Exercise 7: Facsimile linking

Dot Porter

March 2009

In this exercise we will practice linking a transcription to page images representing the transcribed text. The first is a relatively simple method using direct links to image files. The second is a more advanced method using `<facsimile>`.

1 Setting up the templates

Open *whitman_transcription.xml*, which you worked on yesterday, in oXygen. Save it as *whitman_transcription-facs.xml*. (Alternatively, open *whitman_transcription-completed.xml*, located in the *extras* folder.)

When the module for "Representation of Primary Sources" (transcr) is enabled, the `<facsimile>` element and all its children are enabled as well, and the **@facs** attribute is available as a global attribute. This means that any tag in a TEI file may be linked either directly to an image file, or to a Facsimile list that is in turn linked to image files.

2 Direct linking

The most simple method for linking a TEI file to image files is to place links directly on elements using the **@facs** attribute. In whitman_transcription-facs.xml, place a @facs attribute on each `<pb>` tag. For the value of @facs, insert the relevant image file name (for page 27, this would be "whitman_27.jpg").

`<pb n="27" facs="whitman_27.jpg"/>`

Do this for all three pages. Save your file.

3 Displaying the links using XSLT

There are two main types of stylesheets we use to process TEI:

- Cascading style sheets (CSS) process TEI for *display* but does not change the underlying code

- Extensible Stylesheets (XSL) changes the TEI into something else either for display (in which case it is translated into HTML) or for some other purpose (in which case it might be translated into different TEI or another XML sublanguage)

So far in this workshop we have been processing using CSS. In this exercise, in order to take advantage of the facsimile linking we need to process the TEI into HTML, with hypertext links from the page numbers (processed out of the `<pb/>` tags) to the image files (which will open in new tabs or new windows, depending on your browser). To do this, we will use oXygen's built in XSLT processor.

In oXygen, go to the top menu Document > Transformation > Configure Transformation Scenario. On the "Configure Transformation Scenario", click "New". Name the new scenario

"Whitman_transcription-facs". At the "XSLT" tab, the value for XML URL should by default be "${currentFileURL} – do not change this. Next to XSL URL, click the plain folder icon to browse for a local XSL file. Browse to the "styles" folder and select "facsimile.xsl". For "Transformer" select "Saxon-B 9.1.0.3".
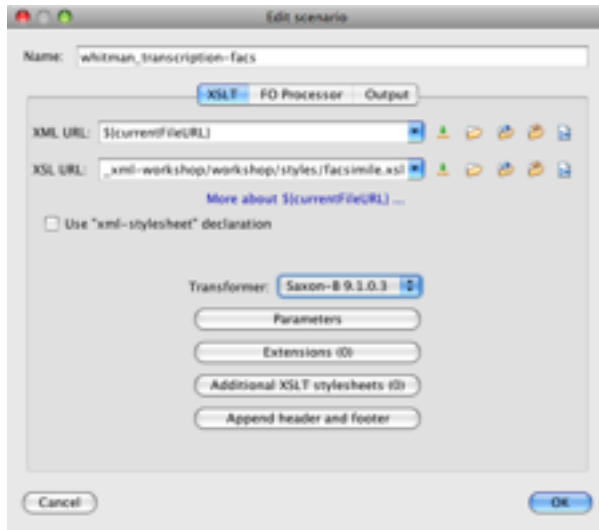
On the "Output" tab, select "Save As" and browse to the folder *output*, saving the file as *whitman_transcription.html*. Check "Open in browser" and "Saved file". Click "OK" in the bottom right corner, and the "Edit scenario" window will close.
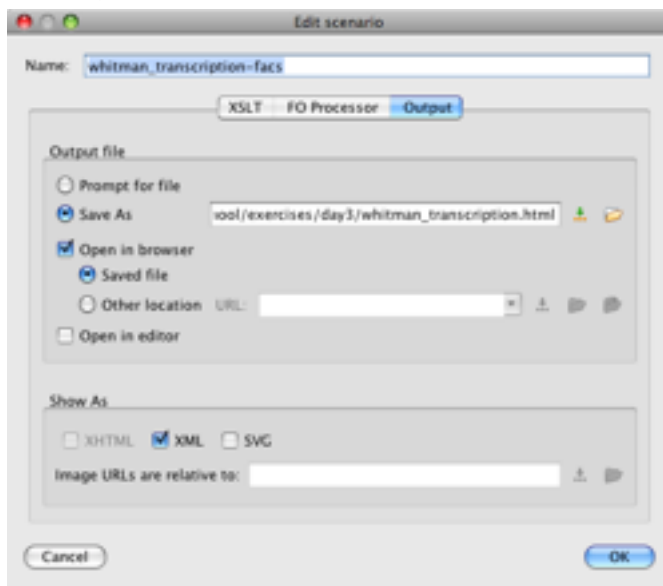
At the "Configure Transformation Scenario" window (which should still be open), click "Transform Now". The new file will open in a browser window.

You'll note that it looks very similar to the view through the CSS file. However if you view the page source, you'll see that the code has been changed completely to HTML.

Hover your mouse over one of the page numbers, and click it. A new tab will open, displaying the image file for that page.

4   Advanced Linking using `<facsimile>`

Moving on, let's try a slightly more advanced method for linking a TEI file to image files.

The general method for linking a TEI file to image files is through **`<facsimile>`**. `<facsimile>` is on the same level as `<text>`, and is used to create an index of all images referenced by a TEI file.

Insert a `<facsimile>` tag between the end of the TEI Header and the opening `<text>` tag.

```
68      </teiHeader>
69      <facsimile></facsimile>
70 ▽    <text>
71 ▽      <body>
```

Inside the facsimile, nest `<graphic>` tags linking to the three image files that correspond to the pages. First insert the `<graphic>` tags, one after the other. Then provide each with a unique xml:id. In the example we use "whitman_pagenumber" but you can use any identification system you like.

```
69 ▽    <facsimile>
70          <graphic xml:id="whitman_27"/>
71          <graphic xml:id="whitman_28"/>
72          <graphic xml:id="whitman_42"/>
73      </facsimile>
```

Next, we need to provide pointers from the `<graphic>` tag to the image files themselves. `<graphic>` uses **@url** for these links. Since the image files and the XML file are in the same directory, we need only provide the image file names, which are in the format *whitman_pagenumber.jpg*. Your finished facsimile section should look similar to this:

```
69 ▽    <facsimile>
70          <graphic xml:id="whitman_27" url="whitman_27.jpg"/>
71          <graphic xml:id="whitman_28" url="whitman_28.jpg"/>
72          <graphic xml:id="whitman_42" url="whitman_42.jpg"/>
73      </facsimile>
```

You now need to link you page breaks to the representative `<graphic>` in the facsimile section. Replace the values of @facs on the `<pb/>` tags (which had been pointing directly to the image files) with pointers to the @xml:id values of the representative `<graphic>`.

Be sure to save your file.

5   Transform the file again.

Since you have already configured the transformation scenario, you need only click the "Apply Transformation Scenario" button (⏺) in the top menu. The resulting page will look and function exactly as the previous one.

6   Linking to areas on an image

Finally, we are going to go a little more advanced and associate each deletion and addition with an area of the image file.

To get the coordinates we'll use the Online HTML Image Mapping Tool (http://www.image-maps.com). Follow the directions to upload the image file for page 27. Create a new rectangle and place it around the "a" that has been crossed out at the start of line 12. It does not matter what values you put for link and title/alt. Save the rectangle and click "Get Your Code" in the right-hand menu. In the next window click "Show Your Code", then scroll all the way down to "Your HTML Code". In that window you will see a lot of code, including `<area>` tags. Find the area tag that matches the values you put in for title/alt. Note that this `<area>` tag has an attribute @coords.

Keep that window open for a moment and return to whitman_transcription-facs.xml in oXygen.

Your `<facsimile>` currently contains a series of `<graphic>` tags. We are now going to associate each graphic with a surface (in this case, corresponding to each individual page) and then subdivide that surface into zones (in this case, corresponding to deletions and additions in the text).

Place a `<surface>` element around each `<graphic>`.

After the graphic representing page 27, but within the `<surface>` element, place a single zone. Give the zone an xml:id based on the xml:id of the graphic – in the example, we use *whitman_pagenumber-number*. **Note: it is vital that you base the zone id on the graphic id or else the stylesheet will not work correctly.**

Return to the browser window and find the `<area>` tag. The value of @coords will be four numbers, separated by commas. Those numbers correspond to the coordinate values on `<zone>` as follows:

@html:coords="@ulx,@uly,@lrx,@lry"

Cut and paste the values from the @coords attribute on `<area>` to the corresponding coordinate attributes on `<zone>` as indicated.

Now go into the transcription and find the deletion on page 27. Add a @facs attribute that points to the new `<zone>`. Your code will look something like this:

```
<surface>
    <graphic xml:id="whitman_27" url="whitman_27.jpg"/>
    <zone xml:id="whitman_27-1" ulx="35" uly="498" lrx="73" lry="537"/>
</surface>
»<del rend="overstrike" facs="#whitman_27-1">a</del> the locust t
```

Save your file!

7   Transform this file to create an HTML image map.

To successfully transform this new code, we'll need to open the *facsimile.xml* file in oXygen and open some of the stylesheet code that has been commented out. Open facsimile.xml (in the *styles* folder) and scroll down until you find the large section of green text (around line 125). Delete the `<!-` and `->` that surround the entire block. Now find the block immediately above the block you and just uncommented (it starts with `<xsl:template match="tei:pb">`). Comment out that entire block (down to `</xsl:template>`, using `<!-` and `->`. Save the stylesheet.

Go back to whitman_transcription-facs.xml and click the "Apply Transformation Scenario" button in the top menu. The transformed file will open in the browser, and you will see that now the image files are present as well in the main view. If you hover your mouse over the deletion on the image, you should see a box appear with the text, pulled from the transcription.

8   Repeat for more pages

Return to oXygen and to the Online HTML Image Mapping Tool, and repeat for pages 28 and 42 what you have done for page 27. Don't forget to save often!