

**RIEF**

**Revue italienne d'études françaises**

Littérature, langue, culture

10 | 2020

La vérité et ses ruses

---

## Les études littéraires entre herméneutique et quantification

**Franco Moretti**

Traducteur : Anne-Frédérique Schläpfer

---



### Édition électronique

URL : <http://journals.openedition.org/rief/6508>

DOI : [10.4000/rief.6508](https://doi.org/10.4000/rief.6508)

ISSN : 2240-7456

### Éditeur

Seminario di filologia francese

### Référence électronique

Franco Moretti, « Les études littéraires entre herméneutique et quantification », *Revue italienne d'études françaises* [En ligne], 10 | 2020, mis en ligne le 10 novembre 2020, consulté le 26 janvier 2021. URL : <http://journals.openedition.org/rief/6508> ; DOI : <https://doi.org/10.4000/rief.6508>

---

Ce document a été généré automatiquement le 26 janvier 2021.



Les contenus de la RIEF sont mis à disposition selon les termes de la Licence Creative Commons Attribution - Pas d'Utilisation Commerciale - Pas de Modification 4.0 International.

---

# Les études littéraires entre herméneutique et quantification

Franco Moretti

Traduction : Anne-Frédérique Schläpfer

---

## NOTE DE L'ÉDITEUR

Nous présentons la transcription de la conférence donnée par Franco Moretti le 23 janvier 2020 à Paris dans le cadre d'une table ronde organisée à la Maison de la Recherche de la Sorbonne Nouvelle. Nous remercions l'auteur de l'avoir révisée pour la publication dans notre Revue.

- 1 Quelle relation, entre l'histoire littéraire quantitative des vingt dernières années et la tradition de l'herméneutique interprétative ? En général, les réponses ont été de deux ordres : pour le camp interprétatif, les deux approches sont incompatibles et la quantification est presque dénuée de valeur ; pour les chercheurs quantitatifs, elles sont parfaitement compatibles et, en fait, complémentaires. Ici, je propose une troisième possibilité, qui émerge d'une comparaison entre les manières dont les deux stratégies opèrent. *Opèrent*, à la lettre : « les pratiques [ayant] souvent de plus vastes implications théoriques que les déclarations théoriques elles-mêmes »<sup>1</sup>. Il s'agit de comprendre ce qu'un paradigme de recherche *fait*, plutôt que ce qu'il déclare vouloir faire.
- 2 Avec une complication car les deux approches rassemblent en réalité de *nombreuses* approches, souvent en forte contradiction entre elles. Je me limiterai donc ici à des travaux dans lesquels je me suis personnellement investi. C'est une décision discutable – dans « Hidden in Plain Sight », Oleg Sobchuck et moi-même avons pris la décision inverse, en examinant environ soixante-dix articles d'une centaine d'auteurs. Mais je la prends, car j'ai été souvent surpris de voir à quel point mon travail est différent suivant les registres dans lesquels il s'inscrit. Dans le cas d'une seule et même personne, cela paraît à la fois bizarre et bon à penser. Peut-être n'est-ce qu'une question

d'incohérence personnelle ; ou, peut-être, le symptôme de quelque chose de plus objectif.

## I

- 3 Commençons par l'herméneutique. Nick Adams, le personnage principal de la nouvelle d'Hemingway intitulée « La grande rivière au cœur double » (1926), s'apprête à partir pêcher :

Nick l'avait tiré de sa boîte à mouche, tandis qu'il était assis, la canne sur les genoux. Il éprouva la solidité du nœud et l'élasticité de la canne en tendant fortement la ligne. C'était une sensation très agréable. Il prit soin de ne pas s'entrer l'hameçon dans le doigt.

Il se mit en route, descendant le courant, la canne à pêche à la main, la bouteille de sauterelles suspendue à son cou par un lacet de cuir noué en demi-clé autour du goulot. Son épuisette pendait à un crochet de sa ceinture. Il portait sur l'épaule un long sac à farine aux coins ficelés en oreilles de lapin. La ficelle passait par-dessus son épaule. Le sac lui battait les cuisses.

Nick se sentait un peu empêtré, mais il ressentait aussi un certain orgueil professionnel avec cet attirail qui lui pendait de tous les côtés. La bouteille à sauterelles ballottait sur sa poitrine. Les sandwiches et la boîte à mouches faisaient bomber les poches de sa chemise contre sa peau.<sup>2</sup>

- 4 Ici a-t-on besoin d'une interprétation ? Pas du tout, si interpréter signifie dissiper l'« obscurité » d'un passage : tout est clair. Mais est-ce vraiment si clair ? L'idée « que la compréhension va de soi », a écrit le fondateur de l'herméneutique moderne, est typique de la « pratique moins rigoureuse » de l'interprétation ; car dans sa version « plus rigoureuse », c'est « l'incompréhension [qui] se produit d'office, et donc la compréhension doit être voulue et recherchée à chaque instant »<sup>3</sup>.

- 5 Voulue à chaque instant... Ces phrases d'Hemingway comportent vingt-cinq propositions prépositionnelles (introduites par une préposition : « from his hook book », « across his lap », etc.)<sup>4</sup>. Vingt-cinq en 149 mots : c'est beaucoup. Mais elles font quelque chose d'essentiel pour le récit : elles rassemblent toutes sortes d'éléments disparates (« The bottle of grasshoppers hung from his neck by a thong tied in half hitches around the neck of the bottle »). Un couteau suisse : un petit monde compressé, bien organisé, et rempli de choses :

Avec la hache, il tailla dans une souche de jeune sapin, fendit l'éclat frais et lisse pour en faire deux piquets de tente. Il les voulait longs et robustes pour tenir solidement dans le sol [...]. Il tendit la toile en enfonçant profondément les piquets avec le talon de la hache jusqu'à ce que les boucles fussent enterrées dans le sol et la toile tendue comme une peau de tambour. (233)

- 6 Un monde de choses, mais pas seulement : Nick veut des piquets « for his tent », ils sont « long and solid to hold the ground », et il les enfonce « until the rope loops were buried ». Chaque action en prépare une autre. C'est le *know how*, le « Savoir-faire », comme Gilbert Ryle a nommé ces chaînes de mouvements silencieux. Nick « tied the rope [...] and pulled the tent up [...] and tied it to the other pine. » Toujours calme et efficace. Mais : le calme – dans un récit ?
- 7 D'ordinaire, nous lisons des récits parce que nos vies ne sont pas assez riches d'événements ; mais si l'expérience clé de toute une génération avait été la Grande Guerre ? Trop d'événements : et c'est ainsi que naît le désir du calme dans la narration. La littérature de guerre, a-t-on écrit, porte sur « des hommes qui [...] n'avaient pas de

contrôle sur les événements qui menaçaient leurs vies »<sup>5</sup>. Pas de contrôle : c'est cela. Le style d'Hemingway est *tout contrôle* : de l'espace, du temps, des gestes, des mots. La vie dans les tranchées alterne entre l'ennui et la terreur ; rien pendant des jours, puis l'apocalypse. La prose d'Hemingway n'est jamais ennuyeuse et jamais effrayante : propre et prudente, c'est parfait pour la *convalescence* (trois ans plus tard, l'épisode central – et le plus heureux – de son grand succès, *L'Adieu aux armes*). Littérature de guerre, parce qu'elle veut *guérir* de la guerre : résoudre la dissonance de l'expérience historique, pour adapter une métaphore de Lukacs. J'y reviendrai.

## II

- 8 De quelques phrases d'une seule nouvelle à la plus évidente nouveauté de l'approche quantitative : c'est l'expansion de l'histoire littéraire bien au-delà d'un petit canon de chefs-d'œuvre. Il fut un temps où un théoricien choisissait un texte – *Don Quichotte*, *Robinson Crusoé*, *L'Idiot* – et érigeait sur cette base toute une théorie du roman. Le « Type thinking », comme l'a nommée Ernst Mayr : « *Tristram Shandy* est le roman le plus caractéristique de la littérature universelle », écrivait Chklovski dans *Sur la théorie de la prose*<sup>6</sup>. Mais face au fourmillement des romans anglais du dix-neuvième siècle de la figure 1, la pensée par type, le « type thinking », est inutile : ici, on doit rendre compte d'une population entière de romans. Pas très grande – 1117 pour être exact – mais tout de même irréductible à un texte unique.

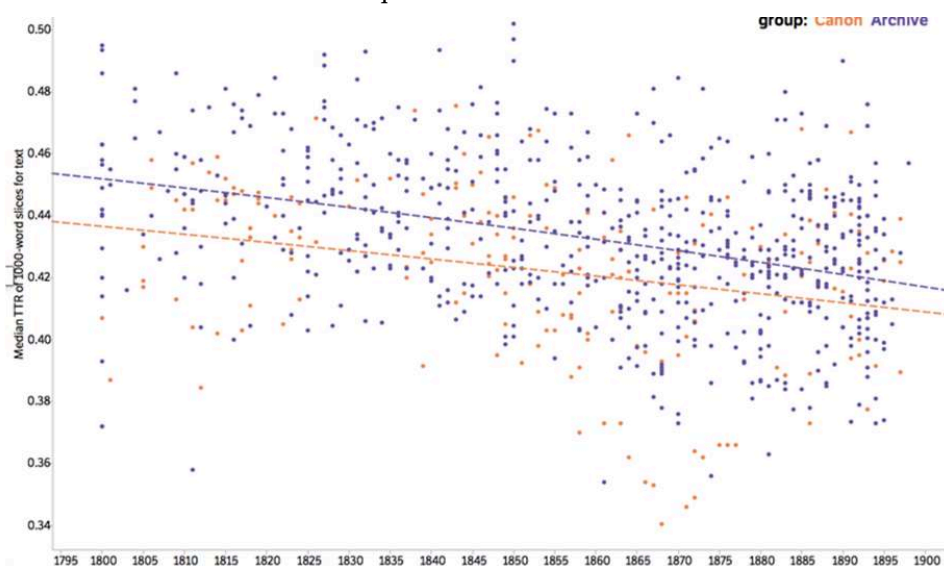


Figure 1.

TIRÉ DE MARK ALGEE-HEWITT, SARAH ALLISON, MARISSA GEMMA, RYAN HEUSER, FRANCO MORETTI, HANNAH WALSER, « LE CANON/L'ARCHIVE: DYNAMIQUE DE LARGE ÉCHELLE DANS LE CHAMP LITTÉRAIRE »<sup>7</sup>.

- 9 Le texte : voilà le point de discord. Il a été l'objet même des études littéraires ; ici, c'est un point. Il a été réduit à un point, tout comme les événements historiques pendant le premier tournant quantitatif d'il y a soixante-dix ans. Les événements étaient aussi centraux pour l'historiographie que les textes pour les études littéraires, et pour la même raison : pour leur unicité. « Les historiens ressemblent à des collectionneurs », écrivait Krzysztof Pomian, « n'amass[a]nt que des choses rares et curieuses, délaissant

tout ce qui était banal, quotidien, usuel... »<sup>8</sup>. Mais une fois que les événements ont commencé à être étudiés par les historiens des *Annales* « en tant qu'éléments d'une série », l'unicité a perdu sa signification, et les événements ont fini par être « relégués à la périphérie de l'histoire, ou complètement ignorés ». Une histoire qui ignore les événements : voilà le tournant quantitatif. Une histoire de l'art sans noms, a écrit Wölfflin. Une histoire de la littérature sans textes.

- 10 Sans textes – *parce qu'il y a trop de textes* : les 250 romans canoniques et les 850 romans oubliés de la figure 1 (respectivement les points orange et les points bleus). Nous voulions vérifier si la richesse linguistique contribuait à la survie ou à l'oubli d'un roman ; tous les textes ont donc été découpés en segments de mille mots, dont on a calculé le ratio type-occurrence<sup>9</sup>. Il s'est avéré que le segment avec le score le plus élevé appartenait à *Arthur Montague, or an Only Son at Sea* (1850) d'Edward Hawker et celui ayant le score le plus bas, à *Adam Bede* (1859) de George Eliot<sup>10</sup>.
- 11 Qu'un roman dont aucun de nous n'avait jamais entendu parler possède une variété lexicale plus riche qu'un texte canonique comme celui d'Eliot allait à l'encontre de ce que nous attendions. Donc nous avons lu ces segments, et beaucoup d'autres segments agrégés autour des extrêmes de ce spectre. Ceux de Hawker relèvent de la description, ce qui fait sens, puisque la description exige des détails, et que les détails augmentent la variété lexicale. Le segment d'Eliot est complètement différent : une jeune femme, avouant avoir abandonné et laissé mourir son enfant, répète sans cesse les mêmes mots, comme si elle était enchaînée – « but I couldn't go away » – à cette scène de mort. Généralisons : un ratio élevé caractérise la prose du narrateur : écrite, analytique, impersonnelle, et presque atemporelle ; un ratio bas, en revanche, caractérise la voix du personnage : brisée, désespérée et hantée par un événement terrible.
- 12 Voilà la lecture de textes considérés comme les éléments d'une série : nous aurions pu partir du segment d'*Adam Bede* pour aller vers la représentation de l'infanticide et de son rôle dans la culture victorienne (comme je l'ai fait dans le cas d'Hemingway) ; mais non : nous sommes partis du segment d'Eliot pour aller vers *d'autres segments romanesques*, et construire des paires conceptuelles abstraites – narrateur/personnage, écrit/oral, analytique/émotionnel, etc. – qui devraient définir l'espace des possibles stylistiques. Pas facile, pas réalisé, hélas, mais la direction était claire<sup>11</sup> : nous n'étudions plus des textes individuels, mais des *séries de textes*. C'est différent.

### III

- 13 Douze phrases, 1 100 romans. L'interprétation d'un texte, la mesure d'un corpus. Et la question revient : quelle relation entre ces dernières ?
- 14 Premièrement, elles sont toutes les deux valides ; à ce sujet, les critiques émanant du camp interprétatif sont complètement infondées. Toutes les deux valides – et avec un moment de chevauchement. Au cours de l'étude quantitative, il y a eu interprétation : nous avons pris la phrase de Hawker, « acres of refreshing greensward, studded with oak, walnut, and hawthorn », et nous l'avons transformée – c'est-à-dire *interprétée* – comme « de la prose analytique et impersonnelle du narrateur romanesque ». À l'inverse, l'interprétation de « La grande rivière au cœur double » a été déclenchée par une forme implicite (et très élémentaire) de quantification : c'était parce que les propositions prépositionnelles étaient si incroyablement nombreuses que je les ai

remarquées – et que « j’ai voulu comprendre », selon l’expression de Schleiermacher. Dans une certaine mesure, chaque méthode s’est appuyée sur l’autre, et des chercheurs quantitatifs parlent à ce propos d’une stratégie « oscillatoire » : un travail « de va-et-vient entre les lectures de près et de loin pour s’approcher d’un centre conceptuel imaginaire »<sup>12</sup>.

- 15 Un travail de va-et-vient... Je ne suis pas d’accord. Oui, j’avais compté jusqu’à vingt-cinq en travaillant sur Hemingway ; pas grand-chose, en termes de quantification, et pourtant, cela m’a paru être tout à fait suffisant. Même chose pour *Adam Bede* : la confession d’Hetty Sorrel est un passage extraordinaire pour l’interprétation ; nous n’avons presque rien fait, et pourtant, à nouveau, cela semblait suffisant. Et en vérité, *c’était* suffisant, parce que les deux études supposaient une *hiérarchie* très claire entre les deux méthodes : la quantification était le moyen et l’interprétation le but dans l’étude sur Hemingway, et inversement pour les 1 100 romans. D’une part, je voulais comprendre comment un récit de pêche à la truite pouvait devenir la nouvelle la plus lue dans les universités américaines ; il devait y avoir *quelque chose de plus* que de la pêche, et quand j’ai vu toutes ces propositions prépositionnelles, j’ai pensé qu’elles étaient un symptôme de ce « quelque chose ». Mais j’essayais de comprendre *les propositions* : qu’il y en ait vingt-cinq, ou dix-huit, ou trente, c’était égal<sup>13</sup>. D’autre part, dans le cas d’*Adam Bede*, nous mesurions le ratio type-occurrence, et ces horribles mots-dièse symbolisaient nos priorités ; ils montraient d’emblée les répétitions, comme nous le voulions (et s’ils ont transformé la lecture en torture, tant pis : la lecture n’était pas le but ici). Conclusion : oui, il y a eu un moment de chevauchement entre les deux méthodes – et puis il a passé ; mieux, on l’a écarté : le travail avançant, l’interprétation s’est faite toujours plus interprétative et la quantification plus quantitative. Pas de va-et-vient, ici ; pas de centre conceptuel.
- 16 Mais peut-être pourrait-on imaginer une étude dans laquelle les deux méthodes auraient le même poids ?
- 17 Je ne connais pas une telle étude ; mais *en principe*, c’est certainement possible. Nous pouvons aisément interpréter les phrases d’Eliot avec la même profondeur que celles d’Hemingway, ou mesurer la fréquence des phrases prépositionnelles dans toutes les nouvelles américaines des années 1920. Rien n’empêche les deux méthodes de travailler l’une à côté de l’autre. Peuvent-elles aussi travailler *ensemble* ? « That is the question ». La quantification peut bien donner de nouveaux objets à l’activité herméneutique ; les interprétations se prêter à la falsification quantitative : cela est clair.<sup>14</sup> Mais ici, la question porte sur la relation entre *les catégories* de l’analyse : est-ce que les catégories quantitative et herméneutique ont la capacité d’interagir afin d’*unifier conceptuellement* les deux approches ? A-t-on la possibilité de formuler une synthèse théorique ?

## IV

- 18 Certaines caractéristiques, écrit Georges Canguilhem dans son étude sur l’épistémologie médicale du dix-neuvième siècle,
- sont qualifiées de normales en tant qu’elles désignent des caractères moyens et les plus fréquents [...]. Mais elles sont aussi qualifiées de normales parce qu’elles entrent à titre d’idéal dans cette activité normative qu’est la thérapeutique. [...] l’état normal désigne à la fois l’état habituel des organes et leur état idéal.<sup>15</sup>

- 19 Le normal comme moyenne-fréquent-habituelle, et le normal comme idéal normatif : un signifiant et deux concepts distincts. Quelle relation, entre les catégories de la quantification et de l'herméneutique ? La même : la première examine les fréquences de la littérature, et la seconde son côté normatif. Normatif, dans le sens que lui, Panofsky, attribuait lorsqu'il parlait de l'art comme « un conflit objectivant, ayant pour but des résultats définitifs, entre la force de la forme et un matériau à maîtriser »<sup>16</sup>. *Bewältigen* : maîtriser, remodeler des matériaux historiques avec la force – *Kraft* – de la forme esthétique. C'est le côté normatif de la littérature : on prend ce qu'il y a pour le transformer en quelque chose d'autre. Et c'est aussi ce sur quoi, ou plutôt *contre* quoi, travaille l'interprétation : en prenant ces « résultats définitifs » pour s'efforcer à *défaire le travail de la forme* : aller du texte, à ses techniques, à son dehors et à la « dissonance » à résoudre. L'interprétation : une compréhension de la littérature qui a toujours la tentation d'aller *au-delà* de la littérature.
- 20 Aller *au-delà* de la littérature est, en revanche, ce que l'approche quantitative ne peut pas faire. Avec le texte individuel qui est l'objet typique de l'herméneutique, la rétro-ingénierie nous indique les aspects de la réalité historique sur lesquels nous devrions nous concentrer, la morphologie, agissant comme un catalyseur de l'intuition historique<sup>17</sup>. Avec des centaines ou des milliers de textes – ou plus encore – cela devient impossible et le lien vertical entre le texte et le monde est remplacé par un lien horizontal entre des textes et des textes, ils sont tous sur le même plan. Les phrases d'Hemingway m'avaient mené à la guerre ; le segment d'Eliot, à d'autres segments romanesques. C'est frappant comme l'approche quantitative est limitée au champ littéraire.
- 21 La quantification emprisonnée dans les livres. Une faiblesse ? Je ne suis pas d'accord. Si, dans le passage de l'herméneutique à la quantification, on perd une dimension, on en récupère une autre : nous ne savons toujours presque rien de la manière dont les systèmes littéraires fonctionnent, ni de la logique de la quantification de diriger notre regard justement vers la structure interne de ces systèmes. Les relations morphologiques mises en lumière par les 1 100 points de notre recherche sont donc un pas, petit, mais réel, vers la solution du gigantesque puzzle du champ romanesque. Un puzzle, il faut le dire, dont la tradition herméneutique, malgré toute sa créativité, n'a jamais été curieuse.

## V

- 22 Quelle relation, entre l'herméneutique et la quantification ? Quand j'ai commencé à réfléchir à cet essai, je ne connaissais pas la réponse. Ayant travaillé souvent avec chacune des méthodes – mais jamais avec les deux ensemble – j'avais imaginé un livre à venir qui serait une sorte de synthèse – une théorie unifiée, micro-macro, de la littérature. Puis, j'ai commencé à décrire ce que j'avais effectivement fait, et le rêve s'est évaporé : le chevauchement des deux pratiques, ou du normatif et du fréquent était trop faible et épisodique pour une véritable synthèse. Faible, parce que la différence entre les deux approches n'est pas une question de niveaux – micro ou macro – mais relève des *questions* qu'elles posent à la littérature.
- 23 J'aurais dû le savoir. L'interprétation *transforme* tout ce qu'elle touche : « ceci signifie en fait cela ». La quantification ne peut pas modifier la moindre de ses données. Les

approches sont antithétiques. Songez à la manière dont elles travaillent sur la forme. L'interprétation va *de la forme au monde*, à la poursuite de la signification historique des œuvres littéraires. La quantification va *de la forme à la forme*, en essayant de dresser un atlas des séries littéraires. Pour la première, la forme est *une force, une activité* : la maîtrise des matériaux historiques qu'il faut approcher avec soupçon, contrer, et démasquer. Pour l'autre, la forme est un produit fini, qu'il faut mesurer avec la tête froide et placer au sein d'un système de relations. L'un, animé par le pathos de la lutte, mène à l'histoire ; l'autre, animé par le pathos de la découverte, à la morphologie. De grandes passions, toutes les deux, la lutte et la découverte. Mais trop exclusives pour travailler ensemble. Elles peuvent certainement – je le répète – travailler l'une à côté de l'autre : s'offrir mutuellement de nouveaux objets d'analyse, ou déclencher de nouvelles questions dans leurs domaines respectifs. Mais elles ne peuvent pas *intervenir* l'une dans l'autre. Nuit et jour : alors que l'une commence, l'autre disparaît. Elles ne cesseront jamais de se poursuivre, et ne deviendront jamais une seule entité.

- 24 Bon. C'est le contraire de ce que j'avais espéré, mais il y a une logique à notre pratique, et nous devrions chercher à la comprendre – à comprendre *ce que nous faisons véritablement lorsque nous étudions la littérature* – plutôt que d'évoquer une synthèse que personne n'a jamais vue. Cela va sans dire : il ne s'agit que de mon opinion sur la question ; quelqu'un d'autre pourrait trouver une synthèse – demain. Alors, les choses changeront. D'ici là, comme Arnold Schönberg l'a dit une fois, le chemin médian est le seul à ne pas mener à Rome.

---

## NOTES

1. « Hidden in plain sight. Data visualization in the humanities », dans *New Left Review*, 118, 2019, p. 86-115. « Hidden in plain sight » et le présent article font partie d'une série de réflexions portant sur l'étude quantitative de la culture, incluant « Operationalizing: or, the function of measurement in modern literary theory » (2013), « Literature, measured » (2016), et « Patterns and Interpretation » (2017), désormais rassemblées dans *Canon/Archive. Studies in Quantitative Formalism from the Stanford Literary Lab*, New York, 2017. Deux de ces publications ont fait l'objet de traduction. Voir : « "L'opérationnalisation" ou, du rôle de la mesure dans la théorie littéraire moderne », dans *Critique*, 8, 2015, p. 712-734 ; « Abstraction, motifs récurrents et forme », dans *Revue d'histoire littéraire de la France*, 116, 3, 2016, p. 521-532.
2. E. Hemingway, « La grande rivière au cœur double », dans Id., *Nouvelles complètes*, tr. de M. Arnaud, M. Duhamel et al., Paris, Quarto Gallimard, 1999, p. 228-244. Ici, p. 237.
3. Les extraits sont issus du « Compendium » de F. Schleiermacher, son texte programmatique de 1819. Voir : *Hermeneutics: the handwritten manuscripts*, dir. H. Kimmerle, Missoula, Montana, Scholars Press, 1977, p. 109, p. 110. « L'obscurité », comme l'observe justement Peter Szondi, n'est guère, pour Schleiermacher, « la seule occasion d'interprétation », *Introduction to literary hermeneutics*, Cambridge, Cambridge University Press, 1995, p. 27.
4. Les voici, en italiques : « Nick took it *from his hook book, sitting with the rod / across his lap*. He tested the knot and the spring *of the rod* by pulling the line taut. It was a good feeling. He was careful not to let the hook bite *into his finger*. He started *down to the stream*, holding his rod, the



bottle of grasshoppers hung from his neck / by a thong tied in half hitches / around the neck / of the bottle. His landing net hung by a hook / from his belt. / Over his shoulder was a long flour sack tied at each corner / into an ear. The cord went over his shoulder. The sack flapped against his legs. Nick felt awkward and professionally happy with all his equipment hanging from him. The grasshopper bottle swung against his chest. In his shirt the breast pockets bulged against him / with his lunch and his fly book. » Cf. E. Hemingway, *The Nick Adams Stories*, Scribner, New York, 1981, p. 190.

5. E. J. Leed, *No man's land*, Cambridge University Press, 1981, p. 33.

6. V. Chklovski, *Sur la théorie de la prose*, tr. G. Verret, Lausanne, L'Âge d'homme, 1973, p. 244.

7. Cette étude est désormais disponible dans *Canon/Archive*, cit. Ce qui suit n'en est qu'un bref résumé. On en trouvera une version française dans F. Moretti (dir.), *La Littérature au laboratoire*, Paris, Ithaque, 2016, p. 217-265.

8. K. Pomian, « L'histoire des structures », dans J. Le Goff, R. Chartier, J. Revel (dir.), *La Nouvelle Histoire*, Paris, Retz-CEPL, 1978, p. 536 et p. 543-544.

9. Le ratio type-occurrence est une mesure standard de variété lexicale qui exprime la relation entre le nombre de mots différents (types) et le nombre effectif de mots utilisés (occurrences). « Bonjour, mon très cher ami » comporte cinq types et cinq occurrences, et donc un ratio type-occurrence de 5/5, ou 1. « Bonjour, Jacques, mon cher Jacques » comporte aussi cinq occurrences, mais seulement 4 types, et donc un ratio type-occurrence de 4/5, ou 0.8.

10. Voici deux sections de ces segments, où les mots-dièse signalent des mots déjà présents dans le même segment, tandis que les astérisques indiquent des mots qui n'appartiennent pas au lexique du roman anglais (environ 230 000 mots), tel qu'il a été établi par Ryan Heuser, qui a écrit le programme pour cette partie de l'expérience.

*Arthur Montague* : « then cut through some acres of refreshing greensward, studded with the oak, walnut, and hawthorn, ascended a knoll, skirted an expansive sheet of# water; afterwards entering an# avenue of# noble elms, always tenanted\* by a# countless host of# cawing\* rooks, whose clamorous conclaves\* interrupted the# stillness that reigned around, and# whose# visits to adjacent corn-fields\* of# inviting aspect raised the# ire and# outcry of# the# yelling urchins employed to# guard them from depredation. »

*Adam Bede* : « And# I# made haste out# of# the# wood#, but# I# could# hear it# crying# all# the# while#; and# when# I# got# out# into# the# fields#, it# was# as# if# I# was# held fast#-- I# could# n't go# away#, for# all# I# wanted so# to# go#. And# I# sat# against# the# haystack# to# watch if# anybody# `ud come#: I# was# very# hungry, and# I#d only a# bit of# bread# left; but# I# could# n't go# away#. »

11. Nous avons échoué parce que des cas extrêmes tels qu'*Arthur Montague* et *Adam Bede* possèdent une clarté épistémologique que les cas normaux n'ont pas. Pour une interaction similaire entre l'extrême et la moyenne, où cette dernière a été analysée avec plus de succès, voir : M. Algee-Hewitt, R. Heuser, et F. Moretti, « On Paragraphs. Scale, Themes, and Narrative Form », désormais dans *Canon/Archive*, cit. Pour la version française, voir : « Des paragraphes : échelle, thèmes et forme narrative », dans F. Moretti (dir.), *La Littérature au laboratoire*, cit., p. 186-216.

12. A. Piper, « Novel Devotions: Conversional Reading, Computational Modeling, and the Modern Novel », dans *New Literary History*, 46, 2015, p. 67-68. De même, Hoyt Long et Richard Jean So ont préconisé « une méthode de lecture qui oscille ou pivote entre l'interprétation humaine et l'interprétation mécanique, chacune fournissant des informations à l'autre et soutenant l'effort du critique qui cherche à extraire du sens des textes » (« Literary Pattern Recognition: Modernism between Close Reading and Machine Learning », dans *Critical Inquiry*, Winter 2016, p. 267).

13. Mieux : le fait qu'il y en ait suffisamment pour être visibles, comme cela a toujours été le cas dans la critique stylistique, a fait une différence. Leur nombre exact pouvait rester vague. Voir, par contraste, la précision avec laquelle Sarah Allison et Marissa Gemma ont établi le lien entre le

registre de la convention dans la *Longman Grammar of Spoken and Written English* (avec un ratio type-occurrence de 30%), et les 500 segments ayant les scores les plus bas de notre corpus, où le ratio type-occurrence oscille entre 27% et 33% (voir : *Canon /Archive*, cit., p. 283). Dans leur réflexion, le pathos de la recherche est indissociable d'une précision inimaginable dans la tradition herméneutique.

14. Voir à ce propos : « Patterns and Interpretation », désormais dans *Canon/Archive*, cit., et pour une discussion approfondie, on se référera à : S. Allison, *Reductive Reading. A Syntax of Victorian Moralizing*, Baltimore, Johns Hopkins University Press, 2018, p. 19ff.

15. G. Canguilhem, *Le Normal et le Pathologique*, Paris, Presses universitaires de France, 1966, p. 98 et p. 101.

16. « Der Begriff des Kunstwollens », dans *Zeitschrift für Ästhetik und allgemeine Kunstwissenschaft*, XIV (1920), p. 339.

17. La liste des phénomènes majeurs (sans parler des phénomènes mineurs) qui pourraient être considérés comme issus « du monde » pour n'importe quelle œuvre littéraire s'avère virtuellement infinie. Pour un Américain vivant en Europe dans les années 1920, cela incluait non seulement les tranchées de la Première Guerre mondiale, mais aussi la révolution socialiste, les voitures et les avions, une décennie de liberté sexuelle sans précédent, les guerres civiles, l'architecture rationaliste, la radio, d'incroyables expérimentations picturales et musicales, le début de l'hyper-inflation... Une forme singulière réagit d'ordinaire à un nombre restreint de ces phénomènes que l'interprétation parvient parfois à isoler avec succès. Mais avec un large corpus, les mécanismes formels croissent dans toutes les directions, et les fils qui lient le monde à l'œuvre deviennent désespérément enchevêtrés.