DHQ: Digital Humanities Quarterly
2018
Volume 12 Number 1

Pour une analyse automatique du jugement critique : les citations modalisées dans le discours littéraire du XIX^{e} siècle [en]

Marine Riguet <marineriguet_at_gmail_dot_com>, Labex OBVIL (Paris-Sorbonne)
Motasem Alrahabi <motasem_dot_alrahabi_at_gmail_dot_com>, Paris-Sorbonne Abou Dhabi

Abstract

Au XIX^{e} siècle, sous l'influence du positivisme et des sciences émergentes, la critique littéraire française cherche à affirmer l'autonomie de son discours. Dans cette entreprise de légitimation, le recours aux citations textuelles vient notamment servir les stratégies d'un jugement critique en quête d'objectivité et de scientificité. À partir d'un grand corpus de textes littéraires, cet article s'attache ainsi à annoter automatiquement les expressions modales employées autour du discours rapporté. Nous présenterons pour ce faire deux applications grâce auxquelles ont pu être identifiés et analysés certains traits caractéristiques du discours de critique littéraire : EXCOM-2, un système d'annotation de catégories sémantiques et discursives ; et E-Quotes, son interface web pour la fouille des citations.

Introduction

Quand, dans la seconde moitié du XIX^{e} siècle, la critique littéraire française cherche à s'affirmer comme discipline autonome, elle doit assurer sa légitimité au milieu de discours scientifiques triomphants tels que la biologie et la médecine. De Sainte-Beuve à Lanson, une part de la critique s'efforce alors d'asseoir son entreprise sur un socle rationnel par le biais de méthodes, de critères et d'emprunts conceptuels, soit, en bref, par la revendication d'une démarche « scientifique » qui l'arrache à la subjectivité des belles-lettres. Or, la citation textuelle nous semble tenir un rôle majeur dans cette stratégie du discours. Elle est d'emblée le cœur d'une critique qui est par définition *métadiscours*, commentaire d'un texte déjà existant ; mais elle se fait aussi l'indice de sources exogènes auxquelles recourt l'auteur pour appuyer son propos. 1

> Claude Bernard dit avec raison : « La conquête intellectuelle de l'homme consiste à faire diminuer et à refouler l'indéterminisme, à mesure qu'à l'aide de la méthode expérimentale il gagne du terrain sur le déterminisme. » Notre vraie besogne est là, à nous romanciers expérimentateurs, aller du connu à l'inconnu, pour nous rendre maître de la nature… [Zola 1881]

En convoquant un discours étranger au sien, comme en témoigne cet extrait du *Roman expérimental*, la critique littéraire orchestre la validité et la normativité de son propre discours. Telle est du moins l'hypothèse de départ qu'il nous faut éprouver avant toute chose.

Ainsi, cette étude à visée exploratoire tend à observer les multiples opérations de « greffe » textuelle [Compagnon 1979] qui viennent alimenter la critique littéraire sur un demi-siècle. Un examen à grande échelle paraît le plus adéquat afin d'identifier la manière dont la critique littéraire construit la cohésion de son discours et prétend façonner, par différents emplois de la citation, sa positivité. En plus du contenu de la citation, c'est sa modélisation qui nous intéresse ici, c'est-à-dire ses formes de récupération et de réappropriation par l'auteur, que ce soit au sein d'un jugement, d'un raisonnement argumentatif, d'une description, etc. La modélisation représente précisément cette opération combinatoire de la citation qui ne consiste pas tant dans la répétition d'un énoncé antérieur, que dans la production d'un nouveau. 2

Dans ce dessein, nous nous sommes tournés vers un système d'annotation de catégories sémantiques et discursives, conçu pour répondre à ce type de questionnement. Nous présenterons, dans un premier temps, les ressources linguistiques que nous avons constituées, puis le système d'annotation et l'interface web adoptés. Ce travail préliminaire nous permettra ensuite d'esquisser des pistes d'analyse et d'ouvrir des perspectives pour une interprétation littéraire plus poussée. 3

Choix méthodologiques et techniques

Notre corpus de travail est constitué d'environ 300 ouvrages de critique littéraire française publiés entre 1830 et 1920[1]. Ces textes ont été préalablement numérisés dans le cadre du Labex OBVIL de Paris-Sorbonne [Alexandre 2016] et sont librement accessibles[2]. 4

Pour l'annotation du corpus, nous avons opté pour une approche orientée connaissances. En effet, en l'absence de ressources linguistiques et de corpus annotés adaptés à nos besoins particuliers, nous nous sommes basés sur un travail précédent sur les modalités énonciatives dans le discours rapporté [Alrahabi 2010]. Cela nous a d'abord permis de réutiliser les marqueurs linguistiques existants et de les enrichir par la suite. Ce choix répond par ailleurs à notre besoin d'effectuer des annotations avec des nuances sémantiques très fines, qui dépassent la simple dichotomie positif/négatif (voir section 1). Dans ce sens, nous avons utilisé EXCOM-2, un outil d'annotation à base de règles [Alrahabi 2010][3] qui nous a permis d'avoir le contrôle sur le processus d'annotation et d'améliorer progressivement la pertinence des ressources linguistiques exploitées. Outre la catégorisation des citations en différentes langues [Alrahabi 2010] [Alrahabi 2015], EXCOM-2 est utilisé dans d'autres tâches de fouille de textes comme l'annotation des *hypothèses* dans les articles de biologie en anglais [Desclés, Alrahabi et Desclés 2011], l'ordonnancement des résultats dans un système de recherche d'information [Atanassova, Bertin et Desclés 2012], la construction des *fiches de synthèse* [Makkaoui, Desclés, Bertin, Jouis et Ganascia 2014] ou le repérage automatique des anecdotes dans les textes littéraires [Bénard et Alrahabi 2017]. 5

Pour la visualisation des résultats, nous avons utilisé E-Quotes[4], une interface web de navigation textuelle guidée par les annotations sémantiques et discursives [Alrahabi 2016]. 6

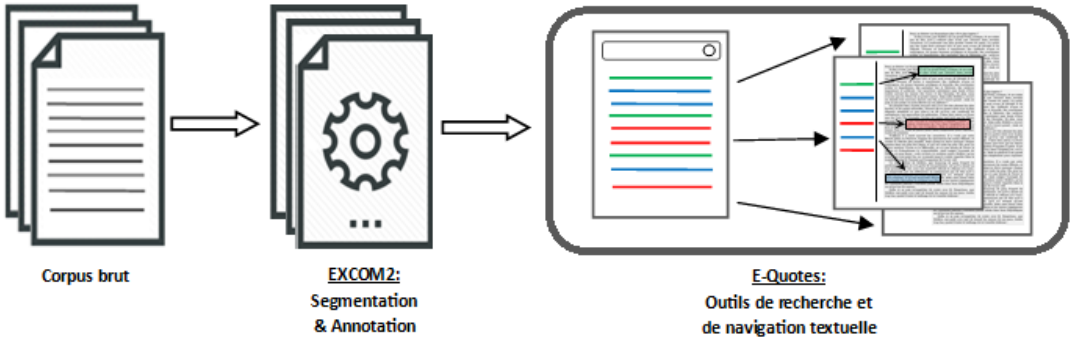

Figure 1. Chaîne de traitement d'E-Quotes

E-Quotes, qui prend les sorties d' EXCOM-2, permet d'effectuer des recherches par mots clés au sein des citations catégorisées et de *naviguer* entre celles-ci.

Construction des ressources linguistiques

Pour la présente étude, nous avons utilisé une grande base de marqueurs linguistiques (plus de 600) manuellement collectés à partir de l'analyse et l'observation de corpus. Il s'agit principalement de marqueurs verbaux permettant d'introduire et de modaliser les citations. Ceux-ci sont classés dans une *cartouche* linguistique selon des critères sémantiques, énonciatifs et discursifs [Alrahabi 2010]. Voici quelques exemples de *groupes* de catégories : 7

- Dire, écrire, déclarer, affirmer, nier, informer, confier…
- Diffuser, transmettre, colporter…
- Observer, décrire, définir…
- Analyser, critiquer, justifier, argumenter, comparer…
- Accuser, attaquer, dénoncer, s'indigner, ironiser, menacer, s'opposer à…
- Encourager, aimer, louer, considérer comme correct, s'accorder avec…
- Penser, croire, douter, déduire…
- Ajouter, répéter, reformuler, exemplification, énumérer, résumer…
- Demander, ordonner, proposer, répondre, permettre, interdire…
- …

Nous avons ensuite enrichi ces ressources par des *adjectifs* et des *adverbes* afin de couvrir de nouveaux phénomènes linguistiques comme, à titre d'exemple, les prises de position concomitantes ou opposées entre auteur et personnes citées (*accord* vs *désaccord*). Les adverbes collectés désignent souvent l'attitude de l'énonciateur par rapport à la totalité de l'acte de locution ; tel est le cas de *certainement, peut-être, à tort, non sans raison, prétendument*, etc. Exemple : « Il ajoute, il est vrai, cette objection : "Mais si je n'ai pas d'oreille ?" » [Guyau 1884]. 8

Il existe également des adverbes (*sévèrement, admirablement, avec raison…*) qui renvoient à d'autres entités énonciatives dans le discours rapporté. Ainsi, dans la phrase suivante, l'adverbe décrit l'attitude du Locuteur : « Donc, dit-il ironiquement, vous ne lirez point Pascal, vous ne lirez point Descartes, vous ne lirez point de Retz. » [Albalat 1905] 9

Quant aux adjectifs, beaucoup plus fréquents, ils peuvent qualifier n'importe quel élément de la phrase. Examinons ces exemples où les adjectifs désignent respectivement le propos rapporté, le locuteur, l'interlocuteur ou n'importe quel autre objet : « Elle a laissé échapper le mot charmant : "Va, je ne te hais point." » [Deschanel 1864] ; « Junon, indignée, dans l'Iliade, lui reproche son "cœur de lionne contre les femmes." » [Saint-Victor 1882] ; « Il lut Montaigne dont il loue la "loquacité" charmante : "Une loquacité qui vient de la joie de tourner d'une façon toujours nouvelle la même chose… " » [Faguet 1904] ; « On peut dire de celle qui a écrit cette admirable lettre, comme d'Élisabeth Ranquet que, "marchant sur la terre, elle était dans les cieux." » [France 1892]. 10

Au début du projet, la base des ressources linguistiques a été régulièrement mise à jour de façon semi-automatique. À partir de différents corpus, EXCOM-2 permet en effet d'extraire la liste des expressions introductrices des citations (expressions candidates), de les classer en soulignant les marqueurs déjà répertoriés, en éliminant les doublons et les expressions antérieurement enregistrées. Notre rôle en tant qu'experts est de filtrer les résultats afin d'ajouter éventuellement de nouvelles entrées ou d'enrichir l'existant. Dans la phrase suivante par exemple, le marqueur *consentir* nous a amené à relever le marqueur *se raviser* : « Henri consent d'abord, puis se ravise : "Non, décidément, ça m'ennuie" » [Lemaitre 1891]. Ce processus nous a permis d'avoir une très bonne couverture au niveau des marqueurs (verbes, noms, adjectifs, adverbes…) et des structures dans lesquelles ceux-ci peuvent apparaître. 11

Il existe différents travaux de recherche sur les marqueurs de la parole et sur les modalités dans le discours rapporté ([Bergler 1992]; [Charron et Jacob 1999]; [Rosier 1999]; [Tuomarla 2000]; [Charaudeau 2005]; etc.). Dans notre analyse des introducteurs et modalisateurs des citations, nous nous plaçons dans une approche énonciative : comment l'énonciateur ancre-t-il le discours rapporté dans la situation d'énonciation ? Comment rapporte-t-il et modalise-t-il le propos du locuteur ? Comment décrit-il l'attitude de celui-ci et comment affiche-t-il sa propre position vis-à-vis de ce qu'il rapporte ? 12

Notre travail s'inscrit également dans une perspective translinguistique où les concepts analysés sont généralement universels (*opinion, jugement, évaluation*…). Les marqueurs utilisés ne dépendent pas non plus du type de corpus ou du genre de texte. Ceux-ci peuvent être employés dans un domaine comme dans un autre (exemples : *déclarer, selon, définition, en conclusion, indignation, être d'accord avec,* etc.), et sont par conséquent facilement traduisibles d'une langue à *l'autre* [Alrahabi, Desclés et Suh 2010]. 13

Annotation automatique

EXCOM-2 est basé sur des marqueurs linguistiques observables et des règles heuristiques afin d'identifier les citations et les catégoriser en fonction des modalités présentes dans leur contexte. Voici un exemple de phrase identifiée comme citation et catégorisée comme « Ironie » : « Il lui faut, dit-on moqueusement, cinq épithètes ! » [Albalat 1905]. La présence dans le contexte *milieu* de la citation de la *virgule* suivie des marqueurs *dit-on* et *moqueusement* permet à EXCOM-2 d'attribuer à ce passage textuel l'étiquette « Citation – Ironie ». 14

Dans EXCOM-2, les règles, dites d'*exploration contextuelle* [Desclés 2006], peuvent être organisées selon un ordre de priorité et utiliser en entrée les résultats d'autres règles. Avant l'étape de l'annotation, seule une phase de segmentation automatique avec EXCOM-2 est nécessaire afin de découper les textes en sections, paragraphes et phrases. 15

Interface de recherche et de navigation

L'interface utilisateur d'E-Quotes (voir figure 2) permet de combiner la technologie classique de recherche d'information (requêtes de mots clés sur un sac de mots) et une recherche autour des citations sémantiquement catégorisées. L'utilisateur peut rechercher, par exemple, un terme dans les citations annotés selon une catégorie bien particulière. 16

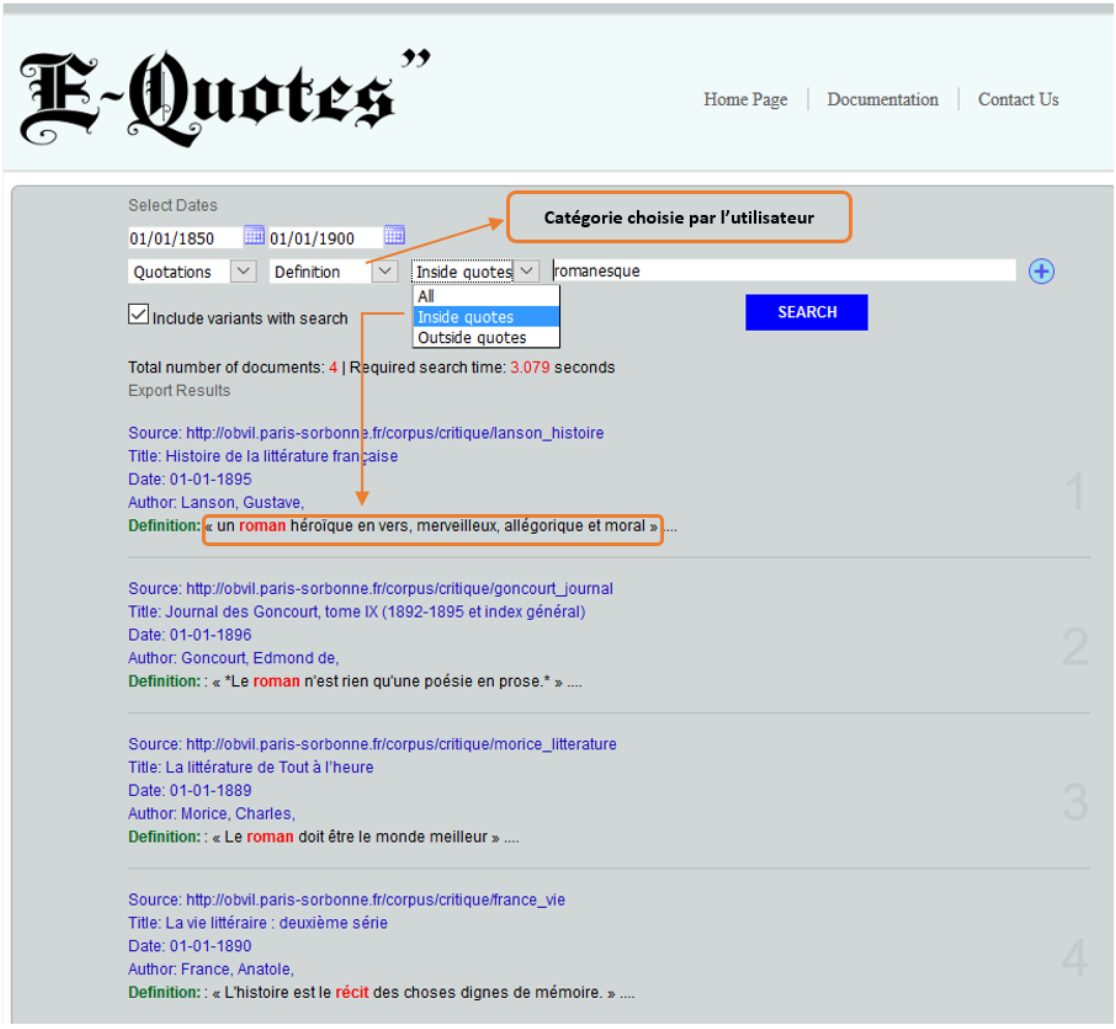

Figure 2. Interface utilisateur d'E-Quotes

Il a également la possibilité d'affiner cette recherche en choisissant de localiser les mots clés à l'intérieur des citations ou bien à l'extérieur de celles-ci. Cette fonctionnalité permet de trouver des réponses à une telle question : *Comment l'énonciateur présente l'attitude d'un locuteur vis-à-vis d'un propos rapporté ?* La position de l'énonciateur (auteur) et l'attitude du locuteur sont représentées par l'ensemble des catégories identifiées dans le contexte d'une citation.

Les termes recherchés dans E-Quotes peuvent être enrichis par leurs synonymes ou des mots du même champ lexical 17

(ex. *darwinisme*, *évolution*, *classification*…) ou bien par leurs équivalents (ex. *Flaubert*, *Gustave Flaubert*, *M. Flaubert*…). Ces unités linguistiques doivent être fournies au système au moment de l'indexation.

L'application permet également d'effectuer des requêtes dans les champs « Titre », « Auteur » ou « Date » des articles du corpus et de combiner plusieurs requêtes ensemble à l'aide des opérateurs ET, OU et NON. Toutes ces fonctionnalités offrent à l'utilisateur le moyen de réaliser des requêtes sophistiquées. Exemple : rechercher dans les articles écrits entre 1850 et 1900 les *Définitions rapportées*, en présence du terme « romanesque » ou ses équivalents, à l'intérieur même de la citation. Voici un des résultats annotés : 18

> Acceptant cette définition de Madame Necker : « Le roman doit être le monde meilleur », Balzac ajoute : « Mais le roman ne serait rien si, dans cet auguste mensonge, il n'était pas vrai dans les détails. » [Morice 1889]

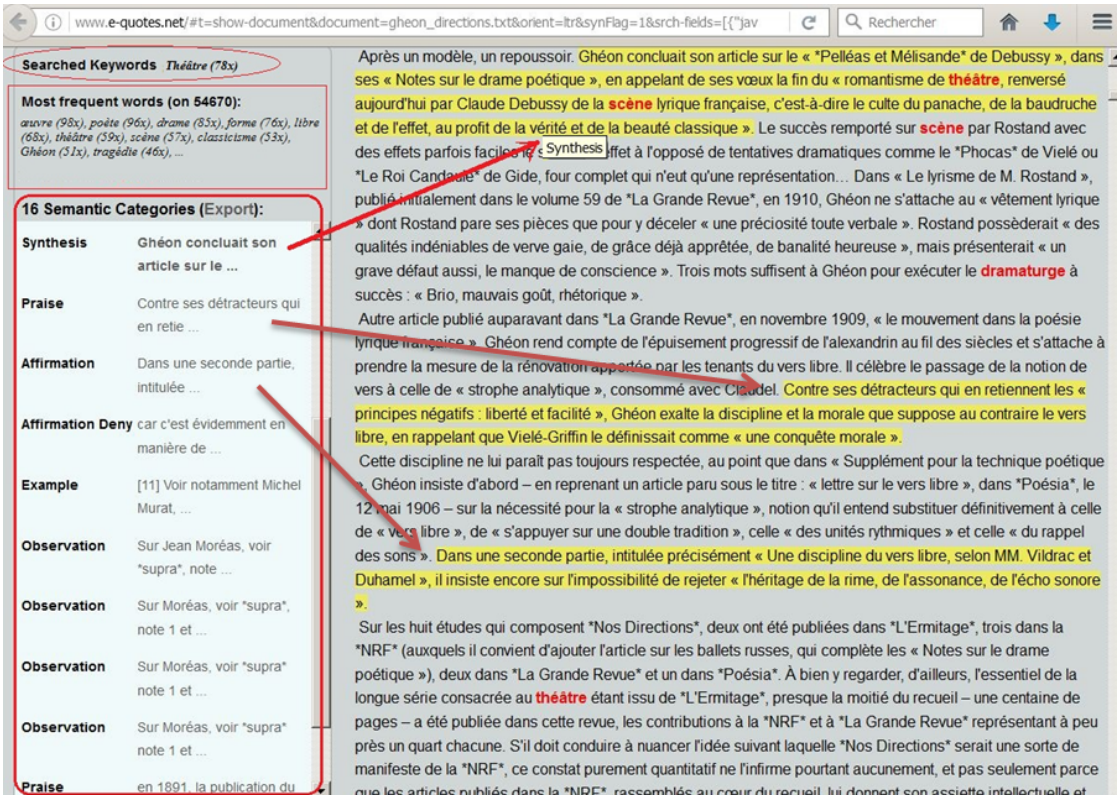

Figure 3. Navigation dans un document annoté à l'aide des catégories sémantiques

Les résultats d'une recherche sont classés par document et chacun d'eux renvoie au contexte de la citation dans l'article d'origine. Toutes les citations catégorisées sont surlignées et les termes de la requête sont coloriés. Afin de faciliter la navigation entre les résultats, une fenêtre dans chaque document affiche toutes les annotations identifiées, reliées à leur contexte dans le document. Cette configuration visuelle où annotations sémantiques et mots clés sont mis en relief offre pour les longs documents une lecture optimale et une interprétation pertinente au niveau du sens et du discours. De cette façon, l'utilisateur a le moyen de parcourir les annotations dans le texte, d'une *Définition* à une *Comparaison*, d'une *Accusation* à une *Indignation*, d'une *Opinion positive* à une autre *négative*, etc. 19

E-Quotes fournit d'autres informations pour chaque document, comme le nombre d'occurrences des termes de la requête et les mots les plus fréquents du document (figure 3). Enfin, les résultats d'une requête peuvent être exportés sous forme de fichier XML ou de tableur pour faciliter leur traitement par d'autres applications. 20

Plusieurs travaux ont en effet abordé la question de l'identification des citations, mais, à notre connaissance, très peu de recherches ont abouti à des applications opérationnelles pour des utilisateurs finaux. Citons le système NewsExplorer[5] du Centre commun de recherche européen #pouliquen2007. À partir d'un nom choisi dans une liste, 21

cette application permet, entre autres, de détecter les citations attribuées à cette personne (locuteur) ou qui parlent de cette personne. NewsExplorer couvre différentes langues et traite quotidiennement des milliers d'articles journalistiques et dépêches. Notre système, contrairement à NewsExplorer, offre la possibilité d'effectuer des recherches dans des citations sémantiquement catégorisées, et de naviguer dans les documents à travers ces catégories très fines au niveau du sens et du discours.

Évaluation métrique des résultats

22 À partir des ressources linguistiques établies et de notre connaissance préalable du corpus, nous avons tout d'abord déterminé un ensemble de catégories sémantiques (voir Table 9) qui nous semblaient les plus à même de rendre compte des différentes nuances modales utilisées par les critiques littéraires. Le choix de ces catégories n'a bien sûr pas vocation à faire autorité dans le domaine et pourrait être sujet à discussion : il n'a d'autre prétention ici que de répondre précisément aux besoins de notre étude. Avec ces catégories, nous avons effectué des tests en boucle sur une partie du corpus (environ 7 %) afin de construire et de stabiliser les règles[6].

23 Ensuite, nous avons annoté avec EXCOM-2 le reste du corpus[7]. Pour l'évaluation en cours, nous nous sommes focalisés dans un premier temps sur le calcul de la *précision* des annotations automatiques obtenues. Nous avons alors demandé à trois personnes natives de niveau universitaire de vérifier manuellement la totalité des annotations. Pour chaque annotation, chaque évaluateur devait choisir entre trois options : « Correct », « Incorrect » ou « Je ne sais pas ». Afin de juger les annotations, un seul critère était formulé aux évaluateurs : le passage entre guillemets doit être introduit par un ou plusieurs termes sémantiquement liés à la catégorie en cours. Vu les nuances sémantiques subtiles entre certaines catégories, nous avons été contraints d'organiser plusieurs séances de sensibilisation pour les évaluateurs. Un guide[8] d'annotation a été créé pour faciliter leur travail ainsi qu'une interface dédiée pour le calcul automatique des résultats (figure 4).

Figure 4. Copie d'écran de l'interface de l'évaluation

24 La règle que nous avons adoptée dans le calcul était de considérer une annotation comme correcte lorsque deux annotateurs au moins la considéraient de la sorte. Ainsi, nous avons obtenu 87,6 % de *précision* pour l'ensemble des catégories évaluées. Voici en détail les résultats pour chaque catégorie :

Catégorie	Nombre des phrases annotées	Nombre des phrases bien annotées	Précision (%)
Accord	215	133	62
Accusation	27	21	77,7
Appréciation	293	281	96
Argumentation	88	85	96
Affirmation	186	165	88
Comparaison	54	47	87
Croyance	221	208	95
Déduction	24	21	87
Définition	169	143	85
Dénonciation	9	9	100
Dépréciation	92	85	93
Désaccord	130	116	89
Doute	37	25	67
Éloge	375	357	95
Évaluation	181	146	81
Erreur	32	25	78
Exemplification	228	211	93
Indignation	138	133	96
Insulte	41	33	81
Ironie	131	126	96
Déni	28	22	78
Observation	301	250	83
Prétention	18	16	88
Plainte	63	61	97
Synthèse	199	181	91
Total	3280	2900	
Moyenne	**87,6**		

Table 1. Résultats de l'évaluation

Cette première évaluation, même si elle ne permet pas encore de mesurer le *silence*, montre de très bons résultats de notre système d'annotation au niveau de la *précision*.

Les taux particulièrement bas de certaines catégories s'expliquent par les points suivants : des marqueurs polysémiques, comme dans la catégorie *Accord* (62 %) :

> Le défendeur nous montre d'abord pour quel motif on adoptait un fils : « Ménéclès, dit-il, ne voulait pas mourir sans enfants ; il tenait à laisser après lui quelqu'un pour l'ensevelir et pour lui faire dans

la suite les cérémonies du culte funèbre. » [Coulanges 1864]

Dans la catégorie *Doute* (67 %), le problème principal est posé par la négation syntaxique qui peut inverser complétement le sens de la phrase :

> Mademoiselle Beaupré, une des premières actrices qui parut sur la scène (car pendant longtemps les hommes tinrent l'emploi des femmes au théâtre), rendait, sans s'en douter, un bien grand hommage à Corneille : « Il nous a fait tort, disait-elle ; nous avions avant lui des pièces pour *trois* écus et nous gagnions beaucoup, aujourd'hui les pièces sont fort cher et nous gagnons peu… » [Du Casse 1864]

.

Nous devons ainsi enrichir notre base de ressources linguistiques par de nouveaux marqueurs pour améliorer la désambiguïsation. 26

Interprétation et analyse littéraire des résultats

Pour l'ensemble des annotations du corpus, nous obtenons la répartition suivante : 27

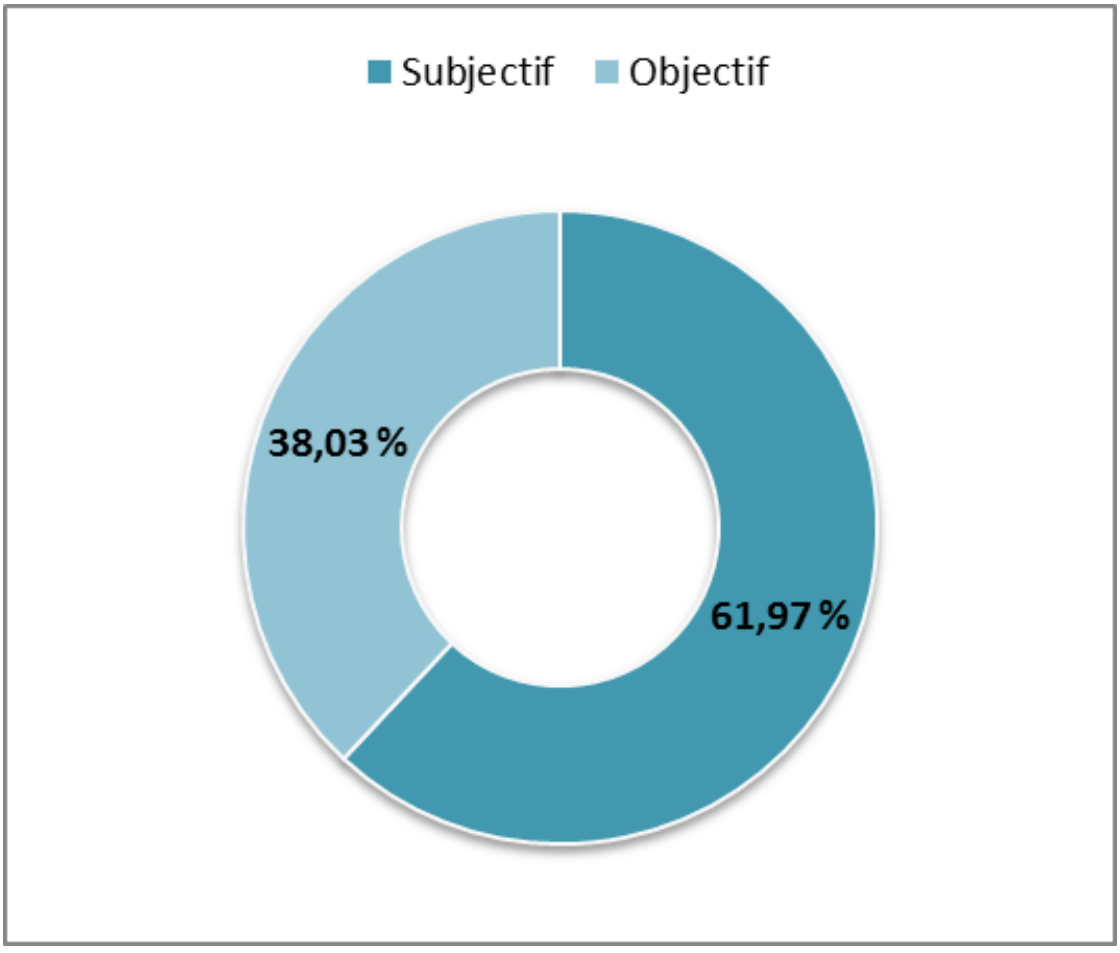

Figure 5. Répartition des annotations par catégories de types objectives et subjectives

Force est de noter que seules 38,03 % des phrases annotées du corpus arborent une pseudo-objectivité, c'est-à-dire une modalisation neutre de l'énonciateur (via les catégories sémantiques *Affirmation*, *Argumentation*, *Comparaison*, *Déduction*, *Définition, Exemplification*, *Observation*, *Synthèse*) ; à l'inverse, prédominent les phrases qui assument pleinement leur subjectivité par l'expression d'une opinion (via les catégories *Accord, Accusation*, *Appréciation*, *Croyance*, *Dénonciation*, *Dépréciation*, *Désaccord*, *Doute*, *Éloge*, *Évaluation*, *Erreur*, *Indignation, Insulte, Ironie, Déni,*

Prétention, Plainte).

En outre, les résultats de certaines catégories prédominent de façon remarquable. 28

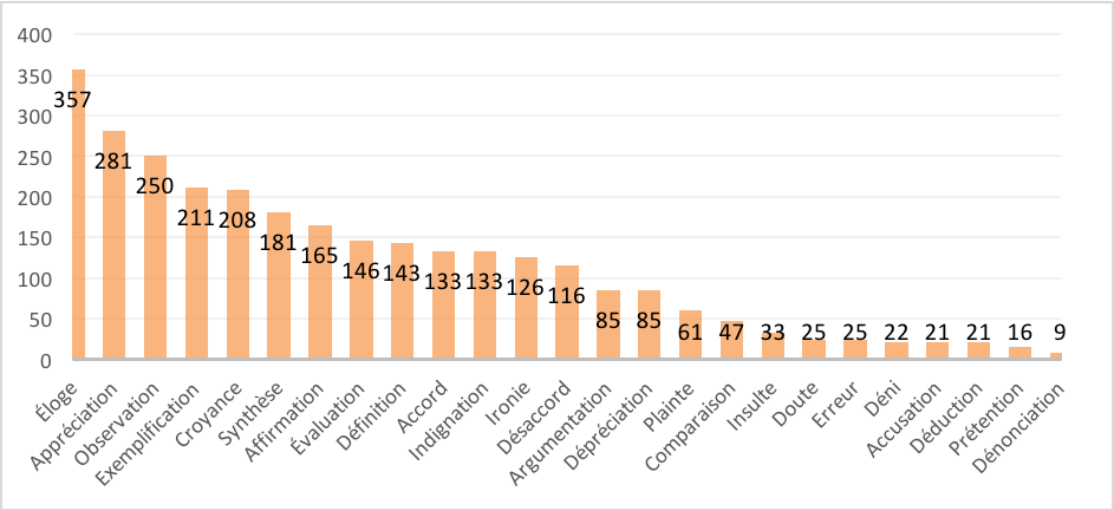

Figure 6. Répartition des annotations par catégorie sémantique dans tout le corpus

Chacune d'elles mériterait une analyse qu'il nous est difficile de tenir dans le cadre de cette étude préliminaire. Mais il nous est tout de même possible d'en esquisser ici les grandes lignes interprétatives. Notons d'ores et déjà l'importance des opinions positives (*Éloge, Appréciation, Accord*) témoignant d'une entreprise de valorisation de la citation dans une construction consensuelle ; ou celle des modélisations neutres (*Observation, Exemplification, Synthèse, Définition*) qui traduisent davantage le rôle rhétorique de la citation, ainsi mise au service de la stratégie argumentative de la critique littéraire.

La citation : élément de définition ?

La définition est un enjeu important de la critique littéraire du XIX^{e} siècle, qui tente à la fois de redéfinir son champ de compétences et de fixer les termes de son application, tout en se refusant au dogmatisme. Brunetière la pose au cœur de l'entreprise critique : « Tout le tort [de l'ancienne critique], c'était de commencer par la fin, et de poser en principe une définition du beau que son objet même est de rechercher. » [Brunetière 1914] Or, au vu des 169 annotations automatiques que compte notre catégorie Définition, il apparaît clairement que la citation intervient comme élément de définition rapportée, c'est-à-dire comme prérequis que le critique pose et à partir duquel il peut développer son propre discours. 29

Cette catégorie sémantique souligne le souci particulier du critique pour définir ses objets, et le besoin, pour ce faire, de recourir à une énonciation exogène qui lui permet d'assurer une assise à son discours. Parmi les sujets récurrents, l'esthétique revient de façon remarquable : plutôt que de définir la littérature elle-même, qui semble moins être un objet en soi qu'un champ de compétence, les critiques s'efforcent de déterminer les caractéristiques des différents courants esthétiques, et les grandes lignes formelles qui se dégagent d'un auteur ou de son œuvre. Pour exemple : 30

Œuvres	Citations annotées
[Albalat 1925]	Larroumet a bien défini Mérimée quand il a dit : *« Il était romantique par les sujets, , classique par la forme serrée, et réaliste par la vie et la crudité. »*
[Faguet 1913]	Si l'on accepte cette définition du romantisme : *« Le romantisme, c'est la prédominance de l'imagination. »*
[Faguet 1903]	Il définit l'œuvre de Zola : « *une épopée pessimiste de l'animalité humaine.* »
[Lemaitre, 1886]	Il définit l'essence du romantisme : *« l'amalgame du passé avec le présent et du présent avec le passé »*.
[Renard 1990]	C'est pour peindre M. Zola, qu'il nomme à cette occasion : *« un monstre puissant, simple et clair »*.

Table 2. Exemples annotés dans la catégorie Définition

Plutôt que de juger, la critique offre ici un véritable examen des composantes littéraires qu'elle cherche à saisir et fixer par le biais de la définition. De même, les genres littéraires font l'objet d'une définition permanente. 31

L'option d'une recherche par champ lexical nous permet de sélectionner directement sur la plateforme E-Quotes les citations modalisées traitant des domaines romanesque, poétique et théâtral. Ainsi avons-nous pu explorer l'ensemble du discours rapporté qui traite de théâtre grâce à l'inclusion de variantes, préétablies selon notre connaissance *a priori* du corpus, telles que *tragédie*, *comédie*, *mise en scène*, *dramaturgie*, etc. ; de la même manière, les définitions propres au domaine poétique sont ciblées à partir des variantes lexicales *poète*, *recueil*, *rime*, *strophe*, etc. Or, si l'on confronte les résultats obtenus pour les citations attachées à chacun des genres littéraires, ce ne sont pas des jugements radicalement distincts qui se dessinent selon le genre, mais une même prévalence de la définition. Force est de remarquer l'effort particulier de la critique pour définir les objets littéraires dont elle traite, questionner leur nature, délimiter leurs traits, leurs fonctions et leurs formes. 32

Œuvres	Citations annotées
[Morice 1889]	Acceptant cette définition de Madame Necker : *« Le roman doit être le monde meilleur »*, Balzac ajoute : *« Mais le roman ne serait rien si, dans cet auguste mensonge, il n'était pas vrai dans les détails. »*
[Rod 1892]	Le Chatterton d'Alfred de Vigny définit ainsi la fonction du poète : *« Le poète montre aux étoiles la route que trace le doigt du Seigneur. »*
[Nisard 1889]	Il définissait la tragédie : *« une action funeste qui doit conduire les spectateurs à la pitié par la terreur. »*
[Faguet 1910]	Il définit la comédie : *« l'ouvrage dramatique qui peint les mœurs des hommes dans une condition privée »*.

Table 3. Exemples annotés dans la catégorie Définition

En recourant à la définition formulée par un autre, c'est-à-dire en faisant emploi d'un pré-discours déjà existant, le discours critique s'inscrit pleinement dans un « dialogisme » au sens bakhtinien [Todorov et Bakhtine 1978] ; ceci présente le double avantage d'ancrer le propos du critique dans une sphère collective, et de lui permettre d'exprimer son opinion vis-à-vis d'un énoncé sans en assumer la pleine autorité. En somme, définir par le biais de l'énonciation rapportée permet au critique de s'engager dans une communauté et de se positionner par rapport à une vérité admise, qu'il relaie sans pour autant la prendre en charge [Desclés 2009]. Recourir à la citation tend dès lors à prétendre, sinon à une objectivité – que l'on pourrait remettre en cause du simple fait qu'il y a prise en charge d'un énoncé – du moins à une certaine neutralité du discours, qui dépasse la subjectivité du locuteur par le biais d'une énonciation plurielle et en partie extérieure. La critique littéraire construit ainsi une posture impartiale à travers laquelle elle peut s'affilier à un discours sur la connaissance. 33

Enfin, concernant les définitions, on peut également noter que la littérature dépasse son champ d'étude propre pour 34

aller puiser hors d'elle des définitions d'objets exogènes mais qui viennent enrichir son discours en élargissant ses frontières.

Œuvres	Citations annotées
[Deschanel 1864]	Peut-être même qu'à prendre pour flambeau cette définition du docteur Halley, « *l'homme est la partie musculaire de l'humanité, la femme en est la partie nerveuse*, » il y aurait lieu, qui sait ? de distinguer le sexe chez les peuples ou les nations comme chez les individus.
[Renard 1894]	C'est Deslon qui définissait le magnétisme *« l'action de la volonté sur la matière animée ».*
[Nordau 1894]	Au degré le plus haut enfin naît l'extase, que M. Ribot nomme *« la forme aiguë de la tendance à l'unité de la conscience ».*
[Gourmont 1925]	Monsieur Tarde, philosophe ingénieux et amer a défini la vie : *« La poursuite de l'impossible à travers l'inutile. »*

Table 4. Exemples annotés dans la catégorie Définition

35 On retrouve ici des notions au confluent de différents domaines, et notamment des sciences en essor, telles que l'*homme* et la *vie*, qui se trouvent alors au cœur de redéfinitions biologiques et médicales ; d'autres références sortent, elles, complètement du champ de compétence de la critique, à l'instar du *magnétisme* ou de l'*extase*, et de tels emprunts soulignent l'important dialogue interdisciplinaire dans lequel la critique littéraire attache son discours. En citant philosophes et scientifiques, elle s'engage là encore dans une communauté, mais une communauté plus vaste, qui dépasse celle du champ littéraire pour entrer dans le champ élargi des savoirs. De la sorte, le positionnement de la critique n'est plus seulement esthétique mais obéit à une ambition épistémologique double, en se voulant, d'une part, discours savant vis-à-vis de l'objet littéraire qu'elle est en charge de définir et, d'autre part, acteur d'un dialogue interdisciplinaire qui vient nourrir le champ de la connaissance.

Quel jugement critique ?

36 En plus de ce souci de définition, la spécificité de notre corpus réside dans la question du jugement critique, c'est-à-dire dans la façon dont le critique évalue les propos d'autrui en l'intégrant au sein d'une appréciation personnelle. À partir des distinctions faites par [Kerbrat-Orecchioni 1980], on peut dégager deux natures d'évaluation : l'un de type bon/mauvais, l'autre de type vrai/faux/incertain. Selon nos résultats d'annotation automatique, les deux types se répartissent au sein du corpus de la façon suivante :

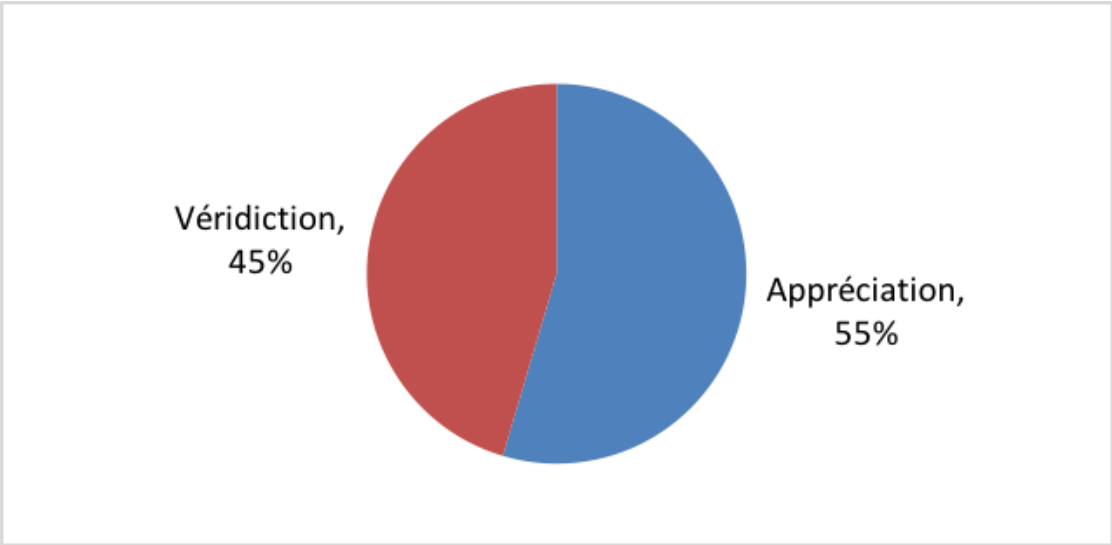

Figure 7. Répartition des annotations du corpus par type d'évaluation

L'évaluation de type bon/mauvais se laisse voir via les catégories sémantiques qui appartiennent plus spécifiquement au domaine axiologique ou appréciatif : *Appréciation*, *Éloge*, *Dépréciation*, *Indignation*, *Insulte.* Ces citations modalisées véhiculent des jugements critiques destinés à établir les qualités et défauts à partir desquels il serait possible de déterminer la valeur d'une œuvre ou d'un auteur, et de distinguer plus généralement une « bonne » d'une « mauvaise » littérature. 37

Œuvres	Citations annotées	Catégories
[Rod 1892]	… mais cela n'empêche qu'il faille tout lui reprocher : *« le manque d'esprit et de délicatesse, l'absence d'âme et de passion, l'abus des descriptions, la prédilection pour les corruptions sociales, un style à la fois laborieux et incolore ».*	Indignation
[Taine 1875]	*« Cela n'a pas de nom, cela ne ressemble à rien »,* tel est le blâme le plus fort.	Indignation
[Lemaitre 1887]	Et là-dessus on cite, à la vérité, de fort beaux paysages, encore que très sobres, de Théocrite, de Lucrèce, d'Horace, de Virgile, surtout l'admirable cri des Géorgiques: *« Oh ! les champs, le Sperchius, le Taygète foulé par les danses des filles de Sparte !… »*	Appréciation

Table 5. Exemples de l'ordre de l'évaluation bon/mauvais

Autour de ces citations, on constate que les critères de l'évaluation critique sont non seulement esthétiques, mais également éthiques, intellectuels, référentiels, etc. Une telle exploration permet donc d'observer les constituantes de la valeur littéraire qui émergent durant le XIX^{e} siècle. Dans la catégorie *Indignation* par exemple, se retrouve exprimée en filigrane la question du bon goût, partie prenante du jugement critique de l'époque. 38

L'expression de l'adhésion, qui s'observe via certaines catégories (*Accord*, *Affirmation*, *Accusation*, *Dénonciation*, *Désaccord*, *Erreur*, *Déni*, *Prétention*, *Croyance*, *Doute*), inscrit en revanche le jugement dans l'ordre de la véridiction, c'est-à-dire que le locuteur évalue la justesse ou l'inexactitude de la citation qu'il introduit par rapport à la notion de vérité : 39

Œuvres	Citations annotées	Catégories
[Janet 1872]	Cependant il reconnaît que Bossuet s'est trompé sur deux points : *« Il s'est trompé quand il a cru le protestantisme incompatible avec de grandes sociétés réglées et prospères ; il s'est trompé quand il a vu l'idéal des gouvernements dans la royauté absolue tempérée par des lois fondamentales. »*	Erreur
[Albalat 1905]	… rien n'est plus faux que de dire : *« Ceci est de l'art parce qu'on ne sent pas la rhétorique, et ceci n'est pas de l'art parce qu'on sent la rhétorique. »*	Désaccord
[Nisard 1889]	Voltaire a raison : *« Les bons ouvrages sont ceux qui font le plus pleurer. »*	Accord

Table 6. Exemples de l'ordre de l'adhésion

Les annotations automatiques rendent ainsi visible la façon dont le critique prend position par rapport à un énoncé non plus en tant que juge mais en tant que spécialiste, voire en tant que savant, dans une dimension non plus esthétique ou éthique, mais épistémique. Sa propre appréciation prend alors d'autant plus de poids qu'elle imite une position déterminée par rapport à une vérité universelle. On voit ici la stratégie par le biais de laquelle la critique littéraire légitime son discours, en l'ancrant sur un terrain épistémique plutôt qu'affectif. De cette manière, on observe un certain nombre de citations modalisées obéissant à ce qu'on pourrait appeler avec Kerbrat-Orecchioni une « subjectivité objectivée », où l'énonciateur prend position sans s'avouer ouvertement comme la source du jugement évaluatif. Lorsque l'annotation automatique de notre système donne lieu à une double catégorisation sémantique, par exemple, on peut clairement observer une subjectivité exprimée sous couvert du discours rapporté, plus discrète, qui perce notamment au travers d'adverbes, mais qui, pourrait-on dire, avance masquée, sous la neutralité apparente de l'affirmation ou de l'observation rapportée. Les citations suivantes en sont une bonne illustration : 40

Œuvres	Citations annotées	Catégories
[Morice 1889]	Comme dit très justement M. Mendès : *« Je ne crois pas qu'il ait jamais existé un homme plus intimement, plus essentiellement poète que M. Léon Dierx. »*	Affirmation Accord
[Heumann 1913]	On y rencontre aussi de gracieux tableaux d'intérieur, et Liebrecht a pu fort justement comparer Kaatje à *« un Terburg en rupture de cadre »*.	Comparaison Accord
[Bazalgette 1898]	Ruskin, comme l'a très bien noté M. J. Milsand, est *« un chaos intérieur de vitalités désordonnées et de déraison indomptable. »*	Observation Accord

Table 7. Exemples annotés dans deux catégories

Plus discrètement encore, la subjectivité peut disparaître de toute manifestation visible ; elle ne transparaît plus qu'implicitement, par le seul fait que l'énonciateur prend en charge le contenu de l'énoncé.

Œuvre	Citation annotée	Catégorie
[Deschanel 1864]	Voltaire, et c'est Victor Hugo qui en fait la remarque, *« Voltaire, si grand au XVIII^e^ siècle, est plus grand encore au XIX^e^… »*	Observation

Table 8. Exemple annoté dans la catégorie Observation

Ces cas sont probablement les plus récurrents dans le corpus critique : autrement dit, ce sont dans les phrases *a priori* neutres, annotées comme des affirmations, définitions ou observations, que se lit l'adhésion du critique qui rapporte la citation, du simple fait qu'il la rapporte et l'intègre à son discours en s'appuyant sur elle sans la contredire. La citation sert d'arme au critique, puisqu'elle appuie sa cause en lui évitant de prendre directement en charge l'énoncé. 41

Un des premiers moyens auxquels a recours la critique littéraire pour légitimer son discours est alors la construction d'un consensus, dans le sens où les critiques n'hésitent pas à prendre position grâce à la citation d'un autre : quand Brunetière se place dans les sillons de Sainte-Beuve ou que Deschanel rapporte l'observation de Taine, l'un et l'autre consolident leur position personnelle par son inclusion au sein d'une communauté. 42

Œuvres	Citations annotées	Catégories
[Albalat 1905]	*« Le style de Taine, confirme M. Faguet, est un miracle de volonté… »*	Affirmation
[Faguet 1898]	Scherer n'a pas eu tort de dire : *« Il y a du Rousseau dans cet homme-là. »*	Accord
[Deschanel 1864]	M. Taine en fait bien la remarque : *« Par-dessous l'amateur du Moyen Age… »*	Observation
[Bourget 1928]	Nous en avons une preuve saisissante dans son article sur Taine à ses débuts : *« Que le savant en lui, »* dit-il à propos des *Essais de critique et d'histoire « ne domine pas trop le littérateur. »*	Évaluation

Table 9. Exemples de citations entre auteurs

Cette construction discursive accompagne la formation d'un métadiscours au travers duquel la critique conquiert son autonomie et s'émancipe des discours environnants. La citation de l'autre devient un argument de poids, comme présupposé incontestable, affirmé, qui prend activement part à la stratégie de validation de toute une communauté. Elle vient redoubler l'opinion du critique en lui offrant une sorte de caution extérieure, une accréditation, au point de faire office d'argument d'autorité et de donner ainsi à l'énoncé une valeur scientifique partagée par l'ensemble des acteurs du domaine littéraire. 43

Mise en perspective : critique et positivisme

Suite à ces premiers constats, il convient de nous interroger sur la nature même de ces stratégies argumentatives. Si l'on retrouve un emploi citationnel proche d'une rhétorique des sciences (Martin, 1996), peut-on en conclure pour autant que la critique littéraire fait ici preuve de singularité ? Vient-elle s'inscrire ouvertement dans une démarche positiviste, prend-elle appui sur certains discours scientifiques qui l'entourent afin de gommer toute critique de goût ? Pour éprouver ces hypothèses, il nous est possible de réaliser un premier sondage à partir de nos résultats, en formant deux sous-corpus : l'un concentrant les critiques dits « positivistes », qui revendiquent d'une façon ou d'une autre l'influence des sciences dans leur entreprise critique (Sainte-Beuve, Renan, Taine, Hennequin, Deschanel, Ghil, Zola, Renard, etc.) ; l'autre réunissant à l'inverse les critiques qui s'érigent contre le scientisme, qu'ils soient romantiques, impressionnistes, ou symbolistes (Lamartine, Anatole France, Barbey d'Aurevilly, Albalat, Lemaitre, etc.). Or, dans ces deux sous-corpus, les annotations automatiques se répartissent de la façon suivante : 44

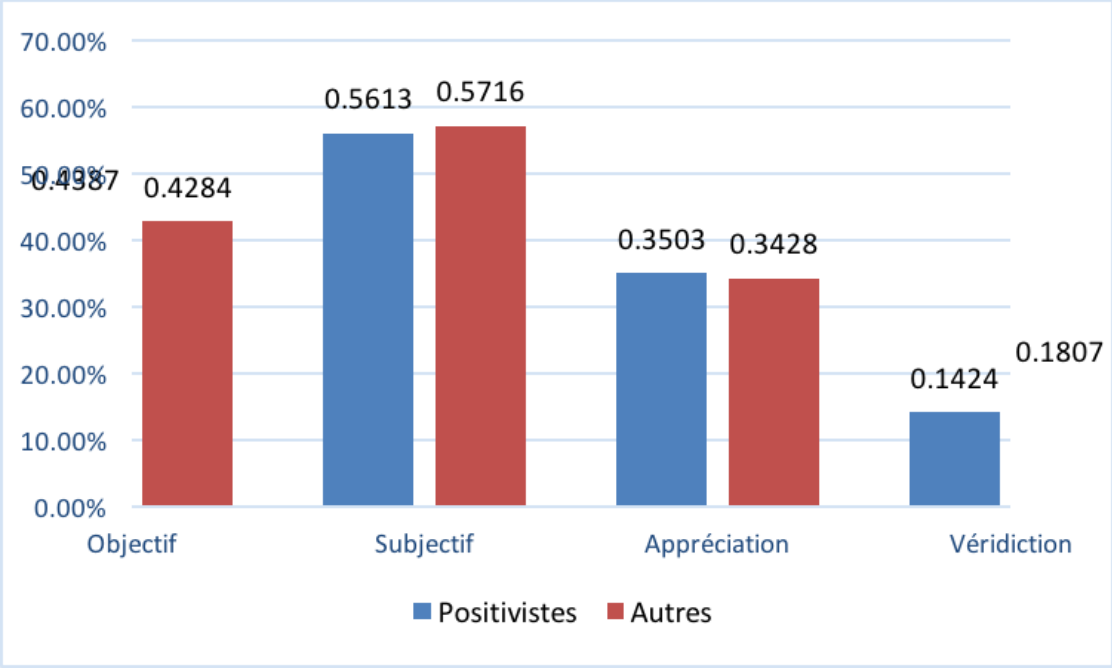

Figure 8. Répartition des annotations par sous-corpus

Bien que cette exploration préliminaire mérite d'être complétée par une étude comparatiste plus large en diachronie, elle nous paraît déjà éloquente dans la mesure où elle révèle une certaine homogénéité entre les annotations des deux sous-corpus. Autrement dit, les critiques d'influence positiviste ont autant recours à la modélisation subjective que les autres (avec une différence infime de 1 %) ; plus encore, en dépit d'une ambition ouvertement épistémologique, ils passent moins par une évaluation de type véridictoire (avec 14,24 % de phrases annotées, contre 18,07 % pour les critiques non positivistes), et davantage par une évaluation de type bon/mauvais. Aussi, ces statistiques invitent à repenser la stratégie discursive de la critique, non pas tant comme une démarche particulière attachée à un courant esthétique et idéologique, mais peut-être davantage comme un phénomène global de formation discursive. La citation serait, pour l'ensemble du discours critique de la seconde moitié du siècle, un moyen d'asseoir sa positivité au sens foucaldien, c'est-à-dire de légitimer son énonciation en l'inscrivant dans un triple domaine de *validité*, de *normativité* et d'*actualité* [Foucault 1969]. 45

Conclusion et perspectives

Une telle méthode, on le voit, sert l'analyse de discours en ouvrant au traitement de larges corpus. Dans le cadre plus spécifique de notre étude, la fouille de texte automatisée permet de faire émerger les stratégies par le biais desquelles la littérature forme son discours. En présentant une dimension épistémique, le jugement critique du XIX^{e} siècle peut pleinement prétendre à une production du savoir. Ainsi retrouve-t-on l'engagement d'une critique moderne qui fait vœu 46

d'impartialité en se fixant la haute tâche d'« analyser, expliquer et classer » [Brunetière 1914] la littérature. Dans un dessein plus global, il s'agirait à terme d'appliquer ces catégories linguistiques à des corpus de différentes époques, de manière à dégager les grandes caractéristiques communes et les singularités de chaque auteur, courant ou siècle.

47 D'un point de vue technique, E-Quotes nous a permis de rendre les informations sémantiques accessibles pour un utilisateur final (*linguistes*, *chercheurs en humanités numériques*, *journalistes*…). Nous l'avons démontré en appliquant les technologies classiques de la recherche d'information (*requêtes de mots clés sur un sac de mots*) à des textes sémantiquement annotés. Par conséquent, l'utilisateur a la possibilité d'effectuer des recherches au sein des citations catégorisées en fonction de l'attitude, l'opinion ou la position du locuteur telle que décrite, vue ou commentée par l'énonciateur. Malgré la simplicité de notre approche – pas d'analyse morphologique ou syntaxique, ni de pré-reconnaissance des entités nommées –, le système EXCOM-2 a pu atteindre un taux de *précision* de 87,6 % sur l'ensemble des catégories évaluées. D'autres travaux nous permettrons d'améliorer encore la couverture des règles heuristiques et des marqueurs linguistiques, notamment pour les catégories *Accord*, *Accusation* et *Doute.*

48 Il reste que les résultats basés uniquement sur la *précision* dépendent des ressources employées (*couverture*, *pertinence*…) et ne permettent pas de découvrir dans un corpus tous les cas d'utilisation des citations. Pour cette raison, nous envisageons de construire un corpus de référence (*Gold Standard Corpus*) pour différentes catégories de notre cartouche linguistique, ce qui nous permettrait, entre autres, de calculer le *rappel*. Notre objectif à moyen terme est de connecter EXCOM-2 et E-Quotes dans une version en ligne libre de droit où l'utilisateur a le moyen de soumettre son propre corpus, de le faire traiter directement (*annotation et indexation*) et d'en exploiter les résultats.

Remerciements

49 Nous remercions les trois évaluateurs anonymes pour avoir contribué à ce travail. Ce travail a bénéficié d'une aide d'État gérée par l'Agence nationale de la recherche dans le cadre des Investissements d'avenir portant la référence ANR-11-IDEX-0004-02.

Notes

[1] Le corpus fait environ 23 millions de mots, 1 200 000 phrases et 140 Mo.

[2] http://obvil.paris-sorbonne.fr/corpus/critique.

[3] Une version basique d'EXCOM-2 est librement accessible à la demande sur l'adresse http://www.excom.fr/.

[4] http://www.e-quotes.net.

[5] http://emm.newsexplorer.eu/NewsExplorer/home/fr/latest.html.

[6] Nous avons construit trois règles en moyenne pour chaque catégorie.

[7] Après les premiers tests, nous avons décidé d'écarter tous les passages entre guillemets qui contenaient moins de quatre mots afin d'éviter les *fausses* citations, notamment les expressions et les mises en valeur typographiques, telles que le terme « villa » dans : *Jean a dit que sa « villa » était en mauvais état*. [Authier-Revuz 1992].

[8] Le guide peut être consulté à l'adresse suivante : https://docs.google.com/document/d/17nedlR3V9kpvmOjjZUuLFPi2MVgCjBwr7V-v6q8Tg7g/edit#heading=h.lxswbydbnspn.

Works Cited

Albalat 1905 Albalat, Antoine. *Les Ennemis de l'art d'écrire: réponse aux objections de MM. F. Brunetière, Émile Faguet, Adolphe Brisson, etc.* Paris, Librairie universelle (1905).

Albalat 1925 Albalat, Antoine. *Comment on devient écrivain.* Paris, Plon (1925).

Alexandre 2016 Alexandre Didier (dir.). "Études littéraires et calcul numérique. Présentation", *Revue d'histoire littéraire de la France*, 3, juillet–septembre 2016, Paris. p. 517–520. (2016).

Alrahabi 2010 Alrahabi, Motasem. *EXCOM-2: plateforme d'annotation automatique de catégories sémantiques. Applications à la catégorisation des citations en français et en arabe*. Thèse de doctorat, Université Paris-Sorbonne (2010).

Alrahabi 2015 Alrahabi Motasem. "E-Quotes: Enunciative Modalities Analysis Tool for Direct Reported Speech in Arabic". *Actes de la 16th International Conference on Intelligent Text Processing and Computational Linguistics* (CICLING). Cairo, Egypt (2015).

Alrahabi 2016 Alrahabi, Motasem. "E-Quotes: un outil de navigation textuelle guidée par les annotations sémantiques". *Actes de TALN 2016*. Paris (2016).

Alrahabi, Desclés et Suh 2010 Alrahabi Motasem, Desclés Jean-Pierre, Suh, Jy-Yeon. "Direct Reported Speech in multilingual texts: Automatic Annotation and Semantic Categorization". *Actes de FLAIRS 2010*. Floride (2010).

Atanassova, Bertin et Desclés 2012 Atanassova Iana, Bertin Marc, Desclés Jean-Pierre. "Ordonnancement des réponses pour une recherche d'information sémantique à partir d'une ontologie discursive". *Recherche d'information sémantique RISE–2012*. 12e Conférence Internationale Francophone sur l'Extraction et la Gestion de Connaissance (EGC-2012). Bordeaux (2012).

Authier-Revuz 1992 Authier-Revuz Jacqueline. "Repères dans le champ du discours rapporté". *L'information grammaticale*, no55, oct.92, p. 38-42 (1992).

Bazalgette 1898 Bazalgette, Léon. *L'Esprit nouveau dans la vie artistique, sociale et religieuse.* Paris, Société d'éditions littéraires (1898).

Bergler 1992 Bergler, Sabine. *The Evidential Analysis of Reported Speech*. Thèse de Doctorat, Brandeis University, Massachusetts (1992).

Bourget 1928 Bourget, Paul. *Quelques témoignages*. Paris, Plon (1928).

Brunetière 1914 Brunetière, Ferdinand. *L'Évolution des genres dans l'histoire de la littérature*. Paris, Hachette et cie (1914).

Bénard et Alrahabi 2017 Bénard, Élodie, Alrahabi, Motasem. "Du petit fait à la doxa: annotation automatique des anecdotes dans le discours critique sur Molière". *Actes de Journées de Linguistique de Corpus*. Grenoble (2017).

Charaudeau 2005 Charaudeau, Patrick. *Le Discours d'information médiatique. L'impossible transparence*. Bruxelles, De Boeck Université (2005).

Charron et Jacob 1999 Charron, Jean, Jacob, Loïc. "Énonciation journalistique et subjectivité: les marques du changement". *Études de communication publique*. 14 (1999): 70.

Compagnon 1979 Compagnon, Antoine. *La Seconde Main ou le travail de la citation*. Paris, Seuil (1979).

Coulanges 1864 Coulanges, Foustel (de). *La Cité antique.* Strasbourg, Silberman (1864).

Deschanel 1864 Deschanel, Émile. *Physiologie des écrivains et des artistes ou Essai de critique naturelle*. Hachette et Cie, Paris (1864).

Desclés 2006 Desclés, Jean-Pierre. "Contextual Exploration Processing for Discourse Automatic Annotations of Texts". *Actes de FLAIRS 2006*. Floride (2006).

Desclés 2009 Desclés, Jean-Pierre. "Prise en charge, engagement et désengagement". *Langue française*. 162 (2009): 29–53.

Desclés, Alrahabi et Desclés 2011 Desclés, Julien, Alrahabi, Motasem, Desclés Jean-Pierre. "BioExcom: Detection and Categorization of Speculative Sentences in Biomedical Literature". *Human Language Technology. Challenges for Computer Science and Linguistics Lecture Notes in Computer Science*. 6562 (2011): 478–489.

Du Casse 1864 Du Casse, Albert. *Histoire anecdotique de l'ancien théâtre en France: Théâtre-Français, Opéra, Opéra-Comique, Vaudeville, Théâtre-Italien, théâtres forains, etc.*, t. I. Paris, E. Dentu (1864).

Faguet 1898 Faguet, Émile. *Politiques et Moralistes du dix-neuvième siècle,* t. II. Paris, Société française d'imprimerie et de librairie (1898).

Faguet 1903 Faguet, Émile. *Zola.* Paris, A. Eymeoud (1903).

Faguet 1904 Faguet, Émile. *En lisant Nietzsche*, Paris, Société française d'imprimerie et de librairie (1904).

Faguet 1910 Faguet, Émile. *Rousseau contre Molière*. Paris, Société française d'imprimerie et de librairie (1910).

Faguet 1913 Faguet, Émile. *La Fontaine.* Paris, Société française d'imprimerie et de librairie (1913).

Foucault 1969 Foucault, Michel. *L'Archéologie du savoir*. Paris, Gallimard (1969).

France 1892 France, Anatole. *La Vie littéraire. Quatrième série*. Paris, Calmann-Lévy (1892).

Gourmont 1925 Gourmont, Remy (de). *Promenades philosophiques. Première série.* Mercure de France, Paris (1925).

Guyau 1884 Guyau, Jean-Marie. *Les Problèmes de l'esthétique contemporaine*. Paris, Alcan (1884).

Heumann 1913 Heumann, Albert. *Le Mouvement littéraire belge d'expression française depuis 1880*. Paris, Mercure de France (1913).

Janet 1872 Janet, Paul. *Les Problèmes du XIX^e^ siècle: la politique, la littérature, la science, la philosophie, la religion*. Paris, Michel Lévy frères (1872).

Kerbrat-Orecchioni 1980 Kerbrat-Orecchioni, Catherine. *L'Énonciation. De la subjectivité dans le langage*. Paris, Armand Colin (1980).

Lemaitre 1887 Lemaitre, Jules. *Les Contemporains : études et portraits littéraires. Troisième série.* Paris, H. Lecène et H. Oudin (1887).

Lemaitre 1891 Lemaitre, Jules. *Impressions de théâtre: cinquième série.* Paris, H. Lecène et H. Oudin (1891).

Lemaitre, 1886 Lemaitre, Jules. *Les Contemporains: études et portraits littéraires. Deuxième série.* Paris, H. Lecène et H. Oudin (1886).

Makkaoui, Desclés, Bertin, Jouis et Ganascia 2014 Makkaoui, Olfa, Desclés, Jean-Pierre, Bertin, Marc, Jouis, Christophe, Ganascia Jean-Gabriel. "Discursive Mining Viewpoints in Building Multi-Document Synthesized Sheets". *Actes de la Twenty-Seventh International FLAIRS Conference*. Floride (2014).

Martin 1996 Martin, Jacky. "Les enjeux du discours scientifique: la stratégie de véridiction". *ASp.* 11–14 (1996). Disponible en ligne: http://asp.revues.org/3414.

Morice 1889 Morice, Charles. *La Littérature de Tout à l'heure*. Paris, Perrin (1889).

Nisard 1889 Nisard, Désirée. *Histoire de la littérature française.* t. IV. Paris, Firmin Didot (1889).

Nordau 1894 Nordau, Max. *Dégénérescence, t. I. Fin de siècle — Le Mysticisme*. Paris, Félix Alcan (1894).

Renard 1894 Renard, Georges. *Critique de combat.* Paris, Dentu (1894).

Renard 1990 Renard, Georges. *Les Princes de la jeune critique: Jules Lemaître, Ferdinand Brunetière, Anatole France, Louis Ganderax, Paul Bourget*. Paris, Librairie de la " Nouvelle Revue " (1890).

Rod 1892 Rod, Edouard. *Les Idées morales du temps présent.* Paris, Perrin (1892).

Rosier 1999 Rosier, Laurence. *Le Discours rapporté. Histoire, théories, pratiques*. Paris-Bruxelles, De Boeck-Duculot (1999).

Saint-Victor 1882 Saint-Victor, Paul (de). *Hommes et Dieux. Études d'histoire et de littérature.* Paris, Calmann-Lévy (1882).

Sainte-Beuve 1864 Sainte-Beuve, Charles Augustin. *Portraits littéraires, 3*. Paris, Garnier frères (1864).

Sainte-Beuve 1867 Sainte-Beuve, Charles-Augustin. *Nouveaux Lundis, 8*. Paris, Michel Lévy frères (1867).

Taine 1875 Taine, Hippolyte. *Les Origines de la France contemporaine. L'ancien régime.* Paris (1875).

Todorov et Bakhtine 1978 Todorov, Tzvetan, Bakhtine, Mikhaïl. *Le Principe dialogique*. Paris, Seuil (1978).

Tuomarla 2000 Tuomarla, Ulla. *La Citation mode d'emploi. Sur le fonctionnement discursif du discours rapporté direct*. Thèse de doctorat, Academia Scientifiarum Fennica. Ser. Humaniora. tom. 308. Saarijärvi, Finland (2000).

Zola 1881 Zola, Émile. *Le Naturalisme au théâtre: les théories et les exemples*. Paris, Charpentier (1881).

This work is licensed under a Creative Commons Attribution-NoDerivatives 4.0 International License.

DHQ: Digital Humanities Quarterly

2018
Volume 12 Number 1

Toward an automatic analysis of critical judgment: modified quotes in the literary discourse of the nineteenth century [fr]

Marine Riguet <marineriguet_at_gmail_dot_com>, Labex OBVIL (Paris-Sorbonne)
Motasem Alrahabi <motasem_dot_alrahabi_at_gmail_dot_com>, Paris-Sorbonne Abou Dhabi

Abstract

In the nineteenth century, under the influence of positivism and emerging sciences, French literary critics sought to assert the autonomy of their discourse. In this work of legitimation, the use of textual citations supported the strategies of a critical judgment in search of objectivity and scientificity. Based on a large corpus of literary texts, this article attempts to automatically annotate the modal expressions used around reported speech. To this end, we will present two applications through which certain characteristic features of the literary critical discourse have been identified and analyzed: EXCOM-2, an annotation system of semantic and discursive categories; and E-Quotes, its web interface for searching citations.

Note on Translation

For articles in languages other than English, DHQ provides an English-language abstract to support searching and discovery, and to enable those not fluent in the article's original language to get a basic understanding of its contents. In many cases, machine translation may be helpful for those seeking more detailed access. While DHQ does not typically have the resources to translate articles in full, we welcome contributions of effort from readers. If you are interested in translating any article into another language, please contact us at editors@digitalhumanities.org and we will be happy to work with you.

Works Cited

Albalat 1905 Albalat, Antoine. *Les Ennemis de l'art d'écrire: réponse aux objections de MM. F. Brunetière, Émile Faguet, Adolphe Brisson, etc.* Paris, Librairie universelle (1905).

Albalat 1925 Albalat, Antoine. *Comment on devient écrivain.* Paris, Plon (1925).

Alexandre 2016 Alexandre Didier (dir.). "Études littéraires et calcul numérique. Présentation", *Revue d'histoire littéraire de la France*, 3, juillet–septembre 2016, Paris. p. 517–520. (2016).

Alrahabi 2010 Alrahabi, Motasem. *EXCOM-2: plateforme d'annotation automatique de catégories sémantiques. Applications à la catégorisation des citations en français et en arabe*. Thèse de doctorat, Université Paris-Sorbonne (2010).

Alrahabi 2015 Alrahabi Motasem. "E-Quotes: Enunciative Modalities Analysis Tool for Direct Reported Speech in Arabic". *Actes de la 16th International Conference on Intelligent Text Processing and Computational Linguistics* (CICLING). Cairo, Egypt (2015).

Alrahabi 2016 Alrahabi, Motasem. "E-Quotes: un outil de navigation textuelle guidée par les annotations sémantiques". *Actes de TALN 2016*. Paris (2016).

Alrahabi, Desclés et Suh 2010 Alrahabi Motasem, Desclés Jean-Pierre, Suh, Jy-Yeon. "Direct Reported Speech in multilingual texts: Automatic Annotation and Semantic Categorization". *Actes de FLAIRS 2010*. Floride (2010).

Atanassova, Bertin et Desclés 2012 Atanassova Iana, Bertin Marc, Desclés Jean-Pierre. "Ordonnancement des réponses pour une recherche d'information sémantique à partir d'une ontologie discursive". *Recherche d'information*

sémantique RISE–2012. 12e Conférence Internationale Francophone sur l'Extraction et la Gestion de Connaissance (EGC-2012). Bordeaux (2012).

Authier-Revuz 1992 Authier-Revuz Jacqueline. "Repères dans le champ du discours rapporté". *L'information grammaticale*, no55, oct.92, p. 38-42 (1992).

Bazalgette 1898 Bazalgette, Léon. *L'Esprit nouveau dans la vie artistique, sociale et religieuse.* Paris, Société d'éditions littéraires (1898).

Bergler 1992 Bergler, Sabine. *The Evidential Analysis of Reported Speech*. Thèse de Doctorat, Brandeis University, Massachusetts (1992).

Bourget 1928 Bourget, Paul. *Quelques témoignages*. Paris, Plon (1928).

Brunetière 1914 Brunetière, Ferdinand. *L'Évolution des genres dans l'histoire de la littérature*. Paris, Hachette et cie (1914).

Bénard et Alrahabi 2017 Bénard, Élodie, Alrahabi, Motasem. "Du petit fait à la doxa: annotation automatique des anecdotes dans le discours critique sur Molière". *Actes de Journées de Linguistique de Corpus*. Grenoble (2017).

Charaudeau 2005 Charaudeau, Patrick. *Le Discours d'information médiatique. L'impossible transparence*. Bruxelles, De Boeck Université (2005).

Charron et Jacob 1999 Charron, Jean, Jacob, Loïc. "Énonciation journalistique et subjectivité: les marques du changement". *Études de communication publique*. 14 (1999): 70.

Compagnon 1979 Compagnon, Antoine. *La Seconde Main ou le travail de la citation*. Paris, Seuil (1979).

Coulanges 1864 Coulanges, Foustel (de). *La Cité antique.* Strasbourg, Silberman (1864).

Deschanel 1864 Deschanel, Émile. *Physiologie des écrivains et des artistes ou Essai de critique naturelle*. Hachette et Cie, Paris (1864).

Desclés 2006 Desclés, Jean-Pierre. "Contextual Exploration Processing for Discourse Automatic Annotations of Texts". *Actes de FLAIRS 2006*. Floride (2006).

Desclés 2009 Desclés, Jean-Pierre. "Prise en charge, engagement et désengagement". *Langue française*. 162 (2009): 29–53.

Desclés, Alrahabi et Desclés 2011 Desclés, Julien, Alrahabi, Motasem, Desclés Jean-Pierre. "BioExcom: Detection and Categorization of Speculative Sentences in Biomedical Literature". *Human Language Technology. Challenges for Computer Science and Linguistics Lecture Notes in Computer Science*. 6562 (2011): 478–489.

Du Casse 1864 Du Casse, Albert. *Histoire anecdotique de l'ancien théâtre en France: Théâtre-Français, Opéra, Opéra-Comique, Vaudeville, Théâtre-Italien, théâtres forains, etc.*, t. I. Paris, E. Dentu (1864).

Faguet 1898 Faguet, Émile. *Politiques et Moralistes du dix-neuvième siècle,* t. II. Paris, Société française d'imprimerie et de librairie (1898).

Faguet 1903 Faguet, Émile. *Zola.* Paris, A. Eymeoud (1903).

Faguet 1904 Faguet, Émile. *En lisant Nietzsche*, Paris, Société française d'imprimerie et de librairie (1904).

Faguet 1910 Faguet, Émile. *Rousseau contre Molière*. Paris, Société française d'imprimerie et de librairie (1910).

Faguet 1913 Faguet, Émile. *La Fontaine.* Paris, Société française d'imprimerie et de librairie (1913).

Foucault 1969 Foucault, Michel. *L'Archéologie du savoir*. Paris, Gallimard (1969).

France 1892 France, Anatole. *La Vie littéraire. Quatrième série*. Paris, Calmann-Lévy (1892).

Gourmont 1925 Gourmont, Remy (de). *Promenades philosophiques. Première série.* Mercure de France, Paris (1925).

Guyau 1884 Guyau, Jean-Marie. *Les Problèmes de l'esthétique contemporaine*. Paris, Alcan (1884).

Heumann 1913 Heumann, Albert. *Le Mouvement littéraire belge d'expression française depuis 1880*. Paris, Mercure de France (1913).

Janet 1872 Janet, Paul. *Les Problèmes du XIX^e^ siècle: la politique, la littérature, la science, la philosophie, la religion*. Paris, Michel Lévy frères (1872).

Kerbrat-Orecchioni 1980 Kerbrat-Orecchioni, Catherine. *L'Énonciation. De la subjectivité dans le langage*. Paris, Armand Colin (1980).

Lemaitre 1887 Lemaitre, Jules. *Les Contemporains : études et portraits littéraires. Troisième série.* Paris, H. Lecène et H. Oudin (1887).

Lemaitre 1891 Lemaitre, Jules. *Impressions de théâtre: cinquième série.* Paris, H. Lecène et H. Oudin (1891).

Lemaitre, 1886 Lemaitre, Jules. *Les Contemporains: études et portraits littéraires. Deuxième série.* Paris, H. Lecène et H. Oudin (1886).

Makkaoui, Desclés, Bertin, Jouis et Ganascia 2014 Makkaoui, Olfa, Desclés, Jean-Pierre, Bertin, Marc, Jouis, Christophe, Ganascia Jean-Gabriel. "Discursive Mining Viewpoints in Building Multi-Document Synthesized Sheets". *Actes de la Twenty-Seventh International FLAIRS Conference*. Floride (2014).

Martin 1996 Martin, Jacky. "Les enjeux du discours scientifique: la stratégie de véridiction". *ASp.* 11–14 (1996). Disponible en ligne: http://asp.revues.org/3414.

Morice 1889 Morice, Charles. *La Littérature de Tout à l'heure*. Paris, Perrin (1889).

Nisard 1889 Nisard, Désirée. *Histoire de la littérature française.* t. IV. Paris, Firmin Didot (1889).

Nordau 1894 Nordau, Max. *Dégénérescence, t. I. Fin de siècle — Le Mysticisme*. Paris, Félix Alcan (1894).

Renard 1894 Renard, Georges. *Critique de combat.* Paris, Dentu (1894).

Renard 1990 Renard, Georges. *Les Princes de la jeune critique: Jules Lemaître, Ferdinand Brunetière, Anatole France, Louis Ganderax, Paul Bourget*. Paris, Librairie de la " Nouvelle Revue " (1890).

Rod 1892 Rod, Edouard. *Les Idées morales du temps présent.* Paris, Perrin (1892).

Rosier 1999 Rosier, Laurence. *Le Discours rapporté. Histoire, théories, pratiques*. Paris-Bruxelles, De Boeck-Duculot (1999).

Saint-Victor 1882 Saint-Victor, Paul (de). *Hommes et Dieux. Études d'histoire et de littérature.* Paris, Calmann-Lévy (1882).

Sainte-Beuve 1864 Sainte-Beuve, Charles Augustin. *Portraits littéraires, 3*. Paris, Garnier frères (1864).

Sainte-Beuve 1867 Sainte-Beuve, Charles-Augustin. *Nouveaux Lundis, 8*. Paris, Michel Lévy frères (1867).

Taine 1875 Taine, Hippolyte. *Les Origines de la France contemporaine. L'ancien régime.* Paris (1875).

Todorov et Bakhtine 1978 Todorov, Tzvetan, Bakhtine, Mikhaïl. *Le Principe dialogique*. Paris, Seuil (1978).

Tuomarla 2000 Tuomarla, Ulla. *La Citation mode d'emploi. Sur le fonctionnement discursif du discours rapporté direct*. Thèse de doctorat, Academia Scientifiarum Fennica. Ser. Humaniora. tom. 308. Saarijärvi, Finland (2000).

Zola 1881 Zola, Émile. *Le Naturalisme au théâtre: les théories et les exemples*. Paris, Charpentier (1881).

This work is licensed under a Creative Commons Attribution-NoDerivatives 4.0 International License.