

ISSN 0268-1145

LITERARY & LINGUISTIC COMPUTING

Journal of the Association for Literary and
Linguistic Computing

Volume 10 · Number 1 · 1995

Editor: Gordon Dixon

Oxford University Press

Mondes lexicaux et univers sémantiques: le logiciel ALCESTE au service de l'étude de l'imaginaire simonien, à partir du traitement de *La Route des Flandres*

PASCAL MOUGIN

Université de la Sorbonne Nouvelle, Paris, France

Abstract

This paper shows how statistic analysis by ALCESTE computing software deals with lexical distribution and points out lists of words frequently co-occurrent, and also draws attention to groups of similar lexically based sentences. Results provided in Simon's *Flanders' Road* show four different lists, widely independent of the fictional unities of the novel. One of them, more precisely studied here, makes a symbolic or figurative system appear (concern the principal male character's self-sensation and desire of woman), which could be seen as a major structural feature of Simon's text.

1. Introduction

Le but de cet exposé est de rendre compte, par un exemple, de l'intérêt qu'un logiciel d'analyse de données textuelles peut présenter pour la recherche littéraire. La question vaut d'être posée dans la mesure où le logiciel dont il sera question, ALCESTE,¹ n'a que peu été appliqué au texte littéraire jusqu'à présent.²

Comme le rappelle son auteur, l'objectif d'ALCESTE n'est pas 'de chercher le sens d'un texte mais de déterminer comment sont organisés les éléments qui le constituent'.³ A partir de calculs relatifs à la ventilation des mots dans un corpus et tenant compte des contextes de chaque occurrence, le logiciel met en évidence l'existence, au sein de ce corpus, de 'mondes lexicaux', i.e. de familles de mots fréquemment co-occurents dans les énoncés. Tout le problème est de savoir en quoi cette structure occurrenceielle, qu'une procédure statistique automatisée aura donc repérée, renvoie à une organisation sémantique, ou tout au moins facilitera la recherche d'une telle organisation par le critique.

Après quelques remarques sur la manière dont fonctionne ALCESTE et la nature des informations qu'il soumet à l'utilisateur, nous donnerons un aperçu et un commentaire des résultats procurés sur *La Route des Flandres*⁴ de Claude Simon.

2. Particularités d'ALCESTE

La plupart des logiciels de lexicométrie peuvent prendre en compte, lors du traitement d'un texte, une parti-

tion de celui-ci en unités (parties et chapitres d'un roman, actes, scènes et rôles d'une pièce de théâtre, etc., ou tout autre découpage décidé par l'opérateur). Les résultats obtenus permettent de comparer ces différentes unités, confirment ou infirment des oppositions thématiques et stylistiques pressenties, et peuvent même en suggérer d'autres. En revanche, le programme ne saurait modifier la partition, déplacer des lignes de démarcation pour mettre en évidence une opposition plus prégnante entre deux passages, ni détecter l'éventuelle hétérogénéité d'une unité et en proposer une subdivision, ni rapprocher des séquences textuelles appartenant à des unités disjointes pour en souligner la similitude. En résumé, il ne pourra pas suggérer l'existence d'une autre structure que celle antérieurement balisée.

ALCESTE, quant à lui, considère le texte comme un ensemble d'énoncés—i.e. d'unités minimales d'énonciation telles que phrases ou propositions⁵—et compare ces énoncés les uns aux autres, indépendamment de leur appartenance à des unités textuelles plus vastes et même *sans tenir compte de leur ordre de succession*. Pour faciliter cette comparaison, chaque énoncé est envisagé non comme une suite de formes,⁶ ce qui compliquerait et fausserait les calculs, mais comme un ensemble de lexèmes, grâce à un procédé de lemmatisation simplifiée.⁷ *Le logiciel regroupe ainsi en quelques familles ou 'classes' (généralement 3 ou 4) les énoncés du texte en fonction de similitudes lexicales*. Ces énoncés rassemblés pourront provenir de différents endroits du texte, mais ils présenteront un certain 'air de famille' du fait de leur vocabulaire constitutif. On peut décrire la procédure autrement et dire qu'ALCESTE *détermine dans l'ensemble du vocabulaire du texte des sous-ensembles de lexèmes fréquemment associés dans la réalisation discursive, autrement dit des sous-ensembles de mots qui s'attirent*.

Les figures suivantes permettront d'y voir plus clair.⁸ Soit L l'ensemble des lexèmes du corpus: $L = \{l_1, l_2, l_3, l_4, l_5, l_6, l_7, \dots, l_n\}$. On peut figurer ces lexèmes par autant de points (●) uniformément répartis sur un plan. Chaque énoncé du corpus, considéré comme un ensemble de lexèmes, peut se représenter, indépendamment de l'ordre de ceux-ci, sous la forme d'un 'fil' qui solidarise les points figurant ses lexèmes constitutifs. La Figure 1 montre, ainsi schématisés, quatre énoncés (E_1, E_2, E_3, E_4) et la partie de L présentant, parmi d'autres, les lexèmes concernés.

Imaginons à présent le schéma qui représenterait de

Correspondence: Pascal Mougin, 57 rue Falguière, 75015 Paris, France.

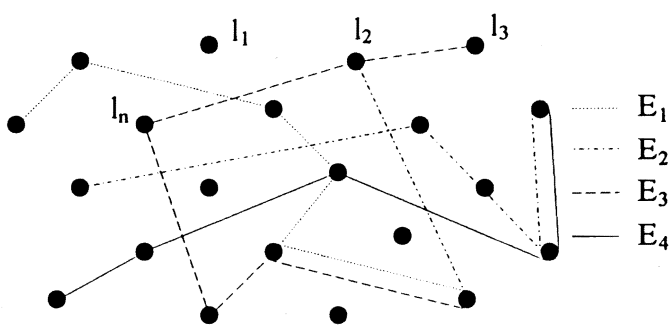


Fig. 1

manière équivalente tous les énoncés d'un corpus donné: il ne serait rien d'autre, on s'en doute, qu'un enchevêtrement illisible. On pourrait toutefois tenter de disposer les lexèmes de manière à démêler les fils entrelacés, et isoler différents foyers nodaux au sein de l'enchevêtrement. Il conviendrait pour cela de réduire au préalable l'ensemble des lexèmes: d'une part en éliminant les mots-outils (qui constituent autant de points de passages obligés de presque tous les fils et rendent l'entrelacs inextricable), et d'autre part les hapax et les mots peu fréquents (autant de points peu empruntés et qui alourdissent l'entrelacs sans intervenir dans sa structure).⁹ Ensuite, on s'arrangerait pour placer dans la même zone de la figure les lexèmes fréquemment cooccurents (i.e. fréquemment reliés par des fils). Ainsi aux cas-limites: deux lexèmes systématiquement associés seraient figurés par deux points superposés, et deux lexèmes qu'on ne rencontrerait jamais ensemble dans un même énoncé seraient situés sur deux côtés opposés. La recherche, certes fastidieuse, d'une telle disposition passerait naturellement par des compromis du fait du nombre des lexèmes et de leurs interrelations, mais elle aurait toute chance d'aboutir à une représentation du type de la Figure 2 (par souci de clarté, nous n'y représentons que les fils).

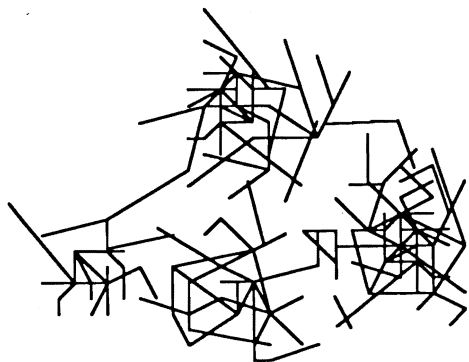


Fig. 2

Les zones d'entrelacement serré—appelons-les pôles d'enchevêtrement—correspondent à autant de classes d'énoncés caractérisés par un matériel lexématique spécifique. On distingue ici quatre classes, d'inégale importance. Cette figure appelle quelques remarques:

—certains énoncés, empruntant leurs lexèmes à plusieurs classes (ils sont figurés par des fils reliant plusieurs pôles d'enchevêtrement), n'appartiennent de ce fait à aucune d'entre elles en particulier: ce sont les laissés pour compte du classement (le

meilleur classement—la meilleure disposition—est justement celui qui présente la plus faible part résiduelle);

- à l'intérieur d'une même classe peuvent se rencontrer deux énoncés n'ayant aucun lexème commun, mais leurs lexèmes respectifs sont fréquemment cooccurents dans d'autres énoncés de la même classe;
- la densité d'enchevêtrement mesure la spécificité et la cohésion d'une classe, autrement dit le degré d'attraction des lexèmes constitutifs des énoncés de cette classe;
- les classes ne sont donc pas des sous-ensembles cloisonnés, mais plutôt des foyers de spécificité par rapport auxquels on pourra situer chaque lexème et chaque énoncé du corpus, selon une échelle continue de distance, à l'aide d'un coefficient d'association à la classe (χ^2).¹⁰

A l'issue du traitement d'un corpus, le logiciel fournit à l'utilisateur, pour chaque classe constituée, la liste de ses énoncés les plus représentatifs et la liste de son vocabulaire spécifique. Ainsi se trouve manifesté un monde lexical.

Nous avons vu que les énoncés rassemblés peuvent provenir de différents endroits du texte, et qu'une classe aura peu de chance de correspondre avec tel ou tel chapitre ou partie de celui-ci. Le degré de coïncidence dépend évidemment de l'organisation du corpus.¹¹ Les résultats procurés par exemple sur *Voyage au bout de la nuit* de Céline montrent une forte correspondance entre les classes établies et des séries de chapitres: les chapitres respectivement consacrés au front, au séjour en Afrique, au voyage à Toulouse, fortement distincts, ont largement pesé dans l'établissement des classes. Il y a un monde lexical de la guerre, un monde lexical africain, et un monde lexical toulousain.¹² Mais il est clair que moins il y aura de coïncidence entre les familles d'énoncés regroupés par ALCESTE et les unités textuelles préexistantes, plus les résultats procurés par le logiciel seront intéressants, car plus spécifiques. Du reste, le travail de rapprochement de passages non nécessairement contigus en fonction d'une ressemblance lexicale ne correspond-il pas à l'activité du lecteur attentif, et *a fortiori* à celle du critique, au moins au point de départ de son étude?

3. Le traitement de *La Route des Flandres*

Aux mondes lexicaux de *Voyage au bout de la nuit*, évoqués rapidement plus haut, il convient à présent d'opposer les résultats obtenus sur *La Route des Flandres*. L'intérêt de ce roman, dans notre perspective, consiste en ceci qu'il ne présente aucune de ces unités textuelles (chapitre ou partie) organisées autour d'un moment, d'un lieu, ou d'un personnage de la fiction. Certes le lecteur apprend progressivement à y distinguer un certain nombre de noyaux fictionnels (épisodes, séquences narratives), dont voici pour mémoire les principaux:

—dans la Meuse, entre l'automne 1939 et le printemps 1940, l'errance d'un groupe de cavaliers, parmi

lesquels le capitaine de Reixach qui les conduit (et qui sera abattu par un soldat embusqué), Iglésia (son jockey dans le civil et amant d'un jour de sa femme), Blum, et Georges (point focal de tout le récit);

- la période de captivité de Georges et de Blum, prisonniers dans un camp allemand d'où Georges parviendra à s'évader ;
- la rencontre de Georges, après la guerre, avec Corinne, la femme de Reixach, vers laquelle l'ont attiré les récits d'Iglésia;
- la course hippique, antérieure à la guerre, où de Reixach, sous les yeux de Corinne, monte une pouliche à la place d'Iglésia.
- la mort, en partie mystérieuse, du Conventionnel de Reixach, ancêtre commun à Georges et au capitaine de Reixach.

Mais, comme on sait, l'ordre du récit n'épouse pas la chronologie de la fiction. Les fragments d'épisodes s'enchevêtrent, reviennent et se superposent comme dans une remémoration en acte, associés autour de pivots de tous ordres (thématiques, sémantiques, lexicaux même) produits ou découverts au fil de l'écriture.

S'agissant donc d'un texte qui n'est pas articulé selon le déroulement de son contenu fictionnel, mais qui, *grosso modo*, 'parle de tout partout', il se peut bien sûr que le logiciel reconstitue des unités distinctes de la fiction sur la base éventuelle de leur vocabulaire spécifique. De fait, on verra que les classes établies par ALCESTE renvoient pour une part à ce type d'unités (cas de la course de chevaux par exemple). Les co-occurrences des mêmes mots dans plusieurs énoncés, qui assurent la cohérence du monde lexical, traduisent alors une relation de proximité au niveau des choses: on pourra parler d'isotopies référentielles (par exemple, on ne s'étonnera pas de la cooccurrence fréquente des mots 'cheval', 'galop' et 'cravache'). Mais on peut espérer aussi que le logiciel mette en évidence des principes d'organisation plus profonds, et que l'attrance statistiquement remarquable entre les éléments d'un monde lexical renvoie à des relations sémantiques plus riches, construites dans et par le langage. C'est d'ailleurs sur ce type de cohérence du texte simonien que Jean Ricardou a d'emblée attiré l'attention des critiques, en remarquant que derrière la dislocation de la fiction,

[...] les mots deviennent des centres *d'irradiation sémantique* qui, sous la croûte de leur sens immédiat, tendent à recomposer entre eux, de proche en proche, les relais d'un langage sous-jacent, libre et mobile, où jouent toutes manières de sens seconds.¹³

Mais là où Ricardou se contente de relever quelques exemples flagrants de ce 'langage sous-jacent' (parce qu'il entend seulement, à l'époque où il écrit, en démontrer l'existence), nous verrons qu'ALCESTE permet d'en esquisser quelques grandes articulations.

Les remarques qui vont suivre résument les analyses des quatre classes constituées à l'issue du traitement de *La Route des Flandres*. Chaque analyse a tenté de caractériser les principales isotopies constitutives de la classe, d'expliquer, à l'aide des contextes, les raisons

de leurs cooccurrences, pour finalement dégager l'unité de l'ensemble. La dernière classe, la plus intéressante dans notre perspective, fait l'objet de remarques plus détaillées.

3.1 Classe A: Le dialogue et l'action

3.1.1 *Vocabulaire*. On trouvera pour chaque classe une liste de ses 'formes réduites' spécifiques, présentées telles que le logiciel les manipule.¹⁴ Chaque forme est suivie de son coefficient de rattachement à la classe (chi²), qu'il faut lire comme un indice de spécificité (plus le nombre est élevé, plus la forme est spécifique des énoncés de la classe). J'ai rassemblé ces formes selon leurs affinités sémantiques (champs lexicaux), ou autre (nature grammaticale, fonction sémiotique, etc.). Il ne s'agit dans chaque cas que d'une proposition de classement, visant à faciliter la lecture de la liste.¹⁵

♦ la parole: dire.¹⁶ (105), voix (45), cri+ (36), ce-que (30) [correspond essentiellement au tour interrogatif 'qu'est-ce que'], parl+ (21), demand+ (17), repondre. (16) oui (16), repet+ (16), ecout+ (16), disput+ (15), jur+ (13), hurl+ (10), repondant (10), entendre. (10), racont+ (9), entendant (7). ♦ **1^{ère} et 2^{ème} personnes**: je (89), vous (75), tu (73), te (70), moi (37), es (26), me (24), avez (24), suis (22), tien< (19), as (19), aurais (14), toi (14), nous (14), avais (12), mon (9). ♦ **vocabulaire oral, familier ou argotique**: dieu+ (39) [correspond à l'interjection 'bon Dieu'], type+ (36), con+ (31), oh (29), pauvre+ (29), parier. (19), gueul+ (19), fout+ (19), salaud+ (10), idiot+ (10). ♦ **verbes d'action**: faire. (21), peser. (14), perdre. (14), prendre. (13), envoy+ (11), ramass+ (8), tirer. (8), detourn+ (8). ♦ **violence, conflictualité**: fusil+ (31), cogn+ (11), coup< (9), mitrail+ (8). ♦ **déplacements**: aller. (76), rentr+ (30), trottin+ (14), descendre. (14), sortir. (11), courir. (10), bouge+ (10), monter. (10), part+ (9), entrer. (8), venir. (8). ♦ **aspect de l'action**: arret+ (23), continu+ (10). ♦ **marqueurs temporels**: toujours (18), dimanche (12), quand (8). ♦ **intellection**: comprendre. (26), croire. (18), malin+ (17), decid+ (12), penser. (10). ♦ **volonté, obligation, possibilité**: vouloir. (36), essay+ (18), peine< (11), besoin+ (9), laiss+ (8), envi+ (8). ♦ **affection, sentiment**: peur (17), reproba+ (17), plaindre. (15), furieux+ (14), fol. (10). ♦ **personnes, personnages**: wack (53), blum (30), iglesia (26), capitaine (22), adjoint (17), georges (16), soeur+ (12), soldat+ (12). ♦ **négation**: pas (67), sans (22), ne (22), ni (7). ♦ **la respiration**: poumons (19), etouff+ (17). ♦ **voir**: regard+ (13). ♦ **logique**: quand-meme (15), pourquoi (11), pourtant (8). ♦ **divers**: boit+ (18), bon+ (18), maison+ (15), hauss+ (13), parapluie+ (11), cafe+ (10), rat+ (10), clef (8), chambr+ (8), camp (7), vill+ (7).

Cette première classe rassemble des énoncés provenant des dialogues ou de leur périphérie immédiate (d'où d'ailleurs une très forte unité tonale imposée par le niveau de langue familier ou argotique). En parcourant la liste de ces énoncés, on s'aperçoit que le dialogue coïncide dans le roman avec les moments d'intensité dramatique, qu'il s'agisse d'une relation conflictuelle entre les personnages (altercation entre les soldats et les paysans, entre Corinne, de Reixach et Iglésia, et entre Corinne et Georges), ou d'une situation dangereuse (menace de mort, attaque de l'ennemi embusqué). Ainsi s'explique que le vocabulaire soit celui de l'action (déplacement, effort, lutte), et de ses aspects (fin/moyens, obstacles/adjuvants, échec/succès). Le vocabulaire est aussi celui des dispositions modales

des acteurs (savoir, capacité, volonté, obligation), car les énoncés renvoient à la manière dont les personnages (principalement Georges) s'interrogent sur ce qui les entoure, délibèrent, interprètent, comme dans cet exemple qui confine à l'autodérision:

[...] nous avons essayé j'ai essayé continué jusqu'au bout faisant semblant de croire que cela pouvait réussir [...] comme si j'espérais réussir à me faire croire que je croyais que c'était possible alors que je savais le contraire (p. 71).

Les relations fonctionnelles entre dialogue et action dans l'économie du récit¹⁷ font que ce premier monde lexical dessine la charpente événementielle du roman: on y trouve la mention des personnages (en tant qu'énonciateurs des dialogues), des principaux contextes fictionnels et même des principaux épisodes. Car c'est à l'occasion du dialogue que l'action se désigne, se récapitule, qu'un personnage apprend d'un informateur une donnée nouvelle ('Le front? Pauvre con! Le front? . . . Y a plus de front, pauvre con, y a plus rien!', p. 102) ou qu'une séquence narrative est embrayée (pour se poursuivre ensuite de manière différente). D'où plusieurs énoncés anaphoriques ou cataphoriques: 'Ecoute: à un moment il nous a payé à boire' (p. 20), suivi de la description de la scène; 'On est perdus on est tombés ce matin dans une embuscade le capitaine vient d'être tué [...]' (p. 84), phrase qui condense l'épisode central du roman.

Ce premier monde lexical est celui de la rigidité classématique, pour reprendre la notion introduite par Greimas:¹⁸ les énoncés renvoient toujours à un personnage ou à un épisode clairement identifiable, grâce à eux le récit délimite et désigne les éléments de la fiction. Il est intéressant de remarquer que ce sont ces énoncés, le plus souvent portés au compte de personnages et empreints d'oralité, qui assument finalement les fonctions classiques de régulation du sens et qui procurent au lecteur les repères fictionnels minimaux.

3.2 Classe B: Le discours: interprétation des signes, recherche de l'identité

3.2.1 Vocabulaire

♦ **vérité/réalité**: admettre. (21), persuader+ (12), réalité (11). ♦ **marqueurs de l'incertitude**: chose+ (27), comment (22), c'est-à-dire (21), suppos+ (19), du-moins (16), quelque-sort (16), certain< (16), au-moins (15), peut-être (10), vague+ (9), sans-doute (8). ♦ **savoir**: savoir (25), apprendre. (19), connaître. (10). ♦ **logique, causalité**: donc (18), au-contre (8), mais (8). ♦ **signes et représentations**: nom+ (24), écrire. (23), représenter+ (22), titre+ (17), lire. (16), livrer+ (16), mot+ (7). ♦ **évaluation**: valoir. (14), tel (12), meilleur+ (8). ♦ **négation**: ne (8), non (7). ♦ **quantification**: cent< (27), vingt (11), nombre+ (8), trois (8). ♦ **démonstratif**: ces (7). ♦ **sentiments, affections**: amour (18), souvenir. (15), instinct+ (15), ruse (15), vertu< (15), plaisir (11), content+ (11), préférer+ (8), esprit+ (8). ♦ **vocabulaire théorique**: loi< (30), noble+ (27), humain+ (21), mourir. (17), posséder+ (17), exister+ (14), détruire. (12), constituer+ (12), absolu+ (9), disposer+ (8), naturelle+ (8), nature+ (8), interdit+ (8), état+ (7). ♦ **l'action**: moyen+ (22). ♦ **mouvements**: répandre. (11), amener. (8). ♦ **histoire militaire**: pistolet+ (22), bataille+ (11), command+ (11), desas-

tre+ (11). ♦ **personnes, personnages**: domestique+ (16), blum (11), le lignage: mère (27), famille+ (24), père+ (11), sabine (7). ♦ **le temps**: suite (23), histoire+ (17), âge+ (17), tradition+ (15), ans (15), année+ (14), jeune+ (14), conserv+ (11), époque (8). ♦ **l'activité commerciale**: échange+ (18), fortune+ (15), profit+ (8). ♦ **divers**: simple+ (17), port+ (16), particulier+ (15), exemple (14), réserve+ (12), grand+ (11), charge+ (11), cuire. (8), cour+ (8), lit (8), rape+ (8), toile+ (8), trouver< (7), mani+ (7).

On reconnaît aisément une spécificité fictionnelle de cette classe: un grand nombre d'énoncés concernent la figure de l'ancêtre de Reixach, les supputations autour de sa mort, la transmission de son souvenir jusqu'à Georges par sa mère, la survivance de ses traditions, de ses vertus et de son nom dans la figure du capitaine, et la possible répétition, d'une guerre à l'autre et à 150 ans de distance, du même suicide: le capitaine 'a trouvé un moyen de trouver ce qu'on appelle une mort glorieuse. Dans la tradition de sa famille' (p. 77). Dans la recherche menée par Georges des analogies entre présent et passé, la question de l'Histoire et de ses 'lois' (le mot est donné pour le plus caractéristique de la classe) débouche sur le problème de la connaissance par les signes (images, documents, récits, dont les parents de Georges sont les emblèmes disqualifiés), et d'une manière générale de la validité de tout discours et de toute représentation:

[...] cette loi qui veut que l'Histoire . . . [...] (ou si tu préfères: la sottise, le courage, l'orgueil, la souffrance) ne laisse derrière elle qu'un résidu abusivement confisqué, désinfecté et enfin comestible, à l'usage des manuels scolaires agréés et des familles à pedigree... Mais en réalité que sais-tu? (p. 173)

Le savoir, dans la constitution de l'individualité, entre en conflit avec l'instinct (l'atavisme?) et l'expérience accumulée, et d'une manière générale Georges constate que tout discours, et le discours par excellence qu'est l'Histoire, a peu à voir avec la réalité, en l'occurrence la guerre vécue, éprouvée par l'individu. D'où le problème longuement évoqué de la contribution des individus à l'Histoire, à la fois irréfutable et dérisoire.

Une modalisation d'incertitude constante caractérise les énoncés de cette classe. Le discours, justement parce qu'il énonce ses propres limites, se cherche et se défie de lui-même, comme dans cet exemple:

[...] c'est-à-dire coupable (lui, le noble de naissance et dont la guerre—c'est-à-dire, en une certaine façon, l'oubli de soi, c'est-à-dire une certaine désinvolture, ou futilité, c'est-à-dire, en une certaine façon, le vide intérieur—était la spécialité) d'avoir voulu mélanger—ou concilier—courage et pensée, méconnu cet irréductible antagonisme qui oppose toute réflexion à toute action (p. 178).

Ici encore, le fonctionnement classématique reste primordial: les énoncés portent de manière univoque sur tel ou tel personnage, épisode ou notion. Mais il est en partie battu en brèche par une remise en cause de ce qui fait les contours et l'unité des uns et des autres.

3.3 Classe C: perception de l'espace et du mouvement

3.3.1 Vocabulaire

♦ **voir**: disparaître< (25), vis (24), apparaître. (13), sembler. (11), voir. (10). ♦ **luminosité**: ombre+ (27), soleil+ (19), clair+ (18), fonc+ (10). ♦ **couleurs**: vert+ (25), jaun+ (15), bleu+ (12), marron (10), roug+ (9). ♦ **iconicité**: dessin+ (17), silhouett+ (16). ♦ **taille**: immense+ (9), petit+ (8), gross+ (8). ♦ **forme**: long+ (38), form+ (34), angle< (32), allonge+ (20), tache+ (23), inclin+ (18), parallèle+ (17), aplati+ (15), pente+ (15), large+ (14), etir+ (13), courb+ (12), cercle+ (11), longue+ (10), raid+ (7), pench+ (7). ♦ **descripteurs de l'espace**: cote< (46), surface (27), vertical+ (17), tapi+ (9), bord+ (8), horizontal+ (8), ligne+ (7). ♦ **position, disposition**: droit+ (63), sur (61), gauche (55), haut+ (27), derrière (25), metr+ (18), loin (18), position (17), vers (12), devant (11), dessus (11), arrière (10), au-bout (8), éloign+ (8), en-bas (7). ♦ **mouvement**: immobil+ (45), retomb+ (42), tourn+ (32), franchir. (28), balanc+ (25), travers+ (22), immobilis+ (20), oscill+ (20), dépass+ (19), bascul+ (18), soulev+ (18), detach+ (18), dirige+ (16), mouv+ (13), march+ (13), recul+ (12), passer. (11), deplac+ (11), remont+ (11), départ (9), rapproch+ (9), suivre. (9), course+ (8), écart+ (7). ♦ **vitesse**: lent+ (30), rapid+ (17), vite+ (8). ♦ **marqueurs temporels divers**: puis (48), cess+ (8), aussitot (7). ♦ **simultanéité, progressivité**: tandis-que (24), fur (16), peu-a-peu (8). ♦ **nombre, ordre, succession**: de-nouveau (39), dernier+ (19), quatre (16), troisième (12), successive+ (10), deux (7). ♦ **voies de communication**: chemin+ (21), rue+ (11), rout< (9). ♦ **moyens de transport**: camion+ (20), voiture+ (8). ♦ **thématique équestre**: cheval+ (46), galop+ (31), cravach+ (25), encolure+ (25), cavalier+ (21), queue+ (20), rene+ (20), sabot+ (19), mousqueton (11), casaque+ (11), jockey+ (7). ♦ **nature**: haie< (39), bois+ (32), nuage+ (30), branche+ (17), colline+ (16), oiseau+ (14), arbre+ (13), pre (10), verger+ (10), feuill+ (9), ciel (7), champ+ (7), herb+ (7). ♦ **constructions**: tribune+ (22), sol+ (15), mur+ (14), toit+ (12), pave+ (10), baraque+ (8). ♦ **acteurs pluriels**: foule+ (19), peloton (16), groupe+ (11). ♦ **parties du corps**: buste+ (32), tet+ (32), main+ (27), jambe+ (23), oreille+ (17), cuisse+ (15), yeux (7). ♦ **liquidité/solidité**: bull+ (11), fondre. (10). ♦ **pesanteur**: léger+ (30). ♦ **l'ordre auditif**: tintement+ (12). ♦ **divers**: se (51), pend+ (17), fer+ (15), ouvrir. (14), engage+ (12), fourche+ (12), double+ (11), poisson+ (11), rid+ (11), identique+ (10), jusque (10), suivant (9), avant (8), pièce+ (8), corinne (7), verre+ (7).

Cette classe s'organise autour de l'ordre visuel. On reconnaît dans les énoncés qu'elle rassemble les aspects qui ont été les premiers signalés comme caractéristiques de l'écriture de Claude Simon,¹⁹ à savoir la représentation de l'espace et du mouvement en tant qu'objets de la perception—ce que Claude Simon appelle 'la surface même des choses'. On ne saurait ici aborder tous les problèmes que soulève cet aspect de son écriture, mais voici quelques indications suggérées par la lecture des énoncés de la classe.

3.3.2 *Une combinatoire picturale*. Trajectoires, formes et couleurs constituent le matériau spécifique de cette classe. Il s'agit de motifs, au sens pictural du terme, qui se trouvent diversement agencés dans différents épisodes du roman (la course de chevaux bien sûr, mais aussi le mouvement du tracteur du métayer chez le père de Georges, l'apparition d'avions de combat dans le ciel, etc.). Un exemple, parmi d'autres possibles, de passages rapprochés par ALCESTE fera mieux

comprendre cette combinatoire. Le premier passage est un extrait de l'épisode de la course hippique où de Reixach, vêtu d'une casaque rose, monte sa pouliche à la place de son jockey Iglésia:

[Les chevaux] apparurent enfin après le dernier arbre, toujours dans le même ordre, la tache, la pastille rose toujours en même position [...], les chevaux agglomérés en paquet paraissant un moment ne plus avancer [...] jusqu'à ce que soudain le premier cheval non pas franchît mais crevât la haie, c'est-à-dire que brusquement il fut là, les deux pattes de devant projetées devant lui, raides, jointes [...], le cheval [...] reposant apparemment sur le ventre comme en équilibre, une fraction de seconde immobile, aurait-on dit, jusqu'à ce qu'il basculât en avant tandis qu'un second, puis un troisième, puis plusieurs ensemble, tous figés successivement en équilibre, dans cette position de cheval à bascule, apparaissent, s'immobilisent, s'inclinent en avant, retrouvant le mouvement en même temps que le contact avec la terre [...] (pp. 160–161).

Le second montre Georges et Iglésia tentant de fuir à l'insu de l'ennemi qu'ils ont aperçu aux abords d'une maison abandonnée:

[...] tout ce qu'il [Georges] pouvait voir c'était la tache claire que faisait le chiffon rose accroché à la haie non loin du cheval, mais pas le cheval, et pas la sentinelle non plus, seulement la tache rose luisant faiblement dans la pénombre, puis même plus le chiffon parce qu'ils avaient franchi encore une autre haie, toujours à reculons, la tête toujours tournée du côté de la route, se heurtant du dos à la haie, la main tâtonnant derrière eux, levant la jambe, un moment à cheval sur la haie, le buste couché dessus, et retombant de l'autre côté (p. 230).

On lit d'un extrait à l'autre le réassemblage de motifs identiques (ce qui a valu à tous ces énoncés d'être regroupés dans la même classe): la tache rose (casaque dans un cas, chiffon abandonné dans l'autre), le franchissement d'une haie (de front et aérien vs à reculons et empêtré), avec une immobilisation au moment crucial (par rémanence rétinienne vs par difficulté). A cela s'ajoute le fait que la première scène est un spectacle (Iglésia et Corinne suivent la course à la jumelle), alors que dans la seconde les personnages en mouvement doivent au contraire échapper au regard de l'ennemi. Il y a donc ici une paradigmativité minutieuse du récit, qui désigne les deux personnages franchissant les haies (de Reixach d'un côté, Georges de l'autre) comme étant dans des situations à la fois équivalentes et opposées. En décidant de monter la pouliche pour emporter la course, de Reixach, mari trompé, entend (en vain) se requalifier aux yeux de Corinne. En s'enfuyant, Georges tente d'échapper à la mort. Les deux enjeux, pareillement essentiels pour les deux personnages, doivent se lire eux aussi comme équivalents, et on voit se profiler ici, notamment à travers la fascination exercée sur Georges par la tache rose, le rapprochement de celui-ci et de Corinne. Il

faudrait, pour en dire plus, commenter les autres occurrences des motifs rencontrés ici (le rose, la haie, le cheval) et reconstituer un système complexe de signification. L'essentiel, dans notre perspective, était de montrer la piste suggérée par ALCESTE.

Cet exemple montre que la combinatoire a tendance à valoriser le niveau sémiologique de la signification,²⁰ au détriment du fonctionnement classématique. L'attention du lecteur est moins attirée sur une scène ou un personnage que sur une forme, un mouvement, ou une couleur réapparaissant dans des contextes différents.

3.3.3 Investissements axiologiques. L'extrême minutie de la description simonienne avertit le lecteur de l'importance de ses motifs privilégiés. Chacun d'eux, au fil de ses occurrences, s'investit de significations particulières, le plus souvent non formulées, imputables à l'état d'esprit de l'observateur virtuel, et qui donnent cette impression caractéristique d'une fascination silencieuse à l'œuvre dans le regard. On sait que les connotations érotiques figurent au premier rang de ces investissements, mais les contextes extraits par ALCESTE en laissent entrevoir d'autres, dont voici deux exemples.

Citons tout d'abord la fonction vitale des ombres projetées et des formes géométriques. En l'absence d'indications plus précises, les cavaliers se fondent sur la direction des rayons du soleil et sur l'orientation des haies et des chemins pour tenter de s'y retrouver. D'où la tentative angoissée de géométrisation du réel, condition de la survie des soldats:

[...] et à ce moment il remarqua que l'ombre projetée par l'angle du mur de brique qui bordait la route recouvrait les membres postérieurs du cheval tout à l'heure en plein soleil, la portion d'ombre projetée par la partie du mur parallèle à la route ne cessant de s'élargir, pensant : 'Mais nos ombres étaient alors sur notre droite, donc le soleil a maintenant franchi l'axe de la route, donc . . .', puis cessant de penser (p. 98).

L'hippodrome, en tant qu'espace idéalement balisé, fournit un modèle de géométrisation possible. Mais, sur les routes du Nord, la question lancinante de Georges, 'Quelle heure est-il?'²¹ (savoir l'heure 'permettant de se diriger d'après la position de son ombre', p. 69), reste sans réponse, et les cavaliers, égarés, tomberont sous le feu de l'ennemi.

Relevons enfin l'ambivalence du mouvement lent et continu. Paisible et sécurisant, il fascine à la fois le scripteur, qui affectionne le motif du glissement horizontal (la progression du tracteur du métayer, le défilement des chevaux de course derrière la haie), et le personnage-narrateur: Georges, après la guerre, montera lui-même sur le tracteur (dont le mouvement se lit comme une image de l'écriture). Mais il est aussi terrifiant, parce qu'il signifie la menace toujours possible de son contraire : la cessation brusque, l'irruption soudaine, l'embuscade (au moment de l'attaque, les chevaux des soldats avancent avec le même mouvement fluide que celui des pur-sang de la course). D'où l'image oxymorique, fascinante et angoissante, de 'la

guerre pour ainsi dire étale pour ainsi dire paisible autour de nous' (p. 289). D'une manière générale, la tranquillité contient en elle, virtuellement, la violence.

3.4 Classe D: sensation et imagination du corps et de la matière

3.4.1 Vocabulaire

♦ **Pordre tactile:** impalpable (34), soyeux+ (19), craquel+ (19), fragil+ (19), précieux+ (15), touch+ (14), tendre< (12), délicat< (7). ♦ **Pordre olfactif, la respiration:** parfum+ (39), acre (29), exhal+ (24), odeur (20), puant+ (19), sentir. (13), respirant (11), cigarette+ (10). ♦ **luminosité:** transparent+ (33), lampe (29), nuit< (24), lumière (13), tenebre+ (8), allum+ (7), jour+ (7), aveugl+ (6), ombre+ (6), ensoleille+ (5), obscur+ (4), sombre+ (4). ♦ **couleurs:** ros+ (19), blanc+ (16), teint+ (15), brun+ (10), mat+ (8), gris+ (5). ♦ **liquidité:** sèche+ (19), coule+ (18), liquide (16), lait+ (13), boire. (7), eau (4), pluie (3). ♦ **thermie:** froid+ (31), tiède+ (20), brûl+ (11). ♦ **l'enveloppe, le voile, le vêtement:** vêtement+ (38), déchir+ (34), repli+ (33), chemise+ (29), envelopp+ (24), toile+ (24), épais+ (20), fine+ (15), mince+ (15), nu (15), étoffe+ (14), étroit+ (12), couch+ (10), drap+ (10), recouvr+ (8), vetu+ (8), masqu+ (7), robe+ (6), costume+ (6), nus (3). ♦ **le corps, l'organique, le biologique:** chair+ (71), seins (44), os (28), organ+ (25), affame+ (25), levre+ (23), ventr+ (17), peau (16), gorge+ (13), visage+ (13), doigt+ (9), cuir+ (8), cheveu< (7), féminin+ (7), sein (7), cou+ (6), joue< (6), bras+ (5). ♦ **pesanteur, constriction:** lourd+ (33), poids (19), écras+ (19), presser. (8). ♦ **profondeur et intériorité:** secret+ (43), contenir. (15), englout+ (11), entour+ (11), enfermer+ (11), profond+ (10), absorber+ (10), impenetrable (8), cach+ (3). ♦ **modalités phénoménologiques:** irréel+ (19), fantom+ (15), sembler, (10), flou+ (10), troubl+ (7), invisible+ (6). ♦ **nature:** fleur+ (28), terre< (23), boue (13), herb+ (11), ciel (6). ♦ **signes et écriture:** papier+ (57), encre (19), légende+ (15). ♦ **la rémanence:** durer. (14), persister+ (14), trace+ (12), sillage (11). ♦ **narcose:** dormir, (4), réveiller+ (4), endormir, (3). ♦ **temps:** aube+ (20), après-midi (19), crépuscule+ (12). ♦ **divers :** comme (47).

3.4.2 Les ordres perceptifs. La perception est ici essentiellement tactile et olfactive. Contrairement à la perception visuelle de la classe précédente, elle n'a lieu qu'à la faveur d'une relation matérielle, mécanique, entre le sujet et l'objet. Au point que la spécificité de chaque ordre s'estompe au bénéfice d'une saisie 'proprioceptive' globale plus proche de la sensation. L'axe sémantique majeur est celui de la consistance (solide vs gazeux, avec les sèmes intercalaires neutres et complexes: gélatinosité, liquidité, moiteur, vapeur, poussière, etc.). A ces valeurs correspondent des équivalents sensoriels (il s'agit en fait de descripteurs qui en permettent le compte-rendu subjectif), correspondance qu'on peut figurer de la manière suivante:

$$\frac{\text{solidité}}{\text{tactilité}} \rightarrow \frac{\text{liquidité, moiteur, etc.}}{\text{tactilité+olfactivité}} \rightarrow \frac{\text{gazéité}}{\text{olfactivité}}$$

L'ordre du visible (couleurs et luminosité) est lui-même rapporté au toucher et à l'olfaction, le plus souvent par image, catachrèse ou comparaison. En voici plusieurs exemples, particulièrement emblématiques de la classe:

[...] l'éblouissante et opaque lumière de cette journée de printemps (comme si la lumière elle-

même était sale, comme si l'air invisible contenait en suspension, comme une eau souillée troublée, cette sorte de crasse poussiéreuse et puante de la guerre) (p. 17)

[...] le chatoyant et impalpable poudroiement de poussière dorée suspendu dans le paisible et vert après-midi aux effluves de fleurs (p. 19)

[...] l'air noir et dur sur les visages comme du métal, de sorte qu'il lui semblait [...] sentir les ténèbres froides adhérer à sa chair, solidifiées (p. 29)

[...] le pénétrant et vert parfum de l'herbe coupée (p. 207)

[...] et avec en plus cette espèce de gravier, de toile émeri conséquence du manque de sommeil sous nos paupières nous ne pouvions voir que la face ombrée noire des arbres [...] (p. 273).

En cela la lumière évoquée ici se distingue de celle de la classe précédente, où elle était avant tout un instrument privilégié de la géométrisation du réel, projetant les ombres et engendrant les formes. Elle entre ici au contraire comme composante de la sensation de matérialité.

Enfin, précisons qu'à cette sensation peut correspondre une indication d'intensité thermique, selon l'opposition suivante:

solidité

sensation thermique intense : gel, feu
vs
liquidité, gazéité

sensation thermique intermédiaire : tiédeur

ainsi que des notations de la dysphorie ou, plus rarement, de l'euphorie.

Tels sont les constituants sémiques élémentaires les plus fréquents dans les énoncés de cette classe. Nous allons voir qu'ils entrent tous dans la constitution d'une opposition thématique majeure, celle du corps 'encarapaçonné' et du corps 'exhalé'.

3.4.3 *Le corps encarapaçonné*. Georges, avançant dans la nuit et sous la pluie avec les autres cavaliers, n'a pour toute conscience du monde et de lui-même qu'une perception de la surface de son corps. Cette perception n'est que la vague sensation d'une épaisseur qui dessine la limite entre une intériorité engourdie et anesthésiée et l'extérieur, agressif et opaque. Le contact seul produit la sensation et le sentiment d'existence. L'exténuation elle-même est décrite comme une dysphorie de cette surface, où s'est formée une 'mince pellicule de saleté et d'insomnie' (p. 36) comme l'effet conjugué d'une exsudation de la fatigue, refoulée depuis l'intérieur, et d'une concrétion exogène, trace du contact prolongé avec l'environnement invisible et immatériel. A l'endroit de la peau, des paupières et des vêtements 'roides et pesants de pluie' (*ibid.*), la sensation confond les extrêmes opposés: brûlure et froideur glacée, détrempage et durcissement. Cette gangue d'inconfort, du fait même de sa dureté, protège imparfaitement le corps du milieu qui l'assaille. La pétrification et la minéralisation menacent, sous la forme d'un épaissement à l'infini de la gangue, gagnant depuis

ses deux faces et solidarissant dans un figement généralisé le corps et le dehors (ce qui sera littéralisé par l'image du glacier). Le froid disparu, lors de la halte dans la grange et dans le camp de prisonniers, la menace devient celle de la putréfaction et de la désagrégation, car la chair qui s'en est remise à sa périphérie pour son maintien n'a plus de fermeté interne.

L'évolution reste à l'état de menace dans le cas de Georges, mais le texte présente quantité d'exemples d'enveloppes de toutes sortes qui échouent dans leur fonction de protection et de délimitation de l'être: carcasse du cheval, carrosserie de voiture, valise, et même la toile du tableau de l'ancêtre, décrite sur le même modèle. L'enveloppe sera rongée, infiltrée, traversée par les éléments naturels (eau, terre, boue), ou encore éventrée, réduite à l'état de loque et de lambeau, et on assistera parallèlement à la délitescence progressive du contenu.

La figure du père de Georges correspond à ce schéma. Voici, à titre d'exemple, plusieurs passages qui s'y rapportent et qui ont été rangés dans la classe par le logiciel. Dans l'extrait suivant, le personnage apparaît comme une masse organique avachie, seulement pré-munie de l'écroulement ultime—métaphoriquement—par la couche de papier qu'il déplace autour de lui:

[Georges] pensant à son père assis dans le kiosque aux vitres multicolores au fond de l'allée de chênes où il passait ses après-midi à travailler, couvrir de sa fine écriture raturée et surchargée les éternelles feuilles de papier qu'il transportait avec lui d'un endroit à l'autre dans une vieille chemise aux coins cornés, comme une sorte d'inséparable complément de lui-même, d'organe supplémentaire inventé sans doute pour remédier aux défaillances des autres (les muscles, les os accablés sous le monstrueux poids de graisse et de chairs distendues, de matière devenue impropre à satisfaire par elle-même ses propres besoins de sorte qu'elle semblait avoir inventé, sécrété comme une sorte de sous-produit de remplacement, de sixième sens artificiel, de prothèse omnipotente fonctionnant à l'encre et à la pâte de bois) (p. 31).

On retrouve l'image de l'enveloppe permanente ('chemise', 'éternelle', 'inséparable') qui, de consubstantielle à l'être dont elle est l'émanation caractéristique ('sécrété'), finit par s'en détacher ('prothèse') à force d'épaississement ('couvrir', 'surchargée', 'cornés' qui peut évoquer la corne, la peau calleuse), laissant pour finir cet être vidé de tout principe interne de tenue. Rien d'étonnant dès lors à ce que le tégument devenu autonome—d'ailleurs les papiers, à partir de ce matin-là, resteront de plus en plus souvent sur la table, menacés de dispersion par le vent du soir (p. 223)—se craquelle, s'écaille, et n'assume plus sa fonction illusoire d'enceinte. L'être lui-même, rendu par cette désquamation à sa flaccidité de tas organique, ne tardera pas à s'effondrer sur lui-même:

[...] la pesante montagne de chair toujours immobile, silencieuse, la lourde et pathétique masse d'organes distendus et usés [...], au point que malgré

sa totale immobilité, malgré sa totale absence de réaction apparente Georges perçut parfaitement [...] comme une sorte de craquement, comme le bruit imperceptible de quelque organe secret et délicat en train de se briser (p. 214).

L'intérieur s'écroule insensiblement, comme s'il pourrissait depuis ses tréfonds (si bien qu'il pourra, en beaucoup d'autres endroits du texte, sembler absorbé, 'bu', par le support où il reposait). De l'être disparu il ne subsistera le plus souvent qu'une vague trainée, une souillure, faite des lambeaux éparpillés de l'enveloppe, ou, dans le cas du père de Georges, 'un insignifiant résidu, l'amas de journaux froissés où depuis longtemps on ne distinguait plus rien' (pp. 34–35).

3.4.4 *Vers le corps exhalé.* Le corps de la femme (la femme entrevue dans la grange et, bien sûr, Corinne) est décrit de manière exactement opposée: la deminuité le découvre, le parfum et la transparence lumineuse le diffusent. Ici les enveloppes prolongent naturellement la chair: c'est l'image de la peau 'soyeuse', émanation duveteuse et diaprée de l'être, c'est le rose de la robe, métonyme du sexe. Car cette plénitude irradiante est aussi attirance, appel et ouverture vers l'intériorité, elle prévoit une circulation de part et d'autre de sa surface (d'où les images de la femme-fleur et de la femme-madrépore). Le mouvement tout en exhalaison suscite le désir, promet le réconfort.

Le désir obsessionnel de Georges pour Corinne n'a de ce fait d'autre sens qu'un salut symbolique: il s'agit d'éviter le pourrissement et la destruction qui le menacent en tant que corps 'enveloppé' (puisqu'il paraît être le pôle imaginaire de la mort dans *la Route des Flandres*), et de s'approcher au plus près du corps exhalé, par l'union à la femme. C'est cette tentative de conciliation paradigmatique qui constitue l'un des fils sous-jacents du roman. Voici, rapidement énumérées, quelques-unes des étapes de cette tentative:

1. La halte dans la grange, moment de la prise de conscience de l'opposition radicale des deux paradigmes (voir p. 36 sqq: Georges vs la femme, froid/dureté/constriction/insensibilité vs tiédeur organique/diffusion de l'être, etc.); Georges sait que la désagrégation le menace s'il retire sa gangue à la fois douloureuse et protectrice, et choisit de dormir dans ses vêtements mouillés; malgré le désir la distance reste maximale entre lui et la femme aperçue.

2. L'échange des vêtements militaires contre des habits civils, autrement dit l'élimination de l'enveloppe, moment de prise de conscience de soi dans le miroir (p. 103), renaissance d'un être réconcilié avec lui-même mais encore maladroit et fragile:

[...] Iglésia et Georges [...] maintenant vêtus comme des valets de ferme, c'est-à-dire vaguement gênés, vaguement mal à l'aise, comme si—au sortir de leur lourde carapace de drap, de cuir, de courroies—ils se sentaient à peu près nus, sans poids dans l'air léger [...] de nouveau dehors, flottant dans cette espèce de vastitude, de vacuité, de vide cotonneux (p. 106).

3. Une mutation interne débouchant sur une plénitude animale du corps:

[...] en un an nous avons appris à nous dépouiller non seulement de cet uniforme qui n'était plus maintenant qu'un dérisoire et honteux stigmaté mais encore pour ainsi dire de notre peau ou plutôt notre peau dépouillée de ce qu'un an plus tôt encore nous nous imaginions qu'elle renfermait, c'est-à-dire même plus des soldats même plus des hommes, ayant peu à peu appris à être quelque chose comme des animaux (pp. 267–268).

Dans cet état finalement salubre, dépouillé de sa carapace, intérieurement reconstitué et revitalisé, Georges peut enfin rejoindre le paradigme du corps exhalé que représente Corinne. A la faveur de l'homogénéité sémiologique enfin conquis,²² le coït peut advenir, et son récit s'entremêle avec celui de l'évasion dans une même euphorie de libération. Notons pour finir une ultime image de Georges sur le modèle de l'intériorité enveloppée, mais cette fois sous la forme harmonieuse—pour un temps du moins—du fruit mûr à la peau tendue par la pulpe:

[...] une fois elle m'atteignit à la figure je sentis l'espèce de saveur bizarre des coups, violente comme si la chair éclatant sur la pommette répandait à l'intérieur en même temps que la douleur comme un jus vert âpre pas désagréable, s'irradiant, pensant à la peau, à la saveur des prunes des reines-claude mûres bleuâtres se fendant et leur jus sucré (p. 271).

Cet exemple d'un élément du 'langage sous-jacent', parmi beaucoup d'autres que mettent en évidence les énoncés de la classe, montre la spécificité de celle-ci. On aura remarqué comment, indépendamment des isotopies référentielles et des noyaux fictionnels, un même substantif peut être employé dans des contextes différents, en particulier par comparaison,²³ et comment un même adjectif peut caractériser plusieurs objets. L'importance des images et des mots à faible valeur classématique détermine ainsi une porosité entre tous les informants de la fiction. Si la classe précédente était celle de la combinatoire des motifs, cette dernière classe est celle de la circulation sémiologique. On observe dans les deux cas la récurrence de motifs ou de sèmes élémentaires (forme, couleur, trajectoire, sensation, dysphorie, etc.) diversement actualisés au fil du texte.²⁴ Mais les investissements fantasmatiques sont plus riches ici qu'ailleurs, parce qu'ils ont pour support la matière, le corps et les éléments naturels, et qu'ils impliquent le niveau le plus archaïque de l'être, à savoir la sensation de soi et le désir de l'autre. C'est ici que l'imaginaire simonien s'invente et s'éprouve.

Conclusion

A. J. Greimas, dans sa *Sémantique structurale*,²⁵ préconisait une procédure d'extraction qui n'est pas très éloignée de la méthodologie ALCESTE:

[...] Le point de départ de l'extraction réside dans la constatation de la redondance de certains lexèmes, représentatifs, a titre d'hypothèse, de l'hypothèse, de l'isotopie à décrire. [...] En considérant [ces lexèmes] comme constituant un premier inventaire restreint, la démarche descriptive débutera par l'extraction de tous les contextes comportant ces lexèmes inventoriés.

Il s'agira ensuite de classer ces contextes et de repérer les principaux lexèmes cooccurrents avec chacune des formes-pivots de départ. Ces lexèmes devront à leur tour être considérés comme des formes-pivots et détermineront autant d'extractions nouvelles, et ainsi de suite, selon un processus arborescent. Greimas poursuit:

La procédure d'extraction apparaît donc, dans son ensemble, comme une série d'opérations successives d'extraction, chaque inventaire de contextes extraits permettant l'extraction et la mise en inventaire d'autres contextes, et ainsi jusqu'à l'épuisement du corpus. [...] Bien plus: on peut supposer que l'élargissement progressif des inventaires extraits, en fournissant de nouveaux éléments d'appréciation, doit permettre de rectifier, si besoin est, l'hypothèse postulée au départ de la description.

La procédure, d'une rigueur exemplaire, nécessite toutefois le choix préalable d'une opposition sémantique pressentie comme cardinale—ainsi que celui de l'opposition lexématique censée la réaliser le mieux, même si ces choix supportent des réajustages en cours de procédure. D'autre part, Greimas présuppose que les séries successives d'extractions permettront de recenser l'intégralité du corpus; or, qui sait si, après quelques inventaires constitués, on ne retomberait pas, lors des extractions suivantes, sur des contextes déjà extraits? On tournerait peut-être la difficulté en augmentant le pas moyen de chaque contexte, mais la quantité d'informations recueillie deviendrait vite difficile à manipuler. En fait, il serait même souhaitable qu'après quelques extractions on ne débouche plus sur des contextes nouveaux: les ensembles de contextes inventoriés, présentant chacun un vocabulaire spécifique et stable, ne seraient rien d'autre, alors, que des classes ALCESTE, avec cet avantage du logiciel sur la procédure manuelle qu'il repère lui-même, par calcul statistique, les termes-pivots les plus discriminants, de même qu'il résout le problème du pas moyen des contextes.

Cette similitude des deux procédures signale une évidence: l'informatique ne vise qu'à automatiser des mécanismes longs et méthodiques, solidement éprouvés par les habitudes critiques.

Notes

1. Analyse Lexicale par Contexte Etablie à partir d'une Segmentation du Texte en Énoncés. Le programme, conçu par Max Reinert (Université de Toulouse-Le Mirail), est commercialisé par la société Image.

2. Deux applications de ce type ont été réalisées par l'auteur du logiciel: voir Max Reinert, 'ALCESTE, une méthodologie d'analyse des données textuelles et une application: *Aurélia* de Gérard de Nerval', *Bulletin de méthodologie sociologique*, 26, 1990, pp. 24–54, et 'La méthodologie ALCESTE d'analyse des données textuelles; application à l'analyse des poésies d'A. Rimbaud', dans G. Maurand, *Poésie et modernité*, C.A.L.S. 1991, pp. 303–325.
3. Max Reinert, *Notice du logiciel Alceste*. Voir aussi, du même auteur, 'Un logiciel d'analyse lexicale [ALCESTE]', *Cahiers de l'analyse des données*, 4, 1986, pp. 471–484.
4. Claude Simon, *La Route des Flandres*, Paris, Éditions de Minuit, 1960. Les références des citations renverront à cette édition.
5. La notion d'énoncé, qui renvoie à une organisation syntaxique mais aussi sémantique du message, reste sujette à caution et en tout état de cause peu opérationnalisable. Le découpage d'un texte (*a fortiori* d'un texte littéraire) en énoncés successifs, à supposer même qu'il soit possible sans arbitraire, ne saurait être réalisé automatiquement. En fait, le logiciel tourne la difficulté en pratiquant plusieurs découpages différents, fondés autant que possible sur la ponctuation, et ne soumet pour finir à l'opérateur que les résultats stables, c'est-à-dire indépendants de ces variations de découpage.
6. Au sens lexicométrique: chaînes de caractères alphanumériques isolées par des délimiteurs (espace ou ponctuation).
7. Il s'agit en fait d'une réduction des formes par reconnaissance des suffixes et des racines courantes du français, ainsi que des terminaisons verbales, des formes conjuguées des verbes irréguliers, et des locutions adverbiales usuelles. La réduction n'intervient qu'en l'absence d'ambiguïté homographique et homonymique. On verra plus loin comment se présentent ces formes réduites (cf p. 59).
8. On se reportera à la notice du logiciel, où Max Reinert expose la méthodologie ALCESTE en décrivant les procédures de calcul utilisées par le programme. La description, rigoureuse et précise, reste d'un accès difficile pour le non-statisticien, et on ne saurait, sans compliquer les choses, en rendre compte dans le cadre de cet exposé. Je recours donc à une analogie, avec le risque d'imprécisions et d'inexactitudes que cela comporte.
9. ALCESTE néglige de même dans les calculs les lexèmes situés aux deux extrémités de l'index hiérarchique. Les seuils d'élimination sont paramétrables, mais on s'aperçoit vite que les modifications ont peu d'incidence sur les résultats.
10. On rejoint ainsi la notion de *typicalité* (*typicality*) mise en place depuis les travaux de E. Rosch (voir Danièle Dubois, éd., *Sémantique et cognition. Catégories, prototypes, typicalité*, Paris, CNRS Éditions, coll. 'Sciences du langage', 1993).
11. ALCESTE donne une indication chiffrée de cette (non-)coïncidence, à condition que la partition en chapitres, parties, etc., ait fait l'objet du codage requis. On pourra savoir dans quelle proportion une classe emprunte ses énoncés à un chapitre particulier.
12. Sans que les classes se réduisent pour autant à ces chapitres. On peut observer comment, par exemple, les énoncés de tel chapitre sur l'Amérique et ceux de tel autre sur Rancy ont été regroupés au sein d'une même classe, signe que les expériences et jugements de Bardamu relativement à ces deux endroits s'énoncent pratiquement dans les mêmes termes.
13. Jean Ricardou, 'Un ordre dans la débâcle', *Problèmes du nouveau roman*, Paris, Éditions du Seuil, coll. 'Tel Quel',

- 1967, p. 52 (article initialement paru dans *Critique*, déc. 1960).
14. C'est-à-dire non accentuées, tronquées à la suite d'une reconnaissance d'un radical ou d'un suffixe (suivies alors du signe + ou <), ou rapportées, dans le cas de certaines formes verbales reconnues, à un infinitif (suivi alors d'un point). Exemples: 'admettre.' correspond à 'admettant', 'admettons', 'admis', 'admise'; 'certain<' correspond à 'certain', 'certaine', 'certains', 'certaines' (reconnaissance du radical et des terminaisons de flexion), et 'reel+' correspond à 'réel', 'réelle', 'réellement', 'réelles', 'réels' (reconnaissance du seul radical). Par ailleurs les formes consécutives reconnues comme formant une locution ont été liées par un trait d'union. Ces manipulations effectuées automatiquement par le logiciel mènent çà et là à des erreurs (par exemple: 'soi+' correspondant à la fois à 'soies' et à 'soir'), auxquelles s'ajoutent les inévitables confusions d'homonymes. Toutes ces erreurs restent néanmoins en nombre limité par rapport à l'ensemble, et n'ont donc pas d'incidence sur les statistiques.
 15. Le classement a été établi à partir de toutes les formes listées par le programme. Mais n'apparaissent ici, par souci de clarté, que les formes pourvues d'un chi2 supérieur à 7 (d'où quelques catégories en apparence peu fournies).
 16. Les formes conjuguées de *dire*, des auxiliaires *être* et *avoir* et des verbes modaux, ainsi que les pronoms et adjectifs personnels, ont été négligés lors de la comparaison des énoncés, et n'apparaissent dans le profil des classes qu'à titre d'éléments illustratifs. Il est frappant de constater ici combien ces formes sont présentes dans les énoncés de la classe alors que les calculs qui ont permis la constitution de cette classe les ont ignorés. Il faut en déduire que les cooccurrences des autres formes sont d'autant plus nombreuses.
 17. S'ajoute à cela le fait que le dialogue, qui le plus souvent inaugure ou ponctue un récit dans le récit, constitue en grande partie l'action du roman, lequel se présente comme une structure—structure certes impossible, brouillée—de récits enchâssés (conversations entre Georges et Blum prisonniers, récits d'Iglésia sur Corinne, etc.).
 18. Voir Algirdas Julien Greimas, *Sémantique structurale* (1966), Paris, PUF, coll. 'Formes sémiotiques', 1986, en particulier pp. 42–54. Le classème est un sème qui indique à quel objet, personne, etc., se rapporte un lexème dans un contexte particulier. L'association du noyau sémiologique d'un lexème (qui est son contenu invariant d'un contexte à l'autre) à tel ou tel classème forme le sémème. Nous recourrons souvent à l'opposition du sémiologique et du classématique.
 19. Voir par exemple les 'Cinq notes sur Claude Simon', de Maurice Merleau-Ponty, *Médiations*, 4, hiver 1961-1962, pp. 5-10, reprises dans *Entretiens*, 31, 1972, ainsi que les articles de Michel Deguy, 'Claude Simon et la représentation', *Critique*, no. 187, 1962, pp. 1009-1032, et de Serge Doubrovsky, 'Notes sur la genèse d'une écriture', *Entretiens*, 31, 1972, pp. 51-64.
 20. Voir note p. 18.
 21. Cf pp. 20, 69, 148, 269 et 272.
 22. Il faudrait repérer ce parcours dans un système global de signification, suggéré par la lecture de la classe mais trop vaste pour qu'on le décrive ici. En voici seulement quelques éléments qui en montreront la cohérence. Iglésia se situe au premier chef dans le paradigme du corps enveloppé. Il n'est même qu'une enveloppe, qu'une surface inerte sans intériorité: un 'masque de cuir et d'os [...], impénétrable, triste, taciturne' (p. 48), 'une feuille de papier déchiré avec deux trous pour les yeux [...] sortant du col raide et mouillé du manteau comme hors d'une carapace' (pp. 37-38). Sa liaison avec Corinne, figure opposée, est ainsi vouée à l'échec. Le Conventionnel, quant à lui, semble payer de sa vie le fait d'avoir retiré trop précipitamment sa 'tunique poussiéreuse [...] comme s'il avait apporté sur lui une impalpable et tenace poussière de décombres' (p. 180), et d'avoir forcé un rapprochement contre nature avec 'la liquide tiédeur du ventre', la 'soyeuse écume' (p. 175) du corps de sa maîtresse. La monstruosité ou le danger du passage brutal d'un paradigme à l'autre imposent donc à la figure de Georges l'évolution que nous venons d'évoquer.
 23. Voir la forte typicité de 'comme'.
 24. Une opération préalable au traitement du texte par ALCESTE permettrait d'ailleurs d'améliorer la mise en évidence de ce type de récurrences. L'idéal serait en effet de marquer comme formes 'illustratives' (formes négligées dans les calculs, comme c'est déjà le cas, par défaut, des mots-outils les plus courants) toutes les formes à référent unique qui sont des informants de la fiction: les noms de personnages, mais aussi beaucoup d'autres ('grange', par exemple, qui désigne 9 fois sur 10 le même endroit). Si on élimine ces mots les oppositions entre les mondes lexicaux correspondront encore moins à des oppositions entre des noyaux fictionnels, et elles seront donc plus immédiatement révélatrices des structures les plus profondes de la signification.
 25. Algirdas Julien Greimas, *op. cit.*, pp. 223 sqq.