DHQ: Digital Humanities Quarterly

2009 Volume 3 Number 4

Conjectural Criticism: Computing Past and Future Texts

Kari Kraus <karimkraus_at_gmail_dot_com>, College of Information Studies and Department of English, University of Maryland

Abstract

Broadly conceived, this article re-imagines the role of conjecture in textual scholarship at a time when computers are increasingly pressed into service as tools of reconstruction and forecasting. Examples of conjecture include the recovery of lost readings in classical texts, and the computational modeling of the evolution of a literary work or the descent of a natural language. Conjectural criticism is thus concerned with issues of transmission, transformation, and prediction. It has ancient parallels in divination and modern parallels in the comparative methods of historical linguistics and evolutionary biology.

The article develops a computational model of textuality, one that better supports conjectural reasoning, as a counterweight to the pictorial model of textuality that now predominates in the field of textual scholarship. "Computation" is here broadly understood to mean the manipulation of discrete units of information, which, in the case of language, entails the grammatical processing of strings rather than the mathematical calculation of numbers to create puns, anagrams, word ladders, and other word games. The article thus proposes that a textual scholar endeavoring to recover a prior version of a text, a diviner attempting to decipher an oracle by signs, and a poet exploiting the combinatorial play of language collectively draw on the same library of semiotic operations, which are amenable to algorithmic expression.

The intended audience for the article includes textual scholars, specialists in the digital humanities and new media, and others interested in the technology of the written word and the emerging field of biohumanities.

In an essay published in the *Blackwell Companion to Digital Literary Studies*, Stephen Ramsay argues that efforts to legitimate humanities computing within the larger discipline of literature have met with resistance because well-meaning advocates have tried too hard to brand their work as "scientific," a word whose positivistic associations conflict with traditional humanistic values of ambiguity, open-endedness, and indeterminacy [Ramsay 2008]. If, as Ramsay notes, the computer is perceived primarily as an instrument for quantizing, verifying, counting, and measuring, then what purpose does it serve in those disciplines committed to a view of knowledge that admits of no incorrigible truth somehow insulated from subjective interpretation and imaginative intervention [Ramsay 2008, 479–482]?

1

2

Ramsay's rich and nuanced answer to that question has been formulated across two publications, of which "Toward an Algorithmic Criticism" is the earliest.^[1] Drawing on his experience as both a programmer and a scholar, he suggests that the pattern-seeking behavior of the literary critic is often amenable to computational modeling. But while Ramsay is interested in exposing the algorithmic aspects of literary methods, he concedes that in searching for historical antecedents, we may be more successful locating them in the cultural and creative practices of antiquity rather than in scholarly and critical practices [Ramsay 2003, 170]. His insight is born of the fact that a text is a dynamic system whose individual units are ripe for rule-governed alteration. While these alterations may be greatly facilitated by mechanical means, historically they have been performed manually, as in the case of the "stochastic" aesthetics of the *I Ching* or the "generative aesthetics" of the Dadaists and Oulipo [Ramsay 2003, 170]. "Ironically," Ramsay writes, "it is not the methods of the scholar that reveal themselves as 'computational,' but the methods of the gematrist and the soothsayer" [Ramsay 2003, 170]. Ramsay's emphasis on prophecy is apt, for it allows us to link the divinatory offices of the poet

and soothsayer with the predictive functions of the computer. It is interesting to note in this context that one of the earliest uses to which a commercial computer was put was as an engine of prediction: the UNIVersal Automatic Computer (UNIVAC), the first digital electronic computer to debut in the United States, successfully predicted the outcome of the 1952 Eisenhower-Stevenson presidential race.^[2] The computer's pedigree in symbolic logic, model-based reasoning, and data processing make it a predictive machine *par excellence*.

With these observations in mind, the principal purpose of this essay is to establish an ancient lineage for what we might call *computational prediction in the humanities*, a lineage that runs counter to Ramsay's assertions insofar as it is distinctly critical or scholarly rather than aesthetic in character. Specifically, I argue that within the field of textual scholarship, conjectural criticism (the reconstruction of literary texts) can be profitably understood as a sub-domain of Ramsay's algorithmic criticism (rule-governed manipulation of literary texts). The essay seeks to do the following:

3

4

5

- First, it challenges the assumption that "computational processes within the rich tradition of interpretive endeavors [are] usually aligned *more with art than criticism* " (my emphasis) [Ramsay 2003, 167]. Although sometimes erroneously considered "pre-interpretive," ^[3] textual scholarship, one of our oldest branches of literary study, is a fundamentally critical rather than artistic endeavor.
- Second, I provide preliminary evidence suggesting not only a formal, but also a genealogical relationship between Mesopotamian prophecy and textual scholarship, one that sanctions a prospective as well as retrospective view of the evolution of a text. Such a proleptic viewpoint has, in fact, already been experimentally adopted in evolutionary biology and historical linguistics, disciplines whose conjectural methods parallel those of textual scholarship (about which see below).
- Third, it offers an alternative theory for why the analogy between science and the digital humanities apparently has so little traction among traditional literary critics. Ramsay's contention is that the tasks at which a computer excels are frequently judged, rightly or wrongly, to be more compatible with the certitude of hard knowledge than the incertitude of soft knowledge. An equally compelling (and complementary) reason is that the computer is too often seen exclusively as a device for processing numbers rather than manipulating text. But if the computer can count numbers like a mathematician, it can also play with letters in ways, as we shall see, that poets and textual scholars alike would recognize.
- Fourth, the essay reconfigures the relationship between humanities computing and science by exploring the common denominator of literacy rather than numeracy. In other words, the essay brings into focus those scientific disciplines that manipulate a symbol system that semiotically *operationally*, we might say more closely resembles letters than numbers, namely linguistics and bioinformatics. The essay further posits that when viewed historically, the relationship among the three disciplines isn't unidirectional, but rather deeply reciprocal. There is compelling circumstantial evidence that the conjectural methods of historical linguistics and evolutionary biology were largely derived from those of textual criticism in the nineteenth century. Subsequently, however, linguistics and biology were (and to some extent remain) more active in exploiting the computer's ability to execute algorithms on strings as a way to predict, model, and analyze evolutionary change or in other words to automate the methods first expounded by textual critics more than a century earlier.

Broadly conceived, then, this essay re-imagines the role of conjecture in textual scholarship at a time when computers are increasingly pressed into service as tools of reconstruction and forecasting. Examples of conjecture include the recovery of lost readings in classical texts and the computational modeling of the evolution of a literary work or the descent of a natural language. Conjectural criticism is thus concerned with issues of transmission, transformation, and prediction (as well as retrodiction).

In the first section of the essay I attempt to define conjecture, often viewed in the humanities as a misguided and antimethodical pursuit, and rationalize it as a form of subjunctive knowledge, knowledge about what might have been or could be or almost was. The object of conjecture is notional rather than empirical; possible rather than demonstrable; counterfactual rather than real. This subjunctive mode, I contend, is not antithetical to the humanities, but central to it. Whether it is a student of the ancient Near East deciphering a fragmented cuneiform tablet or a musician speculatively completing Bach's unfinished final fugue or a literary scholar using advanced 3D computer modeling to virtually restore a badly damaged manuscript, the impulse in each instance — vital and paradoxical — is to go beyond purely documentary states of objects.

The essay develops a computational model of textuality, one that better supports conjectural reasoning, as a counterweight to the material model of textuality that now predominates. *Computation* is here broadly understood to mean the systematic manipulation of discrete units of information, which, in the case of language, entails the grammatical processing of strings^[4] rather than the mathematical calculation of numbers to create puns, anagrams, word ladders, and other word games. The essay thus proposes that a textual scholar endeavoring to recover a prior version of a text, a diviner attempting to decipher an oracle by signs, and a poet exploiting the combinatorial play of language collectively draw on the same library of semiotic operations, which are amenable to algorithmic expression and simulation.

6

7

8

9

10

In an effort to overcome common objections to analogies between the humanities and sciences,^[5] I offer a reflection — part historical, part formal or theoretical — on the parallel importance of the tree paradigm in evolutionary biology, textual criticism, and historical linguistics. By *tree paradigm*, I mean the grouping of manuscripts, languages, or genomes into families and showing how they relate to one another in genealogical terms, and using these relationships to conjecture about lost ancestors or archetypes. My themes are the interdependence and correspondence of tree methods (and their associated algorithms) in three diverse knowledge domains. From tree structures, I turn to a form of causal transmission known as a transformation series, a recognized phenomenon in the three disciplines surveyed. Because of their similarity to a genre of word puzzle known as doublets, transformation series allow us to revisit the connection between wordplay and divination established earlier in the argument.

A Rationale for Conjecture

In textual scholarship, conjecture is the proposal of a reading not found in any extant witness to the text. It is predicated on the idea that words are always signs of other words; the received text also harbors the once and future text. When A. E. Housman alters a line from Catullus's *Marriage of Peleus and Thetis* (c. 62-54 BCE) so that it reads "*Emathiae tutamen, Opis carissime nato*," instead of "*Emathiae tutamen opis, carissime nato*," he is practicing the art of conjecture. ^[6] Housman's small orthographic tweaks produce sizable semantic shifts: by relocating a comma and capitalizing "Opis" so that it is understood not as the Latin term for "power," but as the genitive of "Ops," the mother of Jupiter in Roman mythology, he substantially alters the meaning of the line: "Peleus, most dear to his son, is the protector of the power of Emathia," becomes "Peleus, protector of Emathia, most dear to the son of Ops [Jupiter]." The emendation, as the fictional Housman explains in Tom Stoppard's *The Invention of Love*, restores sense to nonsense by ridding the language of anachronism: "How can Peleus be *carrisme nato*, most dear to his son, when his son has not yet been born?" [Stoppard 1997, 37].

Expressed in grammatical terms, conjecture operates in the subjunctive and conditional moods.^[7] Because its range of motion extends beyond the pale of the empirical, its vocabulary is replete with *coulds*, *mights*, *mays*, and *ifs*. Such a vocabulary reflects not caution — on the contrary, conjecture is a radical, audacious editorial style — but rather a refusal to settle for attested states of texts. In this it is opposed to the current wave of archival models of editing, which, in their rejection of speculative and inferential readings, opt for representation in the indicative mood.

Conjecture has at various moments in history held a place of honor in our repertory of editorial paradigms. Taken collectively, for example, Housman's critical writings offer a magnificent apologia of *divinatio*, as conjecture is still sometimes called. The strength of conviction, the peremptory eloquence, the sheer depth and breadth of knowledge contained in those pages brook no protests about the bugbear of intentionalist editing or the death of the author. But at the same time, conjecture is a foundling, strangely lacking a well-defined history, theory, methodology. Discussion of it, which more often than not occurs incidentally, tends to take place not on an expository or historical plane, but on a homiletic or poetic one. Thus, for example, amid all his scientific discourse on compounded variational formulas, ancestral groups, successive derivation, and other arcane topics pertaining to recension, W. W. Greg falls almost quiet when he turns to the subject of conjecture: "the fine flower" of textual criticism he calls it, choosing metaphor over

definition [Greg 1924, 217]. Treatments of conjecture characterized more by brevity than rigor are the norm: the process is hazily glossed metaphorically or else grounded in the successful critic's abstract qualities of mind, which do not always lend themselves to analysis (e.g., intuition, judgment, confidence, insight, authority, charisma). For these reasons, conjecture has customarily been seen as either a practitioner's art, the conspicuous absence of a meta-literature a symptom of the bottom-up approach it tends to favor, or a prophet's.

The scholarly language can at times suggest the intervention of a supernatural agent. The classicist Robin Nisbet speaks of a "Muse of Textual Conjecture," whom he playfully christens, appropriately enough, Eustochia.^[8] As one response to the exquisitely difficult recovery mission that conjecture sets for its critics, Eustochia might be perceived as a sad contrivance to a misguided endeavor; a *deus ex machina* called in to artificially resolve all textual difficulties. But consider, for a moment, the bleak alternative: a world in which conjectural knowledge never was or could be an article of faith. Samuel Johnson's plaintive verdict that conjecture "demands more than humanity possesses" [Johnson 1765, para. 105] gives our imaginations something to work with; now envision that same despondency on a collective rather than individual scale. Instead of Erasmus proudly proclaiming the reconstruction of ancient texts the noblest task of all, we'd have only Stoppard's Housman making a Faustian pact, born of despair, to suffer the fate of Sisyphus if only each time the stone rolled back he were given another fragment of Aeschylus [Stoppard 1997].

What does this general silence indicate about conjecture, its status in textual criticism, and its practice? E. J. Kenney's position (which he officially co-opts from the nineteenth-century critic Boeckh, but which nevertheless seems to be prevalent among Classical editors) that the solution to the more "difficult cases" of conjecture "comes in a flash or not at all" [Kenney 1974, 147] is tantalizing, but the idea of a lucid eureka moment — eerily similar to the romantic notion of inspired creativity ("At its inception, a [romantic] poem is an involuntary and unanticipated donnée" [Abrams 1953, 24]) — shouldn't curtail discussion; it should open the floodgates. And because the archival model holds conjecture in abeyance, an entire new generation of young textual critics who have cut their critical teeth on documentary precepts are all but forbidden at the outset to engage in it. For this reason alone conjectural criticism should inspire inquiry, but also precisely because it pretends to elude historical, cognitive, or technical understanding.

Would it be possible for conjecture to assume again a position of prominence in editorial theory? It would take a concerted inquiry into history, method, and theory. We would need to seriously engage the psychology of intuition, the AI and cognitive literature of inference, and the history of conjecture as divinatory art and scientific inquiry. We would need to keep apace with advancements in historical linguistics and evolutionary biology (two fields whose relevance to textual criticism is vastly under-appreciated) and the burgeoning psycholinguistic literature of reading. Indeed, we would reap many rewards from outreach to the linguistic community. We would need to dispel the myth that conjecture is a riddle wrapped in a mystery inside an enigma. A curriculum for conjecture would give us better insight into the method underlying the prediction — or retrodiction — of ancient readings of texts. Not least, we would need to work decisively to bring conjectural criticism into the 21st century. Because it has traditionally been described as a balm to help heal a maimed or corrupted text, conjecture is in desperate need of a facelift; the washed-up pathological metaphors long ago ceased to strike a chord in editorial theory.

So what would an alternate language of conjecture sound like?

We might, for starters, imagine conjecture as a knowledge toolkit designed to perform "what if" analyses across a range of texts. In this view, the text is a semiotic system whose discrete units of information can be artfully manipulated into alternate configurations that may represent past or future states. Of course the computing metaphors alone are not enough; they must be balanced by, among other things, an appreciation of the imponderable and distinctly human qualities that contribute to conjectural knowledge. But formalized and integrated into a curriculum, the various suggestions outlined here have the potential to give conjecture a new lease on life and incumbent editorial practices — much too conservative for a new generation of textual critics — a run for their money.

A Formal Definition and Methodological Overview

What do I mean by conjecture, then? Giovanni Manetti's pithy definition — "inference to the imperceptible" — is good 16

11

12

13

insofar as it captures the cognitive leap ("inference") from the known to the unknown.^[9] But it fails to denote the temporal dimension of conjecture so essential to a field like textual criticism, which deals with copies of texts that are produced in succession and ordered in time. Moreover, inference is not a monolithic category but can be further subdivided into inductive inference, deductive inference, abductive, intuitive, probabilistic, logical, etc. Manetti's definition thus needs to be massaged into something more apropos to textual scholarship. For the purposes of this essay, I'll define it as follows: conjecture is allographic inference to past or future values of the sign. Allographic is a semiotic term given currency by the analytic philosopher Nelson Goodman in his seminal Languages of Art and refers to those media that resolve into discrete, abstract units of information — into what might be more loosely called digital units - that can be systematically copied, transmitted, added to, subtracted from, transposed, substituted, and otherwise manipulated. Conjecture as it is understood in this work is indifferent as to whether the mode of allographic manipulation is manual, cognitive, or mechanical. Allographic contrasts with autographic, or continuous, media. Examples of the former include musical scores, alphabetic script, pixels, and numerical notation; examples of the latter, paintings, drawings, and engravings. Painter and theorist Julian Bell summarizes Goodman like this: "Pictures differ from other sign systems, such as writing, by being continuities in which every mark is interdependent, rather than operating through a combination of independent markers like the alphabet. (In computational terminology, they are 'analog' rather than 'digital' representations.)" [Bell 1999, 228]. Words are conventionally allographic, images autographic, though much important work in recent years has contested the validity of these distinctions and examined borderline cases, such as pattern poetry. Indeed, the last two decades of textual criticism have witnessed a wealth of scholarship promoting the text's bibliographic or iconic codes. Despite the virtues of such visual approaches to textuality (and there are many), conjectural criticism has been hamstrung by their success. The general argument in this essay is that conjecture flourishes in an allographic environment, not an autographic one. By "allographic inference," then, I mean the considered manipulation or processing of digital signs with the goal of either recovering a prior configuration or predicting a future or potential one.

If we set aside for a moment the original terms of my definition (allographic inference) and focus instead on the terms used to gloss them (digital processing), then one of the salient points to emerge is that conjecture as formalized in these pages is semiotic in the first instance and computational in the second. Conjecture, that is to say, involves digital signal processing or, to use the more general-purpose term, *computation*.^[10] Moreover, just as digitality, as a concept, extends beyond the boxy electronic machines that sit on our desktops, so too does computation. As David Alan Grier reminds us in When Computers Were Human, homo sapiens has been computing on clay tablets, papyrus, parchment, and paper for millennia [Grier 2005]. A computer is simply someone or something that systematically manipulates discrete, abstract symbols. As conventionally understood, those symbols are numbers, and the manipulations performed on them are addition, subtraction, multiplication, and division. But computers are text manipulators as much as number crunchers. The subfield of computer science that operates on textual rather than numeric data is known as stringology. Reflecting this broader emphasis, the symbols that are the focus of this essay are primarily alphabetic, but also phonemic and molecular. The only criterion for the symbols to function conjecturally, as already stated in our definition, is that they be digital, or allographic. The processes discussed here and throughout likewise depart from computational norms. They are not always or only mathematical operations, but also semiotic or grammar-like operations: transposition, deletion, insertion, substitution, repetition, and relocation. I will have more to say about them later, but for now let me summarize the four poles around which string computation has been organized in this project:

- 1. Input: allographic/digital (texts, words, letters, phonemes, molecular sequences)
- 2. Processor: human or machine
- 3. Processes or Operations: semiotic or grammar-like (transposition, insertion, deletion, relocation, substitution)
- Output: textual errors (scribal copying); conjectural reconstructions or projections (textual criticism, historical linguistics, evolutionary biology); wordplay (poetry); auguries (divination); plain text or cipher text (cryptography)

This framework allows us to associate activities, behaviors, and practices that wouldn't otherwise be grouped together: wordplay, divination, textual transmission, and conjectural reconstruction, for example. It also allows us, as we shall see, to perceive a number of underlying similarities among a set of disciplines that for most of the twentieth century didn't take much interest in one another, namely textual criticism, historical linguistics, and evolutionary biology.

Historically, we have referred to the source of a message about the future as an "oracle" or "sibyl" or "seer"; within the contexts specified in this essay, we can alternately refer to that agent, whether human or mechanical, as a "computer." From this standpoint, divination is not "knowledge in advance of fact" so much as knowledge that is a (computational) permutation of fact [Jonese 1948, 27]. Justin Rye, for example, author of the imaginary language "Futurese," a projection of American English in the year 3000 AD, proposes that the word "build" will be pronounced /bll/ in some American dialects within a century or more as a consequence of consonant cluster simplification [Rye 2003]. Underlying this speculation is a simple deletion operation that has been applied to a factual state of a word to compute a plausible future state. Transition rules of this kind are key ingredients in all of the examples of conjecture discussed in this project: the wordplay of Shakespeare's fools, soothsayers, and madmen; the puns of the Mesopotamian or Mayan priest; the "emendations" of textual scholars; the doublets of Lewis Carroll; the molecular reconstructions of geneticists; the projections of conlangers (creators of imaginary languages); the transformation series of historical linguists, textual scholars, and evolutionary biologists.

These devices and systems are unified by their semiotic and, in some cases, cognitive properties and behaviors: a pun resembles an editorial reconstruction resembles a speech error resembles a genetic mutation, and so forth. Each of these transitive relations is explored in the course of this essay. Consider, as a preliminary example, the analogy between textual and genetic variation. It is the allographic equivalence between the two — the fact that both are digitally encoded — that makes possible a strong theory of translation, allowing one system to be encoded into the other and, once encoded, to continue to change and evolve in its new state. Brazilian-born artist Eduardo Kac, who coined the term *transgenic art* to describe his experiments with genetic engineering, exploits this allographic equivalence in *Genesis* (1999), a commissioned work for the Ars Electronica exhibition. Kac's artist statement reads in part as follows:

The key element of the work is an "artist's gene," a synthetic gene that was created by Kac by translating a sentence from the biblical book of Genesis into Morse Code, and converting the Morse Code into DNA base pairs according to a conversion principle specially developed by the artist for this work. The sentence reads: "Let man have dominion over the fish of the sea, and over the fowl of the air, and over every living thing that moves upon the earth." It was chosen for what it implies about the dubious notion of divinely sanctioned humanity's supremacy over nature. The Genesis gene was incorporated into bacteria, which were shown in the gallery. Participants on the Web could turn on an ultraviolet light in the gallery, causing real, biological mutations in the bacteria. This changed the biblical sentence in the bacteria. The ability to change the sentence is a symbolic gesture: it means that we do not accept its meaning in the form we inherited it, and that new meanings emerge as we seek to change it. [Kac 1999]

Kac's elaborate game of code-switching is enabled by what Matthew Kirschenbaum calls "formal materiality," the condition whereby a system is able to "propagate the illusion (or call it a working model) of immaterial behavior." [Kirschenbaum 2008, 11]. Because this model is ultimately factitious, succeeding only to the extent that it disguises the radically different material substrates of the systems involved, it cannot be sustained indefinitely. But at some level this is only to state the obvious: abstractions leak,^[11] requiring ongoing regulation, maintenance, and modification to remain viable. This fundamental truth does not diminish the power or utility — what I would call the *creative generativity* — of models.

Conjecture as Wordplay

I first became sensitive to the convergence of conjecture and wordplay while studying Shakespeare as a graduate student. The ludic language of Shakespeare's fools, soothsayers, and madmen seemed to me to uncannily resemble the language of Shakespeare's editors as they juggled and transposed letters in the margins of the page, trying to discover proximate words that shadow those that have actually descended to us in the hopes of recovering an authorial text. The pleasure George Ian Duthie, a postwar editor of Shakespeare, shows in permutating variants — juxtaposing

20

21

22

and repeating them, taking a punster's delight in the homophony of stockt, struckt, and struck; hare and hart; nough and nought^[12] — finds its poetic complement in the metaplasmic imagination of Tom O'Bedlam in *King Lear* or the soothsayer Philarmonus in *Cymbeline*. Kenneth Gross points to what he calls Tom's "near echolalia" on the heath in the storm scene, the "strange, homeless babble" that "presses up from within Tom's lists, in their jamming up and disjunctions of sense, their isolation of bits of language[:] . . . toad/tod-pole, salads/swallows, wall-newt/water." [Gross 2001, 184]. There is something scribal and exegetical in Tom's babble, just as there is something poetic and ludic in Duthie's editing. And in the way both manipulate signs, there is also something fundamentally conjectural.^[13] Conjecture, then, can be profitably understood by adopting a semiotic framework — expressed here in computational terms — within which we can legitimately or persuasively correlate the rules and patterns of conjectural transformation with those of mantic codes and wordplay.

Consider, too, the soothsayer Philarmonus, solicited by Posthumus near the close of *Cymbeline* to help decipher Jupiter's prophecy, inscribed on a scroll that serves as a material token of Posthumus's cryptic dream-vision:

When as a lion's whelp shall, to himself unknown, without seeking find, and be embrac'd by a piece of tender air; and when from a stately cedar shall be lopp'd branches which, being dead many years, shall after revive, be jointed to the old stock, and freshly grow; then shall Posthumus end his miseries, Britain be fortunate and flourish in peace and plenty (*Cymbeline*, 5.5.436-443)

The soothsayer responds by translating keywords into Latin, isolating their homophones, and expounding their relations through puns and false etymologies:

[To CYMBELINE]

The piece of tender air, thy virtuous daughter, Which we call *mollis aer*, and *mollis aer* We term it *mulier*, which *mulier* I divine Is this most constant wife, who even now Answering the letter of the oracle, Unknown to you, unsought, were clipp'd about With this most tender air (*Cymbeline*, 5.5.447-453)

Language here is an anagrammatic machine. By sleight of sound, Philarmonus^[14] morphs *mollis aer* (tender air) into *mulier* (woman) and cracks the code. "Every fool can play upon the word," quips *The Merchant of Venice*'s Lorenzo 3.5.41. And so can every prophet. What the incident in *Cymbeline* highlights is that the "capaciousness of ear" with which Kenneth Gross attributes Hamlet applies in equal measure to Shakespeare's soothsayers, clowns, fools, tricksters, madmen — and, I would add, editors [Gross 2001, 13]. It is significant, for example, that Posthumus should initially mistake Jupiter's oracle for the ravings of a madman, for the madman and the prophet hear and speak in a common register. The fool, too, engages in the same digital remixing of sound and sense, and it is worth recalling in this context that Enid Welsford's classic study of the fool's social and literary history devotes an entire chapter to his dual role as poet and clairvoyant [Welsford 1966, 76–112]. It is not only *titular* fools (or madmen or prophets or clowns — the terms quickly proliferate and become functionally equivalent) but also *de facto* fools who, in *King Lear* for example, interface between wordplay and prophecy. The epithet of fool, like that play's ocular and sartorial puns, is passed from one character to another, so that by the end of the play one is hard pressed to identify a single character who *hasn't* tried on motley for size. By turns the fool, Lear, Kent, Cordelia, Gloucester, and Albany bear the brunt of the term, lending credence to Lear's lament that his decrepit world is but a "great stage of fools" (4.6.187).

Let me try to connect the figure of the fool, like that of the prophet, more directly to the motifs of error, wordplay, and

conjecture. A self-styled "corrupter of words" (3.1.34), Feste in *Twelfth Night*, for example, is fluent in the dialect of riddles and puns that constitutes the shared vernacular of Shakespeare's fools. He subscribes to a philosophy of language that views the manipulation of sounds as a form of sympathetic magic — a process of *world*-making through *word*-making (and un-making). In the following exchange with Viola, the bawdy and jocular terms in which Feste expresses this philosophy should not distract us from the radical theory of the sign that underpins it: that there exists an instrumental link between signifier and signified, such that reconfiguring the one provides a means for fashioning, shaping, or foreshadowing the destiny of the other:

Viola:	Nay, that's certain; they that dally nicely with words may quickly make them wanton
Feste:	I would, therefore, my sister had had no name, sir.
Viola:	Why, man?
Feste:	Why, sir, her name's a word; and to dally with that word might make my sister wanton.

Considered within the context of these lines, "corrupter" denotes someone who has the power to pervert the world by deforming ("dallying with") the language used to signify it. But Shakespeare mobilizes other meanings as well. In the domain of textual scholarship, the term *corrupt* has historically been used to describe a text riddled with copy errors. Because Feste's cunning mispronunciations and nonsense words hover between intelligibility and unintelligibility, they closely parallel scribal error: "I did impeticos thy gratillity" (2.3.27), Feste nonsensically proclaims to Sir Andrew, prompting one critic to wryly note that "words here are at liberty and have little meaning apart from that which editors, at the cost of great labor, finally manage to impose upon them" [Grivelet 1956, 71]. The "great labor" to which Michel Grivelet alludes involves exploiting the digitally encoded phonological structure of language to conjecturally recover authentic words from corrupt ones: *peticoat* from *impeticos*, or *gratuity* (perhaps *gentility*?) from *gratillity*. Impressionistically, then, it can often feel as though the play's *textual* "corruptions" — easily confused with its fictional *vocal* corruptions - were originating from within the story rather than from without; as if the text were inverting itself, such that the source of error were imaginary rather than real, with the fool becoming not only the *object* of a flawed transcription but also — impossibly — the *agent* of it. Because they mimic textual corruption, Feste's broken, disfigured puns seem to compress temporality, exposing what the linguist D. Stein would call "diachronic vectors in synchrony." ^[15]

Let me pause here to gather together a number of threads: my central tenet is that the linguistic procedures of Shakespeare's editors often seem to mimic those of Shakespeare's most inveterate "computers" of language. To emend a text — to insert, delete, and rearrange letters, phonemes, or sequences of words — is to ritualistically invoke a divinatory tradition of wordplay that dates back to at least the third millennium BC. For a benign alien power observing human textual rites from afar, the linguistic manipulations of a Duthie or a Housman would, I imagine, be for all intents and purposes indistinguishable from those of a Philarmonus or Feste or Tom O'Bedlam or Chaldean or, for that matter, a computer programmer using string-rewriting rules to transform one word into another. What all these examples have in common are digital units manipulated by computers (programmers, soothsayers, madmen, fools, punsters, poets, scribes, editors) using a small set of combinatorial procedures (insertion, deletion, transposition, substitution, relocation) for conjectural, predictive, or divinatory ends. Duthie and Tom O'Bedlam, computer programmers and Kabbalists, Philarmonus and historical linguists: they all compute bits of language.

As an editorial method, such alphabetic computation is dismissed by Housman as a wanton orthographic game. Finding a kindred spirit in a nineteenth-century German predecessor, he approvingly quotes the following:

Some people, if they see that anything in an ancient text wants correcting, immediately betake themselves to the art of palaeography . . . and try one dodge after another, as if it were a game, until they hit upon something which they think they can substitute for the corruption; as if forsooth

25

26

truth were generally discovered by shots of that sort, or as if emendation could take its rise from anything but a careful consideration of the thought. (Qtd. in Housman 1921.)

And yet the conjectures of, for example, Richard Bentley (1662-1742), one of the few textual critics for whom Housman professes admiration, are as much the product of letter play and combinatorial art as they are of careful deliberation and thought. (Given the structure and workings of the mental lexicon, discussed below, we might reasonably assert that letter play is the *sine qua non* of conjectural thought.) Lamenting the corrupt state of the blind Milton's *Paradise Lost*, which, according to Bentley, was not only dictated to a tone-deaf amanuensis, but also, adding insult to injury, later copy-edited by a derelict acquaintance, Bentley attaches a preface to his edition of the poem that includes a table of the "monstrous Faults" disfiguring the masterpiece, alongside Bentley's proposed emendations [Bentley 1732]. Reflecting the editor's conviction that "Words of a like or near Sound in Pronunciation" were substituted throughout for what Milton intended [Bentley 1732], the table reads like a dictionary of wordplay: "is Judicious" is emended to "Unlibidinous"; "Nectarous" to "Icarus"; "Subtle Art" to "Sooty Chark"; "Wound" to "Stound"; "Angelic" to "Adamic." ^[16] The same "near echolalia" that characterizes Tom O'Bedlam's speech in the storm scene also characterizes Bentley's editorial method.

To adjudicate among variants, Bentley finds evidence and inspiration in diverse knowledge domains. Milton's "*secret* top of Horeb" in the opening lines of *Paradise Lost* is conjecturally restored to "sacred top" through an appeal to literary tradition, geology, meteorology, and logic. According to Bentley, "secret top" finds little precedent in the works of antiquity, while "sacred top" has parallels in the Bible, Spenser, and various classical authors [Bentley 1732, 1n6]. Emending a text by bringing it into alignment with literary antecedents is a technique in which Bentley takes recourse again and again, as though a poem were best thought of as a commonplace book filled with its author's favorite quotations. The model of authorship that underwrites this approach is one that stresses sampling and collage over originality and solitary genius. The more allusive and intertextual the work, the more susceptible it will be to conjectural reconstruction.

Whatever the shortcomings of Bentley's appeal to precedent, it has the effect of helping systemize and guide his wordplay. His paronomastic methods are constrained by other factors as well: the double articulation of natural language, comprised of a first level of meaningful units (called morphemes) and a second level of meaningless units (called phonemes), imposes order and rules on the process. There are in English a total of twenty-six letters of the alphabet capable of representing some 40-45 separate phonemes, and phonotactic and semantic constraints on how those letters and sounds may be combined. The sequence "ptk" in English, for example, is unpronounceable, and the sequence "paf," while pronounceable, is at the time of this writing meaningless, except perhaps as an acronym. Bentley's wordplay is thus bounded by the formal and historically contingent properties of the natural language with which he works, properties that theoretically prevent his substitutions and transpositions from degenerating into mere gibberish.

The importance of digital or allographic units to computation, as I have defined it here, is underscored by Gross's observation that what Tom O'Bedlam manipulates are "*bits* of language" (my emphasis). Likewise, Elizabeth Sewell repeatedly emphasizes that nonsense poetry and wordplay require "the divisibility of its material into ones, units from which a universe can be built" [Sewell 1952, 53]. She goes on to say that this universe "must never be more than the sum of its parts, and must never fuse into some all-embracing whole which cannot be broken down again into the original ones. It must try to create with words a universe that consists of *bits*" (my emphasis) [Sewell 1952, 53–4]. It is precisely the fusion of parts into an indivisible whole that has made pictures — particularly dense, mimetic pictures — historically and technologically resistant to manual computation and, by extension, conjectural reconstruction.

Textual Transmission as Wordplay

The similarities between scribal and poetic computation (understood in the sense in which I have defined it) are brought home by even a casual look at the editorial apparatus of any critical edition of Shakespeare, where editors have traditionally attempted to ascertain whether a particular verbal crux is a poetic device in need of explication or a misprint in need of emendation. The Shakespearean text is one in which an error can have all the color and comedic effect of an intentional malapropism, and an intentional malapropism all the ambiguity and perplexity of an inadvertent error. Is

28

29

30

Cleopatra's "knot intrinsicate" a deliberate amalgam of "intricate" and "intrinsic," one that deploys "half a dozen meanings," or, by contrast, an accidental blend introduced into the text by a distracted compositor? Is Hamlet's "sallied flesh" an alternative spelling of "solid," one that also plays on "sullied," or, more mundanely, a typesetting mistake? Does the Duke's "headstrong weedes" in *Measure for Measure* encapsulate some of the play's core themes, or is it an inadvertent and nonsensical deformation of "headstrong *steeds*"?^[17] We are in general too hasty, insists M. M. Mahood in *Shakespeare's Wordplay*, in dismissing these and other odd readings as erroneous variants rather than appreciating them for what they in many instances are: the artful "twists and turns" of a great poet's mind [Mahood 1957, 17].

That Mahood feels compelled to devote several pages to the problem of distinguishing between errors and wordplay in Shakespeare's text is worth lingering over. It suggests that there is something fundamentally poetic about the changes — or *computations* — a text undergoes as it is transmitted through time and space. This holds true regardless of how we classify those computations, whether as errors, creative interpolations, or conjectural emendations. It is as if a poem, at its most self-referential and rhetorical, were recapitulating its own evolution, or the evolution recapitulating the poem.

To clarify the point, consider the phenomenon of metathesis, the simple transposition of elements. Metathesis is a law of sound change (historical linguistics), a cipher device (cryptography), a poetic trope (classical rhetoric), a scribal error (textual criticism), and a speech error (psycholinguistics). There are entries for metathesis in both *A Handlist of Rhetorical Terms* (a guide to formal rhetoric) and *A Companion to Classical Texts* (a primer on textual criticism, which includes a typology of scribal error).^[18] The overlaps between the two reference works are, in fact, considerable, with entries for epenthesis (the insertion of an element), homeoteleuton (repetition of words with similar endings) and other basic manipulations common to both. While it would be a stretch to claim that the lists contained in them are interchangeable, it wouldn't be much of one. Wordplay, one might conclude, is a poetics of error.

Cognitive evidence supports this view. Clinical experiments, slips of the tongue, slips of the pen, and the speech and writing disorders of aphasics (those whose ability to produce and process language has been severely impaired by a brain injury) all provide insight into the structure and organization of the mental lexicon. Current research models that lexicon as a web whose connections are both acoustic and semantic: like-sounding and like-meaning words are either stored or linked together in the brain [Aitchison 2003]. These networks of relation make certain kinds of slips or errors more probable than others. As natural language users, for example, we sometimes miss a target word when speaking or writing and instead replace it with one that occupies a nearby node, making substitutions like profession for procession or medication for meditation relatively common [Aitchison 2003, 145]. Transpositions like paschetti for spaghetti or blends like frowl (frown plus scowl) are also legion and suggest that the unit of production and manipulation is the individual phoneme or grapheme: discrete, digital, and mobile [Aitchison 2003, 216]. For many aphasics, these errors completely overtake normal speech: clip might become plick; butter, tubber; or leasing, ceiling [Aitchison 2003, 22]. As Jean Aitchison observes, "the problems of aphasic patients are simply an exaggeration of the difficulties which normal speakers may experience" [Aitchison 2003, 22]. But what is uncontrollable in the aphasic is deliberate in the poet. "There is a sense in which a great poet or punster is a human being able to induce and select from a Wernicke aphasia," writes George Steiner in After Babel [Steiner 1998, 297]. To which I would add: or from everyday speech and writing errors.

Steiner's insight repays further attention. Poets, like artists in general, often creatively stress-test the system or medium with which they work, probing its edges, overloading it, and pushing it beyond normal operational capacity. Discovering where language breaks down or deviates from regular use is the business of both poet and neuroscientist, providing a means of gaining insight into the mechanisms of language perception, processing, and production. But whereas the neuroscientist gathers data from aphasic patients in a clinical setting, the poet becomes, as it were, his own research subject, artificially manipulating the cognitive networks of meaning and sound that form his dataset. As Steiner notes, when intentionally created, ordered, and embedded in larger textual or linguistic structures, such distortions become poetry.

Divination as Wordplay

The mantic wordplay in Shakespeare's Cymbeline has forerunners in divination methods indigenous to cultures as

36

35

32

diverse as those of the Quiché Maya of South America and the Sumerians of the ancient Near East. Mayan priests still practice a type of calendrical prophecy that involves uttering the words of a specific date and interpreting their oracle by punning on them [Tedlock 2000, 263]. Notably, the Quiché term for punning is sakb'al tzij, "word dice," which in the context of Mayan daykeeping makes links among wordplay, divination, and games explicit [Tedlock 2000, 263]. Later, we shall examine the kinds of string metrics that one can apply to puns and textual variants to compute optimal alignment, edit distance, and edit operations. What these are and why one might want to compute them are the subjects of another section; for now, it is enough to note that these algorithms rest on the premise that words and texts change procedurally — by adhering to the same finite set of legal operations we've previously discussed.

We can also demonstrate the formalism of divinatory practice by looking at some of the earliest inscribed prophecies of the ancient Near East, which make extensive use of the same conditional blocks found in modern computer programs to control the flow and "output" of the mantic code:

If it rains (zunnu iznun) on the day (of the feast) of the god of the city - [then] the god will be (angry) (zeni) with the city. If the bile bladder is inverted (nahsat) - [then] it is worrying (nahdat). If the bile bladder is encompassed (kussa) by the fat - [then] it will be cold (kussa). [Manetti 1993, 10]

The divinatory apparatus consists of an "if" clause (in grammatical terms the protasis), whose content is an omen; and the "then" clause (the apodosis), whose content is the oracle. We should be wary about assuming, as James Franklin does, that the divinatory passage from omen to oracle is an entirely arbitrary one ("all noise and no laws" [Franklin 2001, 162]). As the examples make clear, the passage is allographically motivated, "formed by the possibility of a chain of associations between elements of the protasis and elements of the apodosis," [Manetti 1993, 7] specifically a phonemic chain in the examples given above. What we are looking at, once again, is a jeux de mots. One word metamorphoses into another that closely resembles it by crossing from the IF to the THEN side of the mantic formula. [19]

The divinatory mechanism may be construed as a program for generating textual variants: the baru, the divinatory priest 39 or technician, inputs the protasis-omen into the system. A finite set of legal operations (substitution, insertion, deletion, relocation, transposition) is performed on its linguistic counters, which are then output in their new configuration to the apodosis-oracle.^[20] Understood algorithmically, the Mesopotamian divinatory code is a proto-machine language, one that precedes by several millennia Pascal or Java or C++.[21]

Textual Criticism as Divination

The allographic operations used to generate a future text in Mesopotamian divination are formally consonant with those 40 that produce variant readings in an open print or manuscript tradition. Mesopotamian divination simulates textual transmission, by which I mean that it seems to accelerate or exaggerate, as well as radically compress, a process that would normally occur over centuries or millennia — and one, moreover, that would in a great many of its particulars occur inadvertently rather than (as in Mesopotamian divination) deliberately. It is as if we were watching a text descend to posterity through time-lapse photography. These conceptual and behavioral associations are visually reinforced by the traditional emblems of the baru, which consist of a writing tablet and stylus.^[22]

The scholarship on Mesopotamian omen sciences published within the last two decades suggests that divinatory systems in the ancient Near East exerted profound influence not only on the content of the literary genres of the ancient world, including the written record of the Israelites, but also on the interpretive practices of their exegetes.^[23] The affinity between Sumerian divination and early textual scholarship of the Hebrew Bible is expertly established by Michael Fishbane in Biblical Interpretation in Ancient Israel, an extraordinary study of the scribal phenomenon of inner-biblical exegesis. Fishbane writes as follows:

Sometimes these [Mesopotamian prophecies] are merely playful jeux de mots; but, just as commonly, there is a concern to guard esoteric knowledge. Among the techniques used are

41

37

permutations of syllabic arrangements with obscure and symbolic puns, secret and obscure readings of signs, and numerological ciphers. The continuity and similarity of these cuneiform cryptographic techniques with similar procedures in biblical sources once again emphasizes the variegated and well-established tradition of mantological exegesis in the ancient Near East — a tradition which found ancient Israel a productive and innovative tradent [i.e., transmitter of the received tradition]. [Fishbane 1989]

Fishbane's work demonstrates through exhaustive research that mantic practice of the ancient Near East is a crucially important locus for understanding early textual analysis and emendation of the Hebrew Bible. Our working hypothesis must therefore be that the connections between these hermeneutic traditions are causal as well as formal. Historically, then, and contrary to popular belief, prediction has been as much a part of the knowledge work of the humanities as of the sciences — as much our disciplinary inheritance as theirs. It is time that we own that legacy rather than disavow it. We have for some time now outsourced conjecture to the Natural, Physical, and Computational Sciences. Paradoxically, then, it is only through interdisciplinarity that we can possibly hope to reclaim disciplinarity.

Trees of History

This section looks at the shared "arboreal habits," to use Darwin's term [Darwin 1871, 811], of conjectural critics in three historical disciplines: textual criticism, evolutionary biology, and historical linguistics. More than a century after the publication of Darwin's *On the Origin of Species*, which helped popularize the Tree of Life, tree methodology remains at the center of some of the most ambitious and controversial conjectural programs of our time, including efforts to reconstruct macrofamilies of languages, such as Proto-Indo-European, and the Last Universal Common Ancestor (LUCA), a single-celled organism from which all life putatively sprang.^[24] At the outposts of the biological and linguistic sciences, conjecture is prospective as well as retrospective: synthetic biologists and language inventors are known to experiment with modeling future states of genomes and languages, respectively.^[25]

In an article entitled "Trees of History in Systematics and Philology," Robert J. O'Hara, a biologist with impressive interdisciplinary credentials, draws attention to the presence of "trees of history," glossed as "branching diagrams of genealogical descent and change," across the disciplines of evolutionary biology, historical linguistics, and textual criticism. Figures 1, 2, and 3 after O'Hara reproduce three such trees [O'Hara and Robinson 1993, 7–17], all published independently of one another within a forty-year span in the nineteenth-century.

43

44



Figure 1. Darwin's Tree of Life (1859)



Figure 1 is Darwin's well-known tree of life from *On the Origin of Species* (1859). Figure 2, the first manuscript stemma ever printed, diagrams the lines of ancestry and descent among a group of medieval Swedish legal texts (1827). And

Figure 3 is a genealogy of the Slavic languages (1853). As an evolutionary biologist — a *systematist* to use the technical term — O'Hara's interest in "trees of history" stems from his work in cladistics, a system of classification based on phylogenetic relationships in evolutionary groups of organisms. Over the last several decades, cladistics has become a highly sophisticated computer-assisted methodology, one that has outpaced stemmatics and its linguistic counterpart technically if not conceptually. It is unsurprising, then, that in addition to satisfying a general theoretical curiosity about the coincidence of trees of history in three seemingly disparate fields, O'Hara's research is also an outreach project. He has, for example, collaborated with Peter Robinson to produce cladistic analyses of Chaucerian manuscripts [O'Hara and Robinson 1993, 53–74]. Such outreach has awakened both textual critics and linguists to the application of bioinformatics to their respective fields. But while O'Hara triangulates all three historical sciences, the link between textual scholarship and linguistics receives comparatively less attention than the links each of them shares with evolutionary biology. It is this third leg of the triangle that I want first to consider.

The evidence suggesting a direct historical relationship between stemmatics in textual criticism and the comparative method in linguistics is relatively hard to come by in the secondary literature, and what little of it there is, is circumstantial. The English-language textual scholarship has almost nothing to say on the subject,^[26] although I suspect the situation with respect to German-language scholarship may be different, and the linguistic and biological literature for the most part falls in step. Standard accounts of the comparative method, as the genealogical approach is known in historical linguistics, are more apt to note the influence of Darwinian principles than textual ones. But a small cadre of linguists and biologists, following the lead of the late Henry Hoenigswald, have set about revising what we think we know about tree visualizations in the nineteenth century.^[27] Hoenigswald, a historical linguist as well as a historian of linguistics, has managed to unearth interesting connections by operating at a granular level of analysis.^[28] He succeeds where others fail because he has something of a detective mentality about him: his work demonstrates that to successfully trace the itinerary of ideas, one sometimes has to be willing to track the movements of the individuals who propagated them. Who knew who when, where, and under what circumstances? His lesson is that it is useful to think in terms of coteries and their membership if the objective is to map the spread of ideas. While other scholars write very generally about the diffusion of the arboreal trope in the nineteenth century, Hoenigswald — and perhaps we see the imprint of the genealogical method here — attempts to discover direct lines of influence.

Any account of the origins of the comparative method in historical linguistics must take into account the achievements of A. Schleicher (1821-1868), the founding father of the *Stambaumtheorie*, or genealogical tree model, by which the relations of the variously known IndoEuropean languages were set forth in a pedigree, and an asterisked hypothetical ancestral form postulated. *Comparative method* is a misnomer to the extent that it suggests comparison is the focal activity of the method, when in fact it is more like a means to an end. Comparisons of variants between two or more related languages are undertaken with the purpose of *reconstructing* their genealogy and proto-language. While the method is partially indebted to the Linnean taxonomic classification system, Hoenigswald is quick to point out a second, perhaps more significant influence: as a student, Schleicher studied classics under Friedrich Ritschl (1806-1876), who, along with other nineteenth-century figures such as Karl Lachmann and J. N. Madvig, helped formalize the science of stemmatology [Hoenigswald 1963, 5]. The affinities between textual criticism and historical linguistics are such that very minor adjustments to the definition of comparative method would produce a textbook definition of stemmatics:

Comparative method: Comparisons of variants between two or more related languages are undertaken with the purpose of reconstructing their genealogy and proto-language.

Stemmatics: Comparisons of variants between two or more related texts are undertaken with the purpose of reconstructing their genealogy and archetype.^[29]

Formal Theories

In order to better understand how an unattested text, language, or genome is inferred from attested forms, I want to briefly look at two competing classes of stemmatic algorithms: maximum parsimony and clustering. Broadly speaking, clustering or distance methods compute the overall similarities between manuscript readings, without regard to whether

47

48

the similarities are coincidental or inherited,^[30] while maximum parsimony computes the "shortest" tree containing the least number of change events still capable of accommodating all the variants. The difference is often expressed in evolutionary biology as that between a phenetic (clustering) versus a cladistic (parsimonious) approach. Here is how Arthur Lee distinguishes the two:

Cladistic analysis is sharply differentiated from cluster analysis by that which it measures. Cluster analysis groups the objects being analysed or classified by how closely they resemble each other in the sum of their variations, using statistical "distance measures." Cladistic analysis, on the other hand, analyses the objects in terms of the evolutionary descent of their individual variants, choosing the evolutionary tree which requires the smallest number of changes in the states of all the variants. ^[31]

In cladistic analyses, the evidentiary gold standard excludes most of the textual data from consideration: only shared derived readings as opposed to shared ancestral readings are believed to have probative value. An ancestral reading, or retention, is any character string inherited without change from an ancestor. A derived reading, by contrast, is a modification of a reading inherited from the ancestor, often motivated rather than arbitrary from a paleographic, metrical, phonological, manual, technological, aesthetic, grammatical, cognitive, or some other perspective (for example, the *Canterbury Tales* scribe who, when confronted with the metrically deficient line "But tel me, why hidstow with sorwe" in the Wife of Bath's Prologue, alters "why" to "wherfor" to fill out the meter). When such readings are contained in two or more manuscripts, they are regarded as potentially diagnostic. Having made it through this initial round of scrutiny, they are then subject to a second round of inspections, which will result in exclusions of variants that have crept into the text via routes not deemed genealogically informative. The transposition of the *h* and the *e* of *the* when keyed into a computer is a mundane example. So pedestrian is this error that your word-processing program most likely automatically corrects it for you when you make it. In the terminology of Don Cameron, it is an *adventitous* rather than *indicative* variants.

Phenetic or clustering analyses involve the use of a distance metric to determine degrees of similarity among manuscripts. The Levenshtein Distance algorithm, for example, named after the Russian scientist who created it, tabulates the number of primitive operations required to transform one variant into another. We can illustrate its application with one of the most notorious variants in the Shakespearean canon. In Hamlet's first soliloquy, we encounter the following lines:

49

50

O that this too too sullied flesh would melt, Thaw and resolve itself into a dew . . . (1.2.129-30)

The first folio reads "solid flesh," while the second quarto has "sallied flesh." The Levenshtein or edit distance between them is determined as follows: let "solid" equal the source reading and "sallied" the target reading. The first task is to align the words so as to maximize the number of matches between letters:^[33]

so-li-d | | | sallied

Sequence alignment, as the method is called, is a form of collation. Identifying all the pairings helps us compute the minimum number of procedures needed to change one string into another. A mismatch between letters indicates that a substitution operation is called for, while a dash signals that a deletion or insertion is necessary. Three steps get us from source to target: a substitution operation transforms the "o" of "solid" into the "a" of "sallied"; an insertion operation supplies the "l" of "sallied"; and another insertion operation gives us the penultimate "e" in our target variant:

solid - >

salid salid - > sallid sallid - > sallied

A distance of three edit operations (one substitution and two insertions) thus separates the two character strings. In

clustering, the greater the edit distance between different readings, the more dissimilar they are assumed to be.

Although heterogeneous in their approach, both clustering and maximum parsimony are *ars combinatoria*. That is, both involve making assumptions about and/or modelling various orders, numbers, combinations, and kinds of — as well as distances between — primitive operations that collectively account for all the known variants in a text's transmission history. Both also take entropy for granted: they assume the depredations of time on texts to be relentless and non-reversible, with the result that the more temporally remote a descendant is from an ancestor, the more dissimilar to the ancestor it will be. In terms of our tree graph, this means that a copy placed proximally to the root is judged more similar to the archetype than a copy placed distally to the root.^[34]

Conjecture as Wordplay (Redux)

In biology, textual criticism, and historical linguistics, a transformation series (TS) is an unbroken evolutionary sequence of character states. The changes that a word or molecular sequence undergoes are cumulative: state A gives rise to state B, which in turn generates C, and so on. The assumption is that each intermediary node serves as a bridge between predecessor and successor nodes; state C is a modified version of B, while D is a derivation of C. The changes build logically on one another and follow in orderly succession. Consider, for example, a hypothetical (and admittedly idealized) textual series: head/heal/teal/tell/tall.^[35] Each word differs from the next by exactly one letter. The transitions are gradual rather than abrupt, striking a balance between continuity and change. "If we can discover such a transformation series," writes H. Don Cameron, "we have strong evidence for the relationship of the manuscripts" in which the readings occur.^[36]

Those readings might relate to one another in the vertical direction, in which case the task of the textual critic is to determine their order and polarity:

53

52



Conversely, they might relate to one another as leaf nodes in the horizontal direction,^[37] under which circumstances the editorial challenge is to infer vertical relationships:



From this diagram, one can reconstruct plausible vertical variants, which recapitulate the horizontal:



In this graph, each internal node generates two children, one of which preserves the reading of the parent, the other of which innovates on it. The innovations, or mutations, collectively constitute the TS and form a single evolutionary path through the tree.

Both distance and parsimony methods shed light on transformation series. The degree of difference between two leaf nodes, for example, reflects their degree of relationship: in our reconstruction, *tail* and *tall*, which differ by only one character, share a more recent common ancestor than *tail* and *head*, which differ by four.^[38] The variants are also ordered so as to minimize the total number of mutations along the vertical axis, consistent with the principle of parsimony.

Now an admission: the source of the head/heal/teal/tell/tall/tall series is the nineteenth-century British author and mathematician Charles Lutwidge Dodgson, better known by the pseudonym Lewis Carroll, who uses it to illustrate not textual transmission or biological evolution, but rather a word game of his own invention, which he called "doublets" [Carroll 1992, 39]. Alternately known as "word ladders," doublets is played by first designating a start word and end word. The objective is to progressively transform one into the other, creating legitimate intermediary words along the way. The player who can accomplish this in the fewest number of steps wins.

The congruity between doublets and biological evolution is the subject of an ingenious essay by scientist David Searls, entitled "From *Jabberwocky* to Genome: Lewis Carroll and Computational Biology." ^[39] Searls begins his section on doublets by quoting the opening lines of Carroll's nonsense poem "Jabberwocky," famous for its portmanteaux, neologisms, and word puzzles:

Twas brillig, and the slithy toves Did gyre and gimble in the wabe; All mimsy were the borogoves, And the mome raths outgrabe.

"Jabberwocky," was originally published in *Through the Looking Glass* (1871), Carroll's sequel to *Alice's Adventures in Wonderland* (1865). In chapter 6, the character of Humpty Dumpty, who comments at some length on the poem's vocabulary, interrupts Alice's recitation to observe of the first verse that "there are plenty of hard words there. 'BRILLIG' means four o'clock in the afternoon — the time when you begin BROILING things for dinner." ^[40] "Upon hearing this explanation," writes Searls, computational biologists "will of course feel an irresistible urge to do something like the following:"

broil-ing | | | | | br-illi-g

[Searls 2001, 340].

That "something" is sequence alignment, a preliminary step to computing the edit distance between the two strings and creating an edit transcript specifying the semiotic operations required to mutate one into the other. "Jabberwocky," manifests the same metamorphic impulse that would later guide Carroll's invention of other word games. Of the gradated form of doublets in particular, Searls remarks that "the notion of the most parsimonious interconversions among strings of letters is . . . at the basis of many string-matching and phylogenetic reconstruction algorithms used in computational biology" [Searls 2001, 341] — and, of course, in textual criticism and historical linguistics.

Interestingly, Carroll himself alludes to the resemblance between doublets and Darwinian evolution in an example that anticipates by more than half a century the linguistic and scriptural foundations of modern molecular genetics. He satirically bridges the evolutionary gap between apes and humans with a six-step transformation series:

Extending Carroll's experiment to sets of words, Searls reconstructs hypothetical ancestors for eight leaf nodes:

55

56

54

57

59



Here the words at the bottom level are used to infer ancestors, but they might also be used to project descendants. 60 They support conjectures that *converge* backward in time or *diverge* forward in time.

While Carroll initially conceived of doublets as a game of substitutions, involving the exchange of one letter for another, he later also admitted transpositions, as in the third step of the following series:

 $\label{eq:icon} \begin{array}{rcl} \text{IRON} \ - \ > \ \text{ICON} \ - \ > \ \text{CORN} \\ & - \ > \ \text{CORD} \ - \ > \ \text{LORD} \ - \ > \ \text{LOAD} \ - \ > \ \text{LEAD} \end{array}$

And although deletions and insertions were never formally a part of doublet puzzles, they were central to another game exploring word generation, which Carroll dubbed "Mischmasch." ^[41] Indeed, taken collectively, Carroll's games incorporate the full spectrum of grammatical operations on strings, almost as if they were intentionally designed to model patterns of descent with variation in cultural and organic systems.

What should we make of the fact that scientific theories of transmission and reconstruction can be so effectively illustrated with a parlor game? Recall that a correspondence between conjecture and wordplay was established earlier in this essay. There we discovered that a textual critic endeavoring to recover a prior text and a diviner attempting to decipher an oracle by signs were often united in their reliance on letter substitutions, puns, anagrams, and other permutational devices. Noting that the procedures underlying such word games were amenable to algorithmic expression, we labeled both the textual critic and the prophet who perform them "computers." At the same time, we maintained that they were computers of a special sort, proficient in the semiotic processing of strings rather than the mathematical reckoning of numbers. Into their ranks we can also admit evolutionary biologists and historical linguists, for whom the grammatical operability of signs is no less germane.

In a post-industrial, Westernized society, an individual with the proper education who displays an aptitude for string manipulation might find gainful employment as an evolutionary biologist or historical linguist or cryptographer. In another milieu, that same individual might instead be inducted into the ways of the poet or prophet. The Mayan priest who practices calendrical divination through puns, the Mesopotamian baru who deciphers oracles through wordplay, and the Shakespearean soothsayer who interprets auguries by means of substitutive sounds and false etymologies find their 21st-century scientific complement in the figure of the computational biologist. Searls remarks that Carroll's doublet puzzles reveal "a turn of mind well suited to methodologies used in modern computational biology." [Searls 2001, 339] This statement has the effect of endowing what might otherwise be seen as an idle pastime with scientific rigor and purpose. But we could also flip the words around to observe that the computational methods of the biologist reveal a turn of mind well suited to wordplay. By thus inverting the language, we begin to see the poet in the scientist as well as the scientist in the poet. The contention of this poet-scientist is that all living things evolve through combinatorial play. The Book of Life, it would seem, is a volume of poetry, a dictionary of wordplay, and a compendium of oracles all bound together.

Coda

In this essay I have tried to advance the proposition that conjectural knowledge in the humanities is a manifestation of

62

the inalienable human need to imagine what might have been or could be or almost was. This mandate of the subjunctive is beautifully conveyed by George Steiner in *After Babel*:

It is the constructive powers of language to conceptualize the world which have been crucial to man's survival in the face of ineluctable biological constraints, this is to say in the face of death. It is the miraculous — I do not retract the term — capacity of grammars to generate counter-factuals, "if"-propositions and, above all, future tenses, which have empowered our species to hope, to reach far beyond the extinction of the individual. We endure, we endure creatively due to our imperative ability to say "No" to reality, to build fictions of alterity, of dreamt or willed or awaited "otherness" for our consciousness to inhabit. [Steiner 1998, xiii–xiv]

For Steiner, that "imperative ability" is more than a psychological or linguistic impulse; more, even, than a creative or ethical force; it is also a life instinct.

In the course of this work I have also tried to show that algorithms on strings — vital to conjecture — are as computationally sound as algorithms on numbers. We are so accustomed to thinking that punsters and textual critics somewhat imprecisely reshuffle symbols, whereas mathematicians and engineers methodically calculate them that it will perhaps take a class of scientists who process strings rather than numbers to convince us otherwise. We have looked at just such a group of professionals, computational biologists, whose work on the chemical "letters" of DNA resembles nothing so much as the parlor game of doublets or word ladders. And yet this form of wordplay is couched in the language of computation rather than poetry, vetted by biologists rather than poets or literary scholars, and published in scientific journals rather than poetry rags or literary monographs. Perhaps most significantly, the similarities between the methods of the computational biologist and the textual scholar — like those between the soothsayer and textual scholar — are not coincidental, but causal and historical.

One of my primary objectives has been to dispute the commonplace assumption that the inner workings of conjecture are ineffable and opaque. R. J. Tarrant, for example, writes that "conjectural criticism advances slowly and unsteadily," proceeding at a "glacial pace" and is not amenable to "discussion of methods and trends" [Tarrant, 121]. Remarks of this sort are scattered throughout the literature of textual scholarship: E. J. Kenney and Martin West profess similar sentiments, as does F. W. Hall, who remarks that, traditionally, conjecture "proceeds from no method and conveys no certainty" [Hall 1913, 151]. My position, by contrast, has been that conjectural knowledge is produced in more determinate ways than is customarily acknowledged, which is to say through not only (or necessarily) the statistical processing of discrete symbols, but also their grammatical or semiotic processing. Stated more generally, there are reciprocal relations among conjecture, computation, and semiotics. This essay, then, has been written with the conviction that a cogent theory of conjecture is a desideratum of textual studies.

But what about conjectural methods in humanistic fields other than textual studies and linguistics that don't rely on systems of inheritance and variation? Take history, for instance, where the historical imaginary is our predominant tool for filling in the missing details of the past; or digital preservation, where the archivist must designate a primary user community for the cultural objects she preserves. By projecting the knowledge base of this user group into the future, the archivist makes decisions about what kind of descriptive metadata and contextual information to include in order to "render the object intelligible" to posterity.^[42] If she misjudges the needs and attributes of this community — if her conjectural faculties falter — then the archive may be less meaningful, accessible, or useful than it would otherwise be. Are these two forms of conjecture so very different, either from each other or from those already described, as to generally make comparisons between them baseless? Or is it possible to create an integrated research narrative capable of accommodating a variety of conjectural projects, perhaps through recourse to a sophisticated typology of conjecture?

My own hunch is that there are deep structures common to many types of conjecture and new methods available to others that await discovery and experimentation. Our success in identifying and exploiting them, however, will depend in no small measure on how committed the digital humanities community is in coming years to building an information infrastructure that supports new models of scholarly communication and research. Patrick Juola, a computational

67

68

linguist at Duquesne University, for example, is currently developing a prototype system for exploring corpus-based approaches to conjecture in the humanities. As Juola explains, his idea is to program the computer "to make essentially random emendations to texts and then rank them for plausibility using a set of agreed-upon criteria," the underlying assumption being that the machine "is capable of devising and examining many more possible readings than even the most prolific human author, and even if 99.9% of the possible readings are flat gibberish, the one in a thousand may include interesting and provocative readings that human authors have missed" (Patrick Juola, email to the author, 3 Oct. 2008.) Noting that this data-driven approach can be extended in a variety of ways, Juola emphasizes the versatility of the comparative framework he is proposing:

I could equally compare American and British literature, or 19th and 20th century literature, or firstperson and third-person novels, or any other interesting category. Essentially, what the prototype I envision will do is generate (and test superficially) random conjectures like [Works in Category 1] use [Semantic category B] more than [Works in Category 2]. (Juola, email to the author, 3 Oct. 2008.)

The promise of Juola's corpus-based conjectural criticism, however, is predicated on the kind of mass digitization from which humanities researchers, as Christine Borgman emphasizes in *Scholarship in the Digital Age*, have thus far benefited very little, at least as compared to their scientific colleagues.^[43] The *data* content layer, moreover, lacks the infrastructural support that libraries, institutions, and publishers have traditionally provided for the *document* content layer.^[44] The problem is exacerbated in the humanities by the fact that the distinction between data and publications (or data and documents) is fuzzy to begin with [Borgman 2007, 219]. If we can create the kind of institutional incentives recommended by Borgman to develop, manage, and curate the data content layer, then perhaps a new practice of conjecture will follow.

Acknowledgments

I would like to express my gratitude to the following friends, colleagues, and mentors for the substantive feedback they provided on earlier versions of this work: Morris Eaves, Kenneth Gross, Matthew Kirschenbaum, Julia Flanders, Alan Liu, and Katie King. Their judicious comments and ideas enriched my thinking about the project and formed the basis for my revisions. I am also indebted to the anonymous readers for DHQ whose detailed suggestions helped strengthen the essay.

Notes

[1] [Ramsay 2008], [Ramsay 2003]. Both essays are part of a larger book project that will be published by University of Illinois Press as part of its Topics in the Digital Humanities series.

[2] See [Ceruzzi 2000].

[3] As a response to the widely held belief that textual criticism and bibliography are technical rather than hermeneutic disciplines, Jerome Mcgann has forcefully argued that creating a critical edition is a fundamentally interpretive endeavor. See [Mcgann 1991].

[4] By "string," I mean a sequence of non-numeric symbols, such as letters. The term, which derives from computer science, is explained below.

[5] As discussed earlier, Ramsay identifies and analyzes these objections in [Ramsay 2003] and [Ramsay 2008], passim.

[6] While this is apparently a genuine conjecture of Housman's, I am relying on the dramatic exposition of it in [Stoppard 1997, 37-8].

[7] Pages 4-5 intermittently incorporate prose I originally drafted for an essay co-authored with Matthew Kirschenbaum in 2000 entitled "Outside the Archive," portions of which were delivered at the Bibliography and the Internet panel, MLA Convention, Marriott Hotel, Washington D. C., 29 Dec. 2000. Several of the themes of this essay intersect with those explored in the earlier paper, which I have cited where appropriate. For a published version of these remarks, see [Kraus 2002].

[8] Eustochia: "a happy conjecture"; [Nisbet 1991]

69

[9] [Manetti 1993, 117] To the extent that it implies specifically human inference, "cognitive" is a misleading word choice. The inferential leap may be either biological or artificial (with "artificial" understood as "computational," "mechanical" or "automated").

[10] Not all computers, however, are digital. The slide rule, for example, is a manually operated analog computer, whereas the abacus is digital. The electronic analog computer, originally developed for military, meteorological, and industrial applications, represents information in voltages that vary continuously rather than discretely. Mechanical analog machines known as differential analyzers were used to solve equations well into the 1960s before being superceded by the von Neumann stored program model. It was the speed, versatility, and (especially) accuracy of digital computers that enabled them to eventually overtake their analog predecessors.

[11] See [Spolsky 2002]. Kirschenbaum also discusses Spolsky's Law in [Kirschenbaum 2008, 242-3].

[12] Duthie is quoted at length in [McLeod 1983, 192-93n22].

[13] Not to mention fundamentally postmodern. With their profusion of puns, ambiguity, and double-coded meanings, Duthie's conjectures exhibit some of the defining traits of literary postmodernism (if we were to transplant them onto the pages of Joyce's *Finnegan's Wake* or Pynchon's *Gravity's Rainbow*, would the results be so very discordant?). There is, more generally, something in the play of conjectural signification that resembles poststructuralist theories of language: each word reverberates with the traces of other words, and it is the promise of these other words that drives the conjectural project.

[14] Note the aptonym: "love of sound."

[15] Qtd. in [Berg 2001].

[16] Bentley n.pag.

[17] See [Mahood 1957, 15-8].

[18] [Lanham 1991]; [Hall 1913].

[19] This process is grounded not in logical or empirical relationships, as A. L. Oppenheim makes clear, but in psycholinguistic principles, both phonological and semantic in character.

Only exceptionally are we able to detect any logical relationship between portent and prediction, although often we find paronomastic associations and secondary computations based on changes in directions of numbers. In many cases, subconscious association seems to have been at work, provoked by certain words whose specific connotations imparted to them a favourable or an unfavourable character, which in turn determined the general nature of the prediction. [Oppenheim 1964, 211]

[20] While the control syntax here is relatively simple, the branching structures of Mesopotamian prophecies were often quite complex. They have a metaphoric affinity with control structures in computer programming. For example, the following are actual Mesopotamian omens and oracles to which I've affixed programming commands: IF, on the Threshold of the Door of the Palace, on the right there is a slit — [THEN . . .]; ELSEIF, on the Threshold of the Door of the Palace, on the right, there is a length-wise slit — [THEN . . .]; ELSEIF, on the Threshold of the Door of the Palace, on the left, there is a slit — [THEN . . .]; ELSEIF, on the ITHEN . . .]; ELSEIF, on the Ithreshold of the Door of the Palace, on the left, there is a slit — [THEN . . .]; ELSEIF, on the Threshold of the Door of the Palace, on the left, there is a lengthwise slit — [THEN . . .]; ELSEIF, in the middle of the Door of the Palace, there is a slit — [THEN . . .]; ENDIF; Manetti analyzes the foregoing sequence of clauses as follows:

As may be seen, all the protases are constructed according to a structural principle of binary opposition between the [Threshold] and [the middle] of the Door of the Palace, between [right] and [left], between [slit] and [lengthwise slit]. It is thus the system itself, understood in an ante litteram structural sense, which has prime importance. The cases which have been observed in the past are no longer the only ones registered, but rather all possible, conceivable cases are laid out according to a system based on oppositions and abstract rules. [Manetti 1993, 11]

[21] There is a question of nomenclature here. I am aware that within tech circles, Java, Pascal, and C++ would be denoted high-level programming languages rather than machine languages. The latter term is reserved for low-level code legible only to the computer. By contrast, a high-level programming language is legible to humans and must be compiled into machine language before the computer can process it. Nonetheless, this narrow definition exists alongside a broader one, which copposes "machine" or "artificial" language to "natural" or "human" language, in this way distinguishing Pascal or C++ or Java from Italian or English or Chinese. It is this more generic sense of the term that I'm drawing on here. The fact that so-called machine or artificial languages in the broader sense are almost always written in English adds a further layer of cultural, technological, and linguistic complexity to these various discriminations.

[22] On the emblems of the baru, see [Manetti 1993, 5].

[23] See, for example, [Bottero 2001], as well as the recent conference (March 6-7 2009) held at the Oriental Institute of the University of Chicago, entitled Science and Superstition: the Interpretation of Signs in the Ancient World http://oi.uchicago.edu/research/symposia/2009.html.

[24] On LUCA, see [Poole 2002].

[25] While my goal in this section to draw parallels among biogenetics, textual criticism, and historical linguistics, I am also sensitive to potential conflicts in their aims or intentions that deserve more attention than I can afford to give them here. One of my aspirations in publishing this article is to raise awareness of and interest in conjecture in the digital humanities. But I am less sanguine about the role it might play in biogenetics, where the potential for abuse of the conjectural ethos is so great. Take, for example, the sub-discipline of synthetic biology, or "making creatures from scratch," which is discussed in [Woolfson 2005]. On the reconstructive possibilities of this emerging discipline, Woolfson muses that "it might be possible to re-create the elusive ancestor of all human life on Earth, a hypothetical organism known as LUCA, or the 'least universal common ancestor' " since "the remnants of LUCA should be scattered across the genomes of all living things." He goes so far as to entertain the possibility of bringing LUCA "back to life." It is for this reason that the material difference between texts and bodies needs to be continually acknowledged. Otherwise we risk glib pronouncements that fail to take into account what is at stake in asserting only their likeness. For discussion of the origins and limits of the analogy between genetic and textual (or informational) code, see [Thacker 2004] and [Kay 2000].

[26] Since first drafting this essay, I have been made aware that Sebastiano Timpanaro's seminal monograph on the German classicist Karl Lachmann has been translated from its original Italian into English by Glenn W. Most under the title *The Genesis of Lachmann's Method*. The book includes an informative chapter on the crosscurrents between textual criticism and linguistics in the nineteenth century, which builds on the earlier work of Henry Hoenigswald, discussed below.

[27] See the landmark volume of essays co-edited by Hoenigswald and Linda F. Wiener, [Hoenigswald and Wiener 1987].

[28] See in particular [Hoenigswald 1963, 1-11].

[29] One could formulate a parallel definition for phylogenetics: comparisons of characters between two or more related taxa are undertaken with the purpose of reconstructing their genealogy and archetype. Until recently, however, many scholars would have inserted a full stop after "genealogy." Less than two decades ago, for example, H. Don Cameron insisted that "For the zoologist the reconstructed ancestor is a fiction. It is a convenient way of presenting the organized information about the relationships of the real animals in question. Nobody tries to reconstruct a living, breathing theocodont or the protodipteron." He contrasted these attitudes and beliefs with those in textual criticism, where the reconstructed text is "the real article, in a literal sense the text that Euripides wrote. Reconstruction is a serious business, and the only point of studying manuscripts at all" [Cameron, 239]. On the contrary, reconstruction has been a desideratum of evolutionary biology from the start: one encounters a conjectural ethos in Darwin's Descent of Man and in the writings of, for example, Thomas Huxley, the eminent Victorian biologist and staunch defender of Darwin's model of evolution. The ethos has always been there, the methodology has not; or rather, the methodology has been hobbled by the lack of a clear unit of comparison and conjecture. But whereas conjectural criticism in recent decades has experienced a downtick in popularity in textual studies, the reverse is true in evolutionary and synthetic biology, where allographic methods have taken hold. With DNA and protein sequences — digital, discrete, and abstract — now firmly established as the clear units of conjectural analysis in evolutionary biology, the reconstructive paradigm has begun to flourish.

[30] In phylogenetics, this difference is expressed as that between so-called homoplastic characters and homologous characters.

[31] Qtd. in [Salemans 1996, 44]. This volume testifies to the distinctly Scandinavian flavor of much new work in stemmatology.

[32] Cameron, in [Hoenigswald and Wiener 1987, 229].

[33] As Julia Flanders has pointed out to me, one could also compute the edit distance between phonological (rather than orthographical) representations of the words.

[34] Although I do not take up the question of the distinction between networks and trees in this article, a fuller exposition of the stemmatic method would need to discuss the issue of conflated texts, which incorporate readings from two or more manuscripts. While tree models cannot accommodate such texts, networks can.

[35] For an example of a genuine rather than hypothetical transformation series, see Cameron, in [Hoenigswald and Wiener 1987, 234].

[36] Cameron, in [Hoenigswald and Wiener 1987, 234].

[37] Unlike interior nodes in a tree, which have both ancestors and descendants, leaf nodes have only ancestors.

[38] Although both head and tail have the letter "a" in common, it occupies a different position in each word and thus doesn't count as a "match."

[39] [Searls 2001, 339–348].

[40] Qtd. in [Searls 2001, 340].

[41] [Searls 2001, 341-2]

[42] [Thomas 2007, 5] The workbook was published online on 19 May 2005 at http://www.paradigm.ac.uk/workbook/introduction/oais-information.html.

[43] [Borgman 2007, 214–5]. Borgman also notes that "generally speaking, the humanities are more interpretive than data-driven," with digital humanists being one obvious exception [Borgman 2007, 213].

[44] Borgman 225.

Works Cited

Abrams 1953 Abrams, M.H. The Mirror and the Lamp. Oxford: Oxford University Press, 1953.

Aitchison 2003 Aitchison, Jean. Words in the Mind: An Introduction to the Mental Lexicon. Oxford: Blackwell, 2003.

Bell 1999 Bell, Julian. What is Painting? New York: Thames and Hudson, 1999.

Bentley 1732 BentleyRichard, ed. Milton's Paradise Lost. London: Jacob Tonson, 1732.

- **Berg 2001** Warning: Biblio formatting not applied. BergAlex. *Review: Berg, Linguistic Structure and Change. Linguist List.* 17 August 2001 . http://linguistlist.org/issues/12/12-2061.html.
- **Borgman 2007** Borgman, Christine L. Scholarship in the Digital Age: Information, Infrastructure and the Internet. Cambridge: MIT Press, 2007.
- Bottero 2001 Bottero, Jean. *Religion in Ancient Mesopotamia*. Edited by Teresa Lavender Fagan. Chicago: Chicago University Press, 2001.
- **Cameron** CameronH. Don. "The Upside-Down Cladogram: Problems in Manuscript Affiliation". In Hoenigswald H. M. and Wiener L. F., eds., *Biological Metaphor and Cladistic Classification*. London: Francis Pinter, 1987. pp. 227-242.
- Carroll 1992 Carroll, Lewis. Lewis Carroll's Games and Puzzles. Edited by Edward Wakeling. New York: Dover, 1992.
- Ceruzzi 2000 Ceruzzi, Paul E. A History of Modern Computing. Cambridge, MA: MIT Press, 2000.
- Darwin 1871 Darwin, Charles. Descent of Man. Princeton: Princeton University Press, 1871.
- Fishbane 1989 Fishbane, Michael. Biblical Interpretation in Ancient Israel. Oxford: Oxford University Press, 1989.
- Franklin 2001 Franklin, James. *The Science of Conjecture: Evidence and Probability before Pascal*. Baltimore: Johns Hopkins University Press, 2001.
- Goodman 1976 Goodman, Nelson. Languages of Art. Indianapolis: Hackett Publishing Co., Inc, 1976.
- Greg 1924 GregW. W. "Massinger's Autograph Corrections in The Duke of Milan, 1623". The Library 4 (1924), pp. 207-218.
- Grier 2005 Grier, David Halan. When Computers Were Human. Princeton: Princeton University Press, 2005.
- Grivelet 1956 GriveletMichel. "Shakespeare as 'Corrupter of Words'"". In Allardyce Nicoll, ed., Shakespeare Survey 16. Cambridge: Cambridge University Press, 1956.
- Gross 2001 Gross, Kenneth. Shakespeare's Noise. Chicago: University of Chicago Press, 2001.
- Hall 1913 Hall, F.W. A Companion to Classical Texts. Oxford: Clarendon Press, 1913.
- Hoenigswald 1963 Hoenigswald, Henry. "On the History of the Comparative Method". Anthropological Linguistics 5 (1963).
- Hoenigswald and Wiener 1987 Hoenigswald, Henry, and Linda F. Wiener. *Biological Metaphor and Cladistic Classification: An Interdisciplinary Perspective*. Philadelphia: University of Pennsylvania Press, 1987.

- Housman 1921 Housman, A.E. "The Application of Thought to Textual Criticism". *Proceedings of the Classical Association* 18 (1921).
- Johnson 1765 JohnsonSamuel. "Preface". In The Plays of William Shakespeare. London: 1765.

Jonese 1948 Jones, Marc Edmund. Occult Philosophy. Victoria: Trafford Publishing, 1948.

Kac 1999 KacEduardo. Genesis: Artist's Statement. 1999. http://www.ekac.org/geninfo2.html.

- Kay 2000 Kay, Lily. Who Wrote the Book of Life: A History of the Genetic Code. Stanford: Stanford University Press, 2000.
- Kenney 1974 Kenney, E.J. The Classical Text. Berkeley: University of California Press, 1974.
- Kirschenbaum 2008 Kirschenbaum, Matthew. *Mechanisms: New Media and the Forensic Imagination*. Cambridge: MIT Press, 2008.
- Kraus 2002 Kraus, Kari. "Conjecture". Performance Research 7 (2002).
- Lanham 1991 LanhamRichard. A Handlist of Rhetorical Terms. Berkeley: University of California Press, 1991.
- Mahood 1957 Mahood, M.M. Shakespeare's Wordplay. London: Methuen, 1957.
- Manetti 1993 Manetti, Giovanni. Theories of the Sign in Classical Antiquity. Bloomington: Indiana University Press, 1993.
- McLeod 1983 McLeod, Randall. "Gon. No more, the text is foolish". In Gary Taylor and Michael Warren, eds., *The Division of the Kingdoms*. Oxford: Clarendon Press, 1983.
- Mcgann 1991 McGann, Jerome. The Textual Condition. Princeton: Princeton University Press, 1991.
- Nisbet 1991 Nisbet, R.G.M. "How Textual Conjectures Are Made". *Materiali e discussioni per l'analisi dei testi classici* 26 (1991).
- O'Hara 1996 O'HaraRobert. "Mapping the Space of Time: Temporal Representation in the Historical Sciences". New Perspectives on the History of Life: Systematic Biology as Historical Narrative. In M. T. Ghiselin and G. Pinna, eds., Memoirs from the California Academy of Sciences 20. San Francisco: California Academy of Sciences, 1996. http://rjohara.net/cv/1996-cas.
- O'Hara and Robinson 1993 O'Hara, R., and P. Robinson. "Computer-assisted Methods of Stemmatic Analysis". In N. Blake and P. Robinson, eds., *The Canterbury Tales Project Occasional Papers Volume I*. London: Office for Humanities Communication Publications, 1993. pp. 53-74.
- **Oppenheim 1964** Oppenheim, A.L. *Ancient Mesopotamia: Portrait of a Dead Civilization*. Chicago: University of Chicago Press, 1964.
- **Poole 2002** Poole, Anthony M. *What is the Last Universal Common Ancestor (LUCA)?* actionbioscience.org, September 2002. http://www.actionbioscience.org/newfrontiers/poolearticle.html#primer.
- Ramsay 2003 Ramsay, Stephen. "Toward an Algorithmic Criticism". *Literary and Linguistic Computing* 18: 2 (2003), pp. 167-174.
- Ramsay 2008 Ramsay, Stephen. "Algorithmic Criticism". In Susan Schreibman and Ray Siemens, eds., A Companion to Digital Literary Studies. Oxford: Blackwell, 2008. pp. 477-491. http://www.digitalhumanities.org/companionDLS/.
- Rye 2003 RyeJustin. Futurese. http://www.xibalba.demon.co.uk/jbr/futurese.html.
- Salemans 1996 Salemans, Ben. "Cladistics or the Resurrection of the Method of Lachmann: On Building the Stemma of Yvain". In Pieter van Reenen and Margot van Mulken, eds., *Studies in Stemmatology*. Amsterdam: John Benjamins Publishing Company, 1996. pp. 3-70.
- Searls 2001 Searls, David. "From Jabberwocky to Genome: Lewis Carroll and Computational Biology". Journal of Computational Biology 8 (2001).
- Sewell 1952 Sewell, Elizabeth. The Field of Nonsense. London: Chatto & Windus, 1952.
- **Spolsky 2002** Spolsky, Joel. *The Law of Leaky Abstractions*. Joel on Software, 2002. http://www.joelonsoftware.com/articles/leakyabstractions.html.
- Steiner 1998 Steiner, George. After Babel. Oxford: Oxford University Press, 1998.

Stoppard 1997 Stoppard, Tom. The Invention of Love. New York: Grove Press, 1997.

Tarrant Tarrant, R.J. "The Editing of Classical Latin Literature". Scholarly Editing 118.

Tedlock 2000 Tedlock, Dennis. "Towards a Poetics of Polyphony and Translatability". In Jerome Rothenberg and Steven Clay, eds., *A Book of the Book*. New York: Granary Books, 2000.

Thacker 2004 Thacker, Eugene. Biomedia. Minneapolis: University of Minnesota Press, 2004.

Thomas 2007 Thomas, Susan. Paradigm Workbook on Digital Private Papers. Oxford: Bodleian Library, 2007.

Welsford 1966 Welsford, Enid. The Fool: His Social and Literary History. Gloucester: Peter Smith, 1966.

Woolfson 2005 Woolfson, Adrian. The Intelligent Person's Guide to Genetics. New York: Overlook Press, 2005.



This work is licensed under a Creative Commons Attribution-NoDerivatives 4.0 International License.