

"Le peu que je sache, je veux le faire connaître afin qu'un autre, meilleur que je suis, découvre la vérité, et que l'œuvre qu'il poursuit sanctionne mon erreur. Je m'en réjouirai pour avoir été malgré tout, cause que cette vérité se fasse jour."

Albrecht Dürer

7. 1. CONCLUSION METHODOLOGIQUE

Nous pensons avoir montré l'intérêt d'appliquer au texte, et à différents paliers, une méthode d'analyse s'appuyant sur une théorie sémiotique du texte et du corpus utilisant un outil statistique probabiliste, donc différentiel. Prendre en compte, dès le départ, la complexité du texte, et la grande variété des genres n'est pas une perte de temps ni d'argent car aux différents paliers les probabilités d'ambiguïté diminuent de manière importante. Ces études ont été menées avec des moyens limités, et dans un corpus peu structuré : dans le domaine des textes électroniques, de nombreuses améliorations peuvent être apportées en utilisant les compétences des différentes communautés intéressées.

7. 1. 1. Enrichissement des corpus

L'importance des normes de discours, de genre et sous-genre que nous avons mises en évidence incite à :

- d'une part, accorder beaucoup de soin à documenter les textes par des en-têtes qui permettent à l'utilisateur de sélectionner des corpus de référence et corpus de travail adaptés à sa recherche (certaines données sont ininterprétables et surtout de nombreux faits sont lissés et passent inaperçus si trop d'hétérogénéités s'additionnent : c'est un peu comme la couleur fade que l'on obtient si on mélange plusieurs couleurs en peinture) .

- d'autre part, à coder les parties de textes, selon les recommandations de la TEI, mais en sachant que chaque corpus ou presque posera des problèmes particuliers : les nouvelles pratiques vont faire avancer les recherches théoriques sur la typologie des genres comme sur la structuration interne des textes.

Pour les mêmes raisons, l'utilisateur d'une station de travail digne de ce nom doit pouvoir choisir son corpus de référence et son corpus de travail (ce qui n'empêche pas de fournir des tables de références toutes prêtes en combinant des critères de période de temps et genres textuels, ex. le roman de la première moitié du XIX^es.).

On doit fournir à l'utilisateur des outils pour effectuer des codage plus fins, par exemple parties de dialogue vs parties descriptives dans le roman, qui ne peuvent pas être repérés dès l'origine sur de vastes corpus (le repérage automatique du dialogue représenté n'est possible qu'en partie).

Des entreprises nombreuses et ciblées de détermination des unités fonctionnelles -au plan sémantique- dans des corpus homogènes arriveront à terme à constituer des dictionnaires de référence pourvus d'un étiquetage "riche".

7. 1. 2. Les documents du SAAS

Ils seront revus au fur et à mesure que des corpus enrichis seront utilisés :

- la présentation des formes regroupées sous le lemme, même dans le cas de traitement sans lemmatisation, doit servir à repérer très rapidement toute particularité dans la distribution (comme le féminin dominant des participes avec le trait /peur/)
- possibilité d'avoir les résultats pondérés sur les parties du discours, et sur les fonctions syntaxiques.

L'hétérogénéité des scores statistiques, qui a fait couler beaucoup d'encre, provient probablement en grande partie de l'hétérogénéité des corpus (puisque des recherches récentes ont montré la meilleure efficacité du passage et de l'étiquetage morpho-syntaxique si on opère dans un corpus homogène en genre : on peut prévoir une réévaluation des seuils avec corpus enrichis, ainsi que de tester aussi d'autres méthodes statistiques, en concevant des documents d'exploitation des résultats.

On prévoit d'effectuer un alignement des scores statistiques obtenus dans différents types d'opérations de contraste :

exemple : aligner les scores du contraste du *Père Goriot* sur le roman du XIX^es, ceux du contraste *Père Goriot* sur le roman des XIX-XX^es, et ceux des contrastes chapitre par chapitre, pour observer l'effet de la surdétermination dans des zones de localité diverses.

Mais même des outils très perfectionnés ne dispenseront pas de la fréquentation et de la proximité des textes : "Notre souci est de constituer des outils polyvalents pour la recherche. Si des recherches autres que lexicales ont été entreprises (morphologiques, syntaxiques, stylistiques, par exemple), ce sont cependant les recherches textuelles et lexicales qu'il importe de mettre en avant, tant leur nécessité paraît évidente. On entre ainsi dans une perspective que l'on pourrait appeler, avec Kierkegaard, celle d'un abandon des certitudes ; mais il s'agit aussi de l'instauration d'une nouvelle dynamique de la recherche. Qu'on y prenne cependant garde : rien ne va de soi et l'informatique ne servira vraiment le chercheur en sciences humaines -philosophe, théologien, historien ou linguiste- que s'il accepte, avec humilité, de mener le combat d'une intégration incessante de l'érudition et de l'informatique, que s'il accepte de vivre avec les textes, de les mettre en lui autant qu'il les consulte et les interroge. Une consultation extérieure des données, en dehors de toute convivialité avec les textes, risque d'aboutir aux pires méprises. il faut "vivre avec" pour comprendre. [...] C'est pour y aboutir que l'informatique nous importe ! Il faut imaginer sans cesse du neuf pour aborder l'inconnu et découvrir ce que souvent l'on croyait connu"¹.

¹ Tombeur P., 1992, p. 247.

7. 2. VERS DE NOUVELLES RECHERCHES

7. 2. 1. Recherches sur *La Comédie Humaine*

A l'aide des programmes du SAAS, et d'autres outils comme Hyperbase, nous nous proposons d'étudier "autrement" ce que Auerbach appelle le "réalisme d'atmosphère" de Balzac : "Dans toute son œuvre (...) Balzac a ressenti les milieux -et ceux-ci sont les plus divers- comme des entités organiques, voire démoniques, et a cherché à communiquer ce sentiment à son lecteur. Il ne s'est pas borné, comme Stendhal, à situer les individus dont il racontait le destin avec sérieux dans leur cadre historique et social exactement déterminé, il a en outre éprouvé cette relation comme nécessaire : tout milieu devient pour lui une atmosphère physique et morale qui imprègne le paysage, l'habitat, le mobilier, les objets, les vêtements, le corps, le caractère, les relations les opinions, l'activité et le destin des individus, et en même temps la situation historique générale apparaît comme l'atmosphère globale qui enveloppe tous ces milieux particuliers²".

- **Revisiter la question de la "description" chez Balzac**

L'importance accordée aux nombreux acteurs "humains" de *La Comédie humaine*³ et une conception un peu superficielle du réalisme balzacien ont contribué à sous-estimer la valeur des descriptions de lieux et d'objets dans le projet esthétique de l'œuvre. Nous faisons l'hypothèse que les lieux et les objets sont décrits pour amorcer et renforcer la perception des principales isotopies du texte.

La description n'est pas une fin en soi, elle est une configuration et il faut chercher sa fonction dans l'économie du texte : une description peut présenter des traits qui renvoient à un autre type de texte, en s'inscrivant dans un projet esthétique inverse. Par exemple, dans une nouvelle libertine du XVIII^es., Reichler C. 1981⁴ montre que ce qui ressemble aux étapes d'une Carte du Tendre, motif du roman amoureux du siècle précédent, débouche très rapidement dans les "Terres inconnues", alors que le roman précieux s'efforce de n'y pas conduire ses héros : la raison en est qu'alors "la psychologie précieuse s'y dévoie en érotologie".

i) évolution, vers la dégradation, de la molécule 'Goriot' (son lieu de vie, avec/sans chauffage, son vêtement, ses habitudes de consommation (tabac, poudre et perruque, nourriture, vaisselle, etc.)

ii) la description de la pension Vauquer et l'adéquation de l'acteur 'Mme Vauquer' avec son milieu (la description "écologique" du milieu est à mettre en relation avec l'admiration de Balzac pour le naturaliste Geoffroy de Saint-Hilaire, auquel le roman est dédié, et qui est cité çà et là dans *La Comédie Humaine*).

² Auerbach E., 1968, p. 468-69.

³ Peu d'œuvres ont donné lieu à la publication d'un "dictionnaire biographique" de leurs personnages.

⁴ Reichler C. 1981, p. 88.

liii) l'isotopie de l'avarice du père Grandet : sa maison, les objets du décor (ou leur absence), le vêtement des trois 'Grandet' et de Nanon, les habitudes d'alimentation frustes que Eugénie tentera de changer pour son cousin Charles, les traits communs à Eugénie et son père, etc.

L'étude des passages concernés, en termes de molécules de traits sémantiques stables et structurés (après calcul statistique) doit permettre de repérer les différentes isotopies aux niveaux méso-générique et macro-générique : ainsi l'emploi de *rafistolé* pour 'Goriot' a été rapproché du lien entre l'acteur 'déjeuner de vermeil' et la molécule 'Goriot'.

- ***Surdétermination, valeurs symboliques et niveau agonistique chez Balzac***

Un procédé récurrent dans *La Comédie Humaine* est que l'isotopie des valeurs familiales est indexée dans des "objets" marqués par une double appartenance et impliqués dans un "transfert" mettant en cause plusieurs acteurs.

i) amour et filiation :

le groupe sculpté par Steinbock dans *La Cousine Bette* est acheté, en secret avec ses rares économies de jeune fille par Hortense Hulot, d'une part pour faire connaître le travail du sculpteur débutant à son père qui est en mesure de lui procurer des commandes, et, d'autre part pour se trouver un mari et résoudre la délicate question de son mariage sans dot (puisque le baron Hulot dilapide dans ses aventures amoureuses l'argent avec lequel il devrait faire vivre sa famille et doter sa fille).

ii) amour conjugal, filiation, amour illégitime :

le dernier objet précieux vendu par le père Goriot est un déjeuner de vermeil sur lequel sont représentées deux colombes, acquis avec ses économies de jeune fille et offert par son épouse juste avant leur mariage : il le négocie auprès de l'usurier Gobseck pour payer la robe de bal de sa fille Anastasia de Restaud. En effet, celle-ci doit apparaître à son avantage au bal de la comtesse de Beauséant où elle portera les diamants de la famille de son mari qu'elle avait vendus pour honorer les dettes de son amant et que son mari a rachetés et exige qu'elle porte à ce bal pour donner fin aux rumeurs sur son compte. De plus, elle ne peut se montrer dans le monde moins bien parée que sa cadette, Delphine de Nucingen, dont c'est la première apparition dans le monde de Saint-Germain où l'aînée était déjà reçue. Cet objet et le père Goriot sont équivalents au plan agonistique : il meurt après avoir transformé et "tordu" ce vermeil en lingot pour le vendre à l'usurier.

iii) filiation, amour illégitime :

les bouquets de fleurs que Félix de Vandenesse prépare pour Mme de Mortsauf sont à double valeur : ils sont des déclarations d'amour en même temps que Félix de Vandenesse feint d'accepter son rôle d'enfant auprès de sa "mère spirituelle" (il prépare les bouquets avec les enfants de celle-ci), mais elle est aussi "épouse de l'âme".

iv) amour filial, amour déçu et amour de Dieu :

quand Eugénie Grandet apprend le proche mariage de Charles rentré de l'étranger, elle fait fondre les bijoux de celui-ci que son père avait achetés en dessous de leur valeur à l'orphelin de son frère qui partait chercher fortune après le suicide de son père, ruiné ; elle en fait faire un ostensor qu'elle offre à l'Église. D'autre part, elle voulait se retirer dans un couvent, projet dont le prêtre l'a dissuadée, parce qu'il comptait sur elle pour ses œuvres charitables. De plus, ces bijoux étaient entrés en sa possession à la suite d'un marché où son père l'a dupée, car il s'en est servi pour s'acquitter à bon compte d'une dette envers elle. On sait d'autre part qu'à la suite de cette déception, Eugénie Grandet acceptera d'épouser un de ses prétendants de longue date, le Président de Bonfons, neveu du notaire Cruchot, à la condition expresse que le mariage ne serait pas consommé.

7. 2. 2. Études sur le Roman d'apprentissage au XIX^es.

Le roman d'apprentissage, forme romanesque particulière au XIX^e s., s'inscrit dans la filiation du roman picaresque (né en Espagne au cours du XVI^es.), et du roman pédagogique d'éducation princière (le *Télémaque* de Fénelon) évoluant en roman d'éducation sociale ou philosophique⁵ : il est marqué également par les œuvres du "bildungsroman" dont une particularité, la thématique artistique (la présence de héros qui appréhendent le monde à travers une vocation artistique⁶), a été mise en évidence par Marcuse, à la suite de Dilthey.

D'autre part, cette forme romanesque constitue un témoin essentiel des évolutions de mentalités dans une période influencée par l'enthousiasme révolutionnaire du siècle précédent, et Napoléon (le "Bâtard sublime"), et qui voit se succéder des troubles sociaux importants dont les auteurs utilisent des faits et des interprétations personnelles⁷.

Une grande partie des textes constituant un corpus de ce sous-genre, tel qu'il est proposé par C. Demay et D. Pernot dans leur étude : *Le Roman d'Apprentissage en France au XIX^e s.*, (Ellipses, 1995, collection Résonances) est actuellement disponible dans Frantext, et on peut envisager d'augmenter ce corpus d'étude, sur le conseil de spécialistes du domaine.

- **Recherches sur genre et sous-genres**

- a) Contraster, à l'aide des programmes du SAAS l'ensemble du corpus "Roman d'apprentissage" ainsi défini sur l'ensemble du corpus "Roman du XIX^es." de Frantext (avec ajouts éventuels pour le rendre plus représentatif).

Contraster chaque texte du corpus "Roman d'apprentissage" :

- b) - sur le corpus "Roman d'apprentissage"

- c) - sur l'ensemble du corpus "Roman du XIX^es."

On pourra de ce fait observer si des régularités thématiques et textuelles existent, dans ce corpus par rapport à l'ensemble du corpus de la production romanesque du XIX^es. : par

⁵ Le mentor ou personnage tutélaire revêt une dimension épique chez Fénelon ; dans les romans picaresques, le style bas transpose par exemple les scènes de bataille en scènes d'auberge.

⁶ Le roman d'apprentissage français met en évidence l'influence négative de la lecture sur la formation du jeune (littérature sentimentale, "bovarysme" ou influence historique -F. Moreau veut devenir le "Walter Scott" français), alors que, pour les Allemands il n'y a pas de formation sans médiation artistique et littéraire.

⁷ Pour Lukacs, Balzac crée dans les *Illusions perdues* un nouveau type qui devient d'une importance capitale pour toute l'évolution du XIX^es., le roman de la désillusion.

exemple dans les traits constituant les "molécules sémiques" des personnages des jeunes (/ambition/, /détermination/, /influençabilité/, etc. ; liens avec des aspects physiques) et des mentors (les femmes qui les aident dans leur initiation, Vautrin, etc.). Étant donné l'importance de la thématique 'Paris vs province' dans le roman français, on pourra observer si l'initiation passe toujours par la capitale et comment elle est décrite par rapport à la formation du jeune.

- ***Recherches sur genre, sous-genres et diachronie***

Disposer d'un corpus de romans français des XVII^e et XVIII^e s. (réuni en fonction de nos hypothèses mais contrastable sur un corpus plus vaste de textes narratifs de la période) permettrait des comparaisons chiffrées en termes de "faits langagiers" pour comparer le roman d'apprentissage avec les types de romans dont il est issu : on pourrait observer et quantifier les différences thématiques (condition sociale des jeunes, des mentors, type d'aventures et d'initiation, etc., lieux) par exemple, mais aussi des évolutions de ton (noble vs burlesque) et plus généralement quelles sortes de différences au plan agonistique marquent l'évolution sur les trois siècles.

"L'œuvre ne doit pas seulement être lue de gauche à droite, mais de droite à gauche et dans tous les sens, et même par bonds ou par plongées, en omettant les intermédiaires ou en joignant hardiment des épisodes ou des textes isolés. En d'autres termes, une œuvre ne doit pas être lue seulement à la façon dont on lit les rubans de mots qui se succèdent au fronton de certaines enseignes lumineuses, qui ne présentent un texte que pour le remplacer aussitôt par le texte suivant. Proust nous a appris à considérer la lecture comme analogue à l'audition d'une sonate ou d'une symphonie, où à l'action continue d'une certaine textualité s'ajoute et s'oppose l'action intermittente et répétée d'une textualité toute différente, dont les combinaisons seules nous donnent la clef de l'œuvre"⁸.

⁸ G. Poulet, "Point de vue critique. Lecture et interprétation du texte littéraire", dans *Qu'est-ce qu'un texte ?* publié sous la direction d'E. Barbotin, Paris, 1975, p. 75, cité dans Tombeur P. 1992, p. 234.