

Transcribing a historical source

Magdalena Turska

September 2014

1 Handy oXygen tricks

Enclose the selection with a tag:

- highlight the characters which you want to tag
- type CTRL+E to display the menu of available tags
- type 'pe' (for <persName>) or 'pl' (for <placeName>) and then press RETURN

Split the long chunk of text into a sequence of elements of the same kind:

- highlight the long chunk and wrap it with your desired element (say <p>)
- move the cursor to a place within that chunk that is the start of next element of the same kind
- type ALT+SHIFT+D to split elements (it inserts closing and starting tags at the cursor position)
- repeat as many times as needed

If you ever forget the key combination to perform the trick try the right-click and see what's in the Refactoring section

It can help a lot to format and indent your work automatically via CTRL+SHIFT+P or clicking the Format and indent icon.

2 The Dantiscus Project

In this and subsequent exercises, we will use the TEI to mark up historical source documents: the letters from the correspondence of the 16th century poet and diplomat Ioannes Dantiscus with the royal house of Poland - King of Poland etc Sigismund I Jagiellon and his wife Bona Sforza. Ca 500 letters between this three important people still exists along more than 5000 other pieces of correspondence with numerous members of European royal houses and other members of Respublica Litteraria of the time. The correspondence is a long-term project of the Laboratory for Editing Sources of the University of Warsaw, by whose kind permission we are using them in this tutorial.

In your Work folder there is a folder called Dantiscus which contains digital images of a couple of letters. You may like to look at them quickly to see the kind of material you will be working on. (The University of Warsaw also makes the whole set available from its web site <http://dantiscus.al.uw.edu.pl>)

In this first session, the aim is to make a transcription which covers:

- the written text and its organization
- some editorially significant aspects such as deletions or additions
- detailed indexing of proper names of persons and places
- expansion of abbreviations
- provision of explanatory notes etc.

We will begin with the letter of Bona Sforza to Ioannes DANTISCUS sent from Cracow 1524-04-29, which you'll find in the files BCz_1601p538.jpg(address) and BCz_1601p537.jpg.

3 Structural tagging of a letter

- Open oXygen and make a new XML document. In the Choose the document type dialogue find All[TEI P5] to associate TEI schema with your file.
- Leave the header alone for the moment. Inside the <body>, add first a <pb> element, to indicate that a new page starts here. Supply the page number (BCz, 1601, p. 538) as the value of the @n attribute, and the name of the image file corresponding with this page (BCz_1601p538.jpg) as the value of the @facs attribute .
- Next, add a <div> element to the body of the <text> element. Give it a @type attribute with the value address
- Add a <p> element to the <div> after the <head> to contain the text of the address.
- Add another <pb> and <div> with a <p> inside to hold the text of the letter itself and link the pb element with the other file BCz_1601p537.jpg

To save you having to type too much, we've thoughtfully prepared a plain text transcription of this page in the file IDL5823.txt. You can cut and paste from this file, or insert the whole of it directly into your encoded version using the File -> Insert File command on the Document menu.

Whichever method you adopt, you'll need to be careful about the ampersands in the text. An XML document cannot contain an ampersand directly: it has to be represented indirectly by means of the short cut & . But you probably knew that already.

4 Editorial markup

Here are some interventions you may wish to mark in the text:

- This text is full of abbreviations: The first line contains 8 words in total and 7 of those used abbreviation in some way. In the plain text transcription we provided you with expanded version of all the words where () mark the expanded text. You should replace them with the appropriate TEI element: <expan> and <ex> to mark whole expanded words and letters that weren't present in the source, respectively
- There is one place on line 7 where the word operante was changed into operaturum after the mark standing for word pro was already written. What's going on is the letter n that was represented by a mark above the a stays behind even though it's not longer needed, and letters ur are squeezed in on the letter e and in small space before pro. That's one hypothesis. The other would be saying it's quite <unclear> what's going on in that crowded space. You could tag the substitution with <subst> <add> and elements while the superfluous n requires a <choice> with <sic> and <corr>, and the hard to read fragments could be marked with <unclear>

5 Editorial markup

And yet some more interventions you may wish to mark in the text:

- There is a big, roughly round hole in the document that affects the beginnings of 4 lines of the text. Similarly in the address there are words missing. In these places the text needs to be supplied in transcription (signalled with [] brackets in the textual transcription). You should replace that with <supplied> and <gap> elements.
- From the middle of line 9 starts the dating sequence: you can mark that with <date> element. You may wish to provide a normalised date such as 1524-04-29 as the value of the @when attribute of the <date>. What are the benefits of doing that?

6 Hands

You can probably note the difference in the handwriting between the message of the letter and the signature. It looks like the letter was penned by a secretary and the queen only signed it afterwards.

If you think it's worth preserving this kind of information you need to use `<handShift>` element to mark the change of hands. It's an empty element, a bit similar to `<pb>` &co that marks the point where the change happened - note that it does not wrap the passages by the same hand

```
<handShift hand="#H1"/>
```

To be able to refer to a hand by it's id like this we need to add respective `<handNote>`s to a `profileDesc` section of the `teiHeader`

```
<profileDesc>
  <handNotes>
    <handNote xml:id="BS1">hand of Bona Sforza</handNote>
    <handNote xml:id="H1">H1 Unidentified secretary of the Bona Sforza</handNote>
  </handNotes>
</profileDesc>
```

7 Semantic markup

This document has names! If you still have the time you can go through it carefully marking up the names of people with `<persName>` and the names of places with `<placeName>`. You can do this quite quickly if you remember the oXygen trick of enclosing the selection with a tag with CTRL+E combination

We will go through more interesting things to do with people and places tomorrow

Don't forget to save the file you have created! You might continue to work on it in the next couple of exercises.

8 Self-assessment

Check if you understand some of the core principles of this exercise by answering the following questions:

- If you want to indicate that an abbreviation and expansion (or correction and error) are linked, what element do you wrap them in?
- If you want to indicate that an addition and deletion are one editorial act, what do you surround them with?
- How do you show that an addition is subsequently deleted?