

Building the Princeton Prosody Archive

Meredith Martin | Princeton University, Department of English

Grant Wythoff | Columbia University, Society of Fellows

Meagan Wilson | Princeton University, Department of English

Travis Brown | University of Maryland, MITH

DH 2014, Lausanne

1 Introduction

1.1 Context

During the 2012-13 academic year, a grant from the down supported the completion of the Princeton Prosody Archive's first phase. The PPA is a database of digitized records on the teaching of poetry as both a popular and a highly specialized genre between 1750-1923. This long paper will reflect on the outcomes of the grant year, as well as some of the challenges and opportunities presented by launching a large-scale digital archive at an institution with no digital humanities infrastructure.

1.2 Content

The Princeton Prosody Archive collects a range of historical documents, including versification handbooks, poetic treatises, how-to-write-a-poem manuals, grammar books, and other materials pertaining to prosody – both the pronunciation of speech and the versification of poetry, rhythm, intonation, and the measure of language. Collecting these materials hosted by the PPA in one place for the first time will not only serve as a touchstone for the burgeoning field of historical poetics (i.e. the study of poetics in all of its literary historical, linguistic, and

educational contexts), the archive will also provide significant advances to the field of digital scholarship by allowing scholars to apply new research methods to the study of poetry rather than concentrating solely on prose. Not a static repository of historical data, it is our hope that the PPA will compel users to rethink the past and future of navigating, conceptualizing, and historicizing large amounts of data in a variety of genres.

Funding from the Mellon Foundation supported the initial compilation of the PPA’s dataset, the alpha testing of user interface features to be completed in the project’s second phase, as well as the investigation of access issues and long term maintenance for the project in its next five years. A summit was hosted in May 2013, bringing together technologists and scholars of historical poetics to discuss these next steps, and the possibilities for future research. These are briefly outlined below.

1.2.1 Curation

One of the most important starting points for our team has been Terry V.F. Brogan’s annotated bibliography *English Versification, 1570-1980* (Brogan (1981)), which collects and cross-references some 6,000 works of poetic criticism. Partnering with the [HathiTrust Digital Library](#) has allowed us to acquire fulltext data from every book listed in Brogan’s volume, as well as thousands more using subject-specific searches related to prosody and versification. One possibility we are considering now is how to cement a workflow whereby subject specialists who receive data from HathiTrust can feed back their curation of that data into the larger HathiTrust ecosystem.

1.2.2 Visualization and “Medium” Reading

Because these books debate the technology of reading a poem and, more broadly, representing the spoken word, many of these texts contain unique page layouts and typographic conventions whose meaning would be otherwise lost to conventional optical character recognition: tablature, gestural diagrams, musical notation, invented spellings and phonetic alphabets. Natalie Houston (University of Houston), a member of the PPA advisory board, is project director for [The](#)

[Visual Page](#), an extension for the OCR software Tesseract that allows the researcher to identify “unique or representative trends in typography, page layout and book design.” Going forward, we will be working with Houston to identify these not quite close, not quite distant, but “medium reading” features of these documents.

1.2.3 Data from Noise

One of our collaborators, Travis Brown (Maryland Institute of Technology in the Humanities), recognized when going over our OCR of mid-nineteenth century texts that attempt to use musical notation to transcribe the sonic qualities of speech that what appears to be “noise” in our OCR is actually regularized patterns. Thus, there is a repeated error for a G-clef, for instance. Further work on these patterns will prove useful for scholars interested in musical notation.

2 Prosody and Historical Poetics

In 2008, the *PMLA* published a series of articles under the rubric “The New Lyric Studies” (January 2008) to explore the ways that our approaches to poetics were changing in the new century. As part of this series, Yopie Prins’s article “Historical Poetics, Dysprosody, and The Science of English Verse” (Prins (2008)), profiles a discussion group dedicated to gaining critical perspectives on the history of reading poetry prior.

Since Prins’s article (which has been cited at least once a year since its publication), a variety of books, special issues, and articles have been published attesting to the increased interest in a more complicated and nuanced understanding of approaches to prosody and poetic form prior to the twentieth-century. Due to the efforts of this group of scholars, historical texts on poetics that have fallen outside of the purview of English poetry are being recognized as crucial to the formation of our modern notion of poetics.

In July 2008, interest in historical approaches to poetic form, and poetic meter, specifically, surged when hundreds of scholars gathered in Exeter for a conference titled “Metre Matters: New Approaches to Prosody, 1780-1914.” The conference generated a cluster of featured papers on the Blackwell’s Literature Compass

website as well as an edited collection (Hall (2011a), Hall (2011b)). The history of prosody is vast and nuanced and touches on the history of English as a literary discipline, the history of Classical pedagogy as it gives way to English as the language of humanistic inquiry, the development of national literatures in America and England – for example, which poets should stand for the national language and literature – not to mention the development of phonology and linguistics coming out of the fields of comparative philology, to name just a few areas of scholarly inquiry.

3 Interface

The Princeton Prosody Archive will:

- provide the user with a unique experience of the thesis that terms like prosody, versification, meter, lyric, rhythm – terms which we take for granted in the twentieth century as stable literary concepts – were both culturally determined and fundamentally in flux over the two-hundred year period covered in the database
- collaborate with Google Books and HathiTrust to collect all relevant texts from 1750- 1923 related to prosody and prosodic theory
- employ optical character recognition that retains document coordinates for individual characters (whose position on the page often conveys important information), capture musical and prosodic notation (macron, breve, ictus), and identify and extract any images on the pages we’ve scanned
- employ robust text markup using new methods in topic modeling, which will allow users to perform a variety of searches, and will allow, perhaps, for unexpected patterns to emerge from our data. Such techniques may include named entity recognition, citation extraction, and syntactic parsing
- present the original page image alongside transcribed text, preserving idiosyncratic characters, diagrams, or diacritical marks that would be lost to conventional OCR software and therefore invisible in conventional searches
- track shifts between multiple editions of books (as in the case of Murray’s much reprinted English Grammar)

- thread together commentary from bibliographers throughout the archive on particular works of poetry and prosodic theory (that is, include discursive commentary from past scholars who have assembled prosodic bibliographies to preserve the history of prosodic discourse. This would include, but would not be limited to, Edwin Guest, Goold Brown, Coventry Patmore, George Saintsbury, and T.S. Omond, all of whom published their bibliographies with commentary prior to 1923 and therefore are out of copyright.
- allow contemporary users to contribute to that historical conversation through threaded comments

4 Collections & Communities

In the process of building the PPA, we asked professors, librarians, and Office of Information Technology staff to think with us about the very basic questions the history of poetry teaching posed in the larger fields of text-encoding and text-mining. In these conversations we realized we would need to build the DH community that could support and sustain our project in the long term, and so we did. The collaborations we began with our project have evolved into the concurrent founding of a fully-fledged Digital Humanities Center at Princeton, for which we are now advertising for an Associate Director.

5 References

Brogan, Terry V. F. 1981. *English Versification, 1570-1980: A Reference Guide with a Global Appendix*. Johns Hopkins University Press.

Hall, Jason David. 2011a. “Materializing Meter: Physiology, Psychology, Prosody.” *Victorian Poetry* 49 (2): 179–197. doi:10.1353/vp.2011.0019. http://muse.jhu.edu/journals/victorian_poetry/v049/49.2.hall.html.

———, ed. 2011b. *Meter Matters: Verse Cultures of the Long Nineteenth Century*. Edited & translated by . 1st ed. Ohio University Press.

Prins, Yopie. 2008. "Historical Poetics, Dysprosody, and \textlessi\textgreaterThe Science of English Verse\textless/i\textgreater." *PMLA* 123 (1) (jan): 229–234. doi:10.1632/pmla.2008.123.1.229. <http://www.mla-journals.org/ezproxy.columbia.edu/doi/abs/10.1632/pmla.2008.123.1.229>.