
"READING BOOKS 'BY THE NUMBERS?'"

STUDYING LITERATURE WITH BIG DATA INSTEAD OF THE CANON

Alexandre JACOBS

Consultant en Gestion Documentaire/Business Analyst

■ L'auteur a reçu le prix ABD-BVD 2017 pour son travail de fin d'études intitulé "*Transmitting information through the pipeline network : reevaluating the gas explosions of San Bruno, Engelhart and Ghislenghien from the perspective of organizational, conceptual or information management-related elements in the pipeline business*" et présenté en juin 2016 à l'Université Libre de Bruxelles (ULB), en vue de l'obtention du titre de Master en Sciences et Technologies de l'Information et de la Communication (MaSTIC). Cet article aborde quelques points forts de ce travail.

■ De auteur mocht de ABD-BVD Prijs 2017 in ontvangst nemen voor zijn eindwerk getiteld "*Transmitting information through the pipeline network : reevaluating the gas explosions of San Bruno, Engelhart and Ghislenghien from the perspective of organizational, conceptual or information management-related elements in the pipeline business*" hetwelk werd verdedigd in juni 2016 in de Université Libre de Bruxelles (ULB) , teneinde het behalen van Master en Sciences et Technologies de l'Information et de la Communication (MaSTIC). Dit artikel haalt een paar sterke punten aan van dit werk.

■ The present theoretical discussion analyzes the impact of big data and statistics upon the study of literature, which is rooted in aesthetics: in other words, it attempts to picture what literature class might look like with big data, determine whether literary criticism could survive this change, and to what extent it would. The article considers both the limits of close reading -which is associated with the literary canon, and therefore inclined to expose some novels, writers or genres as representative of literature on the basis of their quality- and the distanced reading technique -which relies on quantitative observations, that is to say on patterns extracted from literary texts. This article illustrates the interest in resorting to either of these approaches, but also the expense of choosing the one over the other. Also, the viability of a hybrid method based on these two different practices is considered.

■ L'article qui suit analyse d'un point de vue théorique l'impact des big data et des statistiques sur l'étude de la littérature, qui est ancrée dans une tradition d'esthétique : concrètement, il tente d'illustrer ce à quoi l'enseignement de la littérature pourrait ressembler avec les big data d'une part, et de déterminer si (et dans quelle mesure) la critique littéraire pourrait survivre à ce changement. L'article considère pour ce faire les limites de la close reading (qui est associée au canon littéraire, et donc qui exposera certains genres, auteurs, ou romans comme représentatifs d'un gage de qualité) et de la distanced reading (qui repose sur des observations quantitatives, c'est-à-dire des tendances extraites de textes littéraires). L'article se penche sur l'intérêt d'utiliser l'un ou l'autre, mais aussi ce que choisir l'un implique pour l'autre. Pour terminer, l'article discutera la viabilité d'une méthode qui emprunte aux deux pratiques.

■ Het volgende artikel maakt een theoretische analyse van de impact van big data en statistieken op literaire studies, die verankerd zijn in een traditie van esthetica: concreet wordt een poging ondernomen om, enerzijds, te schetsen hoe het literatuuronderwijs er zou kunnen uitzien indien gebruik gemaakt zou worden van big data, en anderzijds, te bepalen of (en in welke mate) de literaire kritiek deze verandering zou kunnen overleven. Om dit te doen bekijkt het artikel de beperkingen van close reading (dat geassocieerd wordt met de literaire kanon, en dus bepaalde genres, auteurs of romans belicht als zijnde representatief voor een blijk van kwaliteit) en distanced reading (wat berust op kwantitatieve waarnemingen, t.t.z. tendensen die afgeleid worden uit literaire teksten). Het artikel buigt zich over het nut dat het gebruik van de ene of de andere methode kan hebben, maar vraagt zich ook af wat het maken van de ene keuze impliceert voor de andere. Ten slotte zal het artikel het hebben over de leefbaarheid van een methode die aan beide praktijken schatplichtig is.

Introduction

Even though some obstacles are still hampering a full-fledged exploitation of big data¹, their potential to revolutionize technologies, business -or, more trivially, dating practices- in the near future leaves little doubt². As a matter of fact, massive amounts of data and applied mathematics are even expected to replace every other tool that might be brought to bear, so that there would finally be "no reasons to cling to our own ways" of resorting to "partial models" such as Newtonian or even Quantum physics, for instance³. And yet, as much as gathering data might be useful in uncountable realms, the added value of collecting,

processing and analyzing data might not come to mind at first sight in some disciplines.

Considering that some fields of studies are intrinsically anchored in aesthetics and the subjective reception of contents (such as fine arts or literature), their ability to withstand the shock of a metric approach based on statistics and quantities deserves attention. True, the application of big data is so broad that fine arts can handsomely benefit from them: a study demonstrates that the association of local metadata with external ones offer remarkable (and inexpensive) opportunities for classification⁴. This being said, applications such as the use of big data

for the study of art, *per se*, raises questions: as for literature, the tradition of the literary canon -i.e. "a selection of works particularly suited for scholarly purposes"⁵ is expected to be in conflict with the big data yardstick, as it is unrelated to the qualitative judgment of the canon. And yet, the surprising possibility of resorting to the big data approach as means to investigate -and evaluate- literature thanks to the drawing of graphs based on novels' characteristics was envisaged as early as 2005 (i.e. more than two years before Anderson coined the term 'big data' on a blog post of the New York Times!) by Franco Moretti in *Graphs, Maps and Trees*.⁶

In order to do justice on the one hand to the graphical illustration of distanced reading (which is a method based on patterns extracted from texts) and, on the other, to the statistical examination of literature, the present paper will consider both the statistical and the canonical methods in the light of what they offer. In order to do so, some salient characteristics of the canon will be examined; in the process, much attention will be devoted to the ambivalent attitude of postmodern and post-colonial writers as these were (as the reader will discover) critical to the literary canon. Once this perspective has been clarified, the paper will investigate the interest in resorting to the quantitative appreciation. In the light of these remarks, the third part will determine the possibility to reach a balanced compromise between these two approaches. This research will be chiefly based on academic articles that discuss literary theory, but will also be performed with the help of material such as an anthology of literature, novels, and Google N-Grams, for instance.

A Qualitative and Quantitative Appreciation of Literary Studies

The Canon

The issue with the canon can be summarized in a statement from Mallarmé's poem "Brise Marine": apparently the French poet "read all the books"⁷. Whatever he meant at the time, this could not have been true (even back then): in Great Britain only, twenty thousand, thirty thousand or more were published at the time⁸! And yet, countless representational literature collections and reading lists have been compiled, even though no one ever read (or simply ever heard about) all the books of a given period. Among the uncountable illustrations for this, the second volume of the Norton Anthology of American Literature between 1820-1865 is interesting: it contains the writings of 45 authors, 28 of whom are male, 17 are female, 3 are Native Americans and, more surprisingly, a few of them are neither poets, nor novelists -e.g., Abraham Lincoln, the political or historical prose of whom offer little from

a purely literary perspective⁹. By no means does this argument denigrate the anthology in particular -and even less Lincoln's writing technique-, but the reader can legitimately wonder why transcriptions of political speeches should be included into the anthology at the expense of, say, some more Algonquin, Cherokee or Apache poetry? Admittedly, this observation is simple and blatant, but it adequately summarizes the type of debates that the canon can raise.

Without explaining the postmodern movement in details, suffice it to mention that adding the works of Lincoln would not be regarded as a fortuitous -or an innocent- choice to them: it illustrates the over-representation of White Americans at the expenses of Natives, which could reflect a vision of US culture assuming that the history of Boston in 1776 is of more importance than the one of Santa Fe the same year¹⁰. By extension, such a choice is "a matter of actual politics, not just academic ones" because the canon is an authoritative "collection of irreducibly individual works which testify in their very uniqueness to the common spirit of humanity", thus forcing to adhere to a dominant representation of culture¹¹. These grievances would also be shared by post-colonial writers for the same reason: in that regard, Homi Bhabha deplores that the literary canon supports "ideological discourses of modernity that attempt to give a hegemonic 'normality' to the uneven development and the differential, often disadvantaged, histories of nations, races, communities, people"¹². True, the canon offers an over-simplified, quality-bound and partitive image of literature that fails to do justice to some authors. And paradoxically, that same canon can offer elements favorable to postmodern and post-colonial views.

Since tastes and interests change over time, the canon is not necessarily at odds with minor genres on a permanent basis: at the time the eminent African-American sociologist WEB Du Bois wrote *Black Reconstruction*, it was legitimate for him to deplore, in the course of the late nineteenth and early twentieth centuries, that college lecturers deliberately encouraged their students to gather thesis material in favor of racism and discrimination towards African-Americans; seventy years later, however, many of his positions -and writings- had nevertheless "become mainstream" as well¹³. True, the canon is authoritative, but its delimitations are sufficiently unclear to easily meet any minor group's agendas: whereas advocates of post-colonialism and postmodernism reject the *social canon* (which consists of a corpus of texts that have a certain educational authority) as representative of the general canon -i.e. "the entire written corpus and every surviving literature"-, they can find in an individual acceptance of the canon a medium particularly well-suited to reflect

*their own personal preferences*¹⁴. In other terms, every individual has his own corpus of references.

In that regard, the illustrious African American literary critic Louis Gates, Jr.'s remark in "What is an African American Classic?" concerning the attachment of minor genres to the canon is unambiguous: he describes the canon as "the texts that a truly well-educated person should have read, and read carefully and closely at least once," and, for years, "[Gates] rued the absence of texts by black authors in this series"¹⁵. Gates acknowledges the selective nature of the canon: "[e]very teacher's syllabus constitutes a canon of some sorts," and his own includes novels on slavery in the United States such as Chesnutt's *The Marrow of Tradition*, Jacobs' *Incidents in the Life of a Slave Girl* or Douglass' *My Bondage and My Freedom*¹⁶. *Regardless of any understanding of the canon, it would be inappropriate to criticize Louis Gates, Jr's for selecting these novels as representatives of a course devoted to African-American literature; true, all these writers are Americans of African descent. Also, Gates honestly exposes these writers as his favorite ones. However, besides not alluding to other less famous writers, it still only emphasizes a unique facet as the common denominator to writers who could be distinctive from other perspectives.*

From a library perspective, resorting to such descriptors is home to interesting observations. Even if indexation can do justice to the multiple origins of an author, using origins remains a highly disputable practice because such descriptions are, needless to mention, a "very partial, temporal, and insufficient characterization at best"¹⁷ and do neither justice to the writer, nor to their novel. This is particularly true in the specific context of literature in a country like the United States, as Americans are a blend of different cultures, heritage and religious beliefs and they are eager to "perceive ethnic distinctions" regardless of "how removed and how artificially selected and constructed" these might be¹⁸. The use of such descriptors result in either making some authors eligible candidates for the purposes of both minor and major agendas alike (an illustration for this is considering literary bibliographies and catalogs can mention "Allen Ginsberg-see Jewish-American literature", "Jack Kerouac-see Franco-American literature, French-Canadian literature"), but sometimes by having an author fit into a category where they would not seem to belong (such as by calling Hemingway as the very general and not most adequate "English-American"¹⁹, implying that his quintessentially American major voice could be perceived from an ethnic perspective)

Statistics

The first part demonstrated the love-hatred relationship that literature scholars (and students) can experience vis-à-vis the canon, not to mention the difficulty to escape from the canon when studying literature. And yet, prior to analyzing the impact of statistical studies upon minor voices, a quantitative approach restricting texts to a succession of words for machine processing purposes does not seem to offer a better option. The conservative philosopher Martin Heidegger would criticize the use of big data for literary analysis very harshly in regard of his skepticism vis-à-vis the cybernetic revolution that occurred several decades before big data: words are used so as to be processed by computers, and language is merely used at a technical level which was, in his opinion, one of the worst aggression that the essence of language could ever undergo²⁰. Interestingly, it also seems that postmodern thinkers (who would not share Heidegger's political views) do not appreciate that manipulation as consistent means of studying literature either: as Bakhtin underlines, "[I]terary language -both spoken and written-, although it is unitary not only in its shared, abstract, linguistic markers but also in its forms for conceptualizing these abstract markers, is itself stratified and heteroglot in its aspect as an expressive system, that is, in the forms that carry its meanings"²¹. In other terms, a mathematical approach fails to grasp all these dimensions -some of them being apparently present at an almost intangible, metaphysical level-, and might therefore leave insufficient room for discussions on literature.

Another reason for which postmodern and post-colonial writers would find a big data approach difficult to accept can be demonstrated thanks to linguistics (Hayakawa) and history (Liu's) besides literature. Hayakawa observes, on the one hand, that science consists of "international systems of weights and measures" allowing "to exchange information with each other, pool our observations and acquire collective control over our environment" whereas literature offers means to help "realize the profound community that exists between us and our fellow human beings": as such, the first one makes mankind able to cooperate, and the second one makes it willing to do so²². On the other hand, Liu's argument demonstrates that new technologies can have a devastating impact upon history: "powerful asynchrony and telepresence [...] together with the sheer abundance of information" tend to create a highly-detailed, but also very partial and segmented notion of history imprisoned in a capsule "that [not only] mobilize[s] across space and time without context in either dimension", but also causes a decontextualizing effect²³.

In other words, a statistical study of literature that borrows from both perspectives would imply, firstly, that numbers can be summoned from a myriad of documents written ages ago on the Web from anywhere and, secondly, that they offer a distinct and accurate value of the most succinct and clearly delimited nature so as to scientifically present literature as a human creation not only directed towards a common goal, but also as one devoid of any intestine conflicts. And this is precisely what postmodernists would never stand: in a 1982 essay about the medias named "The Rediscovery of 'Ideology': Return of the Repressed in Media Studies" -i.e. written a long time prior to the big data phenomenon-, postmodern writer Stuart Hall already expressed his surprise when he remarked with surprise that "[t]he methods of coding and processing a vast corpus of messages in an objective and empirically-verifiable way (content-analysis) were vastly sophisticated and refined" so as to produce oversimplified messages for the masses to consume²⁴. Whether this vision of a consensus based on partial elements seems consistent with the reader's views or not, it seems that more than thirty years later, this observation is still a relevant one.

For these reasons, a quantified approach is expected to be very harmful to literary criticism on several accounts. From a purely representation-oriented standpoint, including the features of African-American literature, for instance, will fail to do justice to the genre; given its minor nature, it would leave a very limited imprint upon the literary production of nineteenth-century America. As such, the contribution of African-American works would be undetectable, and the pattern would highlight their overall resemblance with novels from which they precisely intend to differentiate themselves -including, among others, works indifferent to slaves, or even some overtly in favor of slavery. If merely based on the search for linguistic patterns, the study also silences the obstacles of sociological and political nature -such as racism or slavery to education and social advancement that explains the scarcity of African American literature in the course of history²⁵.

True, the canon gives way to distinguish minor discourses from major ones and gives room for (admittedly unbalanced) dialog. The statistical approach is, however, devoid of that kind of nuances, and does not allow for minor groups to defend themselves: a big data approach to literature might let clear patterns emerge from a graph, they still do not provide more information than numbers and lack interpretation²⁶: indeed, regardless of genres -or works- being significantly more representative than others or not, a statistical study of literature based on the frequency of letters, words or genres does not allow for clear conclusions. This can easily be evidenced by visiting Google N-Grams, and attempting

to demonstrate that the recurrence of a given word is symptomatic of the emergence of African-American novels, for instance. The term "black" comes naturally to mind, but it can apply to much more than simply a color of skin and it is, therefore, not relevant. It would then seem a little more appropriate to use the term "slavery" (which is expected to be used by African-American writers to condemn the institution), but then again, Google N-Grams would not only retrieve the occurrences of that term deprived of their contexts, but also include sentences of racist Southern writers: in short, this does not indicate that a corpus deals with a certain question, but that it mentions a certain word²⁷. This being said, this feature could offer much more appropriate, factual and neutral metadata collection for indexation purposes in library sciences, rather than a partial, canonical-bound tagging transpiring interpretation.

In that regard, Moretti explains that, since no explanations are provided, the context of a trend -and, in an attempt to achieve distanced reading, not of each particular occurrence- has to be found somewhere else²⁸. On that basis, diverging interpretations can be provided for one observation, and such a feature is rather favorable to further confrontations between major and minor perspectives than to the pure and simple extinction of literary criticism. However, "an incredible computing power that can bring to analytical attention patterns of sameness and difference undetectable by the eye of the human reader" can lead to fragile correlations or absurd conclusions, such as the claim that recurring use the letters 'p' and 'b' in Milton's *Aeropagitica* further demonstrating the hypothesis of Presbyterians and Bishops both subscribing to censorship in spite of their diverging religious views²⁹.

And yet, maybe big data would teach future generations of students in literature things that no human ever considered investigating: maybe one of them will demonstrate that John Milton's intention to indulge in a facetious stylistic activity with 'p' and 'b' throughout his literary production -or maybe that the letters 'b' and 'p' religious faiths or ecclesiastical positions could indeed turn men into censors. Still, it seems that the best -and safest- conclusion that a statistical result could offer is limited to figures: in that regard, Moretti prudently states that cyclical interest in some genres does not explain everything in the history of the novel, but can "bring to light its hidden tempo"³⁰ and, that much, at least, is indisputably true.

Towards a Mixed Approach?

As demonstrated, both the quantitative and the qualitative approaches are limited and raise questions, but they also offer interesting insights; it might therefore be interesting to consider how they can intersect so as

to make the study of literature optimal or, in other terms, to let information of a certain quality emerge from texts so as to foster relevant debates. In an attempt to achieve this, Floridi's observations can offer some interesting insights: informational quality "is optimally fit for the specific purposes for which it is elaborated (purpose/depth) and is also easily re-usable for new purposes (purpose/scope)"³¹. The canon offers a more qualitative approach which is useful to a very restricted number of works and not to others -therefore corresponding to the depth of the purpose-, whereas the big data approach will rather focus on the presence of statistical patterns inherent in a non-restricted literary production while offering superficial knowledge -hence the scope. Combining both aspects could be achieved as follows, for instance: it is expected from African-American pieces of work to be much more representative of any subset of literature -a genre- than of literature as a whole. Given this more significant impact, a selection of novels statistically more representative of a clearly defined segment within literature seems to more legitimately deserve an in-depth study -thus coupling both quality and quantity. As a matter of fact, literary classes in universities resort to this method of teaching literature: one class is devoted to a given genre, which can either be a literary movement (e.g., realism), an ethnicity (as discussed earlier) or even a topic (e.g., Pulitzer Prize winners).

Unfortunately, the notion of genre as a balanced compromise is then again problematic from both the canon's perspective as well as the statistical approach. True, the genre is a subset, and this is precisely the nature of the contention: it focuses too much attention upon itself at the expense of other existing ones; as Moretti underlines, the novel consists of the totality of its genres, and whereas this restriction might endow the novel with "its elegance and its power," it still "erased nine tenths of literary history"³². Also, such a restriction might still leave an untold number of novels to read: in other terms, the mixed approach is disputable from a quantitative perspective and, also, from a qualitative one -in that regard, Moretti's use of the term "power" in the quote illustrates once more the qualitative nature of the choice as well. Genres -and sub-genres- accentuate the legitimacy of some discourses since a certain hierarchy emerges out of these classifications -from more to less serious, for instance, or from more to less renowned.³³; therefore, both the inclusion of a novel or a writer into a certain genre -or its exclusion from it- are strongly connoted choices.

Conclusion

In the light of the present arguments, it seems difficult to refute the imperfection inherent in both visions, and

in a hybrid approach as well: relevance in the study of literature either implies a broad overview of the novel with superficial results or a thorough analysis based on a microscopic fraction of literature, but it can not be achieved simultaneously since favoring the one implies affecting the other. Interestingly, it seems that concentrating on a sub-set within big data is contrary to the holistic vision of encompassing everything so as to allow for conclusions to emerge: for that reason, one could also argue that resorting to the genre as a middle ground is the least convincing approach of the three. Because of this impossibility to create a selection of novels ensuring both breadth and depth in scope, an optimal level of informational quality in literary studies can not be expected from a course only teaching any of these three approaches.

However, a class studying literature separately from a quantitative and a qualitative approach -for instance, devoting the first term of the year to a broad study and the other to a deep one deserves a particular attention in the making of twenty-first century literature students -not to mention that such an experimentation would also offer an interesting case study to further analyze the question. Indeed, a strong common denominator characterizes these methods in spite of their differences: their incompleteness, which allows for diverging interpretations, endless debates and, therefore, justify the use of both for investigations. From a literary perspective, the only certainty offered by a statistical analysis based on words, linguistic patterns or genres is of numerical nature, but by no means do quantities provide explanations for themselves in the process: as such, they leave room for a number of more or less convincing understandings that ranges from zero to infinity.

As far as the canon is concerned, the imperfection resides in its being based on the vague and non-demonstrable notion of quality, that favors -and encourages- reading based on subjectivity, if not a determined set of values that meet the reader's beliefs. In other terms, even if distanced reading replaced close reading once and for all, literary criticism would survive the shift, but would certainly not benefit from the process since both methods can lead to different observations -and are, therefore, both respectable in their own rights.

Alexandre Jacobs

Exquando S.A.

Avenue Edison 20

1300 Wavre

alexandre.jacobs@exquando.com

<http://www.exquando.be/>

Mai 2019

References

- ANDERSON, Chris. The End of Theory: The Data Deluge Makes the Scientific Method Obsolete. *Wired: Science*. 2008. Vol. 16.
- BAKHTIN, M. M. *The Dialogic Imagination: Four Essays*. University of Texas Press, 2010. ISBN 978-0-292-78286-0.
- BAYM, Nina et LEVINE, Robert S. The Norton Anthology. *American Literature, 7th ed., vol. A, Beginnings to 1820*. 2007.
- BHABHA, Homi K. *The location of culture*. London; New York : Routledge, 1994. ISBN 978-0-415-01635-3.
- CHESNUTT, Charles W. *The Portable Charles W. Chesnutt*. Penguin, 2008.
- CONN, Peter J. *The American 1930s: a literary history*. Cambridge University Press Cambridge, UK, 2009.
- TERRY EAGLETON. *The Idea of Culture*. Oxford : John Wiley and Sons Ltd, 2013. ISBN 978-0-631-21966-8.
- FISH, Stanley. Mind your Ps and Bs: The digital humanities and interpretation. *New York Times*. 2012. Vol. 23, n° 1.
- Floridi, L. (2006). Information quality. In International Conference on Information Quality was organised in, volume 2006.
- FOWLER, Alastair. Genre and the literary canon. *New Literary History*, 1979, vol. 11, no 1, p. 97-119.
- GUREVITCH, Michael, BENNETT, Tony, CURRAN, James, et al. (ed.). *Culture, society and the media*. London : Methuen, 1982.
- HAYAKAWA, S. J. *Language in Thought and Action*. Oxford, England : Harcourt, Brace, 1949. Language in Thought and Action.
- HEIDEGGER, Martin. Essais et conférences. La question de la technique. *Paris. Ed. Gallimard*, 1958.
- LIU, Alan. Friending the past: The sense of history and social computing. *New Literary History*, 2011, vol. 42, no 1, p. 1-30.
- LOHR, Steve. The age of big data. *New York Times*, 2012, vol. 11, no 2012.
- MENDES, Catulle et DE RICARD, Louis-Xavier. *Le Parnasse contemporain: recueil du vers nouveau*. 1866.
- MORETTI, Franco. *Graphs, maps, trees: abstract models for a literary history*. Verso, 2005.
- SOLLORS, Werner. *Beyond ethnicity: Consent and descent in American culture*. Oxford University Press on Demand, 1986.
- VAN HOOLAND, Seth et VERBORGH, Ruben. *Linked Data for Libraries, Archives and Museums: How to clean, link and publish your metadata*. Facet publishing, 2014.
- VAN HOOLAND, Seth, VERBORGH, Ruben, DE WILDE, Max, et al. Evaluating the success of vocabulary reconciliation for cultural heritage collections. *Journal of the American Society for Information Science and Technology*, 2013, vol. 64, no 3, p. 464-479.
- VAN HOOLAND, Seth, VERBORGH, Ruben, DE WILDE, Max, et al. Evaluating the success of vocabulary reconciliation for cultural heritage collections. *Journal of the American Society for Information Science and Technology*, 2013, vol. 64, no 3, p. 464-479.
- VERBORGH, Ruben, MANNENS, Erik, et VAN DE WALLE, Rik. The rise of the Web for Agents. In : *Proceedings of the First International Conference on Building and Exploring Web Based Environments*. 2013. p. 69-74.
- WOOD, Gordon S. Doing the Continental. *New York Review*, 1997, vol. 20, p. 51-4.
- WOODARD, Komozi. *A nation within a nation: Amiri Baraka (LeRoi Jones) and Black power politics*. Univ of North Carolina Press, 1999.

Notes

1. Verborgh et al., 2013
2. Lohr, 2012
3. Anderson, 2008
4. van Hooland et al., 2013
5. van Hooland and Verborgh, 2014, 161
6. Moretti, 2005
7. Mallarmé, 1866, 168
8. Moretti, 2005, 4
9. Baym, 2007
10. Wood, 1997, 5
11. Eagleton, 2013, 51-55
12. Bhabha, 1994, 171
13. Conn, 2009, 195
14. Fowler, 1979, 98
15. Chesnutt, 2008, Louis Gates, Jr.'s preface, viii
16. Chesnutt, 2008, Louis Gates, Jr.'s preface, xii-xiii
17. Sollors, 1986, 14-15
18. Sollors, 1986, 13
19. Sollors, 1986, 14-15
20. Heidegger, 1958, 40
21. Bakhtin, 288
22. Hayakawa, 1949, 89
23. Liu, 2011, 17
24. Gurevitch et al., 1982, 61
25. Woodard, 1999
26. Moretti, 2005, 26
27. Fish, 2012
28. Moretti, 2005, 9
29. See 26
30. Moretti, 2005, 29
31. Floridi, 2006, 4
32. Floridi, 2006, 30
33. Fowler, 1979, 100