



Séance 4

Structure du document TEI

Anatomie d'un encodage XML TEI

Correction

Encodage du
Misanthrope



Personnages

```
<castList>
  <castItem xml:id="alceste">
    Alceste
  </castItem>
</castList>
...
<sp who="#alceste">
  <speaker>Alceste</speaker>
  <l>...</l>
</sp>
```

Numérotation des vers

```
<sp who="#philinte">
  <speaker>Philinte</speaker>
  <l n="1" part="I">Qu'est-ce donc ? qu'avez-vous ?</l>
</sp>
<sp who="#alceste">
  <speaker>Alceste</speaker>
  <stage>assis</stage>
  <l n="1" part="F">Laissez-moi, je vous prie.</l>
</sp>
```

A decorative graphic on the left side of the slide consisting of two overlapping squares. The bottom-left square is a dark blue, and the top-right square is a lighter blue, creating a cross-like shape.

Structure globale

Composants d'un document TEI

Composants minimaux

Tout document TEI contient a minima deux parties :

- Un **en-tête**, représenté au moyen d'un élément `<teiHeader>` contenant des métadonnées décrivant le document ;
- le **texte** lui-même, représenté par un élément `<text>`.

TEI

`teiHeader`

`facsimile`

`text`

`...`

```
<TEI xmlns="http://www.tei-c.org/ns/1.0">
  <teiHeader>
    <fileDesc>
      <titleStmt>
        <title></title>
      </titleStmt>
      <publicationStmt>
        <p></p>
      </publicationStmt>
      <sourceDesc>
        <p></p>
      </sourceDesc>
    </fileDesc>
  </teiHeader>
  <text>
    <body>
      <p></p>
    </body>
  </text>
</TEI>
```


Header TEI minimal

<titleStmt>

informations
identifiant le
document lui-même

<publicationStmt>

informations sur la façon
dont il est distribué ou
publié

<sourceDesc>

indications sur ses
origines

Contenu de l'élément text

<front>

pour les préfaces, et
tous les éléments
liminaires du texte

<text>

pour le corps du texte
proprement dit

<back>

pour tous les
appendices, épilogues,
postfaces, etc.

A decorative graphic on the left side of the slide consisting of two overlapping squares. The bottom-left square is a dark blue, and the top-right square is a lighter blue, creating a cross-like shape.

TEI header

Structuration des métadonnées du document

Convention de nommage

_Stmt

contenu structuré en
sous-éléments

_Desc

description en prose,
pouvant admettre
des sous-éléments

_Decl

Déclarations associant
sous-éléments et
description en prose

teiHeader

fileDesc

encodingDesc

profileDesc

revisionDesc

fileDesc

OBLIGATOIRE

Description bibliographique du document (ce qu'on pourrait renseigner dans une notice de catalogue (sauf sa langue et date de création)).

<titleStmt>

Déclaration du titre

<publicationStmt>

Déclaration de la publication

<sourceDesc>

Description de la source

fileDesc

titleStmt

publicationStmt

sourceDesc

editionStmt

...

encodingDesc

Description des choix d'encodage

<p>

Prose courante pour décrire la relation entre le document numérique et sa source

<variantEncoding>

Déclaration de la méthode de constitution de l'apparat critique

profileDesc

Description détaillée des aspects non bibliographiques :
informations contextuelles (sujet, contexte de production,
langue, etc.)

<creation>

Information sur la création du texte

<handNotes>

Information sur les mains présentes
dans le document édité

revisionDesc

Information sur les modifications apportées à l'encodage depuis sa version initiale

<listChange>

Liste de modifications effectuées

<change>

Description d'une des modifications effectuées avec sa date, son auteur, sa nature, etc.

Modules du *header*

Le *header* peut être étendu par différents modules optionnels en fonction de la nature du texte encodé : `msdescription` pour les manuscrit, `textcrit` pour l'édition critique

A decorative graphic on the left side of the slide consisting of two overlapping squares. The front square is a medium blue and the back square is a darker blue.

Divisions du document

Contenu de l'élément <text>

Qu'est-ce que le text

Renferme le contenu encodé en tant que tel : texte de toute nature, qu'il soit unitaire ou composite, par exemple un poème ou un drame, un recueil d'essais, un roman, un dictionnaire ou un échantillon de corpus.

text

front

body

back

...

Structuration de la prose

Balise Utilisation

p Paragraphes

seg Subdivisions internes au paragraphe

list Liste contenant des **item**

quote Citations simples, (**q** pour ce qui est entre guillemets)

cit Citations avec référence (**bibl**)

table Tableaux

figure Images

Structuration matérielle

Balise Utilisation

- pb** Changement de page (*page beginning*)
avec @n pour le n° de page qui commence)
- cb** Changement de colonne (*column beginning*)
- lb** Changement de ligne (*line beginning*)

Exercice

Compléter un
teiHeader



Consigne

En vous appuyant sur le template de `teiHeader` fourni dans le *repository*, compléter au maximum le `teiHeader` pour correspondre à un encodage fictif des [*Mystères de Paris* d'Eugène Sue](#) en vous appuyant sur la notice de [catalogue](#), en supprimant les éléments non pertinents pour ladite édition

Exercice

A decorative graphic on the right side of the slide, consisting of a light gray square positioned above a blue square, both of which are partially cut off by the right edge of the frame.

**Éditer numériquement
une pièce de théâtre**

Consigne

À partir de la transcription du *Cid* au format `txt` fourni dans le *repository*, structurer grossièrement les grandes parties de la pièce et de sa préface et compléter le `teiHeader`