

Documentation

encodingDesc

segmentation

Description des principes selon lesquels le texte a été segmenté, par exemple en phrases, en intonèmes (unités tonales), en strates graphématiques (niveaux superposés de signes graphiques), etc.

Toute édition doit être contenue dans un élément `<text>` précisé par les attributs `@xml:id` et `@n` afin d'identification. L'élément `<body>` et son sous-élément `<div type="letter">`¹ contiendront la transcription de la lettre. Un élément `<back>` peut être intégré pour la gestion des notes et des index.

Les éventuelles enveloppes pourront être encodé dans le `<body>`, dans un sous-élément `<div type="enveloppe">` avec un attribut `@facs`

Corps de la correspondance

Il convient de rendre compte, lors de l'encodage et dans la mesure du possible, de la structure diplomatique des dépêches.²

Les formules d'ouverture, seront placées dans un élément `<opener>`. Les mentions de dates et de lieux trouveront leur place au sein d'un sous-élément `<dateLine>` alors que l'apostrophe sera encodée à l'aide de la balise `<salute>`.

Le corps de la lettre sera encodé dans un élément `<p>`. Les changements de ligne, de colonne, de page ou encore de support seront matérialisés par l'emploi respectif des balises `<lb/>`, `<cb/>`, `<pb/>`, `<gb/>`; ils pourront être complétés par l'utilisation des attributs `@n` et `@facs` afin de leur associer une représentation fac-similaire par exemple.

Les formules de fermeture de la correspondance seront placées dans un élément `<closer>` pouvant contenir les formules de politesse (`<salute>`), les mentions de date (`<dateline>`), la signature (`<signed>`) ou encore une adresse (`<address>` avec les sous-éléments `<addrLine>`). Le `<closer>` peut être répété et typé afin de prendre en compte

1. Ils peuvent aussi être complétés par un attribut `@xml:id`.

2. Camille DESENCLOS, *Édition de correspondances*, URL : <http://developpements.enc.sorbonne.fr/diple/schema/correspondance> (visité le 28/05/2015).

les mentions autographes par exemple.³

Le postscriptum (<postscript>) peut être placé avant ou après le <closer> et il peut prendre la même forme qu'une correspondance (utilisation de formules d'ouverture, de fermeture et de signature.).

```
<text>
  <body>
    <div type="letter">
      <opener>
        <dateline>Paris , le 28 mars 1916</dateline>
        <salute>Madame,</salute>
      </opener>
      <p>
        Ici le corps de la lettre
      <lb/>
      <pb/>....
    </p>
    <closer>
      <salute/>
      <signed/>
    </closer>
    <postscript>
      <opener/>
      <p/>
      <closer/>
    </postscript>
  </div>
</body>
<back>
  <div type="note">
    <note/>
  </div>
</back>
</text>
```

3. *Ibid.*

Ponctuation

Les éléments de ponctuation seront encodés avec la balise `<pc>`. À toutes fins de normalisation, la ponctuation ajoutée sera associée à l'attribut `@type="supplied"` ; la ponctuation superflue sera typée `@type="surplus"`. Enfin en cas de modification d'une ponctuation existante, nous utiliserons un élément `<choice>` et ses sous-éléments `<orig>` et `<reg>` afin d'identifier la forme originelle de la forme régularisé.

La modernisation de la ponctuation devra s'accompagner du rétablissements des majuscules et des minuscules à l'aide de la balise `<hi>` et de son attribut `@rend="capitalize/minimize"`.⁴

Correction

Concernant les corrections orthographiques, elles seront encodées avec une élément `<choice>` et les couples `<reg>/<orig>` et `<sic>/<corr>`. Une distinction sera faite entre, les fautes systématiques (`<reg>/<orig>`) et les fautes occasionnelles éparses (`<sic>/<corr>`)⁵.

Normalisation

Afin de pouvoir offrir une lecture modernisée du texte, certains termes doivent être normalisés :

- Les nombres devront aussi être encodés en toutes lettres à quelques exceptions près :
 - En chiffres arabes : les dates, énumérations, numéros d'ordre (adresse, matricule...), calibres, degrés, échelles pourcentages...
 - En chiffre romains : siècles, numéros dynastiques, régions politiques, régions militaires, division d'un ouvrage...
 - Utilisation des balises `<choice>` et `<reg>/<orig>`
- Les majuscules et minuscules doivent être rétablies par l'utilisation de la balise `<hi>` et de sont attributs `@rend="capitalize/minimize"` une attention particulière portera sur :
 - La majuscule en début de phrase ;
 - La majuscule pour l'initiale d'un nom propre ;
 - Les minuscules pour les noms communs.

4. Voir section suivante

5. Christine NOUGARET et Élisabeth PARINET, *L'édition critique des textes contemporains (XIXe-XXIe siècle)*, École nationale des chartes, 2015.

- Il faut fournir, dans la mesure du possible, une version développée des abréviations utilisées (à l'exception des plus courantes) par l'emploi de l'élément `<choice>` et des sous-éléments `<abbr>` et `<expan>`⁶.
- Les coupures et espacement des mots doivent aussi être normalisés.
 - L'élément `<space/>` peut être utilisé pour sémantiser un espace
 - Pour la séparation des mots unis dans la source, l'élément `<w>` avec un attribut `@rend="aggl"` sera utilisé.
Exemple : enrouté > `<w rend="aggl">en</w><w>route</w>`
 - Pour la séparation des mots donnant lieu à une élision, la même balise sera utilisée mais avec l'attribut `@rend="elision"`.
Exemple : lun > `<w rend="elision">l</w><w>un</w>`
 - Enfin, pour la réunion de deux mots séparés dans la source, une simple élément `<w>` sera employé.
Exemple : criti que > `<w>criti que</w>`

hyphénation

les cas d'hyphénation seront matérialisés par l'attribut `@rend="hyphen"` appliqué à l'élément `linebreak` concerné (`<lb/>`).

généralité

Les Dates

Les dates seront encodées avec l'élément `Date`, précisés par les attributs `@when` (format : AAAA-MM-JJ) et en cas d'incertitudes `@cert` et `@nbefore / @notafter` pour donner une échelle.

Noms

Les noms de personne et de lieu seront encodé avec les éléments `<persName>` et `<placeName>` pouvant être complété par l'attribut `@ref` si l'on dispose d'une entrée d'index.

pour les autres noms, l'élément `<name>` pourra être utilisé avec les attributs `@type` pour le préciser et `@ref` pour l'indexation.

6. Il n'est pas besoin d'utiliser l'élément `<ex>` pour encoder les lettre ajouter pour le développement de l'abréviation.

Les ajouts - suppressions - oublis - manques - incertitudes

Les ajouts de l'auteur seront encodé avec l'élément `<add>` complété par les attributs `@place` pour l'emplacement, `@rend` pour l'orientation du texte, `@type` pour en préciser la nature et éventuellement un `@n` pour l'ordre de lecture. les ajouts du transcripteur ou de l'éditeur seront encodé avec l'élément `<supplied>` et ses attributs `@reason` afin d'expliquer l'ajout et `@cert` pour la certitude du terme ajouté. Les ajouts d'une autre main seront matérialisé par l'élément `<handshift/>` aidé des attributs `@medium`, `@rend` et `@resp` lorsque c'est possible.

Les parties de textes effacées, biffées, ou raturées seront encodées avec l'élément `` et les attributs `@rend`, `@hand` `@cause`, `@cert`, `@rest`. Il est ainsi possible d'identifier la main qui a causé l'effacement et d'en préciser les raisons. Les suppressions liées à au transcripteur ou à l'éditeur seront encodées avec l'élément `<surplus>` et l'attribut `@reason`

Les termes peu clairs, seront traités avec l'élément `<unclear>` précisé des attributs `@reason`, `@agent`, `@hand` et `@cert`.

Enfin les manques seront matérialisés par la balise `<gap/>` avec l'attribut `@reason="missing"`

Adresses et cachets - timbres

Les adresses sont encodées avec l'élément `<address>` et le sous-élément `<addrLine>` en cas de répétition par exemple avec une correspondance tenant sur plusieurs carte la balise `<address>` sera complétée par l'attribut `@n`.

Les cachets et timbres seront encodé avec l'élément `<stamp>` et l'attribut `@type` pour les préciser. La balise `<orgName>` peut être utilisée pour préciser l'organisme à l'origine des cachets par exemple.

Autres

Les éléments en exposant seront encodé avec l'élément `<sup>` et l'attribut `@rend="super"`

Les ajoutées dans le textes seront placées dans des éléments `<note>` du back comme nous l'avons vu plus haut et complétés par les attributs `@xml:id` et `@n`. Afin de faire le lien dans le corps de la dépêche, un élément `<ref>` poitant avec sont attribut `@target` vers le note et le même `@n` que cette dernière seront ajoutés.

Les images, photographies ou dessins, seront placés dans un élément `<figure>` comportant une balise `<graphic>` avec un attribut `@url` pour cibler le facsimilé et une balise `<figDesc>` pour la description de l'image.