

Texts and Transmision: Teaching American Literature with Juxta

Steve Marsden, Stephen F. Austin State University

Students, especially undergraduate students, often do not understand authorial or editorial revision or textual scholarship. For many undergraduates, the poem or story is what appears on the page in their anthology. They seldom think except in the haziest of ways of the process by which the author created the work or by which the work was edited for publication or revised for various editions. Only the advanced creative writing students I talked to seem to have a real idea of the perpetual nature of authorial editing, and outside of Early Modern Drama courses, where the variant texts of *Othello*, for instance, might be usually at least acknowledged, undergraduates in many programs seldom encounter the differences between versions of texts. Even in those cases, their knowledge is usually limited to a bullet point or a moment in a lecture. When I have pointed out the (fairly many) typographical errors in the anthology I use in my survey courses, the students seem slightly taken aback. They are used to the text in front of them being unquestioned and unquestionable.

That's not the way it ought to be. It's desirable that students be able to understand that the work as presented in their course text may be the result of a rather complex process, a process that can give insight into the evolution of an author's view of the work, as well as to the process of editing, with its give-and-take and the complex dynamics of power and economics it can open up. Matthew Kirschenbaum argues that this process—for which he uses Randall McLeod's term "transmission," "with its emphasis on technologies of media and mediation"—is central to the concerns of contemporary literary studies (Kirschenbaum 157). In this paper, I will look at some less-than-satisfactory ways of having students visualize differences between editions and get hands-on experience in

comparing editions. I will compare these methods to my latest classroom experiments with using Juxta, free textual-comparison software along with web-available transcriptions of different editions of works. I'll give some examples of how I've used Juxta in the classroom, both as a visualization tool and as a way to let students do fairly sophisticated hands-on textual detective work for themselves. Incorporating Juxta into undergraduate instruction, I will argue, not only increases students' textual-critical sophistication in dealing with both print and electronic versions of texts, but may offer students interesting opportunities for scholarly public service assignments.

Using Variora and Historical Collations

Most scholars use variorum editions or historical collations in scholarly editions to recreate the textual history of a work. A variorum edition presents a base text, and documents other states of the same text by recording the differences between the base text and other witnesses; a historical collation in a scholarly edition does much the same, but frequently records the substantive changes only. These two approaches work fairly well, and provide the advantage of recording the differences between texts without having to reproduce multiple times the things that are the same. In a paper-book world, these editions are efficient ways to record and highlight changes between states of a text. That doesn't make them ideal for the undergraduate classroom, however. The notation system for both are hard to understand, using what seems to the uninitiated to be an arcane shorthand that smacks uncomfortably (for many liberal arts types) of mathematical or logical formulae. Since variorum editions frequently start with the latest text with authorial input as a base text, and record variations from there, they don't have the advantage of easily visualizing changes. It's hard to visualize complete

states of a text from a list of variants alone—students must reconstruct witnesses from the notation, which can be quite time consuming. The historical collations included in scholarly editions frequently exclude accidentals, and seldom provide ways to visualize changes in spacing. In addition, many non-canonical authors have not attracted the extremely time (and hence money) intensive scholarly labor that is required to create a variorum or scholarly edition.

In one senior-level course a few years ago, I managed to get students to understand variorum editions well enough to use them, but it was a difficult slog, losing many students along the way. Even when students were able to extract the differences, they seemed to feel that they had learned an obscure scholarly skill, but what makes students into future scholars—the thrill of independent discovery—was sorely lacking.

Hands-On Collation

The traditional way to have students get hands-on experience with textual criticism is by assigning collation exercises. Unsurprisingly, these kinds of exercises are seldom assigned before the introduction to bibliography and research methods course that begins many graduate programs, and sometimes not even then. Assuming you have all the variant states of a text on hand, physical collation or reading for variation is still tedious, time consuming, and requires a discipline that seems hard to find anywhere, let alone in an average undergraduate classroom. Moreover it costs time that very few professors are willing to commit to a single lesson or project. Variations shown in the page or two you can assign are unlikely to be dramatic—if they are, it means you've set them up to be, and students are probably not doing original research.

Computer Collation in the Undergraduate Classroom

Thanks to the increasing accuracy of Google Books scans and Project Gutenberg's distributed editing, along with other commercial full-text databases and academic ventures such as *The Walt Whitman Archive*, edition-specific electronic texts of usable fidelity have become available easily, especially for 19th Century literature, where many editions will be available and out of copyright.

Software for collation has existed for many years, but has largely had a rather steep learning curve and not much appeal beyond those actually working on scholarly editions or complex textual-critical projects. While I've always had a kind of dabbler's interest in textual criticism, as I had never taken on a serious project before, I looked at even comparatively friendly software with some trepidation.

Juxta is, according to the "About" section of the project website, "an open-source cross-platform tool for comparing and collating multiple witnesses to a single textual work" (*Juxta*). It is free, open-source, and easy to use. It allows easy visualizations of all the changes between any two textual witnesses (including accidentals, if the program is so set), and allows for easy visualization of variant passages across many witnesses.

Juxta as a Classroom Visualization Tool

In my American Literature to 1865 survey courses, I have always liked to give a sense of poems and stories as edited, changing texts. Unsurprisingly, I use the works of Walt Whitman as my exemplar. From its first appearances in 1855, the poems in Whitman's *Leaves of Grass* underwent constant revision until the 1881 edition (the 1891 "deathbed"

Winter 2011 (4:2)

edition being a reprint with additional material). In past semesters, while performing close-readings on one of Whitman's poems (usually "Crossing Brooklyn Ferry"), I have brought the variorum edition to class. Having marked a few places where the work changed across editions, I am able to pause explication at a line, put the variorum on the document camera, mention previous readings, and ask my students to speculate as to why Whitman likely changed them.

This semester, it was a little different. I went to the online repository of Whitman's works, *The Walt Whitman Archive*. Thanks to the work of some very dedicated and forward-looking editors and the proceeds of a series of government grants, the *Archive* contains the texts of the poems and clusters of all of the distinct editions of Whitman's poetry published during his lifetime, as well as a good and growing selection of his known periodical publications, and a growing number of imaged and transcribed manuscripts. I copy-pasted each version of "Crossing Brooklyn Ferry" into a separate Word window, and saved each one under a distinctive name, making sure to save as Unicode UTF-8 (the text format that Juxta prefers). Then, I added each witness in order by the date of the edition using the large + button. Setting the program to ignore spacing and capitalization, but not punctuation, I made it collate. The difference was impressive. I could now project on the wall the text of any one of the versions of Whitman's poem, and compare it head-to-head with the text of any other edition. In the head-to-head view, the changed or omitted lines show up conveniently highlighted, and colored lines connect the texts. Below you can see the comparison view of the 1856 text of "Sun-Down Poem" and the deathbed edition version of "Crossing Brooklyn Ferry." In spite of some false hits due to line-end hyphenation, the changes are clearly highlighted. Then, I could step through the course of the poem, showing the differences between each edition and the next.

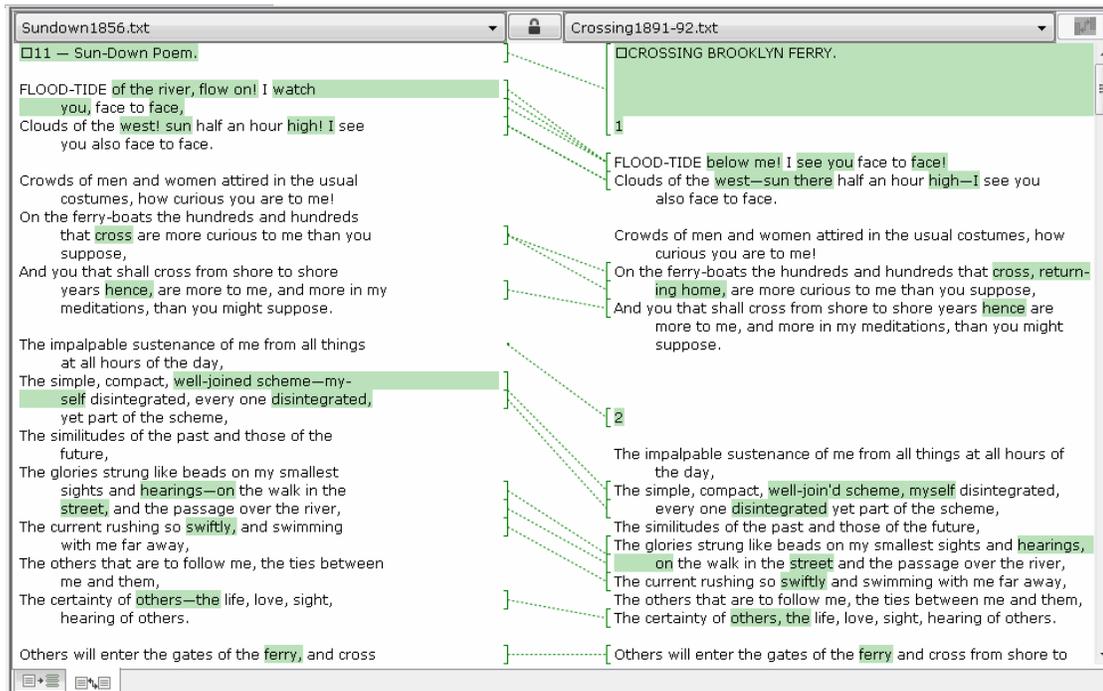


Fig. 1. The 1856 and 1891-92 "Deathbed Edition" versions of Whitman's poem compared

My students seemed a little more impressed than usual as we discussed Whitman's substantive changes, his changing strategies in punctuation, and the effects of each alteration. The more technophilic of my students noted down not just the URL for the *Archive* but the name of the program. One of them later contacted me, letting me know that she'd been using it on her own. This got me thinking: if one of my students in a sophomore-level survey course could learn the program on her own, why couldn't I teach it to seniors?

Juxta Allowing Hands-On Exploration

I determined to see if Juxta was simple enough to allow students working on their own to make their own comparisons between texts. Luckily, I was assigned to teach a senior seminar in Civil War Literature

the next semester. The course had a small enrollment, and was scheduled in a computer laboratory. Since the Civil War and Reconstruction-era literature I meant to teach was all out of copyright (and some of it hard to find in print), I decided to focus on teaching using contemporary texts in digital form. In a fifteen-minute lecture, I showed students how I prepared comparable texts, showed how I loaded them into the program, and demonstrated the end result of collation.

Juxta proved comparatively easy to teach to my students, and using the *Walt Whitman Archive* as a source for easy-to-prepare texts, I had them comparing complete sets of variations within the space of a twenty-five minute in-class exercise. Despite a few misfires (students not choosing the right encoding, and so getting gibberish, or students loading the same files in multiple times, and comparing texts to themselves), the results were encouraging. Walking around, I heard students excitedly saying, "Cool! I got some changes!" or less-excitedly, "It's all punctuation and line break changes here."

I made a textual-critical paper a major assignment in the course. Students had to find relatively reliable electronic texts of multiple author-edited versions of a single short story, poem, or chapter of a novel, link them to their original publications, and compare them using Juxta. Then they had to classify the kinds of changes between versions, report a few specific ones, and write about the possible motivations for the changes and the effects on the meaning of the work as a whole. The objectives of the project were to make students aware

1. That texts exist in different editions, each presenting problems and opportunities for textual, historical, or literary study
2. That texts are produced by a complex process that often involves many individuals with differing goals

Winter 2011 (4:2)

3. That small authorial, editorial, or accidental changes in literary texts can change the readings of those texts significantly
4. That electronic transcriptions or editions of literary texts are not always reliable versions of the print edition they are based on (but often have the advantage of being correctable)

After students chose the works they wanted to focus on, I nudged them towards primary bibliographies if such existed for their authors. For the most part, students were able to find most of the requisite editions on their own with the aid of Google Books and Worldcat. The technological end of the assignment—installing Juxta, copying editions into Unicode UTF-8 format, and collating—went surprisingly easily, with only one student requiring help troubleshooting an installation. In a larger course, this might be a problem, as the current version of Juxta requires administrator privileges to install. According to Juxta's development blog, the program has recently won a Google Digital Humanities reward, and will be developed into a web-based application in the future, removing this obstacle to student use (*Juxta*).

The biggest difficulty students encountered was in finding decent-quality electronic transcriptions of all or most of the relevant editions. It was an opportunity for students to discover the many and varied textual resources of the Internet. The level of difficulty depended on the quality and bibliographical sophistication of the digital editions available for each author. My students rapidly came up with criteria to judge digital editions for their bibliographic usefulness:

1. Texts needed to be based on a particular print edition, and had to clearly indicate which edition they were based on.
2. Texts needed to be relatively free from error. Texts that had seen human editing or error-correction referring to original texts

were much superior to uncorrected scans created by Optical Character Recognition.

3. In order to check the fidelity of transcription (and make sure of the source edition), students needed to have access to page images of the relevant print edition. Digital editions that preserved the original pagination were easier to deal with than those that did not. The best digital sources were those that provided both text and page image in an easily accessible format.

Thanks to *The Walt Whitman Archive*, the many editions of Whitman's poems were very easy to acquire. The texts had obviously been proofread with some care, and clear page images of every page of every edition were available. When students began to work with texts that came from multiple sources, and un-proofread OCRs, however, more problems became apparent. To work with the editions of Ambrose Bierce's Civil War short stories, students had to use texts from a combination of sources to cover the editions. The earliest appearances of the stories, generally in the San Francisco *Examiner*, were not available online: to compare them would require ordering the microfilm, copying or scanning it, then typing by hand (OCR from the messy microfilm would be difficult) and proofing against the copies. The limited time-frame of the assignment made this impractical. The *Ambrose Bierce Project*, an online project out of Penn State dedicated to Bierce, hosts an electronic edition based on the 1891 Steele first edition of Bierce's *Tales of Soldiers and Civilians* prepared by Donald Blume. But it is an edition, not a transcription: it omits the original pagination, and does not include page images. The usual go-to source for page images, Google Books, currently doesn't have a viewable entry for that edition. That was easily overcome: it was a small inconvenience to interlibrary loan the edition and use photocopies to use to proof the electronic text. Google Books had page images and relatively

good OCR plain-text scans from Bierce's 1898 revision of that text, *In the Midst of Life, Tales of Soldiers and Civilians*, though there were inevitable scanning errors. A Project Gutenberg edition had been made from the relevant volume of Bierce's 1909 *Collected Works*, which students could compare to photocopies from a reprint edition in our library. Students working with Louisa May Alcott's *Hospital Sketches* had a fairly decent transcription without page images from *A Celebration of Women Writers* for the first book edition of 1863, but had to deal with a fairly grievous OCR job on the Google Books plain-text scan of Alcott's 1871 revision and expansion, *Hospital Sketches and Camp and Fireside Stories*. They had to interlibrary loan microfilm for the original serialized version. This patchwork solution made student work more difficult, but it also presented to them in a real way the complexities of digital transmission.

After converting texts from various sources into Unicode UTF-8, students tended to find a lot more "false hits" with the widely-sourced texts from other authors than with the consistently prepared Whitman transcriptions. Some texts used "smart quotes" and others "straight quotes." Some electronic texts placed different spacing before and after internal and terminal punctuation. In addition, there were many differences that were obviously the result of OCR problems—"cl" for "d," "thi3" for "this," "Governor ^,ndrew" for "Governor Andrew." In digital sources that included page images, students could check against the images to see if a problem or a change were a part of the original edition or whether it were an artifact of scanning or transcription. In sources that lacked page images, they were forced to either find or inter-library-loan paper copies to check their texts.

Winter 2011 (4:2)

My assignment prompt required a paper with three major sections:

- **Textual History**, which lists the editions and process of publication the work went through, revealing which the student was able to acquire or consult, and how he or she managed to get electronic text and page images
- **Changes**, which gives a criteria for major and minor changes, then lists and speculates on the effects of, or apparent motivations for, major changes, while categorizing and commenting on minor changes such as punctuation or spelling changes
- **Conclusion**, which requires the student to detail his or her process, briefly telling how he or she conducted research, where he or she went for information on editions and process, and how proceeded with comparisons. Students then give their informed conclusion as to the variants

The results were very promising. All of my students managed to come up with an accurate list of editions, and were able to compare the majority of them. A spot-check of their files and, when available, a check against a variorum edition, revealed fairly good accuracy and completeness. Student speculations as to the motivations of changes seemed on the whole reasonable and well-informed. Many of the students had consulted book histories, critical biographies or authorial correspondence relating to editions to help contextualize changes, and came out of the assignment with a greater sense of the events surrounding each publication.

A reflective essay assigned after the textual-critical assignment revealed the assignment met its objectives. I had expected some grumbling about the time expenditure and the close-focus that the assignment required, but student opinion was positive, with all recommending that the program should be used in future undergraduate

classes. Students tended to respond well to the novelty of the approach, one writing that "Getting to use it opened a perspective on literature I have not been exposed to before." They pointed out an increased appreciation for the impact of seemingly trivial variants, one stating that "Even small changes, such as adding or removing a comma or exclamation point, can create a big change in the meaning or the feeling of a single line or the entire work." The idea that changes could be motivated by authors, editors, or accidents in the publication process, the same student wrote, "was something I had never really considered . . . and it really made me think." Tracing the evolution of a Whitman poem through its full run of editions, another student wrote, "brought me a lot closer to the text and its author than conventional methods of critical analysis would have, and I cannot imagine having done it without Juxta." Perhaps most importantly, students seemed braced by the possibility of doing really original research. One wrote "I found some changes during my research that no other critic . . . has announced and those kinds of discoveries are easily made with Juxta."

The assignment could be made much more efficient by ordering and copying some of the newspaper or magazine editions not available on Google Books and having all the relevant primary bibliographies on reserve from the announcement of the project. Likewise, requiring that students report on the work's publication history early as a daily work assignment would encourage them to have all the texts ordered in a timely fashion. Requiring students to produce a list of transcription errors they discovered during the process would make a transition easy from a purely scholarly exercise to a potentially valuable service assignment.

Textual Critical Assignments and Scholarly Public Service

Working with Juxta and multiple editions, the class discovered an incidental effect of the way using the program focuses reader attention. Highlighting only differences across editions, it flags potential proofreading or transcription errors very efficiently, as the same OCR error or transcription error is unlikely to occur across multiple texts. No text sources my students used were free from error. Even in *The Walt Whitman Archive*, where the electronic transcriptions were created and proofread professionally as a part of *Major Authors on CD-ROM: Walt Whitman*, then proofread by project staff, some few scanning or entry errors still persisted (Folsom and Price). Comparing apparent changes to the page images, my students and I were able to discover a number of these errors—from omitted special characters to duplicated words. By corresponding with Ken Price, one of the co-directors of the archive, we as a class were able to see these errors fixed in a matter of days or sometimes hours, where paper critical editions might have had to wait years for another edition or a posted erratum list.

By finding errors in electronic copies, and then using Google and Google Books, students were able to trace the transmission of errors. By searching for the particular error, students could see how a text-entry typographical error in a poetry transcription in a Gutenberg text was replicated in other sites that re-host, replicate, re-edit or re-brand the text, including print-on-demand books for sale on Amazon that were using the Gutenberg-derived text. Then they could go about trying to correct all the incidences of the error.

Every source of electronic texts has its own barriers to user editing. While the plain text portion of Google Books are, unfortunately, not correctable at present, Wikisource, a text repository which operates with the Wikimedia software, enables users to immediately correct errors.

Project Gutenberg texts are correctable by sending a properly formatted email to the site, and most academic text repositories, likewise, can be contacted by email and will correct texts. By seeing the power of their attention and criticism to improve real resources, students begin to feel like real editors, capable of serving the scholarly world.

To that end, in my senior seminar, I have also included a scholarly public service requirement. At the end of the class, the students will document how they've made the information available on their chosen authors and texts more reliable: correcting Wikipedia pages and online texts, as well as creating online study guides or bibliographies. Too often, undergraduate scholarship serves only a pedagogical purpose, consigned to the desk drawer after the semester is over. By correcting the most easily available literary information on the internet, however, a few well-trained and conscientious undergraduates can make an appreciable difference. In 2004, Jerome McGann made a now-famous forecast:

In the next fifty years the entirety of our inherited archive of cultural works will have to be reedited within a network of digital storage, access, and dissemination. This system, which is already under development, is transnational and transcultural. Let's say this prophecy is true. Now ask yourself these question: Who is carrying out this work, who will do it, and who should do it? These turn to sobering queries when we reflect on the recent history of higher education in the United States. Just when we will be needing young people well-trained in the histories of textual transmission and practice of scholarly method and editing, our universities are seriously unprepared to educate such persons. (410)

Unsurprisingly, McGann heads up the list of scholars working with Applied Research in Patacriticism, the group that produced Juxta (Applied

Winter 2011 (4:2)

Research in Patacriticism). One way to create a generation of scholars may be to make scholarly tools with direct appeal and low bars to entry, enabling students to satisfy curiosity while minimizing drudgery.

The development of increasingly viable and popular digital reading technologies, from Kindles to iPads, along with growing resistance to high textbook prices, means more and more students will be going to easily-accessible, public domain, online versions of literary texts: versions that are now often not reliably edited. Inculcating knowledge of textual history, a critical eye to text transmission, and a habit of ferreting out and correcting error can make your class a part of the solution.

Works Cited

- Applied Research in Patacriticism. "Staff." *ARP*. Web. 20 Nov. 2010.
- Folsom, Ed, and Kenneth Price, eds. "Editorial Statements and Procedures." *The Walt Whitman Archive*. Web. 20 Nov. 2010.
- Juxta: Collation Software for Scholars*. Web. 20 Nov. 2010.
- Kirschenbaum, Matthew G. "How Things Work: Teaching the Technologies of Literature." *Teaching Bibliography, Book History, and Textual Criticism*. Ed. Anne Hawkins. London: Pickering & Chatto, 2006. 155-160. Print.
- McGann, Jerome. "A Note on the Current State of Humanities Scholarship." *Critical Inquiry* 30.2 (2004): 409-413. Print.