

## **AML – Guide de transcription « HTML TEI »**

### **1. Les outils : éditeurs HTML et navigateurs**

Le HTML étant très simple à écrire, vous pouvez vous contenter de créer vos fichiers « HTML TEI » via le simple « Bloc-Notes » ou « Notes ». Cependant, un éditeur HTML dédié peut vous faciliter la tâche, principalement grâce aux fonctionnalités suivantes :

- a) formatage automatique du fichier, avec retrait vers la droite à mesure des imbriques et surveillance des couples « balises ouvrantes/fermantes »
- b) auto-complétion et suggestion
- c) détection des erreurs, essentiellement les balises ouvertes et non fermées
- d) pliage de code, où des balises peuvent être repliées (comme dans l'explorateur de vos dossiers) lorsqu'un fichier devient volumineux

Il existe énormément d'éditeurs HTML, dont certains intégralement gratuits. Les plus courants sont, du plus « léger » au plus « lourd » : Notepad++, Sublime Text et Visual Studio Code.

Durant votre transcription, double-cliquez sur le fichier html généré : il s'ouvrira, dans votre navigateur par défaut (Chrome, Firefox, Safari...) qui transformera vos balises en mises en page, de sorte que vous pourrez vérifier de manière « lisible » la progression de votre transcription.

### **2. Structure générale d'un document HTML**

Le principe d'un document HTML est de structurer son contenu via des « balises ». La plupart d'entre elles fonctionnent par couple, comme `<div>....</div>`, mais certaines sont dites « vides » et ne peuvent contenir de texte, comme `<br/>` ou `<hr/>`.

Ces balises peuvent être complétées par des attributs qui en précise le fonctionnement : ex. `<div data-type='page'>...</div>` permet de dire que ce div (qui est une balise très générique) représente spécifiquement une page.

Les balises peuvent s'imbriquer entre elles, mais la fermeture doit se produire dans l'ordre inverse de l'ouverture. Ainsi, un texte souligne (via la balise `<u>`) puis barré (via la balise `<del>`) sera transcrit : `<u><del>texte d'exemple</del></u>`.

### **3. Principales balises « HTML TEI »**

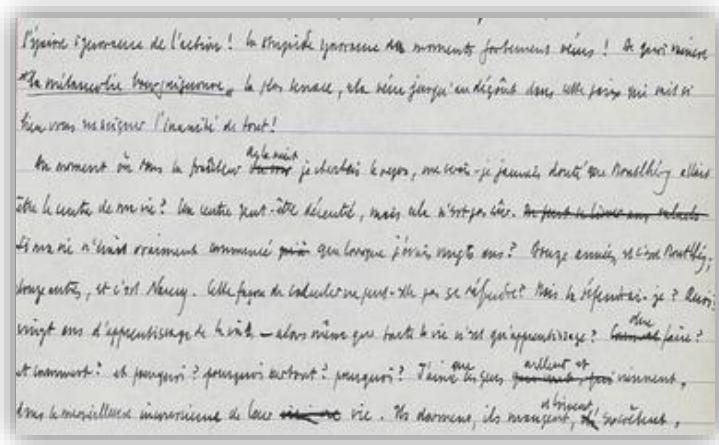
Vous pouvez potentiellement utiliser toutes les balises et tous les attributs HTML universellement supportés (voir la [Référence en ligne](#)). Cependant, afin de faciliter la transposition future de votre HTML TEI intermédiaire en XML TEI, voici celles que nous vous demandons d'utiliser en fonction des cas rencontrés dans le texte que vous transcrivez :

Cas	Balise HTML conseillée	Remarque
Page	<div data-type='page' data-src='url'>...</div>	<p>La balise <code>&lt;div&gt;</code> permet de créer, via l'attribut <code>data-type='page'</code>, de regrouper l'ensemble du texte d'une même page. L'attribut <code>data-src</code>, normalement présent dans cette configuration, permet d'ajouter l'url de l'image de la page transcrive. Comme vous transcrivez un manuscrit AML, il est préférable que cette <code>url</code> soit absolue et renvoie vers nos serveurs (ex. <a href="https://archives.aml-cfwb.be/ressources/public/ML/05841/ELB-AML-ML-05841-025_web.jpg">https://archives.aml-cfwb.be/ressources/public/ML/05841/ELB-AML-ML-05841-025_web.jpg</a>) mais vous pouvez aussi utiliser des urls relatives (ex. \MonDossier1\image1.jpg) du moment qu'elles sont uniques pour votre document.</p> <p>Cette balise n'a aucun rendu spécifique lors de la visualisation via votre navigateur.</p>
Saut de page	<hr/>	Elément vide, <code>&lt;hr/&gt;</code> représente un simple saut de page. Il est généralement rendu par les navigateurs comme une ligne horizontale traversant l'écran, ce qui vous permet de visualiser la structure de votre document à l'écran. Cependant, il ne permet pas de préciser l'url de l'image de la page transcrive ; il est donc préférable de privilégier <code>&lt;div class='page'&gt;</code> ci-dessus, et d'éventuellement la combiner avec <code>&lt;hr/&gt;</code> pour le rendu visuel.
Paragraphe	<p>...</p>	La balise <code>&lt;p&gt;</code> permet de regrouper l'ensemble du texte d'un même paragraphe.
Saut de ligne	 	Elément vide, <code>&lt;br/&gt;</code> représente un saut de ligne au cœur du texte. Il n'est pas nécessaire en fin de paragraphe, car le saut de ligne est automatique entre deux paragraphes.
Emphase	<em data-type='type'>...</em>	Pour signaler toute forme d'emphase d'une partie du texte par l'auteur. Son attribut <code>data-type</code> permet de préciser quelle forme prend cette emphase ( <code>entoure</code> , <code>encadre</code> ...). L'emphase par soulignement est rendue préféablement par la balise <code>&lt;u&gt;</code> .
Souligné	<u>...</u>	En fonction des cas rencontrés, vous pouvez ajouter un attribut <code>data-style</code> pour signaler par ex que le trait est ' <code>double</code> ' ou ' <code>ondule</code> '.
Supprimé	<del>...</del>	Généralement représenté rayé par les navigateurs. En fonction des cas rencontrés, vous pouvez ajouter un attribut <code>data-style</code> pour signaler par ex que le trait est ' <code>double</code> ' ou ' <code>ondule</code> '.
Insertion	<ins data-place ='place'>...</ins>	Cette balise peut être rendu par un soulignement dans un navigateur. Son attribut <code>data-place</code> permet de préciser où est inséré le texte, ex. <code>reecrit</code> , <code>marge</code> , <code>haut</code> , <code>bas</code> ... Ces deux dernières valeurs signalent que le texte a été respectivement ajouté au-dessus ou en-dessous de la ligne de texte. Dans le rendu final, cette insertion sera centrée sur l'endroit elle est située dans votre transcription ou sur la suppression éventuelle qui précède.
Texte en exposant	<sup>...</sup>	Pour rendre <code>XX<sup>e</sup> siècle</code> ou <code>1<sup>er</sup></code> par exemple.
Citation courte (au cœur du texte)	<q>...</q>	La plupart des navigateurs rendront cette balise en ajoutant automatiquement des guillemets, mais il vaut mieux les introduire dans votre texte s'il l'auteur les a utilisés.
Citation longue	<blockquote>...</blockquote>	Vous pouvez y utiliser toutes les balises nécessaires, comme <code>&lt;p&gt;</code> et <code>&lt;br/&gt;</code> par ex.
Illisible	<span data-type='illisible' data-nombre='nb'>...</span>	Cette balise permet de préciser qu'une partie de texte est illisible. Son attribut <code>data-nombre</code> permet en outre de dire combien de caractères compte cette séquence illisible.

## Du manuscrit au XML TEI

Exemple complet : <https://editions.aml-cfwb.be/tei/Temeraire/?page=page%208%20recto>

### 1) Extrait à transcrire



### 2) Transcription en HTML TEI via Visual Studio Code :

```
<div data-type='page' data-src='https://archives.aml-cfwb.be/ressources/public/ML/05841/ELB-AML-ML-05841-025_web.jpg'>
<p>
...
    l'épaisse ignorance de l'action ! la stupide ignorance des moments fortement vécus ! De quoi vaincre<br />
    <u>"la mélancolie bourguignonne,</u> la plus tenace, cela vécu jusqu'au dégoût dans cette paix qui sait si<br />
    bien vous enseigner l'inanité de tout !
</p>
<p>
    Au moment où, dans la fraîcheur du soir je cherchais le repos, me serais-je jamais douté que Montlhéry allait<br />
    être le centre de ma vie ? Un centre peut-être décentré, mais cela n'est pas sûr. On peut se livrer aux calculs<br />
    Si ma vie n'avait vraiment commencé qu'à que lorsque j'avais vingt ans ? Douze années, et c'est Montlhéry;<br />
    douze autres, et c'est Nancy. Cette façon de calculer ne peut-elle pas se défendre ? Mais la défendrai-je ? Quoi ?<br />
    vingt ans d'apprentissage de la vie – alors même que toute la vie n'est qu'apprentissage ? <del>Comment</del><ins data-place='haut'>Que</ins> faire ?<br />
    et comment ? et pourquoi ? et pourquoi surtout ? pourquoi ? J'aime <ins data-place='haut'>que</ins> les gens <del>qui vont, qui</del><ins data-place='haut'>aillent et viennent</ins>,<br />
    dans la merveilleuse inconscience de leur <del>vie</del> : <span data-type='illisible' data-nombre='3' style='font-family: monospace;'>xxx</span></del> ils</del> excretent,
...
</p>
</div>
```

3) Rendu du HTML TEI dans Firefox :

... l'épaisse ignorance de l'action ! la stupide ignorance des moments fortement vécus! De quoi vaincre “la mélancolie bourguignonne., la plus tenace, cela vécu jusqu'au dégoût dans cette paix qui sait si bien vous enseigner l'inanité de tout !

Au moment où, dans la fraîcheur du soir de la nuit je cherchais le repos, me serais-je jamais douté que Montlhéry allait être le centre de ma vie? Un centre peut-être décentré, mais cela n'est pas sûr. On peut se livrer aux calculs Si ma vie n'avait vraiment commencé qu'à que lorsque j'avais vingt ans? Douze années, et c'est Montlhéry; douze autres, et c'est Nancy. Cette façon de calculer ne peut-elle pas se défendre? Mais la défendrai-je? Quoi? vingt ans d'apprentissage de la vie — alors même que toute la vie n'est qu'apprentissage? Comment Que faire ? et comment? et pourquoi? et pourquoi surtout? pourquoi? J'aime que les gens qui vont, qui aillent et viennent, dans la merveilleuse inconscience de leur vie -xxx vie. Ils dorment, ils mangent, et boivent, ils excrètent, ...

4) Rendu final après conversion en XML TEI et transformation XSL dans Firefox

l'épaisse ignorance de l'action ! la stupide ignorance des moments fortement vécus! De quoi vaincre  
“la mélancolie bourguignonne., la plus tenace, cela vécu jusqu'au dégoût dans cette paix qui sait si  
bien vous enseigner l'inanité de tout !  
  
de la nuit  
Au moment où, dans la fraîcheur du soir je cherchais le repos, me serais-je jamais douté que Montlhéry allait  
être le centre de ma vie? Un centre peut-être décentré, mais cela n'est pas sûr. On peut se livrer aux calculs  
Si ma vie n'avait vraiment commencé qu'à que lorsque j'avais vingt ans? Douze années, et c'est Montlhéry;  
douze autres, et c'est Nancy. Cette façon de calculer ne peut-elle pas se défendre? Mais la défendrai-je? Quoi?  
vingt ans d'apprentissage de la vie — alors même que toute la vie n'est qu'apprentissage? Comment faire ?  
que  
et comment? et pourquoi? et pourquoi surtout? pourquoi? J'aime les gens qui vont, qui aillent et viennent,  
et boivent,  
dans la merveilleuse inconscience de leur vie -xxx vie. Ils dorment, ils mangent, ils excrètent,