3 – Structure d'un texte TEI

Tout texte conforme à la TEI comporte

1. un en-tête TEI (balisé comme un élément <*teiHeader*>,
2. la transcription du texte lui-même (balisé comme un élément <*text*>).

L'en-tête TEI contient des informations analogues à celles que l'on trouve sur la page de titre d'un texte imprimé. Il contient jusqu'à quatre parties :

1. une description bibliographique du texte électronique ;
2. une description de la manière dont il a été codé ;
3. une description non-bibliographique du texte (le « profil » du texte) ;
4. un historique de révision.

L'en-tête est décrit avec plus de détails dans la section 20 « Page de titre électronique ».

Un texte TEI peut être unitaire (une œuvre isolée) ou composite (un recueil d'œuvres, comme une anthologie). Dans un cas comme dans l'autre, le texte peut éventuellement comporter des pièces liminaires ou des annexes. Entre les deux se trouve le corps du texte qui, dans le cas d'un texte composite, peut comporter des groupes, chacun contenant encore des groupes ou des textes.

Un texte unitaire sera codé à peu près comme suit[1] :

```
<TEI.2>
    <teiHeader> [informations contenues dans l'en-tête TEI ]
    </teiHeader>,
    <text>
        <front>[ textes préliminaires... ] </front>,
        <body>[ corps du texte... ] </body>
        <back> [annexes... ] </back>
     </text>
</TEI.2>
```

1. Dès cet exemple, on remarquera que les diverses lignes de ce texte-source sont « indentées », c'est-à-dire qu'elles sont en retrait les unes par rapport aux autres, ce qui exhibe la structure sous-jacente ; ce n'est bien sûr pas une obligation, mais une bonne façon de coder les textes-sources ! {NdT}

Un texte composite peut également comporter des pièces liminaires et des annexes. Entre les deux, on trouve éventuellement un ou plusieurs groupes de textes, chaque groupe ayant éventuellement ses propres pièces liminaires et annexes. Un texte composite sera donc codé à peu près comme suit :

```
<TEI.2>
   <teiHeader> [ informations contenues dans l'en-tête
                 du texte composite ] </teiHeader>
   <text>
       <front> [ textes liminaires du texte composite ]
       </front>
       <group>
          <text>
              <front>[textes liminaires du premier texte ]
               </front>
              <body> [corps du premier texte ]  </body>
              <back>[ annexes du premier texte ] </back>
           </text>
           <text>
              <front>[ textes liminaires du deuxième texte ]
              </front>
              <body> [ corps du deuxième texte ] </body>
              <back> [ annexes du deuxième texte ] </back>
            </text>
            [ autres textes ou groupes de textes ici ]
        </group>
        <back> [ annexes du texte composite ] </back>
    </text>
</TEI.2>
```

Il est également possible de définir un ensemble de textes TEI, chacun avec son propre en-tête. Un tel recueil est appelé « corpus TEI », et peut lui-même avoir un en-tête :

```
<teiCorpus>
  <teiHeader> [ information contenue dans l'en-tête du corpus ]
  </teiHeader>
  <TEI.2>
     <teiHeader>[ information contenue dans l'en-tête
                  du premier  texte ] </teiHeader>
     <text> [ premier texte du corpus ]  </text>
   </TEI.2>
   <TEI.2>
   <teiHeader>[ information contenue dans l'en-tête du
```

```
                deuxième  texte ] </teiHeader>
   <text> [deuxiême texte du corpus ]  </text>
 </TEI.2>
</TEICorpus>
```

Cependant, il n'est pas possible de créer un ensemble de corpus, c'est-à-dire plusieurs éléments <*teiCorpus*> combinés les uns avec les autres et traités comme un objet unique. Ceci représente une des limites de la version actuelle des recommandations de la TEI.

Dans la suite de ce document, nous discutons surtout de structures de texte simples. Dans chaque cas, la présentation consiste en une courte liste des éléments TEI pertinents, avec une brève définition de chacun, suivies de la définition des attributs spécifiques à cet élément. Dans la plupart des cas, de brefs exemples sont également fournis.