

Stage d'ecdastique 2019

Atelier pratique XML / TEI

Emmanuelle Morlock (CNRS, laboratoire HiSoMA)

Lyon, Institut des Sources chrétiennes , 20 février 2019

Atelier pratique

- Mise en place des outils
 - GitHub
 - tei-web-editor (Jeffrey C. Witt)
- Introduction au processus de balisage TEI
- Exercice 1
 - balisage simple à partir d'un "Testament de poilu"
- Balises de l'apparat critique
 - <app>, <lem>, <rdg>, <wit>
- Exercice 2 avec l'outil "critical edition toolbox"
 - correction d'erreurs dans un fichier d'édition critique pré-encodé

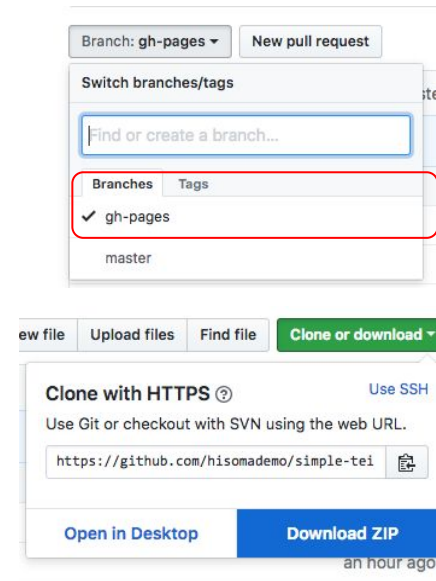
Outils 1 GitHub

Connexion

- Aller sur <https://github.com/>
- Cliquer sur “Login”
- Nom d'utilisateur : hisomademo@gmail.com
- Mot de passe :
- Laisser la fenêtre ouverte

Téléchargement

- Téléchargez la branche ‘gh-pages’ du projet [simple-tei-edition](#)



Outils 2



- Dans un nouvel onglet, ouvrez <https://tei-web-editor.herokuapp.com/>
- Aller dans le menu **views** > **toggle Mirador** pour fermer zone de visualisation
- Cliquer sur le bouton **Login** pour établir la connexion à l'entrepôt GitHub **hisomademo**

Outils 3 et 4



<https://teibyexample.org/>

TEI Validator Service

Pour valider vos encodages (XML bien formé, conformité par rapport à la TEI).



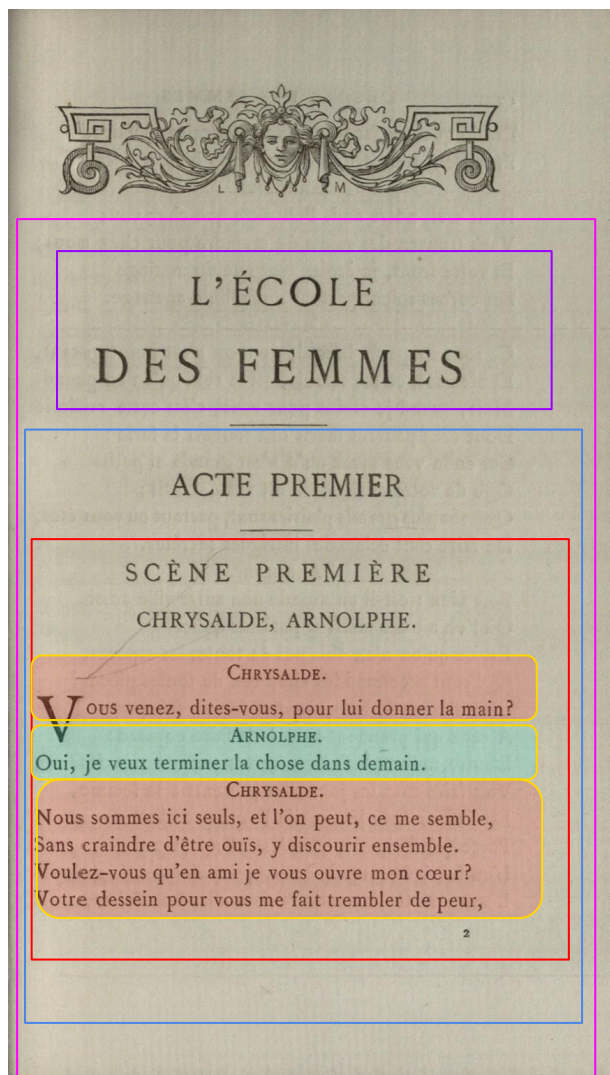
<http://teicat.huma-num.fr>

Marjorie Burghart (CNRS)

Pour préparer une édition critique numérique native, vérifier son travail et afficher l'édition.

**Le balisage TEI est juste
une manière de
représenter votre
“lecture” d’un texte sous
une forme “lisible pour
l’ordinateur”**

Formalisation de la structure logique du texte



Source gallica.bnf.fr / Bibliothèque nationale de France

<text>

<body>

<head>L'École des femmes <lb/><hi rend="i">Comédie</hi></head>

<div1 type="act" xml:id="I">

<head>Acte Premier</head>

<div2 type="scene" xml:id="I01">

<head>Scène Première</head>

<sp who="chrysalde" xml:id="I01-1">

<l>Vous venez, dites-vous, pour lui donner la main ?</l>

</sp>

<sp who="arnolphe" xml:id="I01-2">

<l>Oui, je veux terminer la chose dans demain.</l>

</sp>

<sp who="chrysalde" xml:id="I01-3">

<l>Nous sommes ici seuls, et l'on peut, ce me semble,</l>

<l>Sans craindre d'être ouïs, y discourir ensemble.</l>

<l n="5">Voulez-vous qu'en Ami je vous ouvre mon cœur ?</l>

<l>Votre dessein, pour vous, me fait trembler de peur ;</l>

<l>Et de quelque façon que vous tourniez l'affaire,</l>

<l>Prendre Femme, est à vous un coup bien téméraire.</l>

</sp>

</div2>

</div1>

</body>

</text>

et aussi pour les humains...

<lb ed="T"/>何必能稱美談。」

<persName key="A001583">什</persName>每有疑義，必
共諮決。時偽 <lb ed="T"/>秦主<persName
key="A002097">姚興</persName>專志經法，供養三千餘
僧，<placeName key="PL000000042002">關中
</placeName>舊僧 (...)

Que veut-on représenter ?

Portez ce
vieux ~~cognac~~ **whisky**
au juge blond
qui fume *la pipe*.

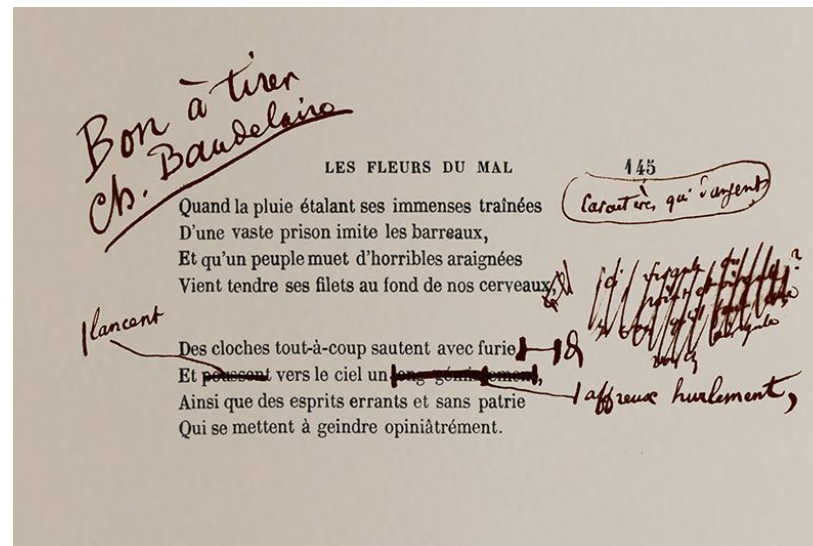
Que peut-on représenter ?

structure du document (titres, paragraphes, strophes)

Structures linguistiques (parties du discours, phrases, mots...)

Références à des personnes ou objets du monde réel, des événements

Métadonnées (auteur, date de création, provenance, genre)



Disposition, mise en page, analyse des formes graphiques et manuscrites

Processus : ratures, campagnes de correction

Annotations, notes de régie...

Le processus de balisage

1. analyse du document
2. sélection des caractéristiques à représenter ou expliciter
3. choix des balises TEI
4. saisie des balises
5. validation

Un exemple commenté



<https://testaments-de-poilus.huma-num.fr/#!/edition/61/141>



Testament de [François Rebeyrotte](#)

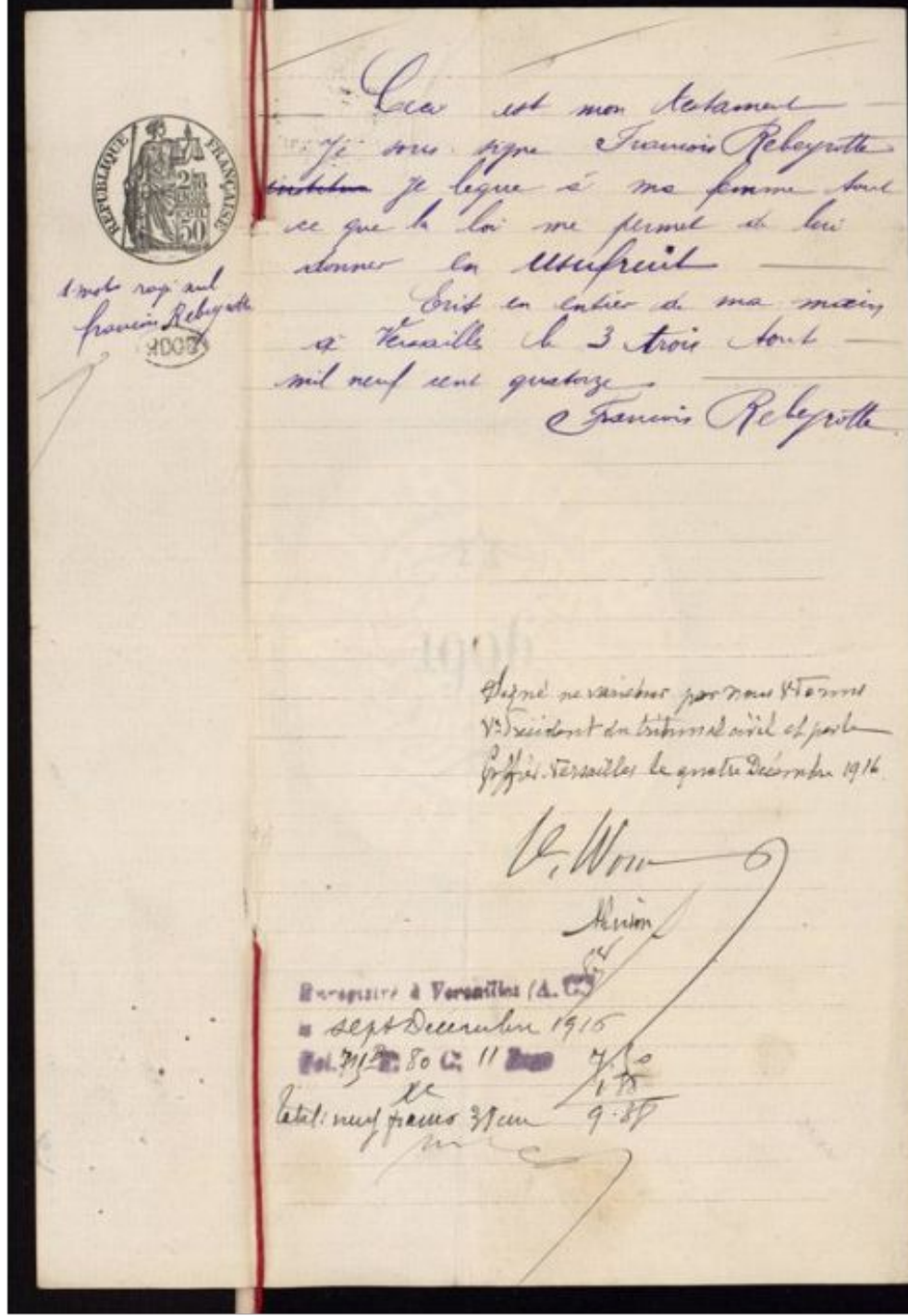
Mort pour la France le 13 juin 1916 | [Baleycourt \(Meuse\)](#)

Testament rédigé le 3 août 1914 | [Versailles](#)

Cote aux [Archives départementales des Yvelines](#) 3E44

1002, minute du 5 décembre 1916

papier timbré réglé, encre violette, 24,9cm x 17,8cm



Transcription

Ceci est mon testament

Je sous signe Francois Rebeyrotte

~~institue~~ Je legue à ma femme tout

ce que la loi me permet de lui

donner en Usufruit.

Ecrit en entier de ma main

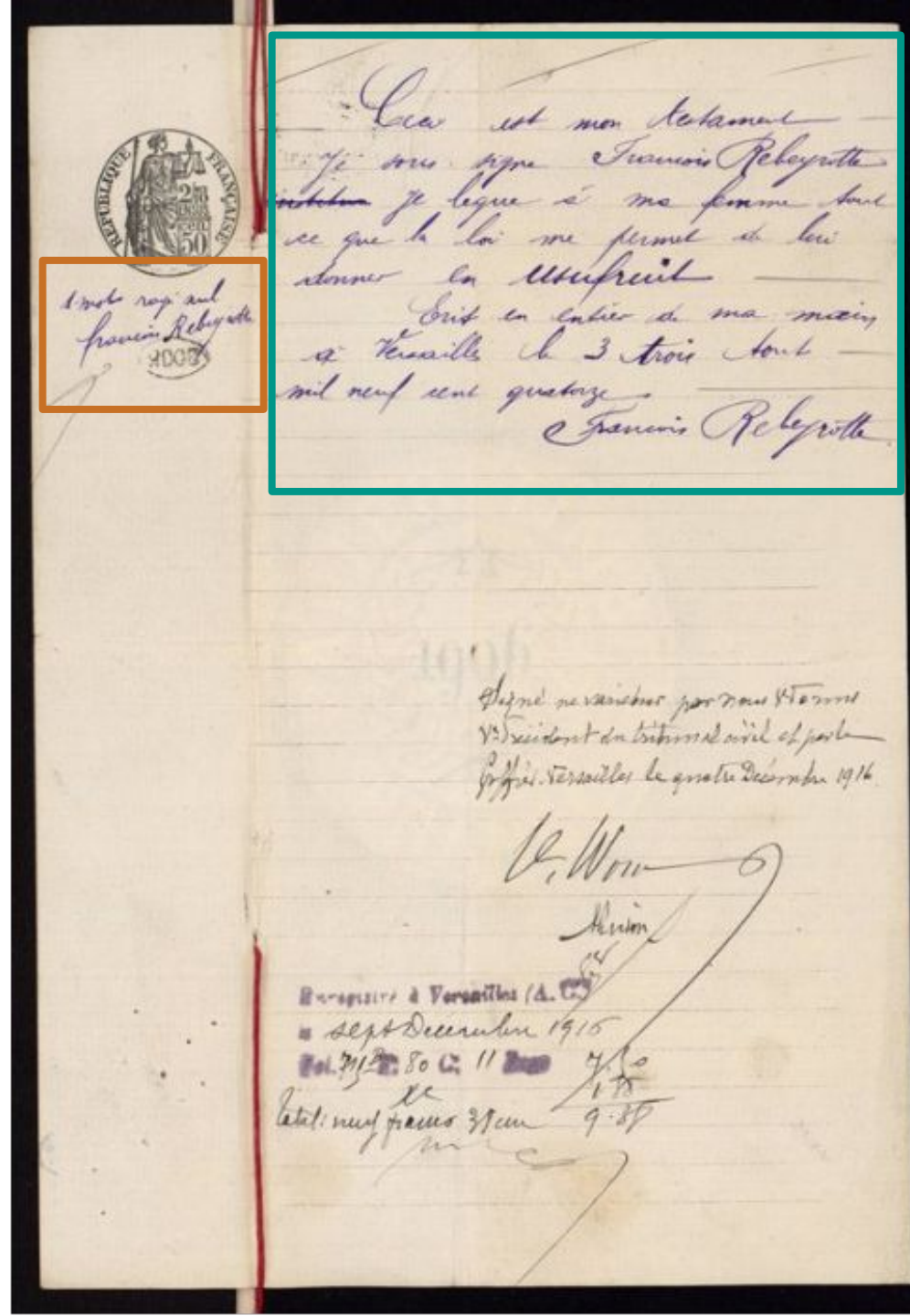
à Versailles le 3 trois Aout

mil neuf cent quatorze.

Francois Rebeyrotte

1 mot rayé nul

Francois Rebeyrotte

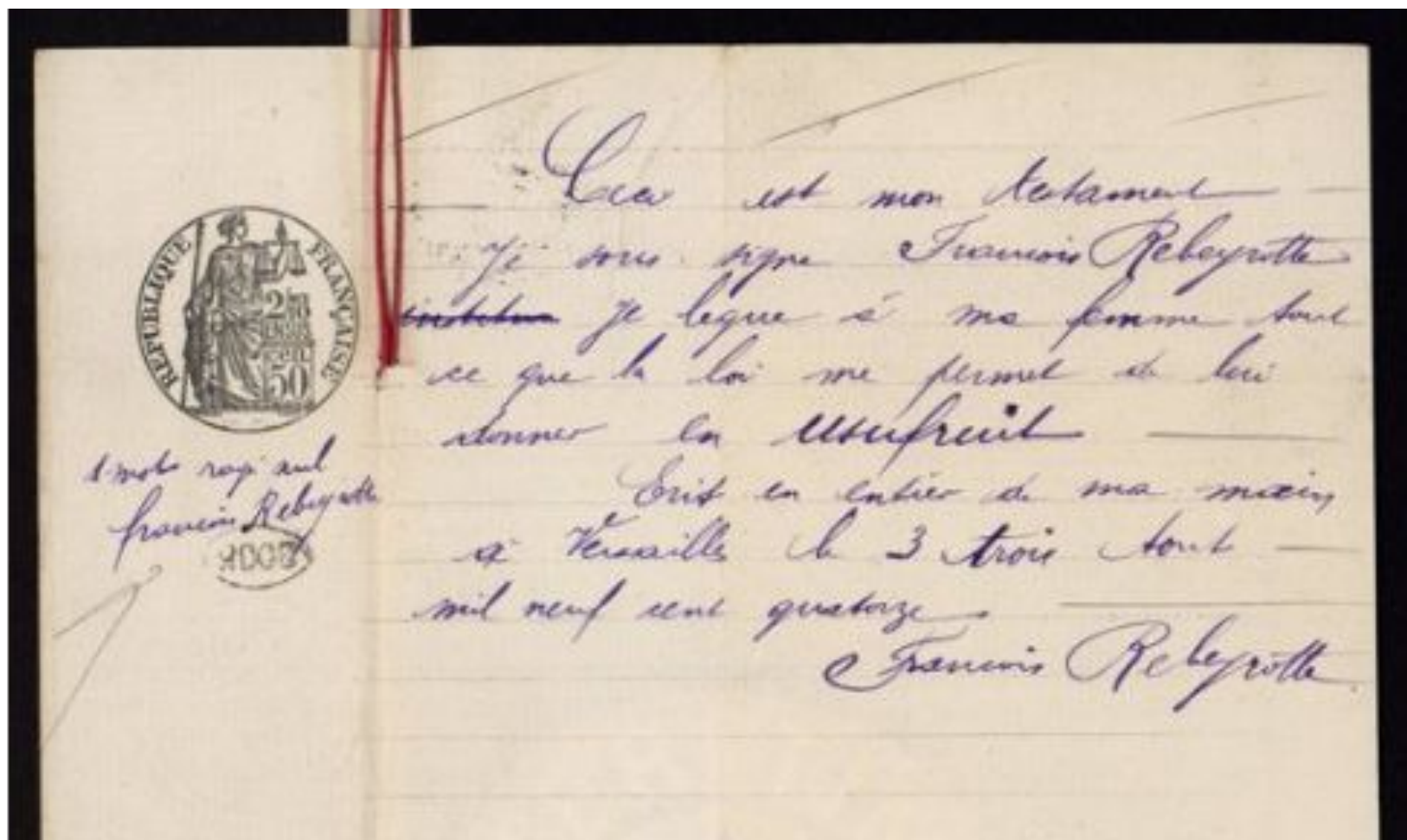


Quelles balises pour quels composants du texte ?

Quelles balises TEI allez-vous utiliser pour représenter :

- le corps de texte et ses parties
- le titre
- les paragraphes
- les lignes
- les noms, les dates,
- la signature...

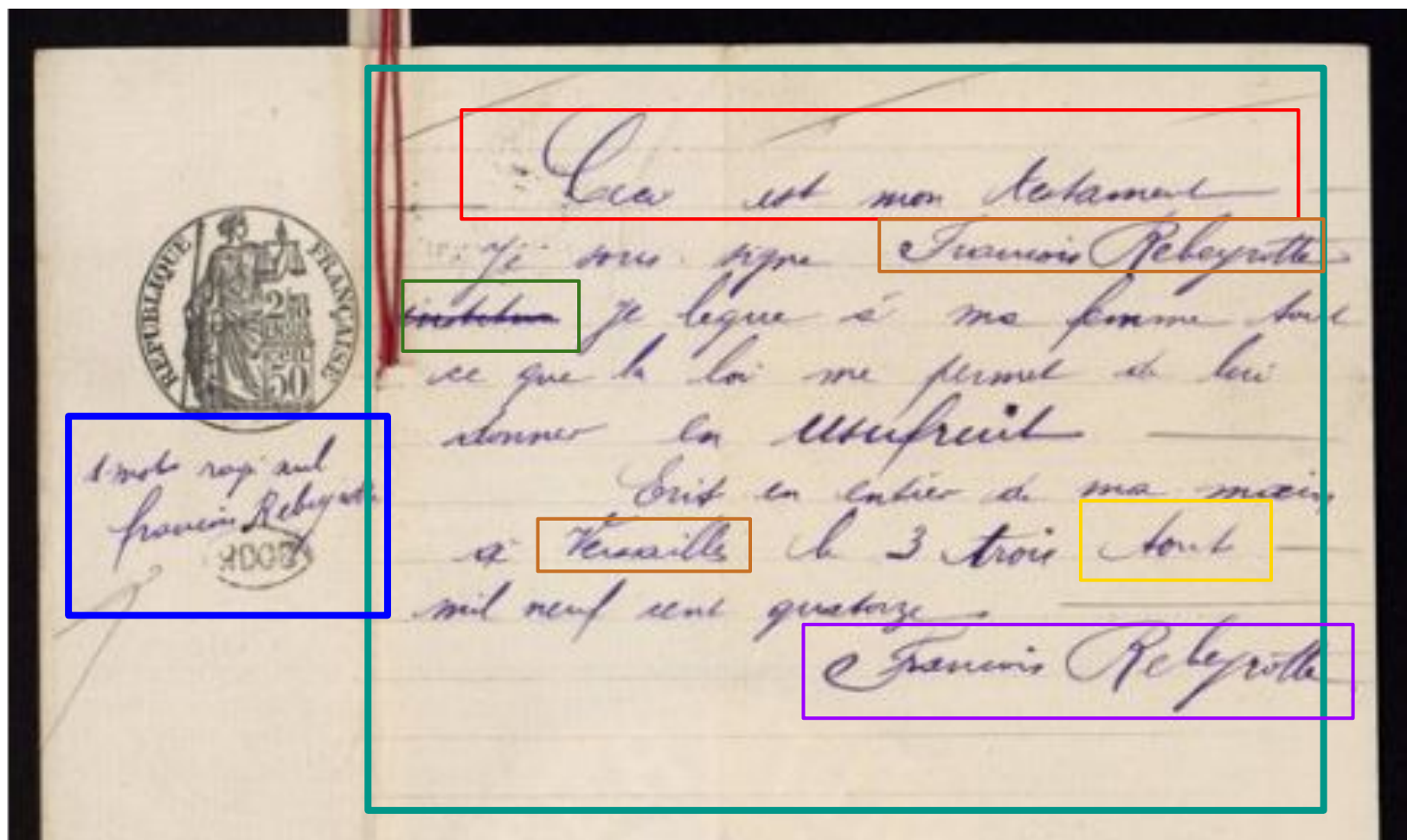
Analyse préalable



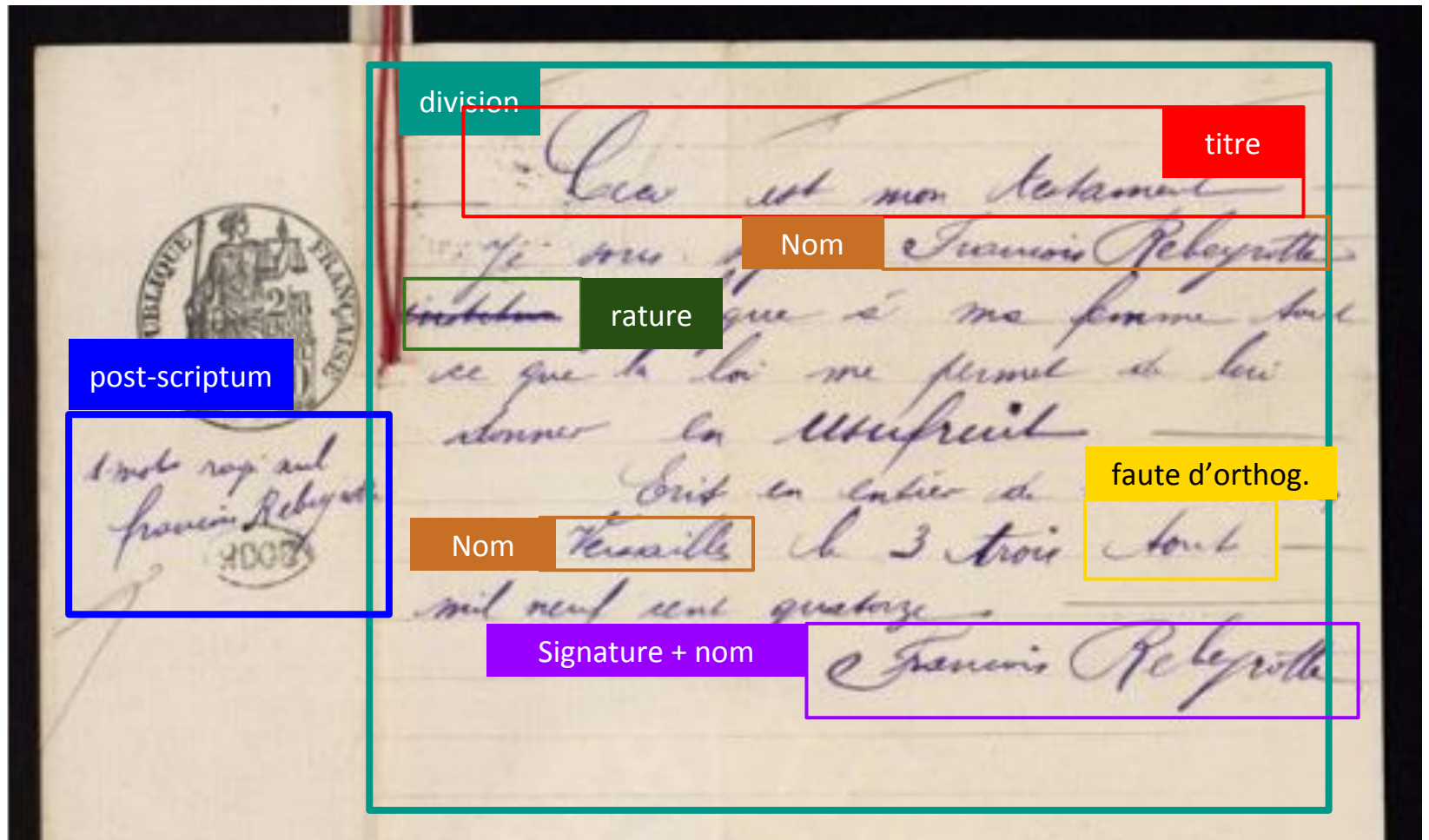
Transcription

Ceci est mon testament
Je sous signe Francois Rebeyrotte
~~institute~~ Je legue à ma femme tout
ce que la loi me permet de lui
donner en Usufruit.
Ecrit en entier de ma main
à Versailles le 3 trois Aout
mil neuf cent quatorze.
Francois Rebeyrotte

Analyse préalable



Analyse préalable



Ceci est mon testament

titre

Je soussigné François P
~~testament~~ je lègue à ma femme tout
ce que la loi me permet de lui
donner en usufruit

paragraphe

Soit en entier de ma
à Versailles le 3 trois tout
mil neuf cent quatorze

paragraphe?

François P

signature

Ceci est mon testament

titre

Je soussigné François
~~testateur~~ je lègue à ma femme tout
ce que la loi me permet de lui
donner en usufruit

paragraphe

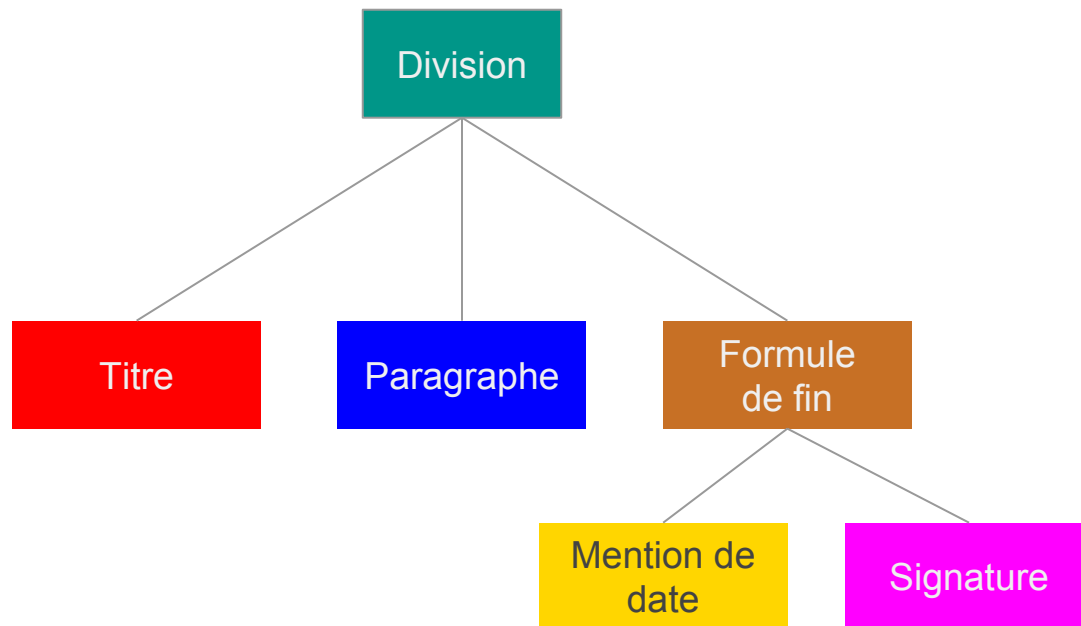
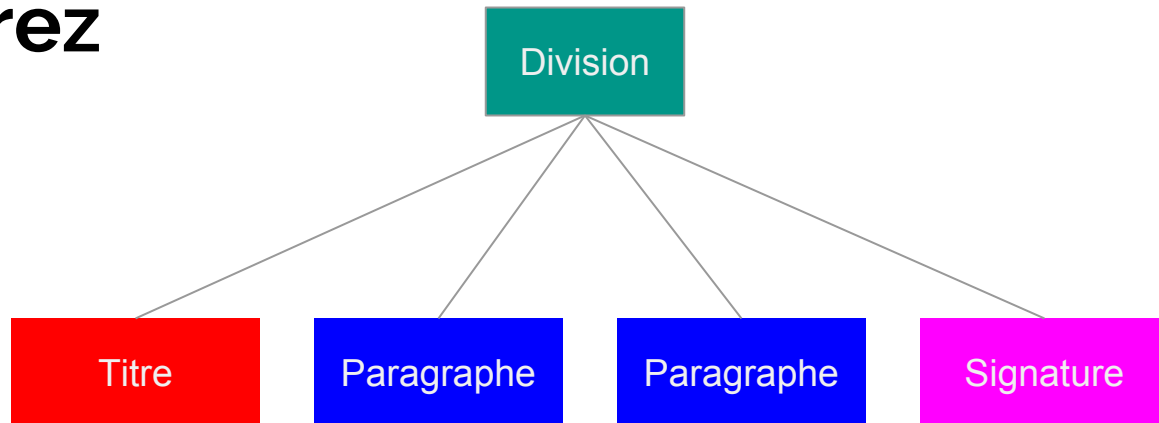
En fait en entier de ma
à Marseille le 3 trois tout
mil neuf cent quatorze

formule de fin

François

signature

Comparez



Définitions TEI

<dateline> (mention de date)

Contient une brève description des lieux, date, heure, etc. concernant la production d'une lettre, d'un article de journal ou d'un autre texte qui, placée au début ou à la fin, lui est associée comme en-tête ou annonce de fin.

<https://www.tei-c.org/release/doc/tei-p5-doc/fr/html/ref-dateline.html>

<closer> (formule finale)

Regroupe une formule de politesse, une indication d'une date et d'autres expressions semblables figurant comme expression à la fin d'une division, en particulier à la fin d'une lettre.

<https://www.tei-c.org/release/doc/tei-p5-doc/fr/html/ref-closer.html>

<signed> (signature)

Contient la dernière salutation, ajoutée à un avant-propos, à une dédicace ou à une autre division du texte.

<https://tei-c.org/release/doc/tei-p5-doc/en/html/ref-signed.html>


Intervention sribale : correction

```
<body>
  <pb/>
  <div>
    <head>Ceci est mon testament</head>
    <p>Je sous signe Francois Rebeyrotte
    <lb/><del>institue</del> Je legue à ma femme tout
    <lb/>ce que la loi me permet de lui
    <lb/>donner en Usufruit.</p>
    <closer><dateline>Ecrit en entier de ma main
    <lb/>à Versailles le 3 trois
    <choice><sic>Aout</sic><corr>août</corr></choice>
    <lb/>mil neuf cent quatorze.</dateline>
    <signed>Francois Rebeyrotte</signed>
    </closer>
  </div>
</body>
```

Exercise 1



Etape 1 = ouvrir le fichier exo_votreprénom.xml

- dans l'éditeur en ligne, cliquez  dans la barre de menus
- puis cliquer sur le lien ci dessous :

My Repositories

<https://api.github.com/repos/hisomademo/simple-tei-edition>

- puis sous “Branch”, cliquer sur “**gh-pages**”
- puis “**xml**”
- ouvrez votre fichier (exemple : “**exo_claire.xml**”)

Etape 2 : Saisie des balises

- D'abord la structure logique principale :
 - <div>, <head>, <p>, <closer>
- Ensuite la structure physique :
 - <pb/>, <lb/>
- Enfin des éléments plus fins :
 - <closer>, <dateline>, <signed>
 -
 - <choice>, <sic>, <corr>
- Vous pourriez avoir besoin de baliser également :
 - les dates (<date>)
 - les chiffres (<num>)
 - les noms de lieux (<placeName>), de personne (<persName>)
- et vouloir expliciter certaines information (via les attributs) :
 - <div type="testament">
 - <date when="1914-08-03">

Etape 3 : enregistrement avec message explicatif

File Path
`https://api.github.com/repos/hisomademo/simple-tei-edition/contents/xml/exo_claire.xml?ref=gh-pages`

Repo
 <https://api.github.com/repos/hisomademo/simple-tei-edition>

Path

Filename

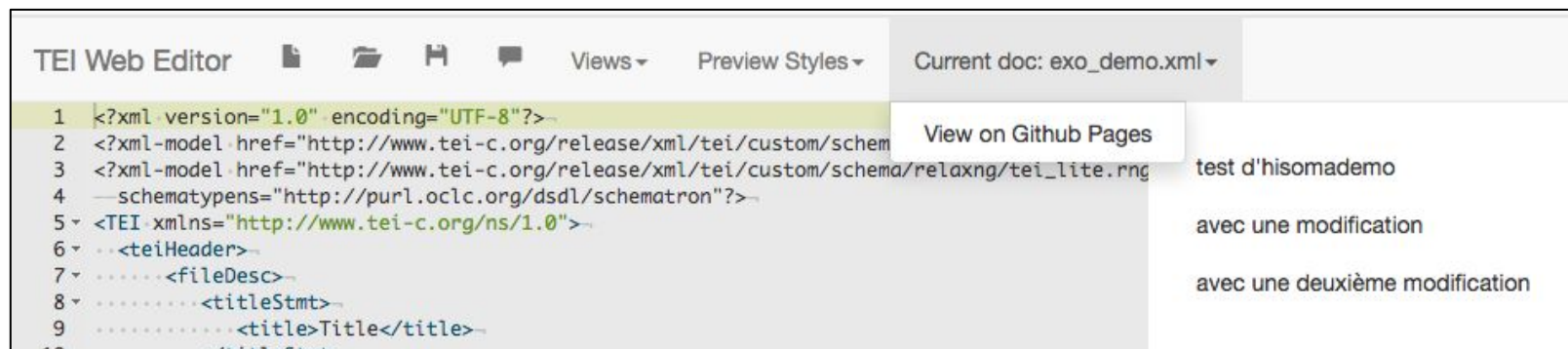
Branch

Last Commit

Message

Create new repository

Etape 4 : visualisation du résultat



- Dans le menu, cliquer sur “Current doc::” et choisir “View on Github Pages”
- Sélectionnez votre fichier dans la page qui s’affiche :

Atelier pratique stage d'ecdote - Sources Chrétiennes 2019

Exercice 1 - fichiers de travail

[Carmine](#)

[Djair](#)

[Élodie](#)

[Elvira](#)

[Federico](#)

[Florence B](#)

[Florence K](#)

[Franck](#)

Etape 5 : validation dans teiByExample

<https://teibyexample.org/xquery/TBValidator.xq>

- Copier le code TEI pour le coller dans la fenêtre du validateur
- Vous pouvez aussi récupérer le fichier sur GitHub et le charger dans l'interface (attention vos fichiers sont dans la branche “gh-pages” et non “master”)

TBE validation service

Enter an XML structure here [Show Help](#) [Maximize](#)

```
<div type="minute">
  <head rend="centered"> Ceci est mon testament</head>
  <p> Je <choice>
    <sic>sous signe</sic>
    <corr>soussigné</corr>
  </choice> <persName ref="60"><choice>
```

Or choose a file to upload

☐ Validate input as autonomous TEI fragment

Validation Result [Copy to input](#) [Maximize](#)

```
<TEI xmlns="http://www.tei-c.org/ns/1.0">
  <teiHeader>
    <fileDesc>
      <titleStmnt>
```



Except where otherwise noted content on this site is licensed under a [Creative Commons Attribution ShareAlike 3.0 License](#)
Contact: teibyexample@kantl.be



Balises pour l'apparat critique



3 balises & 1 attribut à connaître

<witness>	Identification et description des témoins
<app>	note d'apparat critique
<lem>	leçon préférée
<rdg>	témoin
@wit	renvoi depuis <lem> ou <rdg> vers un ou plusieurs <witness>

Les 3 méthodes de balisage TEI pour l'apparat critique

3 manières de relier le texte et l'apparat :

- Location-referenced
- Double-end-point attachement
- Parallel segmentation

Exemple

Vade retro Satana, scandalum es nobis^(a)

a) es nobis] est mihi *B*

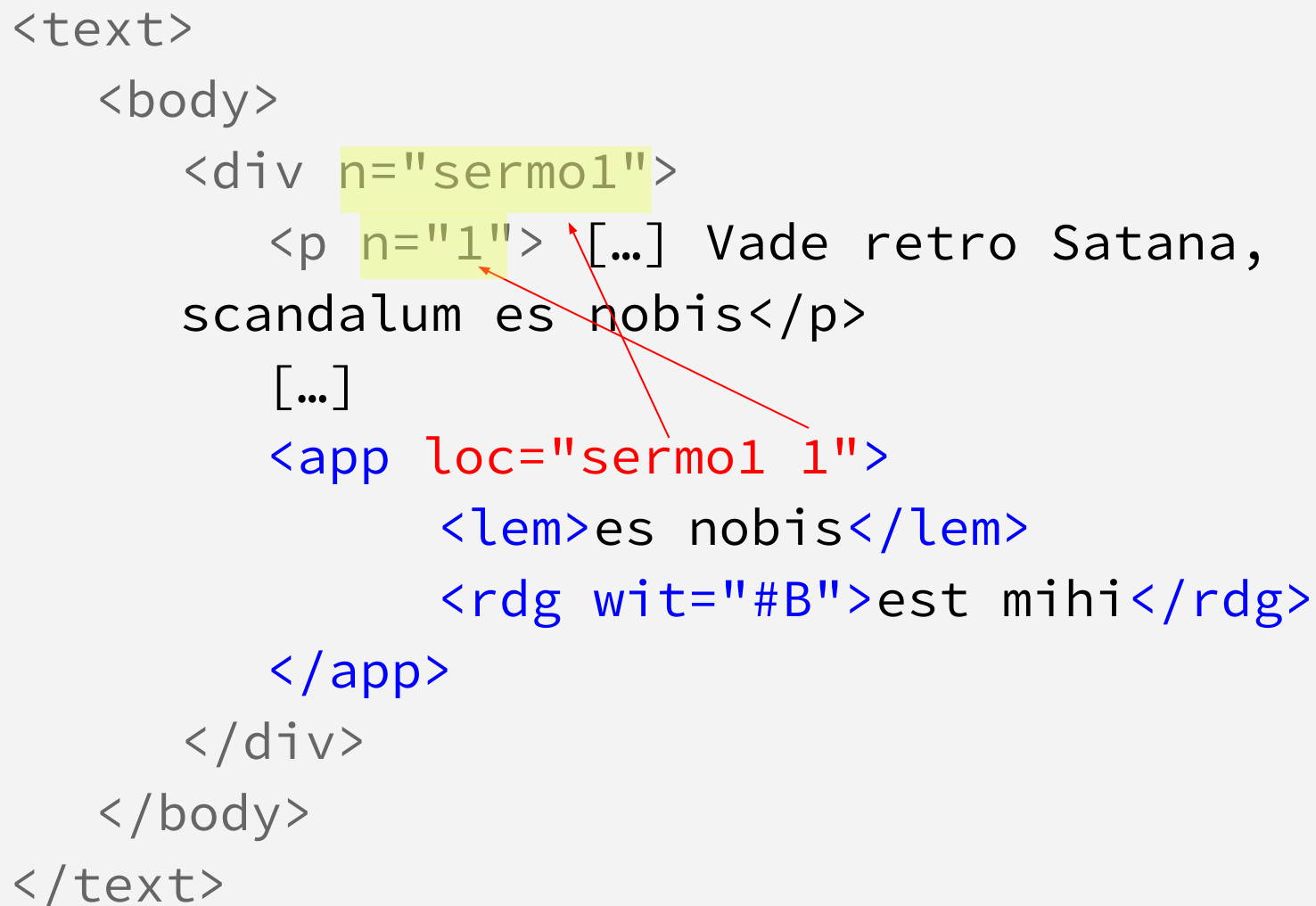
Location referenced

L'apparat est lié au texte de base à l'aide d'un renvoi au bloc de texte contenant la variante.

Le renvoi est noté par une référence formalisée ou un renvoi (“*A 137*” ou “*page 15 ligne 1*”)

Location referenced

```
<text>
  <body>
    <div n="sermo1">
      <p n="1"> [...] Vade retro Satana,
scandalum es nobis</p>
      [...]
      <app loc="sermo1 1">
        <lem>es nobis</lem>
        <rdg wit="#B">est mihi</rdg>
      </app>
    </div>
  </body>
</text>
```



The diagram illustrates location referencing in an XML/TEI structure. Two red arrows originate from the text 'sermo1 1' within the `<app loc="sermo1 1">` tag. One arrow points to the `n="1"` attribute of the `<p>` tag, and the other points to the text 'nobis' within the `<p>` tag, demonstrating how the location is resolved to a specific element and text node.

Double end-point attachement

L'apparat est lié au texte de base en indiquant explicitement le début et la fin du lemme dans le texte de base (avec les attributs **@from** et **@to**).

Les variantes peuvent se chevaucher.

Double end-point attachment

```
<text>
  <body>
    <div n="1">
      <p n="1"> [...] Vade retro Satana,
      scandalum <anchor xml:id="app1b"/>es nobis
      <anchor xml:id="apple"/>
      </p> [...]
    </div>
  </body>
</text>
```

```
<app from="#app1b" to="#apple">
  <rdg wit="#B">est mihi</rdg>
</app>
```

Parallel Segmentation

Tous les témoins sont encodés et mis sur le même plan : les notions de “texte de base” et de lemme peuvent devenir + ou - caduques.

Les variantes ne peuvent pas se chevaucher, mais elles peuvent s’imbriquer.

Parallel Segmentation

```
<text>
  <body>
    <div n="sermo1">
      <p n="1"> [...] Vade retro Satana,
      scandalum <app>
        <lem>es nobis</lem>
        <rdg wit="#B">est mihi</rdg>
      </app> [...]

    </p> </div>
  </body>
</text>
```

Quelle méthode choisir ?

- **Location-referenced**

- Pratique pour **encoder un appareil imprimé** tel qu'il est organisé.
- Reconstitution automatique du texte des différents témoins impossible.

- **Double-end-point attachement**

- Reconstitution automatique du texte des différents témoins possible.
- **Chevauchement entre lemmes possibles.**
- Difficile à encoder sans outils appropriés.

- **Parallel segmentation**

- **Reconstitution automatique du texte des différents témoins possible.**
- Pas de chevauchement de lemmes autorisés, mais on peut placer des variantes à l'intérieur de variantes.
- **Assez facile à encoder.**

Lemme contenant d'autres balises TEI

entrée d'apparat critique

```
<app type="editorial">  
<lem resp="BL 1.328">
```

leçon préférée

```
<choice>
```

```
<reg>Ἀφροδισιάδι</reg>
```

leçon préférée : version régularisée

```
<orig>Ἀφροδισιάτι</orig>
```

leçon préférée : version "diplomatique"

```
</choice>
```

```
</lem>
```

autre leçon

```
<rdg>Ἀφροδισιάτι</rdg>
```

```
</app>
```

fin entrée d'apparat critique

Exemple commenté

fichier simple-tei-edition > xml > valere.xml



**Critical
Apparatus
Toolbox**

<http://teicat.huma-num.fr>

Exercise 2



Corrigez un encodage fautif avec la toolbox

1. Allez sur <https://github.com/hisomademo/simple-tei-edition>
2. Placez-vous sur la **branche “gh-pages”**
3. Cliquez sur bouton vert **“clone or download”**
4. Choisissez de télécharger une archive zip
5. Enregistrez-le sur votre bureau
6. Dézippez l'archive
7. Trouvez le fichier **“valere_fautif.xml”** dans le dossier **“xml”**
8. Chargez le fichier dans la **toolbox** et cliquez sur **“check your encoding”**
9. Utilisez les menus de droite pour repérer les erreurs
10. Vous pouvez les corriger si vous le souhaitez à l'aide de **tei-web-editor**
11. N'oubliez pas de sauvegarder votre travail
12. Pour vérifier à nouveau à l'aide de la toolbox, reprenez à l'étape 1

suppléments



Quelques balises TEI - 1

<code><pb/></code>	saut de page
<code><lb/></code>	saut de ligne
<code><div></code>	division
<code><head></code>	titre
<code><p></code>	paragraphe
<code><opener></code>	formule de début
<code><closer></code>	formule de fin
<code><dateline></code>	mention de date
<code><signed></code>	signature

Quelques balises TEI - 2

<add>	ajout
	suppression
<subst>	substitution
<choice>	balisages alternatifs
<orig>	forme originale
<reg>	forme régularisée
<corr>	correction
<placeName>	nom de lieu
<persName>	nom de personne
<list>	liste (contient des items)
<item>	composant d'une liste
<label>	étiquette attachée à un item