

Practical Exercise: A Slightly Better TEI Document

TEI@Oxford

March 2011

1 Second steps with Oxygen

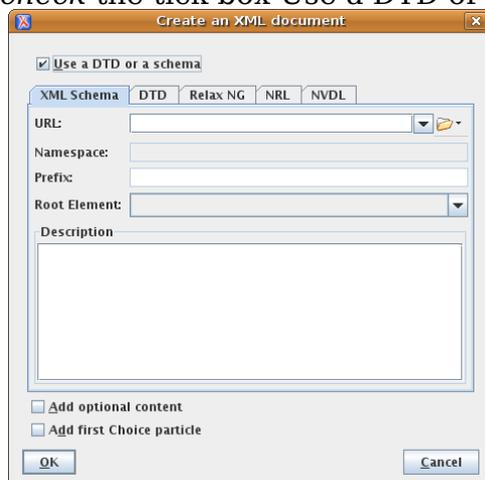
Now that you are a bit more familiar with Oxygen, we will use it to make a complete TEI document, with all its parts. We will also check its validity, and start enriching its mark up.

1.1 Creating a valid TEI document

Open Oxygen again (if necessary) and as before choose the File/New menu item.

Oxygen provides two ways of specifying the kind of document you want to create. You can use one of the predefined *templates* or you can specify directly that you want to create an XML document and supply a schema yourself. Oxygen comes with many predefined templates, including some for several popular TEI schemas, which you may wish to explore at your leisure.

For this exercise, however, we suggest that you select the default 'XML Document' as before. This time when you see the "Create an XML Document" dialog, *check* the tick box Use a DTD or a schema



You need to specify the *schema language* for your schema, by choosing one of the tabs XML Schema, DTD, or Relax NG. It doesn't matter much which you choose, but we recommend the last. Then click on the little yellow folder icon to the right of the empty box labelled URL, choose Browse for local file, and navigate to the location where the required schema file is stored.

For this exercise we've prepared a schema called `tei_ipp` which you should find in your Work folder, as an XSD, DTD, RNG or RNC file. Oxygen will choose the one it wants, depending on the schema language you chose earlier.

Press OK to continue.

Oxygen does its best to take advantage of the information in your schema file, both by providing you with online help, and by adding any mandatory elements for you. It also has a slightly distressing habit of trying to hide

1 SECOND STEPS WITH OXYGEN

the tags from you by default. If you are now looking at a screen like this:



all you need to do is to select the "Text" mode at the bottom left in order to see a more reassuring display, probably like this:



(The details of the display may be slightly different depending on the validator you are using; if you don't see any red lines, click on the red tick on the menu bar in order to force validation of your document)

To make the document valid, you need to insert some minimal metadata. At the very least, you should supply a non-empty `<title>` element, and put a few remarks into a paragraphs inside the `<publicationStmt>` and `<sourceDesc>` elements. It's up to you how much detail you feel like adding, but we suggest you should aim to produce something like this:

```
<?xml version="1.0" encoding="UTF-8" ?>
<?oxygen
RNGSchema="file:/home/lou/Public/TEI-0X/Talks/2010-07-oxford/materials/Work/tei_ipp.rng"
type="xml" ?>
<TEI xmlns="http://www.tei-c.org/ns/1.0">
  <teiHeader>
    <fileDesc>
      <titleStmt>
        <title>Samples from Punch: vol 147</title>
        <principal>Encoded by YOUR NAME HERE</principal>
      </titleStmt>
      <publicationStmt><p>Unpublished exercise</p></publicationStmt>
      <sourceDesc>
        <p>Texts supplied at OUCS TEI Summer School 2010</p>
      </sourceDesc>
    </fileDesc>
    <revisionDesc>
      <change when="2010-07-14">Made header</change>
    </revisionDesc>
  </teiHeader>
  <text>
    <body>
      </body>
    </text>
  </TEI>
```

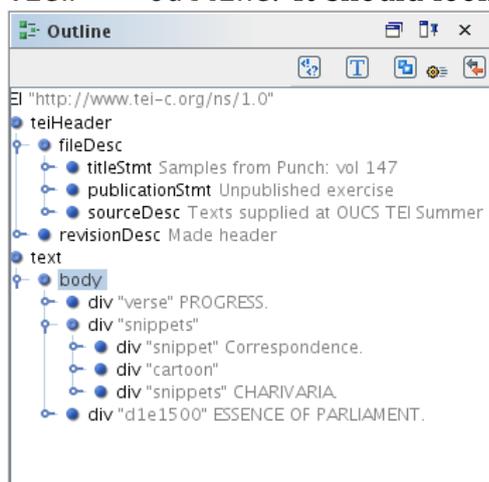
1.2 Adding some content

If your document doesn't have a <body> element inside the <text> element, then add it. But now your document has a <body> but there is nothing in it. Why not insert the <div> element containing the poem which you made earlier? If you still have that open, you can cut and paste between windows in Oxygen. If not, use the Document/File menu to insert the file progress.xml you saved earlier. And if you didn't quite finish the previous part of the exercise we have thoughtfully prepared a suitable file for you: it's in the file called spoilers/progress.xml on your USB key.

1.3 Add some more sections

There is also a file called snippets.xml in your Work directory. Open this with Oxygen, and cut and paste all the <div>s which it contains after the verse <div>. The document should still be well-formed.

In the same way, insert the <div> containing Toby's *Essence of Parliament* in the right place in your document. A lightly marked up version of this is available in the file called essence.xml in your Work directory. Use the 'Outline view' in Oxygen (lower left panel) to check that you have the right structure for your document. If you have turned this off you can turn it back on by going to Perspective -> Show View -> Outline. It should look something like this:



1.4 Validate the file

Up in the menu bar, look for the validation button , and click on it. This will validate your file against the tei_ipp schema we specified when you created it. (You may have noticed in text view that there is an incantation at the top of the file which begins like this: <?oxygen RNGSchema=. This is an Oxygen-specific processing-instruction which indicates the location of the schema to be used to validate the file.) You can configure the 'validation scenario' Oxygen uses in many other ways, which we won't detail in this tutorial.

1.5 Explore!

As you've seen, Oxygen is very good at suggesting elements you might want to use to tag parts of the document. Just type a < at any point, to see a drop-down list of elements legal at that point. For a list of the attributes an element can take, insert a space inside its start-tag. In each case you will also see a small amount of information for each element and each attribute. For full details however, you need to consult the TEI Guidelines.

2 A READING KNOWLEDGE OF THE TEI

Here are some suggestions for ways you could improve the mark up of the material added:

- mark the page boundaries, using `<pb/>`
- mark as `<title>` elements all the titles currently tagged as `<hi rend="it">`)
- mark up the dialogue between Greece and Turkey using `<sp>` elements
- decide what to do about the paragraphs of stars: maybe they should just go? maybe they should become milestones? maybe the paragraphs they separate should become list items?

Make sure you leave your document valid (check for a happy green square in the upper-right corner!) and remember to save the file you've made in the Work folder on your USB key. Call it `exercise-1b.xml` — we will use it again in a later exercise.

2 A reading knowledge of the TEI

You need to be familiar with reading the TEI Guidelines. Start by visiting <http://www.tei-c.org/Guidelines/P5/> and make sure you can

- See the PDF version of the entire Guidelines
- Browse the chapters of the Guidelines in English
- (Browse the chapters of the Guidelines in another language, if you need)
- Navigate the element catalogue at <http://www.tei-c.org/release/doc/tei-p5-doc/en/html/REF-ELEMENTS.html> and follow the reference sections.
- Jump around the reference materials following links

We've include a copy of the whole of the latest release of the TEI Guidelines on your USB key as well, so you don't need to be online to read them.