



**HAL**  
open science

## DHQ: Digital Humanities Quarterly Que mille lectures s'épanouissent... Modélisation du personnage et expérience de “ crowdreading ”

Ioana Galleron, Fatiha Idmhand, Cécile Meynard

### ► To cite this version:

Ioana Galleron, Fatiha Idmhand, Cécile Meynard. DHQ: Digital Humanities Quarterly Que mille lectures s'épanouissent... Modélisation du personnage et expérience de “ crowdreading ”. Digital Humanities Quarterly, 2018. halshs-01815606

**HAL Id: halshs-01815606**

**<https://shs.hal.science/halshs-01815606v1>**

Submitted on 14 Jun 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# DHQ: Digital Humanities Quarterly

Preview

2018

Volume 12 Number 1

## Que mille lectures s'épanouissent... Modélisation du personnage et expérience de « *crowdreading* » [ en ]

Ioana Galleron &lt;ioana\_dot\_galleron\_at\_gmail\_dot\_com&gt;, University of Grenoble

Fatiha Idmhand &lt;fatihaidmhand\_at\_yahoo\_dot\_es&gt;, Université de Poitiers

Cécile Meynard &lt;cecile\_dot\_meynard\_at\_gmail\_dot\_com&gt;, University of Angers

### Abstract

Au cours des dernières années, le distant reading s'est imposé comme l'une des méthodologies les plus intéressantes pour les études littéraires assistées par ordinateur. Toutefois, une autre voie, aussi prometteuse, est actuellement moins explorée : celle de la lecture partagée, que l'on pourrait également nommer *crowdreading*, et qui consiste à soumettre un même texte à de multiples lecteurs, afin d'observer les convergences et les divergences entre leurs compréhensions et interprétations du texte. Cet article présente les résultats d'une première expérience de *crowdreading*, visant à comprendre de quelles façons les lecteurs construisent l'image des personnages littéraires, et dans quelles mesures ils en perçoivent les différentes dimensions qui ont été dégagées par des théoriciens de la littérature comme Ph. Hamon ou V. Jouve.

In some of these cases the model can be improved, in others perhaps not, but even catastrophic breakdown does not imply an epistemological waste of time, as long as good questions come of it. [McCarty 2005]

### Introduction

Au cours des dernières années, le « *distant reading* » (expression pour laquelle un équivalent consensuel ne semble pas encore avoir été trouvé en français<sup>[1]</sup>) s'est imposé comme l'une des méthodologies les plus prometteuses pour les études littéraires assistées par ordinateur. Avec l'analyse de grands volumes de données, il devient possible d'appréhender autrement l'évolution du roman ([Moretti 2013] ; [Jockers 2013]), d'observer la transformation des structures au théâtre [Fischer, Kampkaspar et Trilcke 2015], d'analyser autrement les techniques et procédés littéraires, etc. S'il semble incontestable que la rhétorique, la stylistique, l'histoire de la littérature (et pas seulement celle des grands auteurs) et les sciences du texte en général ont beaucoup à attendre et apprendre de l'exploitation des « *big data* » littéraires, il nous semble que la disponibilité des textes en format électronique et la facilité de créer d'importantes communautés de lecteurs grâce à internet offrent une autre voie, aussi intéressante, qui pourrait être explorée. C'est la voie de ce que nous appelons la « lecture partagée », que l'on pourrait également nommer « *crowdreading*<sup>[2]</sup> », et qui réside d'après nous dans la possibilité de soumettre un même texte à l'interprétation d'un nombre indéfini de lecteurs et de recueillir les résultats de ces multiples lectures à des fins de compréhension du phénomène littéraire.

Il est vrai que cette idée est déjà exploitée, jusqu'à un certain point, dans la plupart des projets de crowdsourcing : ces travaux d'édition électronique adoptent des dispositifs qui permettent aux lecteurs des sources publiées d'apporter leur contribution à l'édition, de réagir, commenter, compléter ou transcrire, de prendre part, en somme, à la construction du savoir. Ces entreprises se distinguent toutefois de celle que nous envisageons, à la fois en raison de leur but (éditorial en premier lieu) et en raison du rôle qu'elles donnent aux

1

2

collaborateurs extérieurs, dont elles attendent moins une interprétation personnelle, qu'informée par la philologie, les théories savantes, et en lien avec les objectifs scientifiques propres au projet. Rares sont donc les projets qui intègrent, à l'heure actuelle, la discussion et l'examen critique, ainsi que la production et la confrontation d'interprétations et de lectures aussi bien savantes que profanes, convergentes ou contradictoires. Plus rares encore sont les projets qui envisagent de se concentrer sur ce matériau second et d'exploiter les métadiscours produits par des collaborateurs d'origines et de cultures diverses, en vue de l'élaboration de nouvelles perspectives sur le fait littéraire.

Fortes de ces constats, nous avons souhaité mesurer la pertinence d'une telle approche en réalisant une expérience de *crowdreading*. Celle-ci visait à caractériser la notion de personnage en littérature et à interroger les enrichissements apportés par de « multiples lectures » à une tentative de modélisation de cette notion. Elle a été menée au sein du consortium CAHIER<sup>[3]</sup>, où elle a pu s'enrichir de différentes réflexions théoriques, méthodologiques et techniques dans le cadre d'une réflexion globale sur les nouvelles approches pour des études littéraires assistées par ordinateur. Cet article recense les premières conclusions du projet (toujours en cours), mais dont il semble d'ores et déjà intéressant de partager les enjeux, les caractéristiques, l'intérêt et les limites. Les premières pistes d'analyse de ces résultats provisoires nous semblent démontrer l'intérêt de développer le *crowdreading* pour les études littéraires et de multiplier les outils et les interfaces afin d'enrichir les théories actuelles.

3

## De la production de données au *crowdreading*

La « première vague » des humanités numériques [Schnapp, Lunenfeld et Presner (s.d.)] s'est principalement consacrée à la création de données, à partir, fondamentalement, de la conversion numérique des objets d'étude des humanités. Dans un second temps, les banques de données et les numérisations massives ont servi à « nourrir » les machines de signes et de caractères, à réaliser des tests et à créer les premiers outils permettant de manipuler les données en question. L'effort réalisé durant ces dix dernières années a donc surtout concerné le développement technique, le renouvellement des méthodes éditoriales, la formation à ces méthodes et aux logiciels, ainsi que l'enrichissement, toujours nécessaire, des bibliothèques. En revanche, la construction de nouvelles analyses littéraires à partir du matériau numérique est toujours balbutiante et peu de résultats sont probants à l'heure actuelle.

4

Ce phénomène peut s'expliquer par plusieurs raisons. Tout d'abord parce que les littéraires ne se sont pas encore, ou très peu, emparés de ces données pour les exploiter de leur point de vue. Il est possible que ce soit par absence d'intérêt et/ou compétence(s) technique(s) mais très souvent, c'est aussi parce qu'ils sont souvent confrontés à des difficultés pour concevoir ou construire des outils d'interprétation « littéraires » à proprement parler. Les recherches menées dans le domaine des études littéraires sont très spécifiques, car l'objet de travail n'est pas seulement le mot ou le groupe de mots – qui, eux, intéressent les linguistes – mais des unités textuelles de longueur variable, aux contours assez flous, qui participent à la construction du sens, de la fiction ; le regard du chercheur en littérature ne porte pas que sur le lexique. Ainsi, si les études littéraires peuvent s'appuyer, jusqu'à un certain point, sur les protocoles mis au point par la textométrie<sup>[4]</sup> ou par d'autres méthodes de text mining<sup>[5]</sup> basées sur l'unité « lexème », ceux-ci montrent vite leurs limites quand il s'agit d'analyser la circulation de la parole dans une pièce de théâtre, les emboîtements des espaces référentiels dans les œuvres de fiction, les structures métriques dans un poème, l'organisation chronologique ou thématique d'une œuvre, etc. Dans le cadre de notre travail, qui concerne la notion de personnage, nous avons eu à nous interroger sur les moyens de capter ces objets de sens, perceptibles au niveau du texte et non des mots, même quand on observe ces derniers dans leur contexte immédiat.

5

Engagées dans différents projets d'édition numérique<sup>[6]</sup> et menant des recherches en littérature assistées par ordinateur, nous voulions, au départ, imaginer les façons d'encoder les personnages afin d'en proposer, dans un second temps, des analyses à partir de l'information ainsi balisée. L'entreprise s'annonçait difficile tant les contours du personnage débordent largement les « entités nommées » pour lesquelles des protocoles de balisage ont été mis au point par la TEI<sup>[7]</sup>. De plus, il nous est apparu très tôt que les difficultés n'étaient pas seulement techniques (identification des balises à employer), mais aussi, et surtout, conceptuelles : est-il possible d'identifier de façon systématique et exhaustive ce qui *fait* un personnage ? Ses caractéristiques profondes ? Comment capter ces caractéristiques par la définition d'entités « balisables » ? En fait, nous sommes rendu compte qu'il nous manquait une description du personnage facile à opérationnaliser.

6

En effet, la littérature scientifique et les théories actuellement disponibles offrent des perspectives globales sur le phénomène ; celles-ci ont nourri notre réflexion, mais n'ont pas apporté de réponse concrète à nos questions. Dans les années 1970, la critique structuraliste a introduit plusieurs concepts utiles, comme ceux de « personnage référentiel », d'« embrayeur » ou d'« anaphore », signalant ainsi l'importance du récepteur dans la co-construction d'un personnage qui est « autant une reconstruction du lecteur, qu'une construction du texte » [Hamon 1977, 119]. À partir de ces idées et concepts, Umberto Eco a ébauché une autre notion importante, celle du *lector in fabula* [Eco 1985, 2006], étayant ainsi, plus largement, le rôle du lecteur dans la fiction<sup>[8]</sup>. Les travaux de James Phelan [Phelan 1989] ont tenté une première catégorisation des personnages littéraires à partir de l'analyse des éléments qui les composent dans leur « sphère » : l'élément mimétique, le synthétique et le thématique. Au début des années 2000, Vincent Jouve s'est proposé d'étudier plus particulièrement ce qu'il a nommé « l'effet-personnage » et les modalités de l'interaction entre texte, lecteur et personnage [Jouve 2002, 27]. Davantage située du côté de la réception, sa proposition a donné lieu à une classification de l'effet-personnage en trois catégories : l'effet-personnel, l'effet-personne et l'effet-prétexte.

7

Toutes ces études et théories proposent des approches pertinentes et éclairantes du personnage, mais restent très « fonctionnalistes » [Weststeijn 2005] car elles considèrent qu'il émerge des mots<sup>[9]</sup>. Surtout, elles reposent sur des intuitions et sur une méthode inductive ; en effet, bien qu'elles fournissent force d'exemples tirés de différents textes pour appuyer les propositions théoriques de leurs auteurs, elles ne proposent toutefois pas d'approche systématique des unités textuelles qui composent le personnage, ni de description, de conceptualisation ou de caractérisation de celles-ci. En d'autres mots, leur approche est largement théorique, alors que nos objectifs requièrent plus de pragmatisme et s'inspirent de l'approche modélisatrice spécifique des humanités numériques [McCarty 2004].

8

C'est la raison pour laquelle nous avons mis en place une expérience de lecture destinée à identifier, sur quelques exemples concrets, les unités textuelles qui concourent, dans l'esprit du lecteur, à la construction du personnage. D'emblée, nous nous sommes interrogées sur la validité d'un modèle ancré dans notre seule lecture, d'autant plus que les premières séances de travail ont immédiatement fait apparaître nos divergences en matière de reconnaissance des personnages et des unités textuelles sélectionnées comme porteuses d'information sur ces objets. Tout en nous attelant à la fabrication d'un modèle partagé, « interne » à l'équipe de recherche dans un premier temps mais destiné à la publication par la suite, et susceptible de donner lieu à une pratique harmonisée d'encodage, il nous a paru important de le confronter à d'autres lectures, différentes, multiples. Celles-ci nous paraissaient être à la fois le moyen de dégager un « consensus » (quelles sont les unités textuelles largement perçues, par différents types de lecteurs, comme « personnophores » ?), et d'éviter la tentation de penser notre travail comme coïncidant avec la « lecture idéale » des œuvres alors qu'elle n'est, à son tour, qu'une lecture parmi d'autres [Iser 1976]. Sans nier l'intérêt de construire un modèle du personnage à partir d'une lecture unifiée de mille œuvres, nous avons ainsi choisi d'en produire un en nous fondant sur mille lectures, nécessairement divergentes, d'un même texte.

9

Les problèmes posés par notre approche sont évidemment très nombreux et sont aussi bien épistémiques que méthodologiques. Si la « lecture distante » effectuée par ordinateur est en principe homogène, tel n'est certainement pas le cas de la « lecture profonde » menée par des individus distincts. Quand nous avons demandé à nos étudiants de répondre de façon individuelle et aussi personnelle possible, en évitant d'échanger entre eux, à un questionnaire portant sur une courte nouvelle de Julio Cortázar, ceci nous a amenée à nous confronter à une hétérogénéité d'autant plus difficile à gérer que, à la différence de nos collègues en sociologie ou psychologie, nous n'avions pas été formées à la réalisation d'enquêtes. Plus largement, il y avait peu d'exemples, dans la littérature, de questionnaires et de travaux de terrain mobilisés pour des objectifs propres aux sciences du texte. À ceci s'est ajouté le constat qu'il existait peu d'outils nous permettant d'accompagner la lecture et de capter la façon dont elle fabrique du sens, et encore moins d'études quant à la façon dont ces outils, lorsqu'ils existent, interfèrent avec ce processus.

10

## Des « lectures partagées » : un tour d'horizon

En dépit des difficultés et questionnements dont le précédent paragraphe ne donne qu'un bref aperçu, notre approche a été stimulée par le fait qu'un certain nombre de projets a déjà intégré et posé le cadre d'une telle « lecture à plusieurs ». Il en est ainsi du projet ENCCRE<sup>[10]</sup> par exemple, qui réunit une équipe internationale, pluridisciplinaire, dont l'objet d'étude est l'*Encyclopédie* de Diderot et d'Alembert. Conformément à la

11

déclaration figurant sur la page d'accueil de son site web, la participation au projet est ouverte « à toute personne ayant déjà travaillé, travaillant ou désireuse de travailler sur l'*Encyclopédie* ». Cependant, à en juger par les extraits qui figurent actuellement sur le site, ENCCRE vise moins le *crowdreading* que la construction d'un texte complet et la proposition d'UNE version annotée de l'*Encyclopédie*. La production et la conjugaison d'interprétations divergentes ne sont donc pas, au moins dans cette première phase, les priorités de l'équipe impliquée dans ENCCRE. Celle-ci est d'ailleurs essentiellement composée de spécialistes de Diderot, de l'*Encyclopédie* et du XVIII<sup>e</sup> siècle, à la différence d'autres projets de crowdsourcing dans lesquels on constate un ratio inversé entre « gens de métier » et « large public ».

Plus proche de notre projet, l'édition enrichie en ligne de *Candide ou l'optimisme*<sup>[11]</sup> que la Bibliothèque nationale de France, la Voltaire Foundation et Orange ont réalisée en 2012, est allée dans le sens du *crowdreading* tel que nous voulons le définir et l'exploiter, à ceci près que l'édition a préparé les « parcours de lecture » et n'a pas eu pour dessein l'observation de différentes lectures telle que nous entendons la pratiquer dans le cadre de notre expérience. Les « chemins » de lecture de *Candide ou l'optimisme* sont déjà tracés et limités en nombre grâce à une application informatique qui permet de « naviguer » dans l'œuvre à partir des quelques choix préétablis. Pensée pour des lectures multi-écrans (tablettes, téléphones, ordinateurs), elle offre néanmoins un environnement élégant, ludique, associant plaisirs de l'audition et de la vue et conduisant le lecteur, quel que soit son âge, à (re)découvrir l'histoire de *Candide* grâce aux apports critiques des spécialistes de *Candide* et de Voltaire, ouvrant ainsi la lecture à des champs thématiques connexes à l'œuvre. Le déploiement des « mondes » du texte à partir des résultats de différentes lectures est une ambition fréquente des éditions dites « enrichies » ou « augmentées ». Rappelons néanmoins que, dans ces cadres-là, l'enrichissement est plutôt synonyme de construction par l'éditeur – et non par les lecteurs – des cheminements dans l'œuvre.

12

Plus près encore de nos intérêts scientifiques, le projet (Dis)similitudes<sup>[12]</sup>, issu d'un partenariat entre le Labex OBVIL et le laboratoire LIP6, a proposé une expérience proche de ce que nous appelons *crowdreading*. Lancée en août 2016, l'expérimentation vise à exploiter la lecture collaborative en incitant les lecteurs à identifier « des comparaisons figuratives dans les textes littéraires<sup>[13]</sup> ». Le but du projet est d'affiner les potentialités d'un outil informatique qui, pour une fois, semble s'inscrire dans la perspective sémasiologique que nous prônons. L'expérience mise sur la participation d'un lecteur-contributeur-crowdsourcer invité à repérer, dans une sélection de textes en prose, des « comparaisons figuratives dans lesquelles le comparant est un nom commun<sup>[14]</sup> ». Nous suivons les résultats de cette expérience et testerons l'outil développé par la suite ; même s'il est encore difficile de tirer des conclusions de cette proposition, nous remarquons que la contribution demandée au lecteur dépasse le cadre de la simple transcription ou de l'annotation, comme c'est souvent le cas dans les projets faisant appel au crowdsourcing. Ainsi, dans une certaine mesure, et bien qu'elle soit fortement restreinte par le cadrage du projet et par le paramétrage du logiciel, (Dis)similitudes cherche à restituer une lecture dans sa dimension interprétative, une voie que nous exploitons également dans notre projet.

13

De façon plus large, nous avons pu remarquer qu'un certain nombre de propositions et de travaux récents de chercheurs en études littéraires s'engagent dans la révision conceptuelle par l'intégration – et c'est une nouveauté dans ce domaine des sciences du texte – d'une approche expérimentale de la littérature assistée par l'ordinateur. Ces expériences, dont il a été question lors de l'atelier annuel de CAHIER (2016<sup>[15]</sup>) et lors de la conférence Digital Humanities à Cracovie (2016<sup>[16]</sup>), tendent au renouvellement de l'herméneutique traditionnellement solitaire des études de lettres : dorénavant, elles sont invitées à adopter d'autres modalités de production du savoir, à la fois ouvertes et partagées, encouragées et facilitées par le numérique.

14

Ces expérimentations et réalisations témoignent d'un intérêt de plus en plus répandu dans le monde des littéraires pour le recours à l'informatique et surtout, pour la conception d'outils et de méthodes facilitant la multiplicité des lectures et les analyses de ces dernières.

15

## **Le *crowdreading* au service de la caractérisation du personnage**

L'expérimentation ici présentée a été menée entre le printemps et l'automne 2016. Elle s'est déroulée en deux étapes (mars-avril 2016, puis septembre-novembre 2016) et a impliqué deux questionnaires différents, dont l'élaboration est décrite ci-après. Les résultats présentés dans cet article proviennent aussi bien de la première

16

que de la seconde enquête, pour les questions ayant été reconduites à l'identique ; en revanche, nous laissons de côté les réponses aux questions que nous avons modifiées entre les deux moutures de l'enquête, et pour lesquelles nous testons encore des formulations et comparons des résultats.

Lors de la première, comme de la seconde vague, la population-cible était formée d'étudiants en études littéraires : licence de lettres modernes et licence d'espagnol. Cette population a été choisie parce qu'elle nous est apparue susceptible de mener des « lectures littéraires » plus élaborées que la moyenne des lecteurs (dans le sens donné à ce syntagme par [Dufays, Gemenne et Ledur 2005]), tout en étant encore loin de la « lecture savante » du théoricien ou de l'historien de la littérature. Leur approche du texte, plutôt « ascendante » que « descendante » [Dufays 2013], nous paraissait par ailleurs adaptée à l'identification des « unités textuelles » du personnage que nous cherchions à identifier à des fins de modélisation.

17

Lors de la première phase, dans le but de recueillir les lectures de nos étudiants, nous leur avons soumis trois nouvelles, de longueurs différentes : *Le Horla* (1887) de Guy de Maupassant, *Les Xipéhuz* (1888) de J. H. Rosny Aîné et *Continuité des parcs* (1964) de Julio Cortázar. Cette dernière a été proposée dans la langue originale, l'espagnol, mais également en traduction française ou anglaise. Ces textes présentent différents types de personnages et posent problème en ce qui concerne au moins une dimension des protagonistes : certains sont non humains, d'autres imaginaires, ou situés sur un plan de la fiction difficile à définir. Dans tous les cas, elles offrent matière à la fois à l'observation de l'effet-personne et de l'effet-personnel que nous pensions, au moins dans un premier temps, pouvoir baliser et explorer.

18

Pour concevoir notre expérience, nous avons nous-mêmes tenté de repérer et d'analyser les éléments qui nous semblaient pouvoir participer à la construction des personnages de ces nouvelles. Sur la base de nos considérations, nous avons construit un questionnaire (voir annexe n°1) destiné à identifier quels personnages nos répondants observent dans les textes donnés et selon quels critères. Ils devaient également signaler les unités textuelles qui déterminent cette identification, et que nous souhaitions analyser, en observant si elles se situent à l'échelle du mot, du groupe de mots ou de la phrase, s'il s'agit plutôt de substantifs, d'adjectifs, de verbes, etc.

19

Le questionnaire a été distribué, dans un premier temps, à un seul groupe d'étudiants de L1 d'une même université, via un formulaire Google<sup>[17]</sup>. En effet, même si la population ciblée était suffisamment restreinte pour que, dans ce premier cas, le recueil des informations soit possible sous la forme d'une enquête papier, nous voulions mettre au banc d'essai les outils dont nous disposions en vue de la réalisation d'une enquête de plus grande envergure, qui avoisinerait, sinon les mille lectures plus d'une fois évoquées dans ce texte, du moins une centaine ou, idéalement, plusieurs centaines. Outre des réponses que les étudiants soumettaient en ligne, le questionnaire leur demandait par ailleurs d'utiliser la fonction « pinceau » de leur lecteur PDF, afin de souligner toutes les unités textuelles participant à la construction de leur propre image des personnages. Après avoir enregistré les réponses aux questions, ils devaient nous envoyer par mail les fichiers PDF surlignés. Seuls des étudiants volontaires étaient sollicités pour cet envoi forcément non anonyme.

20

Dans cette première phase de l'expérience, le questionnaire a été soumis à une soixantaine de répondants et a recueilli trente-quatre réponses, dont vingt-huit se sont révélées exploitables. Le plus souvent, c'est la nouvelle de Julio Cortázar, *Continuité des parcs* (1964) qui a été choisie par les étudiants, assurément en raison de sa brièveté<sup>[18]</sup>. Les quelques réponses données au sujet du *Horla* et des *Xipéhuz* ont donc été laissées de côté comme étant numériquement peu significatives. Le projet d'interroger les étudiants sur plusieurs textes s'est ainsi avéré trop ambitieux.

21

La seconde phase de l'expérience a été ouverte aux étudiants en études littéraires de nos trois universités. Elle se fondait sur un questionnaire remanié (voir annexe 2) et se concentrait uniquement sur la nouvelle *Continuité des parcs*, pour les raisons exposées plus haut<sup>[19]</sup>. Elle a permis de recueillir cent-treize réponses, dont cent quatre se sont révélées exploitables. Quoique les deux enquêtes aient été menées à partir de questionnaires différents, bon nombre de résultats de cette seconde phase restent comparables à la première, et donc exploitables.

22

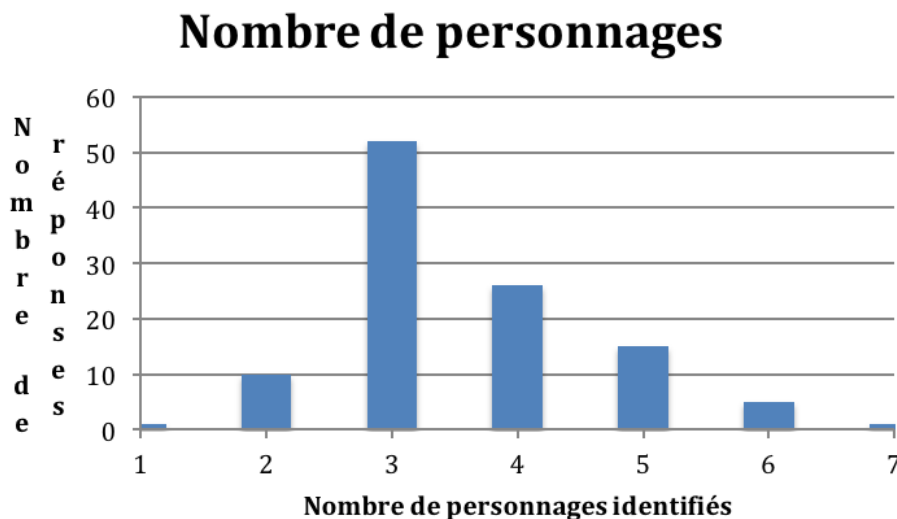
Si la première vague de réponses nous a permis, comme escompté, d'affiner, de revoir et de préciser nos questions, ainsi que la présentation des réponses (majuscules, nombres en chiffres plutôt qu'en lettres, etc.), elle a également fait ressortir les limites des outils utilisés. Googleform et l'export des résultats dans un tableur

23

Excel sont assez efficaces, mais le va-et-vient que nous avons imposé à nos répondants entre le formulaire et le fichier PDF à surligner s'est révélé peu ergonomique. Par ailleurs, il nous est apparu que le surlignage de PDF était loin de constituer une formule idéale ; lorsqu'il s'agit de souligner des termes offrant des informations différentes sur les personnages, mais situés à proximité dans le texte (même ligne, ou même paragraphe), la tentation est grande de les traiter « en bloc », cela sans compter avec la tendance du logiciel à « fusionner » les blocs de couleur proches, même quand on essaye de les séparer par des blancs. Toutefois, en nous demandant quel outil alternatif nous aurions pu utiliser dans le cadre de cette étude (un outil qui permettrait d'observer les différentes stratégies de lecture tout en facilitant leur interprétation semi-automatique), nous avons constaté qu'une telle interface n'existe pas (encore), ou du moins que nous n'en avons pas connaissance. C'est pourquoi nous avons dû nous résoudre à poursuivre avec des fichiers PDF, en attendant la possibilité de créer, avec l'appui d'un ingénieur d'études informaticien, l'outil dont nous aurions besoin. En revanche, afin d'obtenir une image plus claire des unités de sens sélectionnées par les étudiants, nous avons demandé à la seconde vague de répondants de recopier également les syntagmes pertinents dans leurs réponses en ligne. Ceci a toutefois l'inconvénient de réduire le nombre d'unités textuelles retenues : certains répondants ne recopient qu'un ou deux syntagmes par personnage, prennent des raccourcis<sup>[20]</sup>, alors même qu'ils surlignent, en moyenne, beaucoup plus d'éléments par paragraphe (v. *infra*).

La première remarque que l'on peut faire, comme le montre le graphique ci-dessous, est que le nombre de personnages identifiés varie assez considérablement à l'intérieur du groupe de répondants, même si une majorité d'entre eux en indique trois.

24



**Figure 1.** Graphique fondé sur les résultats fusionnés des deux vagues d'interrogation des répondants

La lecture des réponses rédigées montre que les étudiants interprètent de façon restrictive la consigne de compter les personnages, seuls certains éléments du texte étant jugés dignes d'être retenus. Ainsi, l'homme et la femme sont mentionnés dans toutes les réponses et, de façon presque aussi fréquente, on relève la présence de l'homme qui lit, assis dans un fauteuil<sup>[21]</sup>. En revanche, l'intendant (que certains désignent comme « le majordome », selon l'appellation qu'il a dans le texte en espagnol) n'est mentionné que par un quart des répondants, tandis que le « fondé de pouvoir » (ou « l'agent d'affaires ») n'est signalé que par moins d'un sixième. Par ailleurs, peu d'étudiants font l'hypothèse d'une différence entre le « il » du premier paragraphe et le mari de la femme adultère, ou, si l'on veut, entre le lecteur de la nouvelle et le lecteur dans la nouvelle<sup>[22]</sup>. Un seul répondant indique les « chiens » comme personnages potentiels.

25

Il est intéressant d'observer que les surlignements réalisés dans les fichiers PDF montrent que certains étudiants identifient « l'intendant » sans le mentionner dans la liste des personnages. Parfois, ils ne surlignent le substantif qu'à l'occasion de sa seconde « apparition », au paragraphe 3, alors que le personnage « n'est pas là », comme nous le dit le narrateur ; sans doute est-ce la position syntaxique du substantif, en fonction sujet, qui explique cette visibilité accrue du personnage dans le troisième paragraphe. À l'inverse, alors même

26

qu'ils l'ont identifié au paragraphe 1 et qu'ils n'hésitent pas à relever d'autres éléments redondants, plusieurs étudiants ne le surlignent pas à cet endroit.

De même, plusieurs surlignent le « fondé de pouvoir », sans nécessairement l'inclure dans la liste des personnages, tandis que la plupart observe « le poignard » qui devient « tiède au contact de la poitrine » alors que d'autres personnifications ne sont pas relevées, comme « le souffle du crépuscule [qui] semblait danser sous les chênes ». Cette différence entre les éléments recopiés dans le formulaire et les surlignements suggère que les lecteurs construisent un large spectre de « personnages potentiels » au cours de la lecture, mais que seules certaines possibilités sont jugées dignes de recevoir, au final, l'étiquette de « personnage », en fonction de critères et de seuils que notre enquête ne permet pas, dans son état actuel, de capter. Cette expérience nous suggère que plutôt que d'encoder « des personnages », qui sont autant de constructions ancrées dans des lectures individuelles, plus ou moins originales ou, au contraire, consensuelles, notre objectif doit être d'identifier des éléments « personnophores », quitte à ne se pencher que dans un second temps sur les critères supplémentaires grâce auxquels les lecteurs perçoivent un personnage « à part entière ».

27

Quelles sont les unités que les étudiants retiennent comme porteuses d'une information sur le personnage ? Pratiquement toute suite de mots du texte a vocation à participer, pour un lecteur ou pour un autre, à la construction d'un personnage. Hormis la phrase sur les chiens<sup>[23]</sup>, seuls les segments de phrases « la double et implacable répétition était à peine interrompue le temps qu'une main frôle une joue », « Il commençait à faire nuit » et « à la fin » ne sont relevés par aucun répondant. Cependant, certaines unités sont, comme on pouvait s'y attendre, plus fréquemment relevées que d'autres, et c'est sur celles-ci que nous nous sommes concentrées pour notre réflexion sur la modélisation du personnage.

28

Avant d'en proposer une analyse, observons toutefois que la délimitation de ces unités n'est pas allée de soi, tant leurs contours varient d'une réponse à une autre. On y trouve aussi bien des mots isolés, que des syntagmes, des phrases ou des paragraphes entiers. Ainsi, dans une phrase comme « il fut ainsi témoin de la dernière rencontre dans la cabane parmi la broussaille », certains étudiants surlignent uniquement le mot « rencontre », d'autres « dernière rencontre », certains « témoin de la dernière rencontre » et d'autres l'ensemble de la phrase. Devant cette hétérogénéité des pratiques à l'intérieur du groupe test, mais aussi chez chaque répondant, nous nous sommes demandé si, pour l'étape suivante de l'enquête, il ne serait pas pertinent de proposer aux répondants un découpage préétabli en mots ou en syntagmes nominaux/verbaux. La soumission d'un tableau dans lesquels les répondants cocheraient les cases de leur choix que nous aurions délibérément construites sous formes d'unités de base, plus ou moins grandes, au lieu d'un fichier PDF, serait-elle plus efficace ? Il nous est toutefois apparu que ceci reviendrait à introduire un filtre dans la perception du « spectre personnophore » évoqué plus haut, et à pré-baliser des chemins de lecture, comme cela a été fait par le projet « Candide ». Certes, cela présente des avantages en termes de rigueur, mais un tel découpage, pré-imposé, rend moins bien compte des pratiques individuelles de lecture que nous voulons capter et analyser pour comprendre ce qui « fait » personnage dans une perspective individuelle ou une autre. Aussi avons-nous décidé de continuer à laisser les répondants nous signaler librement les unités de sens de leur choix et de poursuivre les recherches dans le but de trouver les moyens de systématiser la lecture des réponses, à des fins d'analyse.

29

Dans la phrase « il fut ainsi témoin de la dernière rencontre dans la cabane parmi la broussaille », il apparaît que le terme « rencontre » figure invariablement dans tous les soulignements, alors que « cabane », « broussaille » ou le verbe « assista » ne sont jamais soulignés seuls, mais avec l'ensemble de la phrase. Nous retenons dès lors « rencontre » comme pivot d'une unité de sens qui peut prendre de multiples formes : « Il fut ainsi témoin de la dernière rencontre », « témoin de la dernière rencontre », « rencontre dans la cabane », « témoin de la dernière rencontre dans la cabane parmi la broussaille », etc. En revanche, « il se retourna un instant pour la voir courir, les cheveux dénoués » semble formé de quatre unités participant à la construction de personnages différents : « il », « se retourna (un instant) », « pour la voir courir », « les cheveux dénoués ». Ces quatre unités sont plus ou moins systématiquement séparées par les étudiants lors des soulignements, à en juger par les espaces qu'ils laissent entre les différents blocs de couleur.

30

En confrontant les différentes pratiques, on parvient à proposer un découpage virtuel de la nouvelle de Julio Cortázar en 121 unités textuelles « personnophores » – chacune ayant été désignée comme telle par un répondant ou un autre (voir annexe n° 4). Parmi ces 121 unités possibles, les répondants surlignent entre 4 et 63 unités de sens, avec une moyenne de 27 unités relevées. La plupart de ces unités concernent le

31



personnage « lecteur » (47/121), avec une ambiguïté intéressante des réponses puisque nos répondants ont tendance à identifier, comme on l'a vu plus haut, le lecteur du premier paragraphe avec celui du dernier paragraphe, dont la tête dépasse du fauteuil vert et qui sera la victime potentielle de l'amant. Le plus grand nombre d'éléments de sens est surligné dans le second paragraphe : en moyenne, 10,5 unités – à comparer avec les 9,5 unités relevées au premier paragraphe, et avec les 6,9 du troisième paragraphe : ceci paraît justifié par la structure même de la nouvelle qui introduit deux nouveaux personnages au paragraphe 2 (l'homme et la femme), alors que rien de tel n'arrive au paragraphe 3, sauf si l'on considère que la victime potentielle de l'amant, cet homme assis dans un fauteuil dont on ne voit que la tête, est un nouveau personnage, distinct du lecteur dont on entrevoit la présence au premier paragraphe. Cette hypothèse, qui témoigne de la complexité de la nouvelle, a été formulée par quelques rares lecteurs, comme on a pu le voir plus haut. De fait, au paragraphe 3, la plupart des annotations se concentrent sur la deuxième partie du paragraphe, à partir de l'entrée de l'amant assassin dans la maison ; les observations concernant la disposition des pièces, l'ameublement, etc. semblent avoir été vues comme des indications sur l'identité de la victime. Nous avons envisagé plusieurs hypothèses pour expliquer ce comportement : ce nombre plus réduit d'unités de sens relevées au troisième paragraphe traduit probablement une certaine « lassitude » des répondants même si cette hypothèse est quelque peu invalidée par la brièveté du texte à analyser ; ou à l'inverse, un déplacement de l'attention vers l'action qui devient, à ce moment-là, plus haletante.

Comme on pouvait s'y attendre, les éléments relevés le plus systématiquement sont les désignateurs qui servent à construire une première identité des personnages : « il », « l'homme », « la femme ». D'ailleurs, comme pour éviter la confusion, certains étudiants ne surlignent que le premier « il », tandis que d'autres surlignent les « il » se référant au propriétaire (surtout dans le premier paragraphe, donc), mais non ceux qui renvoient à l'amant. De la même manière, seul un étudiant sur deux surligne le pronom « elle », alors qu'aucun autre ne lui fait concurrence dans le texte ; il en est de même de « lui ». Encore moins nombreux sont les étudiants qui relèvent « ils », ou bien « les protagonistes », traduisant la difficulté de traiter ces désignations collectives : doit-on considérer qu'elles créent un nouveau personnage, ou bien qu'elles « renforcent » la perception de chaque personnage individuel auquel elles réfèrent ?

32

Presque aussi fréquemment que les désignateurs sont relevés « le visage griffé », « méfiante » et « la tête de l'homme en train de lire un roman » par les étudiants ; ces groupes nominaux, et cet adjectif, permettent de préciser une image, de fixer chaque personnage dans une posture significative. Tous les éléments se rapportant au corps sont ainsi, de façon assez systématique, notés : « le corps de l'amant », « les cheveux dénoués » ou les « cigarettes à la portée de [la] main » du propriétaire-lecteur. Cependant, adjectifs et autres allusions au corps ne font pas non plus l'objet d'un relevé exhaustif, comme si certains avaient été perçus comme peu significatifs (la « main qui frôle une joue », par exemple). La question qui se pose alors concerne la façon d'encoder cette intensité « caractérisante » des éléments, qui ne participent clairement pas à part égale à la construction de l'image du personnage. Plus particulièrement, on peut se demander si cette attention aux particularités physiques se retrouverait à propos d'autres textes, ou bien si elle s'avère plus élevée ici en raison du silence du récit quant aux noms des personnages. Il sera donc pertinent, à l'avenir, de mener une enquête plus large sur différents types de récits pour voir si ce phénomène est également observable.

33

L'une des surprises du test a surtout été de constater que pratiquement tous les étudiants soulignent des verbes d'action comme participant à la construction des personnages : « étanchait », « se dérobait », « (pour la voir) courir », « lire », « entra la première », « reprit sa lecture », etc. Ces scores élevés sont d'autant plus étonnants qu'on les attendait plutôt *a priori* à propos d'expressions comme « jouissait du plaisir » ou « tout en demeurant conscient », qui nous paraissent esquisser une certaine psychologie du personnage principal, le lecteur, mais non à propos d'unités textuelles indiquant ce que les personnages font, plutôt que ce qu'ils sont. En fin de compte, il s'avère que les éléments concernant les sentiments ou les valeurs des personnages, sur lesquels nous pensions initialement diriger nos efforts d'encodage et d'extraction, sont bien moins fréquemment relevés que les verbes d'action : un personnage se définit avant tout, pour nos répondants, par son action.

34

Tout comme la diversité des éléments textuels surlignés, l'analyse des réponses à la question 5, qui demandait à nos répondants de caractériser le personnage de « il » en dix mots, fait pleinement ressortir la variété des lectures, et l'intérêt qu'il y a à pousser plus loin l'exploration, avec les outils du numérique, de cette diversité perceptive<sup>[24]</sup>. En effet, à part trois caractéristiques qui sont mentionnées systématiquement (« il » est « passionné », « homme » et « riche »), le reste du vocabulaire utilisé pour sa description est caractérisé par un

35

haut degré de dispersion, même si dans certains cas on peut saisir des affinités conceptuelles entre les termes employés par les répondants : ainsi, ils sont nombreux à le définir, par exemple, comme « propriétaire », « riche », « aisé », « travaille dans la finance », etc. Il est tout à fait clair que « il » a de multiples visages, presque autant que de lecteurs. Plus encore, on constate un véritable décalage entre les portraits proposés et les éléments relevés dans le texte, comme si pour construire leur image du personnage, nos répondants prenaient véritablement de la distance par rapport au texte et bâtissaient leurs propres représentations par extrapolation<sup>[25]</sup>, à partir d'implications souvent surprenantes<sup>[26]</sup>.

## Conclusions et perspectives

Au terme de ces travaux, un premier bilan est possible, en prenant acte des limites du projet dans son état actuel. S'il faut sans doute envisager d'élargir l'enquête, afin d'augmenter le nombre de réponses exploitables, ce que nous avons exposé plus haut nous conforte dans notre volonté de repenser l'analyse littéraire, de façon à ne pas chercher à comprendre et à interpréter uniquement « l'aventure d'une écriture » [Ricardou 1967, 111], mais bel et bien celle d'une écriture et d'une lecture, ou plutôt, pour revenir à notre proposition introductive, de mille lectures. Le *crowdreading* tel que nous l'avons pratiqué, en essayant de le mobiliser pour identifier les éléments textuels personnophores, susceptibles d'être encodés de façon systématique, permet de s'attaquer à deux chantiers qu'il était difficile d'envisager jadis à une grande échelle : d'un côté la possibilité de tirer parti d'une multiplicité d'approches, de points de vue et de sensibilités autour d'un même matériau littéraire et, de l'autre, celle d'approcher, sur de nouvelles bases, le phénomène esthétique, en fournissant de nouvelles catégories d'analyse. Notre démarche, qui rejoint dans une certaine mesure d'autres travaux en cours sur la création de « communs » en lien avec le développement des cultures de l'informatique/du numérique [Merzeau 2014], ouvre une double réflexion : d'une part, sur la possibilité d'apporter des contributions complémentaires au processus d'actualisation, ou de (re)construction, de certaines notions de théorie littéraire à l'aune de la révolution numérique en cours, et, d'autre part, sur les moyens d'informatiser les approches sémasiologiques et herméneutiques qui sont au cœur des études littéraires.

36

Au vu de nos résultats, encore partiels, et des quelques hypothèses stimulantes qu'ils suscitent d'ores et déjà (et qui seront à vérifier), il nous faut maintenant envisager de repenser le questionnaire et les modalités de réponse de façon plus méthodique. L'expertise des spécialistes de science de l'éducation, de sociologues et de psychologues sera recherchée dans cette troisième phase. En tout état de cause, les cent-trente réponses déjà obtenues tendent à prouver que la construction du personnage par les lecteurs est d'abord le résultat d'une sensibilité marquée à l'effet-personne, lequel est visiblement aisé à catégoriser. Il sera dès lors important de repenser l'enquête pour permettre aux lecteurs d'identifier et étiqueter plus facilement l'effet-personnel et l'effet-prétexte. Il nous faut également concevoir un outil de recension des réponses performant, en ligne, de façon à élargir l'enquête à d'autres textes pour enrichir nos analyses sur ce qui « fait » personnage aux yeux des lecteurs, ainsi que sur l'analyse des « scénarios communs » définis par Eco et auxquels se superposent, plus ou moins exactement, ce que l'on pourrait identifier comme des « scénarios individuels ». Enfin, il nous semble très enrichissant de demander aux lecteurs d'accompagner leurs réponses de documents iconiques et/ou sonores qui leur semblent pouvoir compléter leur idée du personnage<sup>[27]</sup> ; cela permettrait même éventuellement de créer une banque de données commune dans laquelle les lecteurs pourraient venir sélectionner des éléments afin d'attribuer des caractéristiques non textuelles aux personnages des récits qu'ils lisent en ligne. Mais ceci est une autre histoire – à venir...

37

## Annexes

### Annexe 1.

Questionnaire 1

38

TITRE DE LA NOUVELLE : .....

39

PARTIE 1

40

1. Combien de personnages y a-t-il dans l'œuvre?

.....

2. Pouvez-vous les énumérer?

.....

3. Pouvez-vous surligner dans le texte des éléments qui vous permettent d'identifier et d'imaginer chacun de ces personnages. Utilisez les pinceaux (outils du .pdf) et pensez à nous renvoyer le .pdf. En cas de difficultés, imprimez le texte !

4. Pouvez-vous regrouper les caractéristiques et identifier des groupes de personnages sur cette base?

.....

5. Maintenant, nous allons nous intéresser à un personnage en particulier. Les questions, selon le texte concerné, portent sur *Il* (chez Cortázar), *le Horla* (chez Maupassant) ou *Bakhoûn* (chez Rosny)

Pouvez-vous le définir en 10 mots?

44

Pouvez-vous classer ces mots en catégories?

45

.....

46

PARTIE 2 : Informations sur le lecteur/la lectrice

47

1. Vous êtes : un homme/une femme ?

.....

2. Votre âge

.....

3. Votre catégorie socio-professionnelle

.....

4. Vos habitudes de lecture

Nombre de livres lus par an

51

Genres (Roman, BD, Poésie, Théâtre, Documentaire, Science-Fiction, Témoignage, Autobiographie, Aventure, Amour, Fantasy, Psychologie, autres).....

52

## Annexe 2

Questionnaire 2

53

Chers étudiants, chers lecteurs,

54

Dans le cadre d'un projet de recherche portant sur la conceptualisation des personnages littéraires, nous vous invitons à compléter le questionnaire suivant, après avoir lu la nouvelle «Continuité des parcs»/«Continuidad de los parques» de Julio Cortázar. Vous pouvez la télécharger ici >>> <https://mon-partage.fr/f/z0ftZXk3/>

55

Nous tenons à vous préciser que les données seront traitées de façon anonyme. Votre participation à ce questionnaire vaut autorisation d'utilisation de vos réponses anonymisées, pour les fins du projet de recherche.

56

Nous vous remercions par avance pour votre contribution !

57

Ioana Galleron (Université de Lorient), Fatiha Idmhand (Université de Poitiers), Cécile Meynard (Université d'Angers)

58

1° Combien de personnages repérez-vous dans l'œuvre (saisissez un chiffre, s'il vous plaît ; par ex. : « 23 personnages », écrivez « 23 »).

2° Pouvez-vous les énumérer ?

3° Pouvez-vous recopier ici les éléments qui vous permettent d'identifier chacun de ces personnages ? Indiquer, entre parenthèses, le ou les personnage(s) au(x)quel(s) renvoie chaque élément.

4° Pouvez-vous regrouper ces éléments en catégories ? Indiquez, entre parenthèses, les

personnages auxquels vous pensez.

Maintenant, nous allons nous intéresser à un personnage en particulier : « Il », le lecteur de Julio Cortazar.

59

- 5° Pouvez-vous le définir en 10 mots ?
- 6° Pouvez-vous classer ces mots en catégories, en groupes ?
- 7° Comment imaginez-vous le personnage « il » (caractérisation psychologique, physique, comportementale...) ? Décrivez-le en quelques lignes.
- 8° En lien avec la question précédente, avez-vous une (des) image(s) ou un/ des son(s) en tête ? Si c'est le cas, copiez et collez ici le(s) lien(s) URL vers cette (ces) image(s) ou ce(s) son(s).

Informations sur le lecteur/ la lectrice

60

10° Vous êtes : un homme/ une femme ?

11° Votre âge (saisissez un chiffre s'il vous plaît).

12° Votre université.

13° Quelle est votre formation et votre année d'études ?

14° Nombre de livres lus par an.

15° Genre (roman, BD, poésie, théâtre, documentaire, science-fiction, témoignage, autobiographie, aventure, amour, fantasy, psychologie, autres...)

### Annexe 3

Julio Cortazar, « Continuidad de los Parques », *Fin d'un jeu* (1956), traduit de l'espagnol par C. et R. Caillois, Gallimard, 1963.

61

Il avait commencé à lire le roman quelques jours auparavant. Il l'abandonna à cause d'affaires urgentes et l'ouvrit de nouveau dans le train, en retournant à sa propriété. Il se laissait lentement intéresser par l'intrigue et le caractère des personnages. Ce soir-là, après avoir écrit une lettre à son fondé de pouvoir et discuté avec l'intendant une question de métayage, il reprit sa lecture dans la tranquillité du studio, d'où la vue s'étendait sur le parc planté de chênes. Installé dans son fauteuil favori, le dos à la porte pour ne pas être gêné par une irritante possibilité de dérangements divers, il laissait sa main gauche caresser de temps en temps le velours vert. Il se mit à lire les derniers chapitres. Sa mémoire retenait sans effort les noms et l'apparence des héros. L'illusion romanesque le prit presque aussitôt. Il jouissait du plaisir presque pervers de s'éloigner petit à petit, ligne après ligne, de ce qui l'entourait, tout en demeurant conscient que sa tête reposait commodément sur le velours du dossier élevé, que les cigarettes restaient à portée de sa main et qu'au-delà des grandes fenêtres le souffle du crépuscule semblait danser sous les chênes.

62

Phrase après phrase, absorbé par la sordide alternative où se débattaient les protagonistes, il se laissait prendre aux images qui s'organisaient et acquéraient progressivement couleur et vie. Il fut ainsi témoin de la dernière rencontre dans la cabane parmi la broussaille. La femme entra la première, méfiante. Puis vint l'homme le visage griffé par les épines d'une branche. Admirablement, elle étanchait de ses baisers le sang des égratignures. Lui, se déroba aux caresses. Il n'était pas venu pour répéter le cérémonial d'une passion clandestine protégée par un monde de feuilles sèches et de sentiers furtifs. Le poignard devenait tiède au contact de sa poitrine. Dessous, au rythme du cœur, battait la liberté convoitée. Un dialogue haletant se déroulait au long des pages comme un fleuve de reptiles, et l'on sentait que tout était décidé depuis toujours. Jusqu'à ces caresses qui enveloppaient le corps de l'amant comme pour le retenir et le dissuader, dessinaient abominablement les contours de l'autre corps, qu'il était nécessaire d'abattre. Rien n'avait été oublié: alibis, hasards, erreurs possibles. À partir de cette heure, chaque instant avait son usage minutieusement calculé. La double et implacable répétition était à peine interrompue le temps qu'une main frôle une joue. Il commençait à faire nuit.

63

Sans se regarder, étroitement liés à la tâche qui les attendait, ils se séparèrent à la porte de la cabane. Elle devait suivre le sentier qui allait vers le nord. Sur le sentier opposé, il se retourna un instant pour la voir courir, les cheveux dénoués. À son tour, il se mit à courir, se courbant sous les arbres et les haies. À la fin, il distingua dans la brume mauve du crépuscule l'allée qui conduisait à la maison. Les chiens ne devaient pas aboyer et ils n'aboyèrent pas. À cette heure, l'intendant ne devait pas être là et il n'était pas là. Il monta les trois marches du

64

perron et entra. À travers le sang qui bourdonnait dans ses oreilles, lui parvenaient encore les paroles de la femme. D'abord une salle bleue, puis un corridor, puis un escalier avec un tapis. En haut, deux portes. Personne dans la première pièce, personne dans la seconde. La porte du salon, et alors, le poignard en main, les lumières des grandes baies, le dossier élevé du fauteuil de velours vert et, dépassant le fauteuil, la tête de l'homme en train de lire un roman.

#### **Annexe 4. Unités de sens participant à la construction des personnages**

(par ordre décroissant de fréquence dans les soulignements ; seuls les résultats du premier groupe test sont reproduits ici)

65
----

Il	24
(Puis vint) l'homme	24
La femme	23
le visage griffé (...branche)	22
méfiante	20
(le corps de l')amant	16
(la tête de) l'homme en train de lire	16
les cheveux dénoués	15
elle	12
étanchait... (le sang...égratignures)	12
que sa tête...élevé	11
(que les) cigarettes (...main)	11
Lui	11
se dérobaux caresses	11
pour la voir courir	11
(avait commencé à) lire (le roman...auparavant)	10
avec l'intendant	10
il	9
jouissait (du plaisir presque pervers)	9
entra la première	9
ils	9
Il	8
reprit sa lecture	8
dans la tranquillité...chênes	8
Il	8
se séparèrent à la porte (de la cabane)	8
il	8
il	8
l'intendant	8
Installé dans son fauteil favori	7
il	7
laissait ... caresser..vert	7
se mit à lire les derniers chapitres	7
de s'éloigner (petit...l'entourait)	7
Le poignard	7
devenait tiède...poitrine	7
se retourna un instant	7
se mit à courir	7
Il	6
se laissait... Intéresser...personnages	6
Il	6
(fut...) rencontre (...broussaille)	6
Elle	6
à son fondé de pouvoir	5

Sa mémoire retenait sans effort	5
Dessous...liberté convoitée	5
(Jusqu'à...) ces caresses qui	5
enveloppaient	5
comme pour le retenir	5
et le dissuader	5
devait suivre le sentier	5
Sur le sentier opposé	5
d'affaires urgentes	4
après avoir écrit une lettre	4
et discuté	4
une question de métayage	4
le dos à la porte	4
pour ne pas être gêné	4
par une irritante...divers	4
tout en demeurant conscient	4
les protagonistes	4
Il	4
n'était pas venu pour répéter	4
Un dialogue(.. se déroulait..pages)	4
les contours de l'autre corps(...abattre)	4
qui allait vers le nord	4
le poignard en main	4
et, dépassant le fauteuil,	4
les noms et l'apparence des héros	3
(et qu'au-delà...) souffle du crépuscule(...chênes)	3
absorbé	3
par la sordide...se débattaient	3
Admirablement	3
le cérémonial...furtifs	3
dessinaient abominablement	3
se courbant...haies	3
A travers le sang ... (oreilles)	3
La porte du salon, et alors,	3
les lumières des grandes baies	3
le dossier élevé... vert	3
quelques jours auparavant	2
et l'ouvrit de nouveau	2
Ce soir-là	2
L'illusion... le prit..aussitôt	2
étroitement liés à la tâche	2
À son tour	2
il	2
distingua dans la brume(mauve du crépuscule)	2

ne devait pas...là	2
Il	2
les paroles de la femme	2
l'abandonna à cause	1
dans le train	1
(en retournant) à sa propriété	1
Phrase après phrase	1
il	1
se laissait prendre...vie	1
Il	1
comme un fleuve de reptiles	1
et l'on sentait...toujours	1
Rien n'avait été oublié...calculé	1
La double	1
une main	1
frôle	1
une joue	1
Sans se regarder	1
l'allée qui conduisait à la maison	1
A cette heure	1
monta les trois marches du perron	1
et entra	1
lui parvenaient encore	1
D'abord une salle bleue	1
puis un corridor	1
puis un escalier avec un tapis	1
En haut, deux portes	1
Personne dans la première pièce	1
personne dans la seconde	1
et implacable... le temps qu'	0
Il commençait à faire nuit.	0
A la fin	0
Les chiens... n'aboyèrent pas.	0

Table 1.

## Notes

[1] Le concept de « *distant reading* » est imaginé par Franco Moretti [Moretti 2000] dans son article « Conjectures on world literature » : appliquant le concept de « système-monde » (défini par Immanuel Wallerstein pour penser l'économie) à la littérature, il définit cette dernière comme un noyau à la périphérie duquel émergent les romans modernes sous l'influence d'autres cultures. Il propose donc une méthode de lecture à distance qui permette aux chercheurs de comprendre le fonctionnement de cette littérature à l'échelle globale (article repris dans *Distant reading*, 2013).

[2] Nous distinguons ici ce néologisme du « multiple reading » qui réfère à des lectures réitérées par une seule personne d'un même texte, avec une perspective d'appropriation et de meilleure compréhension de ce texte (voir [Gallagher 2004]). Pour nous, le « *crowdreading* » désigne la lecture par plusieurs personnes d'un même texte, avec comme objectif d'observer les points de convergence et de divergence dans l'appréhension et la réception de ce texte.

[3] Corpus d'auteurs pour les humanités : informatisation, édition, recherche. Pour des informations supplémentaires, voir :



<http://cahier.hypotheses.org/>.

[4] Pour une présentation rapide de ses méthodes et techniques, voir : [Pincemin, Heiden, 2008].

[5] Comme, par exemple, la topic modeling ; pour une présentation de cette technique, voir : [Mohr et Bogdanov 2013].

[6] Professeur de littérature française et humanités numériques, spécialiste du théâtre du XVIII<sup>e</sup> siècle, Ioana Galleron est responsable de l'édition électronique du théâtre de Louis de Boissy et travaille actuellement à la numérisation en XML/TEI des registres de la Comédie-Française. Professeur de littérature française, spécialiste du roman du XIX<sup>e</sup> siècle, Cécile Meynard a été entre 2005 et 2015 co-directrice du projet d'édition en ligne des Manuscrits de Stendhal ([www.manuscrits-de-stendhal.org](http://www.manuscrits-de-stendhal.org)). Elle est actuellement co-responsable du projet de transcription en crowdsourcing du manuscrit de *Mon évasion* de Benoîte Groult, dont le prototype est disponible en ligne : <http://www.espace-transcription.org/>. Professeur de littérature latino-américaine contemporaine, spécialiste de génétique des textes et des œuvres, Fatima Idmhand a été responsable du projet ANR-CHispa (2014-2017) dont l'objectif a été le développement d'outils et logiciels destinés à faciliter l'édition numérique de manuscrits modernes ([chispa.hypotheses.org](http://chispa.hypotheses.org)). Actuellement, elle travaille sur le développement de méthodes pour le codage/balispage des courants culturels et littéraires (dans les archives, les manuscrits et les textes) en vue de leur exploitation informatisée.

[7] Non sans que ces protocoles ne présentent leurs propres difficultés, et ne suscitent à leur tour de multiples questions. Voir, par exemple : [Ciotti, Lana et Tomasi, 2014-2015].

[8] Rappelons qu'Eco définit le texte comme une « machine paresseuse », exigeant de la part du lecteur une collaboration active pour remplir les « blancs ».

[9] Voir la définition déjà ancienne du personnage comme « tas de mots » par [Forster 1927].

[10] <http://enccre.academie-sciences.fr/>

[11] <https://candide.bnf.fr>

[12] <http://dissimilitudes.lip6.fr:8180/#/project>

[13] *Ibid.*

[14] *Ibid.*

[15] Le support de cette conférence est actuellement consultable sur Academia, il devrait être prochainement disponible sur le site de HALSHS. [https://www.academia.edu/27046668/De\\_cahIER\\_à\\_cahierR.\\_Pour\\_des\\_études\\_littéraires\\_assistées\\_par\\_ordinateur](https://www.academia.edu/27046668/De_cahIER_à_cahierR._Pour_des_études_littéraires_assistées_par_ordinateur).

[16] Quelques présentations ont été consacrées à ce sujet. Voir : [Eder, Jannidis, Rybicki, Schöch et van Dalen 2016] ; [Galleron 2016] ; [Rockwell, Mandell, Sinclair, Wilkens, Capinatu et Downie 2016] ou [Riguet et Mpouli 2016].

[17] Pour des questions de simplicité de gestion, le recours au formulaire de Google s'est imposé à nous. Il permet aisément de mener une telle enquête entre des chercheurs de différentes institutions impliquant un grand nombre de participants étudiants. L'idée étant, à long terme, de réaliser un travail analogue à échelle internationale, nous n'avons pas eu recours à des outils plus locaux (dans les « Environnements Numériques de Travail » par exemple) ou nationaux (du « Réseau National de télécommunications pour la Technologie l'Enseignement et la Recherche » par exemple).

[18] Le texte est joint en annexe n° 3.

[19] Il est à noter que nous n'avons pas abandonné complètement le projet de revenir, plus tard, aux autres textes initialement choisis.

[20] Plutôt que de recopier « il se laissait lentement intéresser », « il laissait sa main gauche caresser », etc., un répondant préfère ainsi renvoyer à tous les « il + imparfait » comme rendant compte de ce qu'il appelle « le lecteur », par opposition à « il + passé simple », qu'il considère comme renvoyant au tueur – ce qui n'est pas toujours le cas. Ainsi, la réponse, plutôt fine et argumentée par ailleurs, souffre des limitations de l'outil employé pour la recueillir.

[21] Il n'est laissé de côté que dans deux cas, par des répondants qui considèrent clairement qu'il fait partie d'un autre niveau du texte, et que seuls les personnages du livre qu'il lit sont de « vrais » personnages.

[22] Il est possible que les étudiants n'aient pas pu percevoir en français et en anglais le jeu introduit par l'auteur sur l'identité de ce « il ». En effet, les traductions anglaise et française de la nouvelle ont gommé l'ambiguïté volontairement introduite par Julio Cortázar sur l'identité du sujet. Si la langue espagnole n'impose pas la présence du pronom personnel sujet avant le verbe – le récit commence par « Había empezado a leer » et pourrait être « il », « elle » voire même « on » –, les traductions française et anglaise ont choisi de redonner une identité précise au sujet, et proposé « he » et « il ».

[23] Le répondant ayant identifié les chiens comme personnages ne surligne pas la phrase en question !

[24] Pour traiter ces réponses, nous avons utilisé le logiciel AntConc (voir : <http://www.laurenceanthony.net/software.html>), après les avoir « fondues » en un seul document, au format RTF.

[25] C'est le cas à propos du caractère « passionné », clairement attribué à « il » parce qu'il se laisse absorber par sa lecture, et qu'il s'organise de façon à ne pas être dérangé.

[26] Il en est ainsi du caractère « épris de liberté » qui est attribué parfois à l'amant. Interrogés sur cette description plutôt surprenante du personnage, quelques étudiants ont expliqué qu'ils l'avaient perçu ainsi parce qu'il est question à propos de son cœur qui bat au rythme de la « liberté convoitée ».

[27] Il existe déjà des projets pédagogiques au niveau scolaire et secondaire qui vont dans ce sens, en recourant par exemple au logiciel <https://animoto.com/>. Nous avons nous-mêmes tenté une expérimentation dans notre première campagne d'enquête, que nous n'avons pas retenue dans la deuxième, mais que nous réintroduirons dans la prochaine phase car le peu de résultats déjà obtenus permet déjà d'émettre nombre d'hypothèses stimulantes.

## Works Cited

**Bernard 1999** Bernard, Michel. *Introduction aux études littéraires assistées par ordinateur*. Paris, PUF (1999).

**Bernard 2003** Bernard, Michel. *Études littéraires assistées par ordinateur. Réflexion sur les conditions d'émergence d'une discipline*. Dossier d'habilitation à diriger des recherches, sous la direction du professeur Henri Béhar, Paris (2003).

**Ciotti, Lana et Tomasi, 2014-2015** Ciotti, Fabio, Lana, Maurizio, Tomasi, Francesca. "TEI, Ontologies, Linked Open Data: Geolat and Beyond". *Journal of the Text Encoding Initiative* [Online], 8 (2014–2015). DOI: 10.4000/jtei.1365

**Cortázar 1964** Cortázar Julio, *Continuité des parcs*.

**Dufays 2013** Dufays, Jean-Louis. "Sujet lecteur et lecture littéraire: quelles modélisations pour quels enjeux?". *Recherches et travaux*, ELLUG, 83: 77–88 (2013).

**Dufays, Gemenne et Ledur 2005** Dufays, Jean-Louis, Gemenne, Louis, et Ledur, Dominique. *Pour une lecture littéraire. Histoire, théories, pistes pour la classe*, Bruxelles: De Boeck – Duculot (2005).

**Eco 1985** Eco, Umberto. *Lector in fabula*. Paris, Livre de Poche (1985) (1<sup>ère</sup> éd. Grasset, 1979)

**Eder, Jannidis, Rybicki, Schöch et van Dalen 2016** Eder, Maciej, Jannidis, Fotis, Rybicki, Jan, Schöch, Christof, van Dalen, Karina (2016). "Literary Concepts: The Past and the Future". In *Digital Humanities 2016: Conference Abstracts*. Kraków, Jagiellonian University & Pedagogical University: 63–65 (2016). Disponible en ligne: <http://dh2016.adho.org/abstracts>.

**Fischer, Kampkaspar et Trilcke 2015** Fischer, Frank, Kampkaspar, Dario, Trilcke, Peer. "Digitale Netzwerkanalyse dramatischer Texte" (2015). Disponible en ligne: <http://gams.uni-graz.at/o:dhd2015.v.040>.

**Forster 1927** Forster, Edward Morgan. *Aspects du roman* (trad. Sophie Basch). Paris, C. Bourgois (1993).

**Gallagher 2004** Gallagher, Kelly (2004). *Deeper reading*. Portland, Maine: Stenhouse Publishers.

**Galleron 2016** Galleron, Ioana. "Playing with French Drama: from Old Research Questions to New Research Tools". In *Digital Humanities 2016: Conference Abstracts*. Kraków, Jagiellonian University & Pedagogical University (2016): 522–523 <http://dh2016.adho.org/abstracts>

**Hamon 1977** Hamon, Philippe. "Le statut sémiologique du personnage". In *Poétique du récit*. Paris, Seuil (1977).

**Hoover 2013** Hoover, David L. "Textual Analysis". In *Literary Studies in the Digital Age. An Evolving Anthology* (2013). Disponible en ligne: <https://dsanthology.commons.mla.org/textual-analysis/>. DOI: 10.1632/lstda.2013.3 [consulté le 26 septembre 2016].

**Iser 1976** Iser, Wolfgang. *L'Acte de lecture. Théorie de l'effet esthétique* (trad. Evelyne Sznycer). Paris, Mardaga (1985).

**Jannidis 2013** Jannidis, Fotis. "Character". In Hühn, Peter et al. (éd.). *The living handbook of narratology*. Hamburg, Hamburg University Press (2013): paragraphe 10. Disponible en ligne: <http://wikis.sub.uni-hamburg.de/lhn/index.php/Character> [consulté le 26 septembre 2016]

**Jockers 2013** Jockers, Matthew L. *Macroanalysis. Digital Methods and Literary History*. Chicago & Springfield,

University of Illinois Press (2013).

**Jouve 2002** Jouve, Vincent. *L'effet-personnage dans le roman*. Paris, Presses universitaires de France (1<sup>ère</sup> éd. 1992).

**Maupassant 1897** Maupassant, Guy de, *Le Horla*.

**McCarty 2004** McCarty, Willard. "Modeling. A Study in Words and Meanings". In Schreibman, Susan, Siemens, Ray et Unsworth, John (éd.). *A Companion to Digital Humanities*. Wiley (2004). DOI: 10.1002/9780470999875.ch 19

**McCarty 2005** McCarty, Willard. *Humanities Computing*. Basingstoke, Palgrave MacMillan (2005).

**Merzeau 2014** Merzeau, Louise. "Culture numérique, médias, commun et vivre ensemble". Disponible en ligne: <http://merzeau.net/culture-numerique-media-communs-et-vivre-ensemble/> [consulté le 8 septembre 2017]

**Mohr et Bogdanov 2013** Mohr, John W., Bogdanov, Petko. "Introduction – Topic Models: What They Are and Why they Matter", *Poetics*, 41–6: 545–569 (2013). DOI: <http://doi.org/10.1016/j.poetic.2013.10.001>.

**Moretti 2000** Moretti, Franco. "Conjectures on world literature", *New Left Review*, January - February 2000, p. 54-68.

**Moretti 2013** Moretti, Franco. "Conjectures on World Literature". *Distant Reading*. Londres et New York, Verso (2013).

**Phelan 1989** Phelan, James. *Reading People, Reading Plots. Character, Progression and the Interpretation of Narrative*. Chicago et London (1989).

**Pincemin, Heiden, 2008** Pincemin, Bénédicte, Heiden, Serge (2008) – "Qu'est-ce que la textométrie? Présentation", *Site du projet Textométrie*. Disponible en ligne: <http://textometrie.ens-lyon.fr/spip.php?rubrique80>. [consulté le 8 septembre 2017]

**Ricardou 1967** Ricardou, Jean. *Problèmes du Nouveau roman*. Paris, Éditions du Seuil (1967).

**Riguet et Mpouli 2016** Riguet, Marine, Mpouli, Suzanne. "À la croisée des discours littéraire et scientifique: la comparaison comme haute figure dialogique". In *Digital Humanities 2016: Conference Abstracts*. Kraków, Jagiellonian University & Pedagogical University: 330–333 (2016). Disponible en ligne: <http://dh2016.adho.org/abstracts>

**Rockwell, Mandell, Sinclair, Wilkens, Capinatu et Downie 2016** Rockwell, Geoffrey, Mandell, Laura, Sinclair, Stéfan, Wilkens, Matthew, Capitanu, Boris, Downie, J. Stephen. "The Trace of Theory: Extracting Subsets from Large Collections". In *Digital Humanities 2016: Conference Abstracts*. Kraków, Jagiellonian University & Pedagogical University: 85–92 (2016). <http://dh2016.adho.org/abstracts>

**Rommel 2004** Rommel, Thomas. "Literary Studies". In Schreibman, Susan, Siemens, Ray, Unsworth, John (éd.). *A Companion to Digital Humanities*. Oxford, Blackwell Publishing (2004). Disponible en ligne: <http://www.digitalhumanities.org/companion/> [consulté le 26 septembre 2016]

**Rosny aîné, 1888** Rosny Joseph Henri, dit Rosny aîné, *Les Xipéhuz*.

**Schnapp, Lunenfeld et Presner (s.d.)** Schnapp, Jeffrey, Lunenfeld, Peter, Presner, Todd. "The Digital Humanities Manifesto 2.0" (s.d.). Disponible en ligne: [http://www.humanitiesblast.com/manifesto/Manifesto\\_V2.pdf](http://www.humanitiesblast.com/manifesto/Manifesto_V2.pdf).

**Weststeijn 2005** Weststeijn, Willem G.. "Towards a Cognitive Theory of Character". In *Amsterdam International Electronic Journal for Cultural Narratology* (2005). Disponible en ligne: [http://cf.hum.uva.nl/narratology/a07\\_weststeijn.htm](http://cf.hum.uva.nl/narratology/a07_weststeijn.htm) [consulté le 26 septembre 2016]