

ILiADS Institute Week Proposal

The Restoration of Nell Nelson: An Investigation of *The Chicago Times*' Series: "City Slave Girls"

Team Information:

Rebecca Parker, University of Pittsburgh at Greensburg, Social Sciences and English Literature double major (class of 2016), project manager/chief editor – will attend ILiADS

Rob Spadafore, University of Pittsburgh at Greensburg, Creative and Professional Writing major (class of 2016), editor – will attend ILiADS

Nicole Lottig, University of Pittsburgh at Greensburg, Anthropology major and Gender Studies minor (class of 2016), transcription and encoding team – will attend ILiADS

Brooke Stewart, University of Pittsburgh at Greensburg, English Literature major (class of 2018), transcription and encoding team – will attend ILiADS

Project Description:

Helen Cusach's muckraking that exposed the unsanitary working conditions and cruel mistreatment of the women and children in Chicago's manufacturing industry was published under the pseudonym Nell Nelson in her series "City Slave Girls" printed by *The Chicago Times* in late July 1888. Cusach's satirical and witty exposé created an immediate social agitation causing some of Chicago's most prominent men to publicly react. This included responses from Chicago's governor John Altgeld, the president of the Chicago Board of Trade, Charles Hutchinson, the vice president of the First National Bank, Lyman Cage, among others. Owners of the textile mills, meat-packing factories and the other commodity-based industries that Cusach scrutinized recognized the threat, with one manufacturer ultimately suing the paper claiming misrepresentation. Later in 1888 the Barkley Publishing Company printed an abridged version of the newspaper series in a book titled *City Slave Girls*. Author John McEnnis also included parts of the series in his 1888 pro-union book *The White Slave Girls of Free America*. Reprinted in the *New York World* and referenced by labor and women's rights groups, such as the Illinois Woman's Alliance, the series gained nationwide coverage. Yet, presently Nelson's series is infrequently referenced in the discussion of American industrialization or investigative reporting; her work superseded by that of Upton Sinclair and Nellie Bly.

There are three main facets of this project. First and foremost the project requires a complete digitization of the three available sources where Nelson's "City Slave Girls" series were published. Each of these sources provide their own hurdles in obtaining code-ready, plain text. This includes working with messy OCR texts and extensive transcriptions of low-quality photographs of the original newspaper articles. The project team considers this the most important aspect of the project because there is currently no complete archive of this series digitally available as searchable text. The second facet of the project is to maintain a logical tag

set developed within the TEI guidelines that will govern the markup of all three sources. Finally, this project intends to provide data visualizations and conclusions to related research questions.

One of the beginning research questions posed was what kind of information was altered and excluded. Why if all of the sources were published within the same year (1888) do both of the book versions exclude important aspects used by Nelson in the original newspaper series detailing the corruption? What was left out and/or changed? How did the book publications provide a different variation to Nelson's clearly disgusted image portrayed in the original *Chicago Times* publication? In the fall of 2014 a co-editor of the project, Shane Daube, showed interest in representing the connotation of the three main character types within the series: the male personae, the female personae, and Nelson. He wanted to compare (using the minimally transcribed data in development) the three different voices and the variation of voice connotations within the original articles versus the article parts in *The White Slaves of Free America*. Using XQuery and SVG Shane developed a graph visualization that provided interesting conclusions in regard to the extreme alterations made between versions. He provided variation data that otherwise would only have been derived and made obvious after a close reading of both sources.

In the 2015 fall semester, the major focus was overhauling the code behind the articles. We worked to establish a system consistent with the Text Encoding Initiative that was as objective as possible. When evaluating our code, we considered three major ideas: the quantitative measure of dialogue dependent on gender and speaker, the relationship between possessive adjectives and their corresponding nouns, and Nelson's distinct vocabulary. We also changed the CSS, added SSI's, and established a schematron. Along with these changes, we found a second home at newtFire.org within the Pittsburgh Supercomputing Center.

Our first step was bringing our system of coding in line with the Text Encoding Initiative (TEI). This was a major task; in working with the prior tagging system, we found that nearly every element needed to be replaced or renamed. On top of that, the initial markup was inherently subjective by design. The prior code was intended to markup points of connotation (such as sexual, sarcastic, positive, or negative). As a team, we discovered significant inconsistency in this interpretation of text from coder to coder. When dealing with a single example, what one person considered sarcastic, another considered sexual. This also posed the question of how much of the text needed to be wrapped. If a whole paragraph is sarcastic, should all of the sentences be wrapped? What if just one word seemed sarcastic? Our initial concept was to map these inconsistencies between coders. We thought that this too would cause an issue with time forcing every individual within the project to read and mark up each article. After spending time bringing this inconsistent set of tags into the TEI, we made the ultimate decision to change the focus of our system. In doing so, we removed all tags representing the subjective connotations.

The first major research question developed out of these tags: if Nell's intention was to unveil the plight of the 'city slave girls,' whose voice does she give the most dialogue to? We were interested in evaluating the quantity of dialogue between males and females and between speakers. We produced a bar graph comparing the amount of speech characters from each speaker and another graph comparing the quantity of speech characters from each gender, over the collection of transcribed articles. We also implemented these tags into our JavaScript and

CSS of the articles themselves. When viewing the articles on our website, a clickable box appears on the page, allowing viewers to toggle the highlighted text of the different types of dialogue, dependent on the speaker.

Our next goal was to consider a more objective way to markup Nelson's articles. Thanks to a suggestion from Prof. Greg Bondar, we began marking up grammatical elements. We decided to begin with both nouns and adjectives. This alone proved to be an immensely time-consuming task.

With these rules and tagging system established, we created two more sets of data visualization. The first was a list, pulling all of the adjectives inside the entire collection of articles and outputting a count for all the adjectives that appear more than twice. This is to display visually the interesting choice of language that Nelson used in exposing Chicago industry. We also created HTML tables for each article using XSLT that highlight the association of possessive pronouns, nouns, and adjectives. Using our system of 'archetypes,' these tables help to reveal how the nouns and adjectives of each article are being associated with the working women and in some cases revealing an interesting possessive action enforced on the working women by the employers.

The Restoration of Nell Nelson project aims to make the work of Cusach known again through the digitization of her series. We are making the entire text of the articles available as searchable TEI constrained XML. Using XSLT, SVG and JavaScript we highlight Cusach's distinct use of language and her representation of gender and archetypal speakers. By comparing the original series with the republications we plan to indicate how changes in Cusach's distinctive language altered the series original intent for complete exposure of corrupt Chicago industrialization. This project is free to access by anyone who wants to know more about the conditions women and children in Chicago's manufacturing industry were exposed to by visiting <https://nelson.newtfire.org>

The Role of ILiADS:

We wish to attend ILiADS to collaborate with and get some new ideas from the expert coaches and other project teams about our encoding technique. We are interested in learning how to use new and different technology to enhance our project, especially through the addition of a greater variety of data visualizations. We have been thinking about how to represent where all of the factories that Nelson visited in Chicago by creating a digital map, but we are unsure how to start this and help from experts in the field would be key to displaying this visualization. We are also interested in charting or graphing the possessive ownership of and by our defined 'archetypes.' We believe that ILiADS will be able to help our team develop our project into a very useful resource for historians interested in muckraking, industrialization, and women of the Progressive Era.

Project Timeline (by semester):

Spring 2014: Project originated as Rebecca Parker's History Capstone (extensive research paper and campus-wide presentation) at the University of Pittsburgh at Greensburg.

Fall 2014: The commencement of making this project digital with the first transcriptions and preliminary subjective encoding.

Spring 2015: Continued article transcription and edits to the website.

March 2015: Project represented at the Sigma Tau Delta International Convention in Albuquerque, New Mexico.

Fall 2015: Markup shift to meet TEI standards; more objective tag set determined; continued transcription and encoding of articles; new data visualizations; housed on newtFire (<https://nelson.newtfire.org>) in the Pittsburgh Supercomputing Center.

November 2015: Project represented at the first Undergraduate Network for Research in Humanities (UNRH) conference.

Spring 2016: New data visualizations; continued article transcriptions and encoding.

March 2016: Project represented at the Sigma Tau Delta International Convention in Minneapolis, Minnesota.

Summer 2016: Continued development of site's user experience; continued article transcriptions and encoding.

June 2016: Project presentation at Keystone Digital Humanities Conference in Pittsburgh, Pennsylvania. (TBD)

July 2016: ILiADS Institute: More data visualization and friendlier user experience.

December 2016: Project nearing completion.

Sustainability:

The project is housed on newtFire.org at the Pittsburgh Supercomputing Center. The management and correspondence of the project team relies on the innovative Git version control system and the social network GitHub. The project is student run and will be maintained by all of the team participants, as well as any incoming future members. A few of the team members have applied to graduate schools for Digital Humanities and hope to build the project by what they learn in furthering their education.