



UNIL | Université de Lausanne

Faculté des lettres

Centre NUCLEUS



Formation de deux jours

Initiation à l'encodage XML-TEI

Lausanne, février 2025

Sonia Solfrini et Simon Gabay

(Université de Genève)



Plan du cours 2 : encodage de base du text

- L'encodage de base de la prose
- L'encodage de base de la poésie
- L'encodage de base du théâtre
- Quelques exercices

L'encodage de base de la prose

- `<pb/>` (page beginning) indique le début d'une page.
- `<p>` (paragraph) indique un paragraphe.
- `<lb/>` (line beginning) indique le début d'une ligne.
- `<hi>` (highlight) permet d'encoder n'importe quel élément différent du reste du texte (italique, gras, majuscules, petites capitales, etc.) avec l'attribut `@rend` : italic, bold, uppercase, lowercase, etc.
- ...

Plus de détails ici : <https://teibyexample.org/exist/examples/TBED03v00.htm>

L'encodage de la prose

Exemple :

La « *Text Encoding Initiative* » (TEI) est l'un des projets les plus durables et influents du champ aujourd'hui appelé « humanités numériques ». Son but est de fournir des recommandations pour la création et la gestion sous forme numérique de tout type de données créées et utilisées par les chercheurs en sciences humaines, comme les sources historiques, les manuscrits, les documents d'archives, les inscriptions anciennes et bien d'autres.

```
<?xml version="1.0" encoding="UTF-8"?>
<TEI xmlns="http://www.tei-c.org/ns/1.0">
  <teiHeader>
    [...]
  </teiHeader>
  <text>
    <body>
      <pb/>
      <p>
        <lb/>La « <hi rend="italic">Text Encoding Initiative</hi> » (TEI) est l'un des projets les plus durables et influents du champ aujourd'hui
        <lb/>appelé « humanités numériques ». Son but est de fournir des recommandations pour la création et la
        <lb/>gestion sous forme numérique de tout type de données créées et utilisées par les chercheurs en sciences
        <lb/>humaines, comme les sources historiques, les manuscrits, les documents d'archives, les inscriptions
        <lb/>anciennes et bien d'autres.
      </p>
    </body>
  </text>
</TEI>
```

L'encodage de base de la poésie

- `<lg/>` (line group) indique une strophe ou un groupe de vers.
- `<l/>` (line) indique un vers. N.B. : Il ne faut pas confondre un vers avec un retour à la ligne (`<lb/>`) !
- L'attribut `@n` peut indiquer le numéro de la strophe ou du vers.
- L'attribut `@type` peut indiquer le type de strophe ou de vers (e.g. sonnet, quatrain, etc.).
- `<head>` peut indiquer le titre d'un poème, etc.
- `<div>` (division du texte) contient une subdivision du texte, par exemple un poème.
- ...

Plus de détails ici : <https://teibyexample.org/exist/examples/TBED04v00.htm>

L'encodage de la poésie

Exemple :

Trois Anneaux pour les rois elfes sous le ciel,
Sept pour les seigneurs nains dans leurs demeures de pierre,
Neuf pour les hommes mortels destinés au trépas,
Un pour le Seigneur des Ténèbres sur son sombre trône,
Au pays de Mordor où s'étendent les ombres
Un Anneau pour les gouverner tous
Un Anneau pour les trouver
Un Anneau pour les amener tous,
Et dans les ténèbres les lier
Au pays de Mordor où s'étendent les ombres.

```
<?xml version="1.0" encoding="UTF-8"?>
<TEI xmlns="http://www.tei-c.org/ns/1.0">
  <teiHeader>
    [...]
  </teiHeader>
  <text>
    <body>
      <lg>
        <l n="1">Trois Anneaux pour les rois elfes sous le ciel,</l>
        <l n="2">Sept pour les seigneurs nains dans leurs demeures de pierre,</l>
        <l n="3">Neuf pour les hommes mortels destinés au trépas,</l>
        <l n="4">Un pour le Seigneur des Ténèbres sur son sombre trône,</l>
        <l n="5">Au pays de Mordor où s'étendent les ombres</l>
        <l n="6">Un Anneau pour les gouverner tous</l>
        <l n="7">Un Anneau pour les trouver</l>
        <l n="8">Un Anneau pour les amener tous,</l>
        <l n="9">Et dans les ténèbres les lier</l>
        <l n="10">Au pays de Mordor où s'étendent les ombres.</l>
      </lg>
    </body>
  </text>
</TEI>
```

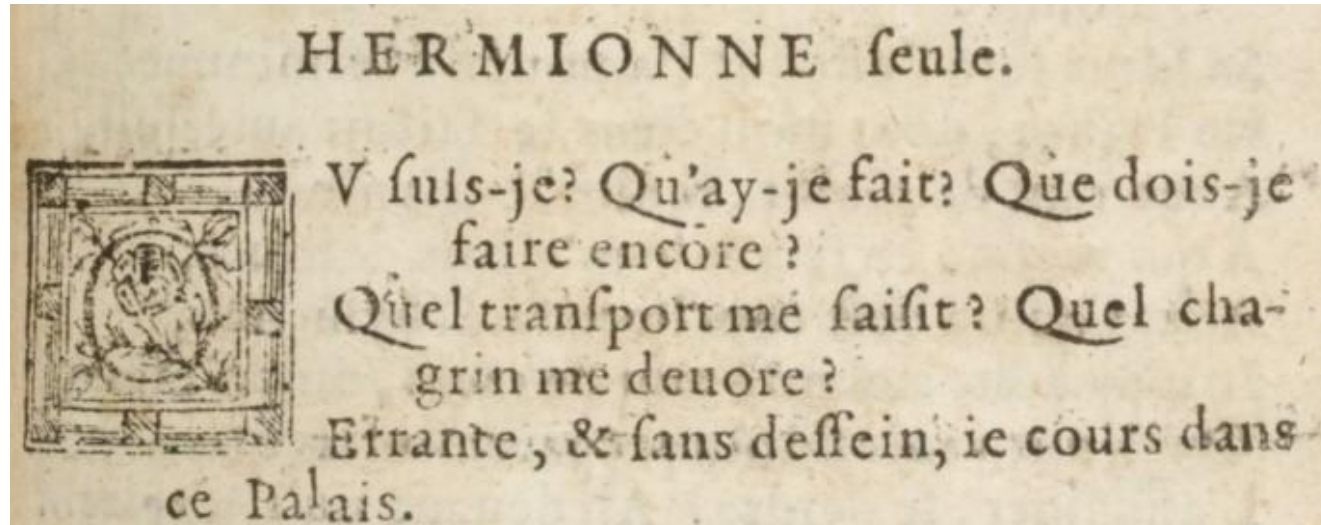

L'encodage de base du théâtre

- `<lg/>` (line group) indique un groupe de vers.
- `<l/>` (line) indique un vers.
- `<head>` peut indiquer le début d'une scène ou d'un acte (e.g. Acte premier).
- `<div>` (division du texte) contient une subdivision du texte, par exemple une scène ou un acte.
- `<sp>` indique une prise de parole.
- `<speaker>` indique le nom d'un personnage qui parle.
- `<stage>` indique une didascalie.
- ...

Plus de détails ici : <https://teibyexample.org/exist/examples/TBED05v00.htm>

L'encodage du théâtre

Exemple :



```
<?xml version="1.0" encoding="UTF-8"?>
<TEI xmlns="http://www.tei-c.org/ns/1.0">
  <teiHeader>
    [...]
  </teiHeader>
  <text>
    <body>
      <sp>
        <speaker>HERMIONNE</speaker> <stage>feule.</stage>
        <l>Ov luis-je ? Qu'ay-je fait ? Que dois-je faire encore ?</l>
        <l>Quel tran[por]t me faifit ? Quel chagrin me deuore ?</l>
        <l>Errante, & fans deffein, ie cours dans ce Palais.</l>
      </sp>
    </body>
  </text>
</TEI>
```

Nota Bene : quelques signes particuliers

- Les signes &, < et > appartiennent au code et doivent donc être *échappés*.
- < signale le début d'une balise et donc il ne peut aussi signifier “plus petit que”. Pour ce dernier sens, il faut utiliser **<** ;
- > signale la fin d'une balise et donc il ne peut aussi signifier “plus grand que”. Pour ce dernier sens, il faut utiliser **>** ;
- Pour le signe &, il faut utiliser **&** :

Exemple : <|>Errante, & fans deffein, ie cours dans ce Palais.</|>

Plan du cours 2 : encodage de base du text

- L'encodage de base de la prose
- L'encodage de base de la poésie
- L'encodage de base du théâtre
- Quelques exercices

```

<text>
  <body>
    <div type="sonnet" n="1" xml:id="s1">
      <head>Sonetz</head>
      <lg type="quatrain" rhyme="abba">
        <l n="1">QVI voudra voyr comme vn Dieu me <rhyme label="a">furmonte</rhyme>,</l>
        <l n="2">Comme il m'affault, comme il fe fait <rhyme label="b">vainqueur</rhyme>:</l>
        <l n="3">Comme il r'enflamme, & et renglace mon <rhyme label="b">cucœur</rhyme>,</l>
        <l n="4">Comme il reçoit un hõnneur de ma <rhyme label="a">honte</rhyme>.</l>
      </lg>
      <lg type="quatrain" rhyme="abba">
        <l n="5">Qui voudra voir une ieuneffe <rhyme label="a">prompte</rhyme></l>
        <l n="6">A fuivre en vain l'obiet de fon <rhyme label="b">malheur</rhyme>,</l>
        <l n="7">Me vienne voir: Il voirra ma <rhyme label="b">douleur</rhyme>,</l>
        <l n="8">Et la rigueur de l'Archer qui me <rhyme label="a">donte</rhyme>.</l>
      </lg>
      <lg type="tercet" rhyme="ccd">
        <l n="9">Il cognoistra combien la raifon <rhyme label="c">peult</rhyme></l>
        <l n="10">Contre fon arc, quand vne fois il <rhyme label="c">veult</rhyme></l>
        <l n="11">Que nostre cucœur fon efclauē <rhyme label="d">demeure</rhyme>:</l>
      </lg>
      <lg type="tercet" rhyme="ccd">
        <l n="12">Et fi voirra que ie luis trop <rhyme label="c">heureux</rhyme>,</l>
        <l n="13">D'auoir au flanc l'aiguillon <rhyme label="c">amoureux</rhyme>,</l>
        <l n="14">Plein du venin dont il fault que ie <rhyme label="d">meure</rhyme>.</l>
      </lg>
    </div>
  </body>
</text>

```

Nota Bene : les rimes

- Pour indiquer un schéma métrique, on peut ajouter l'attribut **rhyme** à l'élément **lg** et/ou l'élément **rhyme** avec l'attribut **label** :

```
<lg type="quatrain" rhyme="abba">  
  <l n="1">QVI voudra voyr comme vn Dieu me <rhyme label="a">furmonte</rhyme>,</l>  
  <l n="2">Comme il m'affault, comme il fe fait <rhyme label="b">vainqueur</rhyme>:</l>  
  <l n="3">Comme il r'enflamme, & et renglace mon <rhyme label="b">cuœur</rhyme>,</l>  
  <l n="4">Comme il reçoit un hōnneur de ma <rhyme label="a">honte</rhyme>.</l>  
</lg>
```

```

<text>
  <body>
    <div type="play">
      <head>ANDROMAQUE, </b><hi rend="italic">TRAGEDIE.</hi></head>
      <div type="act" n="1">
        <head>ACTE PREMIER.</head>
        <div type="scene" n="1">
          <head>SCENE PREMIERE.</head>
          <stage>ORESTE, PYLADE.</stage>
          <sp>
            <speaker>ORESTE.</speaker>
            <lg>
              <l n="1"><hi rend="initiale">O</hi>VY, puis que ie retrouue vn Amy fi </b>fidelle,</l>
              <l n="2">Ma Fortune va prendre vne face nou <b break="no" rend="-"/>uelle;</l>
              <l n="3">Et déjà fon courroux femble s'eftre </b>adouci,</l>
              <l n="4">Depuis qu'elle a pris foin de nous rejoindre ici.</l>
            </lg>
          </sp>
        </div>
      </div>
    </div>
  </body>
</text>

```


Nota Bene : la coupure de mot en fin de ligne

- Pour indiquer une coupure de mot en fin de ligne, on peut ajouter `break="no" rend="-"` à l'élément `<lb>` marquant le saut de ligne :

```
<sp>
  <speaker>ORESTE.</speaker>
  <lg>
    <l n="1"><hi rend="initiale">O</hi>VY, puis que ie retrouve vn Amy fi <lb/>fidelle,</l>
    <l n="2">Ma Fortune va prendre vne face nou <lb break="no" rend="-"/>uelle;</l>
    <l n="3">Et déjà fon courroux femble s'eftre <lb/>adouci,</l>
    <l n="4">Depuis qu'elle a pris foin de nous rejoindre ici.</l>
  </lg>
</sp>
```