

Mary Erica Zimmer and Douglas Duhaime
Chicago Colloquium on Digital Humanities and Computer Science (DHCS)
23-24 October 2014
Northwestern University
Evanston, IL
Poster Submission

From search to serendipity: computational approaches to intertextuality in the work of Geoffrey Hill

The poetry of Geoffrey Hill is consistently noted for its depth and breadth of reference—as well as the perceived barrier to entry this combination can present, especially for those new to Hill’s *œuvre*. While many have considered Hill’s debts to individual authors and texts, further dimensions of his reading’s living presence within his work remain to be explored. Here, as David-Antoine Williams has noted, computational approaches hold promise. Frequently, criticism of Hill’s poetry has emphasized its “difficulty” for present-day audiences, given what many describe as the poems’ obscurities of language and reference, as well as their oft-compressed syntax.¹ Yet as Hill’s wryly expressed frustrations with the term “difficulty” suggest, such comments incompletely engage both matter and manner of his work.

Wrestling with Hill’s poems involves grappling with both their substance and their manner of marshaling reference. Particularly when phrases or items recur, possessing some sense of intertexts Hill may have in mind when writing proves useful. As nearly eighty years of reading imbue Hill’s works, however, audiences drawn to consider his complexities can lack orientation—a sense of where to begin, that is, when contexts surrounding and informing the work remain recondite. For those seeking interpretive purchase, attempting to fix meaning of a word or phrase can also prove problematic, as such an approach may not address fully ways key terms and collocations animate Hill’s works over time. In these instances, as Matthew Sperling has observed, non-serial forms of reading can shed light.

Corpus-based approaches provide a window into not only the breadth of Hill’s reading, but also the multiple registers in which a word or phrase may prove significant. Through varied forms of search, one may illuminate historical and linguistic environments—“contextures,” in Hill’s parlance—that become visible as key words and phrases intersect with other texts. Perceiving these verbal settings allows one to hear nuances of resonance such words and phrases may hold for Hill. The rewards are often rich. For Hill, engaging truly challenging literature involves assessing its originality in light of literary and etymological history. One might understand Hill’s own work in similar terms.

As Sperling has written, the ability “to point out” (*indicare*) provided by an index can serve as point of departure for further exploration. In this case, the project’s evolving methods have moved towards modeling habits of mind one might

¹ These aspects are discussed in an unpublished paper by Martin Dodsworth.

form over a lifetime of reading and thought. Advances have been made through development of a full-text-searchable index, which facilitates rapid performance of fuzzy searches as well as direct string matches. Using the Python package Whoosh to create lemmatized, searchable indexes has allowed for non-serial approaches that have significantly decreased query time while conserving data that may aid in identifying fuzzy text reuse. Through iteration, underlying cruces informing Hill's thought can emerge.

In light of particular challenges posed by Hill's work, this poster presents a case study in exploring influence in and across corpora, as well as in and through time. Hill's extensive, ongoing, and frequently discussed engagement with early modern literary culture prompts inclusion of the EEBO-TCP corpus, as its 45,000+ texts from 1475-1700 present a range of concerns and texts specifically noted as germane to Hill's thought and writing. Intersecting with this early modern register is a far more contemporary vein, one Hill himself has noted as shaping what his work has both been and become. These more proximate literary currents are represented through the 46,000+ texts of Project Gutenberg, which contain not only recognizable High Modernist influences Hill has cited, but also texts ranging further afield, along lines often unexpected in their emergence. Given Hill's commitment to exploring non-English texts in their original languages, the Latin Library is also used to shed light on potential patterns in Hill's practice.

In the main, our approaches help human readers note veins of textual influence, ones likely to inform the Hillian work or works being considered. The project's indicial reading strategies show textual indebtedness while revealing patterns of emphasis and reconsideration within Hill's own substantial corpus. Far from decreasing the need for human perspective, however, these computer-assisted methodologies orient attention, along paths where qualitative exploration may well engender insight. By presenting approaches to date in light of ongoing questions, the poster will illuminate "affordances" of various computational approaches, in terms complementing Hill's own sense of his works' development.

--

Mary Erica Zimmer is a Ph.D. candidate in Editorial Studies in The Editorial Institute at Boston University whose dissertation will serve as a companion to selected works of Geoffrey Hill. Through the NEH "Early Modern Digital Agendas" Summer Institute at the Folger Shakespeare Library, she has developed dual digital projects concerned with textual networks and literary influence and is particularly drawn to intersections between the two.

Douglas Duhaime is a doctoral student studying digital approaches to eighteenth-century literature and science at the University of Notre Dame. His research focuses on the intersections of early modern natural philosophy and classical political economy in eighteenth-century literary works—relationships he has pursued by using natural language processing techniques, running hand-written scripts on text data from EEBO, ECCO, and the Philosophical Transactions to trace patterns in Enlightenment-era literary history.