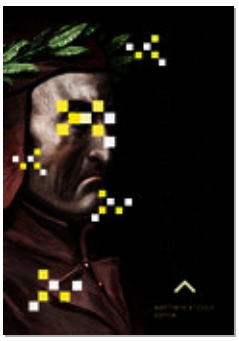
Debates in the Digital Humanities

Matthew K. Gold



Published by University of Minnesota Press

The Ancestral Text

MICHAEL WITMORE

In this post I want to understand the consequences of "massive addressability" (Witmore) for "philosophies of access"—philosophies that assert that all beings exist only as correlates of our own consciousness. The term "philosophy of access" is used by members of the speculative realist school: it seems to have been coined largely as a means of rejecting everything the term names. Members of this school dismiss the idea that speculative analysis of the nature of beings can be replaced by an apparently more basic inquiry into how we access the world, an access obtained through either language or consciousness. The major turn to "access" occurs with Kant, but the move is continued in an explicitly linguistic register by Heidegger, Wittgenstein, Derrida, and a range of poststructuralists.

One reason for jettisoning the priority of access, according to Ray Brassier, is that it violates "the basic materialist requirement that being, though perfectly intelligible, remain irreducible to thought" ("The Enigma of Realism"). As will become clear, I am sympathetic to this materialist requirement and more broadly to the speculative realist project of dethroning language as our *one and only* mode of access to the world. (There are plenty of ways of appreciating the power and complexity of language without making it the wellspring of Being, as some interpreters of Heidegger have insisted.) Our quantitative work with texts adds an unexpected twist to these debates: as objects of massive and variable address, texts are "handled" in precisely the ways usually reserved for nonlinguistic entities. When taken as objects of quantitative description, texts possess qualities that—at some point in the future—could be said to have existed in the present, *regardless* of our knowledge of them. There is thus a temporal asymmetry surrounding quantitative statements about texts: if one accepts the initial choices about what gets counted, such statements can be "true" now even if they can only be produced and recognized later. Does this asymmetry, then, mean that language itself, "though perfectly intelligible, remain[s] irreducible to thought" (Brassier)? Do iterative methods allow us to satisfy Brassier's materialist requirement in the realm of language itself?

Let us begin with the question of addressability and access. The research described on this blog involves the creation of digitized corpora of texts and the mathematical description of elements within that corpus. These descriptions obtain at varying degrees of abstractions (nouns describing sensible objects, past verb forms with an auxiliary, etc.). If we say that we know something quantitatively about a given corpus, then we are saying that we know it on the basis of a set of relations among elements that we have provisionally decided to treat as countable unities. Our work is willfully abstract in the sense that, at crucial moments of the analysis, we foreground relations as such, relations that will then be reunited with experience. When I say that objects of the following kind—"Shakespearean texts identified as comedies in the First Folio"—contain more of a certain type of thing (first- and second-person singular pronouns) than objects of a different kind (Shakespeare's tragedies, histories), I am making a claim about a relation between groups and what they contain. These groupings and the types of things that we use to sort them are provisional unities: the circle we draw around a subset of texts in a population could be drawn another way if we had chosen to count other things. And so, we must recognize several reasons why claims about these relations might always be revised.

Every decision about what to count offers a caricature of the corpus and the modes of access this corpus allows. A caricature is essentially a narrowing of address: it allows us to make contact with an object in some of the ways Graham Harman has described in his work on vicarious causation. One can argue, for example, that the unity "Shakespeare's Folio comedies" is really a subset of a larger grouping or that the group can itself be subdivided into smaller groups. Similarly, one might say that the individual plays in a given group aren't really discrete entities and so cannot be accurately counted in or out of that group. There are certain words that *Hamlet* may or may not contain, for example, because print variants and multiple sources have made *Hamlet* a leaky unity. (Accommodating such leaky unities is one of the major challenges of digital text curation.) Finally, I could argue that addressing these texts on the level of grammar—counting first- and second-person singular pronouns—is just one of many modes of address. Perhaps we will discover that these pronouns are fundamentally linked to semantic patterns that we haven't yet decided to study but should. All of these alternatives demonstrate the provisional nature of any decision to count and categorize things: such decisions are interpretive, which is why iterative criticism is not going to put humanities professors out of business. But such counting decisions are not—and this point is crucial—simply another metaphoric reduction of the world. Principal component analysis (PCA), cluster analysis, and the other techniques we use are clearly *inhuman* in the number of comparisons they are able to make. The detour through mathematics is a detour *away* from consciousness, even if that detour produces findings that ultimately converge with consciousness (i.e., groupings produced by human reading).

Once the counting decisions are made, our claims to know something *in a statistical sense* about texts boils down to a claim that a particular set of relations

pertains among entities in the corpus. Indeed, considered mathematically, the things we call texts, genres, or styles simply *are* such sets of relations—the mathematical reduction being one of many possible caricatures. But counting is a very interesting caricature: it yields what is there now—a real set of relations—but is nevertheless impossible to contemplate at present. Once claims about texts become mathematical descriptions of relations, such statements possess what the philosopher Quentin Meillassoux calls ancestrality, a quality he associates primarily with statements about the *natural* world. Criticizing the ascendance of what he calls the Kantian dogma of correlationism—the assumption that everything that can be said "to be" exists only as correlate of consciousness—Meillassoux argues that the idealist or critical turn in Continental philosophy has impoverished our ability to think about anything that exceeds the correlation between mind and world. This "Great Outdoors," he goes on to suggest, is a preserve that an explicitly speculative philosophy must now rediscover, one that Meillassoux believes becomes available to us through mathematics. So, for example, Meillassoux would agree with the statement, "the earth existed 4.5 billion years ago," precisely because it can be formulated mathematically using measured decay rates of carbon isotopes. The statement itself may be ideal, but the reality it points to is not. What places the Great Outdoors out of doors, then, is its indifference to our existence or presence as an observer. Indeed, for Meillassoux, it is *only* those things that are "mathematically conceivable" that exceed the post-Kantian idealist correlation. For Meillassoux, "all those aspects of the object that can be formulated in mathematical terms can be meaningfully conceived as properties of the object in itself." Clearly such a statement is a goad for those who place mind or natural language at the center of philosophy. But the statement is also a philosophical rallying cry: be curious about objects or entities that do not reference human correlates! I find this maxim appealing in the wake of the "language is everything" strain of contemporary theory, which is itself a caricature of the work of Wittgenstein, Derrida, and others. Such exaggerations have been damaging to those of us working in the humanities, not least because they suggest that our colleagues in the sciences do *nothing but* work with words. By making language everything—and, not accidentally, making literary studies the gatekeeper of all disciplines—this line of thought amounts to a new kind of species narcissism. Meillassoux and others are finding ways to not talk about language all the time, which seems like a good thing to me.

But would Meillassoux, Harman, and other speculative realists consider texts to be part of the Great Outdoors? Wouldn't they have to? After all, statements about groupings in the corpus can be true now even when there is no human being to recognize that truth as a correlate of thought. Precisely *because* texts are susceptible to address and analysis on a potentially infinite variety of levels we can be confident that a future scholar will find a way of counting things that turns up a new but as yet unrecognized grouping. Human reading turned up such a thing when scholars in the late nineteenth century "discovered" the genre of Shakespeare's late

romances. (Jonathan Hope and I have, moreover, redescribed these groupings statistically [Witmore and Hope].) Like our future mathematical sleuth might do a century from now, nineteenth-century scholars were arguing that romance was already a *real* feature of the Shakespearean corpus, albeit one that no one had yet recognized. They had, in effect, picked out a new object by emphasizing a new set of relations among elements in a collection of words. Couldn't we expect another genre to emerge from this sort of analysis—a genre X, let's say—given sufficient time and resources? Would we accept such a genre if derived through iterative means?

I can imagine a day, one hundred years from now, when we have different dictionaries that address the text on levels we have not thought to explore at present. What if someone creates a dictionary that allows me to use differences in a word's linguistic origin (Latinate, Anglo-Saxon, etc.) to relate the contents of one text to another? What if a statistical procedure is developed that allows us to "see" groupings we *could* recognize today but simply have not developed the mathematics to expose? When you pair the condition of massive addressability with (1) the possibility of new tokenizations (new elements or strata of address) or (2) the possibility that all token counts past and future can be subjected to new mathematical procedures, you arrive at a situation in which something that is arguably true now about a collection of texts can only be known in the future.

And if something can be true about an object now without itself being a correlate of human consciousness, isn't that something part of the *natural world,* the one that is supposed to be excluded from the charmed circle of the correlation? Does this make texts more like objects in nature, or objects in nature more like texts? Either way, the Great Outdoors has become larger.

NOTE

This chapter originally appeared as "The Ancesteral Text" (http://winedarksea.org/?p=979).

BIBLIOGRAPHY

Brassier, Ray. "The Enigma of Realism: On Quentin Meillassoux's After Finitude." In *Collapse: Philosophical Research and Development.* Vol. 2. 2007.

Meillassoux, Quentin. *After Finitude: An Essay on the Necessity of Contingency*. New York: Continuum, 2008.

Witmore, Michael. "Text: A Massively Addressable Object." In *Wine Dark Sea.* December 31, 2010. http://winedarksea.org/?p=926.

Witmore, Michael, and Jonathan Hope. "Shakespeare by the Numbers: On the Linguistic Texture of the Late Plays." In *Early Modern Tragicomedy*, edited by Subha Mukherji and Raphael Lyne, 133–53. Woodbridge, Suffolk, UK: D. S. Brewer, 2007.

This page intentionally left blank