

Swarthmore College

Works

English Literature Faculty Works

English Literature

4-1-2018

Search And Replace: Josephine Miles And The Origins Of Distant Reading

Rachel Sagner Buurma
Swarthmore College, rbuurma1@swarthmore.edu

L. Heffernan

Follow this and additional works at: <https://works.swarthmore.edu/fac-english-lit>



Part of the [English Language and Literature Commons](#)

Let us know how access to these works benefits you

Recommended Citation

Rachel Sagner Buurma and L. Heffernan. (2018). "Search And Replace: Josephine Miles And The Origins Of Distant Reading". *Modernism/modernity*. Volume 3, Issue 1.
<https://works.swarthmore.edu/fac-english-lit/377>

This work is brought to you for free by Swarthmore College Libraries' Works. It has been accepted for inclusion in English Literature Faculty Works by an authorized administrator of Works. For more information, please contact myworks@swarthmore.edu.

Search and Replace: Josephine Miles and the Origins of Distant Reading

M modernismmodernity.org/forums/posts/search-and-replace

Apr 11, 2018 By: [Rachel Sagner Buurma and Laura Heffernan](#)

Volume 3, Cycle 1

From this day forward, every time you see the name Roberto Busa invoked as a—or *the*—founding scholar of either quantitative or computational method in the humanities, we want you to mentally search and replace with another name: Josephine Miles.

Miles was a poet and an English professor at Berkeley. In the 1930s, as a graduate student at Berkeley, she completed her first distant reading project: an analysis of the adjectives favored by Romantic poets. In the 1940s, with the aid of a Guggenheim, she expanded this work into a large-scale study of the phrasal forms of the poetry of the 1640s, 1740s, and 1840s. In all of this distant reading work, Miles created her tabulations by hand, with pen and graph paper. She also directed possibly the first literary concordance to use machine methods. In the early 1950s, Miles became project director of an abandoned index-card-based Concordance to the Poetical Works of John Dryden. Partnering with the Electrical Engineering department at Berkeley, and contracting with their computer lab and its IBM tabulation machine, Miles used machine methods to complete the concordance. It was published in 1957, six years after she and several woman graduate students and woman punch-card operators began the work. It was thus begun around the time that Busa circulated early proof-of-concept drafts of his concordance to the complete works of St. Thomas Aquinas, and published 17 years before the first volumes of the 56-volume *Index Thomasticus* began to appear.

There are good reasons, of course, that scholars and journalists like to begin with Busa: he was the first concordance-maker to automate all five stages of the process, in 1951. Busa also sought out high-profile partnerships with Thomas Watson of IBM and others; he foregrounded the innovative nature of his work, and his index incorporated new programming approaches as they developed through the later 1950s and 60s. Miles, on the other hand, worked with the engineers and machines that were close at hand in nearby Cory Hall; she credited the woman typists and punchcard operators she collaborated with; she valued “imaginative” programmers.”[1] And Miles came to the discarded Dryden concordancing project from her own difficult experiences attempting to convince the University of California Press to publish her data along with her literary-critical and literary-historical interpretations of it. When she started work on the Dryden project, Miles was already deeply experienced with the problems of publication and workflow that swirl around the history of concordancing, both manual and machine.[2]

What we stand to gain from Miles is therefore twofold. Most importantly, we find in her work a literary genealogy for distant reading to stand alongside other long genealogies that track the rise of “quantitative or empirical approaches to literary history” through literary sociology, for example, as Ted Underwood does. Miles understood the latest in computational concordancing as influenced by the “inventive” concordancing tradition of the last generation, including “the great Lane Cooper Concordance for Wordsworth, Cornell 1911” (Miles, review, 290). Miles saw concordances and machine indexing as a core part of literary criticism, for they could help scholars to a broader view of comparisons between poems and poets. And Miles’s distant reading work was not only literary, it was in an important sense modernist: her work tested and overturned some of her generation’s defining accounts of modernist and metaphysical poetry as “hard” or “concrete.” Miles’s distant reading projects are therefore part of the history of twentieth-century poetry. And she did not limit her “tabular view” to literary history; her quantitative work also influenced her own poetic style and shaped the Verse Composition course she taught at Berkeley to poets like A. R. Ammons and Jack Spicer.

Replacing Busa with Miles has a second function: it can stand as an example of how we might write a history of literary scholarship that does not center originality and individual accomplishment. Though Busa had the help of many associates—Melissa Terras has tracked down and credited Busa’s female punch-card operators—their names were not credited in his published work. Miles, by contrast, quite carefully credited the women who worked on her projects. She gave the names of her two graduate student workers, Mary Jackman and Helen S. Agoa, on the cover of the Dryden index. She spoke in the preface and in later interviews about the importance of her collaboration with Penny Gee of Berkeley’s Cory Hall computing lab. And she emphasized that her *Renaissance, Eighteenth-Century, and Modern Language in English Poetry: A Tabular View* (1960) would never have been finally printed had not “an heroic woman” and “master typist” at the University of California Press offered to type the charts.[3]

Substituting “Miles” for “Busa,” then, allows us to imagine an origin story for computationally-assisted reading that comes from within the discipline of literary study, and involves a woman project lead who preserved the names of her female collaborators. Miles’s career also offers us another origin point for distant reading and quantitative methods of literary analysis. And it helps us see how Miles’s computational concordance project and her distant reading work informed one another, and together informed her own composition of poetry and teaching of poetic composition.

The Dryden Concordance

À								
AARON	WB	340	ABIDE	AE8	600	ABODES	AE1	91
	HAP	187		AE10	1147		AE1	492
AARON'S	BR	297		KT2	158		AE1	736
	AA	525		KT2	208		AE2	337
ABANDON	SAA	1024		KT3	618		AE2	393
	AE2	472		M8	108		AE2	473
	KT3	287		CAM	139		AE2	1073
ABANDONED	CI	527		C	301		AE3	151
	EAZ	37		C	311		AE3	616
	BR	126		C	342		AE4	674
	BR	266		FL	324		AE4	964
	AE1	841		M15	362		AE6	510
	AE2	90	ABIDES	G1	270		AE6	761
	AE2	762		G4	261		AE6	911
	AE2	769		KT1	369		AE6	1007
	AE2	915	ABJECT	VP7	58		AE7	165
	AE3	793		HIF	228		AE7	315
	AE4	281		HIF	347		AE7	329
	AE10	1216	ABLE	AU	24		AE7	696
	AE11	399		AM	187		AE8	54
	AE11	1072		EDGA	34		AE9	342
	KT3	764		TA	73		AE10	8
ABANDONING	B	277		HAP	1002	ABOLISH	AE11	1153
	AM	695		HAP	1004	ABOLISHED	AE12	264
	HAP	1612		J6	572		AF	144
	J10	153		J10	236		HIF	330
	MC	67		M9	119		C	604
	AE2	1015		M15	2		G3	636
	AE12	723		G3	120		M1	491

Fig. 1. From the first page of Guy Montgomery's *A Concordance to the Poetical Works of John Dryden*. Miles took Montgomery's 64 boxes of index cards and made a computerized index with the help of Helen S. Agoa and Mary Jackman.

The Dryden project initially entered Miles's life as something of a "departmental obligation and chore." [4] When Miles's colleague Guy Montgomery died in 1951, he left behind him 64 shoeboxes full of index cards towards a concordance to Dryden's complete works. George Potter, the English department chair at Berkeley at the time, wrote to Montgomery's former students, but none could take on the project. Potter asked Miles to do the job because her scholarship had focused for over a decade on concordances and tabulation. Miles remembers Potter saying, "You use concordances so much, and so much counting, you ought to be able to handle this sixty-three shoeboxes of cards for the Dryden concordance which Guy Montgomery left when he died." That project, Miles recalls, "led me into years of studying computers, and I did make a computer concordance" (*Poetry, Teaching*, 75). To transform the index-card concordance into a computer concordance, Miles applied for funding from the Faculty Research Committee, which she used to contract with the Electrical Engineering Department to use their IBM tabulation machine and their punch-card staff to create a new, corrected set of cards for tabulation. University President Robert Sproul granted funds for the photolithographic reproduction of the results for publication.

Writing to Montgomery's one-time graduate student collaborator Lester A. Hubbard in 1957, Miles described the "experimental" and collaborative nature of the project: "[t]he Concordance underwent so many metamorphoses . . . that at least seven names would have had to be listed as editors" (letter to Hubbard). In her preface to the published volume, Miles thanked faculty from the English, French, and Speech Departments, but she especially singled out the female staff members from the computer lab who "worked under the guidance of Mr. Gordon Morrison and Mr. Boyd Judd at the Computer Laboratory," including Shirley Rice, Odette Carothers, and Penny Gee, who had punched each card in the Concordance with a single word, a symbol corresponding to the title of the poem in which it appeared, and a line number. [5] Later, Miles remembered Gee as "very smart and good" and

—most importantly—a true collaborator, as opposed to those “IBM people from San Jose” who would arrive periodically to flatly ask, “What can we do to help you?” “I’ve never been able to connect with them,” Miles explains, “ though I did with Penny Gee. She really taught me” (*Poetry, Teaching*, 126).

For the concordance’s publication, Miles elected to give editorial credit to the two graduate students, Mary Jackman and Helen S. Agoa, who had worked intensely on the concordance in its final stages, declining to place her own name on the cover; her name appears instead as the author of the *Concordance’s* Preface. Later, Hubbard would engage a lawyer to demand that he be acknowledged as co-author on the basis of the work he had done to generate the initial set of index cards with Guy Montgomery. Thin slips with Hubbard’s name on them were distributed to be pasted into the already-published volumes.

Miles’s Distant Reading

By the time the Dryden concordance was published in 1957, Miles had been using her own handmade word counts to analyze poetic language for almost twenty years. As a graduate student in the late 1930s at Berkeley, Miles’s dissertation examined Wordsworth’s language — specifically, whether he expressed emotions literally or through metaphor. “I developed a method for doing this,” Miles later explained, “which involved counting, because I wanted to show actual proportions, that he did very little else but just state literally” (*Poetry, Teaching*, 65). *Wordsworth and the Vocabulary of Emotion* appeared in print in 1942, and was followed by *Pathetic Fallacy in the Nineteenth Century: A Study of a Changing Relation Between Object and Emotion* (1942), and *Major Adjectives in English Poetry From Wyatt to Auden* (1946).



Fig 2. Miles teaching teachers of literature and composition through the Bay Area Writing Project. Image courtesy of Carolyn H. Smith.

Miles employed graduate students at the rate of \$25 for 20 hours of work to help her to tabulate the adjectives, nouns, and verbs of past poetry and their place within the syntax of the poem; at times students like “Miss Jean Warren” suggested modifications to “methods of analysis.”[6] In the following decade, *The Continuity of Poetic Language* (1951), *Eras and Modes in English Poetry* (1957), and *Renaissance, Eighteenth-Century, and Modern Language in English Poetry: A Tabular View* (1960) appeared.

The star of large-scale literary data-gathering rises and falls in different eras, and Miles worked at a time in which the making of data sets was regarded as feminized and mechanical.[7] Accordingly, she experienced difficulties convincing publishers to reproduce her data, and she had trouble finding a reception among literary critics who viewed her datasets as merely preparatory to the true work of evaluation. As Miles described it, her projects fell between the “two stools” of linguistic research and literary study. She didn’t identify with the linguists, for linguistic analysis was “too alien to the text.” Meanwhile, literary critics wished to absorb the “interesting” claims of her books while discarding “all the tables and charts that she throws in there” (*Poetry, Teaching*, 133).

The Tabular View and Mid-century Criticism

Miles's intent was to bring criticism and data together; she viewed her tabulations as correctives to the critics who "turn discussions of what poetic language is to prescriptions for what it ought to be," adding that "not only theorists like Richards but also historians like Bateson do more evaluating than tabulating" (*Major Adjectives*, 305). Christopher Rovee describes her as a "rogue formalist of the early new-critical period," and indeed, she was very much in conversation with the poet-critics and the formalist theorists of mid-century.[8] Miles's tabulations of the grammatical tendencies of past poets revealed not only the forms of past poetry, but showed the degree to which poetic values of the mid-twentieth-century overdetermined critics' sense of the past, turning poets like Blake or Donne or Wordsworth or Yeats into "mirrors" of present poetry.[9]

We may think of Blake, for example, as "active, rebellious, and eccentric," but Miles's work showed how he "as wholeheartedly as any accepted the material of his kind and time," using language in which "nouns and adjectives . . . strongly dominated verbs," sentences that were "phrasally compounded," and rhymes that "stressed not limits, periods bound and conclusions . . . but rather interior units and correspondances, in echo and onomatopoeia" (*Eras*, 87).

In the case of Donne, Miles described how "modern critics" like Cleanth Brooks and I.A. Richards, who "emphasize the poem as aesthetic object," necessarily ignore or distort the "language of conceptual reference and argument" which the tabular view reveals to be dominant in Donne's poetry (31). Likewise, Miles shows mid-century's Wordsworth to be almost a complete fabrication, noting that modern critics "ignore the mass of Wordsworth's work" and cherry pick unrepresentative poems like "A Slumber Did My Spirit Seal" with its "rocks, and stones, and trees" that represent "our own great concreteness" rather than Wordsworth's more typical style of subtle generalization (128). Yeats, too, is praised in the terms of the present, yet Miles's work revealed that the language of his later volumes—*Michael Robartes and the Dancer* (1921) *The Tower* (1928), and *The Winding Stair* (1933)—became *less* predicative, bending back instead to the balanced and classical mode of his earliest work.

Modern Poetry and the "Middle Distance"

Eras in English Poetry

	1500-40	1540-70	1570-00	1600-40	1640-70	1670-00	1700-40	1740-70	1770-00	1800-40	1840-70	1870-00	1900-40
PHRASAL				Fletcher	Milton		Thomson Somerville Dyer Philips Prior	T. Warton Armstrong J. Warton Collins Lyttleton Mason Gray Akenside	Bowles Blake	Hemans Keats	Tennyson Whitman		Moore Pound
				Sylvester	More	Wesley Blackmore	Blair					Wilde	
					Waller			Shenstone	Southey Goldsmith Crabbe	Shelley	Poe Arnold	Swinburne Bridges Hopkins Phillips Thompson Dickinson	E. Sitwell Yeats Stevens Eliot
BALANCED	Hawes Dunbar Douglas	Sackville	Spenser Shakespeare Hall			Dryden Marvell Addison Pomfret	Pope Parnell Young	Johnson	Wordsworth Rogers Burns Chatterton	Bryant Campbell Macaulay Scott	Longfellow Emerson		
				Sandys	Crashaw								
CLAUSAL	Barclay Lydgate Chaucer Sternhold Skelton Ballad Langland	Lindsay Gascoigne Baldwin Heywood Wyatt Surrey Breton Googe Turberville	Sidney Drayton Chapman Marlowe Daniel Donne Jonson	Quarles Carew Herrick Shirley Harvey Wither Herbert	Lovelace Denham Vaughan Cleveland Sackling Cowley		Roscommon Gay	Creech Walsh Oldham	Coleridge	Byron Moore	Meredith E. Browning R. Browning	Hardy Housman	Cummings Frost Auden Millay

Josephine Miles

Fig 3. From Josephine Miles, "Eras in English Poetry," *PMLA* 70 no. 4 (1955): 861.

Miles's tabular views not only rescued past poets held hostage by present-day poetic values, they revealed entirely new genealogies linking past to present. Nearly all of Miles's scholarly essays contain illuminating and original asides about modern poets. She notes, for example, how the "Donne tradition" lives on in "the Cavalier lyricism of Cummings and Millay" as much as "the metaphysical meditation of Frost and Auden" (*Eras*, 27). She describes how the dominance of the "family relations of father, mother, son" to be found in the old ballads reappears in "such poets as Auden or Lowell," while Ezra Pound and Robert Penn Warren and Federico Garcia Lorca develop upon "Coleridge's ballads of night and strangeness" (107). She sees T. S. Eliot as attempting to "strike a balance" between Miltonic phrasal poetry (qualitative, coordinate) and Donnic predicative poetry (clausal, conceptual, full of logical subordination) (24).

Against her era's critical truisms—its emphasis on the image, its separation of poetry from prose, its figuration of the poem as object—Miles carved out an alternative view of modern poetry's challenges and strengths. In her view—one that looked at poems as sentences, and traced the pendulum swings of each century from verbs to adjectives and back again—poets of the 1950s faced the same challenges as "Pope or Thomson" did in the 18th century: they had at their disposal a "stifling amount of device to deal with a stifling amount of objects and sensations." Modern poetry, contended Miles, needs "a Wordsworth of its own, to be the generalizer and steadfast interpreter of its own terms" (*Eras*, 125).

Miles's own poetry reflected her aspirations to pull back enough from poetic matter to become a "generalizer and steadfast interpreter of . . . terms" but not so much as to become "alien to the text." Reviewing several volumes of Miles's poetry in a 1959 article titled "Distance and Surfaces: The Poetry of Josephine Miles," Robert Beloof described Miles's achievement in terms of her point of view. She "sees her world from one of the most difficult

of artistic points of view . . . the middle distance,” Beloof writes. From this vantage, “she is never aware of it as a great metaphysical scheme on the one hand, nor, on the other, as a stream of detail which is itself the the ultimate reality.” The “middle-distance viewer is close enough to see the figures, and their principal attitudes, but not to feel an indissoluble unity with them. Nor is he far enough away that his impression is totally of the over-reaching structure. There is no ease. He is drawn in both directions, to the individuals of the community, and to the great abstractions which order them.”[10]

Miles, drawn in two directions, made a third—and deserves to have pride of place in our methodological origin stories.

Notes

[1] Josephine Miles, review of *A Concordance to the Poems of Matthew Arnold* by Stephen M. Parrish; *A Concordance to the Poems of W. B. Yeats* by Stephen M. Parrish; *Concordance to the Poems of Wallace Stevens* by Thomas F. Walsh, *Victorian Studies* 8 no. 3 (1965): 290-92, 292.

[2] Robert Oakman’s 1980 *Computer Methods for Literary Research* (Athens: University of Georgia Press, 1980) credits neither Miles nor Busa, but rather cites Stephen Parrish’s *A Concordance to the Poems of Matthew Arnold* as the first concordance to use machine methods (69).

[3] Josephine Miles, *Poetry, Teaching, and Scholarship: Oral History Transcript and Related Material, 1977–1980* (Berkeley: University of California Press, 1980), 123.

[4] Josephine Miles, letter to Lester A. Hubbard, October 12, 1957, Josephine Miles Collection, Bancroft Library, University of California, Box 8, Folder 13.

[5] Josephine Miles, “Preface” to Guy Montgomery and Lester A. Hubbard, eds. *Concordance to the Poetical Works of John Dryden* (Berkeley: University of California Press, 1957), ii-iii.

[6] Josephine Miles, *Major Adjectives in Romantic Poetry* (University of California Publications in English 12 no. 3 (1946), 305.

[7] See Janet Abbate, *Recoding Gender: Women's Changing Participation in Computing* (Cambridge, MA: MIT Press, 2012), 1-2.

[8] Christopher Rovee, “Counting Wordsworth by the Bay: The Distance of Josephine Miles,” *European Romantic Review*, 28 no. 3 (2017), 405-412, 406

[9] Josephine Miles, *Eras and Modes in English Poetry* (Berkeley: University of California Press, 1957, 1964), 128.

[10] Robert Beloof, "Distance and Surfaces: The Poetry of Josephine Miles" *Prairie Schooner* 32 no. 4 (1958-1959): 276-284, 276, 284.