

Interim Report

PW-51672-14

Coptic SCRIPTORIUM: Digitizing a Corpus for Interdisciplinary Research in Ancient Egyptian

Project Directors:

Caroline T. Schroeder, University of the Pacific

Amir Zeldes, Georgetown University

Institution: University of the Pacific

November 28, 2015

Coptic SCRIPTORIUM has met most of the projected milestones for the grant period thus far. This report outlines the activities and accomplishments during Phase 3 of the work plan for the grant period. Please see our May 2015 interim report for some of the milestones in Phase 3 completed earlier this year. Additional activities are described at the end of the report.

Continued transcription and digitization of expanded corpus (Shenoute, Sayings) and completion of import, tokenizing, and tagging of papyri.

We have now completed the milestone of publishing a full book of the Gospel of Mark annotated with the tools developed under NEH ODH grant HD 51907 and manually edited by annotators/editors on the team. In addition, the annotators/editors have published additional material in the *Sayings of the Desert Fathers* corpus, and have published new documents by Shenoute based on manuscript transcriptions provided to the project by David Brakke (Ohio State University). (Papyri were completed in the previous reporting period.)

Discussion and planning regarding creating an online interface for researchers to contribute text and/or annotations to the corpus; research the potential of adapting LAUDATIO, SoSOL, or creation of our own.

We have developed a robust environment on GitHub for adding transcription files and annotations to our corpora. After conversations amongst the project directors and annotators/editors about potential features of an online tool for transcription and/or annotation, we decided to continue development of an online transcription editor and online annotation tool in conjunction with other scholars in Germany under the aegis of the bilateral NEH-DFG grant for KELLIA (HG-229371-15). Dr. Zeldes has worked with the LAUDATIO team in Berlin to archive our text corpora.

Meeting of Advisory Board in Summer 2015

The Coptic SCRIPTORIUM advisory board met via Skype with the project directors on July 7, 2015. Board members reviewed our activities, especially the newly published corpora and the web application that resolves persistent URNs for our data to provide visualizations of and links to the most recent versions of that data (at <http://data.copticscriptorium.org>). Board members supported the activities and provided additional suggestions for improving the papyri annotations and for providing guidance to users on how to cite the project and corpora. We have made the changes to the papyri, and Schroeder is in ongoing consultation with the board about citation practices. We also consult the board regularly about linguistic annotation issues.

Additional Activities

In addition to the originally planned goals and milestones, we have several further achievements:

- **Lemmatizer:** one of our editors/annotators (Elizabeth Davidson) compiled lexical information for a Coptic lemmatizer, which will annotate words in our corpora with their dictionary headwords. A. Zeldes incorporated this lexical data into the lemmatizer tool under HD 51907.

- Hackathon: In July 2015 in California, C. Schroeder hosted a hackathon. Two editors/annotators (Elizabeth Davidson, Christine Luckritz Marquis) and an undergraduate intern funded by the University of the Pacific (David Sriboonreuang) attended. We annotated and published a number of documents to our public corpora, updated our project wiki, and planned for the fall.