

Cahier des spécifications techniques

I. SPÉCIFICATIONS GÉNÉRALES

I.1. OBJET

- Conversion texte (OCR ou saisie) et structuration de 195 fichiers image (format PDF).

I.2. VOLUME GLOBAL

- 195 fichiers PDF, représentant 42 388 pages et 70 975 300 signes hors balisage (estimation). Textes correspondant soit à des volumes entiers (livres), soit à des extraits de volumes (articles, chapitres, etc.). Liste des textes jointe (avec indication des lots et des tranches).

I.3. LOTS

- 3 lots en fonction des spécifications techniques et de la qualité des documents.

Lot 1

- Volume : 147 textes, représentant 26 882 pages et 37 194 000 signes hors balisage (estimation). Textes des 16^e-18^e siècles ou textes du 19^e avec qualité PDF médiocre.
- Qualité requise : 99,9 % (avec possibilité d'indiquer par un balisage les manques et doutes).
- Mode d'acquisition : saisie (par défaut). OCR possible, mais uniquement si l'objectif qualité est atteint.
- Format du livrable : XML-TEI : spécifications indiquées (voir *infra*), schéma et css fournis. Voir *infra* pour les spécifications.
- Langues : français 15^e-16^e siècles : 5 textes (28 pages), français 17^e-18^e siècles : 84 textes (12 626 pages), français 19^e-20^e siècles : 29 textes (725 pages), latin/français-latin : 7 textes (1 913 pages), latin-grec : 1 texte (7 pages), italien : 19 textes (5 627 pages).

- Particularités : balisage indiqué en marge du texte si nécessaire (niveaux de titres, titres non hiérarchiques, citations, vers, notes tournantes, etc.).

Lot 2

- Volume : 10 textes, représentant 3 108 pages et 6 073 500 signes (estimation). Textes des 15^e-16^e siècles. Qualité des PDF bonne (numérisation récente).
- Qualité requise : 99 % (avec possibilité d'indiquer par un balisage les manques et doutes).
- Mode d'acquisition : saisie.
- Format du livrable : traitement de texte structuré (titres, blocs, raccrochage des notes, etc.). Respect de la pagination et maintien des sauts de ligne. Voir *infra* pour les spécifications.
- Langues : français 15^e-16^e siècles : 9 textes (2 873 pages), latin : 1 texte (235 pages).
- Particularités : présence de nombreux caractères spéciaux (abréviations, sigles, caractères gothiques, etc.). Tableau de l'encodage requis fourni avec chaque fichier.

Lot 3

- Volume : 38 textes, représentant 12 398 pages et 27 707 500 signes (estimation). Textes contemporains (19^e-20^e siècles). Qualité des PDF correcte.
- Mode d'acquisition : OCR avec correction des mots en doute et maintien des enrichissements (surtout l'italique).
- Qualité requise : 99,5 %.
- Format du livrable : traitement de texte (odt). Document structuré (titres, nature des blocs, insertion des notes, etc.). Voir *infra* pour les spécifications.
- Langues : français 19^e-20^e siècles : 29 textes (9 955 pages), français-latin : 1 texte (435 pages), italien : 7 textes (3 453 pages), anglais : 1 texte (557 pages).

I.4. MISE À DISPOSITION DES PDF

- Évaluation : fichiers PDF bruts disponibles sur Dropbox. Lien fourni sur demande.

- Production : fichiers PDF préparés et métadonnées complètes fournis sur Dropbox. Lien fourni sur demande. Possibilité d'échanges par FTP.

I.5. ENVELOPPE BUDGÉTAIRE MAXIMALE

Lot	Budget (€ HT)
Lot 1	66 000
Lot 2	12 000
Lot 3	25 000

I.6. CALENDRIER DE LIVRAISON

Lot 1

- 6 tranches mensuelles, représentant chacune environ 4 500 pages PDF.

Tranche	Date de livraison	Paiement (€ HT)
Lot 1-A	15/11/2018	10815
Lot 1-B	15/12/2018	11041
Lot 1-C	15/01/2018	10889
Lot 1-D	15/02/2018	11483
Lot 1-E	15/03/2018	11100
Lot 1-F	15/04/2018	10672

Lot 2

- 3 tranches mensuelles, représentant chacune environ 1 000 pages PDF.

Tranche	Date de livraison	Paiement (€ HT)
Lot 2-A	15/11/2018	3629
Lot 2-B	15/12/2018	3765
Lot 2-C	15/01/2018	4606

Lot 3

- 2 tranches mensuelles, représentant chacune environ 7 000 pages PDF.

Tranche	Date de livraison	Paiement (€ HT)
Lot 3-A	15/11/2018	12240
Lot 3-B	15/12/2018	12760

II. SPÉCIFICATIONS PAR LOT

II.2. LOT 1

II.2.A. Format du livrable (indications générales)

- Fichier XML-TEI, conforme à la restriction Teinte appliquée par le LABEX OBVIL (voir ci-dessous, Documentation).
- Respect des graphies originales. Trois exceptions :
 - 1) le s long (ſ), à transcrire = s ;
 - 2) l’esperluette (&) à transcrire = & ;
 - 3) les guillemets droits, à transcrire en guillemets anglais doubles (“ ”).
- Pas de dissimilation u/v, i/j.
- Rétablissement du flux du texte : texte en blocs de niveau paragraphe ; ne pas tenir compte des sauts de ligne (sauf indication contraire), ne pas tenir compte des césures, ne pas tenir compte des retours dans les vers longs.
- Notes au point d’appel. Rétablissement du flux du texte en cas de note tournante (sur plusieurs pages).

- Indication des numéros de page. Balise <pb/> en début de page, format <pb n="i"/>, <pb n="ii"/>, <pb n="1"/>, <pb n="2"/>, etc.
- Ne pas transcrire =
 - 1) La page de titre du document (sauf indication contraire). Les informations seront transcrites dans le TEI-Header.
 - 2) Les titres courants.
 - 3) Les signatures (numéros de cahiers) et les réclames (premier mot ou début du premier mot de la page suivante) sous le texte.
 - 4) Les guillemets répétés en début de ligne dans les citations (lignes guillemetées).
- Pour plus de précisions, voir la section Balisage ci-dessous.

II.2.B. Documentation disponible

Schéma : teinte.rng.

Téléchargeable à cette adresse : <https://github.com/OBVIL/Teinte/blob/gh-pages/teinte.rng>.

CSS : tei2html.css.

Téléchargeable à cette adresse : <https://github.com/OBVIL/Teinte/blob/gh-pages/tei2html.css>.

XSLT : tei2html.xsl.

Téléchargeable à cette adresse : <https://github.com/OBVIL/Teinte/blob/gh-pages/tei2html.xsl>.

Description des éléments : <https://obvil.github.io/Teinte/teinte.html>.

II.2.C. TEI-Header (en-tête de document)

Un TEI-Header sera fourni avec chaque PDF, au format XML.

Exemple :

```
<?xml version="1.0" encoding="UTF-8"?>
<?xml-model href="http://obvil.github.io/Teinte/teinte.rng" type="application/xml"
schematypens="http://relaxng.org/ns/structure/1.0"?>
<?xml-stylesheet type="text/css" href="../../Teinte/opentei.css"?>
<?xml-stylesheet type="text/xsl" href="../../Teinte/tei2html.xsl"?>
<!--Odette: 2018-04-09T14:03:31+02:00-->
<TEI xmlns="http://www.tei-c.org/ns/1.0" xml:lang="fr">
  <teiHeader>
    <fileDesc>
      <titleStmt>
```

```

<title>Sur les gens de lettres (Considérations sur les mœurs de ce siècle)</title>
<author key="Duclos, Charles (1704-1772)">Charles Duclos</author>
</titleStmt>
<editionStmt>
  <edition>OBVIL</edition>
  <respStmt>
    <name></name>
    <resp></resp>
  </respStmt>
</editionStmt>
<publicationStmt>
  <publisher>Sorbonne Université, LABEX OBVIL</publisher>
  <date when="2018"/>

  <idno>http://obvil.paris-sorbonne.fr/corpus/critique/duclos_gens-de-lettres_1751/</idno>
  <availability status="restricted">
    <licence target="http://creativecommons.org/licenses/by-nc-nd/3.0/fr/"><p>Copyright ©
2018 Sorbonne Université, agissant pour le Laboratoire d'Excellence « Observatoire de la vie
littéraire » (ci-après dénommé OBVIL).</p><p>Cette ressource électronique protégée par le code
de la propriété intellectuelle sur les bases de données (L341-1) est mise à disposition de la
communauté scientifique internationale par l'OBVIL, selon les termes de la licence Creative
Commons : « Attribution - Pas d'Utilisation Commerciale - Pas de Modification 3.0 France (CC BY-
NC-ND 3.0 FR) ».</p><p>Attribution : afin de référencer la source, toute utilisation ou publication
dérivée de cette ressource électroniques comportera le nom de l'OBVIL et surtout l'adresse
Internet de la ressource.</p><p>Pas d'Utilisation Commerciale : dans l'intérêt de la communauté
scientifique, toute utilisation commerciale est interdite.</p><p>Pas de Modification : l'OBVIL
s'engage à améliorer et à corriger cette ressource électronique, notamment en intégrant toutes
les contributions extérieures. La diffusion de versions modifiées de cette ressource n'est pas
souhaitable.</p></licence>
  </availability>
</publicationStmt>
<sourceDesc>
  <bibl>Charles <author>Duclos</author>, <title>« Sur les gens de lettres »</title>, in <hi
rend="i">Considérations sur les mœurs de ce siècle</hi>, <pubPlace>Paris</pubPlace>,
<publisher>Pault</publisher>, <date>1751</date>, <biblScope>p. 243-269</biblScope> PDF :
<ref target="http://gallica.bnf.fr/ark:/12148/bpt6k9636372g">Gallica</ref>.</bibl>
  </sourceDesc>
</fileDesc>
<profileDesc>
  <creation>
    <date when="1751"/>
  </creation>
  <langUsage>
    <language ident="fre"/>
  </langUsage>

```

```
</profileDesc>  
</teiHeader>
```

II.2.D. Balisage

Balises XML-TEI à utiliser. La documentation complète sur les éléments autorisés est disponible à l'adresse : <https://obvil.github.io/Teinte/teinte.html>.

Les balises obligatoires ou très fréquentes sont précédées d'un point d'exclamation.

II.2.D.1. Problèmes de saisie

! <seg>

Balisage des doutes sur la saisie : mot difficilement lisible ou ambiguïté.

Exemple :

Dans un texte ancien, en présence de la graphie **fine**, difficulté de distinguer entre le s long (l) et le f = sine ?
fine ?
→ <seg>fine</seg>

En cas de passage manquant ou illisible, utiliser plutôt <gap>.

La balise <seg> peut également être utilisée pour un encodage provisoire, par exemple pour un système d'identifiants et de renvois. L'information à entrer est alors indiquée directement dans le PDF.

- Documentation : https://obvil.github.io/Teinte/teinte.html#el_seg.

! <gap>

Texte non transcrit. Indiquer obligatoirement une raison et un ordre de grandeur.

Le contenu étant destiné aux seuls éditeurs, il n'est pas nécessaire de le formaliser outre mesure, juste d'être clair.

Exemples :

- 1) Page manquante = <gap>mq page 148 (lacune PDF)</gap>.
- 2) Deux lignes illisibles = <gap>mq 2 lignes (illisible)</gap>.
- 3) Mot manquant = <gap>mq 1 mot (lacune PDF)</gap>

- Documentation : https://obvil.github.io/Teinte/teinte.html#el_gap.

II.2.D.2. Structure générale

! <div1>, <div2>, <div3>, etc.

Section du texte, hiérarchie réursive.

Les divisions supportent divers attributs dont l'attribut @type. Par exemple : <div1 type="preface">, <div1 type="chapter">, <div2 type="section">, <div1 type="act">, <div2 type="scene">. Par défaut, pas d'attribut, sauf indication précise en ce sens dans le PDF.

- Documentation : https://obvil.github.io/Teinte/teinte.html#el_div.

II.2.D.3. Entrées et sorties de section

! <head>

Titre, niveau déterminé par le contexte hiérarchique.

N.-B. : attribut @type possible. L'élément <head> n'est pas typé par défaut, mais il peut s'avérer nécessaire de préciser son type, avec trois valeurs possibles : "sur" (surtitre), "main" (titre proprement dit), "sub" (sous-titre).

Exemples :

```
<head>Chapitre II</head>
```

```
<head>Titre du chapitre</head>
```

```
<head type="sur">Chapitre II</head>  
<head type="main">Titre du chapitre</head>
```

```
<head type="sur">Chapitre II</head>  
<head type="main">Titre du chapitre</head>  
<head type="sub">Sous-titre du chapitre</head>
```

```
<head type="main">Titre du chapitre</head>  
<head type="sub">Sous-titre du chapitre</head>
```

- Documentation : https://obvil.github.io/Teinte/teinte.html#el_head.

<argument>

Résumé d'une section, contient des blocs. Il s'agit généralement d'un résumé en début de partie ou de chapitre. Extrêmement rare, indiqué le cas échéant dans le PDF.

Exemple :

II

Projet de journal pour les *jeunes*. — Les autographes de V. Hugo. — Adelphe Froger, la *République des Lettres*. — Le *Sherry-Cobbler*. — Quelques chansons. — Les *Vivants*.

```
<div1>
<head>II</head>
<argument>Projet de journal pour les <hi rend="i">jeunes</hi>. — Les autographes de V. Hugo. — Adelphe
Froger, la <hi rend="i">République des lettres</hi>. — Le <hi rend="i">Sherry-Cobbler</hi>. — Quelques
chansons. — Les <hi rend="i">Vivants</hi>.</argument>
```

- Documentation : https://obvil.github.io/Teinte/teinte.html#el_argument.

! <epigraph>

Épigraphie, citation en exergue de section. Indiqué le cas échéant dans le PDF.

*Semper ego auditor tantum, nunquamne
reponam
Vexatus toties rauci Theseide Codri ?
Juven. Sat. I.*

```
<epigraph>
<l rend="i">Semper ego auditor tantum, nunquamne reponam</l>
<l rend="i">Vexatus toties rauci Theseide Codri ?</l>
<bibl>Juven. Sat. <hi rend="sc">i</hi>.</bibl>
</epigraph>
```

- Documentation : https://obvil.github.io/Teinte/teinte.html#el_epigraph.

<byline>

Mention de responsabilité (dans un article de presse, une lettre...). Peu fréquent, équivaut à l'élément <signed> (signature) utilisé essentiellement pour les auteurs des préfaces, introductions, etc. autres que l'auteur principal du texte. Indiqué le cas échéant dans le PDF.

- Documentation : https://obvil.github.io/Teinte/teinte.html#el_byline

<salute>

Dédicace courte (en début d'ouvrage : de type « À mon ami Frédéric Dupont »). Peu fréquent, indiqué le cas échéant dans le PDF.

Salutation (dans une lettre). Peu fréquent, indiqué le cas échéant dans le PDF.

Exemple de salutation :

MONSEIGNEUR,

*Les Beaux Arts ne furent jamais
ingrats envers leurs bienfaiteurs. Vous*

<salute><hi rend="sc">Monseigneur</hi>,</salute>
<p rend="i"><hi rend="sc">Les</hi> Beaux Arts ne furent jamais ingrats envers leurs bienfaiteurs. Vous
...

- Documentation : https://obvil.github.io/Teinte/teinte.html#el_salute.

<dateline>

Mention de date, de temps ou de lieu, ex : donné le ... à ... (« datum », dans une lettre, une charte...).

Peu fréquent, ne baliser qu'en cas d'indication en ce sens dans le PDF.

- Documentation : https://obvil.github.io/Teinte/teinte.html#el_dateline.

<signed>

Signature. Peu fréquent, indiqué le cas échéant dans le PDF.

- Documentation : https://obvil.github.io/Teinte/teinte.html#el_signed.

! <trailer>

Titre de fin (livre, chapitre, page, acte, scène...). En fin de section, généralement centré : FIN ; FIN DE L'ACTE III, FIN DU LIVRE PREMIER, etc.

- Documentation : https://obvil.github.io/Teinte/teinte.html#el_trailer.

II.2.D.4. Blocs

! <p>

Paragraphe de texte.

Documentation : https://obvil.github.io/Teinte/teinte.html#el_p.

N.-B. : en cas de mise en forme particulière concernant un paragraphe entier, il est possible d'utiliser l'attribut @rend.

Exemples :

- 1) paragraphe centré : <p rend="center"></p>
 - 2) paragraphe au fer à droite : <p rend="right"></p>
 - 3) paragraphe avec retrait gauche : <p rend="margin"></p>
 - 4) paragraphe composé en sommaire (indentation négative) : <p rend="hanging"></p>
 - 5) paragraphe non indenté : <p rend="noindent"></p>
 - 6) paragraphe entièrement en italique : <p rend="i"></p>

7) paragraphe entièrement en petites capitales : `<p rend="sc"></p>`

! `<p/>`

Paragraphe vide (ligne de blanc). Uniquement si la ligne de blanc sépare deux paragraphes de prose. Les espaces avant et après les titres, les citations, les ornements, les illustrations, les répliques de théâtre, les strophes, sont générés automatiquement par la CSS.

Documentation : https://obvil.github.io/Teinte/teinte.html#el_p.

! `<l>`

Un vers, généralement contenu dans une strophe `<lg>`.

Règles particulières :

1) L'indentation des vers doit être marquée à l'aide d'espaces insécables (sauf dans le cas des vers coupés au théâtre, voir ci-dessous).

	Le sort plein d'injustice M'ayant, enfin, rendu Ce reste un pur supplice, Je ferais plus heureux si j'avais tout perdu.
<code><l></code>	Le sort plein d'injustice <code></l></code>
<code><l></code>	M'ayant enfin, rendu <code></l></code>
<code><l></code>	Ce reste un pur supplice, <code></l></code>
<code><l></code>	Je serais plus heureux si j'avais tout perdu. <code></l></code>

2) Les vers longs peuvent être reportés à la ligne suivante, à gauche ou à droite, avec une indentation. Rétablir l'unité du bloc et ne pas inscrire de saut de ligne.

Exemples :

	<i>non che farmi Cieco su' miei stessi capricci , ardisco Contro de' vizii miei darmi bat- taglia ,</i>
<code><l rend="i"></code>	non che farmi <code></l></code>
<code><l rend="i"></code>	Cieco su' miei stessi capricci, ardisco <code></l></code>
<code><l rend="i"></code>	Contro de' vizii miei darmi battaglia, <code></l></code>
	Tant de portraits brillans formaient plus d'un ta- bleau ! Toutefois, dans sa rage Athalie implacable ; Phedre, en proie aux remords, & malgré foi coupable ;
<code><l></code>	Tant de portraits brillans formaient plus d'un tableau ! <code></l></code>
<code><l></code>	Toutefois, dans sa rage Athalie implacable ; <code></l></code>

<l>Phedre, en proie aux remords, & malgré soi coupable ;</l>

3) En cas de mise en forme particulière concernant un vers entier, il est possible d'utiliser l'attribut @rend.

Exemples :

1) vers centré : <l rend="center"></l>
 2) vers entièrement en italique : <l rend="i"></l>
 3) vers entièrement en petites capitales : <l rend="sc"></l>

• Documentation : https://obvil.github.io/Teinte/teinte.html#el_l.

! <l/>

À proscrire, car correspond généralement à un changement de strophe : voir élément <lg> ci-dessous.

! <lg>

Strophe de vers.

Ne pas employer cet élément pour toutes les citations en vers (rien ne prouve qu'un groupe de vers constitue une strophe). À utiliser uniquement quand une ligne de blanc délimite clairement deux groupes de vers.

Exemple :

<pre> <lg> <l>Bientôt sur la tragique scene</l> <l>L'art tragique s'éclipsera</l> <l>Je vois travestir Melpomene</l> <l>En machiniste d'opéra.</l> </lg> <lg> <l>Bientôt une ivresse indiscrette</l> <l>Séduira cent jeunes Auteurs.</l> <l>Je vous, pour un seul vrai Poète,</l> <l>Vingt futiles décorateurs.</l> </lg> </pre>	<pre> Bientôt sur la tragique scene L'art tragique s'éclipsera . Je vois travestir Melpomene En machiniste d'opéra. Bientôt une ivresse indiscrette Séduira cent jeunes Auteurs. Je vois, pour un seul vrai Poète, Vingt futiles décorateurs. </pre>
--	---

N.-B. : en cas de mise en forme particulière concernant une strophe entière, il est possible d'utiliser l'attribut @rend.

Exemples :

- 1) strophe centrée : `<lg rend="center"></lg>`
- 2) strophe entièrement en italique : `<lg rend="i"></lg>`
- 3) strophe entièrement en petites capitales : `<lg rend="sc"></lg>`

- Documentation : https://obvil.github.io/Teinte/teinte.html#el_lg.

! `<quote>`

Citation, notamment les paragraphes en retrait ou/et en italique.

Citation en prose :

```
<quote>
  <p></p>
  <p></p>
</quote>
```

Citation en vers :

```
<quote>
  <l></l>
  <l></l>
  <l></l>
</quote>
```

- Documentation : https://obvil.github.io/Teinte/teinte.html#el_quote.

! `<bibl>`

Référence bibliographique.

À utiliser pour la source d'une citation.

**Elle peint des amans la joie & la tristesse,
Flatte, menace, irrite, apaise une maîtresse.
Mais pour bien exprimer ces caprices heureux,
C'est peu d'être Poète il faut être amoureux.**
Boil. Art Poët.

```
<quote>
  <l>Elle peint des amans la joie & la tristesse,</l>
  <l>Flatte, menace, irrite, apaise une maîtresse,</l>
  <l>Mais pour bien exprimer ces caprices heureux,</l>
  <l>C'est peu d'être Poète il faut être amoureux.</l>
  <bibl><hi rend="i">Boil.</hi> Art Poët.</bibl>
</quote>
```

- Documentation : https://obvil.github.io/Teinte/teinte.html#el_bibl.

! `<ab>`

Un bloc « anonyme ». Utilisé pour les séparateurs centrés de type ornement (*, ***, etc.).

Indiqué le cas échéant dans le PDF. Dans ce cas, utiliser l'attribut @type, valeur "ornament" = <ab type="ornament"/>.

- Documentation : https://obvil.github.io/Teinte/teinte.html#el_ab.

! <label>

Étiquette, pour un intitulé court, de niveau paragraphe ou de niveau caractère. Utilisé par exemple pour les intertitres, pour des listes de définitions, ou dans le théâtre.

Indiqué le cas échéant dans le PDF. Deux usages principaux :

- 1) les titres non hiérarchiques (titre d'un poème cité, par exemple), dans ce cas <label type="title">;
- 2) les locuteurs dans les citations de théâtre ou de dialogues, dans ce cas <label type="speaker">.

Exemples :

*Autre sur la mort d'un Ecclésiastique puissant
& détesté.*

Je fais bien qu'un homme d'Eglise
Qu'on redoutait fort en ce lieu
Vient de rendre son ame à Dieu :
Mais je ne fais si Dieu la prise.

```
<quote>  
<label type="title"><hi rend="i">Autre sur la mort d'un Ecclésiastique puissant & détesté.</hi></label>  
<l>Je fais bien qu'un homme d'Eglise</l>  
<l>Qu'on redoutait fort en ce lieu</l>  
<l>Vient de rendre son ame à Dieu :</l>  
<l>Mais je ne fais si Dieu la prise.</l>  
</quote>
```

Daphné

Pour te donner un baiser, soit;
La civilité me l'ordonne.

Damon.

Si la civilité le donne
C'est mon amour qui le reçoit.
Baise, baise je te supplie,
Daphné veux-tu me refuser?

```
<quote>  
<label type="speaker"><hi rend="i">Daphné</hi></label>  
<l>Pour te donner un baiser, soit ;</l>  
<l>La civilité me l'ordonne.</l>  
<label type="speaker"><hi rend="i">Damon.</hi></label>  
<l>Si la civilité le donne</l>  
<l>C'est mon amour qui le reçoit.</l>  
<l>Baise, baise je te supplie,</l>  
<l>Daphné veux-tu me refuser?</l>  
</quote>
```

```
<l>Si la civilité le donne</l>  
<l>C'est mon amour qui le reçoit.</l>  
<l>Baise, baise je te supplie,</l>  
<l>Daphné veux-tu me refuser ?</l>  
</quote>
```

- Documentation : https://obvil.github.io/Teinte/teinte.html#el_label.

II.2.D.4bis. Blocs (cas du théâtre)

Les pièces de théâtre (6 textes du lot) constituent un cas particulier, avec appel à trois balises dédiées : `<sp>` = tour de parole ; `<speaker>` = locuteur ; `<stage>` = didascalie. La balise `<l>` fait également l'objet d'une instruction particulière. Ces textes feront l'objet d'instructions spécifiques liées à chaque PDF.

`<sp>`

Théâtre : tirade, tour de parole.

Délimite le tour de parole dans son intégralité, y compris le nom du locuteur (`<speaker>`) et le texte proprement dit.

Uniquement dans le théâtre.

- Documentation : https://obvil.github.io/Teinte/teinte.html#el_sp.

`<speaker>`

Nom du locuteur dans une pièce de théâtre.

Uniquement dans le théâtre.

- Documentation : <http://www.tei-c.org/release/doc/tei-p5-doc/fr/html/ref-speaker.html>.

`<stage>`

Théâtre, didascalie, pouvant apparaître à différents niveaux du texte.

La didascalie peut figurer en début ou en fin de section (personnages, décor, mouvements), à la suite du nom du locuteur, dans le corps du texte.

Uniquement dans le théâtre.

- Documentation : https://obvil.github.io/Teinte/teinte.html#el_stage.

`<l part="">`

Vers coupés entre plusieurs locuteurs dans les pièces de théâtre = balisage avec l'attribut `@part` et l'indication I (début), M (milieu), F (fin) (`<l part="I">`, `<l part="M">`, `<l part="F">`). Les cas seront précisément indiqués sur le PDF et le nombre de textes concernés est limité (6 textes du lot).

Exemples :

	CHARLES.
	J'en appelle à mes droits.
	HUGUES.
	J'en appelle à la France.

```

<sp>
<speaker>CHARLES.</speaker>
<l part="i">J'en appelle à mes droits.</l>
</sp>
<sp>
<speaker>HUGUES.</speaker>
<l part="f">J'en appelle à la France.</l>
</sp>
    
```

	GEORGETTE.
	Vas-y toy.
	ALAIN.
	Vas-y toy.
	GEORGETTE.
	Ma foy je n'iray pas.
	ALAIN.
	Je n'iray pas aussi.
	ARNOLPHE.
	Belle ceremonie,
	Pour me laisser dehors. Hola ho je vous prie.

```

<sp>
<speaker>Georgette.</speaker>
<l part="i">Vas-y-toy.</l>
</sp>
<sp>
<speaker>Alain.</sp>
<l part="m">Vas-y-toy.</l>
</sp>
<sp>
<speaker>Georgette.</speaker>
<l part="f">Ma foy je n'iray pas.</l>
</sp>
<sp>
<speaker>Alain.</sp>
<l part="i">Je n'iray pas aussi</l>
</sp>
<sp>
<speaker>Arnolphe.</speaker>
<l part="f">Belle ceremonie,</l>
<l>Pour me laisser dehors. Hola ho je vous prie.</l>
</sp>
    
```


II.2.D.5. Figures, tables et listes

<figure>

Boîte hors du flux, parfois encadrée, pouvant contenir une image, une table, une formule, ou autre objet susceptible d'être référencé dans une table dédiée. Permet notamment d'envelopper une image avec sa légende.

Utiliser uniquement pour indiquer la présence d'une illustration. L'insertion ou non de l'image source sera réglée ultérieurement par les équipes de recherche.

Exemple :

```
<figure>Illustration p. 56</figure>
```

- Documentation : https://obvil.github.io/Teinte/teinte.html#el_figure.

<list>

Liste. Indiqué le cas échéant dans le PDF. Comprend des éléments <item>.

Exemple :

```
<list>
  <item></item>
  <item></item>
  <item></item>
</list>
```

- Documentation : https://obvil.github.io/Teinte/teinte.html#el_list.

Tableaux

<table>

Tableau.

- Documentation : https://obvil.github.io/Teinte/teinte.html#el_table.

<row>

Ligne d'un tableau.

Documentation : https://obvil.github.io/Teinte/teinte.html#el_row.

<cell>

Cellule d'un tableau.

Documentation : https://obvil.github.io/Teinte/teinte.html#el_cell.

Exemple d'un tableau à trois colonnes et deux lignes :

```
<table>
<row>
<cell></cell>
<cell></cell>
</row>
<row>
<cell></cell>
<cell></cell>
</row>
<row>
<cell></cell>
<cell></cell>
</row>
</table>
```

II.2.D.6. Fonctions

! <hi>

Mise en forme de niveau caractère, sans interprétation sémantique (ex : italique, gras, couleurs...).

Utilisé avec l'attribut @rend, restreindre aux valeurs suivantes : "b" = gras, "big" = lettrine, "i" = italique, "sc" = petites capitales, "u" = souligné, "sup" = exposant.

Exemples :

```
Quichotte = <hi rend="b">Quichotte</hi>
Quichotte = <hi rend="i">Quichotte</hi>
QUICHOTTE = <hi rend="sc">Quichotte</hi>
Quichotte = <hi rend="u">Quichotte</hi>
Napoléon Ier = Napoléon I<hi rend="sup">er</hi>
Tables de La Revue blanche 1889-1903 = <hi rend="i">Tables de <hi rend="i">La Revue blanche</hi>
1889-1903</hi>
```

- Documentation : https://obvil.github.io/Teinte/teinte.html#el_hi.

II.2.D.7. Points

! <pb/>

Saut de page « page break ».

Obligatoire. Format de saisie <pb n=""/>. Les chiffres romains sont autorisés. En l'absence d'un numéro de page, n="n.p.".

Attention :

- 1) En cas de césure de page à page (mot coupé d'une page sur l'autre), poser la balise avant le mot concerné.
- 2) Dans le cas d'une nouvelle section (<div>), le <pb> peut figurer entre l'ouverture de la section (<div>) et son intitulé (<head>), mais il ne doit en aucun cas figurer avant la <div>.

Exemple : un chapitre commençant en haut de page :

```
<div2>  
<pb n="48"/>  
<head>Chapitre VI</head>
```

MAIS PAS

```
<pb n="48"/>  
<div2>  
<head>Chapitre VI</head>
```

- 3) En général, faire figurer le <pb> avant l'intitulé d'un titre, hiérarchique ou non (<head>, <label>).

Documentation : https://obvil.github.io/Teinte/teinte.html#el_pb.

! <cb/>

Saut de colonne.

Mêmes règles d'utilisation que <pb>.

Documentation : https://obvil.github.io/Teinte/teinte.html#el_cb.

! <lb/>

Saut de ligne « line break ».

Documentation : https://obvil.github.io/Teinte/teinte.html#el_lb.

! <note/>

Une note, encodée dans le flux du texte, au point de l'appel.

S'il existe, faire figurer l'appel (1, 2, 3, etc. ; a, b, c, etc. ; *, **, ***, etc.) dans l'attribut @n. Exemple : <note n="1">.

Position des notes : utiliser l'attribut @place, avec les valeurs "bottom" et "margin". Note en bas de page : <note place="bottom">. Note en marge : <note place="margin">. S'il est difficile de déterminer le point d'appel d'une note (cas fréquent pour les notes marginales des textes 16^e-18^e), suivre les indications portées sur le PDF.

Documentation : https://obvil.github.io/Teinte/teinte.html#el_note.

II.3. LOT 2

II.3.A. Format du livrable (indications générales)

- Traitement de texte (de préférence odt).

- Éléments à restituer lors de la saisie :

- Éléments à restituer lors de la saisie :

1) Graphies originales, sauf le s long (ſ), à transcrire par s.

2) Lignes : passage à la ligne = saut de ligne

3) Caractères spéciaux (abréviations) : utiliser les équivalences de la table des caractères livrée avec chaque document. Voir II.3.D.

4) Coupures de mot (césures) : rendre par (-) ou trait d'union conditionnel . Voir table des caractères en II.3.D. Si indiqué dans le PDF passer les mots à cheval sur deux pages en page suivante.

Exemple simple de saisie :

iay dediée ces fascetieux deuitz, lesquelz sont de leur
 naturel, uaieusnir les vieulx emerites, font rire les pi
 teux, allument le feu de ioye, font croistre lencre au
 cornet seruent demplastres au besoing, & font ou(-)
 blier toutes rencunes, maledictions & iniures, car el(-)
 les ont esté tainctes au fleue de l'esté. Et generale(-)
 ment à toutes personnes de quelque qualité & con(-)
 dition qu'ilz soi(e1) t sains & malades, bossus, boyteux,

- 5) Numéros de page ou de folio : inscrit en début de page/folio, en style <pb>. Voir II.3.C. Les numéros de page ou folio sont systématiquement indiqués sur le PDF.
- 6) Soulignements : souligner le mot.
- 7) Lettrines : insérer une lettrine.
- 8) Lettres rubriquées : mettre en **gras**.
- 9) Paragraphes centrés : centrer le paragraphe.
- 10) Culs-de-lampe : rendre les espacements par des espaces insécables.
- 11) Illustrations : indiquer [figure].

II.3.B. Métadonnées (en-tête de document)

Un cartouche de métadonnées sera fourni avec chaque PDF et devra figurer en tête du document.

Exemple :

title : Le Decameron de Messire Iehan Bocace Florentin, nouvellement traduit d'Italien en François par Maistre Anthoine le Maçon conseiller du Roy & tresorier de lextraordinaire de ses guerres [traduction d'Antoine Le Maçon]
author : Boccace (1313-1375)
publisher : Sorbonne Université, LABEX OBVIL
issued : 2018
idno : http://obvil.paris-sorbonne.fr/corpus/faceties/boccace_decameron-nouvellement-traduit_1545/
source : Boccace, *Le Decameron de Messire Iehan Bocace Florentin, nouvellement traduit d'Italien en François par Maistre Anthoine le Maçon conseiller du Roy & tresorier de lextraordinaire de ses guerres* [traduction d'Antoine Le Maçon], Paris, Estienne Roffet, 1545, 254 f. + 8 f. prélim. ; portrait en tête du volume, fol. Chantilly, Bibliothèque Condé, Cabinet des livres, IV-H-38.
created : 1545
language : frm

II.3.C. Stylage

Appliquer une feuille de style simplifiée :

- <pb> [style de caractère] = numéro de page ou de folio, format [p.] ou [f.]. Indiqué en marge du PDF = [f. 1r], [f. 1v], [f. 2r.], etc.
- Titre 1, Titre 2, Titre 3 = niveau de titre hiérarchique. Indiqué T1, T2, etc. en marge du PDF.
- Corps de texte = paragraphe de texte en prose.
- <l> = paragraphe de texte en vers.

- <dateline> = date et lieu en début ou en fin de section. Indiqué <dateline> en marge du PDF.
- <label> = titre non hiérarchique. Indiqué <label> en marge du PDF.
- <signed> = signature en fin de section. Indiqué <signed> en marge du PDF.
- <explicit> = paragraphe indiquant la fin d'une section. Indiqué <explicit> en marge du PDF.
- <trailer> = mention indiquant la fin d'une section. Indiqué <trailer> en marge du PDF.

II.3.D. Exemples d'instructions spécifiques pour un document

II.3.D.1. Exemple 1 : fichier faceties_001

Lettrines












Insérer une Lettrine (dans le texte) :

P

Exemples	Convention provisoire	Encodage final	Rendu
	P _{LURES}	<hi rend="initial">P</hi>Lures	—
	H _{OMO}	<hi rend="initial">H</hi><hi rend="big">O</hi>mo	—

Rubrication

Les majuscules en début de phrase sont rubriquées en deux couleurs, rouge et orange. Ne pas tenir compte de la différence de couleur. À la saisie se contenter de mettre la lettre en gras. Après la conversion XML, les <hi rend="bold"> seront remplacés par des <hi rend="big">.

Exemples	Convention provisoire	Encodage final	Rendu
	A	<hi rend="big">A</hi>	—
	C	<hi rend="big">C</hi>	—
	C	<hi rend="big">C</hi>	—
	D	<hi rend="big">D</hi>	—
	E	<hi rend="big">E</hi>	—
	H	<hi rend="big">H</hi>	—
	I	<hi rend="big">I</hi>	—
	N	<hi rend="big">N</hi>	—
	O	<hi rend="big">O</hi>	—
	T	<hi rend="big">T</hi>	—
	V	<hi rend="big">V</hi>	—

Passages soulignés



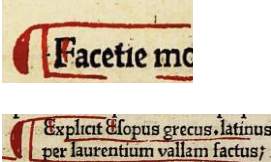
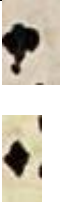



Un certain nombre de passages sont soulignés en rouge. Les souligner à la saisie.








Symboles, abréviations, caractères









Tableau fourni avec chaque document. Deux possibilités pour la plupart des caractères :

- 1) utiliser une convention provisoire ;
- 2) utiliser le rendu final.

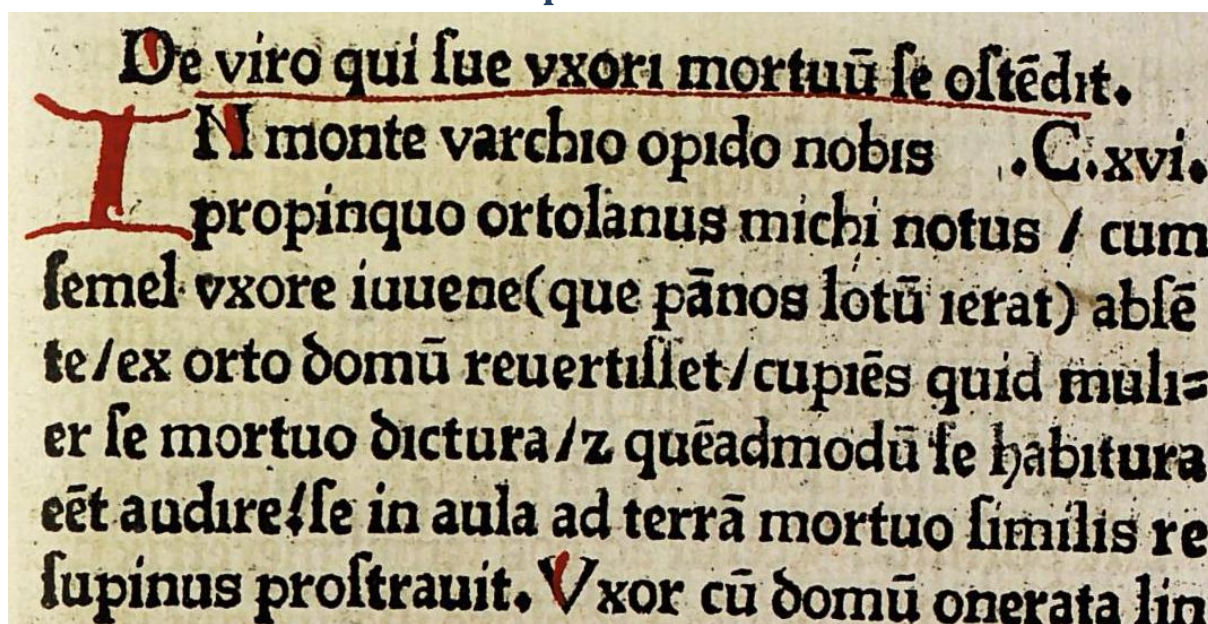
Le prestataire peut choisir l'une ou l'autre solution quand elles sont toutes deux disponibles, en indiquant le choix effectué. Le résultat sera réencodé après la conversion au format XML et fera l'objet d'un balisage spécifique (élément <choice>) pour le développement des abréviations.

Exemple	Convention provisoire	Encodage final	Rendu	Descriptif
	(-)	<lb rend="hyphen"/> [ligne suivante]		trait d'union conditionnel (U+00AD)
	(pdm)	¶	¶	pied-de-mouche
	(ipdm)	⁋	⋈	pied-de-mouche réfléchi [convention adoptée en l'absence d'un capitulum en C)
	(period)	.	.	point
	(sol)	/	/	solidus = virgula suspensiva
	(sol)	<hi rend="b">/</hi>	/	solidus (couleur rouge) = virgula suspensiva double (marque de paragraphe)
	(!)	!	!	point d'exclamation

Exemple	Convention provisoire	Encodage final	Rendu	Descriptif
	(i!)	¡	i	point d'exclamation culbuté
	(;)	;	;	point-virgule
	(?)		?	punctus interrogativus
	(a1)	<am>ā</am>	ā	a macron minuscule
	(con)	<am>ↄ</am>	↵	signe d'abréviation con minuscule
	d	d	d	d minuscule
	(e1)	<am>ē</am>	ē	e macron minuscule

Exemple	Convention provisoire	Encodage final	Rendu	Descriptif
	(e2)	<am>e̱̱</am>	ē̇	e minuscule + double macron souscrit
	G	G	G	G majuscule
	(i1)	<am>ī</am>	ī	i macron minuscule
	(o1)	<am>ō</am>	ō	o macron minuscule
	(q2)	<am>q̃</am>	q̃	q minuscule + tilde combiné
	(q3)	<am>qꝫ</am>	q₃	ligature q minuscule + lettre finale et (A76B)
	(q4)	<am>ȹ̃</am>	q̃p	ligature minuscule latine qp + tilde combiné
	(u1)	ū	ū	u macron minuscule

Exemples de saisie



Convention provisoire

De	viro	qui	sue	vxori	mortu(u1)	se	ost(e1)dit.			
.C.xvi.										
I										
N		monte		varchio		opido	nobis			
propinquo		ortolanus		michi	notus	(/)	cum			
semel	vxore	iuuene	(que	p(a1)nos	lot(u1)	ierat)	abs(e1)(-\$)			
te	(/)	ex	orto	dom(u1)	reuertisset	(/)	cupi(e1)	quid	muli(-)	
er	se	mortuo	dictura	(/)	z	qu(e1)admod(u1)	se	habitura		
e(e1)t	audire	(!)	se	in	aula	ad	terr(a1)	mortuo	similis	re
supinus		prostrauit(.)		Vxor	c(u1)	dom(u1)	onerata	lin		

Rendu final

De	viro	qui	sue	vxori	mortuū	se	ostēdit.			
.C.xvi.										
I	N	monte		varchio		opido	nobis			
propinquo		ortolanus		michi	notus	/	cum			
semel	vxore	iuuene	(que	pānos	lotū	ierat)	absē\$			
te	/	ex	orto	domū	reuertisset	/	cupiē	quid	muli	
er	se	mortuo	dictura	/	z	quēadmodū	se	habitura		
eēt	audire	!	se	in	aula	ad	terrā	mortuo	similis	re
supinus		prostrauit.		Vxor	cū	domū	onerata	lin		

II.3.D.2. Exemple 2 : fichier faceties_009

Lettrines

Insérer une Lettrine (dans le texte) :

H

Exemple	Convention provisoire	Encodage final	Rendu
		<code><hi rend="initial">H</hi></code>	
		<code><hi rend="initial">T</hi></code>	



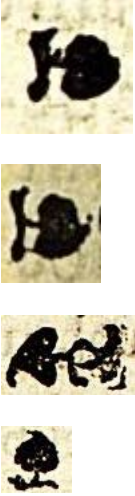





Symboles, abréviations, caractères










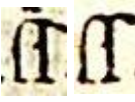



Tableau fourni avec chaque document. Deux possibilités pour la plupart des caractères :

- 1) utiliser une convention provisoire ;
- 2) utiliser le rendu final.

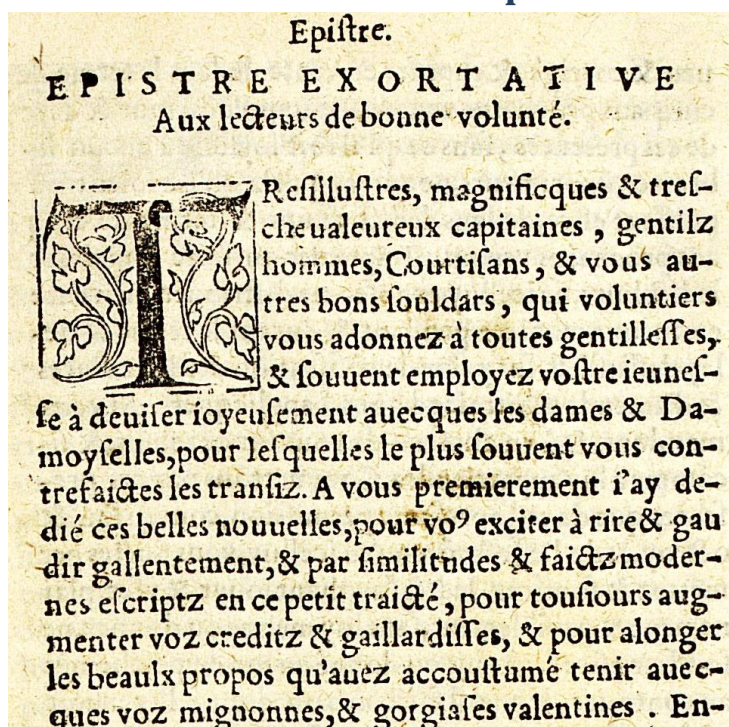
Le prestataire peut choisir l'une ou l'autre solution quand elles sont toutes deux disponibles, en indiquant le choix effectué. Le résultat sera réencodé après la conversion au format XML et fera l'objet d'un balisage spécifique (élément `<choice>`) pour le développement des abréviations.

Exemple	Convention provisoire	Encodage final	Rendu	Descriptif
	(-)	<code><lb rend="hyphen"/></code> [au début de la ligne suivante]		trait d'union conditionnel (U+00AD)
	(seg)car(/seg)			[caractères difficilement lisibles : signes de ponctuation ? Marquer comme

Exemple	Convention provisoire	Encodage final	Rendu	Descriptif
				seg]
	,	,	,	virgule [même si elle semble réfléchie]
	(hedera)	❧		fleuron (1 seule convention quel que soit le sens)
	(pdm)	¶	¶	pied-de-mouche
	&	&	&	esperluette
	(a1)	<am>ā</am>	ā	a macron minuscule
	ct	ct	ct	ligature c + t

Exemple	Convention provisoire	Encodage final	Rendu	Descriptif
	(e1)	<am>ē</am>	ē	e macron minuscule
	ff	ff	ff	ligature f + f
	fi	fi	ff	ligature f + i
	(i1)	<am>ī</am>	ī	i macron minuscule
	(o1)	<am>ō</am>	ō	o macron minuscule
	(p3)	<am>ꝑ</am>	p̅	p minuscule barré à travers la descendante
	(q1)	<am></am>	q̃	q macron minuscule
	s	s	s	s long
	si	si	si	ligature s long + i
	ss	ss	ss	ligature s long + s long
	st	st	st	ligature s long + t
	(u1)	<am>ū</am>	ū	u macron minuscule
	(us)	<am>ꝰ</am>	⁹	lettre modificative us

Exemples de saisie



Exemple A

Convention provisoire

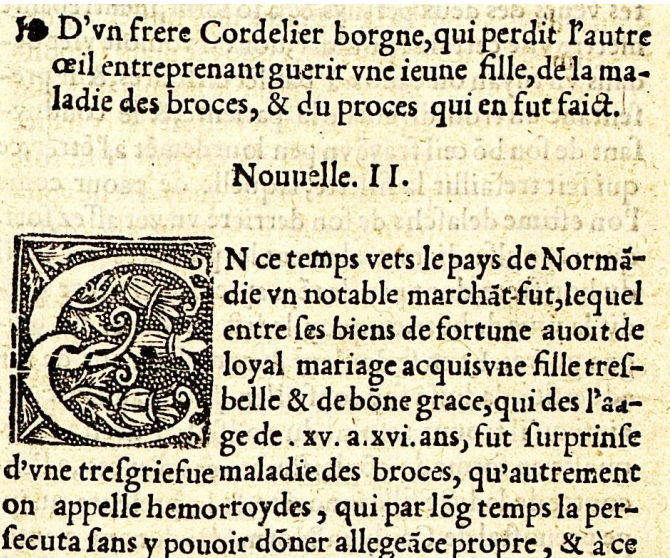
<p>T Resillustres, magnifiques & tres(-)</p> <p>cheualeureux capitaines, gentilz hommes, Courtisans, & vous au(-) tres bon souldars, qui volontiers vous adonnez à toutes gentilleses, & souuent employez vostre ieunes(-) se à deuiser ioyeusement avecques les dames & Da(-) moyselles, pour lesquelles le plus souuent vous con(-) trefaictes les transiz. A vous premierement i'ay de(-) dié ces belles nouuelles, pour vo(us) exciter à rire & gau(-) dir gallentement, & par similitudes & faicts moder(-) nes escriptz en ce petit traicté, pour tousiours aug(-) menter vos creditz & gaillardises, & pour alonger les beaulx propos qu'avez accoustumé tenir avec(-) vos mignonnes, & gorgias valentines. En(-)</p>	<p>EPISTRE EXHORTATIVE Aux lecteurs de bonne volonté.</p>
---	---

Rendu final

<p>T Resillustres, magnifiques & tres</p> <p>cheualeureux capitaines, gentilz hommes, Courtisans, & vous au-</p>	<p>EPISTRE EXHORTATIVE Aux lecteurs de bonne volonté.</p>
---	---

tres bon souldats, qui voluntiers
vous adonnez à toutes gentilleses,
& souuent employez vostre ieunes-
se à deuiser ioyeusement avecques les dames & Da-
moyselles, pour lesquelles le plus souuent vous con-
trefaictes les transiz. A vous premierement i'ay de-
dié ces belles nouuelles, pour vo⁹ exciter à rire & gau-
dir gallentement, & par similitudes & faicts moder-
nes escriptz en ce petit traicté, pour tousiours aug-
menter vos creditz & gaillardises, & pour alonger
les beaulx propos qu'auiez accoustumé tenir avec-
vos mignonnes, & gorgiases valentines. En-

Exemple B



Convention provisoire

(hedera) D'un frere Cordelier borgne, qui perdit l'autre
œil entreprenant guerir vne ieune fille, de la ma(-)
ladie des broces, & du proces qui en fut faict.

Nouuelle. II.

E N ce temps vers le pays de Norm(a1)(-)
die vn notable march(a1)t fut, lequel entre ses biens de fortune auoit de
loyal mariage acquis vne fille tres(-)
belle & de b(o1)ne grace, qui des l'aa(-)
ge de. xv. à. xvi. ans, fut surprinse
d'vne tresgriefue maladie des broces, qu'autrement
on appelle hemorroydes, qui par l(o1)g temps la per(-)
secuta sans y pouoir d(o1)ner allege(a1)ce propre, & à ce

Rendu final

☛ D'un frere Cordelier borgne, qui perdit l'autre
œil entreprenant guerir vne ieune fille, de la ma
ladie des broces, & du proces qui en fut faict.

Nouuelle. II.

EN ce temps vers le pays de Normā
die vn notable marchāt fut, lequel entre ses biens de fortune auoit de
loyal mariage acquis vne fille tres
belle & de bōne grace, qui des l'aa
ge de. xv. à. xvi. ans, fut surprinse
d'une tresgrieue maladie des broces, qu'autrement
on appelle hemorroydes, qui par lōg temps la per
secuta sans y pouoir sdōner allegeāce propre, & à ce

II.4. LOT 3

II.4.A. Format du livrable (indications générales)

- Traitement de texte (de préférence odt).
- Ne pas conserver :
 - 1) Les titres courants.
 - 2) Les signatures (numéros de cahiers) et les réclames (premier mot ou début du premier mot de la page suivante) sous le texte.
- Indiquer les numéros de page (format [p.]).
- Restituer le flux du texte (pas de coupure à chaque saut de page).
- Insérer les notes au point d'appel. Rétablir l'unité des notes tournantes (courant d'une page sur l'autre).

II.4.B. Métadonnées (en-tête de document)

Un cartouche de métadonnées sera fourni avec chaque PDF et devra figurer en tête du document.

Exemple :

title : Marie Taglioni
author : André Levinson [Levinson, André (1887-1933)]
publisher : Sorbonne Université, LABEX OBVIL
issued : 2018
idno : http://obvil.paris-sorbonne.fr/corpus/danse/levinson_marie-taglioni/
source : André Levinson, *Marie Taglioni*, Paris, Félix Alcan, 1929, 1 vol. (143 p.-
[VIII] f. de pl.) : ill. ; in-8. PDF : [Gallica](#).
created : 1929

II.4.C. Stylage

Appliquer une feuille de style simplifiée :

- <pb> [style de caractère] = numéro de page ou de folio, format [p. n] ou [f. n].
- Titre 1, Titre 2, Titre 3 = niveau de titre hiérarchique. Indiqué T1, T2, etc. en marge du PDF.
- Corps de texte = paragraphe de texte en prose.
- <quote> = citation, paragraphe en retrait ou dans un corps plus petit.
- <l> = paragraphe de texte en vers.
- <dateline> = date et lieu en début ou en fin de section. Indiqué <dateline> en marge du PDF.
- <label> = titre non hiérarchique + locuteur dans une pièce de théâtre. Indiqué <label> en marge du PDF.
- <signed> = signature en fin de section. Indiqué <signed> en marge du PDF.
- <trailer> = mention indiquant la fin d'une section. Indiqué <trailer> en marge du PDF.