transcription: more than meets the eye

a workshop by Catherine Addington, ca2bb@virginia.edu

Materials

➢ presentation slides (below)
➢ access to a chalkboard or whiteboard
➢ scratch paper and pencils for participants to use for transcription (may also transcribe on a computer/tablet if preferred)

Scope

➢ 60 minutes
➢ Ideal for a small group, under 20 participants
➢ Can be expanded or contracted by removing or adding sample exercises

Learning Objectives

After this workshop, participants will:

1. Be familiar with the process of print-to-digital transcription and able to articulate the intellectual labor involved
2. Have inductively arrived at the basic concepts of textual scholarship, especially the distinction between text and document
3. Understand the definitions of the three major methods of transcription (diplomatic, linear, and timed) and what factors should be taken into consideration when choosing a strategy
4. Be aware of TEI markup as a possible tool for their own projects

Lesson Plan

Introduction

• Begin by introducing the objectives for the workshop.
• Hint at the theoretical discussion to come but focus on project implementation.

Sample #1: Declaration of Independence

• Ask participants to work in pairs transcribe the heading. Explain that they can use either paper or a word processor depending on the tools they have available, but the idea behind this activity is that our transcription is digital.
• Discuss different approaches. Any discrepancies among participants’ strategies?
• Show Archives.gov transcription. Ask participants to note the differences from the original manuscript. What’s different? What’s missing? What’s added?
  ○ Case different
  ○ Font changed
  ○ Line break eradicated
  ○ Capitalization preserved
  ○ Spacing/alignment altered
• To the extent possible, guide the conversation toward articulating the principle behind this transcription. What is it aiming at? Who is the audience? What is this being used for?
Sample #2: Notes on the State of Virginia

- Ask participants to work in pairs to transcribe the fragment.
- Discuss different approaches. Any discrepancies among participants’ strategies?
  - What did we do with the unusual characters (long s, connected ct)? Why?
  - Did anyone make note of the smudge? Why or why not?
- Show Notes on the State of Virginia transcription. Ask participants as a group to note the differences from the original manuscript. What’s different? What’s missing? What’s added?
  - Structure maintained, far more than in the declaration example
  - Spacing truncated in heading
  - Marginalia displayed on left instead of right
  - Outdated characters made readable
- To the extent possible, guide the conversation toward articulating the principle behind this transcription.
  - What is it aiming at? Who is the audience? What is this being used for?
  - Give context that Scholars’ Lab digitized Jefferson’s personal copy, go to website as example of completed digital edition.

Sample #2B: Notes on the State of Virginia, page 1 to 2

- Introduce the complication of a page break. Explain that “men-” at bottom of page 1 to “mentioned” at top of page 2 was print convention at the time for page breaks.
- How do we transcribe this? Do we care where the pages originally were if we’re reading online? Is our way of including print marks to make them functional or to represent them? Guide them through making a list of pros/cons as a group.
- Go to the next slide with the Scholars’ Lab edition’s solution to this.
  - Discuss what we think of this solution as a group.
  - Show on website that these page breaks are relevant to their edition because of page-by-page scans in margins as well. (Think about your audience!)
Sample #3: First line of Pride & Prejudice

- Ask participants to work in pairs to transcribe the fragment.
- Ask someone to write theirs on the board.
  - Have them briefly explain their strategy.
  - Do they keep small caps or not? Do they keep hyphens/line breaks or not? Why?
- Then ask the group: what is the relationship between these two things (i.e. the original on the screen and the transcription on the board)?
  - We’re a culture of materialists so we hate this but at some point we have to admit that this is a question about metaphysics.
  - This is best demonstrated rather than said.
- Make changes to the transcription. First a punctuation mark, then a word, then a sentence. Ask for reactions from the group each time.
  - What level of intervention is okay?
  - How does each intervention alter the relationship between the original and the transcription?
  - At what point does the transcription lose its relationship to the original and become something different?
  - Why do we feel the impulse to maintain the relationship between these two independent products? → introduce the concept of text vs. document.
- Introduce TEI as transcription method.
  - Show P&P first line markup. What is the document here? What is the text?
  - If we cared about preserving more elements of the print document here, what are some things we could add? (Linebreaks could be noted.)
  - Explain that there are also tags for other units: pages, chapters, marginalia, columns...it can get as complicated or as simple as you want.
Conclusion

- Show Lavagnino quote ("you need to believe in transcription"): there's no such thing as a complete representation.
  - You have to set limits for your project about how much transcription is actually necessary for your purposes.
  - Something is always going to be lost in remediation, and something is always going to be added. It's a question of setting priorities about what your audience needs.
- Introduce three theories of transcription: diplomatic, linear, timed. Ask participants to categorize our three samples. What strategy does each seem to use? (It's a continuum, none completely fits into any of the three.)
  - Declaration of Independence: closest to "linear," seeks to provide a readable text more than a digital twin to the document
  - Notes on the State of Virginia: closest to "diplomatic," though it does have plenty of linear elements. Preservation of marginalia and page breaks = this edition cares about representing the document.
  - Pride & Prejudice: if we were to transcribe our process of making changes to the original (i.e. transcribing the crossed-out words and insertions), that would be an example of "timed"
- How to choose a strategy, then? Show concluding quote from Robinson & Solopova on transcription as "fundamentally interpretive."
  - Choose your strategy (diplomatic, linear, timed, or some combination of the three) based on what your audience needs.
  - What information, if preserved, could prove useful and to whom?
  - Remind participants to be transparent about whatever methodology they choose.
- Share resources for further learning.
- Thanks / Q&A

Credits

This workshop was inspired by transcription exercises and discussions in Allison Bigelow and Rafael Alvarado's course SPAN 7559: Latin American Digital Humanities and Jim Ambuske and Loren Moulds's course HIST 5559: History in a Digital Age.

Samples for the above exercises were taken from Archives.gov, notes.scholarslab.org, and the BBC.

Sources: Lavagnino, 2006; Pierazzo, 2009; and Robinson & Solopova, 1993, all of which were found via the Lexicon of Scholarly Editing.

Other resources to be mentioned include the Text Encoding Initiative and FromThePage.
transcription
more than meets the eye

Catherine Addington
University of Virginia
ca2bb@virginia.edu
objectives

1. through a series of sample exercises, learn how to navigate decision-making in transcription projects
2. become familiar with the basic concepts of textual scholarship
3. acquire tools, resources, and models for implementing transcription projects of your own

warning: this workshop is nominally about transcription but secretly about metaphysics
IN CONGRESS, JULY 4, 1776.

The unanimous Declaration of the thirteen united States of America.
In Congress, July 4, 1776.

The unanimous Declaration of the thirteen united States of America, When in the Course of human events, it becomes necessary for one people to dissolve the political bands which have connected them with another, and to
QUERY I.

A

An exact description of the limits and boundaries of the state of Virginia?
QUERY I.

AN exact description of the limits and boundaries of the state of Virginia?
Mississippi, to latitude 36°. 30'. North: and on the South by the line of latitude last men-

tioned. By admeasurements through nearly the whole of this last line, and supplying the unmeasured parts from good data,
the West by the Ohio and Mississippi by the line of latitude last-mentioned. By admeasurements the unmeasured parts from
It is a truth universally acknowledged, that a single man in possession of a good fortune, must be in want of a wife.
“It is a truth universally acknowledged, that a single man in possession of a good fortune, must be in want of a wife.”

Chapter 1

It is a truth universally acknowledged, that a single man in possession of a good fortune, must be in want of a wife.
<head>Chapter I.</head>

<p><hi rend="smallcaps">It</hi> is a truth universally acknowledged, that a single man in possession of a good fortune, must be in want of a wife.</p>
In order to use the TEI approach you need also to believe in transcription. It is impossible for a transcription to reproduce the original object; it is always a selection of features from that object...

—Lavagnino, 2006
<table>
<thead>
<tr>
<th>Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>diplomatic</strong></td>
<td>“the goal of such transcription is the reproduction of the actual appearance of the document”</td>
</tr>
<tr>
<td><strong>linear</strong></td>
<td>“the goal of the transcription is the description of the text”</td>
</tr>
<tr>
<td><strong>timed</strong></td>
<td>“the goal of the transcription is the description of the process”</td>
</tr>
</tbody>
</table>

Source: Pierazzo, 2009
Transcription for the computer is a fundamentally interpretative activity, composed of a series of acts of translation from one system of signs (that of the manuscript) to another (that of the computer).

Accordingly, our transcripts are best judged on how useful they will be for others, rather than as an attempt to achieve a definitive transcription of these manuscripts.

—Robinson & Solopova, 1993
resources

➢ TEI: Text Encoding Initiative
documentation & criticism

➢ Lexicon of Scholarly Editing
multilingual index of
basic terminology

➢ FromThePage
storage & collaboration tool
thanks!

questions?

Catherine Addington
ca2bb@virginia.edu