

Talk 2: Overall TEI Structure

TEI @ Oxford

14 February (Valentine's Day)

Default Text Structure

All TEI documents are structured in a particular manner. This section attempts to describe the different variations on this as briefly as possible.

Structure of a TEI Document

There are two basic structures of a TEI Document:

- `<TEI>` (TEI document) contains a single TEI-conformant document, comprising a TEI header and a text, either in isolation or as part of a `teiCorpus` element.
- `<teiCorpus>` contains the whole of a TEI encoded corpus, comprising a single corpus header and one or more TEI elements, each containing a single text header and a text.

TEI basic structures (1)

```
<teiCorpus xmlns="http://www.tei-c.org/ns/1.0">
  <teiHeader>
    <!-- required -->
  </teiHeader>
  <TEI xmlns="http://www.tei-c.org/ns/1.0">
    <!-- required -->
  </TEI>
  <!-- More <TEI> elements -->
</teiCorpus>
```

TEI basic structures (2)

```
<TEI xmlns="http://www.tei-c.org/ns/1.0">  
  <teiHeader>  
    <!-- required -->  
  </teiHeader>  
  <facsimile>  
    <!-- optional-->  
  </facsimile>  
  <sourceDoc>  
    <!-- optional -->  
  </sourceDoc>  
  <text>  
    <!-- required if no facsimile or sourceDoc-->  
  </text>  
</TEI>
```

<text>

What is a text?

- A text may be unitary or composite
 - unitary: forming an organic whole
 - composite: consisting of several components which are in some important sense independent of each other
- a unitary text contains
 - optional front matter
 - <body> (required)
 - optional back matter

<text>

What is a text?

- A text may be unitary or composite
 - unitary: forming an organic whole
 - composite: consisting of several components which are in some important sense independent of each other
- a unitary text contains
 - optional front matter
 - <body> (required)
 - optional back matter

<text>

What is a text?

- A text may be unitary or composite
 - unitary: forming an organic whole
 - composite: consisting of several components which are in some important sense independent of each other
- a unitary text contains
 - optional front matter
 - <body> (required)
 - optional back matter

Composite texts

A composite text contains

- optional front matter
- `<group>` with `<text>` inside (required)
- optional back matter

A corpus is a collection of text and header pairs that also has its own header.

`<group>` tags may self-nest.

TEI text structure (1)

```
<text>
  <front>
  <!-- optional -->
    </front>
    <body>
  <!-- required -->
    </body>
    <back>
  <!-- optional -->
    </back>
</text>
```

TEI text structure (2)

```
<text>
  <front>
  <!-- ... -->
  </front>
  <group>
    <text>
      <body>
        <p>...</p>
      </body>
    </text>
  </group>
  <back>
  <!-- ... -->
  </back>
</text>
```

Another Grouped Text Example

```

<TEI xmlns="http://www.tei-c.org/ns/1.0">
  <teiHeader>
    <!-- header information for the whole collection -->
    </teiHeader>
    <text>
      <!-- optional front matter -->
      <group>
        <text>
          <!-- optional front matter -->
          <body>
            <!-- First Body -->
            </body>
          </text>
        <text>
          <!-- optional front matter -->
          <body>
            <!-- Second Body-->
            </body>
          </text>
        </group>
      </text>
    </TEI>
  
```

Next

And now we're going to break for lunch, but when you come back we'll talk about the kinds of metadata you can put in the the `<teiHeader>` element!