

# 1. Guide d'encodage de l'édition des *Seint Confessor* de Wauchier de Denain d'après le manuscrit 412 de la BNF

Ce petit guide a pour but d'expliquer de quelle manière a été encodée en XML-TEI l'édition. Des principes stricts d'encodage ont été déterminés pour produire des données pérennes, interopérables et pour faciliter leur transformation. L'intégralité de l'encodage proposé est donc *TEI conformant*. L'édition des *Seint Confessor* réunit les textes suivants :

- *La Vie de saint Martin*
- *Les Dialogues sur les Vertus de saint Martin*
- *La Vie de saint Brice*
- *La Vie de saint Gilles*
- *La Vie de saint Marcel de Limoges*
- *La Vie de saint Nicolas*
- *La Vie de saint Jérôme*
- *La Vie de saint Benoît*
- *La Vie de saint Alexis*

Le corpus numérique est constitué de neuf fichiers XML TEI distincts correspondant chacun à un texte du recueil. Ces derniers sont réunis dans un `teiCorpus` nommé *CorpusLiSeintConfessor.xml*

## 1.1. Structurer le fichier XML

La structuration des textes contenus dans les fichiers XML suit les recommandations du service CTS, *Canonical Text System* (développé pour le Homer Multitext Project grâce à Christopher Blackwell et Neel Smith). Ce système permet de mettre en place un ensemble de services basés sur l'identification des textes ou de leurs fragments grâce à une référence canonique qui prend la forme d'une URN CTS [Note: Uniform Resource Name, nom d'un standard informatique qui permet d'identifier une ressource indépendamment de sa localisation et de son accessibilité par internet, ce qui permet à cet identifiant d'être pérenne]. Ainsi, cette édition est aisément citable et archivable grâce au respect de normes internationales.

Dans le recueil de Wauchier de Denain, chacune des Vies est encodée dans un fichier XML différent évitant ainsi de faire des fichiers trop longs et permettant de bien séparer chacun des travaux tout en respectant la norme CTS. Les fichiers de l'édition sont nommés et identifiés de la manière suivante : « urn:cts:froLit:jns915:jns1856.ciham-fro1 ». La première partie de l'URN : « urn:cts » indique à quel système de référence appartient l'URN ; ici, la norme CTS. « froLit » signifie que le texte appartient à un corpus des textes en ancien français. Les éléments précédés de « jns » indiquent les identifiants de l'œuvre dans le portail Jonas développé par l'IRHT qui rassemble des répertoires de textes et manuscrits médiévaux en langue d'oc et d'oïl, « 915 » est l'identifiant de l'auteur Wauchier de Denain sur le portail, tandis que 1856 désigne l'œuvre, dans ce cas, la *Vie de saint Martin de Tours*. Enfin la dernière partie de l'URN désigne l'instance éditrice : « ciham-fro1 », dans ce cas nous avons mis le nom de notre laboratoire de rattachement : CIHAM, « fro1 » faisant référence à un répertoire de textes en ancien français n°1, dans le cas où le laboratoire du CIHAM adopterait cette norme pour d'autres projets.

### 1.1.1. Structurer le `teiHeader`

Le `<teiHeader>` comporte trois grandes sections :

#### 1.1.1.1. Le `fileDesc`

Le `<fileDesc>` comporte lui-même : Le `<sourceDesc>` contient toutes les informations nécessaires sur le manuscrit de base, C1 [Note: Le sigle correspond au manuscrit 412 de la Bibliothèque Nationale de France]. L'ensemble des données sont contenues dans un `<msDesc>`

```
<msIdentifier xml:id="ms412">
  <settlement>Paris</settlement>
  <repository>Bibliothèque nationale de France</repository>
  <idno>fr. 00412</idno>
</msIdentifier> [...]
<msItemStruct>
  <locus>fol.103-fol.170</locus>
  <author>Wauchier de Denain</author>
  <title>Oeuvres hagiographiques 'Li Seint Confessor' de Wauchier
    de Denain</title>
  <incipit>De saint Martin mout doit on doucement et volentiers le
    bien oïr etentendre</incipit>
  <explicit/>
</msItemStruct>
```

Pour permettre l'établissement de l'apparat critique, l'ensemble des témoins de la tradition est déclaré dans le `<sourceDesc>` dans un élément `<listWit>`. Chaque témoin est contenu dans un élément `<witness>` auquel est attribué un identifiant pour identifier dans l'apparat le manuscrit auquel se réfèrent les différentes leçons. Cet identifiant est ajouté grâce à l'attribut `@xml:id`.

```
<listWit>
<witness xml:id="C1" corresp="#ms412">Paris, manuscrit fr.
  412, Bibliothèque nationale de France, <date when="1285">1285</date>
</witness>
<witness xml:id="C2">Paris, manuscrit fr. 411, Bibliothèque
  nationale de France, <date notBefore="1300" notAfter="1400">14e siècle</date>
</witness>
<witness xml:id="C3">London, manuscrit, Royal 20.D.VI,
  British Library, <date notBefore="1250" notAfter="1300">13e siècle</date>
</witness>
<witness xml:id="G1">Bruxelles, manuscrit 9225 bibliothèque
  royale, <date notBefore="1300" notAfter="1400">14e
  siècle</date>
</witness>
<witness xml:id="M1">Paris, manuscrit fr. 23112,
  Bibliothèque nationale de France, <date notBefore="1200" notAfter="1300">13e siècle</date>
</witness>
</listWit>
```

### 1.1.1.2. L'encodingDesc

L'<encodingDesc> permet de donner des informations sur l'encodage du fichier XML. Dans notre cas, il a permis de déclarer la méthode choisie pour constituer l'apparat critique.

```
<variantEncoding method="parallel-segmentation"
  location="internal"/>
```

L'utilisation de la norme CTS pour structurer le corpus y est également déclarée afin que les fichiers XML puissent être visualisés grâce à l'utilisation d'outils de visualisation ou de gestion de corpus comme les outils Capitains du projet *Perseus* [Note: Thibault Clérice, Bridget Almas et Stella Dee, « Capitains A tool suite for the CTS Norm »], notamment *Capitains Ahab* pour générer la base des données des textes et *Capitains Nemo* [Note: Thibault Clérice, « Capitains Poster at Grenoble's conference Humanites Numeriques#: L'exemple de l'antiquité »] qui permet de générer une interface utilisateur à partir d'un fichier XML utilisant les normes CTS.

```
<refsDecl n="CTS">
  <cRefPattern n="section"
    matchPattern="(.)+(.)"
    replacementPattern="#xpath(/tei:TEI/tei:text/tei:body/tei:div[@n='$1']/tei:div[@n='$2'])">
    <p>This pointer pattern extracts chapter and section</p>
  </cRefPattern>
  <cRefPattern n="chapter"
    matchPattern="(.)"
    replacementPattern="#xpath(/tei:TEI/tei:text/tei:body/tei:div[@n='$1'])">
    <p>This pointer pattern extracts chapter.</p>
  </cRefPattern>
</refsDecl>
```

### 1.1.1.3. Le profileDesc

Le <profileDesc> contient les notices de tous les noms de personnage et de lieu apparaissant dans le fichier. Chacune des notices possède un identifiant vers lequel vient pointer chacune des apparitions du lieu ou du personnage dans la Vie. Cette méthode permet de générer automatiquement à partir des fichiers XML les index des noms de lieux et des noms propres.

Les notices de nom de personnage sont regroupées dans le <particDesc> et <listPerson>. Afin de constituer les notices des noms de lieux nous avons utilisé les balises suivantes :

Elles contiennent un renvoi vers la notice BnF du personnage concerné quand elle existe, le nom du personnage et une courte notice explicative. Si le personnage est un personnage historique, ses dates de naissance et de mort sont spécifiées quand cela est possible.

```
<person corresp="http://catalogue.bnf.fr/ark:/12148/cbl19423647/PUBLIC"
  xml:id="martin">
  <persName>Martin (saint)</persName>
  <birth when="0316">316</birth>
  <death when="0397-11-08">397</death>
  <note type="biographical"
    source="http://dbpedia.org/page/Martin_of_Tours">
    <surname type="complex">Saint Martin de Tours</surname> est né
    dans en Hongrie dans la colonie romaine de <placeName ref="#sabarie">Sabaria</placeName> en 316. Il est mort à
    <placeName ref="#conde">Candes</placeName> en France le
    8 novembre 397.</note>
</person>
```

Les notices de noms de lieux, regroupées dans le <settingDesc> et <listPlace> contiennent le nom du lieu, un identifiant grâce à l'attribut @xml:id et une courte notice explicative si nécessaire. Les occurrences qui apparaissent dans le texte sont signalées par la balise <placeName> et identifiée grâce au pointeur @ref.

Afin de constituer les notices des noms de lieux nous avons utilisé les balises suivantes :

Quand le lieu est une ville identifiée, ses coordonnées sont ajoutées afin de pouvoir à terme faire une carte des lieux cités dans le corpus.

```
<place corresp="http://pleiades.stoa.org/places/423025"
  xml:id="rome">
  <placeName full="yes">Rome</placeName>
  <country type="pays">Italie</country>
  <country type="civilisation">Occident</country>
  <location>
    <geo>41.891775, 12.486137</geo>
  </location>
  <note>Ville italienne, capitale de l'empire romain d'occident.</note>
```

```
</place>
```

Pour assurer la bonne identification du lien, nous avons ajouté avec *@corresp* un lien vers une notice externe de préférence vers le site *Pleiades : Ancient Places* quand cela était possible, sinon vers *data.bnf*

Enfin, afin de pouvoir classer les différents lieux cités par pays, mais aussi en fonction de leur rattachement à l'Orient ou à l'Occident nous avons utilisé l'élément *<country>* avec *@type* de valeur pays pour signaler le pays et *@type* de valeur civilisation pour indiquer l'appartenance à l'Orient ou à l'Occident.

```
<place corresp="http://pleiades.stoa.org/places/109321"
  xml:id="amiens">
  <placeName>Amiens</placeName>
  <country type="pays">France</country>
  <country type="civilisation">Occident</country>
  <location>
    <geo>49.8936075, 2.297948</geo>
  </location>
  <note>Ville du Nord de la France.</note>
</place>
```

### 1.1.2. Structurer les différentes parties du texte

Le corpus est structuré à l'aide des balises suivantes :

Le texte est structuré à l'intérieur d'une balise *<text>* avec un attribut *@n* de valeur *edition* pour signaler que le texte encodé est le texte de l'édition et afin de le différencier, le cas échéant, du fichier contenant la traduction. *<text>* englobe l'élément *<body>* qui possède deux attributs. Le premier, *@n*, permet de spécifier l'URN CTS, tandis que le second, *@xml:lang*, indique la langue majoritaire du corpus, ici, l'ancien français signifié par la valeur *fro*. Le *<body>* présente des *<div>* de deux niveaux. La *<div>* parent possède un attribut *@type* de valeur *chapter* qui correspond à un paragraphe de notre édition. Si le texte présente un titre rubriqué, alors on le fera apparaître dans l'édition dans un élément *<head>*, enfant de la *<div>* de premier niveau. La *<div>* de niveau 2 possède un attribut *@type* de valeur *section* qui correspond à un sous-paragraphe dans l'édition. Chacune de ces divisions est numérotée, constituant ainsi un texte divisé en paragraphes et sous-paragraphes identifiés par des numéros afin de permettre une navigation plus aisée et un système de citation de l'édition efficace. À l'intérieur de la *<div>* de niveau 2, le texte est englobé dans une balise *<p>*. Cette méthode permet d'établir un lien pérenne, une référence de stable pour citer les textes du corpus grâce au respect des normes de citations CTS et à l'encodage TEI.

```
<div type="chapter" n="1">
  <head>De saint Martin</head>
  <div type="section" n="1">
    <p>[...]</p>
  </div>
</div>
```

Au sein même du corpus, en raison de la particularité stylistique de l'auteur qui consiste à ajouter des vers à l'intérieur de son texte en prose, les groupes de vers ont été réunis dans une balise *<lg>* ayant un attribut *@n*. Chacun des vers a été englobé dans une balise *<l>* avec un attribut *@n*.

```
<div type="section" n="2">
  <p>
    <lg n="1">
      <l n="1">Que demande qi a assez ?</l>
      <l n="2">Ja riches hom n'iert assazez,</l> [...] <l n="14">Qui q'il soit bel ne qui soit lait,</l>
      <l n="15">Por ce s'i devroient tenir.</l>
    </lg>
  </p>
</div>
```

Enfin, l'intégralité des parties dialoguées est englobée dans une balise *<said>* avec un attribut *@who* quand l'interlocuteur est un personnage nommé dans le récit. La valeur de l'attribut est alors un pointeur qui renvoie vers la notice du personnage dans le *<teiHeader>*.

```
Seinz Beneoiz lor dist :
<said who="#benoit">Proions nostre Signor qe vos puissiez ve<lb/>oir
celui qi cest moine en meine.</said>
```

Les éléments du texte dans une langue étrangère, latin ou grec dans notre cas, sont signalés à l'aide de l'élément *<seg>* et de son attribut *@xml:lang*

```
Si fist <seg xml:lang="lat">signum crucis</seg> seur lui.
```

## 1.2. De l'édition facsimilaire à l'édition normalisée

La réalisation nativement numérique de cette édition nous a permis de créer une édition multifacette qui peut proposer différentes visualisations du texte, une édition facsimilaire et une normalisée.

### 1.2.1. Encodage de la mise en page du manuscrit

Afin de pouvoir produire à terme une vue facsimilaire de l'édition, des informations sur la disposition du texte dans C1 ont été encodées en utilisant les balises suivantes :

Chaque saut de page est indiqué avec élément *<pb>* qui contient, grâce à l'attribut *@n*, le numéro du folio et, grâce à l'attribut *@facs*, un lien vers le folio concerné du manuscrit C1, numérisé sur le site Gallica. L'indication recto ou verso n'est pas ajoutée, car elle sera donnée par la lettre de la colonne.

```
<pb n="104">
```

```
fac="http://gallica.bnf.fr/ark:/12148/btv1b84259980/f218.highres"/>
```

Chaque saut de colonne est indiqué avec l'élément `<cb>` et contient, grâce à l'attribut `@n`, la lettre de la colonne à suivre, a et b désignant la première et la deuxième colonne du recto, tandis que c et d désigne la première et la deuxième colonne du verso.

```
<cb n="a"/>
```

Les lettrines, ainsi que les mots rubriqués ont été indiqués grâce à la balise `<hi>`.

```
<hi rend="decorated-initial">A</hi>donec
```

Ce sont les valeurs de l'attribut `@rend` qui indiquent le type de décoration présent dans le manuscrit. Voici la liste des valeurs utilisées dans le corpus :

**rub** désigne les lettres écrites à l'encre rouge.

**cat-  
ed**

**dec** désigne les lettrines décorées qui ouvrent les chapitres.

**o-  
rat-  
ed-ini-  
tial**

**ini-** désigne les majuscules bleues qui viennent parfois ouvrir une série de vers dans le corpus.

**tial-  
blue**

On trouve également dans le corpus des majuscules à l'encre rouge, mais elles semblent décoratives et ne donnent pas d'indications de structuration du texte, c'est pourquoi nous ne les avons pas signalées dans l'encodage.

Le cas des titres rubriqués est traité de manière un peu différente pour des problèmes de transformation et afin de structurer plus aisément l'édition normalisée. Nous indiquons dans le texte, à l'aide de la balise `hi` et de la valeur *rubricated* de l'attribut `@rend`, le fait que le titre est en rouge, mais nous ajoutons également la valeur d'attribut *orig* pour indiquer que c'est un élément du manuscrit afin que le texte rubriqué n'apparaisse pas dans l'édition normalisée. Toutefois, la rubrique du manuscrit est copiée comme titre du chapitre dans la balise `<head>` qui précède la `<div>` concernée.

```
<head>Miracles de saint Martin</head>
<div type="section" n="1">
  <p>
    <hi rend="rubricated orig">Miracles de saint <lb/>Mar<pc type="orig">.</pc>
    <lb/>tin</hi> [...] </p>
  </div>
```

C'est donc le texte de la balise *head* qui apparaîtra dans l'édition normalisée. Nous avons dû opter pour ce système, car il s'est avéré être le plus simple et juste d'un point de vue sémantique face à des rubriques qui pouvaient être discontinues sur plusieurs lignes. Si l'œil humain parvient très aisément grâce aux couleurs d'encre différentes à regrouper les mots du titre ensemble, l'opération devenait beaucoup plus difficile à opérer à l'aide de balise XML.

### 1.2.2. Encodage des graphies d'origine et des normalisations

L'un des buts de cette édition est de pouvoir proposer une vue facsimilaire et une vue normalisée du texte et donc de pouvoir aisément passer de l'une à l'autre. Cette plasticité est rendue possible par l'utilisation des balises : Ce balisage permet de conserver dans le code des graphies originales et de les restituer si besoin.

Les régularisations concernent les espaces entre les mots, la graphie des u/v et des i/j, l'ajout de signes diacritiques (accents et trémas) et enfin l'utilisation des majuscules dont l'usage dans le manuscrit ne correspond pas à nos usages modernes. Nous avons donc régularisé leurs emplois selon les usages en rigueur pour l'édition des textes médiévaux afin de faciliter la lecture.

```
<placeName ref="#france">
  <choice>
    <orig>f</orig>
    <reg>F</reg>
  </choice>rance
</placeName>
```

Le signalement et le développement des abréviations fonctionnent sur le même principe. Les caractères spéciaux des abréviations sont signalés par le *Numerical Character Reference Decimal* du signe concerné.

```
<choice>
  <abbr>#</abbr>
  <expan>com</expan>
</choice>
```

Afin d'alléger le code, mais aussi de procéder de manière homogène pour le développement des abréviations, les balises `<choice>` et leur contenu sont dans la plupart des cas appelés par une entité dans les fichiers XML. La liste des entités est contenue dans le fichier *.dtd entity-Wauchier.dtd* du projet.

### 1.2.3. Ponctuation originale et ponctuation normalisée

Le texte du manuscrit C1 présente des marques de ponctuation, points simples ou doubles, toutefois l'usage qu'en fait le scribe ne correspond en rien à l'usage moderne. Ainsi, il peut utiliser ces signes pour séparer des propositions relatives de leurs antécédents ou pour séparer des syntagmes au sein même d'une proposition, sans que sa pratique

n'ait rien de systématique dans notre corpus. La ponctuation a donc dû être normalisée pour faciliter la lecture du texte. Sur toutes les marques de ponctuations qui ont dû être modifiées, la balise `<pc>` a été ajoutée.

Le corpus présente trois cas de figure. Dans le premier cas, la ponctuation originale est supprimée dans l'édition normalisée. L'ajout de l'attribut `@type` de valeur *orig* sur l'élément `<pc>` signale que le signe est issu de la ponctuation du manuscrit et qu'il ne doit pas apparaître dans la version normalisée.

```
car <lb/>adonc estoit
costume en
<placeName ref="#france">
<choice>
  <orig>f</orig>
  <reg>F</reg>
</choice>rance
</placeName>
<pc type="orig" ana="completive">.</pc> qe li vilein<lb/> de la contree prenoient les ymages de
lor <lb/>deables
```

Dans le deuxième cas, le manuscrit ne présente aucun signe de ponctuation, mais il est nécessaire d'ajouter un signe pour faciliter la lecture et la compréhension du texte. L'ajout de l'attribut `@type` de valeur *reg* sur l'élément `<pc>` signale que le signe est une régularisation de l'éditeur et qu'il ne doit pas apparaître dans la vue facsimilaire de l'édition. Cette méthode est également beaucoup utilisée dans le corpus pour ajouter les apostrophes nécessaires afin de signaler les élisions.

```
l<pc type="reg">'</pc>oneur
```

Dans le troisième cas, le manuscrit présente bien un signe de ponctuation, mais celui-ci ne correspond pas à nos usages modernes. Ainsi, pour faciliter la lecture et la compréhension du texte, le signe de ponctuation doit être modifié. Dans ce cas-là, nous avons opté pour la même méthode que pour les régularisations en utilisant la balise `<choice>`.

```
<choice ana="interrogation">
  <orig>.</orig>
  <reg> ?
</reg>
</choice>
```

De même afin d'alléger la syntaxe du fichier XML, ces balises sont appelées dans le fichier grâce à une entité contenue dans le fichier dtd *entity-Wauchier.dtd* du projet.

Dans les trois cas, l'ajout de l'attribut `@ana` est utilisé pour signaler la cause du changement de ponctuation.

### 1.3. L'apparat critique

Cette édition propose un appareil critique afin de montrer les lieux de variance entre les différents témoins de la tradition et rendre compte de la modularité du texte au Moyen Âge. L'intégralité du corpus est accompagnée du relevé des variantes qui apparaissent dans les manuscrits de la famille C. En raison des grandes différences qui peuvent apparaître entre les leçons des différentes familles de manuscrits, mais aussi par manque de temps, seule l'édition de la Vie de saint Martin présente une vision complète de la tradition dont nous avons collationné les variantes de la famille C, ainsi que les variantes d'un représentant de chacune des sept autres familles de manuscrits qui comportent une version de la Vie.

#### 1.3.1. Les corrections

Le texte du manuscrit C1 est extrêmement peu fautif. Nous ne sommes intervenus que dans les cas où la leçon proposée rendait le texte difficilement compréhensible. Les corrections opérées sont dans la majorité des cas issues d'une leçon présente dans les manuscrits de la même famille, la plupart du temps C2 sur lequel nous sommes appuyés en priorité. Elles ont été encodées grâce aux balises suivantes :

Trois types de corrections ont été opérés dans le corpus. Quand le texte du manuscrit est difficilement compréhensible en raison d'un manque dans le texte, nous avons ajouté les lettres manquantes dans une balise `<corr>` avec un attribut `@type` de valeur *add*.

```
a cel coup en cha<corr type="add">ça</corr> il le deable
```

Quand le texte du manuscrit est difficilement compréhensible en raison d'un ajout de texte, la plupart du temps lié à une faute par dittographie dans le texte, les lettres en trop sont englobées dans une balise `<corr>` avec un attribut `@type` de valeur *del*.

```
Li pseudom <corr type="del">qi</corr>
<lb/>qi vit le beau liu
```

Enfin, quand le texte présente une erreur ponctuelle dans le texte qui demande le recours à un autre témoin pour remplacer la graphie "fautive" de C1, nous avons utilisé le doublet `<sic>` et `<corr>` englobé dans une balise `<choice>`. Ainsi la vue facsimilaire présentera la graphie originale du manuscrit, tandis que la vue normalisée donnera la version corrigée du texte avec entre crochets les endroits substitués par une leçon alternative issue de C2 ou C3.

```
<app>
  <lem wit="#C2 #C3">adjuto <choice>
    <sic>z</sic>
    <corr type="add">r</corr>
  </choice>
</lem>
<rdg wit="#C1">adjutoz</rdg>
</app>
```

L'intégralité des corrections opérées est également signalée dans l'apparat critique pour renseigner la graphie fautive présente dans le manuscrit de base.

```
<app>
  <lem wit="#C2 #C3">
    <corr type="add">es</corr> escriptures</lem>
    <rdg wit="#C1">escriptures</rdg>
  </app>
```

### 1.3.2. Collations des variantes et mise en place de l'apparat critique

L'apparat critique a été conçu nativement en XML. La méthode de la parallel segmentation a été choisie pour encoder l'apparat, méthode qui présente l'avantage d'être simple et lisible directement dans le code, car l'apparat est signalé dans la linéarité du texte. Elle permet également d'aisément comparer les leçons entre elles.

Cette édition fait apparaître dans l'apparat une hiérarchie entre deux types de variantes textuelles de la tradition : le lemme et les leçons. Le lemme correspond au texte édité, la plupart du temps équivalent au texte de C1, lui donnant dans l'édition la primauté sur les autres leçons. Ainsi cette édition ne propose pas de parcourir le texte de tous les témoins, mais donne à lire les différentes leçons présentes dans la tradition, à titre indicatif comme d'autres réalisations textuelles existantes du recueil de Wauchier de Denain. Les particularités de mise en page ou graphiques des manuscrits originaux, comme la présence de lettrines ou d'abréviations, n'ont donc pas été signalées contrairement au texte du lemme.

L'apparat de l'édition suit les prescriptions d'un apparat négatif, exceptionnellement mixte où nous signalons pour le lemme le manuscrit de base si ce dernier diffère du texte proposé dans C1. L'apparat ne présente que les leçons avec un texte sémantiquement différent du lemme. Chaque leçon est rattachée au manuscrit dont elle est issue par un identifiant qui renvoie dans le document XML à l'élément <witness> correspondant. Cette méthode permet toutefois de rétablir, à tout moment, les indications nécessaires à l'établissement d'un apparat positif par comparaison entre les sigles de manuscrits cités pour les leçons divergentes et la liste des manuscrits collationnés présente dans le <teiHeader> dans l'élément <listWit> du document XML.

Afin d'encoder l'apparat nous utilisons les balises suivantes :

La balise <lem> permet de délimiter le texte soumis à des variations et contient toujours le texte édité. Chaque apparat ne peut contenir qu'un élément *lem*.

```
<app>
  <lem wit="#C2 #C3">
    <corr type="add">es</corr> escriptures</lem>
    <rdg wit="#C1">escriptures</rdg>
  </app>
```

Les balises <rdg> permettent de signaler les variantes. Leur attribut @wit permet de déclarer le ou les manuscrits proposant le même texte que celui de la variante concernée. Il se réfère à l'identifiant déclaré dans le <listWit>. Quand un manuscrit présente une omission à l'emplacement où devrait se trouver le lemme, le manque est signalé par un attribut @cause de valeur om. pour une omission simple et de valeur lacune pour une omission qui s'étant sur plusieurs vers ou un paragraphe.

```
<app>
  <lem> De saint Martin </lem>
  <rdg wit="#C2">Ci commence la vie de monseigneur saint
    Martin</rdg>
  <rdg wit="#C3">Ci comence la vie saint Martin</rdg>
  <rdg wit="#G1" cause="om." />
</app>
```

Nous avons également utilisé @type sur les balises <rdg> pour signaler des types de variantes textuelles particulières. @type n'accepte qu'une liste de valeur close pour assurer l'homogénéité de l'encodage :

- ajout pour la variante qui consiste en des ajouts
- casAbs pour les leçons où le texte propose une variation sur l'expression du complément du nom avec ou sans l'expression de la préposition "de".
- etym. pour les leçons où la variante propose un terme identique avec un changement de base étymologique. Nous avons ainsi signalé les formes du verbe être étymologiques et les formes refaites sur \*estare.
- erreur pour les leçons où la variante propose une leçon issue d'une erreur évidente de copie.
- inv. pour les inversions
- outil pour les leçons où la variante consiste en la suppression, l'addition ou le remplacement d'un mot outil de type conjonction de coordination, adverbe de phrase, préposition ou conjonction de subordination et qui n'a pas d'influence majeure sur la syntaxe ou le sens de la phrase
- plur. pour les passages du singulier au pluriel
- proPer signale les leçons qui consistent en l'ajout d'un pronom personnel sujet
- sing. pour les passages du pluriel au singulier
- tps. pour les changements de temps ou de mode
- réduction. pour les endroits où le texte est réécrit de manière plus synthétique sans changement majeur de sens

```
<app>
```

```
<lem>en escripture a metre</lem>
<rdg wit="#G1" type="inv.">a metre en escripture</rdg>
</app>
```

Les leçons rejetées, très peu nombreuses, sont toujours signalées dans le texte comme des corrections ou des suppressions. La graphie fautive présente dans C1 est alors présentée dans l'apparat comme une leçon qui apparaît toujours en tête des variantes

```
<app>
<lem>
<corr type="add">es</corr> escriptures</lem>
<rdg wit="#C1">escriptures</rdg>
</app>
```

## 2. Tableau des éléments