

Search

- [Issues](#)
- [Contemporaries](#)
- [Book Series](#)
- [Conferences](#)

[How Cultural Capital Works: Prizewinning Novels, Bestsellers, and the Time of Reading](#)

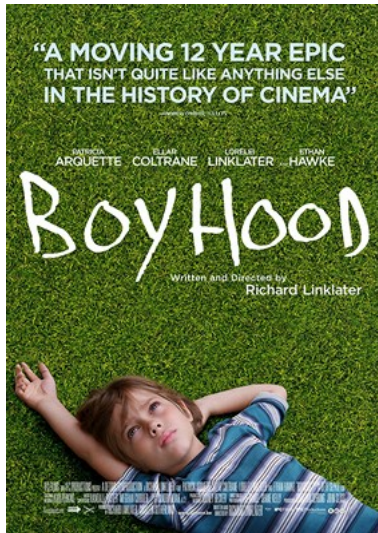
Andrew Piper and Eva Portelance / 05.10.16

There came an acute awareness of time.

— André Alexis, *Fifteen Dogs*, Winner of the Giller Prize 2015

When the movie *Boyhood* appeared, it was widely regarded as a cinematic breakthrough. Time, that great object of modernist film criticism, was no longer simulated through make-up and montage, but rendered on the screen as something that could directly be observed. After *Boyhood*, the "long film" didn't signal an epic of run-time — as in Fassbinder's 15-hour *Berlin Alexanderplatz* — but a sustained capturing of time itself, an epic of filmed-time.

And yet there was also something deeply familiar about *Boyhood*. The nostalgic, masculine point of view, the look back to recuperate boyish things past — these moves felt very much of their moment, our moment; they are potent signs of the way "important" cultural objects continue to distinguish themselves today. The movie's iconic film poster of the boy lying in the grass daydreaming seemed to capture something like a high-cultural semiotic bull's-eye — if not more broadly within contemporary life, then certainly within the literary field, which will be our focus here. The film's widespread appeal was at least partially attributable to what we hope to show is a pre-existing horizon of expectation, one in which the tropes and language of nostalgia have become a key feature of what is deemed culturally significant today. The long-forgotten modernist mantra of "make it new" appears increasingly, if not entirely, to have given way to a greater high-cultural sensibility of retrospection.



The poster from *Boyhood* (2014).

Few genres have come to stand as icons of literary taste more than prizewinning novels. However ambivalent we may feel about literary prizes — and, as James English argues, ambivalence is built into the prize experience itself — the literary prize is designed above all else to distinguish.¹ It not only draws a line that cordons off a piece of writing through the act of institutional validation. It also reinserts this work back into the economic marketplace in a more central location. By endowing the literary work with a newfound cultural capital, the prize centralizes the work within the economic marketplace.² It allows novels to coexist with other works that only ever claimed economic value for themselves — the various genres that make up "bestsellers." The prize is what allows for two very different types of writing to simultaneously occupy a central place within the same economic space. While the two can overlap (prizewinners can be bestsellers and vice versa), the point of the prize is to help us maintain social distinctions within an otherwise monolithic marketplace.

This essay is about the relationship between prizewinning novels and their economic counterparts, bestsellers. It is about the ways in which social distinction is symbolically manifested within the contemporary novel and how we read social difference through language. As the more recent work of English has shown, not only can we observe very strong differences between bestselling and prizewinning writing, but the process of cultural distinction also appears to revolve most strongly around the question of time.³ The interlocking nexus of nature, childhood, and retrospection that so marked the film *Boyhood* does indeed appear to mark out the high cultural work of prizewinning novels, just as the diligent attention to the present on the part of bestselling novels comes into sharper relief through the lens of prizewinners. It is these temporal frameworks, or what Bakhtin might have called "chronotopes," that emerge as some of the more meaningful ways to distinguish the work of cultural capital from that of economic capital.

The approach we will be using here draws on the emerging field of textual analytics within the framework of Bourdieu's theory of the literary field.⁴ Our interest lies in exploring not simply a larger population of works than a select handful of novels, but also the ways in which groups of works help to mutually define one another through their differences.⁵ As Bourdieu writes, "Only at the level of the field of positions is it possible to grasp both the generic interests associated with the fact of taking part in the game and the specific interests attached to different positions."⁶ Similar to English, we are interested in understanding the ways in which "bestsellers" and "prizewinners" cohere as categories and the extent to which this coherence is based on meaningful, and meaningfully distinguishing, textual features. Unlike English and his team, who work with a long historical view of these two categories (including 50 years' worth of data), we focus instead on a more synchronic portrait of different kinds of socially valued writing from the last decade. And whereas English and his team hand-labeled works according to a three-part scheme of temporal setting

(historical, contemporary, futuristic), we use machine learning and statistical modeling to better understand the linguistic differences of these categories, especially as they relate to the category of time. Rather than rely on the static construct of the novel's "setting" as the analytic lens, we try to understand the temporal patterns that run through these works.⁷ How might the lexical striations of time within novels help differentiate between these types of writing? Do we indeed see connections between processes of social selection — a committee designating something a prizewinner, large numbers of readers buying a book — and the intrinsic qualities of the works selected by these groups? If so, what are the stylistic and symbolic norms that govern distinctions within the literary field? How is something like Jauss's notion of a "horizon of expectation" a powerful agent in the creative process, helping guide the process of creation and selection of individual works?

English's work has identified profound trends surrounding the shift of narrative setting in popular versus serious fiction.⁸ Our hope is that computational text analysis can add nuance to our understanding of this process of differentiation and provide us with important information about the linguistic foundations of these social processes. At its best, it can aid us in identifying the values and ideological investments that accrue around different cultural categories — the act of "position taking," in Bourdieu's words — and the ways in which these horizons of expectation help maintain positions of power and social hierarchy. Our project is ultimately about asking how forms of social and symbolic distinction correspond and how that knowledge may be used to critique normative assumptions about what counts as significant within the literary field. The value of computational analysis for us is the way it enables a position of critique with respect to the texts and fields under review, the kind of critical distant reading that critics have repeatedly been calling for.

The collection of novels we will be using represents a data set curated by our lab that consists of roughly 1,200 novels in English that have been published within the last decade.⁹ Instead of taking a longer historical view, we try to understand better the contemporary field of literary production and the extent to which it can be understood in Bourdieu's terms as a field. "The science of the literary field," writes Bourdieu, "is a form of analysis situs which establishes that each position...is subjectively defined by the system of distinctive properties by which it can be situated relative to other positions."¹⁰ The novels are broken down by six different categories, which we briefly summarize in Table 1 and describe below. Overall, the collection has been designed to compare socially distinctive writing (what we might call "literary fiction") with socially popular writing ("bestselling fiction"). We define popularity in two different ways, both according to a pure economic categorization (weeks on the bestseller list) as well as a more diverse array of genres based on different types of economically successful writing (what we might call "genre fiction"). Taken together, these groups have been chosen with an eye to understanding the social filters, whether market-based, journalistic, institutional, or even generic, that construct symbolic fields of meaning.

Table 1. Breakdown of the different collections of novels by category. Best-sellers tend on average to be the longest genre, while romances are the shortest. Prizewinners have the highest vocabulary richness (ratio of unique words to total words), while romances have the lowest.

	Prize	Bestsellers	NYTimes	SciFi	Mystery	Romance	All
Authors	194	100	191	162	150	175	972
Works	216	200	193	200	200	200	1,209
Words (Unique)	215,969	198,859	185,399	204,850	161,160	105,567	579,179
Words (Total)	21,997,136	24,931,580	19,095,475	23,090,002	20,132,621	16,432,668	125,679,482
Mean Length	101,800	124,700	98,940	114,300	100,700	82,160	103,767

The prizewinner category consists of works shortlisted for any of five different prizes in three different English-speaking countries over the past decade (the National Book Award (US), PEN/Faulkner Award (US), Governor General's Award (Canada), Giller Prize (Canada), and the Man Booker Prize (UK)). Instead of focusing on a single prize, we try to understand how prizewinners may or may not cohere even across different national contexts (thinking about cross-linguistic similarities would be a challenging next step).

The bestseller category consists of novels published since 2001 that have spent the longest aggregate amount of time on the New York Times bestseller list, with a minimum threshold of at least ten weeks and maximum of over 200 weeks. These are works that have achieved a stunningly high degree of cultural saturation. As we can see in the table, their list is also disproportionately occupied by a much smaller range of authors. In delightfully tautological fashion, being a bestselling-author is one of the best predictors of being a bestselling-author. The New York Times set consists of novels that were reviewed in the New York Times in order to capture a collection of novels that have passed through a high cultural filter, though not the very highest. If only a tiny fraction of novels ever wins prizes, only a slightly larger fraction ever gets reviewed in a major news outlet. These novels thus represent a second tier of cultural awareness. Finally, we constructed three sets of "genre fiction" that were drawn from "bestselling" novels according to rankings on Amazon.com.

In our essay, we will be moving through three distinct steps: first, we show the degree to which these groups do or do not differ from one another. To what extent do categories of social distinction like bestsellers and prizewinners behave like genres such as romance, mystery, and science fiction — and to what extent might we want to consider them to be genres in their own right? How does contemporary fiction differentiate itself when seen from a computational perspective? Second, we explore the extent to which bestsellers and prizewinners separate on the question of time. As we show, temporality appears to be one of the strongest and most meaningful ways that these two categories differ from one another, lending further support to the work of English and his team. In the conclusion of our essay, we then offer a critique of this "nostalgic turn" within literary prizewinners. Rather than search for underlying causes — whether the move away from contemporaneity in the field of serious literature is the work of "late capitalism" (Jameson) or a handful of social actors (English) — our interest lies in bringing to light the semantic constriction that such a nostalgic stance implies. The bias towards a particular type of temporal framework is not in our view value neutral. Computation can help us see the linguistic underpinnings of these positions and critique the unstated norms associated with the "significant," "important," or "serious" within cultural production today.

Representing Genre Space

We begin with a series of visualizations of the relationships between novels from our different collections (figs. 1-4). These relationships are based on eighty different features drawn from the Linguistic Inquiry and Word Count Software (LIWC), a popular text analytics package (which is of course not without its critics).¹¹ Using LIWC as a basic first step, we gain a sense of how similar the novels are to one another across a variety of stylistic features, including grammar, punctuation, and a series of what LIWC calls psychological and social processes (expressions of cognitive insight or hesitation, or a focus on bodies or sensory perception). LIWC offers a simple-to-use tool that can give us a first look at the quantitative differences between texts that are based on a robust variety of relevant literary categories.

As we can see from the different figures, romance not surprisingly looks the most distinctive when compared to other types of writing, especially science fiction (fig. 1). Prizewinners overlap most with novels reviewed in the New York Times, an overlap suggesting a great deal of similarity between the filtering mechanisms of major media outlets and prize selection committees (fig. 2). It is important to emphasize here that we are not showing causality — if being reviewed in the New York Times determines whether you end up on a literary prize short list — but rather the ways in which novels that get chosen by these different selection mechanisms appear to have a great deal in common. Indeed, as we have shown elsewhere, it is incredibly difficult to determine what actually makes a novel move from the review pile to the short list; this difficulty suggests not only many extrinsic as well as unknown factors at work, but also potentially a great deal of chance.¹² Finally, prizewinners and bestsellers appear to exhibit a relatively strong degree of polarity (fig. 3), though not nearly as strong as the differences within genre writing. In other words, a first glance suggests that genre is a much stronger organizing force than social distinction, but that social distinction, understood as two different forms of valuation, appears to exhibit stylistically different norms.

Figs. 1-4. These graphs represent the similarity between novels in our collection based on the eighty different features contained in the Linguistic Inquiry and Word Count Software (LIWC). Multi-dimensional scaling was used to render the coordinates in two-dimensional space.



Fig. 1. Comparing romances to science fiction, the most distinctive two genres in our collection.



Fig. 2. Prizewinners compared to novels reviewed in the New York Times. The high degree of similarity suggests that these are being drawn from similar populations.

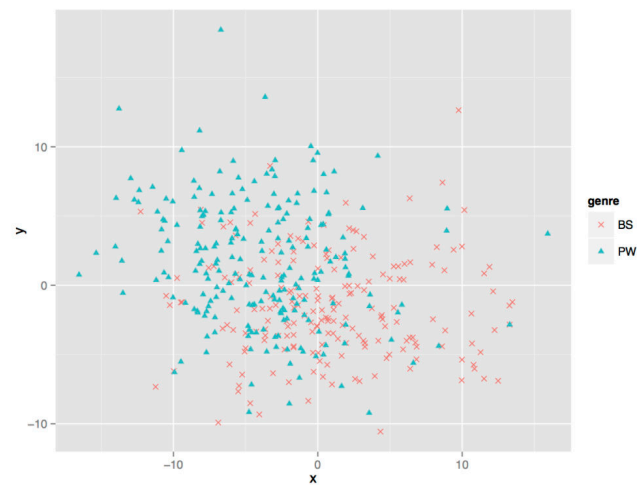


Fig. 3. Comparing prizewinners to bestsellers. While not as strong as the generic differences between romances and science fiction, there is still a relatively strong degree of polarity. As we show below, we can correctly predict prizewinners with about 81% accuracy when compared to bestsellers.

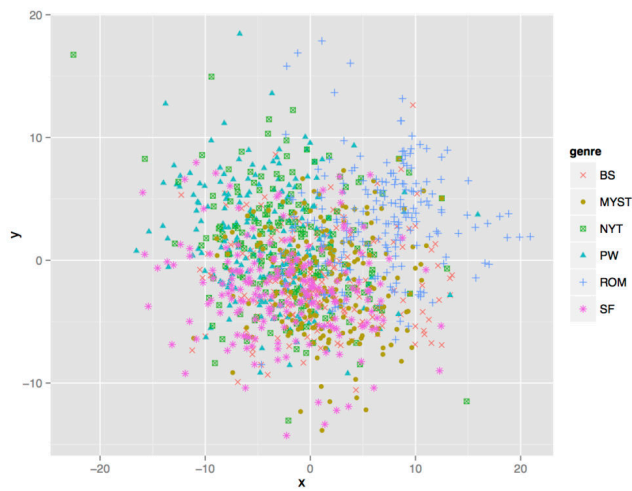


Fig. 4. This represents the entire genre space of our data set. We find the overall centrality of the mystery genre of potential interest.

To further test these relationships, we used the process of machine learning to build a classifier whose aim is to predict which category a given novel belongs to.¹³ Based on novels the learning algorithm has seen, how well can it properly predict novels it has not seen? As Table 2 shows, the strongest distinctions exist between prizewinning novels and romances. Just about 99 times out of 100, the computer can correctly predict whether a novel is either a romance or has been shortlisted for a literary prize, which suggests that if you want to win a literary prize, writing a romance, or anything resembling one, is a very bad idea.

As we descend the list, we see how the genre groups overall are more easily classifiable than the social value groups; this tendency confirms the insights from the clustering graphs. Nevertheless, with just one exception, the social value classes do exhibit coherence as classes, with the strongest difference being exhibited between bestsellers and prizewinners, which can be differentiated with a relatively high degree of accuracy (about 81%). Interestingly, although the social value groups may contain multiple genres — and they do indeed contain some works from our three categories (science fiction, romance, mysteries) — they nonetheless exhibit strong predictability when compared either to the genre sets or other classes of social value (Table 3). The exception is the relationship between prizewinners and novels reviewed in the New York Times, where the classifier's performance is basically just above chance. Indeed, the classifier consistently over-predicted New York Times novels to be prizewinners; perhaps it has covert feelings about the unfair nature of the selection process.

Table 2. Pairwise comparisons of our different novel sets using an SVM classifier. The F1 score is the balanced accuracy of how well group A was predicted in comparison to group B. Percentages rounded to the nearest tenth.

Group A	Group B	Accuracy (F1) in %
PW	ROM	98.8%
NYT	ROM	97.7%
ROM	SF	97.5%
MYST	ROM	94.5%
BS	ROM	93.5%
MYST	SF	90%
MYST	PW	89%
PW	SF	89%
NYT	SF	87.2%
MYST	NYT	84.7%
BS	SF	84.5%
BS	PW	81.8%
BS	MYST	76.5%
BS	NYT	76.4%
NYT	PW	56.9%

Table 3. Summary of the average predictive accuracy of a given class based on all pairwise classification scores.

Genre	Avg. Accuracy
Romance	96.49%
SciFi	89.47%

Genre	Avg. Accuracy
Mystery	87.06%
Prizewinners	83.16%
Bestsellers	82.54%
NY Times	80.45%

Natural Time v. Clock Time

Given that we can predict distinctions between different types of social valuation with a relatively high degree of success, what are the literary features that might underlie these categorizations? Our first step was to run a test to discover the most distinctive words for each corpus and then assess these words for the presence of more unified categories.¹⁴ Essentially what we are asking is: what words appear with greater consistency in one group over another and, second, in what ways might these words semantically cohere? While we do not present the entire output here, Table 4 lists three groups of words that are statistically unique to one set over another, their frequency within the more dominant collection, and the relative increase of their usage over the comparison corpus. This allows us to see how important a word is both to a given corpus and its relative over-usage when compared with another corpus. While there are different ways that different readers might interpret the linguistic differences between these groups, we have identified what we see as three highly connected fields, which we label nature, childhood, and time.

Table 4: The Language of Prizewinners

Nature (PW)			Childhood (PW)			Time (PW)			Time (BS)		
Term	Freq	Increase	Term	Freq	Increase	Term	Freq	Increase	Term	Freq	Increase
Village	1.66	180%	Childhood	0.64	99%	Winter	1.44	108%	Tonight	1.74	94%
Birds	1.62	158%	Boy	6.95	69%	Youth	0.58	107%	Minute	5.86	82%
Dust	1.08	125%	Father	10.94	68%	Afterward	0.72	76%	Second	5.12	54%
Sea	1.76	122%	Child	3.4	66%	Whenever	0.82	73%	Tomorrow	1.44	54%
Horses	2.5	109%	Children	4.09	61%	Sometime	3.94	53%	Immediately	1.57	48%
Winter	1.44	108%	Dream	2.52	61%	Old	9.99	49%	Yesterday	0.76	48%
Forest	0.77	108%	Imagine	3.29	50%	Born	1.07	48%	Fast	1.52	42%
River	2.22	100%	Mother	11.33	46%	Spring	0.97	44%	Today	2.22	35%
Grass	1.24	97%	Become	2.77	44%	Young	4.18	41%	Quick	3.6	35%
Earth	1.55	85%	Remember	6.36	26.5%	Older	1.41	39%	Hour	5.26	29%
Trees	3.86	83%				Once	7.19	37%	Busy	3.3	27%
Mountain	1.2	77%				Summer	1.79	25%	Next	6.5	13%

Table 4. A list of the most distinctive words of prizewinning novels when compared with bestsellers using the Wilcoxon rank sum test. "Freq" represents the average number of occurrences per 10,000 words while "Increase" represents the percentage increase of that term in prizewinning novels versus bestsellers. Together, these statistics allow us to assess how common a word is in a given collection and also how much more prevalent it is in one collection over another. Only words that appeared in more than half of all documents and at least 1,000 times were considered. All words listed have p-values below 0.0001. The full list is available at <https://dx.doi.org/10.6084/m9.figshare.2061996>.

How to understand these lists? Take for example the word "father" (which includes the word "fathers" as well). This word form appears 10.94 times per 10,000 words on average or just over 100 times per novel (if we assume an average novel length of about 100,000 words). The increase of 68% means that the average for bestsellers is 65 times per novel, so that fathers appear about 40-50 more times per novel in the prizewinning set, or roughly 1x more every 4-5 pages (if we assume a novel to be about 200 pages).

These individual differences by themselves may not feel all that dramatic. And certainly the effect for less frequent words is considerably smaller ("dust," for example, appears only about 10x per novel). But when we begin to look at these words in aggregate, when we think about them as a semantic field that suffuses a given novel or collection of novels, their disproportionate presence in one collection versus another starts to feel significant — not just statistically (though we'll return to statistics later), but in the more literal sense of signifying something. And when we begin to look at them in relation to another group of novels, that significance becomes even more pronounced.

According to these lists, the temporal world of prizewinning novels looks more seasonal, but also dichotomous, divided by a strong sense of a before and after, of old and young. "Strange" is interestingly one of the more distinctive words of prizewinning novels. Nature's cyclicalities serves as a kind of backdrop to reflect on the polarizing aspects of past and present within human consciousness. The estrangement from what was compels a desire to look back. Time in bestsellers, on the other hand, appears to be deeply urgent: here it is the time of minutes, seconds, hours, where clock time reigns. Time's framework is bounded by a single day: tonight, today, yesterday, tomorrow. Things happen immediately, fast, quick, and people are in a hurry or are busy. This is the world of the next.

If we look more closely at some of our novels, it is clear that there are numerous ways that prizewinning novels abound in moments of nostalgia. In Philip Roth's *Everyman* (2007), winner of the PEN/Faulkner Award, we see how the protagonist begins to think about his daughter and her childhood as he awaits another operation; his reminiscences themselves are a routine and yet essential mental operation, much like the medical one he is about to undergo. David Chariandy's *Soucouyant* (2007), shortlisted for the Governor General's Award, tells the story of a young man returning to take care of his mother, who is now suffering from Alzheimer's. He remembers the first indications of his mother's illness when he was a child, using her weakness as a way of satisfying his needs by sneaking "peanut butter and corn syrup, lime pickle and molasses" from the cupboards. "We were never caught," he says in a very sad twist on childhood inventiveness.¹⁵ In Johanna Skibsrud's *The Sentimentalists* (2010),

winner of the Canadian Giller Prize, we learn of the story of a daughter who returns home to retrieve her father, who is too old to live alone. In addition to the familiar scenes of looking at old family photographs, this retrospective framework infects even the present of her past. While driving her father away from their family home in the narrative's present, she remembers a family road trip from her childhood, a road trip in which she imagines what that moment will feel like when remembered in the future.

In reminder to myself that although it felt like it might there was no way that a two-day trip could last forever, I imagined the different ways I might recount its events even as they occurred. When my father, for example, shifted in his seat and Helen bumped into my hand on the stick as I drove, when he interrupted our long deliberate silences by saying, "Tell me stories, my sweethearts," I would silently address an invisible, future audience...[16](#)

The past that is recounted here is one in which the present is already retrospectively coded. The narrator is remembering how she experienced life ready to be remembered: "I imagined the different ways I might recount [these] events as they occurred." She remembers how her father asked her to tell him stories, as she imagines a future in which she recounts this past, just as she is doing now in the car ride with her father, a father who can himself no longer remember. Nostalgia seeps into the very crevices of past and present alike, as we become ever more fearful of retrospection's extinction through the process of aging.

The question that emerged for us was not so much whether we could find particular instances of nostalgia among prizewinners, but whether nostalgic was a good way to categorize the group overall when compared with bestselling writing. Is it fair to say that prizewinning novels are significantly more defined by a nostalgic temporal mentality than bestselling writing, or indeed, any other kind of writing? And if so, what are the implications of such a stance? What does this say about the process of social valuation that helps construct this category or the field of readerly identification that is framed by it?

To test this, our first step was to construct a dictionary that we felt represented a nostalgic state drawn from our list of most distinctive words.[17](#) To what extent did words that potentially indicated a nostalgic stance suffuse one type of writing versus another?

We found that prizewinners used these words in aggregate to a significantly greater degree than all other types of writing in our collection, even an additional collection of nineteenth-century realist fiction (Fig. 5). On average, prizewinners use this vocabulary about 50% more often than any of the other market-driven novels. Just under 10% of the bestsellers use an amount equal to or more than the average frequency of nostalgia words in prizewinners. And while the New York Times novels were the closest in frequency to the prizewinners, even in this case there was statistical significance in the differential usage of this vocabulary.[18](#) As we move up the cultural scale, we see an increase of nostalgic narrativity at work, with prizewinners representing something like a high-cultural apex.[19](#)

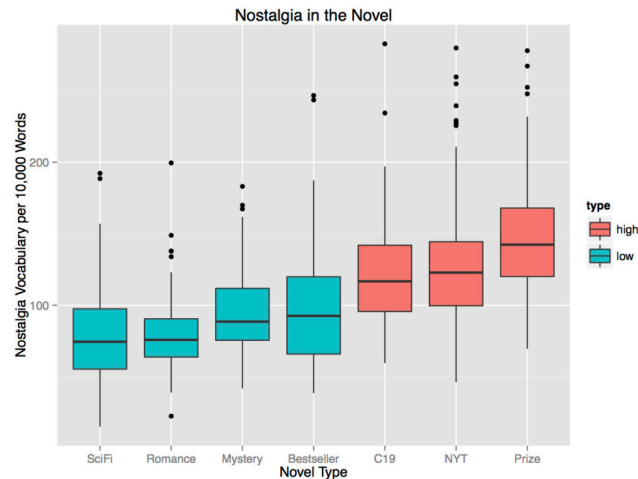


Fig. 5. Boxplot of the average nostalgia scores for each of our novel collections. "High" and "low" refer to cultural status. The black horizontal bands represent the median rate of nostalgic vocabulary while the boxes capture the values between the first and third quartiles (i.e. the range of half of the values). Here we see that the top three-fourths of prizewinners have more nostalgic vocabulary than the median rate for New York Times novels.

It is worth pausing here and asking whether these features, while different, represent a particularly strong signal. To what extent can we say that prizewinning novels cohere around the language and tropes of nostalgia and retrospection? Perhaps there are other features that are more important to prizewinning novels than this one. We found that prizewinners, taken in aggregate, use these words about 1.5% of the time, which amounts to roughly 6-8 times per page. In order to better understand the relative strength of this presence, we compared these words to the suite of LIWC dictionaries mentioned above. Of the forty-six non-grammatical dictionaries contained in LIWC, our nostalgia words, at 1.5%, ranked close to the middle, in twentieth place. However, we need to keep in mind that our dictionary is considerably smaller than all of the LIWC dictionaries. With just 28 word types, for example, our nostalgic vocabulary is almost as present in prizewinning novels as negative emotions are with five hundred word types (1.76%). At the same time, no other dictionary represented as extreme an increase between prizewinners and bestsellers than ours, other than "colons" and "semi-colons." In other words, besides two rare forms of punctuation, the nostalgia lexicon captured the largest increase between the collections out of all eighty dictionaries in LIWC, again with just these 28 word types. Finally, when we combined the nostalgia vocabulary with the time words indicative of bestsellers, we were able to predict prizewinning novels with 83.7% accuracy, a slight performance increase over using all LIWC features, which suggests that just this single feature does as good a job of identifying prizewinners as all 80 features combined.

As with any [lexicon-based approach](#), however, it is possible that our dictionary might be capturing something more or less than nostalgia (so-called false positives). Is the more frequent appearance of these words truly indicative of "nostalgia," or perhaps something else entirely, or maybe just semantic noise? As a second step, we used the process of machine learning to try to predict the number of nostalgic passages per novel in our collection. Here, [we identified 300 nostalgic passages by hand, where "passage" was defined as a unit of 1,000 words of text and "nostalgic" was defined as "any scene that includes a retrospective looking back at one's own life, ideally in a longing kind of way."](#)[20](#)

The passages were drawn equally from all six of our genres and thus represent something like a non-genre-specific discourse of nostalgia.

Using these passages, we then trained our classifier in two different methods, one consisting of the LIWC dictionaries and one using Bag of Words (BoW), which means we analyze all words in a passage. Employing the same [10-fold process of cross-validation](#), we were able to predict nostalgic passages with greater than 90% accuracy.[21](#) As we can see in Fig. 6, once again we find a significantly higher number of nostalgic passages within prizewinning novels. The strength of the difference between the prizewinners and New York Times novels appears to decrease in this model (though the upper bound is still considerably higher in prizewinners). It suggests that the differences between these collections are not necessarily occurring at the structural level of "retrospection" but more at the level of specific words—a greater emphasis on the specific semantics of childhood, nature, and memory.

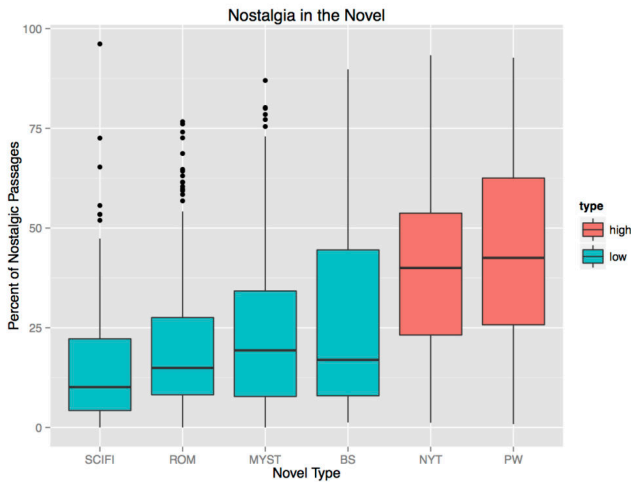


Fig. 6. Boxplot of the average number of nostalgic passages for each of our novel collections as identified by the BoW method. As with figure 5, "high" and "low" refer to cultural status. Once again, we see how prizewinners contain significantly more nostalgic passages than bestselling writing (though we now see a greater range within the bestsellers than before — while the median is low (less than 25%), the third quartile extends close to the prizewinners' median).

Pretty Little Girls with Hair, or The Latent Conservatism of Nostalgia

So what does this all mean? If we can agree that nostalgia is indeed significantly more present in "serious" literature than the "popular" kind, how are we to understand this? Students of literary history will know that the association of Literature (with a capital L) with a nostalgic stance, with modes of reverie and return, has been a poignant cultural constellation since at least the Romantic Age. As Jonathan Sachs has argued, the "slow time" promoted by Romantic literary culture was a conscious reaction against the increasing commercialization of the printed book trade.²² Romantic nostalgia was designed to generate modes of attention and habits of reading that stood in stark contrast to the market's incessant generativity and over-proliferation. Romantic retrospections were a kind of resistance to a new form of writing to the moment. By the turn of the nineteenth century, the "momentary" was no longer understood as a form of occasionality, a periodic focalization on the now as it had been for older, aristocratically inflected models of creative writing. Instead, it was increasingly associated with a sense of an accelerating and dissipating attention to the new.

James English has convincingly argued that this shift to the non-contemporaneity of serious fiction dates more concretely back to the 1980s, when we can see one kind of writing, bestsellers, begin to write to the present and another, prizewinners, write to the past.²³ As English points out, the timing of this shift is remarkably consistent with the ascendance of a neoliberal paradigm in the Anglophone world under Reagan, Thatcher, and later Mulroney. Despite the "fit" of the data, however, English cautions us about having recourse to these large social externalities as explanations. He encourages us in the spirit of Bruno Latour to focus instead on the micro-politics of the literary field — the editors, journals, and review systems that helped promote this shift towards retrospection.

The problem with both explanations, however plausible, is the recourse to causality. Whether it is the neoliberal paradigm shift of the 1980s, the Romantic turn to intensive versus extensive reading habits, or the social networks of prize committees, in each case we are suddenly in need of vastly different kinds of data, which we usually don't have. Is the post-1980 turn to non-contemporaneity actually a return to a longer tradition dating back at least to Romanticism (or maybe Augustine?), which would make the immediate postwar period look more like an anomalous blip? Are the writers or editors in these imagined social networks actually writing or thinking about global capitalism? How can we reliably draw those lines of intermediation between actors that Latour wants us to? And then there is the question of institutional infrastructure: what role does the conglomeration of the publishing industry or the rise of art subsidies play in places like Canada during this period? These are all good questions, but they require entirely different kinds of research, and importantly new kinds of data, which take us well beyond the novels and the literary field of writing.

We want to focus here not on the production of an ad hoc origin story — and all of the attendant problems that causation brings about — but instead on providing a critique of a given state of affairs. Far from celebrating high culture's resistance to the progressive and the now, we want to conclude our paper with an effort to show how these commitments to nostalgia can at some level also contribute to what we see as a kind of latent conservatism. There is a default belief in our discipline that the serious must be good and the popular less so. We simply want to point out that as a retrospective attention to their characters' youth has come to, if not monopolize, then certainly dominate works of serious fiction, we can also see ways in which we simultaneously become semantically trapped in the familiar. The preference for the past that has come to mark serious literature has in our view serious consequences. Computation provides us with a useful tool to better understand the values associated with these literary fields as they manifest themselves on a larger scale.

By way of conclusion, we focus on the semantic fields associated with childhood in our novels as a way of understanding the larger ideological investment of nostalgia. If nostalgia is indeed a powerful framework through which serious fiction differentiates itself, we want to better understand the implications of these novels' investment in such semantic spaces.

A great deal of recent work in the field of child studies has focused on understanding how childhood has been portrayed at different points in the past in different kinds of documents.²⁴ Some of this work has drawn on novels and the depiction of children within them.²⁵ Most often what we find are portraits of childhood that are nuanced, "conflicting," and "critical" of their moment.²⁶ What this field has not been able to do so far is give us a larger aggregate view of the different discourses on childhood that manifest themselves consistently across a broad array of texts.

In order to do so, we employ techniques drawn from the field of distributional semantics, in which vector representations of words are used to better understand the associations that accrue around given key words or concepts within a particular corpus of writing.²⁷ Vector representations are valuable for understanding larger semantic patterns for two reasons: first, they are grounded in the principle that a word's meaning is defined by those words that appear more frequently around it. Rather than use a fixed taxonomy to understand a word's meaning, such as a dictionary or thesaurus, distributional semantics argues that a word's meaning depends on the semantic neighborhoods that accrue around it over time.²⁸ Second, vector representations of words are valuable because they allow us to directly compare semantic fields. Using techniques of addition, we can model aggregate concepts, like "childhood," for example, by combining the vectors for "boy" and "girl." Similarly, using vector subtraction allows us to extract one vector's features from another to isolate those features that are unique to a given concept — by removing the "girl" vector from "boy" we gain a clearer portrait of what is unique about the semantic field of "boy," i.e. we lose what children in general seem to be associated with and can focus more specifically on the semantics associated with a particular subject position.

In Table 5, we see the collocate vectors for "boy" and "girl" in prizewinning novels based on their mutual information scores, which privilege those words that are more likely to appear near the keywords than elsewhere in novels.²⁹ These vectors give us a general portrait of the semantics of childhood, one that is both familiar, but also interesting in its details. Fathers, for example, are much closer to boys than mothers are to girls; girls are more likely to be "little" than boys, while boys are more

associated with "play;" girls are more generally located (trapped might be a better word) in a semantic field of "hair," "beauty," "marriage," "love," and "babies." Indeed, "love" appears 27% more often on the same page as girl than it does with boy. This is a partial view of childhood that our most culturally prized writing today gives us.

If we subtract out those elements that are common to each of the child-positions, i.e. if we subtract boy from girl and girl from boy, we see how the field for girl becomes more clichéd, while the boy-field starts to appear more random (Table 6). It suggests that there is less distinctiveness, but perhaps also less consistency, to the depictions of boys, where girls exist in a more lexically constrained universe. The different levels of mutual information, especially at the top end of the distributions (the M.I. scores remain higher longer), offer another piece of evidence to support this. Fewer words account for more of the semantic consistency of girls.

It is hard not to feel the rather Victorian nature of these fields, especially for girls, and especially when compared to the ways in which girls are depicted more generally within contemporary culture (Table 7). When compared to the semantic field for girls within contemporary magazines drawn from the Corpus of Contemporary American English (COPA), for example, the vector for prizewinning novels feels decidedly more precious and even antiquated. To test this, we looked at girls within a collection of nineteenth-century novels that consisted of roughly 66,000 pages randomly selected from 675 novels. Table 8 shows the semantic field for girls in the nineteenth century, its intersection with the semantic field of prizewinning novels, and what happens when you subtract the Victorian words from prizewinning ones; the table gives us some idea of the unique contribution of prizewinners today towards the semantic framing of girls. While the association of "boys" with girls in prizewinning novels is a unique difference with the nineteenth century (suggesting more inter-gender freedom for girls today), the rest — hair, school, dance, little, blonde, girlfriend, etc. — hardly makes it seem like we have made much of a departure from the antiquated gender constructs of the past. Finally, we looked at the differences between girls in prizewinners and girls in bestsellers (Table 9). It is hard to say that either list feels liberatory. But that is also the point.

These vectors do not discount the important contributions that serious fiction makes to contemporary culture. There are undoubtedly all sorts of subtle, nuanced, ineffable things that these novels also convey. But such analysis does highlight some of the ways in which the stronger orientation around childhood and retrospection in prizewinners means that these novels are spending more time circling around semantic fields that appear to be highly constrained and conventional, especially when looked at in gendered terms. Vector semantics can show us what the horizons of expectation look like for different kinds of concepts or topics within novels. When looked at in this way, the recurrent recourse to childhood and nostalgia brings with it a system of values that are at the very least in need of questioning. The more often we look back, the more we are entwined in a linguistic universe of familiarity, the less we break through convention and the conventional.

The Simple and the Complex

Our work has tried to offer an initial attempt to explore the semantic and symbolic conventions that underlie social forms of cultural value. Our aim has been to better understand the ideological investments that accrue around the processes of valuation that help determine distinctions within the fields of contemporary culture. In bringing these practices to light — in showing the formulas of both high and low culture and the way they help determine one another — our goal is to think more creatively about questions of value and the ways we identify and institutionalize it. By making explicit the implicit judgments and orientations that people in positions of authority exercise, these kinds of macroscopic views can provide useful frameworks for challenging an existing critical consensus about what matters and how.

There are of course numerous limitations to the kind of analyses we are able to offer. Vector semantics can only tell us so much about how language is being used in a given novel. There is a great deal more local complexity than our models can convey. Prizewinners too only represent one kind of "serious" literature. One can imagine constructing more differentiated versions of high culture than we are currently able to. And certainly there is more going on in this process of cultural differentiation than just time — a host of other smaller features coalesce to help us identify "entertainment" from the "important."

And yet taken together, what we have here is a portrait of a diverse sample of some of the most highly valued works of literature written over the past decade, whether it is writing that has sold massively or has made it through the narrowest of critical filters. The language of temporality materializes as one of the most salient ways of distinguishing between these two forms of writing. There is a remarkable alignment between symbolic and economic positions. Texts written for the market celebrate the present and the next, while texts written for critical judgment celebrate a resistance to this very same logic. And yet as critically celebrated artifacts continue to emphasize the semantics of childhood and natural time, we find their expressive universe contracts. It is an important reminder that beneath our most elevated rhetoric about the subtlety and nuance of great fiction, there are a host of under-recognized, far more banal, and ultimately far more conservative features holding these objects together. There is a simplicity to the complex that is important not only to observe, but also to question. The kinds of techniques we have used here can wake us up to the biases or horizons of expectation that we may not know we have and that, upon closer scrutiny, are not all that prize-worthy.

Andrew Piper is associate professor in the Department of Languages, Literatures and Cultures at McGill University. He is also the director of [txtLAB](#), a digital humanities laboratory at McGill.

Eva Portelance is a student in the Department of Linguistics at McGill University.

Tables 5-9. These tables present the mutual information scores in descending order for the top 20 words associated with boys and girls for our various novel collections. They have been normalized to make the scores easier to interpret. The window of association is a single page, which is represented as 250 words of text. The higher the score the more likely a given word is to appear near the keyword than it is elsewhere in the text.

Table 5. The semantics of girls and boys in prizewinning novels

GIRL (PW)	MI Score	BOY (PW)	MI Score
littl	3.01	son	3.66
boy	2.52	girl	2.52
daughter	2.31	father	1.86
hair	2.3	school	1.81
young	2.16	young	1.78
dress	1.79	said	1.21
school	1.78	littl	1.16
danc	1.7	mother	1.08
pretti	1.56	old	1

GIRL (PW)	MI Score	BOY (PW)	MI Score
woman	1.54	children	0.94
said	1.51	child	0.93
mother	1.42	older	0.89
beauti	1.39	brother	0.75
child	1.19	age	0.72
look	1.11	man	0.71
marri	1.09	dead	0.69
love	1.08	kid	0.65
like	1.04	play	0.64
skirt	1	say	0.63
babi	0.99	phone	0.63

Table 6. The unique semantics of girls and boys in prizewinning novels. The left column shows what happens when you subtract boys from girls and the right column shows girls subtracted from boys. The girl vector appears to get more clichéd while the boy vector appears increasingly random.

PW (GIRL-BOY)	MI	PW (BOY-GIRL)	MI
boy	2.52	girl	2.52
daughter	2.31	father	1.24
hair	1.88	littl	1.16
dress	1.57	old	0.68
pretti	1.56	blind	0.56
danc	1.46	oldest	0.54
woman	1.18	ran	0.51
beauti	1.13	brother	0.5
marri	1.09	stick	0.48
skirt	1	stood	0.47
girlfriend	0.79	pig	0.46
love	0.78	fish	0.46
nice	0.73	person	0.46
boyfriend	0.64	complet	0.44
blond	0.64	dead	0.42
like	0.61	uncl	0.41
sister	0.6	yell	0.4
look	0.56	phone	0.39
wore	0.56	fine	0.38
music	0.56	lesson	0.38

Table 7. The semantics of girls in prizewinning novels compared to the semantics of girls drawn from magazine articles in the Corpus of Contemporary American English.

GIRL (PW)	MI	GIRL (COPA)	MI
littl	3.01	gossip	5.6
boy	2.52	poster	5.49
daughter	2.31	cosmo	5.43

GIRL (PW)	MI	GIRL (COPA)	MI
hair	2.3	meets	5.34
young	2.16	teenage	5.02
dress	1.79	boy	4.48
school	1.78	scouts	4.33
danc	1.7	girl	4.15
pretti	1.56	14-year-old	4.09
woman	1.54	sexy	3.99
said	1.51	birthday	3.9
mother	1.42	12-year-old	3.8
beauti	1.39	cute	3.78
child	1.19	shy	3.73
look	1.11	daddy	3.43
marri	1.09	marry	3.29
love	1.08	hungry	3.27
like	1.04	little	3.26
skirt	1	mama	3.06
babi	0.99	17-year-old	3.04

Table 8. Column 1 = the semantics of girls in Victorian novels; column 2 = the intersection of the semantics of girls in prizewinners and Victorian novels, meaning the words that are in both vectors; and column 3 = the unique semantics of girls in prizewinners in relationship to Victorian novels, meaning we subtracted the vector for girls in Victorian novels from prizewinners to arrive at those words that are unique to prizewinners.

GIRL (Victorian)	MI	GIRL (Victorian \cap Prize)	MI	GIRL (Prize - Victorian)	MI
young	5.5	littl	3.01	boy	1.97
mother	2.94	daughter	2.31	hair	1.43
love	2.85	hair	2.3	school	1.22
daughter	2.65	young	2.16	danc	1.19
child	2.6	school	1.78	littl	1.06
poor	2.41	pretti	1.56	skirt	1
marri	2.29	woman	1.54	dress	0.96
miss	2.17	said	1.51	blond	0.95
woman	2.14	mother	1.42	girlfriend	0.79
sister	1.96	beauti	1.39	older	0.75
littl	1.95	child	1.19	teenag	0.72
mrs	1.94	look	1.11	boyfriend	0.64
pretti	1.93	marri	1.09	babi	0.59
said	1.76	love	1.08	name	0.56
beauti	1.65	like	1.04	wore	0.56
kiss	1.63	babi	0.99	giggl	0.55
dear	1.58	older	0.98	stage	0.55
ladi	1.51	sister	0.91	arm	0.51
know	1.49	mrs	0.87	kid	0.5
father	1.36	eye	0.72	rifl	0.5

Table 9. The semantics of girls in prizewinning novels and bestsellers arrived at by subtracting one category from another (removing bestseller words from the prizewinner vector and vice versa).

GIRL (PW - BS)	MI	GIRL (BS - PW)	MI
danc	0.79	littl	2.1
dress	0.78	babi	1.32
hair	0.74	pet	1.27
pretti	0.68	daddi	0.94
woman	0.61	secur	0.85
beauti	0.59	boy	0.82
skirt	0.56	mother	0.78
stage	0.55	school	0.74
death	0.52	teenag	0.65
said	0.52	colleg	0.59
mrs	0.5	friend	0.58
rifl	0.5	case	0.58
dancer	0.47	agent	0.55
left	0.46	vehicl	0.53
film	0.45	judg	0.53
shy	0.44	cri	0.52
kill	0.44	summer	0.5
round	0.42	lawyer	0.48
leg	0.42	name	0.48
girlfriend	0.41	mom	0.46

References

1. James English, *The Economy of Prestige: Prizes, Awards, and the Circulation of Cultural Value* (Cambridge: Harvard University Press, 2005). [2]
2. This is what English calls the process of "capital intraconversion." [2]
3. James English, "Now, Not Now: Counting Time in Contemporary Fiction Studies," *MLQ* (forthcoming). We discuss the methodological differences as well as differences in data between our approaches below. [2]
4. See the two major works, Pierre Bourdieu, *Distinction: A Social Critique of the Judgment of Taste*, trans. Richard Nice (Cambridge: Harvard University Press, 1984) and Pierre Bourdieu, "The Field of Cultural Production, or: The Economic World Reversed," in *The Field of Cultural Production*, ed. Randal Johnson (New York: Columbia University Press, 1993). [2]
5. As researchers in other fields have repeatedly pointed out, quantity is not by itself a solution. Representativeness and relationality are still essential components to literary analysis. How we sample and how we contextualize will continue to play a major role in what we think we know. As English points out, the contemporary literary field is both massive and hard to define. Where does the "contemporary" begin and what should be included within it? The extent to which a collection is representative of a given population is a crucial aspect of the validity of the insights, just as the data to which it is compared will govern what one finds. For a recent example of this problem within the field of social media, see Derek Ruths and Jürgen Pfeffer, "Social media for large studies of behavior," *Science* 346, no. 6213 (2014): 1063-1064. [2]
6. Bourdieu, *Distinction*, 12. [2]
7. For a good introduction to thinking about linguistic patterns and machine learning, Hoyt Long and Richard Jean So, "Literary Pattern Recognition: Modernism Between Close Reading and Machine Learning," *Critical Inquiry* 42, no. 2 (Winter 2016): 235-267. [2]
8. As English writes, "Bestsellers halt their retreat from the present in the early 1980s, pausing for a few years and then rising rapidly back to the earlier ratio and beyond, to the point where 80%-90% of them are set in the present day, a level that has held steady now for two decades. In contrast, shortlisted novels accelerate their abandonment of the present from the late 1970s onward. By the mid-1980s fully half of them are set in the past or the future, and by the late 1990s contemporary settings have clearly become—in the precincts of high critical esteem—a minority taste." English, "Now, Not Now," forthcoming. [2]
9. A full list of the novels can be accessed at <https://dx.doi.org/10.6084/m9.figshare.2061990>. [2]
10. Pierre Bourdieu, "The Field of Cultural Production," 30. [2]
11. We use LIWC both because of its intrinsic relevance to literary study (with its emphasis on psychological and social vocabulary) and its value as a form of dimension reduction. Instead of building a model based on all words in our novels (upwards of 500,000 dimensions), LIWC combines these words into a smaller set of higher-level categories. There are of course other possible methods, such as using principal component analysis. While there may be no best way, a deeper understanding of the different ways of reducing the high dimensionality of literary vector space models is an important area of much needed research. For a review of LIWC, see Yla R. Tausczik and James W. Pennebaker, "LIWC and Computerized Text Analysis Methods," *Journal of Language and Social Psychology* 29, no. 1 (2010): 24-54. For those interested in studying the dictionaries used by LIWC, see their language manual: <http://www.liwc.net/LIWC2007LanguageManual.pdf>. [2]
12. Andrew Piper, "How I Predicted the Giller Prize," *.txtLAB*, <http://txtlab.org/?p=581>. [2]
13. We use an SVM algorithm through the kernlab package in R. For an introduction, we recommend the excellent handbook by Brett Lantz, *Machine Learning with R* (New York: Packt Publishing, 2013). [2]

14. We used a Wilcoxon rank sum test to identify statistically distinct words for each corpus. For a discussion of the appropriateness of this test, see Adam Kilgarriff, "Comparing Corpora," *International Journal of Corpus Linguistics* 6, no. 1 (2001): 97-133. [↗]
15. David Chariandry, *Soucouyant* (Vancouver: Arsenal Pulp Press, 2007), Kindle edition. [↗]
16. Johanna Skibsrud, *The Sentimentalists* (New York: W.W. Norton & Co., 2010), Kindle edition. [↗]
17. The dictionary contains the following 28 word stems: afterward, age, born, boy, child, childhood, children, daughter, dream, famili, father, girl, life, littl, live, memori, mother, old, older, onc, recal, rememb, school, sometim, son, winter, young, youth. We use word stems because they capture a greater level of generality (family and families is considered one unit rather than two, for example). [↗]
18. The prizewinners (mean = 145.34 ± 37.3 per 10K words) were significantly higher than the New York Times novels (mean = 125.78 ± 38.59 per 10K words), $t(398.581) = 5.21$, $p = 3.035e-07$). The data is normally distributed with equal variance. [↗]
19. It is not the case, however, that prizewinners exceed the nostalgic narrativity of their shortlisted counterparts in any significant way. Winning does not correlate with being more nostalgic. As we mentioned earlier, understanding what textual features distinguish winners from non-winners is to date very hard to tell. However, it does appear that women tend to use nostalgic vocabulary to a somewhat greater degree than men (by about 18 more nostalgia words per 10,000 words or 180 words for an entire novel, $p\text{-value} = 0.0005871$). Finally, there was some slight difference for the different prizes. An analysis of variance test suggests that the Man Booker Prize (UK) and National Book Award (US) have slightly lower levels of nostalgia within their collections ($p\text{-value} = 0.0219$). The amount of nostalgia overall has fluctuated from year to year with no clear trend in a particular direction. [↗]
20. Here is a sample passage from our nostalgia training set. It comes from Justin Cronin's science-fiction novel, *The Passage* (2010):

She was just a little girl, not more than six or seven, when she'd asked Sister Margaret, who ran the convent school in Port Loko, what she was hearing, and Sister laughed. Lacey Antoinette, she said. How you surprise me. Don't you know? She lowered her voice, putting her face close to Lacey's. That's nothing less than the voice of God. But she did know; she understood, as soon as Sister said it, that she'd always known. She never told anyone else about the voice, the way Sister had spoken to her, as if it was something only the two of them knew, told her that what she heard in the wind and leaves, in the very thread of existence itself, was a private thing between them. There were times, sometimes for weeks or even a month, when the feeling receded and the world became an ordinary place again, made of ordinary things.

Justin Cronin, *The Passage*, Kindle edition. [↗]

21. Here we present the results of the classification of the training data. We also show the correlation between the predicted passages for each method and the respective dictionary scores for each novel. This shows how well the results correlate between the machine learning and the dictionary-based approach.

Method Accuracy (F1) Correlation p-value

LIWC 97.3% 0.751, $p < 2.2e-16$

BoW 90.1% 0.534, $p < 2.2e-16$

Both methods show very strong correlations, but they also identify discrepancies between novels identified as being highly nostalgic. We found that when there was a strong divergence between the methods in their identification of nostalgic novels (ranking them higher) that both methods were accurate. In other words, each method picked up a different aspect of nostalgia, and thus a combined average would provide the best coverage of nostalgia detection overall. [↗]

22. Jonathan Sachs, "The Glimmer of Futurity, 1811-1815," in *The Regency Revisited*, ed. Tim Fulford and Michael E. Sinatra (Basingstoke: Palgrave Macmillan, 2016), 17-30. [↗]
23. See note 8. [↗]
24. The field of childhood studies dates back to the work of Philip Ariès, *Centuries of Childhood: A Social History of Family Life*, trans. Robert Baldick (New York: Vintage, 1962). For a recent overview of the field, see *The Palgrave Handbook of Childhood Studies*, ed. Jens Qvortrup, William A. Corsaro, and Michael-Sebastian Honig (Basingstoke: Palgrave Macmillan, 2009). [↗]
25. *The Child in British Literature: Literary Constructions of Childhood, Medieval to Contemporary*, ed. Adrienne Gavin (New York: Palgrave, 2012); Ellen Pifer, *Demon or Doll: Images of the Child in Contemporary Writing and Culture* (Charlottesville: University of Virginia Press, 2000); Maria Tatar, *Off with their Heads! Fairy Tales and the Culture of Childhood* (Princeton: Princeton University Press, 1993). [↗]
26. "More specifically," writes Katherina Dodou, "[this essay] seeks to demonstrate that novels at this time have set out to examine critically the meanings attached to the idea of childhood." Katherina Dodou, "Examining the Idea of Childhood: The Child in the Contemporary British Novel," in *The Child in British Literature*, ed. Adrienne E. Gavin (New York: Palgrave Macmillan, 2012), 238. [↗]
27. For an excellent overview of approaches to distributional semantics that includes a discussion of mutual information that will be used here, see Peter D. Turney and Patrick Pantel, "From frequency to meaning: Vector space models of semantics," *Journal of Artificial Intelligence Research* 37, no. 1 (2010): 141-188. For a discussion of vector word representations in particular, see Tomas Mikolov, Wen-tau Yih, and Geoffrey Zweig, "[Linguistic Regularities in Continuous Space Word Representations](#)" (paper presented at the Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, 2013). [↗]
28. We use the measure of mutual information to determine words' association with a given keyword. Mutual information allows us to identify those words that are more likely to appear in proximity to a given word than by mere chance. In more technical language, mutual information provides a measure of how much semantic "uncertainty" is accounted for by a given word relative to the keyword. The higher the score the more a particular word accounts for the possible meaning(s) of the keyword. For example "the" will appear very often near "boy" or "girl," but because it also appears very often elsewhere in novels it does not receive a high mutual information score, while "little" and "school" do receive high scores because they appear more often near boy and girl than they do elsewhere in novels generally. Measuring the mutual information of collocates should not be confused with the technique of word embeddings, which models words' analogous relationship to each other (that "oyster" is often used in similar contexts as "fish"). For a more in-depth discussion of mutual information and its limitations, see again Peter D. Turney and Patrick Pantel, "From frequency to meaning," 141-188. [↗]
29. We define proximity as being on the same page, here modeled as 250 word units. [↗]

P45

Post45 is a collective of scholars working on American literature and culture since 1945. The group was founded in 2006 and has met annually since to discuss new work in the field.

- [Peer Reviewed](#)
- [Contemporaries](#)
- [Book Series](#)
- [Conferences](#)