



PROJECT MUSE®

Statistical Analysis at the Birth of Close Reading

Yohei Igarashi

New Literary History, Volume 46, Number 3, Summer 2015, pp. 485-504 (Article)

Published by Johns Hopkins University Press

DOI: 10.1353/nlh.2015.0023



➔ For additional information about this article

<https://muse.jhu.edu/article/601625>

Statistical Analysis at the Birth of Close Reading

Yohei Igarashi

The effect is like taking the cover off a machine.

—William Empson, “Basic English and
Wordsworth”

HOW THE DEFINING INTERPRETIVE technique of literary studies, close reading, might relate to non-close methods of textual analysis has become a vexing question in the present swirl of discourse occasioned by the digital humanities. It may be instructive to remember, then, that close and non-close reading, far from crossing paths for the first time recently, encountered one another in the early twentieth century on the terrain of educational “word lists.” The genre of the word list, a list of words usually taking word frequency as its elemental measure, shaped close reading at its founding. In fact, the founder of close reading, I. A. Richards, believed that one such word list, Basic English—a parsimonious but usable version of the English language reduced to only 850 words—was the best tool for teaching students to be better close readers. Granted, word frequency counts were not alone or even predominantly responsible for the development of close reading. Yet the significant role that such statistical analyses played at the birth of close reading is especially relevant and underappreciated today.

The inception of close reading confronted the same question that Alan Liu has identified as the “meaning problem” at the heart of the digital humanities: what is the relation between “quantitative interpretation and humanly meaningful qualitative interpretation”?¹ Engaging the same problem as Liu, but taking the opposite route, this essay inspects not the status of meaning in today’s “distant” or “quantitative” analyses, but rather how “distant” or “quantitative” analyses functioned in Richards’s meticulous theorization of “the *original* difficulty of all reading, the problem of *making out the meaning*.”² In other words, our concerns today about human interpretation and statistical data recapitulate those that attended early theorizations of close reading itself. Neither a defense nor a critique of close reading or non-close text analyses, this essay brings up for consideration a historical episode, hoping it will enrich our understanding of different modes of analysis in our discipline, and

concludes by pointing to the field of reading research as a direction toward which explorations of close reading, statistical analyses, and their relation might in the future go.

The earlier intersection of close reading and statistical analysis will hardly be news to anyone familiar with the contours of close reading's pre-New Critical theorization, C. K. Ogden and Richards's co-invention and promotion of the weird (and weirdly enthralling) Basic English project, or Basic English's direct influence on our machine translation systems.³ Yet the relation between close reading's advent and statistical analysis has thus far remained hidden in plain sight in the discourse associated with the digital humanities. I would surmise a few possible reasons. First, some important accounts that announce the arrival of digital text analysis methods decontextualize close reading. In Franco Moretti's case for "distant reading," for example, close reading is the "very solemn treatment of very few texts taken very seriously"; in N. Katherine Hayles's overview of the digital humanities, close reading means "learning to read complex texts" and represents the "Traditional Humanities."⁴ Close reading may indeed involve these things; it is worth acknowledging too that such accounts never claim to offer detailed histories of close reading. But the generality of these descriptions and their rhetoric of progress have unfortunately precluded the recognition that the origins of close reading in Richards's work in the late 1920s, as well as his inextricable advocacy of Basic English, were steeped in the non-close reading methods of that era, even in the latest work in statistics.⁵ Moreover, Richards was able to view close reading and word frequency lists as complementary because both could be oriented toward the social aim of improved widespread literacy. The discussion below sets out by adopting a delimited, historically specific understanding of close reading, returning to what John Guillory has called its "inaugural moment," in order to retrieve how close reading was actually formulated as well as its interaction with theoretical projects continuous with it.⁶

Second, digital humanities discourse has tended to focus on digital technology, with many accounts rehearsing the same story of the digital humanities' origin in the 1940s with Roberto Busa's computer-generated concordance for the works of Thomas Aquinas.⁷ As soon as one steps back from the fact of digitality, however, one sees a longer history involving the application of statistical methods or nineteenth-century "statistical thinking" to linguistic artifacts, including literary ones, long before the digital humanities or corpus linguistics.⁸ I share Lisa Gitelman's conviction that "there must be other . . . histories," namely those that take into account the "predigital history of the digital humanities."⁹ For this reason, the discussion below prefers "statistical analysis" over some other

terms currently in use. For example, “quantitative” circulates widely as a byword for digital textual analyses today and would seem to operate similarly to “statistical.” Yet I prefer “statistical” slightly since within literary studies, the term “quantitative”—like “numbers”—already denotes an entirely distinct set of objects and analyses, with vast historical precedent, concerning prosody. The problem with “quantitative,” then, begins with its ambiguity and extends to the worrisome fact that the term can unwittingly overwrite the long history of imaginative discourse’s relation to the quantitative when it is used to celebrate a purportedly novel methodological quality in today’s research.¹⁰

“Statistical analysis” can also be distinguished from Moretti’s “distant reading” and Matthew Jockers’s “macroanalysis.” Both terms seek to differentiate themselves from close reading through their commitment to mining a large quantity of texts: “The sheer quantity of available data makes the traditional practice of close reading untenable as an exhaustive or definitive method of evidence gathering.”¹¹ While Moretti and Jockers are reacting strongly against close reading’s traditional focus on single texts, “statistical analysis” accommodates the full range of objects that historically have been amassed and analyzed, including phonemes, letters, and words, not only discrete texts. Again, this is only a matter of preference, but “statistical analysis” allows us to recognize the continuities between predigital statistical readings of single texts—William Empson’s “Honest in *Othello*,” which is based on the frequency criterion—and digital counterparts like Stephen Ramsay’s illuminating analysis of Virginia Woolf’s *The Waves*.¹²

Indeed, unlike “machine reading” or “algorithmic criticism,” “statistical analysis” does not depend on digital or electronic computing technologies for its definition. The latter name, proposed by Ramsay, is probably the closest to “statistical analysis,” but Ramsay too traces the origins of his method, and the digital humanities in general, to Busa.¹³ The historical record, however, reveals many statistical analyses done through manual tabulation and computation, from concordances to image clusters to the genre of the word list, which play an important part in my account below.¹⁴ Dismissing predigital statistical analyses by understanding or celebrating digitality as a condition for non-close reading means subscribing to the severest forms of presentism and neglecting the most canonical cautions from media studies against conclusions of this kind.¹⁵

Finally, it will be necessary, moving forward, to recover more fully those statistical analyses performed in the interdisciplinary field of reading research that are highly relevant to literary studies and to our ongoing discussion of interpretation and statistics. Many statistical tools

and findings oriented to elementary or secondary education had, and continue to have, significant implications for the discipline of English at the university level. Edward Lee Thorndike's *The Teacher's Word Book* (1921), for example, was a landmark modern resource in the language arts and, as we will see, cast its shadow over Richards's early theories of close reading.

Thorndike's *The Teacher's Word Book* is a forgotten forebear to those impressive statistical analyses performed in literary studies today. A student of the psychologists William James and James McKeen Cattell, Thorndike was a key figure in the modern development of the discipline of educational psychology, particularly due to his efforts at scientizing and bringing statistics into psychology. Thorndike's most famous refrain encapsulates his conviction that psychology required the total adoption of statistical methods, a refrain repeated with variation across his oeuvre, realized by his textbook, *An Introduction to the Theory of Mental and Social Measurements* (1904), and enshrined in his obituary: "everything that exists, exists in quantity."¹⁶ *The Teacher's Word Book* lies somewhat to the side of the main current of his works in educational psychology, but it is altogether consistent with his belief that human learning could be quantified, reading ability being no exception: "How many English words should the ordinary boy or girl know the meanings of at the end of Grade 8? Which words should all or nearly all pupils know at that stage? In what grades and in what connections should they be learned? . . . It appears that one notable cause of our inability to answer them correctly is our lack of knowledge of the frequency of occurrence of words in the talk our pupil and graduate will or should hear, and the books, articles, letters, and the like, which he will or should read."¹⁷

The Teacher's Word Book produced a list of the 10,000 most frequently and widely occurring words in its total count of about four million running words. Thorndike used forty-one sources, these divided into five corpora: (1) "Children's Reading" (e.g., chapters from *Black Beauty* and *Little Women*, plus primers, readers, and textbooks from several subjects); (2) "Standard Literature" (literary concordances and excerpts from Bartlett's *Quotations*); (3) "Common Facts and Trades" (e.g., the US Constitution, cookbooks, almanacs); (4) "Newspaper Reading" (counts from newspapers); and (5) "Correspondence" (counts from letters). While Thorndike's count incorporated some sources that were already partly processed by scholars (e.g., concordances of the Bible, Milton, Wordsworth, and Tennyson) and some preexisting frequency lists (e.g., W. A. Cook and M. V. O'Shea's count of the private correspondence of thirteen American letter-writers, published in *The Child and His Spelling* [1914]), the list was nevertheless a massive endeavor that took Thorndike and his assistants a decade of counting and tabulation.¹⁸

Word lists existed before Thorndike. Most notably, there is the deep history of shorthand systems and their word lists, which extend back to antiquity, although the first English system appears in the Renaissance with Timothy Bright's *Characterie: An art of short, swift, and secret writing by character* (1588).¹⁹ Shorthand systems either intuited or counted the most commonly occurring sounds, letters, or words in a language, since those high-frequency letters, sounds, or words deserved the easiest-to-write symbols. Let us pause to note, then, that a real history of non-close or "distant reading" involves in large part the history of stenography, and implicates *handwriting* just as much as different modes of reading.²⁰ Inspired by F. W. Kaeding's milestone German word frequency count of 1898, also for the purpose of shorthand, English word lists for non-stenographic uses began to emerge in the early twentieth century. For example, Leonard P. Ayres's influential *A Measuring Scale for Ability in Spelling* (1915) and the aforementioned Cook-O'Shea correspondence count sought to ascertain the most commonly used words, since these were the words that students most needed to learn how to spell.²¹ Yet Thorndike's *The Teacher's Word Book* was the first word list to move beyond word frequency alone by computing a word's frequency while factoring in its likely dispersion throughout "the talk our pupil and graduate will or should hear, and the books, articles, letters . . . which he will or should read."²² To accomplish this aim, Thorndike devised a "credit" measure. Words in a specific work were assigned credits for their frequency: e.g., 1–4 occurrences were worth 1 credit, 5–9 occurrences were 2 credits, and so on for most of the sources in "Children's Reading." But the credit values differed across the corpora: up to 19 occurrences of a word earned only 2 credits for sources in the "Correspondence" corpus, which Thorndike weighted less than the "Children's Reading" that young readers were much more apt to encounter.²³ After adding up each word's credit from the 41 sources—Thorndike was keeping tabs on about 20,000 words—he offered in his list the 10,000 most "important" words for elementary school students to know. "In" was first on the list with a credit sum of 211. "Mean," "nothing," "order," "sea," and "seem" all had a credit of around 100 and sat within the top 500 words. "Gurgle" and "zest" were near the bottom of the list with only 3 credits each.

The Teacher's Word Book was chiefly motivated by the pedagogical idea that the words students encountered in elementary school ought to be more carefully controlled and keyed to their grade levels. *The Teacher's Word Book* would allow teachers to look up the statistical prominence or rarity, in children's reading in general, of certain words in Percy Bysshe Shelley's "Ode to a Skylark" (Fig. 1), for example, and then decide how to teach these words to their students: "some [words] should be explained

at the time to serve the purpose of the story or poem, but then left to their fate; some should be thoroughly taught and reviewed."²⁴

The list served, then, as the standard against which basal readers and other elementary textbooks systematically selected and graded the vocabulary they used. To be sure, there are serious irregularities in Thorndike's word list, most notably concerning derivatives. The preexisting counts he relied on handled derivatives differently, and Thorndike's list is itself only partly lemmatized; this explains the low ranks of Shelley's "sunken" and "bright'ning," even though "sink" is actually in the top 3,000 words and "sunk" in the top 5,000, while "bright" is in the top 2,000 and "brighten" in the top 5000. And then there was the intractable question of polysemy: as Thorndike's critics confirmed, *The Teacher's Word Book* could not distinguish between, for example, Shelley's "strains" and the word's other meanings. Nevertheless, by and large, *The Teacher's Word Book* answered many educational needs and proved to be tremendously influential, selling more than 20,000 copies.²⁵ A second edition followed in 1931, this time with a list of 20,000 words based on a count of ten million words; yet another expanded edition of 30,000 words came after that. And in response to the meaning problem, Thorndike and his student, Irving Lorge, published *A Semantic Count of English Words* (1938), which featured 570 selected words and broke down each word's frequency by its denotations as demarcated by the *Oxford English Dictionary*.²⁶

While Thorndike was undertaking his count, Ogden and Richards hit upon the idea of a word list of their own. In 1918, Ogden and Richards started discussing and drafting *The Meaning of Meaning* (1923). As they worked on their chapter on "Definition," they noticed that most objects in the world ("referents") are defined by their relation to other referents, and that only a few fundamental kinds of relation were necessary for most rudimentary definitions: for example, similarity (a "pear" is like an apple); spatial and temporal relations ("yesterday" is before today); and causation ("steam" is caused by heating water). But these reflections led to a curious—and more narrowly lexicographical—experiment. As Richards remembered it, while writing that chapter, he and Ogden "suddenly stared at one another and said, 'Do you know this means that with under a thousand words you can say everything.' If a word can be defined in a descriptive phrase of not more than ten words, you can substitute the descriptive ten words for the word and get rid of it."²⁷ Ultimately, this algorithm of "word substitution" would excise the English language of most of its words, leaving only a set of function words (which describe the fundamental kinds of relation among referents considered in *The Meaning of Meaning*) along with a working vocabulary of several

RATINGS OF THE WORDS FROM THE "ODE TO A SKYLARK" BY THE LIST. THE NUMBERS REFER TO SUCCESSIVE THOUSANDS. "NOT" MEANS THAT THE WORD IS NOT IN THE 10,000.							
blithe	7	harmonious	7	rainbow	3	unbidden	not
bright'ning	not	hidden	4	rapture	4	unbodied	not
chaunt	10	hymeneal	not	shrill	4	unpremeditated	not
chorus	7	hymns	4	soar	5	vaunt	7
divine	2	ignorance	4	spirit	1	wert	4
dost	3	madness	4	sprite	6	wherein	4
forth	1	matched	2	strains	3	wrought	3
fountain	2	melody	4	sunken	10		
gladness	5	panted	3	thine	3		
heeded	3	profuse	7	triumphal	7		

Fig. 1. From Thorndike's "Word Knowledge in the Elementary School." "Forth" and "spirit" are high frequency terms, among the first 1,000 words on the list, "divine" is in the second thousand, and so on.

hundred indispensable words. "Word substitution" was the germ of Basic English—infelicitously called at first "Panoptic English"—which Ogden went on to develop for the next ten years, drawing on his study of Jeremy Bentham's theories of language and early modern universal languages.²⁸ In 1929, he published the earliest iteration of the Basic word list. Ogden, a polymath, brought an excessive philosophical and theoretical apparatus to bear on Basic, but its main contours, for present purposes, can be described in further detail as follows.

Unlike constructed or "artificial" universal languages like Esperanto, Basic thus began with the existing English language, and, via the "word substitution" process, arrived at an 850-word list. The relative shortness of the list mattered: Basic claimed, like all of the shorthand scripts and universal languages that came before it, that its system could be mastered easily and quickly, actually in about thirty hours of study;²⁹ equally important, the Basic word list was short enough that it could be columned and printed on one side of a single sheet of paper and still be legible (Fig. 2). Basic achieves much of its reduction through the use of 100 "operator" words, which included only sixteen verbs, two auxiliary verbs, and twenty prepositions (among other essential words), but which could replace over 4,000 verbs (*SBE* 6). For instance, since Basic removes "enter" or "ascend" from its version of English, a Basic writer uses "go in" and "go up" instead. The remaining 750 words consisted of 600 carefully chosen "Things" (from "account" to "year") and 150 "Qualities" (e.g., "private," "sharp"). An excerpt from a Franklin D. Roosevelt speech, "It has been wonderful to me to catch the note of confidence from all

over the country," is translated into Basic's protracted style as, "It has been a great experience to me to be made conscious of the signs of an increasing belief in the future all over the country." If one wanted to write about "sole"—the fish, not of the foot—Ogden provided, "Small flat-fish with delicate taste."³⁰

Basic English had two aims: first, "to serve as an international auxiliary language . . . a second language for use throughout the world in general communication, commerce, and science"; and second, "to provide a rational introduction to normal English; both as a first step . . . for those whose natural language is not English, and as a grammatical introduction, encouraging clarity of thought and expression, for English-speaking people at any stage of proficiency" (*SBE* 4). From the start, then, Basic had two incommensurate and ultimately incompatible ambitions.³¹ The first aim, discernible in Basic as acronym ("British American Scientific International Commercial"), was to serve as a global linguistic medium facilitating the communication of information from many discourses—scientific, political, and more—toward the utopian idea of world peace so salient in the internationalism of the interwar period. In this regard, Basic was no different from its seventeenth- and eighteenth-century precursors, which were also driven, amid violent conflict, by the dream that a universal language would lead to universal love.³² The other aim, indicated by the word "basic," was to be a word list of relatively simple terms to assist in the learning of English by native as well as non-native speakers. In short, Basic aspired to be at once an "island vocabulary"—a miniature but comprehensive tongue—and a word list for educational purposes, just like Thorndike's.

Despite appearances, Basic was as statistically driven as *The Teacher's Word Book*. True, Ogden disparaged word lists: he insisted that Basic's "words are not the words most commonly used, as determined by word-counts" and snorted at "the army of word counters" inspired by *The Teacher's Word Book* (*SBE* 3, 30). Yet after Ogden was alerted to *The Teacher's Word Book* on inquiring about "word counts" in *Notes and Queries*, he studied Thorndike's findings³³; a great deal of his writing on Basic, as he developed the system, was devoted to positioning his list in relation to Thorndike's and others' statistical data.³⁴ After all, Thorndike had already suggested something very much like Basic's goal as an international language, claiming that the top 500 words of his count could, with slight modification, serve as "a *basic* list of great value in teaching foreign adults to read English."³⁵ And Basic's second goal—"to provide a rational introduction to normal English"—revealed Ogden to be attempting a systematization of English language learning similar to Thorndike's. Thus scholars of Basic like Rita Raley and Lydia H. Liu rightly identify Basic

BASIC ENGLISH WORD LIST									
*THINGS									
OPERATIONS	100	400 General	500	600	700	800	900	1000	1100
COME	ACCOUNT	EDUCATION	GENERAL	SENSE	200	300	400	500	600
GET	ACT	EFFECT	MIDDLE	SERVANT	KNIFE	KNIFE	KNIFE	KNIFE	KNIFE
GIVE	ADDITION	END	MILK	SIX	APPLE	KNOT	KNOT	KNOT	KNOT
GO	ADJUSTMENT	ERROR	MIND	SHADE	ARCH	KNOW	KNOW	KNOW	KNOW
KEEP	AGREEMENT	EXAMPLE	MINUTE	SHAKE	ARM	LEG	LEG	LEG	LEG
LET	AIR	EXCHANGE	MIST	SHAME	ARMY	LIBRARY	LIBRARY	LIBRARY	LIBRARY
MAKE	AMOUNT	EXISTENCE	MONEY	SIDE	BAG	LIP	LIP	LIP	LIP
FUT	AMUSEMENT	EXPANSION	MONTH	SILENCE	BALL	LOCK	LOCK	LOCK	LOCK
SEEM	ANIMAL	EXPERIENCE	MORNING	SILK	BAND	MAP	MAP	MAP	MAP
TAKE	BE	ANSWER	MOTHER	SILVER	BASIN	MATCH	MATCH	MATCH	MATCH
DO	APPARATUS	FACT	MOTION	SISTER	BASKET	MONEY	MONEY	MONEY	MONEY
HAVE	APPROVAL	FALL	MOUNTAIN	SIZE	BATH	MOON	MOON	MOON	MOON
SAY	ARGUMENT	FAMILY	MOVE	SKY	BED	MOUTH	MOUTH	MOUTH	MOUTH
SEE	ART	FATHER	MUSIC	SLEEP	BELL	MUSCLE	MUSCLE	MUSCLE	MUSCLE
SEND	ATTACK	FEAR	NAME	SLEP	BELL	NAIL	NAIL	NAIL	NAIL
MAY	ATTEMPT	FEELING	NATION	SLOPE	BERRY	NECK	NECK	NECK	NECK
WILL	ATTENTION	FICTION	NEED	SMASH	BIRD	NEEDLE	NEEDLE	NEEDLE	NEEDLE
ABOUT	ATTENTION	FIELD	NEWS	SMELL	BLADE	NERVE	NERVE	NERVE	NERVE
ACROSS	AUTHORITY	FIGHT	NIGHT	SMILE	BOARD	NET	NET	NET	NET
AFTER	BACK	FIRE	NOISE	SNOKE	BOAT	NOSE	NOSE	NOSE	NOSE
AGAINST	BALANCE	FLAME	NOTE	SNEEZE	BOOK	OFFICE	OFFICE	OFFICE	OFFICE
AMONG	BASE	FLIGHT	NUMBER	SNOW	BOAT	NOSE	NOSE	NOSE	NOSE
AT	BEHAVIOR	FLOWER	OBSERVATION	SOAP	BOAT	ORANGE	ORANGE	ORANGE	ORANGE
BEFORE	BELIEF	FOLD	OFFER	SOCIETY	BOTTLE	OVEN	OVEN	OVEN	OVEN
BETWEEN	BIRTH	FOOD	GIL	SONG	BOY	PARCEL	PARCEL	PARCEL	PARCEL
BY	BIT	FORCE	OPERATION	SPICE	BOY	PEN	PEN	PEN	PEN
DOWN	BITE	FORM	OPINION	SORT	BRAIN	PENCIL	PENCIL	PENCIL	PENCIL
FROM	BLOOD	FRIEND	ORDER	SOUND	BRICK	PICTURE	PICTURE	PICTURE	PICTURE
IN	BLOW	FRONT	ORGANIZATION	SOUP	BRIDGE	PIPE	PIPE	PIPE	PIPE
OFF	BOY	FRUIT	OWNER	STAGE	BRUSH	PLANE	PLANE	PLANE	PLANE
ON	BRASS	GOLD	PAGE	STUCK	BURN	PLATE	PLATE	PLATE	PLATE
OVER	BREAD	PAIN	PAINT	THICK	BURN	PLATE	PLATE	PLATE	PLATE
THROUGH	BREATH	PAIN	PAINT	THICK	BURN	PLATE	PLATE	PLATE	PLATE
UNDER	BROTHER	PARIS	PAINT	STEAM	BURN	PLATE	PLATE	PLATE	PLATE
UP	BUILDING	PARIS	PAINT	STEAM	BURN	PLATE	PLATE	PLATE	PLATE
WITH	BURN	GRIP	PART	STEP	CAKE	POT	POT	POT	POT
AS	BURST	GROWTH	PASTE	STITCH	CAMERA	POTATO	POTATO	POTATO	POTATO
FOR	BUTTER	GUIDE	PEACE	STOP	CART	PUMP	PUMP	PUMP	PUMP
OF	CANVAS	CAUTION	PEACE	STOP	CART	PUMP	PUMP	PUMP	PUMP
TILL	CARE	HARMONY	PLACE	STRETCH	CAT	RAT	RAT	RAT	RAT
THAN	CAUSE	HATE	PLANT	STRUCTURE	CHAIN	RECEIPT	RECEIPT	RECEIPT	RECEIPT
A	CHALK	HEARING	PLEASURE	STAIR	CHURCH	ROOF	ROOF	ROOF	ROOF
THE	CHANGE	HEAT	POINT	SUGGESTION	CHEST	ROD	ROD	ROD	ROD
ALL	CHANGE	HEAT	POINT	SUGGESTION	CHEST	ROD	ROD	ROD	ROD
ANY	CLOTH	HISTORY	POISON	SUMMER	CHURCH	ROOF	ROOF	ROOF	ROOF
EVERY	COAL	HOLE	POLISH	SUPPORT	CIRCLE	SAIL	SAIL	SAIL	SAIL
NO	COLOR	HOPE	PORTER	SURPRISE	CLACK	SCHOOL	MALE	WHITE	
OTHER	COMFORT	HOUR	POSITION	SWIM	CLOUD	SCISSORS	MARRIED	WRONG	
SOME	COMMITTEE	HUMOR	POWDER	SYSTEM	COAT	SCREW	MATERIAL		
SUCH	COMPANY	ICE	PRICE	TASTE	COLLAR	SEED	MEDICAL		
THAT	COMPARISON	IDEA	PRINT	TAX	COMB	SHEEP	MILITARY		
THIS	COMPETITION	IMPULSE	PROCESS	TEACHING	COW	SHIP	NECESSARY		
YOU	CONNECTION	INDUSTRY	PRODUCE	TEST	CUP	SHIRT	NEW		
WHO	CONTROL	INSECT	PROPERTY	THEORY	CUSHION	SKIN	OPEN		
AND	COFFER	INSTRUMENT	PROSE	THING	DOG	SKIRT	NORMAL		
BECAUSE	COPY	INSURANCE	THOUGHT	THOUGHT	DOOR	SNAKE	PAST		
BUT	CORK	INTEREST	FULL	THUNDER	DRAIN	SOCK	POLITICAL		
OR	CORN	INVENTION	PURPOSE	TOP	DRAWER	SPADE	POOR		
IF	COUGH	JELLY	COVER	TOUCH	DRESS	SPONGE	POSSIBLE		
THOUGH	COUNTRY	JOURNEY	QUESTION	TRADE	EAR	SPRING	PRESENT		
WHILE	CRACK	RAIN	TRANSPORT	TRADE	EGG	SQUARE	PRIVATE		
HOW	CREDIT	RANGE	TRICK	TRICK	ENGINE	STAMP	PROBABLE		
WHEN	CRIME	RAY	TURN	TRICK	EYE	STAR	QUICK		
WHERE	CRUSH	KNOWLEDGE	REACTION	TURN	FACE	STATION	READY		
WHY	CRY	LANGUAGE	REASON	TURN	FARM	STEM	RED		
AGAIN	CURRENT	REAR	REWARD	TURN	FEATHER	STICK	REGULAR		
EVER	DAMAGE	REASON	REASON	TURN	FINGER	STOCKING	RESPONSIBLE		
FORWARD	DANGER	REAR	REAR	TURN	FISH	STOMACH	RIGHT		
HERE	DAUGHTER	LEAD	RELATION	TURN	FLAG	STREET	ROUND		
NEAR	DAY	LEATH	REPRESENTATIVE	TURN	FLOOR	SUN	SAME		
NOW	DEATH	LETTER	REQUEST	TURN	FOOT	TABLE	SECOND		
OUT	DEBT	LEVEL	RESPECT	TURN	FORK	TAIL	SEPARATE		
STILL	DESIGN	LIFT	REST	TURN	THEATRE	WALK	SERIOUS		
THEN	DESIRE	LIMIT	RHYTHM	TURN	FRAME	THUMB	SMOOTH		
THERE	DETAIL	LIXEN	RICE	TURN	GARDEN	GIRL	TICKET		
TOGETHER	DESTRUCTION	LIQUID	RIVER	TURN	GLOVE	TOE	STIFF		
WELL	DEVELOPMENT	LOSS	ROOM	TURN	GOAT	TONGUE	STRAIGHT		
ALMOST	DIRECTION	LOVE	RULE	TURN	GUN	TOOTH	STRONG		
ENOUGH	DIGESTION	MAN	SALT	TURN	HAIR	TRAIN	SWEET		
EVEN	DISTANCE	MARK	SAND	TURN	HEAT	TRAY	TALL		
LITTLE	DISCOVERY	MEASURE	SEAT	TURN	HEAD	TRICK	TIGHT		
MUCH	DISCUSSION	METAL	SECRETARY	TURN	HEART	UMBRELLA	TIED		
NOT	DIVISION	MEAT	SELF	TURN	HOCK	WALL	TRUE		
QUITE	DOUBT	MEASURE	SEAT	TURN	HORN	WATCH	VIOLENT		
SO	DRINK	MEAT	SECRETARY	TURN	HORSE	WAITING	WARM		
VERY	DISTRIBUTION	MEAT	SELF	TURN	HOSPITAL	WHIP	WARM		
TOMORROW	DIVISION	MEAT	SECRETARY	TURN	HOUSE	WHISTLE	WET		
YESTERDAY	DIVISION	MEAT	SECRETARY	TURN	ISLAND	WINDOW	WIDE		
NORTH	DIVISION	MEAT	SECRETARY	TURN	JEWEL	WING	WISE		
SOUTH	DIVISION	MEAT	SECRETARY	TURN	JEWEL	WING	WISE		
EAST	DIVISION	MEAT	SECRETARY	TURN	JEWEL	WING	WISE		
WEST	DIVISION	MEAT	SECRETARY	TURN	JEWEL	WING	WISE		
PLEASE	DIVISION	MEAT	SECRETARY	TURN	JEWEL	WING	WISE		
YES	DIVISION	MEAT	SECRETARY	TURN	JEWEL	WING	WISE		

Fig. 2. The Basic English Word List

as (in Liu's words) "a conception of English as a statistical system."³⁶ By Liu's argument, Ogden understood that the 850 words recurring in Basic writing were the opposite of the enormous Joycean lexicon of *Finnegans Wake*, where each word reappeared less frequently (despite the novel's lengthiness, which is otherwise a strong determinant of low lexical va-

riety). And Ogden's intuition anticipated Claude Shannon and Warren Weaver's stochastic analysis of printed English in *The Mathematical Theory of Communication* (1948), which invokes none other than Basic English and Joyce's novel as its polar examples of information redundancy and entropy.³⁷ It is safe to say that Liu's argument, as much as mine, necessitates reading Ogden's overt dismissals of statistical word lists, and his claims for the superiority of Basic's non-quantitatively, "analytically" selected 850 words over frequency-based vocabularies, for what they really are: agonistic remarks made about rivals on the crowded ground, dominated by Thorndike, of statistical word lists.

While Ogden maintained that "the real statistical task of linguistic[s] is not so much the determination of the number of words actually used by any particular class of persons as . . . how a given field of reference may be covered with the greatest economy," his original collaborator Richards could, with more disinterest, see the affinities between *The Teacher's Word Book* and Basic English.³⁸ According to Richards, the main innovation of Basic English was the high-redundancy 100 "operator" words, Basic's syntactical workhorses used over and over again (e.g., "put," "give") to replace a huge variety of English words—and all of these words were also high in Thorndike's count, actually among the top thousand. "Whether we are content, as with the early word counts of E. L. Thorndike . . . to determine which words occur most frequently in various types of writing and base our list on frequency and range (the number of different sources using the word), or whether, as with Ogden's work, we take vastly more complex considerations into account," observed Richards, "there is no substantial difference in the result as regards these one hundred operation or structure words."³⁹

Evidently, Richards was immersed and fluent in such word lists and the data they presented. What is more, in 1928, as Ogden was finalizing the first versions of the Basic vocabulary, and Richards was approaching completion of *Practical Criticism* (1929), the two writers met, and Richards left their meeting convinced that the word list he had helped think up nearly ten years prior had a significant bearing on the founding document of close reading he was writing: "Much stuff on the borderline of his Simplified English and my *Practical Criticism* to discuss. They have a good deal to do with one another. His stuff being about 'How to say everything with the use of a minimum of words and constructions?'

E.g., get in} in place of {embark etc.
get off} {disembark" ⁴⁰

Richards was looking for a tool exactly like Basic as he sought a better way to implement paraphrase in teaching students to become better readers. As is well known, in his “practical criticism” experiments Richards asked students (and other willing participants, including T. S. Eliot) to perform evaluative “readings” of poems he distributed to them after removing the author’s name, the poem’s title, and obvious historical markers.⁴¹ The single most consistent problem for his students was the inability first to grasp the underlying sense or “thought” of a poem: “The most disturbing and impressive fact brought out by this experiment is that a large proportion of average-to-good (and in some cases, certainly, devoted) readers of poetry frequently and repeatedly *fail to understand it*. . . . They fail to make out its prose sense, its plain, overt meaning, as a set of ordinary, intelligible, English sentences, taken quite apart from any further poetic significance. They would travesty it in a paraphrase. They fail to construe it just as a schoolboy fails to construe a piece of Caesar” (*PC* 12). The source of the problem, according to Richards, is the “compression of poetic language”: this compactness allows poetry to communicate feelings quickly and effectively, but it also “tends to obstruct the discursive intelligence that works by spreading ideas out and separating their parts” (*PC* 204). Richards found that his students produced commentary and judgments determined predominantly by the feelings elicited by the poem or its form, but based on a faulty understanding of the poem’s sense, or worse, they entirely neglected the task of construing the meaning of the poem—a tendency responsible for several of the types of misreading taxonomized in *Practical Criticism*. Hence, although Richards does not in *Practical Criticism* outline a complete, formalized method for close reading, he implies a recommended reading strategy where understanding the sense of the poem is first in importance and in sequence. Insofar as the reader’s inability to construe a poem’s sense derails the entire interpretive process, Richards thought that “spreading ideas out and separating their parts” ought to be the first deliberate and disciplining activity after feelings—not to mention stock responses and irrelevant personal associations—have affected the reader’s response. “All respectable poetry invites close reading,” writes Richards, a process that always first requires “attention to its literal sense” (*PC* 195).

In a perfect world, paraphrase would trigger such attention, since a reader decompressing the poem via prose translation is compelled to reckon with “its plain, overt meaning” and unfold any ambiguities, all the while keeping at bay the distortions introduced by human interpretation, that is, by “the personal situation of the reader” (*PC* 227). Yet paraphrase, as teachers typically used it, was counterproductive: Richards observed that students would either resort to mindlessly “shuffling synonyms,”

thereby “leaving all the doubtful parts of the meaning unilluminated,” or they would “write a poem—in prose—on a partially similar theme.”⁴² And since students increasingly were not proficient enough in Latin or French to use either language as a medium into which to paraphrase English literary writing, the solution had to be a form of intralingual translation. In brief, Richards was convinced that paraphrase was the initial and most decisive component exercise subsumed within the technique of close reading, but teachers everywhere needed a better “reasoned general technique for construing” (*PC* 294).

This is where Basic English came in. As Richards finished *Practical Criticism* and subsequently wrote several books on Basic English—including *Practical Criticism*’s little-read sequel, *Basic in Teaching: East and West* (1935)—he came to believe, from his experiences teaching at Cambridge, Harvard, Bryn Mawr, and National Tsing Hua University in China, that the first step toward the skill he identified as “close reading” in *Practical Criticism* could only be taught to British, American, and Chinese students through Basic English.⁴³ A word list, like Basic, could prevent students from working around paraphrase by “shuffling synonyms”: the “very limitations of its vocabulary,” writes Richards, “forces you so constantly to explore into the meaning of the original. You cannot transpose it into Basic without close inquiry.”⁴⁴ With Basic, a student could not merely swap in synonyms word by word, and it would be more difficult for a student not to observe a poem’s diction or ambiguities, which were thrown into relief by the flat and prolix baseline version in Basic. As Richards elaborates elsewhere, “after a very little practice, the difficulty of writing Basic versions ceases altogether to be a difficulty in finding Basic turns of phrase and becomes a difficulty in deciding just what the original is, may, or might be saying . . . the work becomes, that is, very nearly a pure exercise *in interpretation*.”⁴⁵

Notice, then, that Richards was above all else attempting to minimize the unruly, subjective, and affective—in a word, human—interpretive tendencies that disrupt the analysis of a poem. And he reached for a statistical resource, the Basic word list, because it was a “limited medium”: limited in the sense that Basic is a finite pool of words, and limited also because the medium would curb (so Richards hoped) those “erratic associations, the interference of emotional reverberations” that lure readers away from the task of first decoding the underlying literal sense of a text, the task from which all good interpretations and judgments spring.⁴⁶ In other words, Richards worked hard to come up with an interpretive method that would accomplish exactly the same thing sought by today’s digitally performed statistical analyses: “to banish, or at least crucially delay, human ideation at the outset of interpretation,”

as Alan Liu has described the dream of “tabula rasa interpretation” of the digital humanities.⁴⁷ Of course, the word list more likely to be used today in tandem with close reading might be one statistically derived from the work(s) or author(s) in question: one digital humanities project has performed close readings of Emily Dickinson’s poetry with the aid of a statistical analysis of her letters.⁴⁸ This is a different approach than that of Richards or his student Empson, both of whom believed that the results of Basic’s statistical analysis of *the entire English language* should be used to translate, and construe, a poem in a manner that ensured the reader would be relatively unbiased by his or her feelings and associations.⁴⁹

Even if close readers do not—and relatively few close readers ever did—consult the single-page Basic sheet and formulate controlled-vocabulary paraphrases as a first step, Basic raises important theoretical questions about close and non-close reading: How would one even perform—or ever communicate to others—a close reading without something like Basic’s roster of simple, high-frequency words operating in one’s mind? Are not literary works, in part, statistical objects generated using Basic and non-Basic words, where the non-Basic words could by the “word substitution” formula be translated into Basic? This is precisely how Richards and Empson viewed poems. At the very least, then, in historical terms, it is now possible to recognize that statistical analyses worked in tandem with—as a component to facilitate, or a “plug-in” for—close reading; the inaugural theorization of the technique of close reading was itself shaped by that era’s “distant reading.”⁵⁰ By the same token, those who object to “distant reading” on principle, because of its perceived antagonism to close reading, will have to concede the important function played by statistical approaches at this interpretive procedure’s birth. Believed to be complementary, the two had “a good deal to do with one another,” and it may be productive for us to integrate the two once again in an equally or more thoughtful manner.

What happened thereafter to statistical analyses and close reading, once so intimate? I would like now to trace their gradual divergence into the field of reading research on the one hand, and Cleanth Brooks’s transformation of Richardsian close reading on the other. To begin with the former: *The Teacher’s Word Book* and similar statistical data were the conditions of possibility for the emergence of a body of research, beginning in the 1920s, in educational psychology on the problem of “readability.” An ongoing area of research today, readability is concerned with the “ease of understanding, owing to the style of writing.”⁵¹ After *The Teacher’s Word Book*, a spate of mathematical readability formulas appeared—including Thorndike’s own measure—from the 1920s until

computerized word counts of the 1960s.⁵² George R. Klare has written useful historical surveys of these readability formulas, noting that Rudolf Flesch's "Reading Ease" formula from 1948 is perhaps the most influential. Flesch believed that *The Teacher's Word Book* had wrought confusion between word frequency and the idea of easiness; on this issue only, Flesch concurred with Ogden that a word's frequency was not as significant as Thorndike and graders of textbook vocabularies assumed, let alone a measure of easiness for readers.⁵³ On the contrary, for Flesch, "simplicity . . . is mainly a question of sentence structure and concreteness of expression."⁵⁴ Flesch thus participated in the notable trend in post-Thorndikian readability formulas of taking frequency counts (Thorndike's or others') out of the equation altogether. The simplest version of Flesch's formula is a linear equation based only on the two variables of word shortness and sentence shortness:

$$\text{RE} = 206.835 - .846\text{wl} - 1.015\text{sl}$$

RE = Reading Ease, on a scale of 0-100, where 100 is very easy and 0 very difficult

wl = the number of syllables per 100 words

sl = the average number of words per sentence⁵⁵

For example, the Flesch "Reading Ease" of the first paragraph of this essay ("How the defining interpretive technique . . .") is approximately: $206.835 - (.846 \times 154) - (1.015 \times 29.167) = 46.946$, which is college-level readability. To the extent that "reading ease" here depends entirely on the shortness of words and sentences, it is not surprising that Flesch criticized Basic English, with its interminable sentences and awkward circumlocutions. At the same time, and as many instances of poetry testify, words of few syllables can be strung together into short sentences with confounding results. If *The Teacher's Word Book* confused word frequency and ease, and Basic English mistook the economy of its vocabulary for simplicity, then Flesch's ideal of "plain talk" errs too in assuming that brevity equals readability.⁵⁶ Yet Flesch's formula remains the most widely adopted readability equation in the world outside the confines of reading research, and versions of the formula continue to be featured today within "Readability Statistics" in Microsoft Word's "Spelling and Grammar" function.

Meanwhile, Brooks's study of Richards's work is amply documented, and it suffices to say that he was cognizant of Richards's work on Basic English and, by extension, the controversy over word lists.⁵⁷ In any case, Brooks knew for himself the genre of the word list: his first book, a purely philological study called *The Relation of the Alabama-Georgia Dialect to the Provincial Dialects of Great Britain* (1935), drew heavily from

a 1909 word list amassing dialectal and colloquial expressions of East Alabama.⁵⁸ Moreover, Brooks and his collaborator, Robert Penn Warren consulted Flesch's popular manual on readability, *The Art of Plain Talk* (1946), as they worked on *Modern Rhetoric* (1949), eventually borrowing Flesch's paraphrasing exercises—i.e., rewriting convoluted passages into “plain” style—for their composition textbook.⁵⁹ Once Brooks's reading is recollected in this way, one wonders: where are all these word lists and “statistical conceptions of language” in Brooks's theoretical works on interpretation and in New Criticism more broadly? The absence of these topics is nowhere more conspicuous than when Brooks addresses, throughout *The Well Wrought Urn*, the various ways by which poetry can be “difficult” without once mentioning vocabulary—the category subject to constant statistical analysis after Thorndike, central to Richards's Basic paraphrase exercise, and ubiquitous in Brooks's reading and immediate historical context.⁶⁰

The fate of the statistical analysis/close reading conjunction so prominent in Richards can be clarified by a brief look at Brooks's famous notion of the “heresy of paraphrase” in *The Well Wrought Urn*. The “heresy of paraphrase,” which of course becomes a signature tenet for New Critical interpretation, is the error of thinking that “the poem constitutes a ‘statement’ of some sort,” the notion that a poem can be reduced to a prose message.⁶¹ But the foregoing account reveals that Brooks is, paradoxically, of two minds when it comes to paraphrase. One view is signified by the “heresy of paraphrase”: the epistemological anxiety caused by science forces Brooks to insist that a poem's meaning resides not in any “content” extractable in the form of proposition—as in scientific fact—but rather in the paradoxical structure of a poem. The idea is that poetry has a greater “truth” than science, one that cannot be reduced to a verifiable statement and that can only be approached by the close reading of a poem's structure or form.⁶² On this count, Brooks's principle of the “heresy of paraphrase” severely undermines Richards's original outline of close reading: the noticing of tensions in a poem displaces paraphrase as the central constituent process of close reading.

But Brooks also understands paraphrase another way, one rooted in pedagogical reality. He is no more averse to the actual exercise of paraphrase than Richards. As we have seen, he relies on paraphrase exercises in *Modern Rhetoric*, but he understands its indispensability in the reading of poetry, too: he acknowledges that paraphrase can be “a necessary preliminary step in the reading of a poem,” and even recognizes that his interpretations in each chapter of *The Well Wrought Urn* are nothing other than prose adumbrations, perhaps even Basic-like.⁶³ From this perspective, paraphrase is hardly heresy. Rather, Brooks is

only repeating the caution that we have seen Richards already advise, namely that students asked to paraphrase a poem can be prone to “write a poem—in prose—on a partially similar theme,” and thereby to overlook other facets of the poem (such as tone); recall that this was, for Richards, the major rationale for paraphrasing in Basic.

Thus Brooks’s memorable catchphrase, “the heresy of paraphrase,” ultimately obscures his quite ambivalent stance on paraphrase and leads to the dismissal of paraphrase as a whole. And with momentous consequences: reading at the elementary and secondary levels, the problem of vocabulary, the work of Thorndike and his heirs, readability, and the joining of statistical analyses and interpretation are all excised from New Critical interpretation. This point is confirmed as soon as we contrast this development with Richards’s own conviction that the problem of “close reading” would have to be pursued into the lower levels of the educational system. “I decided to back out of literature, as a subject, and go into elementary education,” wrote Richards, who would later in his career focus on literacy, even writing children’s textbooks toward this end.⁶⁴

How might we, at present, avoid replicating the New Criticism’s failure to incorporate reading research at the very moment when its concerns overlapped most with those of reading research? Digital affordances allow for enlightening, valuable forms of statistical analyses of literary works and literary history, but this account suggests that there is also an opportunity for those in literary studies once again to learn from and contribute to reading research, which has studied interpretation and used statistics all along. English professors like to gripe about the extent to which their students are unprepared for college-level reading and writing. We acknowledge, then, that instruction in the kinds of knowledge and ability involved in literary studies does not begin on the first day of college. What if we invested more time studying literacy, reading, and writing at the elementary, secondary, and university levels—attending not only to the last, but integrating all three—aiming ultimately at one of Basic’s original goals, “encouraging clarity of thought and expression”?

UNIVERSITY OF CONNECTICUT

NOTES

1 Alan Liu, “The Meaning of the Digital Humanities,” *PMLA* 128, no. 2 (2013): 415. See also the work that Liu treats at length, Ryan Heuser and Long Le-Khac, “A Quantitative Literary History of 2,958 Nineteenth-Century British Novels: The Semantic Cohort Method,” *Stanford Literary Lab Pamphlet* 4 (2012), for Heuser and Le-Khac’s own valuable reflections on their “hypothesis-testing mode of interpretation” (49).

2 I. A. Richards, *Practical Criticism: A Study of Literary Judgment* (New York: Harcourt, Brace & Co., 1929), 174 (hereafter cited as *PC*).

3 Rita Raley has convincingly shown how machine translation, from its advent at midcentury to the present day, has adopted Basic English's logics of functionality, performativity, and universalism. See Raley, "Machine Translation and Global English," *The Yale Journal of Criticism* 16, no. 2 (2003): 294–95. As it will become clear, my account also concurs with David Simpson, "Prospects of Global English: Back to BASIC?" *The Yale Journal of Criticism* 11, no. 1 (1998): 301–7, which decries the relative neglect of Basic English by those in literary studies. Many thanks to Christine Mitchell for alerting me to the link between Basic English and machine translation, and for other helpful thoughts in response to an earlier iteration of this essay.

4 Franco Moretti, "Conjectures on World Literature," *New Left Review* (2000): 57; N. Katherine Hayles, *How We Think: Digital Media and Contemporary Technogenesis* (Chicago: Univ. of Chicago Press, 2012), 11, 29.

5 For example, C. K. Ogden, and most likely Richards too, knew the work of their contemporary, the Cambridge statistician G. Udny Yule, including *The Function of Statistical Method in Scientific Investigation* (1924) and his influential later work, *The Statistical Study of Literary Vocabulary* (1944).

6 John Guillory, "Close Reading: Prologue and Epilogue," *ADE Bulletin* 149 (2010): 8. This essay is also informed by Guillory's other work on close reading, particularly a chapter on the origins of close reading circulated for a colloquium at New York University in 2007 and a talk on Richards and neurophysiology, most recently delivered at the 2015 Modern Language Association convention.

7 This origin story is repeated, for example, in Hayles, *How We Think*, 23 and Matthew L. Jockers, *Macroanalysis: Digital Methods and Literary History* (Urbana: Univ. of Illinois Press, 2013), 3.

8 With "statistical thinking," I allude to Theodore M. Porter, *The Rise of Statistical Thinking, 1820–1900* (Princeton, NJ: Princeton Univ. Press, 1986), which charts the history of predisciplinary statistical thought as it flourished in the contexts of the social and biological sciences. As I mention below in connection with stenography, there are many interesting intersections between statistical thinking and languages before the twentieth century. On the emergence of corpus linguistics, see Tony McEnery and Andrew Hardie, "The History of Corpus Linguistics," in *The Oxford History of Linguistics*, ed. Keith Allan (Oxford: Oxford Univ. Press, 2013), 728.

9 Lisa Gitelman, *Paper Knowledge: Toward a Media History of Documents* (Durham, NC: Duke Univ. Press, 2014), 56.

10 For "quantitative," the OED provides, "relating to, concerned with, or based on vowel length. Also in *Prosody*: (of meter, etc.) based on syllable length rather than the placing of stress." The fact that any number of methods and topics are located at the intersection of the "quantitative" and literature—e.g., formalist approaches concerned with meter, numerology, the relation of political economy to literary discourse, the dual career of Wallace Stevens—suggests that the idea of the quantitative or quantification is less than ideal as a name for the non-close kinds of reading it typically describes today.

11 Jockers, *Macroanalysis*, 9.

12 On the question of the quantity of texts engaged, my account could converge with Stephen Best and Sharon Marcus's concept of "surface reading," insofar as surface reading too can be performed on a single text or many. In forswearing "symptomatic reading," Marcus and Best offer the very large category of "surface reading," which encompasses a diverse range of extant and emergent scholarly methods and practices (book history, close reading, cognitive approaches, narratology, discourse analysis, digital methods), all of which are uninterested in symptomatic ideology critique. Statistical analyses can be a

kind of surface reading, but presumably they can just as easily serve a symptomatic reading; there is nothing intrinsically “surface” about literary applications of statistics. See Stephen Best and Sharon Marcus, “Surface Reading: An Introduction,” *Representations* 108, no. 1 (2009): 1–21.

13 Stephen Ramsay, *Reading Machines: Toward an Algorithmic Criticism* (Urbana: Univ. of Illinois Press, 2011), 1.

14 It merits remembering, too, that although Moretti’s work has come to signify the textual analysis methods of the digital humanities, the important “distant reading” experiments that he recounts in “The Slaughterhouse of Literature” were performed by human reading and tabulation. See Moretti, “The Slaughterhouse of Literature,” *MLQ* 61, no. 1 (2000): 220.

15 Raymond Williams, for example, explains how technologies are for a long time “looked for and developed with certain purposes and practices already in mind” in *Television: Technology and Cultural Form* (London: Routledge, 2003), 7; David Thorburn and Henry Jenkins discuss how older practices persist through and after media transition in “Introduction: Toward an Aesthetics of Transition,” in *Rethinking Media Change: The Aesthetics of Transition*, ed. Thorburn and Jenkins (Cambridge, MA: MIT Press, 2003), 1–16. See also Alan Liu, “The Meaning of the Digital Humanities,” 416–18, which advocates for a “science and technology studies” approach with a longer view of technology and media.

16 “Dr. E. L. Thorndike, Psychologist, Dies,” *The New York Times*, 10 Aug. 1949. Accessed 12 Jan. 2014.

17 Edward L. Thorndike, “Word Knowledge in the Elementary School,” *Teachers College Record* 22, no. 4 (1921): 334.

18 Thorndike, “Word Knowledge,” 334–41. While it is unclear who exactly assisted Thorndike on the first edition of his word book, Thorndike and Irving Lorge’s later semantic count relied on WPA workers. See, for example, Works Progress Administration, *Index of Research Projects*, Vol. 1 (Washington: U.S. Government Printing Office, 1938–39), 67. Whether involving Depression-era workers, “female information workers,” Amazon Mechanical Turks, students, or others, the clerical tasks that go into massive endeavors of counting, tabulation, typing, reproduction, corpus creation, OCR correction, and so on, all raise questions about unjust labor practices. Regarding early twentieth-century “female information workers,” see Natalia Cecire, “Ways of Not Reading Gertrude Stein,” *ELH* 82, no. 1 (2015): 281–312.

19 Charles C. Fries and A. Aileen Traver, *English Word Lists: A Study of Their Adaptability for Instruction* (Ann Arbor, MI: The George Wahr Publishing Co., 1950), 2.

20 In addition to shorthand, concordance building also involves handwriting; see Lane Cooper’s description of how he compiled his Wordsworth concordance, “The Making and the Use of a Verbal Concordance,” *The Sewanee Review* 27, no. 2 (1919): 191–95. Cecire’s “Ways of Not Reading Gertrude Stein” shows, via a series of brilliant readings and connections, that a real history of “distant reading” would also involve feminized labor, including typing, as well as the controversy around the style that came to be known as “Steinese.” And Benjamin Morgan, in his book project, *The Outward Mind: Materialism, Science, and Aesthetics in Nineteenth-Century Britain* (email correspondence with author), recovers the important Victorian-era “literametrics” of Thomas Corwin Mendenhall, L. A. Sherman, and Robert Moritz, among others, which compared word- and sentence-length distributions for different authors and genres.

21 Fries and Traver, *English Word Lists*, 4–18.

22 Fries and Traver, *English Word Lists*, 21.

23 Thorndike, “Word Knowledge,” 351, 354. Thorndike is opaque about the credit system he devised to account for a word’s dispersion across different kinds of reading, and asks his readers “to accept arbitrarily these credits” (“Word Knowledge,” 335n1).

- 24 Thorndike, "Word Knowledge," 355.
- 25 Geraldine Joncich, *The Sane Positivist: A Biography of Edward L. Thorndike* (Middletown, CT: Wesleyan Univ. Press, 1968), 393.
- 26 Lorge, "The English Semantic Count," *Teachers College Record* 39, no. 1 (1937): 65–77; Lorge, *The Semantic Count of the 570 Commonest English Words* (New York: Institute of Psychological Research, Teachers College, Columbia University, 1949), v–ix.
- 27 Richards and Reuben Brower, "Beginnings and Transitions," in *I. A. Richards: Essays in His Honor*, ed. Brower, Helen Vendler, and John Hollander (New York: Oxford Univ. Press, 1973), 34.
- 28 C. K. Ogden, "The Universal Language," *Psyche* 9, no. 3 (1929): 1.
- 29 Ogden, *The System of Basic English* (New York: Harcourt, Brace and Co., 1934), 4 (hereafter cited as *SBE*).
- 30 *SBE* 151; Ogden, "Progress of Basic," *Psyche* 10, no. 2 (1929): 36. On the one hand, Basic can be highly effective at cutting through jargon, bureaucratese, and obscurantism—and poetry. On the other hand, its limited vocabulary and the wordiness it generates are impractical, and Ogden and Richards dealt with this aspect of Basic in amusingly compulsive ways: they churned out translation after translation of excerpts and whole texts in Basic (the Bible, *Carl and Anna*, *Robinson Crusoe*), as well as glossaries and other metadiscursive guides; and oftentimes, Ogden and Richards, in their books promoting Basic, secretly start writing in Basic at a certain point in order to ask readers later to concede how unnoticeable the transition from English to Basic English was, i.e., how smoothly Basic can in fact read (see, for example, Richards, *Basic English and Its Uses* [London: Kegan Paul, Trench, Trubner & Co., 1943], 20).
- 31 Ogden and Richards would become divided, even estranged, by these two aims. Ogden fantasized about the global adoption of Basic, and he refers often to "debabelization." Richards, as well as allies such as William Empson, saw Basic's greatest use in its capacity to facilitate the comprehension of texts. Eventually, Richards officially severed his ties from Ogden's Basic movement, although he never stopped championing Basic's usefulness as a tool toward close reading. See, for example, Richards, *Selected Letters of I. A. Richards*, ed. John Constable (Oxford: Clarendon, 1990), 117–20.
- 32 Murray Cohen, *Sensible Words: Linguistic Practice in England, 1640–1785* (Baltimore: Johns Hopkins Univ. Press, 1977), 19, and Kevis Goodman, *Georgic Modernity and British Romanticism: Poetry and the Mediation of History* (Cambridge: Cambridge Univ. Press, 2004), 24–26.
- 33 Ogden, "Editorial," *Psyche* 9, no. 5 (1929): 5–6.
- 34 Ogden, "Progress of Basic," 22–25; *SBE* 22–32.
- 35 Thorndike, "Word Knowledge," 364, emphasis added. Despite this ambiguous early comment, Thorndike later distances his statistical word lists from lingua franca projects. See *Interim Report on Vocabulary Selection for the Teaching of English as a Foreign Language* (London: P.S. King & Son, 1936), 1.
- 36 Lydia H. Liu, *The Freudian Robot: Digital Media and the Future of the Unconscious* (Chicago: Univ. of Chicago Press, 2011), 91; Raley, "Machine Translation and Global English."
- 37 Liu, *The Freudian Robot*, 99.
- 38 Ogden, "Editorial," 9.
- 39 Richards, *Basic English and Its Uses*, 62.
- 40 Richards, *Selected Letters*, 46.
- 41 John Paul Russo, *I. A. Richards: His Life and Work* (London: Routledge, 1989), 294–316. Since Richards knew well Thorndike's psychological and educational-psychological writings, one wonders whether Richards's practical criticism experiments were influenced by Thorndike's reading comprehension experiments, as well as Thorndike's classification of reading errors (on which see Richard L. Venezky, "The History of Reading Research," in

Handbook of Reading Research, ed. P. David Pearson, Rebecca Barr, Michael L. Kamil, and Peter Mosenthal [New York: Longman, 1984], 16–17).

42 Richards, *Basic in Teaching: East and West* (London: Kegan Paul, Trench, Trubner & Co., 1935), 56–57.

43 Richards, *Basic in Teaching*, 49; Russo, I.A. *Richards*, 298.

44 Richards, *Basic English and Its Uses*, 107–9.

45 Richards, *Basic in Teaching*, 84.

46 Richards, *Basic English and Its Uses*, 94; PC 13.

47 Alan Liu, “The Meaning of the Digital Humanities,” 414. For a lyrical dismantling of this dream and some aspects of “surface reading,” see Ellen Rooney, “Live Free or Describe: The Reading Effect and the Persistence of Form,” *differences* 21, no. 3 (2010): 112–39.

48 See Matthew Kirschenbaum, “The Remaking of Reading: Data Mining and the Digital Humanities,” <http://www.csee.umbc.edu/~hillol/NGDM07/abstracts/talks/MKirschenbaum.pdf>.

49 Empson provides a good example of how Basic English can be used toward close reading in “Basic English and Wordsworth (A Radio Talk),” *The Kenyon Review* 2, no. 4 (1940): 449–57.

50 I borrow Alan Liu’s memorable use of the word “plug-in” in “The Meaning of the Digital Humanities,” 415.

51 George R. Klare, “Readability,” in *Handbook of Reading Research*, ed. Pearson et. al., 681.

52 Jeanne S. Chall, “The Beginning Years,” in *Readability: Its Past, Present, and Future*, ed. Beverley L. Zakaluk and S. Jay Samuels (Newark: International Reading Association, 1988), 2–7; Venezky, “The History of Reading Research,” 24.

53 Rudolf Flesch, *The Art of Plain Talk* (New York: Harper & Brothers, 1946), 173.

54 Flesch, “How Basic is Basic English?” *Harper’s Magazine* (1944): 341.

55 George R. Klare, *The Measurement of Readability* (Ames: Iowa State Univ. Press, 1963), 58–59. There are various other measures offered by Flesch himself (see *The Art of Plain Talk*, 58–65) or built upon Flesch’s original measure, e.g., the Flesch and Kincaid measure (see Klare, “Readability,” 693).

56 Guillory, “The Memo and Modernity,” *Critical Inquiry* 31, no. 1 (2004): 123–32, and especially 126n49.

57 Cleanth Brooks, “I. A. Richards and *Practical Criticism*,” in *The Critics Who Made Us: Essays from Sewanee Review*, ed. George Core (Columbia: Univ. of Missouri Press, 1993), 46.

58 Brooks, *The Relation of the Alabama-Georgia Dialect to the Provincial Dialects of Great Britain* (Baton Rouge: Louisiana State Univ. Press, 1935), 4.

59 Brooks and Robert Penn Warren, *Cleanth Brooks and Robert Penn Warren: A Literary Correspondence*, ed. James A. Grimshaw, Jr. (Columbia: Univ. of Missouri Press, 1998), 134–137.

60 Brooks, *The Well Wrought Urn: Studies in the Structure of Poetry* (New York: Harcourt, 1947), 76 *inter alia*.

61 Brooks, *The Well Wrought Urn*, 196.

62 Guillory, *Cultural Capital: The Problem of Literary Canon Formation* (Chicago: Univ. of Chicago Press, 1993), 159–160.

63 Brooks and Warren, *Understanding Poetry* (New York: Henry Holt and Co., 1938), xi; Brooks, *The Well Wrought Urn*, 206.

64 Richards and Brower, “Beginnings and Transitions,” 29.