

TP-4 : transcription de manuscrit

TEI @ Oxford

July 2012

1 Objets de l'exercice

Dans cet exercice nous allons explorer :

- comment effectuer une transcription d'un manuscrit
- l'application des balises `<add>`, `` and `<subst>`
- comment se servir de l'element `<choice>` pour traiter les abbreviations
- comment se servir de l'element `<unclear>` pour les morceaux de texte peu lisible

2 Sommaire

Nous allons prendre une transcription déjà faite, mais sans balisage et l'enrichir avec des indications de

- ajouts, ratures, substitution
- abbréviations et leur expansions
- passages difficile ou impossible de lire

Vous aurez besoin du logiciel oXygen, et des fichiers `preface-ms.xml` et `preface-ms.jpg` dans votre dossier Travaux

3 Pour commencer...

Lancez oXygen, et ouvrez le fichier `preface-ms.xml`. Vous verrez un document balisé au minimum, avec un entete minimale.

Visualisez le fichier image `preface-ms.jpg`

Nous allons attirer votre attention sur les différences entre les deux, et proposer des améliorations du balisage...

4 Usage de rend

Le premier mot du manuscrit est souligné pour indiquer que c'est le titre. Notre encodage le représente comme un `<head>`, mais n'indique pas qu'il est souligné.

- En oXygen, mettez le curseur à la fin de la balise ouvrante `<head>` (juste avant le `>`) et tapez un blanc
- une liste des attributs disponibles s'affiche. Sélectionnez `rend` et fournissez la valeur `underline`
- Changez en mode auteur et regardez l'effet.

5 Deuxieme paragraphe

Comparez la deuxième paragraphe dans l'image avec la version XML.

- Après le mot 'it', le petit v indique que les mots 'about glory, honour' au-dessus sont ajoutés à ce point. Donc on insert

```
<add place="below">about glory, honour</add>
```

apres ce mot.

- Ensuite il y a une *substitution*, c'est à dire, une rature (des mots 'battles, and glory of battles') et un ajout ('deeds or lands') qui constituent un seul acte: suivi d'une rature (de 'or land,'). On peut représenter ces actions ainsi:

```
<subst>
  <del rend="stoked">battles, and glory of battles</del>
  <add place="above">deeds or lands</add>
</subst>
<del rend="stoked">or land,</del>
```

- Le mot 'or' est raturé et les mots 'or anything about' sont ajoutés au dessous du ligne. On représente cela ainsi:

```
<subst>
  <del rend="stoked">or</del>
  <add place="below">or anything about</add>
</subst>
```

- Le mot 'majesty' n'est pas très lisible, et pourrait être erroné. On peut signaler cela avec l'élément `<unclear>` ainsi:

```
<unclear reason="scribbled">majesty</unclear>
```

- Le mot 'dominion' n'est pas présent dans le manuscrit: Owen écrivait simplement 'domin'. On peut traiter cela comme une abbréviation ainsi:

```
<choice>
  <abbr>domin</abbr>
  <expan>domin<ex>ion</ex>
</expan>
</choice>
```

- The mot 'whatever' a été raturé à la fin de la phrase: encodez cela ainsi:

```
<del rend="stoked">whatever</del>
```

- Votre représentation de ce deuxième paragraphe devrait ressembler à ceci:

```
<p> Nor is it <add place="below">about glory, honour,</add> about <subst>
  <del rend="stoked">battles, and glory of battles</del>
  <add place="above">deeds or lands</add>
</subst>
```

```

<del rend="stoked">or land,</del>
<subst>
  <del rend="stoked">or</del>
  <add place="below">or anything about</add>
</subst> any might, <unclear reason="scribbled">majesty</unclear>,
<choice>
  <abbr>domin</abbr>
  <expa>domin<ex>ion</ex>
</expa>
</choice> or power <lb/>
<del rend="stoked">whatever</del> except War. </p>

```

6 Troisième paragraphe

- Au debut du 3ème para, il y a un ajout (Above all its), dans lequel le dernier mot (its) a été raturé. Ceci est suivi d’une substitution : ‘This book’ a été remplacé par ‘I am’.

```

<add place="left">Above all </add>
<del rend="stoked">Its </del>
<subst>
  <add place="above">I am </add>
  <del rend="stoked">This book</del>
</subst>

```

- Il y a un autre ajout, au dessus du ligne, qui a été raturé, et qui contient un mot guère lisible mais qui pourrait etre lu ‘certainly’ :

```

<add place="above">
  <del rend="stoked">
    <unclear>certainly</unclear>
  </del>
</add>

```

- Le mot ‘My’ a été ajouté à gauche, le mot ‘Its’ a été raturé, le mot ‘The’ a été ajouté et raturé :

```

<subst>
  <add place="left">My </add>
  <del rend="stoked">It's </del>
</subst>
<add place="above">
  <del rend="stoked">The </del>
</add>

```

- Les mots ‘of’ et ‘it’ ont également été raturés.

Voila notre proposition pour l’encodage de ce 3eme paragraphe:

```
<add place="left">Above all </add>
<del rend="stoked">Its </del>
<subst>
  <add place="above">I am </add>
  <del rend="stoked">This book</del>
</subst> is <add place="above">
  <del rend="stoked">
    <unclear>center</unclear>
  </del>
</add> not concerned with Poetry.
<lb/>
<subst>
  <add place="left">My </add>
  <del rend="stoked">It's </del>
</subst>
<add place="above">
  <del rend="stoked">The </del>
</add> subject <del rend="stoked">of</del> is War, and the pity of
<del rend="stoked">it</del> War. <lb/> The Poetry is in the pity.
```

7 Quatrieme et cinquième paragraphes

Le quatrième para commence avec plusieurs ratures indépendants, et une substitution. Ensuite il y a un ajout ('to this generation') contenant un autre ajout 'past', qui est raturé..

Mais sans doute vous arrivez à comprendre le jeu, donc nous vous laissons jouer! N'oubliez pas que vous avez la possibilité à tout moment de visualiser l'effet de votre balisage, en vous servant du mode Author, mais nous vous conseillons de rester en mode Text pour effectuer votre balisage.

Le fichier `preface-final.xml` contient une version complète si vous vous renoncez à le compléter vous même!