

Libraries at University of Nebraska-Lincoln
Library Conference Presentations and
Speeches

University of Nebraska - Lincoln

Year 2006

The 'Meat-ax' on the Desktop: A Digital
Edition of Cather's Early Journalism

Andrew Jewell
University of Nebraska - Lincoln, ajewell2@unl.edu

M/MLA Presentation
November 10, 2006
Andrew Jewell

“The ‘Meat-ax’ on the Desktop: A Digital Edition of Cather's Early Journalism”

In 1921, Will Owen Jones, Willa Cather's old editor in Lincoln, Nebraska, recalled the novelist's early career as a theatrical reviewer for the *Nebraska State Journal*. “[S]he wrote dramatic criticisms of such biting frankness that she became famous among actors from coast to coast,” he wrote, “Many an actor of national reputation wondered on coming to Lincoln what would appear the next morning from the pen of that meatax young girl of whom all of them had heard.”¹ Many readers of Cather, those familiar with her as the author of *Death Comes for the Archbishop* and *My Ántonia*, are surprised that anything emerging from her pen could be called “meatax”; her subtle and suggestive novels rarely inspire critics to use metaphors involving brutal, sharp objects. And yet, that cutting voice is there in the nearly six hundred columns, reviews, features, and articles Cather wrote between 1891 and 1904, pieces that contain sentences as pointed and funny as these: “The Dramatic Mirror announces that Maggie Mitchell will revive ‘Jane Eyre’ and ‘Fanchon’ next season. Yes, but who on earth or in the waters under the earth is to revive Maggie Mitchell?”² And another: “She was quite pretty — when she had her veil on.”³

Willa Cather's work as a journalist, writing for mostly Lincoln and Pittsburgh newspapers when she was in her twenties, is a remarkable and wide-ranging body of work that offers significant insights into major critical questions about Cather and American literary and cultural history. Since her writings were largely critical responses to theatrical productions, books, and music, the articles provide a detailed and rich record

of her emerging aesthetic sensibilities during this crucial period of development. Additionally, since she covered such a wide variety of performances, companies, and works, her journalism offers a detailed view of the culture of the 1890s and early 1900s. In responding critically to so much material, she had opportunity to comment on a wide range of topics, providing relevant and compelling texts that can inform many dominant approaches to Cather's work and to literary study generally, including issues of gender, sexuality, race, class, the marketplace, performativity, biography, aesthetics, and cultural studies. And yet, at this moment, no complete edition of this writing exists in an acceptable, scholarly edited form. If readers want to access Cather's journalism, they either have to consult the highly selective volumes by Bernice Slote and William Curtin from the mid-1960s, *The Kingdom of Art* and *The World and the Parish*, respectively, or they have to wind their way through reel after reel of microfilm.

Literary scholars are beginning to recognize the centrality of periodical writing in literary history. In their recent article "The Rise of Periodical Studies," Sean Latham and Robert Scholes note that scholars are "finding in periodicals both a new resource and a pressing challenge to existing paradigms."⁴ This richness of intellectual discovery in the pages of periodicals has led to the publication of many critical articles and books and is a driving force behind mass digitization efforts that make complete runs of various periodicals available to the growing audience of interested scholars and readers. Scholarly editions of the journalistic writings of important authors ought to be a natural outgrowth of this vigorous study of periodicals. To our knowledge, no scholarly edition of the complete journalism of an American author has ever been completed, even though many authors of central importance--Walt Whitman, Mark Twain, Theodore Dreiser,

Ernest Hemingway, and John Dos Passos, for example--have had journalistic careers. The absence of these editions is part of what Shelley Fisher Fishkin calls the historical "neglect" of American literary scholars who have "missed an important aspect of American literary history and biography" in "glossing over" the journalistic careers of imaginative writers,⁵ a neglect many are now eager to remedy. To address that need, the Cather Project and the Center for Digital Research in the Humanities at UNL have partnered to collect, edit, and publish, for the first time, the complete journalistic writings of Willa Cather. We are creating an edition of Willa Cather's journalism that will provide a model for editions of other authors, demonstrating to scholars the unprecedented access and comprehensibility that a digital scholarly edition can provide.

Led by Kari Ronning, Assistant Editor of the Willa Cather Scholarly Edition, and myself, the project will locate previously unidentified writing by Cather; transcribe Cather's work from microfilm and/or original paper copies of the periodicals; richly encode the transcriptions in eXtensible Mark-up Language (XML) according to the Text Encoding Initiative's document type definition (the standard for digital scholarly editions); capture high-quality digital page images of Cather's work from the original publication when available; research and write annotations of Cather's references to people, places, events, and culture; provide an introduction placing the journalism in relation to the historical period and to Cather's mature work; develop a sophisticated user interface that allows for searching, sorting, and browsing the entire corpus of articles; work with catalogers to create full records for integration into local and international library catalogs; and publish the edition digitally on the free, educational website, the *Willa Cather Archive*.

Cather's journalistic writings, with their wide-ranging topics, critical expressiveness, and bold voice, are not texts for scholars devoted to Cather esoterica, but are fundamental texts for understanding Cather's development as a writer. Spanning her late teenage years into her thirties, Cather's journalism is a record of her imaginative and critical growth during her young adulthood. As such, it provides a record of her engagement with hundreds of literary, musical, and theatrical works, as well as with key figures in late nineteenth-, early twentieth-century culture. In her articulation of her critical perspective, the young Cather established ideas and themes that would be further explored in her later fiction: the cost of the creative life and of artistic integrity, the difficult balance of the ideal with the pragmatic, and the superficial materialism of much of modern American culture. Additionally, in Cather's journalism one can discover prototypes for many of the people and places that populate her fiction. For example, Cather's description of a performance of *Camille* in her 1918 novel *My Ántonia* is based on a performance she reviewed in the November 23, 1893 issue of the *Nebraska State Journal*. Or, more subtly, Cather's sketch of a couple in a strained marriage sitting in a theater watching the same play that they had seen years before when intoxicated with one another (December 3, 1893) is echoed in her 1925 novel, *The Professor's House*. From the conceptual to the specific, the background to Cather's fiction is visible in her journalism. To have such a record means that scholars can study the transitions of Cather's mind as she matured and better understand her remarkable imagination.

One hurdle to any previous ambitions to publish a complete edition of Cather's journalism has been page limits: any print-based edition would have to be a multi-volume publication in order to be comprehensive, and presses are very reluctant to

undertake such projects. In a digital environment, we can provide a wealth of content absolutely inconceivable in a print edition: for each article, we present a fully-edited electronic text, a high-quality page image of the original publication (typically scanned at a high resolution from a microfilm copy), and dozens and dozens of annotations which are filled with images of the people and places Cather writes about. The amount of and range of the annotations, something for which Kari Ronning is almost exclusively responsible, is astounding: for the first fifty-five articles alone, nearly 1,500 names, titles, allusions, places, and cultural references have been identified for explication.

The digital environment has other advantages as well. Unlike print editions, our digital edition will be very flexible. If research uncovers previously unknown Cather-authored work in the future—even if it is years after our initial editorial efforts are done—our edition can accommodate and immediately include the new work. Also, once the entire corpus of the journalism is encoded, we can develop tools that will allow readers to interact with and learn about these writings, for example, allowing them to track word usage over time, analytically compare the journalism stylistically to Cather's mature fiction, or automatically generate an alphabetized index of all the people referenced in Cather's writings.

The two most obvious advantages of digital presentation at this stage, however, are access and searchability. Anyone with an internet connection and a web browser will be able to read these writings. Not only will a reader be able to go through all six hundred chronologically, but she will also be able to do sophisticated searching across the entirety of the texts. A scholar exploring Cather's connections to Clara Morris or Shakespeare or Herrmann the Magician or productions of *Tosca* will be able to

immediately locate all of her references to these subjects. Furthermore, our encoding practices will enable users to automatically generate sophisticated indices, moving beyond simple names and titles, which Slote and Curtain were limited to, to include theatrical roles, musical titles, dramatic titles, and more.

Many important works of Cather scholarship have been published essentially uninformed by the large corpus of Cather's early journalism. Critical arguments that consider, for example, Cather's racial attitudes do not consider certain texts that would seem fundamental to the argument, such as Cather's rather startling review of an exhibit of "racial types" in a 1901 article written from Washington, D.C. Though one could blame scholars for not properly doing their research when ignoring such materials, I tend to be more forgiving: is it really practical to expect an academic living on the east coast to seek out a text that exists only on a reel of microfilm in the Nebraska State Historical Society? The level and nature of access directly affects scholarship, for scholars, like all humans, must live within the confines of budgets and deadlines.

I think it worth pausing for a moment to explain the process of creating a digital edition. Computers have been so heavily marketed as a convenience tool, and so heavily criticized as a product that allows students to cheat or simply do bad research, that many seem to regard digital work with some suspicion. Digital scholarship is getting more and more respect—and funding—but it still lacks the sacred aura of the printed text. Too often, people think undertaking a digital project involves merely typing something out on a word processor and slapping it up on the web. Though digital projects—very, very bad digital projects—can be done that way, that quick-and-easy method falls far short of the international standards developed over the past few decades by consortia of scholars and

technical specialists from around the world. I will spare you the technical details, but reflecting on our process demonstrates how humanities computing work, though avant-garde in many respects, is absolutely grounded in traditional scholarly practices.

Our first step was to identify the copy-texts and our sources for obtaining the copy-texts. Given that the vast majority of Cather's journalistic writing exists in only one form, its publication in the original periodical, selection of copy-text wasn't a major issue. What was and continues to be a challenge is getting those texts in a usable form. As far as we know, most of the runs of the *Nebraska State Journal*, the *Lincoln Courier*, the *Pittsburgh Leader*, and other newspapers that published Cather's work exist only on microfilm. It is difficult to do accurate transcription directly from the microfilm; as our early efforts demonstrated, the subtle distinctions between characters, as between a comma and a period, are difficult to discern among the scrunched letters dimly projected on a microfilm reader. Our solution was to rely on high-quality digital scans of the microfilm, using one of the microfilm scanners available in the Center for Digital Research in the Humanities at the University of Nebraska-Lincoln. Though the scanning process is relatively time-consuming, the images we derive can serve multiple purposes: they are superior versions of the original that can be used for transcribing and editing the text, they can be stored on an accessible server so different project team members can consult them from different locations, and they can be delivered to the public along with our digital transcriptions.

The creation of the digital transcription is itself a highly-involved and crucially important process. Like many digital projects, our texts are encoded in Text Encoding Initiative (TEI)-conformant Extensible Markup Language (XML), which means that we

put the transcriptions into an electronic form that will be stable over time and can contain a good deal of editorially supplied information; this markup, in Susan Hockey's phrase, "puts intelligence into texts, providing information to help computer programs perform more meaningful operations on them."⁶ The "intelligence" we add, at its most basic level, informs the brainless computer what the human mind is trained to detect when looking at a page of print: that the words exist in complicated relationship to one another, that certain words serve as a heading, that groups of words broken up with hard returns and indentation constitute a paragraph, that certain words are the names of people or plays, etc. The amount of markup can vary tremendously, so early on in the project we had to decide what information was important within the text and worth singling out. To do that, we had to begin to envision what sort of questions we might want to ask of this group of texts, or, more pointedly, what kinds of queries we wanted to make it possible for a computer to ask. We decided that, in addition to the information the annotations will supply as one reads each article individually, we may want to note every person Cather references, and every play, book, piece of music, and character. Moreover, we recognized that it would be very important to regularize the spelling of each of those categories, so that if we asked the computer to generate a list of people referenced, Sarah Bernhardt wouldn't show up repeatedly under multiple names: "Sarah" with an "h", "Sara" without an "h", "Bernhardt," "Sarah Bernhardt," or "the great French actress."

Encoding in XML is a very time-consuming process, for it requires that a well-informed human mind read each article carefully and make judgments about the application of the markup. Regularization requires identification, which isn't always a piece of cake, and it requires absolute consistency. We can use some tools to encourage

consistency, but markup is ultimately a product of human skill, discernment, and time. And markup of names is only a part of the process: the encoder must also make decisions about the structure of the text, insert references to the annotations, and describe and reference accompanying illustrations. And after the transcription and initial encoding are done, others do multiple proofreadings to remove all errors and present a text that is accurate and as rich as possible.

The work that this requires is, in the end, worth it, for it creates a body of electronic texts that are both powerful and stable. Of course, we have not had enough history with digital research yet to know exactly how things will age, but we are following widely-adopted procedures that were designed by intelligent people to withstand the rapidly-changing technological world. This means separating all intellectual content from design and layout, for we do know that, no matter how great things look to us today, in about five years we'll think it looks way too 2005. The hundreds of XML files will not need to be changed at all. Instead, another file written in Extensible Stylesheet Language for Transformation or XSLT—one file—can be altered to reflect new design decisions. The XSLT stylesheet works by making transformations in response to the XML markup; for example, we tell the XSLT file to make the headings look bold and appear in the center of the screen. It is this XSLT file that allows the XML to be transformed to HTML for easy viewing in a web browser.

Creation of the texts is only part of the story, however. The other part is the research and authorship of the hundreds and hundreds of annotations done by Kari Ronning, annotations so large and various and illustrated that they would be impractical in a print edition. Consider Cather's article "Between the Acts," published April 29, 1894

in the *Nebraska State Journal*. An extensive section entitled "Cues" offers a generous if brief glimpse into the theater business of the later nineteenth century in its length recitation of the latest theater news and gossip. Cather's voluminous and quick references Ronning identified one-hundred and fifty-two items that needed annotation.

Though this article represents a dramatic example of the richness and extent of the annotations, it is not unusual. Cather's journalism, particularly in the period we are currently working on, is highly referential; it expects an audience interested in and conversant with theatrical life. Without this depth of annotation, our edition of the journalism would require a highly-specialized readership, one already steeped in both the life and writings of Willa Cather and late nineteenth-century culture. That said, I doubt there are many people in the world that could recognize a dramatic situation in an article later echoed in Cather's mature fiction, catch references to Byron and the Bible, know the plot of the play "Brother John," and tell you how many Italian fruit sellers were in Lincoln, Nebraska, in 1893. This wide referentiality was a major quality of the journalism we sought to make visible. The selective editions published in the 1960s privileged Cather's discussions of art and artists, and as a result, one could easily assume that Cather was only interested in that topic. Though certainly art dominated her imagination and is the source of some of her most impassioned writings of the period, she has much more to say. In its representation of Cather's January 28, 1894 article, "One Way of Putting It," Bernice Slotte's edition included Cather's disparaging remarks on the artistic merit of the Lansing Theater's drop curtain, but excluded Cather's lampoon of egotism within the Nebraska state government: "It is strange, but there seems to be something in the air of that state house which causes an inevitable enlargement of the

head. The very janitor who mops the floor—sometimes he really does mop it—imagines that if he should lay off a day the reign of chaos and old night would begin immediately."⁷ Our edition of the full article, which includes an annotation, photographs of the Lincoln capitol building, and the drawing—perhaps by Cather—that accompanied the original article, gives a much richer and more complete view of Cather's writing.

In "The Novel D meubl ," Cather wrote, "One does not wish the egg one eats for breakfast, or the morning paper, to be made of the stuff of immortality."⁸ In essence, I agree with her: not every common thing needs to be overstuffed with self-importance and profundity. But I would quibble with the details of her comment. In looking back, fully aware of the remarkable career and works of this American writer, some parts of the morning paper do seem much more than transitory. Though I stop short of declaring our edition as the stuff of immortality, I do believe that the digital edition of Cather's journalism, the first full and scholarly edition of the material that emerged from Cather's "meatax" pen, will reclassify these writings out of egg-dom and into the body of Cather's well-read, well-studied work.

¹ Quoted in Bernice Slotte, *The Kingdom of Art: Willa Cather's First Principles and Critical Statements, 1893-1896* (Lincoln: U of Nebraska P, 1966), 16-17.

² Willa Cather, "With Plays and Players," *Nebraska State Journal* (March 11, 1894): 13. *The Willa Cather Archive*. 14 August 2006

http://libtextcenter.unl.edu/examples/servlet/transform/tamino/Library/cather?&_xmlsrc=http://libtextcenter.unl.edu/cather/writings/journalism/cat.j00033.xml&_xslsrc=http://libtextcenter.unl.edu/cather/xslt/cather_j.xsl

³ "Amusements," *Nebraska State Journal* (April 20, 1894): 6. *The Willa Cather Archive*. 14 August 2006

http://libtextcenter.unl.edu/examples/servlet/transform/tamino/Library/cather?&_xmlsrc=http://libtextcenter.unl.edu/cather/writings/journalism/cat.j00004.xml&_xslsrc=http://libtextcenter.unl.edu/cather/xslt/cather_j.xsl

⁴ Sean Latham and Robert Scholes, "The Rise of Periodical Studies," *PMLA* 121.2 (March 2006): 517.

⁵ Shelley Fisher Fishkin, *From Fact to Fiction: Journalism and Imaginative Writing in America* (Baltimore: Johns Hopkins U P, 1985), 3.

⁶ Hockey, Susan. "The Reality of Electronic Editions" in *Voice, Text, Hypertext: Emerging Practice in Textual Studies*. Ed. Raimonda Modiano, Leroy F. Searle, and Peter Shillingsburg. (Seattle: Walter

Chapin Simpson Center for the Humanities in association with the University of Washington Press, 2004). 362.

⁷ Cather, Willa. "One Way of Putting It," *Nebraska State Journal* (January 28, 1894): 13. *Willa Cather Archive*

http://libtextcenter.unl.edu/examples/servlet/transform/tamino/Library/cather?&_xmlsrc=http://libtextcenter.unl.edu/cather/writings/journalism/cat.j00032.xml&_xslsrc=http://libtextcenter.unl.edu/cather/xslt/cather.j.xsl#ref_n00401, April 10, 2006.

⁸ Cather, "The Novel D meubl ," in *Willa Cather on Writing* (Lincoln: U of Nebraska P, 1988): 36.