On beyond Paul the Simple

David J. Birnbaum
University of Pittsburgh
2011-08-20 Sofia

Outline
1. Goals and premises
2. History and features of 2008 pilot edition of the “Vita of Paul the Simple”
3. Requirements for new edition
4. What can be automated and what cannot
5. What will change for the upcoming edition of the entire Codex Suprasliensis, and why

Goals
- Support for research
  - Textological
  - Linguistic
- Accessibility
- Preservation
- Support for teaching

Premises
- An electronic edition is not just a paper edition on screen
  - Multipurposing (views, indices, etc.)
  - Integration of resources
  - Dynamic content
- Project as workstation
  - Each response should anticipate and facilitate whatever the user is likely to do next

Paul the Simple
- Genesis: Spring 2008, introductory graduate practicum on Old Church Slavonic, University of Pittsburgh
- Edition
  - http://clover.slavic.pitt.edu:8080/exist/paul/data/paul_main.html
  - Will move to: http://paul.obdurodon.org
- Principal reports
  - Presentation at the fourteenth International congress of Slavists (Ohrid, 2008)
Browse the text: main view

- Paragraphed OCS
  - Helsinki ASCII transcription converted to Unicode and corrected manually against facsimile in Zaimov/Capaldo edition
  - Paragraphing introduced by editors
- Linked to
  - Parallel Greek (from Zaimov/Capaldo edition)
  - Parallel English (new translation by practicum students)
  - Folio views (see below)
- Linguistic commentary
  - Word-for-word, scrolls and highlights
  - Back-linked
- Font-switcher (freely-distributed downloadable fonts)

Browse the text: folio view

- Interlinear OCS and Greek on left
  - Mouse-over linguistic commentary
- Facsimile on right
  - Magnifying glass
- Links to other folios
- Links to main view and searches
- Same OCS and Greek text and linguistic annotation as main view
  - Single source = multipurposing

Search “Paul the Simple”

- Menu populated dynamically on load
- Returns list of forms with linguistic annotations, manuscript locations, Greek parallels
Search the text: grammatical search

- Menu lists populated dynamically on load
- Faceted searching
- Returns list of forms with linguistic annotations, manuscript locations, Greek parallels

Search the text: full text search

- Wild cards / regular expressions
- `антони` retrieves all words that begin with the string `антони`
- Returns locations, Greek, parallels, lemmata
Requirements for new edition

- Technical requirements
- Access and distribution requirements

Technical requirements

- Internet based (no CD/DVD)
- OS-independent (Windows, MacOS, iOS, Linux)
- Browser based: current versions of
  - Windows: Internet Explorer, Firefox, Chrome, Opera, Safari
  - MacOS: Safari, Firefox
  - iOS: Safari
  - Linux: Firefox
- Standards based: Unicode, HTML, CSS (Javascript)
  - No Flash, Silverlight, etc.
  - No proprietary applications

Access and distribution requirements

- Publication on mirrored servers
  - Long-term preservation
  - Redundant access
- Sites accessible to all at no cost
- All source files available to all at no cost
  - Creative Commons Attribution Non-Commercial Share-Alike License
  - http://creativecommons.org/licenses/by-nc-sa/3.0/

What can be automated

- (“Automated” = machine-assisted)
- Transformation of Helsinki plain ASCII text to XML and Unicode
- Creation of multiple materials from a single source
  - Main (paragraphed) view, folio view, linguistic and full-text databases, user interfaces for databases
- Linking of OCS and Greek at the line level
- Linking of folio view transcription and facsimile image at the line level
- Linguistic markup (PROIEL)
  - (Reports [tabular, graphic, etc.])

What cannot be automated

- Proofreading and correction of errors in Helsinki corpus and Zaimov/Capaldo Greek texts
- Paleographic and other diplomatic annotations
- Linking of folio view transcription and facsimile image at the word level
  - Automated at the line level
- Linking of OCS and Greek at the word level
  - Automated at the line level

Changes introduced in new edition

1. Content: framework for linguistic annotation
2. Development: method of linguistic annotation
3. User interface: frames and panels
4. User interface: characters and font
5. User interface: image linking
6. User interface: mouseover
Content: linguistic annotation

• Then
  – Idiosyncratic citation forms (based on Swan OCS Grammar)
  – вьꙁьпи: for въꙁъпи, 3.p.sg.aor. of въꙁъпьь-
    j, agr.w. нава understood
• Now
  – Standard reference system
  – To be determined
• Why
  – Different goals, different audience

Development: linguistic annotation

• Then
  – Manually inserted lemma, grammatical annotations
• Now
  – PROIEL machine-assisted linguistic analysis and markup
• Why
  – Paul the Simple = 3 folios
  – Codex Suprasliensis = 285 folios
• But
  – Linking to Greek at word level must still be done manually

User interface: frames and panels

• Then
  – HTML frameset (frames)
• Now
  – CSS pseudo-frames (panels)
• Why
  – Better support, easier user interface, more mature CSS standard

User interface: fonts

• Then
  – Client-based fonts and font-switcher
    – Download and install fonts, select
• Now
  – CSS3 web fonts
• Why
  – Easier for technologically inexperienced users
    – The only way to render fonts on closed platforms (iOS)

User interface: images

• Then
  – Image linking at page level (side by side, no internal links)
• Now
  – Image linking at line level (transcription to image and image to transcription)
  – Under consideration
    – Image linking at word level
• Why
  – Easier user navigation

User interface: touch devices

• Then
  – Hover (CSS) or mouseover (Javascript) generates a different event than click
• Now
  – Touch-sensitive alternative (under investigation)
• Why
  – Your finger is not a mouse
User interface: workstation

- Then
  - Limited linking across views
- Now
  - Reconceptualization of linking
- Why
  - Original goal was pedagogical, for students
  - New goals prioritize research resource, for scholars

The digital humanities perspective

- Digital humanities is humanities
  - It’s not about the technology
- A digital edition should facilitate philological work that would be impossible (or so impractical as to be effectively impossible) without it
- Design principles for digital humanities have their own rationale (not the same as for e-commerce)
- The target audience is repeat visitors: mid-level to expert users
- The user interface should be intuitive and unobtrusive

On beyond Paul the Simple

David J. Birnbaum
University of Pittsburgh
djbpitt@gmail.com
http://www.obdurodon.org

Design considerations

- Limited screen real estate
  - Number and location of windows and panels
  - Pop-ups and tool tips
- Limited inventory of user-interface events
  - Click
  - Mouse movement, double click, right click, etc.