



# Objectifs

- La structuration des données :
  - Pourquoi et comment structurer les données
- L'imprimé :
  - Structure d'un livre
  - Principes de balisage
- Encodage formel des différentes structures :
  - La prose
  - Le théâtre
  - La poésie

# Structurer des données

- C'est convertir des données (textuelles) dans un format exploitable par un ordinateur
- Dans le contexte des documents digitalisés, c'est :
  - Utiliser un langage formel de marquage pour décrire un texte sous la forme d'une suite de phénomènes textuels et paratextuels qualifiant et **structurant** le texte.
- Le document numérique est alors exploitable par l'humain ET par la machine.
- Exemple d'un format compris par la machine :  
HTML : site du CESR !

# Histoire et informatique

Pourquoi structurer les données ?

Pourquoi ne pas utiliser une base de données ?

Comment structurer ces données ?

# Pourquoi structurer en histoire ?

- Pourquoi encoder ? Quel est le but ultime ?
  - Structurer pour structurer
    - Classer comme Buffon
    - Pour une aide intellectuelle, une meilleure compréhension : structurer la pensée
  - Pour éditer les textes
    - Comment éditer en ligne ?
  - Pour analyser les textes
    - Quels sont les outils pour explorer les données ?

# Les inconvénients des bases de données

Voir <http://www.lespetitescases.net/node/1088>

- Explosion de l'information en tables et en champs
  - *Atomisation du texte : décomposition*
- Reconstruction du texte par des requêtes
- Pas de fichier contenant l'ensemble du texte structuré
- *Pas d'édition incrémentale et cumulative*
- *Le retour au texte original est impossible*

# L' « ecdotique »

(Marjorie Burghart – EHESS Lyon)

- Science de l'édition de textes (rien de numérique)
- Les données textuelles sont compliquées :
  - Mots composés de caractères
  - Concepts
- Édition critique à l'apparat savant
  - Choix de l'éditeur pour l'édition des textes
  - Notes érudités qui accompagnent les textes



# Structure d'un document : le manuscrit

- Structure du registre de la Chancellerie
  - Description du registre :
    - Critique externe : fabrication du registre, état du registre
    - Critique interne : le contenu du registre (durée, différence des pièces...)
  - Dresser la typologie des actes : évocation, mandement, placet, répit, sauvegarde, acte exécutoire, lettre d'office, relief, lettre de pardon, commutation de peine, abolition.

 **LA DEF-**  
**FENCE, ET IL-**  
**LYSTRATION DE LA**  
*Langue Francoyse.*

Par I. D. B. A.



Imprimé à Paris pour Arnoul l'Angelier,  
tenât sa Bouticque au second pillier  
de la grand' sale du Palays.

1 5 4 9.

**AVEC PRIVILEGE.**

Un document imprimé

# Structure du document : le livre imprimé

- La page de titre
  - Le titre, l'auteur, l'éditeur, etc.
- avec les pièces liminaires
  - Privilège
  - Dédicace
  - Préface...
- Le texte
  - Le corps du texte
  - Un groupe de textes : recueils
- Les pièces annexes
  - Table des matières, postface, etc.

# Structure du document : le livre imprimé

- Le corps du texte
  - Les divisions avec leur en-tête ou titre (chapitre, livre, partie, etc.)
    - Les paragraphes
    - Les strophes...
  - La mise en valeur typographique : le rendu
    - Les notes
    - Les marges
    - Les tableaux et les listes
    - Les illustrations

# Structure du document

- Structure physique du document :
  - Ce n'est pas un texte linéaire
    - Feuillet ou la page
    - L'aspect physique : déchirure
    - Reliure

La structure physique interfère dans l'interprétation du texte

# Du document au texte

- Le document est composé :
  - De textes
  - D'images
  - De sons
- Le texte est composé de caractères qui forment des mots
- Pour structurer, il faut baliser !
  - Pour baliser, il faut utiliser des éléments et des attributs

# Principes de balisage

- **L'élément** est la structure de base marquée par une balise de début et une balise de fin
  - La balise ouvrante se marque avec des chevrons : `< . . . >` ; pour la balise fermante une barre oblique est ajoutée `</ . . . >` :
    - exemple `< . . . > . . . </ . . . >` ;

si le contenu est vide, les balises se raccourcissent :

- Exemple `< . . . / >`

soit une barre oblique APRES le contenu

- Exemple :

```
<title><lb/>Le Tiers Livre</title>
```

# Principes de balisage

- Deux règles essentielles dans XML :

- la syntaxe des balises
- Règle d'imbrication stricte : il est interdit de faire chevaucher les éléments, par exemple :

```
<title><p>Le Tiers Livre</title></p>
```

- Hierarchie : parents/enfant

- une certaine catégorie d'éléments est imbriquée dans une autre catégorie :

- `<addrLine>` est toujours imbriqué dans `<address>`
- `<date>` est parfois utilisé seul ou parfois imbriqué dans d'autres éléments comme `<bibl>`

# Principe de balisage

J-D. Fekete and N. Dufournaud. « Analyse historique de sources manuscrites : application de TEI à un corpus de lettres de rémission du XVIe siècle. » Document Numérique, numéro spécial “Les documents anciens”, 3(1-2):117-134, 1999.

- Balisage formel et typographique qui permet :
  - une transcription structurelle :
    - `<div> et <head>`
    - `<p>`
  - Un respect de la mise en forme :
    - `<lb/>`
    - `<fw>` pour *Forme Work* pour les en-tête et pied de page
- Balisage sémantique : éléments analytiques succincts
  - `<persName><date> <geogName><address>`

# Principe de balisage

- **L'attribut** est un paramètre associé à un élément
  - Il fournit une description de l'élément
    - Exemple : @n
    - `<pb n="1 recto"/>` 1 recto étant la valeur de l'attribut
  - Autres attributs courant :
    - @type
    - @rend
- Certains éléments sont également des attributs:
  - Exemple : `<label>` et @label

# Présentation du fichier

```
<teiHeader>
```

```
...
```

```
</teiHeader>
```

```
  <text>
```

```
    <front>...</front>
```

```
    <body>
```

```
      <div>
```

```
        ...
```

```
      </div>
```

```
    </body>
```

```
  </text>
```

# Du texte au document

- Le sonnet extrait de La Défense et illustration de la langue française de Du Bellay
  - Exercice Oxygen de Lou Burnard
- Le sonnet dans l'oeuvre de Du Bellay

A l'Ambicieux,  
ET AVARE ENNEMY  
DES BONNES LETTRES.

Sonnet.

*Serf de Faueur, Esclaué d'Auarice,  
Tu n'beus iamais sur toy mesmes pouuoir,  
Et ie me veux d'un tel Maître pouuoir,  
Que l'Esprit libre en plaisir se nourisse.  
L'Air, la Fortune, & l'humaine Police  
Ont en leurs Mains ton malheureux Auoir.  
Le Iuge auare icy n'a rien à voir.  
Ny les troys Seurs, ny du Tens la malice.  
Regarde donc qui est plus soubaitable  
L'ayse, ou l'ennuy, le certain, ou l'instable.  
Quand à l'honneur, i'espere estre immortel:  
Car un cler Nom soubz Mort iamais ne tombe.  
Le tien obscur ne te promet rien tel.  
Ainsi, tous deux serez soubz mesme T umbe.*

# ☰ Exercice sonnet (Lou Burnard)

- Utiliser les éléments et attributs suivants :
  - `<div>` *pour la structure générale*
  - `<head>` *pour l'en-tête*
  - `<lb/>` *pour marquer les lignes*
  - `<lg>` *pour le découpage en strophe*
  - `<l>` *pour la précision des vers*
  - `@rend="it"` *pour le rendu en italique*
  - `@rend="center"` *pour le centrage*

# Structure de La Défense et illustration de la langue française de Du Bellay

```
<text>
  <pb/>
  <front>... </front>
  <pb/>
  <body>
    <div n="1" type="book">
      <head>...</head>
      <div n="1" type="chapter">
        <head>...</head>
        <p>...</p>
        ...
      </div>
      <div n="2" type="chapter">...</div>
    </div>
    <div n="2" type="book">...</div>
  </body>
  <back>
    <div type="sonnet">...</div>
    <div type="postface">...</div>
  </back>
</text>
```

 **LA DEF-**  
**FENCE, ET IL-**  
**LUSTRATION DE LA**  
*Langue Francoyse.*

Par **I. D. B. A.**



Imprimé à Paris pour Arnoull'Angelier,  
tenât sa Boutique au second pillier  
de la grand' sale du Palays.

**1 5 4 9.**

**AVEC PRIVILEGE.**

# Méthodologie

## Exercice\_tp-3.xml : Du Bellay, *La deffence, et illustration de la langue francoyse.*

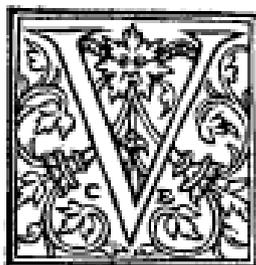
### Préface de l'auteur

- Les éléments et attributs
  - `<pb/>`
    - `@n`
  - `<fw>`
    - `@place="top-right top-left bot-center bot-right"`
    - `@rend="it"`
    - `@type="pageNum sig catch"`
  - `<hi>` *pour highlighted*
    - `@rend="larger it"`
  - `<closer>` et `<salute>`
    - `@rend="center"`

# A MONSEI-

GNEVR LE REVE-  
rendissime Cardinal  
du Bellay.

S.



**V**E le Personnage,  
que tu ioues au Specta-  
cle de toute l'Europe,  
noyre de tout le Mon-  
de en ce grand Thea-  
tre Romain, non tant  
d'affaires, & telz, que  
seul quasi tu soutiens  
l'Honneur du sacré Col-  
lege!

pecheroy-ie pas (comme dit le Pindare Latin) contre le bien publicq, si par longues paroles i'empeschoy le tens, que tu donnes au service de ton Prince, au profit de la Patrie, & à l'accroissement de ton immortelle renommée. Espiant doncques quelque heure de ce peu de relax, que tu prens pour respirer soubz le pesant faix des affaires Francoyses (charge uerayement digne de si robustes epaules, non moins que le Ciel de celles du grand Hercule) ma Muse a pris la hardiesse d'entrer au sacré Cabinet de tes saintes, & studieuses occupations: & la entre une

4 ij de

de riches, & excellens uoiez de iour en iour de-  
diez à l'image de ta grandeur, pendre le sien  
humble, & petit: mais toutesfois bien heureux  
s'il rencontre quelque faueur deuant les yeux  
de ta bonté, semblable à celle des Dieux im-  
mortelz, qui n'ont moins agreables les pauures  
presentz d'un bien riche uoloir, que ces su-  
perbes, & ambicieuses offrandes. C'est en ef-  
fect la Deffence & illustration de nostre Lan-  
gue Francoyse. A l'entreprise de laquelle rien  
ne me induyt, que l'affection naturelle enuers  
ma Patrie, & à te la dedier, que la grandeur  
de ton nom. Afin qu'elle se cache (comme  
soubz le Bouclier d'Alax) contre les traitz en-  
uenimez de ceste antique Ennemye de uertu,  
soubz l'ombre de tes esles. De toy dy-ie, dont  
l'incomparable Scauoir, Vertu, & conduyte  
toutes les plus grandes choses, de si long tens  
de tout le Monde sont experimentées, que ie  
ne les scauroy plus au nif exprimer, que les  
courant (suyuant la ruse de ce noble peintre  
Tymante) soubz le uoyle de silence. Pour ce,  
que d'une si grande chose il uault trop myeux  
(comme de Carthage disoit T. Liue) se taire  
du tout, que d'en dire peu. Recoy doncques a-  
uecques ceste accoutumée Bonté, qui ne te rend  
moins ayable entre les plus petiz, que ta Ver-  
tu, & Autorité uenerable entre les plus grâds,  
les premiers fruietz, ou pour myeux dire les

premieres fleurs du Priantens de celuy, qui en  
toute Reuerence, & Humilité bayse les mains  
de ta R.S. Priant le Ciel te departir autant de  
heureuse, & lögue uie, et à tes haultes entrepri-  
ses estre autât fauorable, comme enuers toy il a  
esté liberal, noyre prodigue de ses Graces.

A Dieu, De Paris ce. 15. de  
Feurier. 1549.

**L'**auteur pryé les Lecteurs differer leur  
iugemēt iusques à la fin du Liure, & ne le cō-  
damner sans auoir premierement bien veu, &  
examiné ses raisons.

à iii

# Correction de l'exercice\_tp- 3.xml

```
321 <fw place="top-right" type="pageNum">[2]</fw>
322 <div type="preface">
323   <head rend="center">
324     <lb/>A MONSEI-
325     <lb/>LE REVE-
326     <lb/>rendissime Cardinal
327     <lb/><persName>du Bellay</persName>.
328     <lb/>S.
329   </head>
330 <p>
331   <lb/><hi rend="larger">V</hi>EU <hi rend="it">le Personnage,
332   <lb/>que tu joues au Spec-
333   <lb/>tacle de toute l'<geogName>Europe</geogName>,
334   <lb/>voyre de tout le Mon-
335   <lb/>de en ce grand Thea-
336   <lb/>tre Romain, veu tant
337   <lb/>d'affaires, & telz, que
338   <lb/>seul quasi tu soutiens: ô
339   <lb/>l'Honneur du sacré Col-
340   <lb/>lege! pecheroy'-je pas (comme dit le <name>Pindare</name>
341   <lb/>Latin) contre le bien publicq', si par longues
342   <lb/>paroles j'empeschoy' le tens, que tu donnes au
343   <lb/>service de ton Prince, au profit de la Patrie, &
344   <lb/>à l'accroissement de ton immortelle renommée?
345   <lb/>Epiant donques quelque heure de ce peu de re
346   <lb/>laiz, que tu prens pour respirer soubz le pesant
347   <lb/>faiz des affaires Francoyses (charge vraiment
348   <lb/>digne de si robustes epaules, non moins que le
349   <lb/>Ciel de celles du grand <name>Hercule</name>) ma Muse a pris
350   <lb/>la hardiesse d'entrer au sacré Cabinet de tes sain-
351   <lb/>ctes, & studieuses occupations: & la entre tant
352   <fw rend="it" place="bot-center" type="sig">a ij</fw>
353   <fw rend="it" place="bot-right" type="catch">de</fw>
354   <pb n="" />
355   <fw place="top-left" type="pageNum">[2v]</fw>
356   <lb/>de riches & excellens voeuz de jour en jour de-
357   <lb/>diez à l'apos;image de ta grandeur, pendre le sien
358   <lb/>humble, & petit: mais toutesfois bien heureux
359   <lb/>s'apos;il rencontre quelque faveur devant les yeux
360   <lb/>de ta bonté, semblable à celle des Dieux im-
361   <lb/>mortelz, qui n'apos;ont moins agreables les pauvres
362   <lb/>presentz d'apos;un bien riche vouloir, que ces su-
363   <lb/>perbes, & ambicieuses offrandes. C'apos;est en ef-
364   <lb/>fect la Deffence & Illustration de notre Lan-
365   <lb/>gue| Francoyse. A l'apos;entreprise de laquelle rien
366   <lb/>ne m'apos;a induyt, que l'apos;affection naturelle envers
```

# Exercice\_tp-4.xml : autre sonnet de Du Bellay

- Balisage :
  - Logique
    - <div>
    - <head>
    - <lg>
    - <|>
  - Formel
    - <fw>
    - <hi>

À TRES ILLUSTRE.

Princesse Madame Marguerite

Seur Vnique du Roy,

Luy presentant ce liure.

SONNET.

PAR Vn sentier inconnu à mes yeux  
Vostre grandeur sur ses ailes me porte  
Ou de Phebus la main scauant & forte  
Guide le frein du chariot des cieulx.  
La eleué au cercle radieux  
Par Vn Demon beureux, qui me conforte,  
Celle fureur tant douce i' en rapporte,  
Dont vostre nom i' egally aux plus haultx dieux.  
O Vierge donc, sous qui la Vierge Astrée  
A fait en cor en nostre sieclx entrée!  
Prenez en gré ces poëtiques fleurs.  
Ce sont mes vers, que les chastes Carites  
Ont emaillez de plus de cent couleurs  
Pour aler voir la fleur des MARGVERITES.

COELO MVSA BEAT.

Sonnet tiré de l'Olive de  
1550

exercice\_tp-4.xml

```

279 <pb n="2v" />
280 <div type="epistle">
281 <fw place="top-left" type="pageNum">[2v]</fw>
282 <head rend="center it">
283 <lb/>A <persName>TRES ILLUSTRÉ.
284 <lb/>Princesse Madame Marguerite
285 <lb/>Seur Unique du Roy</persName>,
286 <lb/>Luy presentant ce livre.
287 </head>
288 <div type="sonnet">
289 <head rend="center">
290 <lb/>SONNET.
291 </head>
292 <lg>
293 <l>PAR <hi rend="it">un sentier inconnu à mes yeux</hi></l>
294 <l><hi rend="it">Vostre grandeur sur ses ailes me porte</hi></l>
295 <l><hi rend="it">Ou de <name>Phebus</name> la main scavante et forte</hi></l>
296 <l><hi rend="it">Guide le frein du chariot des cieulx.</hi></l>
297 </lg>
298 <lg rend="it">
299 <l>Là élevé au cercle radieux</l>
300 <l>Par un Demon heureux, qui me conforte,</l>
301 <l>Celle fureur tant douce j'&apos;en rapporte,</l>
302 <l>Dont vostre nom j'&apos;egalle aux plus haulx dieux.</l>
303 </lg>
304 <lg rend="it">
305 <l>O Vierge donc, sous qui la Vierge Astrée</l>
306 <l>A faict encor&apos; en nostre siecle entrée!</l>
307 <l>Prenez en gré ces poëtiques fleurs.</l>
308 </lg>
309 <lg>
310 <l><hi rend="it">Ce sont mes vers, que les chastes Carites</hi></l>
311 <l><hi rend="it">Ont emaillez de plus de cent couleurs</hi></l>
312 <l><hi rend="it">Pour aler voir la fleur des</hi> MARGUERITES.</l>
313 </lg>
314 <closer rend="center">
315 <signed>
316 <lb/><foreign xml:lang="lat">COELO MVSA BEAT</foreign>.
317 </signed>
318 </closer>
319 </div>
320 </div>

```

105  
104  
103  
102  
101  
100  
99  
98  
97  
96  
95  
94  
93  
92  
91  
90  
89  
88  
87  
86  
85  
84  
83  
82  
81  
80  
79  
78  
77  
76  
75  
74  
73  
72  
71  
70  
69  
68  
67  
66  
65  
64  
63  
62  
61  
60  
59  
58  
57  
56  
55  
54  
53  
52  
51  
50  
49  
48  
47  
46  
45  
44  
43  
42  
41  
40  
39  
38  
37  
36  
35  
34  
33  
32  
31  
30  
29  
28  
27  
26  
25  
24  
23  
22  
21  
20  
19  
18  
17  
16  
15  
14  
13  
12  
11  
10  
9  
8  
7  
6  
5  
4  
3  
2  
1

Monsieur, ie vous ai escript vers la S. Martin, et  
envoie nostre arpenterie, et ensemble a Monsieur  
Loysel, pour vous faire entendre, que se i'avois quelq  
chose de bon, ie vous en ferois volontiers part. Je  
vous ai aussi escript, qu'on avoit imprimé quelqz nouveaux  
glozeurs Alemans avecque Foquelm<sup>us</sup>, sur Perse: desquelz  
vous vous pouvez passer fort bien le livre de Monsieur  
de la Scala, n'a pas est si tost arrive, qu'on la eut eue.  
Monsieur de Candale, qui a Pjem est en Medouc, a  
eu nouvelles de son livre, escript au pape, touchant  
la reformation des Equivoques. &c. Le Cardinal Borro-  
meo lui a escript, que Clavius Jesuite, qui fait a  
Rome grand estat et profession des Mathematiques.  
(il a ~~gloze~~ gloze ~~sur~~ Euclides, escript de hono-  
reurs plus que Orontius ni Montanus, et fait com-  
mentaires sur le petit livre de Desacroboseo, plus long  
qu'un iour sans pain). que ce bon seigneur, fromont a  
dire quelq chose (il ne <sup>par</sup> fai que <sup>par</sup> oindire. Car ie n'ai veu le  
dit seigneur depuis la dite response receu de Rome)  
au discours de Monsieur de Candale: et dit M<sup>onsieur</sup> de  
Candale, que ~~il~~ tout i est: mais quil n'a leu tout  
son livre, ou sil l'a leu, quil ~~ne~~ ne la entendu. Il  
est apres a droiser sa response apud Medulos.  
Mais quil soit de retour, ie tendrai le tout, et en sau-  
rez, q ce qu'en ~~apud~~ avrai appris. Je n'ai autre cho-  
se, de q nouveau. Si monsieur Loysel voit les p-  
sesces, il pourroit entendre, que M<sup>onsieur</sup> de Montague  
a receu son livre. Il est en cete ville des Dimanche  
dernier. Je prie Dieu, M<sup>onsieur</sup>, quil vous ait ~~so~~  
en sa garde. De Bourd, ce 26 de Nouemb. 1583.

☉ Vostre humble serviteur  
Elie Vinet.

Lettre  
d'Elie  
Vinet

# La lettre d'Elie Vinet

## exercice\_tp-2.xml

- Utiliser les éléments et attributs suivants :
  - <div>
  - <opener> *pour marquer le début de la lettre*
  - <ab> *pour préciser le bloc de texte*
  - <lb/>
  - <closer> *pour marquer la fin de la lettre*
    - <salute>
    - <signed>
  - @rend="center"
  - @n="" *pour préciser les numéros de ligne*

# Le théâtre

- Les éléments
  - **<castList>** *pour la distribution*
    - **<castGroup>** *liste de personnages*
    - **<castItem>** *personnage*
      - **<role>** *description d'un rôle*
  - **<sp>** *monologue*
    - **<speaker>** *locuteur*
  - **<stage>** *pour l'indication scénique*

# Discours fantastiques de Justin Tonnelier de Florence. Exemple\_tp-5.xml

- Structure du texte de Justin Tonnelier

```
<fw>...</fw>

<div>

<head>...</head>

<castList>

  <castGroup>

    <head>...</head>

    <castItem><role>...</role></castItem>

  </castGroup>

</castList>

<div>

  <head>...</head>

  <sp>

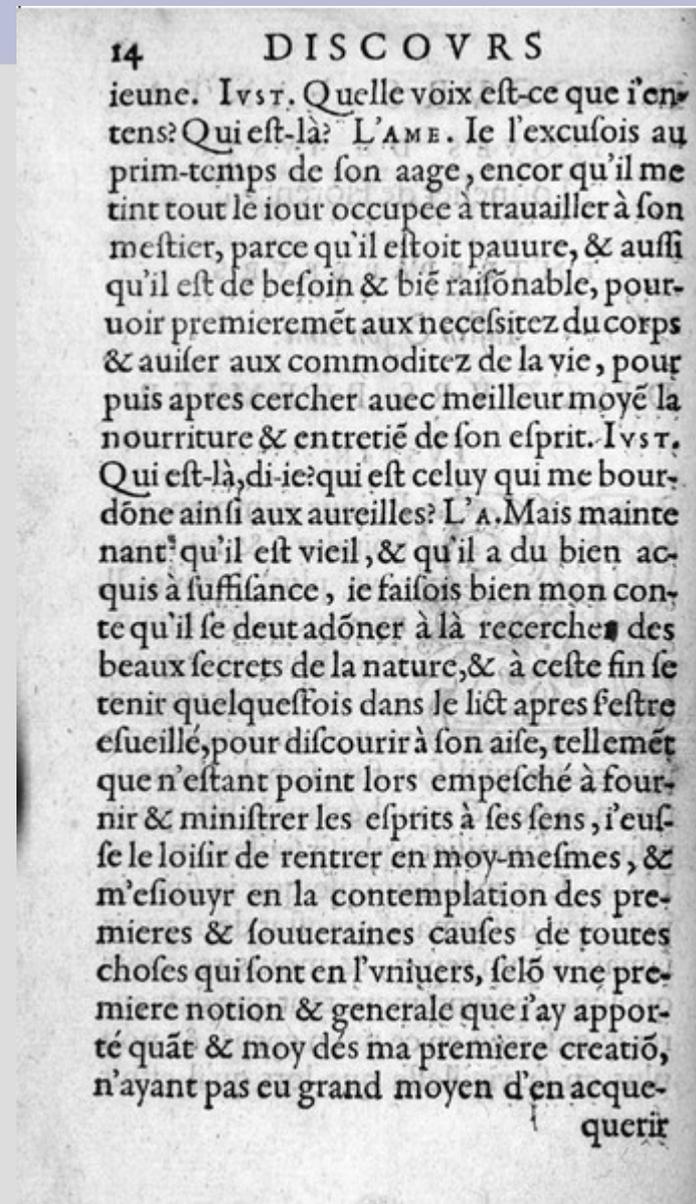
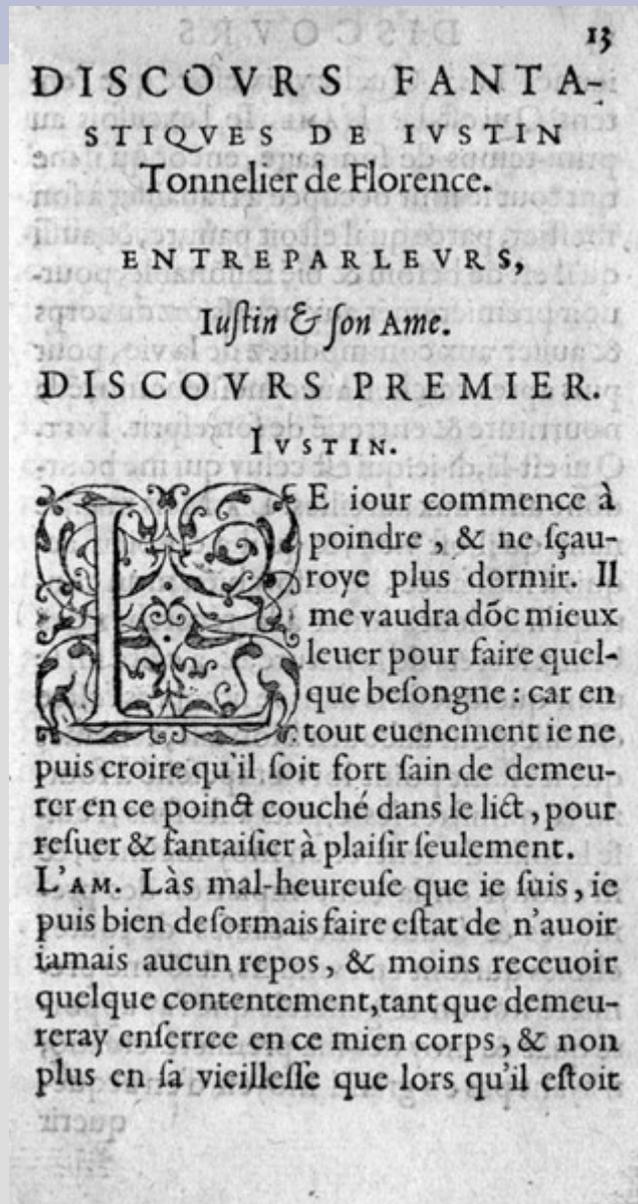
    <speaker>...</speaker>

    <ab>...</ab>

  </sp>

</div>
```

# Exercice\_tp-5.xml : Discours fantastiques de Justin Tonnelier de Florence.



```
33 <fw place="top-right" type="pageNum">13</fw>
34 <head rend="center">
35 <lb/>DISCOURS FANTA-
36 <lb/>STIQUES DE JUSTIN
37 <lb/>Tonnellier de Florence.
38 </head>
39 <castList rend="center">
40 <castGroup>
41 <head><lb/>ENTREPARLEURS,</head>
42 <castItem><lb/><role xml:id="justin"><hi rend="it">Justin</hi></role> <hi rend="it">&amp;</hi></castItem>
43 <castItem><role xml:id="ame"><hi rend="it">son Ame.</hi></role></castItem>
44 </castGroup>
45 </castList>
46 <div type="discours" n="1">
47 <head rend="center"><lb/>DISCOURS PREMIER.</head>
48 <sp who="#justin">
49 <speaker><lb/>JUSTIN.</speaker>
50 <ab>
51 <lb/>Le jour commence à
52 <lb/>poindre, &amp; ne scau-
53 <lb/>roye plus dormir. Il
54 <lb/>me vaudra donc mieux
55 <lb/>lever pour faire quel-
56 <lb/>que besongne: car en
57 <lb/>tout evenement je ne
58 <lb/>puis croire qu'il soit fort sain de demeu-
59 <lb/>rer en ce poinct couché dans le lict, pour
60 <lb/>resver &amp; fantaisier à plaisir seulement.
61 </ab>
62 </sp>
63 <sp who="#ame">
64 <speaker><lb/><abbr>L'AM.</abbr></speaker>
65 <ab>
66 Làs mal-heureuse que je suis, je
67 <lb/>puis bien desormais faire estat de n'avoir
68 <lb/>jamais aucun repos, &amp; moins recevoir
69 <lb/>quelque contentement, tant que demeu-
70 <lb/>rera enserree en ce mien corps, &amp; non
71 <lb/>plus en sa vieillesse que lors qu'il estoit
72 <pb/>
73 <fw place="top-left" type="pageNum">14</fw>
74 <fw place="top-center" type="title">DISCOURS</fw>
75 <lb/>jeune.
76 </ab>
77 </sp>
78 <sp who="#justin">
79 <speaker><lb/><abbr>JUST.</abbr></speaker>
80 <ab>
81 Quelle voix est-ce que j'en-
```