ॐ

Writing, Reception, Intertextuality: Networking Women's Writing

Sarah Connell and Julia Flanders
Northeastern University
Boston, Massachusetts

Reading has come under renewed scrutiny in the digital age, a result of the defamiliarization of the medium that has also brought about a thorough rethinking of what is meant by "text," "book," "author," "publish," and other terms that map our understanding of how ideas are circulated through technologies of inscription. The meaning and method of reading within digital spaces have vacillated between two extremes of scale: the vastness of aggregation made possible by large-scale text digitization (exemplified early on by text corpora such as the *Thesaurus Linguae Graecae* and the *Middle English Dictionary*, and later by *Early English Books Online*), and the minute scrutiny made possible by the detailed modeling of individual instances, as for instance in digital scholarly editions like the *Piers Plowman Electronic Archive*. Since the early 2000s, fascination with large-scale data analysis has shifted attention toward modes of reading that sample the source to produce a derived text—a statistical artifact from which we can in turn read clusters of words, shifts in topic or register, and changes in orthographic habits.[1] These remote reading practices, however, fail to capitalize on valuable modeling of the individual text, but more recently researchers have been exploring ways of bringing these two ends of the digital spectrum into closer conversation. This rapprochement is particularly intriguing for the study of textual reception, since it opens up for analysis the points of connection between an individual text and the full network of cultural connections that constitute its circulation, readership, and reception: authors, publishers, readers, reviewers, and collectors. This essay explores the study of readership and reception of women's writing through the lens of these emerging digital methods, examining two corpora related to women's writing with large-scale analytical methods that are dependent on and informed by the detailed textual models in these collections' metadata and markup. Our focus is on the late eighteenth century, but the collection whose reception we are study-

ing extends back to the sixteenth century, and these methods apply more broadly still.

The *Women Writers Project* (*WWP*) at Northeastern University is itself a contemporary experiment in the recirculation and reawakening of readership for early women's writing, and it takes on a special challenge by attempting to overcome a major discontinuity in that circulation. We are partly talking here about the perception, even among scholars of literature, that early women's writing generally speaking did not exist except as a set of isolated exceptions, a pervasive misperception that amounts to a thorough erasure of women's writing as a literary category in both popular and scholarly awareness. But we are also talking about the loss of a frame of reference for reading these writings—for understanding their position within our sense of genre, for assessing their typicality or exceptionalism, for hearing their dialogues with contemporary authors and issues. The *WWP* is an attempt to recover from both of these forms of loss in two ways: first, by putting early women's writing back into wide circulation (via *Women Writers Online*, the project's primary publication), taking advantage of digital media to increase the dissemination and reading of these materials; and second, by providing environments for exploration to help us gain back (or create afresh) some understanding of those missing frames of reference, including contemporary readership.[2]

This essay focuses on the *Cultures of Reception* project, a grant-funded initiative that concluded in 2016, and its chief outcome: *Women Writers in Review*, a collection of periodical reviews and other reception materials that respond to women's writing.[3] It may be helpful first to situate that project within the overall contours of the *WWP*'s work.[4] At the core of the *Women Writers Project* is *Women Writers Online* (*WWO*), a collection which now includes about 400 texts by women, in English, from the early sixteenth century through 1850. These texts are transcribed and digitally encoded using the TEI Guidelines, an international language for representing scholarly research materials in digital form.[5] This encoding captures the essential structural components of the text, a wide range of more specific aspects of the content (such as quotations, annotations, names, citations, and dialogue), and details relating to the transcription and editing of the text (including handling of typographical errors, spelling regularization, and unclear or deleted text). *WWO* stands at the center of a growing set of related open-access digital materials that now includes *Women Writers in Context* (a collection of contextual essays and exhibits), *Intertextual Networks* (a comprehensive bibliography, still in development, of sources quoted or

cited in *WWO*), and *Women Writers in Review*.[6] These materials all link to the relevant *WWO* works, and are in turn linked to from *WWO*. Overall, the *WWP* seeks to build an interconnected web of information in which individual women's texts are shown not in isolation, but instead are contextualized in multiple ways. Clearly, this context is not a comprehensive reconstruction of some imagined contemporary web of connections; the coverage is not fully representative, and the interlinking is focused on a few specific channels of connection. But although it can't restore in full those lost cultural textures, this contextualizing endeavor can offer a sense of women's writing as something abundant, self-aware, embedded, even commonplace.

Women Writers in Review (*WWiR*) is the part of this ecology that speaks most directly to the themes of this special issue, and also most fully exemplifies the project's contextualizing motives. *WWiR* includes about 600 items from the period 1770–1830, predominantly periodical reviews but also advertisements, biographical materials, letters, and other genres. (For simplicity in this essay we refer to all of the *WWiR* contents as "reviews.") The design of *WWiR* is motivated by an emphasis on interconnectivity. The data is designed to provide many points of potential interconnection, including the text under review, the author of the text under review, the periodical in which the review was published, the date and place of publication of the text under review and of the review itself, the genre of the review, the evaluation provided by the review (negative, positive, neutral, mixed), and specific thematic keywords of particular relevance to our research. The publication's interface is structured around these interconnections, offering paths of exploration that traverse these different categories of analysis. For instance, a reader interested in negative reviews might focus on negative reviews of a group of authors, and then notice that they were all published in the same periodical, which might lead them to next explore the other authors reviewed there.

Several clusters of research questions animate this project. In very broad terms, we seek to understand how to model what we mean by "reception history" in the digital representation of a collection of documents, and to use that modeling to make visible patterns from the reviews themselves. From these patterns, we hope this material will enable researchers to shed light on specific questions: How do periodical reviews in this period imagine the relationship between the local space of writing (an author's situatedness within a specific region or nation) and the increasingly national and transnational space of reading? On what stage are women writers presumed to be performing, and what are the consequences for their reception? Reciprocally,

how do reviews work to constitute for women authors a sense of a reading public, and to what extent is that public itself constituted as a located, situated readership or as more broadly construed and geographically vague (or far-flung)? For this article, we focus our attention on the ways in which accounts of female readership and authorship are implicated in the larger operations of literacy, and we explore several interconnected methods for teasing out the implications of these accounts.

The reviewing enterprise and the topic of readership

Like so many digital collections that support complex cultural analysis, the *WWiR* collection of reviews is a proxy. What we learn from this collection is not (directly at least) about women's writing or how this writing was read by readers; rather, *WWiR* allows us to better understand what these readers wanted to read (and hence would buy from reviewers) concerning their own reading practices—real, imagined, or idealized. In reading through even a small number of reviews in the *WWiR* collection, it becomes evident that the authors of these texts position themselves at the outset as operating in the service of readers who have too much to read, and in so doing they bring to public visibility the very phenomenon of readership as a social fact. Periodical reviewers in the mid-eighteenth century began self-consciously as "tasters to the public"—sampling everything (at a time of increasing overabundance and widely varying quality), and providing summaries to help readers make choices.[7] As a result, reviews from this period emphasize long textual extracts that are presented as typical of the work (and in fact the language of those quotations is in itself an important subset of the material for study). Reviewers serve as a kind of imagined proxy for public responses, or for how the public ought to respond; the reviewer operates as an ideal reader, setting a good example and telling the public what it should approve and disapprove, how it should take its readerly pleasure, and what forms of that pleasure are legitimate and illegitimate. These reviewers express themselves through a "we" that makes common cause between the reviewer and the reading public ("we should expect it," "we should at least feel *inclined* to respect the religious opinions of such a character"), but it also marks out the reviewer as having a distinct (and exemplary) moral or critical position that is grounded in the pluralized identity of the periodical itself ("we were going to object to the extreme youth of Lady Julia," "we frankly confess ourselves puzzled").[8]

Because these reviews are often highly conventional in their rhetoric, and because their judgments of writing and writers are deeply bound up with shared cultural values (or, at least, with values that are presented as shared), scanning even a modest subset of individual reviews provides a surprisingly clear picture of the whole collection. However, methodologies for large-scale text analysis enable us to move beyond such selective observations and pursue more focused inquiries into eighteenth- and nineteenth-century transatlantic review culture as represented in the collection. These methods also bring into focus patterns of language that can direct our attention back to local examples and opportunities for closer interpretation. For this study, we first sought to elicit what we might term a "verbal thematics" of eighteenth- and nineteenth-century review culture, working with a technique called "topic modeling."[9] Then, using the results of that analysis to guide our attention to an especially salient vocabulary cluster, we went back to the texts and examined some of these key terms in their contexts to understand more fully how these reviews mobilize a politics of gendered literacy and readership.

For the purposes of this discussion, we can understand topic modeling as a technique for discovering groupings of words ("topics") that frequently appear together across a collection of documents, and for showing how prominent each topic is within a given document. As many scholars have noted, the term "topic" is something of a misnomer, since these clusters of words result from a probabilistic analysis of word distributions within the document rather than from any actual understanding on the part of the algorithm of what the document is "about"; in some cases, the clusters may simply be artifacts of the document's organization or oddities of typography. In many cases, however, these clusters can both reveal specific preoccupations of a collection of texts, and also identify the lexicon through which these preoccupations are expressed. For example, we have identified in the *WWiR* collection a topic that comprises language for describing and reviewing drama, using terms such as: *acted*, *comedy*, *drama*, *dramatic*, *farce*, *plays*, *representation*, *stage*, *theatre*, and *tragedy*.

Our experimentation with topic modeling has also contributed to our larger-scale research into methods that can usefully combine text encoding and text analysis. Because the *WWiR* corpus has been encoded following the *TEI Guidelines*, we can use markup to select specific texts, or segments of texts, as the focus of analysis. For the initial phase of our research into *WWiR*, we created three corpora. The first contains the full set of review

materials in the *WWiR* collection: approximately 605,000 words from 632 texts.[10] The second corpus was created by extracting all quoted materials from the full corpus: approximately 275,000 words from 370 texts. This corpus includes the (often lengthy) extracts from the work under review that are quoted within the reviews, along with a smaller number of quotations from other texts, including other periodical reviews.[11] This second corpus is useful as a way of studying the parts of the reviewed works that reviewers find particularly notable or representative. The third corpus focuses on review language. For this we selected only the texts that were marked in their metadata as being either theatrical or literary reviews, rather than publication notices, biographies, letters, poems, or the other non-review genres represented in *WWiR*. We then removed all of the quotations and headings, producing a corpus restricted to the language of the literary and theatrical reviews themselves: 310,000 words in 515 texts.[12] For all three corpora, we also included some basic metadata for each text: reviewed author, reviewed work, periodical publication title, and review publication date.

Topic modeling is a complex technique with various parameters that can be adjusted (including the number of topics and the resulting level of granularity in topic differentiation); in addition, because it is probabilistic rather than deterministic, even with identical parameters two models trained on the same corpus will not be identical. In the materials accompanying this article, we provide the full details of our model training process as well as the data sets we used, for those who are interested in continuing these experiments.[13] Here it may be sufficient to note that we undertook significant experimentation to determine a number of topics that effectively distinguished interesting thematics; we tested our parameters until we found those that delivered broadly consistent results; we trained multiple instances of each model and aggregated word-level analysis from multiple models (in order to identify the areas where they agree); and we focused on patterns that appear in multiple models.[14]

The topics generated give a revealing conspectus of the thematic concerns of these reviews. For example, the reviews corpus has a topic comprising language for describing and evaluating poetry, with the following terms appearing consistently:

> *beautiful, compositions, elegance, elegant, fancy, genius, imagery, line, lines, muse, pathetic, pieces, poem, poems, poet, poetic, poetical, poetry, prose, publication, robinson, select, sonnet, sonnets, specimen, verse, verses, volume.*[15]

Similarly, the quotations corpus (a distillation of exemplary literary language from the reviewer's perspective) has a topic that we can identify as poetic language:

> *ah, breast, cold, day, er, eye, eyes, fair, form, head, heart, hope, life, light, love, nature, pale, rose, soft, soul, st, sweet, thee, thine, thou, thy, time, tis, vain, woe, youth.*[16]

Even more significantly, these two complementary topics are distributed in similar proportions across the individual texts in their respective corpora—that is, the texts that have a higher proportion of the poetic language topic also tend to have a higher proportion of the poetry review language topic, which makes sense but also reinforces the idea that there is some genuine "topicality" in evidence here.[17]

A similar effect occurs with a set of topics representing the language of novelistic description. These topics share not only distributions but also some vocabulary; several terms appear consistently across all three corpora—*father, friend, house, mother, son*—and some further terms appear very frequently—*family, sir, woman, daughter, lord, master, hand, poor, time, wife, heart, lady.* Focusing on the two mutually exclusive corpora—reviews and quotations—is particularly illuminating. The terms most strongly associated with this topic in the quotations corpus but not the reviews corpus are those used for setting and character descriptions, as well as dialogue markers:

> *air, country, day, dear, door, eyes, fine, found, friends, half, heard, honour, hope, ladies, left, life, manner.*

For example, in the quotations corpus, the phrase "my dear" appears 113 times and the construction ", dear" (which signals its use as an apostrophe as in "'tis love, dear babe") appears 56 times, largely marking instances where characters are speaking to each other.[18] By contrast, the key terms for this topic in the reviews corpus and not the quotation corpus represent the language that reviewers use for plot summary:

> *brother, castle, child, consequence, death, falls, fortune, husband, london, lover, marriage, married, mistress, received, return, takes, visit.*

For example, the construction "in consequence" appears 30 times in the collection of reviews, often to relate events from reviewed works ("in consequence of his horse having fallen" and "in consequence of her misconduct").

"Falls" is often used in present-tense plot summary; the phrase "falls in love" (with various adverbs) appears twelve times. There are also present-tense descriptions of characters falling into each other's arms or grasps, off their horses, and at each other's feet. The shared vocabularies of these two topics are thus largely those that name common descriptors and roles of persons in novels, while the distinct terms reveal differences in how the reviews and the quoted material describe or relate the stuff of novels: plots, characters, and settings.

Of particular significance for our examination of reviewing culture is the language concerning reviewing, readership, and authorship. Both the full *WWiR* corpus and the reviews corpus showed the strong presence of a topic in this area, with such striking overlap in vocabulary that we feel justified in treating these as essentially the same topic. In the full *WWiR* corpus, the topic was represented in the following cluster of terms:

> *art*, *author*, *character*, *effect*, *found*, *language*, *manners*, *means*, *natural*, *pleasure*, *public*, *read*, *reader*, *readers*, *story*, *style*, *taste*, *times*, *writer*.

In the reviews corpus (where the language of the reviewed works and of non-review genres is excluded), the topic was represented by a very similar set of terms (words in bold are common to both sets):

> *account*, *appears*, **author**, **character**, *characters*, *common*, *degree*, *description*, **effect**, *english*, *fair*, *introduced*, *lady*, **language**, **means**, **pleasure**, *productions*, **public**, **read**, **reader**, **readers**, **story**, **style**, *tale*, *volumes*, **writer**.

This alignment is more striking than might at first appear: the presence of the topic in the full *WWiR* corpus is not simply an artifact of its presence in the reviews-only corpus but an independent result. This is a consistently prominent topic in both the full *WWiR* corpus and the reviews corpus, in the sense that almost every text in the collection has some proportion of the cluster of words that belong to the topic. For example, in one model trained on the reviews corpus, all but two of the 515 texts (99.6%) have this topic in proportions of at least 10%; 96% of the texts have this topic in proportions of at least 20%; and the average proportion across the entire set of texts is 37%. It is not, perhaps, surprising that language about reading, writing, and reviewing dominates in a collection of reviews, or that terms like *read*, *reader*, *readers*, *writer*, and *author* would be so closely associated.[19] But

this topic also suggests some of the more specific thematics of reviewing, in the set of terms that keep company with these. Many of the terms associated with this pair of topics indicate the reviewers' roles as tastemakers for the reading public, as discussed above (*art*, *effect*, *manners*, *merit*, *pleasure*, *public*, *read*, *reader*, *readers*, *style*, *spirit*, *taste*, *times*). We also see foregrounded terms for discussing gender (*fair*, *lady*), nationality (*english*), and the work of authorship (*account*, *character*, *characters*, *description*, *language*, *productions*, *story*, *tale*, *volumes*). What we see here is thus a vocabulary for transatlantic periodical reviews' descriptions of the acts of reading, writing, and reviewing itself. The high proportion of this topic within the full *WWiR* corpus (which also contains language from the texts under review), along with the fact that the topic includes both language for discussing authorship *and* language for discussing review and evaluation, reflects the highly self-referential nature of reviewing in this period and reinforces the close links between reviewers, authors, and readers.[20]

Reading the gendered "reader"

Topic modeling, in this experiment, directs our attention to the ways a semantic field—the "aboutness" of a document or collection—is instantiated in specific words, even though the topic modeling technique does not itself detect this "aboutness," but instead highlights verbal phenomena that suggest it to us. What topic modeling does not show us is how the broad vocabulary of the topic (words like *reader* or *author*) is inflected by local contexts of usage, by the words that do not show up as part of the topic. We get a hint from the vocabulary of this "reviewing/readership/authorship" topic that gender and nationality are somehow associated with its semantic field. But in order to understand more specifically how the thematics of gender come into play, we need to look closely at words in their contexts and we need a different set of reading techniques for this examination. For reasons of space we will focus here on readership, but this approach could be extended to the full set of terms suggested by the review language topic to produce a very broad and detailed view of its operations across the corpus.

As a first step, we extracted all instances of the words *reader* and *readers* in *WWiR*, and identified the associated adjectives, verbs, and descriptive phrases.[21] To provide some contrast, we also did the same analysis of *WWO*, showing how readership is figured in the literary language to which reviewers respond. The most striking pattern is a fundamental difference between the adjectives associated with *reader* and *readers*: the words asso-

Table 1. Possessive adjectives associated with the plural readers in the WWO *and* WWiR *collections.*

Term	Count and frequency in *WWO*	Count and frequency in *WWiR*
our readers	42 (9.7% of all adjectives with "readers")	152 (52.0% of all adjectives with "readers")
her readers	9 (2.1%)	33 (11.3%)
his readers	18 (4.2%)	8 (2.7%)
my readers	172 (39.8%)	7 (2.3%)
their readers	4 (0.9%)	7 (2.3%)
your readers	44 (10.2%)	7 (2.3%)
Total	288 (66.9%)	214 (73.3%)

ciated with the plural *readers* are overwhelmingly possessives ("my," "our," "her," "your," "his," "their"), which together represent between 67% and 73% of all adjectives used with *readers* in these collections (see table 1). By comparison, possessives constitute just 13% (*WWO*) and 12% (*WWiR*) of the adjectives used with the singular *reader*. Within this set of possessives, there are also some noticeable differences between the *WWiR* and *WWO* collections. When invoking readers in the aggregate, authors of both reviews and *WWO* texts overwhelmingly do so through a phrasing that marks a connection ("our readers," "my readers"); moreover, the language of reviewing does so in a way that identifies the reviewer as speaking from a kind of collective viewpoint rather than as an individual.[22]

Because the descriptors applied to the singular *reader* are quite diverse, and because no individual term has widespread use, we can most clearly see broader characterizations when we aggregate groups of thematically related terms (see table 2). Although these thematic characterizations are shared between the language of *WWO* and *WWiR*, from their distributions within and between the collections we may discern something interesting about how readership is imagined in these two contexts.[23] First, within *WWO*, the adjectives grouped under the "dear reader," the "friendly reader," and the "virtuous reader" together account for nearly half of the adjectives used with *reader* in *WWO*, compared with only 12.8% of usage in *WWiR*. *WWO* authors seem to hail the readers of their own works predominantly in terms that call the reader into a relationship of emotional rapport. In comparing the specific readerly characterizations between *WWO* and *WWiR*, not all of the differences are statistically significant, but notably those that are occur in a higher proportion of instances in *WWiR*: the "judi-

Table 2. Thematically related adjectives associated with the singular reader in the WWO *and* WWiR *collections.*

Concept	Adjectives	Count and frequency in *WWO*	Count and frequency in *WWiR*
The dear reader	gentle, my, dear, beloved	72 (24.1% of all adjectives with "reader")	3 (6.4% of all adjectives with "reader")
The friendly reader	courteous, constant, friendly, good-natured, agreeable, indulgent	44 (14.7%)	2 (4.3%)
The judicious reader	impartial, judicious, candid, ingenuous, sensible, unprejudiced, liberal, unbiassed, unpartial, cautious, fastidious, justly thinking, understanding	33 (11.0%)	7 (14.9%)
The religious reader	Christian, godly, Catholic, Protestant, lay, religious	26 (8.7%)	NA
The virtuous reader	good, kind, worthy, chaste, benevolent, virtuous	22 (7.4%)	1 (2.1%)
The attentive reader	patient, curious, diligent, inquisitive, incessant, interested, attentive	16 (5.4%)	NA
The dim or inattentive reader	simple, unlearned, indifferent, tired, injudicious, hapless, feeble sighted, overwearied	13 (4.3%)	3 (6.4%)
The general reader	every, general, ordinary, common	11 (3.7%)	4 (8.5%)
The intelligent reader	ingenious, intelligent, wise, enlightened	10 (3.3%)	6 (12.8%)
The emotive reader	susceptible, longing, melting, sympathizing, tender, feeling, impatient, sentimental, romantic	7 (2.3%)	4 (8.5%)
Total		277 (92%)	46 (98%)

cious reader," the "dim reader," the "general reader," the "intelligent reader," and the "emotive reader." It thus seems that in *WWiR*, where reviewers are hailing the readers of others' works, some of the most salient categories are those that place the reader in relation to the project of assessment. The "intelligent reader" and the "judicious reader" are in an implied partnership with the reviewer to distinguish between what is and is not worth public attention. The "dim" or "inattentive reader" solidifies this partnership by negation—where we posit good readers, we must also posit bad ones. The

"general reader" is invoked commonly in gestures of assumed consensus, not only when signalled adjectivally ("surely every reader must agree," "almost every reader will feel"), but also through an attribution of responses that achieve the same effect of consensus: "the reader must have been particularly struck," "the reader will perceive," "the reader will find his fancy amused," "the reader will probably suspect," "the reader will here recognise."

Where the "intelligent" and "judicious" readers are included in the project of assessment, the "sentimental," "feeling," and "romantic" readers invoked in the "emotive reader" category are situated outside the economy of critical judgment. The emotive reader responds not to the text but to the situation it describes, beguiled by feeling into losing critical distance. Sentimental readership results in misreading: "the sentimental reader will be disposed to think that the portrait of Leonora is too cold to be natural, and consequently the interest in her favor, which she really deserves, will be much lessened."[24] Here, the "sentimental reader" is guilty of nothing more than a kind of literary naïveté, but the stakes of that innocence are made clearer when the author is cast as a seducer whose cover is precisely this appeal to the reader's sympathies, as in *The Annual Review*'s account of Mary Robinson's *Poetical Works*. The reviewer accuses Robinson of "sentimental misrepresentations" through which she creates a self-justifying fiction that conceals a private character of doubtful virtue. These misrepresentations pose a specific risk to "juvenile ignorance," typified as a "tender-hearted young lady" who lacks the readerly skill to understand their fictiveness and is vulnerable precisely to their sentimental appeal.[25] Such unwary traversal of the space of representation (and misrepresentation) is a risk engendered by new literacies among "the herd of readers," and in particular female readers, who may be taken in by the story of Robinson's sensibility and misfortunes. The reviewer counters this appeal to sensibility by realigning it: through an extended narrative of female loss of virtue, "sensibility" is recast as a cover story for "that guilty, but much enduring class of women, who rashly bartering away the good opinion of the world . . . receive nothing in exchange but some vague and ineffectual claims on the gratitude, tenderness, or pity, of the most base, selfish, and profligate portion of mankind! Such a one was poor Mrs. Robinson." In other words, if as readers we fall into the trap of sympathy through "juvenile ignorance," we have the example of "that guilty . . . class" and Robinson's own troubled life before us as an object lesson in the perils of such thought.

The "emotive reader" thus emerges in a set of scenarios which rehearse the risks of readership, figured specifically as the risks of female readership and female authorship. In contexts where the author's own moral

standing is more sound, however, the sentimental reader's tendency to pass through the space of representation to a direct emotive response carries a different valence. The *British Critic's* review of West's *The Mother* approves of the "feeling reader":

> It might be expected that the Mother who wrote the Letters alluded to in the title page would have much to say . . . that must prove her truly worthy of that honourable name . . . but it could not have been expected of any writer, before trial, that she would do so ample justice to an arduous subject; and so often call forth the sympathy and even the tears of the attentive and feeling reader.[26]

Rather than weakening an intellectual defense mechanism that is considered morally necessary in the age of the novel, the sympathy and tears that accompany this reading instead evince a virtuous response to a text in which the gender of the author aligns with both religion and motherhood—a text against which no defense is needed.

These examples direct our attention to a multivalent problem of female readership and female authorship. The other reader characterizations (the "judicious," "intelligent," and "general" readers) situate the activity of readership within a distinctive morality that entails candor, education, judgment, and an appropriately scoped emotional connection between reader, text, and author. They represent reading as participation in a kind of social contract, an undertaking that strives to make a fair assessment through activities of attention, discernment, judgment, and taste. In other words, as reading becomes a private, personal activity, it is simultaneously being given higher public stakes; the success of reading as a private critical exercise is the index of its success in shoring up public morality (or, at least, resisting its decline). In this context, the language of "emotive readership" offers a space in which to rehearse anxieties about reading in an age of increasing literacy, where the female reader also stands in for "the herd of readers" and the broadening of literacy across class lines.

If we now turn back to our topical cluster of words that foreground readership, authorship, and reviewing (*author, character, effect, language, means, pleasure, public, read, reader, readers, story, style, writer*), what can we learn about the ramifications of gender within this topic more broadly? To start with, the characterizations of authorship are overwhelmingly associated with gender. "Fair" is the only adjective appearing with "author" in

the reviews corpus (33 instances), and the word "authoress" appears a further 86 times; together they account for 13% of all references to authors (out of 913 instances of "author" or "authoress"). While an emphasis on the author's gender is not perhaps surprising in a collection of reviews of women's writing, it reminds us that this marking served to make these reviews (and their subjects) visible within periodicals primarily devoted to reviews of male-authored texts. Related words strengthen this reminder: the word "female" (occurring 341 times) modifies a set of words that includes (in rank order) *character(s)* (25), *mind* (19), *writer(s)* (18), *sex* (15), *pen* (12), *readers* (5), *author(s)* (5). These writing-related words together account for 29% of the instances of "female." Here, gender and authorship are tightly connected and highly visible.

Another key word from this topic cluster—*character*—takes us further in unpacking how gender operates in the transactions of literate culture. "Character" is understood in these materials as the qualities that distinguish human beings, considered as historical, political, and social creatures. Thus, accurate character representation is also a test of authorship: the successful management of character, like the successful management of emotional investment, is a skill with high (and gendered) stakes in the age of expanding literacies. Reviewers disagree about whether this skill is within the scope of female authors. According to the *Quarterly Review*, women are "generally speaking, gifted with a nice perception of the various shades of character and manners," but the *Monthly Review* praises Hannah Adams's handling of character by distinguishing her from most authors of her gender:

> The greater number of female productions . . . are sadly wanting in proofs of good sense and the qualities of good writing; and the best of them . . . are deficient in that discernment into characters . . . found in similar productions of the other sex.[27]

Similarly, Maria Edgeworth is repeatedly praised in terms that align this kind of discernment with the appropriate handling of the domain of sentiment; one typical review writes that she is "furnished with every important requisite for the task she has undertaken—with a sound and cultivated understanding—constitutional freedom from the diseases of sentiment—a clear, discriminating perception of character."[28] The feminized domain of sentiment is firmly distinguished from the good sense and discrimination of character attributed to Edgeworth and Adams, regardless of the authors' gender.

Our cluster of focal terms—*reader, readers, author, character*—became visible through their shared membership in a topic of special intensity and ubiquity across the collection of reviews. As this exploration has shown, however, this word cluster is a sparse verbal symptom of a topic in a much deeper sense, one which could not have been taken for granted at the outset. The move from probabilistic collocation to deep cultural preoccupation is in this case made by recontextualizing the words that the topic model had at first treated as wholly independent and unordered quanta, and from these words-in-context deriving another kind of model of late eighteenth-century preoccupations with gender and literacy, one that operates at a cultural level where terms like "reception" and "readership" gain their purchase. This model helps to reveal how the reception of women's writing in late eighteenth-century periodical reviews staged an acute cultural preoccupation with the emergence of a publicly visible intellect on the part of women, as both readers and writers. All of the reviewer responses called out by this topic—consternation about the effects of women's writing on public morality, claims about the skills that are within the purview of female readers and writers, gendered attention to techniques of representation and readerly traversal thereof, and so forth—express the conviction that the presence of women within literate culture requires notice, response, and critical management.

.   .   .

Authorial gender is salient in eighteenth- and nineteenth-century reviews not simply in relation to the negativity or positivity of the review, but also as a pervasive lens for evaluation along many vectors. The gender of these authors is invoked equally in praise and in censure; authors are lauded for possessing both masculine qualities of discretion and feminine moral delicacy, and they are condemned for failing to adhere to the expected standards of femininity as well as for conforming to the poor writing habits ascribed to all women. Gender anxiety is woven into all aspects of these reviews—how quotations from the reviewed texts are selected and contextualized, how the authors' biographies are discussed, and most certainly in how authors' works are evaluated. Because this anxiety permeates these reviews so thoroughly, we have adopted methodologies that move between different contexts to get at the subtler textures of language usage on multiple levels. Combining topic modeling with close reading and with corpus-wide examination of patterns in word usage enables us to take advantage of the ways that topic modeling can recontextualize individual terms, and identify new approaches to inter-

rogating the collection as a whole. We have found that, despite the changing fortunes of topic modeling in the digital humanities, there are ways of using this technique that do not mistake the model-generated "topics" for topicality, but instead treat topic modeling as another mechanism for understanding language usage, a mechanism that can be fruitfully combined with other methods of inquiry that return us to the text.[29]

Extending these approaches to other collections could yield interesting results, although careful consideration needs to be given to the nature of the texts under scrutiny and the representation of the underlying data. For instance, topic modeling in our analysis made visible the vocabulary cluster associated with the reviewing enterprise, but the thematic salience of this cluster was strongly reinforced by the generic homogeneity of the reviews corpus. Topic modeling is also well suited to shorter documents of fairly uniform length, such as the *WWiR* collection. Moreover, the presence in this collection of structural markup (for instance, to mark quoted material) and detailed metadata (to differentiate between genres of review materials) made possible a more nuanced analysis including a comparison of the language of reviewers with that of the text under review. A resource like the *Early English Books Online–Text Creation Partnership* digital collection also includes structural markup and metadata that would support this kind of comparative analysis, but texts would need to be carefully selected for length in order for topic modeling to yield useful results.

Nevertheless, all of the analyses carried out in this essay are incomplete: they demonstrate approaches that could yield fuller insight if pursued with larger sets of texts, more words, more iterations. By their nature they are not perfectible, since the information they use must always be a limited proxy for the "culture" we are trying to discover and understand. However, there are several specific directions for further work that seem especially promising in bringing us closer to that understanding. Comparison of the *WWiR* collection with reviews of male-authored works would be illuminating, to see what topics emerge there and to explore how the language of *reader*, *author*, *character*, and associated descriptors operates in reference to male authors and their readership. With a larger collection of reviews, we could experiment with word vector analysis to see whether the word relationships foregrounded in the topic models are discernible in the different algorithmic space of word embedding models. The *Women Writers Project* is now in the early stages of a project to develop parallel corpora of male- and female-authored texts to support comparative text analysis, starting with collections to parallel *WWO* but extensible to *WWiR* as well. These corpora

will be exposed for exploration via the *WWP Lab*, where we have published a prototype interface to support the use of word embedding models on these corpora.[30] Moreover, as it becomes increasingly possible to aggregate data sets like these using technologies of linked open data (whereby multiple individual collections can be queried and virtually aggregated), the boundaries of individual collections will be easier to traverse, even though the realities of funding may still dictate that specific research projects focus on distinct timespans, genres, genders, and countries. The internet itself—essential to the forms of reception and circulation that operate for modern readers—is a network of separate fragments between which linked open data seeks to build navigable connections.

ॐ

Notes

1   For example, the *Early Modern Print* project has developed tools (including an n-gram browser and keyword-in-context explorer) through which readers can compare terms of interest across the *Early English Books Online–Text Creation Partnership* collection; see *Early Modern Print*, Humanities Digital Workshop, Washington University in St. Louis, at earlyprint.wustl.edu; and *Text Creation Partnership*, at textcreationpartnership.org/tcp-eebo. Scholars such as Ryan Heuser and Cameron Blevins have employed natural language processing models to examine conceptual relationships in *Eighteenth Century Collections Online* and to study patterns in "recurrent themes" across the 10,000 entries in an early American midwife's diary; see see Heuser's "Word Vectors in the Eighteenth Century," at ryanheuser.org/word-vectors; and Blevins's "Topic Modeling Martha Ballard's Diary," Apr. 1, 2010, at cameronblevins.org/posts/topic-modeling-martha-ballards-diary.

2   See *Women Writers Online* (*WWO*), at the *Women Writers Project* (*WWP*), Northeastern University, wwp.northeastern.edu/wwo. A license is required for access, but free trials are available upon request.

3   For an overview of the "Cultures of Reception" research project, see *WWP*, at wwp. northeastern.edu/research/projects/reception/index.html; and see *Women Writers in Review* (*WWiR*), Northeastern University, at www.wwp.northeastern.edu/review.

4   The "Cultures of Reception" project, the *Intertextual Networks* project, and the development of the *WWP* itself were made possible with generous funding from the National Endowment for the Humanities, the Andrew W. Mellon Foundation, and the Gladys Krieble Delmas Foundation.

5   See *Guidelines for Electronic Text Encoding and Interchange*, at *Text Encoding Initiative*, www.tei-c.org/P5.

6   See *Women Writers in Context*, Northeastern University, at wwp.northeastern.edu/context.

7   Unknown reviewer, *The Monthly Review* 13 (1755): 399; qtd. in Antonia Forster, *Index*

to Book Reviews in England, 1749–1774 (Carbondale: Southern Illinois University Press, 1990), 5. While nearly all the reviews in the *WWiR* collection are anonymous, we hope in the future to use stylometry, in combination with bibliographic and historical research, to identify and analyze the works of individual reviewers.

8   Unknown reviewer, "1812–10: *The Eclectic Review* on Edgeworth's *Tales of Fashionable Life*," *WWiR*, Nov. 16, 2016, at wwp.northeastern.edu/review/reviews/edgeworth .fashionlife.eclecticrev.2.

9   Many excellent explanations of topic modeling already exist, for example Ted Underwood's introduction "Topic Modeling Made Just Simple Enough," *The Stone and the Shell*, Apr. 7, 2012, at tedunderwood.com/2012/04/07/topic-modeling-made-just-simple-enough; and Scott Weingart's "Topic Models for Humanists: A Guided Tour," July 25, 2012, at scottbot.net/HIAL/index.html@p=19113.html. The models described here were created using the R implementation of the MALLET package developed at the University of Massachusetts, Amherst.

10 Of course, many other methods for subsetting the collection are possible and could form the basis for future investigation. For example, we chose to include all of the *WWiR* items in our full-collection corpus because the advertisements and other non-review materials in *WWiR* were themselves a significant aspect of transatlantic periodical culture and because we were invested in exploring the collection as a whole. Nevertheless, another useful corpus for analysis might be constructed by selecting only those items identified in their metadata as literary or theatrical reviews.

11 For example, an 1826 review of the *Works of Maria Edgeworth* in the *Boston Monthly Magazine* notes, "We should have gone on to say something more of this admirable author if we had not, by chance, cast our eyes upon a well written criticism upon Miss Edgeworth's powers and productions, published in the North American, for January, 1818, on reviewing 'Harrington, a Tale, and Ormond, a Tale' from her pen. This criticism is worthy resuscitation for its truth and elegance," after which follows a lengthy extract from the earlier review. See unknown reviewer, "1826–03: *Boston Monthly Magazine* on Edgeworth's *Works of Maria Edgeworth*," *WWiR*, Nov. 16, 2016, at wwp.neu.edu/review/reviews/edgeworth.works.bostonmonthly magazine.1.

12 The reviews and quotations corpora are mutually exclusive, but they are not fully parallel in the texts they include since not all reviews have quotations and the quotations corpus also contains published extracts which account for 34 texts and 6,000 words in that corpus.

13 All three corpora are available for experimentation at *WWP*, wwp.northeastern.edu /research/publications/datasets/recirc2017.

14 Briefly, we determined that 25 topics and 800 iterations provided the best results for our analyses and we used hyperparameter optimization in training our models.

15 To avoid reducing our topics to the top ten words from a single model, we have aggregated results from multiple models. We trained four models on each of the three corpora and extracted each topic's top fifty words (the threshold at which our tests showed that we were focusing on the most clearly prominent terms for each topic). Note that all terms are lowercased during the corpus preparation process. In some cases, as with the review and authorship terms, it was clear that we were

seeing the "same" topic from one model, and even one corpus, to another, because many of the terms were the same, as were the patterns in topic proportions across the collection. Combining the top fifty terms from various versions of models enabled us to determine which were appearing consistently in each model, regardless of the variability inherent in topic modeling.

16   The form "er" is from constructions such as "e'er" and "when'er," which the tokenizer we used did not recognize; and the term "st" is from constructions such as "would'st" and "should'st." For the purposes of this discussion, we have focused on a few topics, but these are far from exhaustive; for example, there is another poetry topic comprising language that is not just poetic, but even romantic: *blood*, *bright, dark, deep, earth, free, hath, pass, sight, sleep, voice, wave.*

17   These patterns can be seen in clearer detail in visualizations included in the supplementary materials for this article at *WWP*, wwp.northeastern.edu/research /publications/datasets/recirc2017. These visualizations also make it possible to compare proportions for all the topics discussed here across the documents in each corpus.

18   By contrast, there are only ten instances of "dear" in the reviews corpus.

19   It is also unsurprising, given the overall high proportions of this topic across the *WWiR* collection, that its terms appear quite frequently within the corpus. For example, among the words listed above for the full *WWiR* corpus's review and authorship topic, "character" appears 649 times and "author" appears 569 times; most terms in this set appear two to three hundred times. For a full count (with stopwords included) of terms in the collection, see *WWP*, wwp.northeastern.edu/research /publications/datasets/recirc2017.

20   For more on "the complex relationship between professional authors and their audiences" in periodical writing, see Manushag N. Powell's *Performing Authorship in Eighteenth-Century English Periodicals* (Lewisburg, Pa.: Bucknell University Press, 2012), 4.

21   Although the analysis of associated phrases could have been done automatically using natural language processing techniques, that approach exceeded our expertise; given the manageable number of instances, hand analysis was feasible and provided for greater immersion in the data.

22   All but one instance of "my readers" in the *WWiR* collection is from quoted material.

23   Because the *WWiR* collection is so much smaller than the *WWO* collection and because of the low frequencies of all these terms, the differences in frequency of each category are most useful as qualitative characterizations rather than as numbers for precise comparison.

24   Unknown reviewer, "1807–08: *The Monthly Review* on Edgeworth's *Leonora*," *WWiR*, Nov. 16, 2016, wwp.northeastern.edu/review/reviews/edgeworth.leonora .monthlyreview.1.

25   Unknown reviewer, "1806: *The Annual Review* on Robinson's *The Poetical Works of the Late Mrs. Mary Robinson*," *WWiR*, Nov. 16, 2016, wwp.northeastern.edu/review /reviews/robinson.robinsonpoetical.annualreview.1.

26   Unknown reviewer, "1809–06: *The British Critic* on West's *The Mother*," *WWiR*, Nov. 16, 2016, wwp.northeastern.edu/review/reviews/west.mother.britishcritic.1.

27 Unknown reviewer, "1814–01: *The Quarterly Review* on Edgeworth's *Patronage*," *WWiR*, Nov. 16, 2016, wwp.northeastern.edu/review/reviews/edgeworth.patronage.quarterlyrev.1; unknown reviewer, "1799–12: *The Monthly Magazine* on Adams's *A Summary History of New England*," *WWiR*, Nov. 16, 2016, wwp.northeastern.edu/review/reviews/adams.histnewengland.monthlymagreview.1.

28 Unknown reviewer, "1812–10: *The Eclectic Review* on Edgeworth's *Tales of Fashionable Life*," *WWiR*, Nov. 16, 2016, wwp.northeastern.edu/review/reviews/edgeworth.fashionlife.eclecticrev.2.

29 See, for example, Benjamin Schmidt, "Words Alone: Dismantling Topic Models in the Humanities," *Journal of the Digital Humanities* 2, no. 1 (2012), at journalofdigitalhumanities.org/2-1/words-alone-by-benjamin-m-schmidt.

30 See the *WWP Lab* at *WWP*, wwp.northeastern.edu/lab; and the *Women Writers Vector Toolkit*, at *WWP*, lab.wwp.northeastern.edu/wwvt/.