



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 5 : l'intervention éditoriale et la mise en page

- Les corrections
- Les normalisations orthographiques
- Ajouts et suppressions
- Les abréviations
- Les détériorations
- La mise en page

# Les corrections

Il est possible de corriger le texte avec l'élément `<corr>` :

Pylade est un home → Pylade est un hom`<corr>m</corr>`e

Au cas où l'on perd de l'information lors de la correction, on peut aussi la conserver avec un `<choice>` :

Pylade est un

`<choice>`

`<sic>home</sic>`

`<corr>homme</corr>`

`</choice>`

En cas d'oubli manifeste, on peut toujours rajouter un mot avec `<supplied>` :

Pylade un homme → Pylade `<supplied>est</supplied>` un homme

# Les normalisations orthographiques

Pour normaliser tout en conservant le texte original, il est possible d'utiliser les balises `<orig>` et `<reg>`, encadrées par une balise `<choice>` :

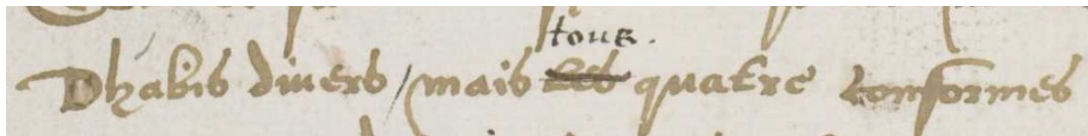
```
<choice>
  <orig>Hermione a Pyrrhus prodiguoit tous...</orig>
  <reg>Hermione à Pyrrhus prodiguait tous...</reg>
</choice>
```

On peut aussi opérer au niveau du mot ou au niveau de la lettre :

```
<choice>
  <orig>prodiguoit</orig>          prodigu<choice><orig>o</orig><reg>a</reg></choice>it
  <reg>prodiguait</reg>
</choice>
```

## Ajouts et suppressions

- `<del>` : une portion de texte effacée ou raturée avec `@type` pour le type de suppression ;
- `<add>` : une portion de texte ajoutée avec `@type` pour le type d'ajout ;
- `<subst>` : substitution d'une portion de texte par une autre.



```
<p>
D'habis divers mais
<subst>
  <del>les</del>
  <add type="above">tous</add>
</subst>
quatre conformes
</p>
```

# Les abréviations

- <choice> : proposition d'un encodage alternatif.
- <abbr> : la forme abrégée ;
- <expan> : son expansion ;
- <am> : les lettres abrégées ;
- <ex> : les lettres ajoutées pour résoudre l'abréviation.

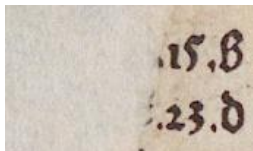


```
<p>
  blabla
  <choice>
    <abbr>chāter</abbr>
    <expan>chanter</expan>
  </choice>
  blabla.
</p>
```

```
<p>
  blabla
  <choice>
    <abbr>ch<am>ā</am>ter</abbr>
    <expan>ch<ex>an</ex>ter</expan>
  </choice>
  blabla.
</p>
```

## Les détériorations

- `<gap/>` : portions de textes illisibles.
- `<damage>` : portions de textes abîmées (eau, trous, coupe...), mais que l'on peut encore lire avec un certain degré de certitude.
- `<unclear>` : portions de textes qui ne peuvent pas être transcrites avec certitude.
- `<supplied>` : texte ajouté par l'éditeur scientifique pour n'importe quelle raison.

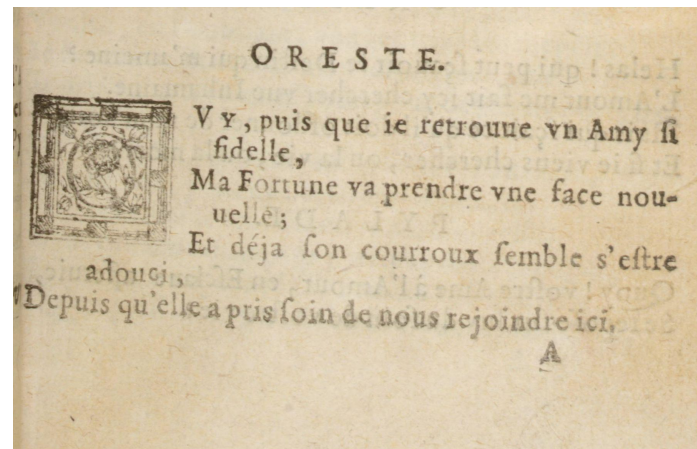


```
<p>  
  <gap/>.15.b  
  <gap/>.23.d  
</p>
```

# La mise en page

Tous les éléments de mise en page (titre courant, réclame, signature de cahier...) sont encodés avec la balise `<fw>`.

```
<sp>  
  <speaker>ORESTE</speaker>  
  <l>OVY, puis que ie retrouve vn Amy fi  
    <lb/>fidelle,</l>  
  <l>Ma fortune va prendre vne face nou-  
    <lb/>uelle;</l>  
  <l>Et déjà son courroux femble s'estre  
    <lb/>adouci,</l>  
  <l>Depuis qu'elle a pris soin de nous rejoindre ici.</l>  
  <fw>A</fw>  
</sp>
```





# La mise en page

Comme l'élément `<fw>` peut encoder de nombreuses choses, il est possible de :

- regrouper tout dans un même `<fw>` :

```
<fw>  
  <lb/>TRAGEDIE 75  
  <lb/>CLEONNE  
</fw>
```

- Distinguer plusieurs `<fw>` :

```
<fw type="runningTitle" place="top-centre">TRAGEDIE</fw>  
<fw type="pageNum" place="top-right">75</fw>  
<fw type="catchword" place="bot-centre">CLEONNE</fw>
```

