texte est en français de la Renaissance à l'exception des passages dans une autre langue signalée par l'élément `<foreign>`. La liste des langues et des dialectes que nous utilisons est donnée en annexe de ce manuel.

Parfois, le mot ou l'expression dans une langue autre que celle de l'œuvre est mise en italiques ; on utilise alors l'attribut `@rend` et la valeur « it » soit dans l'élément `<foreign>` soit en ajoutant l'élément `<hi>`. Ainsi, page 72 des *Discours*, plusieurs mots sont exprimés dans une autre langue que le français du XVI^e siècle et le marquage peut se faire sous deux formes :

```
<foreign xml:lang="la" rend="it">restare</foreign>
```

```
<foreign xml:lang="la"><hi rend="it">restare</hi></foreign>
```

Les deux balisages sont équivalents. Nous recommandons le second : il est plus long, mais il différencie nettement un balisage formel d'un balisage sémantique. De plus, la première solution complique le travail sur la feuille de style, le marquage formel et typographique devant être déclaré dans l'élément sémantique.

¹⁶ Voir la troisième partie du présent manuel.

¹⁷ Cf. <http://xml.coverpages.org/>

(3) L'onomastique : noms de personnes, dénomination sociale et toponymes

- `<name>`

L'élément `<name>` est générique ; nous l'utilisons pour les noms de personnes, qu'ils soient réels ou fictifs ; par exemple la déesse Minerve sera marquée ainsi :

```
<name>Minerve</name>
```

ou mieux encore :

```
<name type="mythology">Minerve</name>
```

- `<persName>`

L'élément `<persName>` pour *personal name* est plus précis que `<name>` car il permet une déclinaison importante dans la dénomination. Le nom de François Rabelais peut être marqué par l'élément `<name>` ou être précisé grâce à ce nouvel élément :

```
<persName>François Rabelais</persName>
```

Prenons un exemple plus complexe, celui de « maistre François de Connan, maistre des Requestes de l'hostel » :

```
<persName>maistre François de Connan, maistre des Requestes de  
l'hostel</persName>
```

L'élément `<persName>` utilise d'autres éléments imbriqués les uns dans les autres qui sont :

1. `<surName>` correspond au nom de famille (ou patronyme) ;
2. `<foreName>` correspond au prénom ;
3. `<genName>` pour *generational name* est employé pour désigner une hiérarchie avec l'attribut utile `@type` pour « l'aîné » par exemple ;
4. `<roleName>` comprend des attributs comme `@type` pour « religion » ou pour « titre » par exemple ;
5. `<addName>` pour *additional name* concerne les épithètes avec pour attribut `@type` pour « honorable femme » par exemple ;
6. `<nameLink>` encode les particules nobiliaires.

Reprenons le dernier exemple et complétons le balisage :

```
<persName>  
  <foreName>François</foreName>  
  <nameLink>de</nameLink>  
  <surName>Connan</surName> ,
```

```
<roleName>maistre des Requestes de l'hostel</roleName>
</persName>
```

Prenons deux autres exemples. Le nom de Pline le Jeune peut être encodé ainsi :

```
<persName>
  <foreName>Pline</foreName>
  <genName>le jeune</genName>
</persName>
```

ou encore saint Augustin :

```
<persName>
  <roleName type="religion"»saint</roleName>
  Augustin
</persName>
```

Parmi les attributs, notons les plus utilisés comme @type et @xml:id ou @ref que nous connaissons déjà, et l'attribut @ref pour *reference* qui peut indiquer un lien vers une notice, soit localement, soit dans un fichier externe grâce à un identifiant. Par exemple :

```
<persName>François Rabelais</persName>
```

peut devenir :

```
<persName ref= ".../personnes.xml#FR">François Rabelais</persName>
```

qui renvoie à la notice sur François Rabelais dans un fichier externalisé « personnes.xml ».

Le choix éditorial se pose ici : encoder le plus simplement possible les noms ou effectuer une première analyse et marquer d'une manière plus élaborée les noms de personnes¹⁸.

- Toponymes

Les toponymes sont encodés grâce aux éléments <placeName> et <geogName>. Le premier correspond à la désignation des villes alors que le second concerne tout type de toponymes. Par exemple, une ville sera encodée ainsi :

```
<placeName>Lion</placeName>
```

alors qu'un pays – même imaginaire – sera marqué ainsi :

```
<geogName>Angleterre</geogName>
```

¹⁸ Il est même possible de réaliser une prosopographie grâce à un encodage XML/TEI. Cf. [http://www.tei-c.org/chapitre 13](http://www.tei-c.org/chapitre_13) du manuel P5.

Il est bien entendu possible d'utiliser l'attribut `@type` pour dresser une typologie : pays, régions, rivières, etc. Deux attributs sont spécifiques : `@notBefore` et `@notAfter` dont la valeur est la date à partir de laquelle ou jusqu'à laquelle le nom de ce toponyme est utilisé.

Par ce biais, il est même possible de dresser des cartes grâce aux éléments `<location>` et `<geo>` imbriqués dans l'élément `<place>`. Les coordonnées, longitude et latitude, sont indiquées dans l'élément `<geo>`. À ce stade du projet, ces éléments n'ont pas été utilisés¹⁹.

(4) *La datation calendaire* : `<date>`

L'élément `<date>` contient une date dans n'importe quel format. Quatre attributs permettent d'en préciser la valeur.

- L'attribut `@when` donne la valeur de la date sous la forme standard suivante : `aaaa-mm-jj`. Par exemple :

```
<date when="1556-05-13">
```

- Au XVI^e siècle, il est fréquent que la date précise ne soit pas connue. Nous utilisons alors les attributs `@notBefore` et `@notAfter` :

```
<date notAfter="1500">1500</date>-<date notBefore="1543">1543</date>
```

- Un quatrième attribut, `@cert`, précise le degré de certitude de la date indiquée ; il peut avoir comme valeur `high`, `medium`, `low` ou `unknown` :

```
<date cert="low">1543</date>
```

L'attribut `@calendar` n'a pas été utilisé mais il serait utile pour signifier à quel système la date se rattache, notamment pour indiquer les dates antérieures à la réforme du calendrier : l'édit de 1564 fait commencer l'année au 1^{er} janvier (auparavant entre janvier et Pâques). Nous donnons un exemple imaginaire :

```
<date when="1531-02-02" calendar="julian">deux febvrier 1530</date>
```

Le choix éditorial est de simplifier le marquage du temps et de n'utiliser en général que l'élément `<date>` avec ses attributs `@when` ou `@cert`.

2. Images typographiques

Le choix éditorial a été de n'encoder que les illustrations liées à une banque d'images grâce à l'adresse « url ». Par exemple, la page de titre des *Discours non plus mélancoliques* contient une marque d'imprimeur :

¹⁹ Voir le chapitre 13 dans les Guidelines <http://www.tei-c.org/>

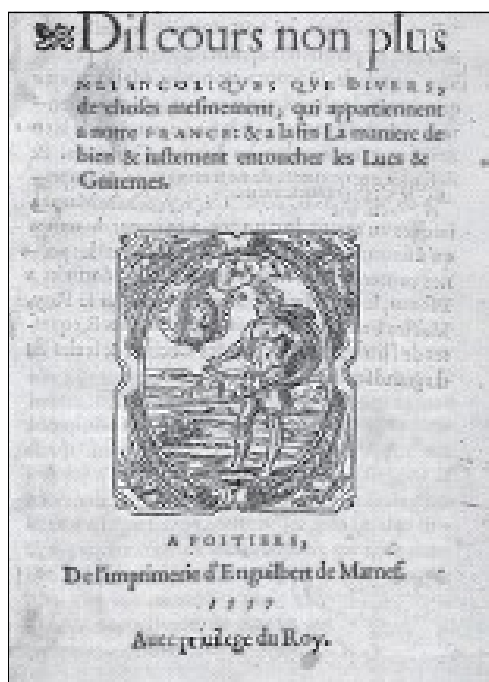


Illustration 22: Illustration de la page de titre des Discours

Pour encoder la marque d'imprimeur, l'élément `<graphic>` est utilisé avec l'attribut `@url` dont la valeur est l'adresse http ; il indique la localisation en ligne de l'image, à savoir :

```
<graphic url="http://www.bvh.univ-
tours.fr/B861946101_DP1139/img/DP1139_001_img_01.jpg"/>
```

L'affichage de l'image peut poser parfois des problèmes de proportion et demande un encodage spécifique qui n'est pas étudié ici.

L'encodage de la page de titre peut être résumé ainsi²⁰ :

```
<front>
  <titlePage>
    <docTitle>
      <titlePart>Discours...</titlePart>
    </docTitle>
    <docImprint>
      <graphic url="http://www.bvh.univ-
tours.fr/B861946101_DP1139/img/DP1139_001_img_01.jpg"/>
      ...
    </docImprint>
    <imprimatur>Avec privilege du Roy.</imprimatur>
```

²⁰ L'encodage complet de la page titre des *Discours* est étudié ultérieurement dans la quatrième partie du manuel.

```

</titlePage>
<div type="privilege">
  <p>...</p>
</div>
</front>

```

Grâce à une connexion au serveur, l'image est récupérée. À l'affichage, le résultat est proche de la page de titre originale :

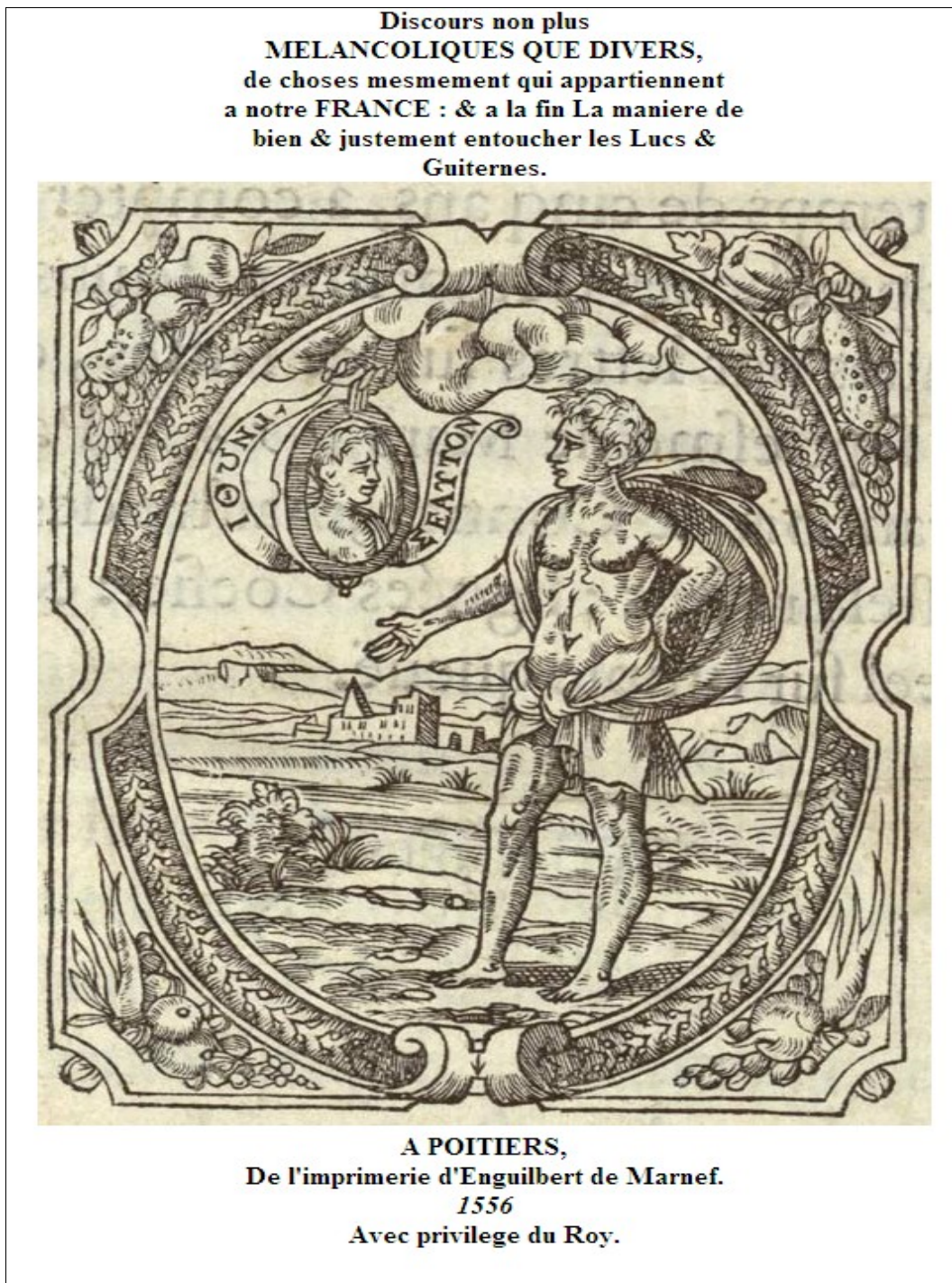


Illustration 23: Affichage de l'illustration de la page de titre des Discours

Dans le projet BVH, la typologie des données graphiques est établie à partir de l'extraction automatique. L'encodage TEI en dépend et est en cours de développement.

III En-tête TEI <teiHeader>

Tout texte de format TEI est composé d'un en-tête TEI suivi du corps du texte lui-même. L'en-tête contient des informations indispensables à l'identification du fichier :

1. <fileDesc> une description bibliographique du fichier,
2. <encodingDesc> le rapport entre le texte électronique et la source originale,
3. <profileDesc> une description des aspects non bibliographiques,
4. <revisionDesc> un historique des révisions du fichier électronique.

Cet en-tête doit prendre en compte l'aspect juridique lié à la propriété intellectuelle et au droit d'auteur. Il peut être réduit à quelques lignes : le titre électronique, l'éditeur électronique et une bibliographie sommaire de la source originale. Cependant, il est fortement recommandé de le compléter soit en apportant des informations sur la source originale, soit en donnant les modalités de l'édition électronique. Le plus grand nombre de renseignements est requis pour diffuser et échanger les fichiers électroniques.

1. Description bibliographique du fichier : <fileDesc> *file description*

La description bibliographique du fichier commence par le titre électronique ainsi que l'identité des responsables de son contenu intellectuel. La source, à partir de laquelle le fichier a été généré, fait ensuite l'objet d'une description bibliographique.

L'élément <titleStmt> pour *title statement* contient les éléments suivants :

- <title> désigne le titre électronique de l'ouvrage décidé par l'équipe éditoriale.
- <author> cite l'auteur de la source originale.
- <respStmt> pour *statement of responsibility* donne le nom du responsable de la collection électronique, ainsi que ceux de l'encodage.

L'élément <editionStmt> pour *element statement* contient les éléments éditoriaux suivants :

- <edition> avec son attribut *n* et sa valeur, donne le numéro de la présente édition électronique (souvent le numéro « 1 ») ; on y ajoute la date, l'url du fichier en ligne ainsi que son identifiant.
- <respStmt> pour *statement of responsibility* donne le nom du responsable de l'édition électronique et le domaine de sa responsabilité.

L'élément <extent> donne la taille matérielle du texte ; nous l'utilisons pour indiquer le nombre de kilooctets du fichier.

L'élément `<publicationStmt>` pour *publication statement* contient les informations sur l'organisme responsable de la publication.

- `<authority>` le nom des organismes responsables de l'œuvre électronique.
- `<address>` l'adresse des organismes.
- `<availability>` la restriction juridique de sa distribution : copyright.

L'élément `<seriesStmt>` pour *series Statement* n'a pas été encore utilisé dans le projet BVH. Il permet de regrouper les informations concernant une collection. Il pourra être ajouté s'il est envisagé d'encoder une collection : dictionnaires, œuvres complètes d'un auteur, minutes notariales d'une étude, etc.

L'élément `<sourceDesc>` pour *source description* est une partie obligatoire de l'en-tête car il consigne toutes les informations sur la source originale. Il contient les éléments :

- `<biblFull>` pour *fully-structured bibliographic citation* ; il reprend toutes les informations d'une source électronique du fichier TEI telle qu'une première édition électronique grâce aux éléments `<titleStmt>`, `<publicationStmt>`.
- `<biblStruct>` pour *structured bibliographic citation* ; cet élément comprend d'autres éléments sur les informations bibliographiques contenues principalement dans la page de titre.
- `<monogr>` pour *monographic level* ; imbriqué dans l'élément `<biblStruct>`, il contient les éléments bibliographiques comme `<title>`, `<author>`, `<imprint>`, `<publisher>`, `<date>`, `<pubPlace>` pour *publication place* qui reprennent les parties de la page de titre de la source originale.

Plusieurs problèmes se posent dans cette partie de l'en-tête de TEI. Le premier concerne l'auteur quand celui-ci n'est pas formellement désigné. Par exemple, les *Discours non plus mélancoliques...* est une œuvre attribuée à Elie Vinet ; nous proposons l'encodage suivant :

```
<monogr>
  <title>...</title>
  <respStmt>
    <resp>attribué à</resp>
    <name>Elie Vinet</name>
  </respStmt>
</monogr>
```

La même façon d'encoder pourra être utilisée quand un ouvrage fait l'objet de traductions : on précisera alors dans l'élément `<resp>`, « traduit par ».

- `<listBibl>` pour *citation list*, cet élément contient une liste de citations bibliographiques de la source.

- `<msDesc>` pour *manuscript description* ; dans la version de P5 de septembre 2007, cet élément a remplacé `<msDescription>` ; il concerne à l'origine la description des manuscrits mais on peut l'étendre aux imprimés anciens. L'élément `<origDate>` peut s'imbriquer et indiquer la date de la première version connue.
- `<msIdentifier>` pour *manuscript identifier* imbriqué dans l'élément `<msDesc>` ; il contient les informations pour retrouver la source : lieu de conservation, identifiant et ville. Dans le cas des *Discours*, l'encodage est le suivant :

```
<msIdentifier>
  <settlement>Poitiers</settlement>
  <repository n="861946101">Médiathèque François Mitter-
rand.</repository>
  <idno>DP 1139</idno>
</msIdentifier>
```

L'élément `<settlement>` est préféré à celui de `<placeName>` car il s'applique à tous les cas. Le numéro de la bibliothèque issu du répertoire des centres de ressource géré par l'ABES (Agence bibliographique de l'enseignement supérieur) est spécifié dans l'élément `<repository>`, lieu de conservation de la source. L'élément `<idno>` donne le numéro de catalogue de l'œuvre au sein de la collection de la bibliothèque.

- `<physDesc>` pour *physical description*, imbriqué dans l'élément `<msDesc>` contient la description physique de la source grâce à deux éléments principaux : `<objDesc>` et `<bindingDesc>`.

Voici un exemple type de la description du fichier électronique :

```
<fileDesc>
  <titleStmt>
    <title>Nouvelles récréations et joyeux devis de Bonaventure des
Périers, Lyon, 1561: édition électronique format XML/TEI P5.</title>
    <author>
      Bonaventure Des Périers
      <date notAfter="1500" type="birthday">1500</date>--<date when="1543"
type="death">1543</date>
    </author>
    <respStmt>
      <name xml:id="MLD">Marie-Luce Demonet</name>
      <resp><date>2008</date>Responsable BVH CESR</resp>
    </respStmt>
    <respStmt>
      <name></name>
      <resp><date>2008</date>Transcription</resp>
    </respStmt>
    <respStmt>
      <name xml:id="TV">Thierry Vincent</name>
      <resp>Dissimilation</resp>
    </respStmt>
    <respStmt>
      <name xml:id="VLR">Vanessa Le Rolle</name>
      <resp><date>2009</date>encodage</resp>
    </respStmt>
  </titleStmt>
```

```

    <respStmt>
      <name xml:id="ND">Nicole Dufournaud</name>
      <resp><date>2009</date>Contrôle de l'encodage</resp>
    </respStmt>
  </titleStmt>
  <editionStmt>
    <edition n="1"><date>2009</date></edition>
    <respStmt>
      <name xml:id="TU">Toshinori Uetani</name>
      <resp>Révision</resp>
    </respStmt>
  </editionStmt>
  <extent>853 ko</extent>
  <publicationStmt>
    <authority>Université François-Rabelais</authority>
    <address>
      <addrLine>3 rue des Tanneurs</addrLine>
      <addrLine>37041 TOURS Cedex 1 France</addrLine>
    </address>
    <authority>Centre National de la Recherche Scientifique (CNRS)</authority>
    <address>
      <addrLine>3 rue Michel-Ange</addrLine>
      <addrLine>75794 Paris France</addrLine>
    </address>
    <authority xml:base="http://www.cesr.univ-tours.fr">Centre d'Etudes
    Supérieures de la Renaissance (CESR)</authority>
    <address>
      <addrLine>59 rue Néricault-Destouches</addrLine>
      <addrLine>BP 11328</addrLine>
      <addrLine>37013 Tours Cedex</addrLine>
    </address>
    <availability>
      <p>Utilisation non commerciale libre. Copyright. Université
      François Rabelais Creative Commons
      <date when="2008">2008</date>
      </p>
    </availability>
  </publicationStmt>
  <sourceDesc>
    <biblStruct>
      <monogr>
        <author>Des Périers, Bonaventure</author>
        <title>Nouvelles recreations et joyeux devis de feu
        Bonaventure des Periers, valet de Chambre de la Royné de
        Navarre.</title>
        <imprint>
          <pubPlace>Lyon</pubPlace>
          <publisher>
            <roleName>libraire</roleName>
            <name>Guillaume Rouillé</name>
          </publisher>
          <date>1561</date>
        </imprint>
      </monogr>
    </biblStruct>
    <listBibl>
      <bibl>
        <author>Henri Baudrier</author>
        <title>Bibliographie lyonnaise.</title>
        <pubPlace>Lyon</pubPlace>

```

```

        <date></date>, t. IX, p. 282.
    </bibl>
    <bibl>
        <author>Sybille von Gültlingen</author>
        <title>Bibliographie des livres imprimés à Lyon au
        seizième siècle.</title>
        <pubPlace>Baden-Baden</pubPlace>
        <date when="2006">2006</date>, t. X, p. 134 (Guillaume
        Rouillé, n° 508.
    </bibl>
</listBibl>
<msDesc>
    <msIdentifier
xml:base="http://gallica2.bnf.fr/ark:/12148/bpt6k791287.r=les+nouvelles+recrea-
tions+.langFR#">
        <country>France</country>
        <settlement>Lyon</settlement>
        <institution>Bibliothèque municipale de Lyon</institution>
        <repository n="B693836101"></repository>
        <collection></collection>
        <idno type="ouvrage">Rés A 489170</idno>
        <altIdentifier
xml:base="http://iris.lib.virginia.edu/speccol/gordon/gordonimages/Gordon1561_D4
7/">
            <country>USA</country>
            <settlement>Charlottesville</settlement>
            <institution>University of Virginia</institution>
            <collection>Douglas H. Gordon Collection</collection>
            <idno type="ouvrage">Gordon1561_D47</idno>
            </altIdentifier>
        </msIdentifier>
        <physDesc>
            <objectDesc>
                <supportDesc>
                    <extent>in-4</extent>
                    <collation>239 p.</collation>
                </supportDesc>
            </objectDesc>
        </physDesc>
    </msDesc>
</sourceDesc>
</fileDesc>

```

Dans la version de septembre 2007 de TEI, plusieurs éléments nouveaux imbriqués dans `<msDesc>` permettent d'affiner la description de la source comme `<msContents>` dans le cas d'une collection ; `<history>` permet de noter les origines, la provenance et l'acquisition de l'œuvre.

2. Description de l'encodage : `<encodingDesc>` *encoding description*

La description relative au codage précise les méthodes et les principes éditoriaux qui ont permis la transcription du texte et l'élaboration du fichier.

- L'élément `<projectDesc>` pour *project description* comprend la description succincte du projet électronique.

- L'élément `<editorialDecl>` pour *editorial declaration* comprend les pratiques éditoriales de l'encodage dont celles du présent manuel :
 - `<correction>` Les erreurs détectées dans la source sont listées.
 - `<hyphenation>` Les consignes au sujet des coupures de mots en fin de ligne sont écrites.
 - `<normalization>` Normes de transcription pour l'édition électronique. Les descriptions des symboles graphiques sont définies ici.

Voici un exemple type tiré des *Discours* :

```
<encodingDesc>
  <projectDesc>
    <p>BVH : programme "bibliothèques virtuelles humanistes"</p>
  </projectDesc>
  <editorialDecl>
    <correction>
      <p><title>LISTE DES VARIANTES ET DES CORRECTIONS</title></p>
      <p>...</p>
    </correction>
    <hyphenation>
      <p>La division des mots en fin de ligne respecte l'original.
      </p>
    </hyphenation>
    <normalization>
      <p>Cette transcription suit l'édition de 1556.</p>
    </normalization>
  </editorialDecl>
</encodingDesc>
```

3. Description non bibliographique : `<profileDesc>` *text-profile description*

Cette partie concerne les informations non bibliographiques décrivant un document.

L'élément `<langUsage>` spécifie la ou les langues utilisées. S'y associe l'élément `<language>` : l'identification est indiquée par l'attribut `@ident` auquel est associé le code de la langue. La nomenclature des codes standards correspond à la norme ISO 639²¹.

De nombreuses langues figurent dans les *Discours* et doivent être déclarées dans l'en-tête TEI avant de les encoder dans le texte, ainsi :

```
<profileDesc>
  <langUsage>
    <language ident="frm">Français de la
      Renaissance</language>
    <language ident="la">Latin</language>
    <language ident="el">Grec</language>
```

21 Cf. <http://xml.coverpages.org/>

```

        <language ident="de">Allemand du XVIIe
            siècle</language>
        ...
    </langUsage>
</profileDesc>

```

Les informations sur la création du texte peuvent être données ici grâce à l'élément `<creation>` à la place de `<history>` dans `<fileDesc>`. D'autres éléments comme `<textClass>` pour *text classification* ou `<handList>` peuvent être incorporés dans la liste des éléments susceptibles d'être utilisés pour établir une taxonomie ou une liste des différentes écritures ou caractères d'imprimerie.

4. Journal de modifications : `<revisionDesc>` *revision description*

À chaque révision du fichier, une description des modifications apportées est marquée grâce à l'élément `<revisionDesc>`. Elle alimente un journal des modifications.

- L'élément `<change>` enregistre les séquences. On peut y ajouter l'élément `<date>` avec son attribut `@when` dont le standard est « année - mois - jour » :

```
<change><date when="2007-05-31">31 mai 2007</date></change>
```

- Un second élément `<name>`, imbriqué dans `<change>`, indique le nom de l'encodeur et de la personne qui contrôle l'encodage. On peut ajouter le motif du changement en texte libre.

```

<revisionDesc>
  <change>
    <date when="2007-05-18">18 mai 2007</date>
    <name xml:id="ND">Nicole Dufournaud</name>
    Création du fichier en TEI P5 pour encodage
  </change>
</revisionDesc>

```

IV Texte de l'œuvre

L'élément `<text>` englobe l'ensemble du texte. Il peut être unitaire ou composite comme un recueil. Le texte dans son ensemble est composé au maximum de quatre parties :

1. `<front>` regroupe les pièces situées avant le début du texte lui-même comme la page de titre de l'ouvrage
2. `<group>` regroupe plusieurs textes
3. `<body>` regroupe le corps entier d'un texte unitaire
4. `<back>` regroupe les annexes.

Au début du manuel, nous avons exposé une structure simple, à laquelle nous ajoutons l'élément `<back>` :

```
<text>
  <front>
    ...
  </front>
  <body>
    ...
  </body>
  <back>
    ...
  </back>
</text>
```

Cette structure simple ne convient pas à toutes les œuvres du fait de leur complexité, comme celle des *Euvres* de Louise Labé.

1. Structure complexe : les *Euvres* de Louise Labé

Comme son titre l'indique, dans l'exemple des *Euvres* de Louise Labé, ce sont plusieurs œuvres assemblées, et même d'auteurs différents, ce qui implique l'utilisation de l'élément `<group>`. L'analyse de cette structure particulière donne le résultat suivant :

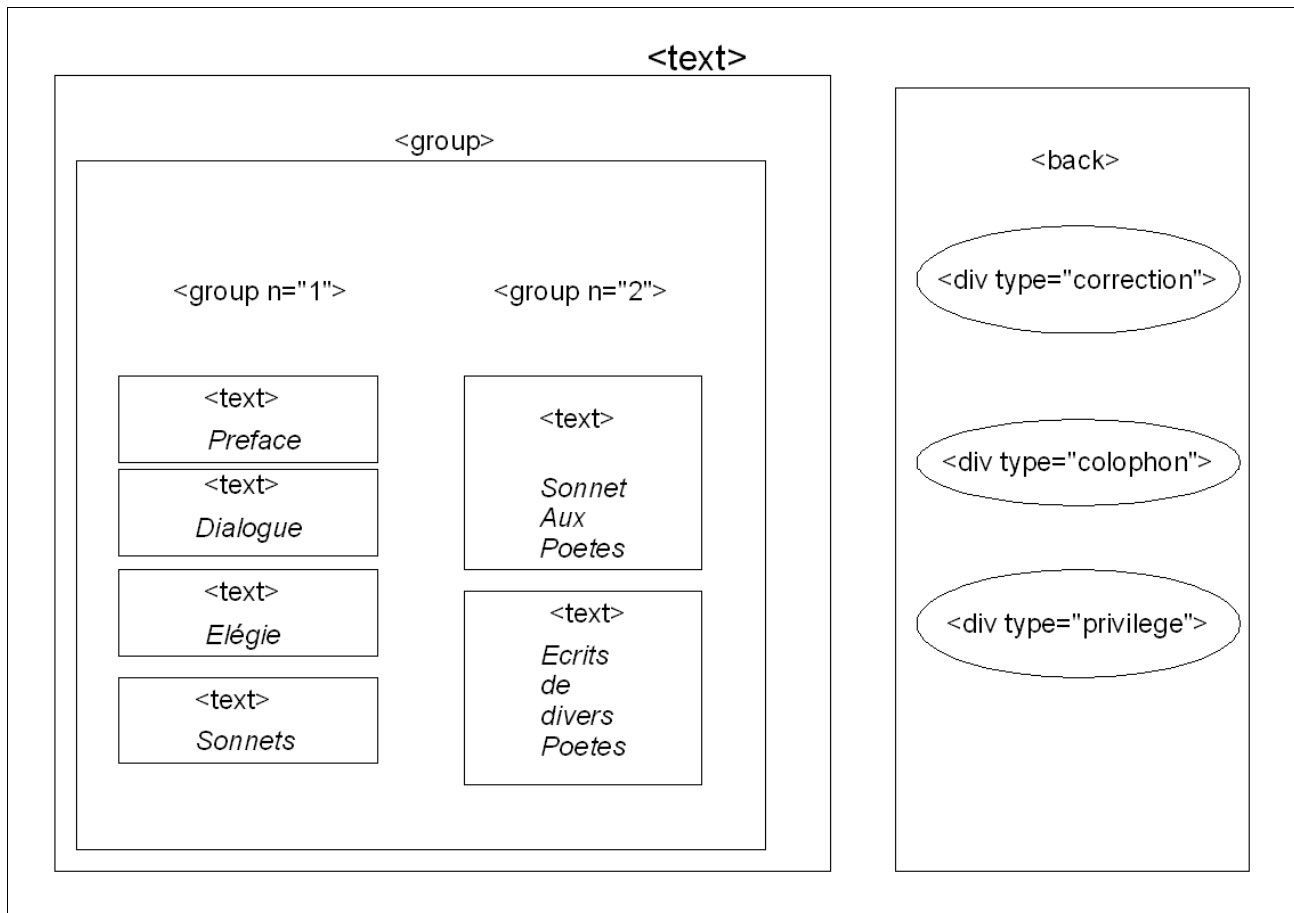


Illustration 24: Structure des *Euvres de Louise Labé*

La structure montre que le texte est coupé en deux groupes principaux : les *Euvres* de Louise Labé et les « hommages » qui lui ont été rendus par d'autres auteurs. Le tout forme un seul texte avec deux ensembles distincts rassemblés dans un même groupe. Le second ensemble regroupe plusieurs hommages d'auteurs différents : nous avons dissocié le premier sonnet s'adressant aux divers poètes de leurs propres hommages.

```
<text>
  <front>[titre des Oeuvres]</front>
  <group>
    <group n="1">
      <text n="1" type="preface">
        <front>[titre de la Préface]</front>
        <body>
          ... [corps du texte]
        </body>
      </text>
      <text n="2" type="dialogue">...</text>
      <text n="3" type="elegie">...</text>
      <text n="4" type="sonnet">...</text>
      <trailer>FIN DES EUVRES...</trailer>
    </group>
    <group n="2">
      <text type="hommage">
        <body>
```

```

        <div n="1" type="sonnet">...</div>
    </body>
</text>
<text type="hommage">
    <body>
        <div n="1" type="ode">...</div>
        ...
        <div n="24" type="ode">...</div>
    </body>
</text>
</group>
<trailer>FIN</trailer>
</group>
<back>
    <div type="correction">...</div>
    <div type="colophon">...</div>
    <div type="privilege">...</div>
</back>
</text>

```

Nous utilisons l'élément `<trailer>` pour signifier la fin de chaque groupe. Après le second groupe, l'élément `<back>` clôture l'ensemble du texte avec des divisions qui reprennent les « corrections », le « colophon » et le « privilège ».

2. Page de titre et pièces liminaires

L'élément `<front>` contient tous les éléments qui précèdent le texte : la page de titre et les pièces liminaires d'un texte. Un exemple complexe nous est donné par la page de titre des *Discours non plus mélancoliques* qui est encodée ainsi :

```

<front>
    <titlePage>
        <docTitle>
            <titlePart>
                <lb/>Discours non plus
                <lb/><hi rend="sc">MELANCOLIQUES QUE DIVERS,</hi>
                <lb/>de choses mesmement qui appartiennent
                <lb/>a notre FRANCE : &amp; a la fin La maniere de
                <lb/>bien &amp; justement entoucher les Lucs &amp;
                <lb/>Guiternes.
            </titlePart>
        </docTitle>
        <docImprint>
            <graphic url="http://www.bvh.univ-
            tours.fr/B861946101_DP1139/img/DP1139_001_img_01.jpg"/>
            <lb/>A <placeName>POITIERS</placeName>,
            <lb/>De l'imprimerie d'<persName>Enguilbert de Marnef</persName>.
            <lb/><docDate><hi rend="italic">1556</hi></docDate>
        </docImprint>
        <imprimatur>
            <lb/>Avec privilege du Roy.
        </imprimatur>
    </titlePage>
    <pb n="a i v°"
    <div type="privilege">
        <p>...</p>
    </div>

```



```
</div>
</front>
```

Nous y distinguons les éléments suivants :

- `<titlePage>` reprend tous les éléments de la page de titre de l'œuvre,
- `<docTitle>` reprend le titre général de l'œuvre,
- `<titlePart>` contient un fragment du titre de l'œuvre,
- `<docImprint>` donne les éléments concernant l'édition et l'impression de l'œuvre,
- `<docDate>` est l'élément de la date de l'œuvre,
- `<imprimatur>` avise le lecteur de l'autorisation de publication de l'œuvre ; ici, il s'agit d'un privilège royal.

Le privilège du roi est une pièce liminaire de l'œuvre ; il est encodé grâce à l'élément `<div>` dans lequel est ajouté le type « privilege ».

La partie graphique a été étudiée précédemment avec les images typographiques.

Les pages de titre intermédiaires ne nécessitent pas l'élément `<titlePage>` qui caractérise la page de titre du livre, mais sont marquées uniquement avec l'élément `<docTitle>` puis `<titlePart>`. Les *Euvres* de Louise Labé en offrent un exemple avec la préface :

```
<text>
  <front>
    <titlePage>
      ...
    </titlePage>
  </front>
  <group n="1">
    <text type="preface">
      <front>
        <docTitle>
          <titlePart>...</titlePart>
        </docTitle>
      </front>
      ...
    </text>
  </group>
</text>
```

Deux chapitres et deux annexes sont en cours d'élaboration :

- un chapitre sur les liens et les pointeurs
- un chapitre sur les images
- une annexe sur les outils
- une annexe avec une feuille de style

Remarque importante

Pour les textes en langue française (unicode UTF-8), le fichier doit commencer par la ligne de code suivante :

```
<?xml version="1.0" encoding="UTF-8"?>
```

Elle permet de résoudre les problèmes posés par les caractères accentués et le « ç ». Dans les balises, que ce soient les éléments ou les attributs, il est fortement recommandé de ne jamais mettre ni d'accent ni de caractère français ; il faut renoncer aux caractères suivants : « é », « è », « ç », « à » et « ù ». Ces recommandations valent également pour les noms de fichiers. En revanche, dans le texte, ces mêmes caractères peuvent être utilisés grâce au standard Unicode²². Nous soulignons des exceptions. Dans les fichiers XML, trois caractères ne peuvent pas être utilisés tels quels : « < », « > », « & » ; deux autres caractères ne peuvent pas être utilisés dans les attributs : « ' » et les guillemets. Les raisons sont évidentes pour les deux premiers : ce sont les chevrons d'ouverture et fermeture de balises XML ; pour le troisième, la raison est identique car le « & » est utilisé pour définir les entités ; les guillemets sont également utilisés ainsi que l'apostrophe. Exemples :

- pour & : saisir `&` ; dans le texte ;
- pour ' : saisir `'` ; dans l'attribut de la balise ;
- pour " : saisir `"e;` ; dans l'attribut de la balise.

Ces développements sont invisibles dans l'affichage en HTML.

Pour les textes de la Renaissance, le caractère « & » pose particulièrement problème car il est utilisé de façon très régulière par les scripteurs de l'époque.

Si l'on part d'un texte déjà transcrit à l'aide d'un traitement de texte, il est recommandé de le convertir d'abord au format « texte unicode » (ou d'utiliser OpenOffice) pour éviter les problèmes des caractères accentués et spéciaux.

²² Les ordinateurs codent les lettres et autres caractères sous formes de nombres. Il existe plusieurs systèmes de codage pour couvrir toutes les langues en usage. Unicode est un standard qui attribue spécifiquement un numéro unique à chaque caractère quel que soit le logiciel et quelle que soit la langue. Cf. <http://www.unicode.org>

Annexe 1 : Liste des livres encodés (juin 2009)

1. Béroalde de Verville François, *Le Moyen de Parvenir*, ca. 1616
2. Béroalde de Verville François, *Le Voyage des Princes Fortunez*, Paris, Claude de la Tour, 1610
3. Calenzio Eliseo, *Les Grandes et fantastiques batailles des grans Roys Rodilardus, & Croacus*, Blois, Julian Angelier, 1554
4. Castiglione Baldassare, traduit par. Jacques Colin, *Le Courtisan*, Lyon, François Juste, 1538
5. Colonna Francesco, traduit par François Béroalde de Verville. *Le Tableau des riches inventions*, Paris, Mathieu Guillemot, 1600
6. Corrozet Gilles, *Le Tableau de Cebes de Thebes*, Paris, 1543
7. Des Périers Bonaventure, *Nouvelles recreations et joyeux devis*, Lyon, Guillaume Rouillé, 1561
8. Du Bellay Joachim, *L'Olive*, Paris, Arnoul L'Angelier, 1550
9. Du Bellay Joachim, *La Deffence et illustration de la langue francoyse*, Paris, Arnoul L'Angelier, 1549
10. Du Bellay Joachim, *L'Olive*, Paris, L'Angelier, 1549
11. Garzoni Tomazo, traduit par Gabriel Chappuys, *Le Theatre des divers cerveaux du monde*, Paris, Félix Le Mangnier, 1586
12. Gelli Jean-Baptiste, *Les Discours fantastiques de Justin tonnelier*, Lyon, La Salemandre, 1566
13. Horace, *Art poetique*, traduit par Jacques Peletier du Mans, Paris, 1545
14. Labé Louise, *Euvres*, Lyon, Jean de Tournes, 1555
15. Lery Jean de, *Histoire d'un voyage fait en la terre de Bresil*, La Rochelle, 1578
16. Lucien de Samosate, *La Mouche*, traduit par Geofroy Tory de Bourges, Paris, Geofroy Tory de Bourges, s.d., 1533
17. Marot Clément, *Les Blasons anatomiques du corps femenin*, Paris, Charles Langelier, 1543
18. Marot Clément, *L'Adolescence clémentine*, Lyon, Etienne Dolet, 1538
19. Rabelais François, *Briefve declaration*, Paris, 1552
20. Rabelais François, *Cinquiesme Livre*, s.l.n.d., 1564

21. Rabelais François, *Gargantua*, Paris, 1542
22. Rabelais François, *Pantagruel*, Lyon, François Juste, 1542
23. Rabelais François, *Quart Livre*, Paris, 1552
24. Rabelais François, *Tiers Livre*, Paris, Chrétien Wechel, 1546
25. Rabelais François, *Tiers Livre*, Paris, 1552
26. Tory Geoffroy, *Champ fleury*. Paris, Geoffroy Tory et Giles Gourmont, 1529
27. Vinet Elie (attribué à), *Discours non plus melancoliques que divers*, Poitiers, Enguilbert de Marnef, 1556

Annexe 2 : Liste des langues et dialectes utilisés

Langues ou dialectes	Code ISO 639
Langues artificielles et latin macaronique	art
Allemand	de
Arabe	ar
Breton	bre
Gascon	oci-gascon
Grec	el
Hébreu	he
Italien	it
Latin	la
Lorrain	fr-lorrain
Picard	fr-pcd
Poitevin	fr-poit

Annexe 3 : liste des éléments et des attributs

Attribut.....	
ana=" ".....	39
calendar=" ".....	40
cert=" ".....	40
cols=" ".....	23, 24
hand=" ".....	30, 31, 34
ident=" ".....	48, 49
n=" ".....	11, 16, 18, 19, 20, 45, 46, 51, 52, 53
notAfter=" ".....	40, 45
notBefore=" ".....	40
place=" ".....	20, 21, 22, 31, 34
reason=" ".....	29, 30
ref=" ".....	39
rend=" ".....	17, 21, 25, 36, 37, 52
resp=" ".....	11, 30
role=" ".....	24, 25
rows=" ".....	23, 24
source=" ".....	30, 32
target=" ".....	28
type=" ".....	16, 17, 20, 21, 22, 26, 28, 37, 38, 39, 40, 42, 45, 47, 51, 52, 53
url=" ".....	41
when=" ".....	11, 36, 40, 45, 46, 49

who=" "	18, 37
xml:base=" "	30, 46, 47
xml:id=" "	30, 39, 45, 46, 49
xml:lang=" "	37

Elément.....	
<abbr>.....	28, 29, 34
<add>.....	29, 31, 34
<addName>.....	38
<address>.....	11, 44, 46
<addrLine>.....	11, 46
<altIdentifier>.....	47
<author>.....	30, 36, 43, 44, 45, 46, 47
<authority>.....	13, 44, 46
<availability>.....	44, 46
<back>.....	12, 14, 15, 50, 52
<bibl>.....	11, 30, 36, 46, 47
<biblFull>.....	44
<biblStruct>.....	44, 46
<bindingDesc>.....	45
<body>.....	12, 14, 15, 16, 17, 50, 51
<cell>.....	23, 24, 25, 26, 27
<change>.....	49
<choice>.....	33, 34
<cit>.....	35, 36, 37
<collation>.....	47

<collection>.....	47
<corr>.....	32, 34
<correction>.....	48
<country>.....	47
<creation>.....	49
<date>.....	11, 30, 36, 40, 44, 45, 46, 47, 49
.....	29, 30, 34
<desc>.....	30
<div>.....	12, 16, 17, 18, 26, 28, 42, 52, 53
<docAuthor>.....	33
<docDate>.....	53
<docImprint>.....	15, 41, 52, 53
<docTitle>.....	15, 33, 41, 52, 53
<edition>.....	43, 46
<editionStmt>.....	43, 46
<editor>.....	30
<editorialDecl>.....	48
<encodingDesc>.....	13, 43, 48
<expansion>.....	29, 34
<extent>.....	43, 46, 47
<fileDesc>.....	13, 14, 43, 45, 49
<foreign>.....	37
<foreName>.....	38, 39
<front>.....	12, 14, 15, 17, 33, 41, 50, 51, 52, 53
<fw>.....	20, 21

<gap>.....	29, 30
<genName>.....	38, 39
<geo>.....	40
<geogName>.....	39
<graphic>.....	41
<group>.....	14, 50, 51, 53
<handList>.....	49
<head>.....	16, 17, 18, 22, 23, 26
<hi>.....	21, 36, 37, 52
<history>.....	47, 49
<hyphenation>.....	48
<idno>.....	45, 47
<imprimatur>.....	41, 52, 53
<imprint>.....	44, 46
<institution>.....	47
<item>.....	22, 28
<l>.....	18, 22, 35, 36
<label>.....	23
<language>.....	48, 49
<langUsage>.....	48
<lb>.....	19, 22, 24, 25, 26, 27, 28, 29, 31, 35, 36, 37, 52
<lg>.....	18, 22
<list>.....	22, 28
<listBibl>.....	44, 46
<location>.....	40

<monogr>.....	44, 46
<msContents>.....	47
<msDesc>.....	45, 47
<msDescription>.....	45
<msIdentifier>.....	45, 47
<name>.....	9, 38, 44, 45, 46, 49
<nameLink>.....	38
<normalization>.....	30, 48
<note>.....	21, 22
<objDesc>.....	45
<objectDesc>.....	47
<orig>.....	31, 32, 33
<origDate>.....	45
<p>.....	9, 10, 13, 14, 17, 18, 19, 22, 29, 30, 31, 32, 34, 37, 42, 46, 48, 52
<pb>.....	10, 20, 52
<persName>.....	38, 39, 52
<physDesc>.....	45, 47
<place>.....	40
<placeName>.....	39, 45, 52
<profileDesc>.....	13, 43, 48
<projectDesc>.....	47, 48
<publicationStmt>.....	13, 14, 44, 46
<publisher>.....	44, 46
<pubPlace>.....	30, 44, 46, 47
<quote>.....	35, 36, 37

<ref>.....	28
<reg>.....	31, 32, 33
<repository>.....	47
<repository>.....	45
<resp>.....	44, 45, 46
<respStmt>.....	43, 44, 45, 46
<revisionDesc>.....	13, 43, 49
<roleName>.....	38, 39, 46
<row>.....	23, 24, 25, 26, 27
<said>.....	37
<seriesStmt>.....	44
<settlement>.....	45, 47
<sic>.....	31, 32, 34
<sourceDesc>.....	13, 14, 44, 46
<sp>.....	17, 18
<space>.....	17
<speaker>.....	17, 18
<stage>.....	18
<supplied>.....	29, 30
<supportDesc>.....	47
<surName>.....	38
<table>.....	22, 23, 24, 26
<teiHeader>.....	12, 13, 14, 30, 43
<text>.....	12, 14, 15, 17, 30, 37, 50, 51, 52, 53
<textClass>.....	49

<title>.....	9, 10, 13, 14, 30, 31, 36, 43, 44, 45, 46, 47, 48
<titlePage>.....	15, 33, 41, 52, 53
<titlePart>.....	15, 33, 41, 52, 53
<titleStmt>.....	13, 14, 43, 44, 45
<trailer>.....	18, 51, 52
<unclear>.....	29, 30

Index des illustrations

Illustration 1: Chapitre 23 du Tiers Livre de Rabelais, édition 1546.....	6
Illustration 2: Transcription du chapitre 23 du Tiers Livre.....	7
Illustration 3: Fins de ligne dans le Tiers Livre de Rabelais 1546, prologue (p. 2)	19
Illustration 4: Titre courant des Euvres de Louise Labé (p. 19).....	20
Illustration 5: Pied de page des Euvres de Louise Labé (p. 3).....	21
Illustration 6: Appel de note dans les Euvres de Louise Labé (p. 154).....	22
Illustration 7: Tableau simple à deux colonnes tiré des Discours (p. 1).....	23
Illustration 8: Affichage d'un tableau simple à deux colonnes tiré des Discours (p. 1).....	24
Illustration 9: Dépliant des Discours (p. 9).....	24
Illustration 10: Affichage du dépliant des Discours (p. 9).....	25
Illustration 11: Table des matières des Discours.....	26
Illustration 12: Affichage d'un extrait de la table des matières des Discours.....	27
Illustration 13: Extrait de la table des matières du Tiers Livre de Rabelais 1552.....	28
Illustration 14: Euvres de Louise Labé (p. 11).....	29
Illustration 15: Rature dans les Discours (p. 66).....	29
Illustration 16: Ajout dans le privilège du roi des Discours.....	31
Illustration 17: Extrait de la page de titre des Euvres de Louise Labé.....	32
Illustration 18: Ajout et rature dans le Tiers Livre de Rabelais.....	34
Illustration 19: Expression tirée du Tiers Livre de Rabelais (p. 83 verso).....	35
Illustration 20: Référence bibliographique dans les Discours (p. 8).....	36
Illustration 21: Citation avec une référence bibliographique (Tiers Livre de 1552 p. 38).....	36
Illustration 22: Illustration de la page de titre des Discours.....	41

Illustration 23: Affichage de l'illustration de la page de titre des Discours.....	42
Illustration 24: Structure des Euvres de Louise Labé.....	51