

**François RASTIER**

CNRS, INaLF, Paris / CLF, Paris IV

## LA SÉMANTIQUE DES THÈMES OU LE VOYAGE SENTIMENTAL

*Il aurait fallu juger des paroles par les sentiments et  
non pas des sentiments par les paroles. (Bossuet)*

Malgré sa place, cette étude n'est pas une conclusion ni une synthèse. Simple contribution personnelle, elle n'engage aucunement les autres collaborateurs du volume. L'occasion que j'ai eue de lire l'ensemble me permet simplement un regard rétrospectif.

### **I. Pour définir le thème**

La notion de thème tient une place particulière dans le paysage intellectuel français. La critique thématique issue de Bachelard a été si bien acceptée par les autorités académiques que les programmes des concours l'honorent régulièrement, et les sujets de mémoire ou de thèse qui s'en inspirent se comptent par centaines chaque année. Elle s'est tant diffusée que bien des recueils de textes à l'usage des écoliers, du petit classique au manuel, sont maintenant organisés par thèmes.

Fort utilisée, la notion de thème reste en général intuitive<sup>1</sup>. Du moins, les définitions ordinaires n'ont pas de rapport précis avec les sciences du langage. Par exemple J.-P. Richard, dont les études thématiques se signalent par leur finesse inégalée, définit le thème comme « un principe concret d'organisation, un schème ou objet fixe autour duquel aurait tendance à se constituer un monde » (1961, p. 24). Ce vocabulaire d'inspiration philosophique rend hommage à Merleau-Ponty, et plus généralement à la phénoménologie existentielle.

La notion philosophique de thème nous paraît insuffisante. Soit elle désigne une catégorie transcendante, ou un schème, voire un archétype au sens jungien, alors que notre objectif est de montrer la formation et l'évolution des thèmes au

sein d'une sémantique historique et comparée, partie intégrante d'une sémiotique des cultures — non d'une anthropologie philosophique<sup>2</sup>. Soit elle s'applique à la relation d'un sujet au monde, et veut en faire la cause de son langage, alors même que cette relation est médiatisée par son expression. Dans les deux cas l'autonomie relative du sémiotique se trouve éludée, le monde ou le sujet philosophique gouvernant la problématique du thème.

Définir plus généralement le thème comme un concept ou comme une « relation affective » le maintiendrait à l'écart des sciences du langage. Du moins, les définitions qui ont cours doivent-elles être rapportées à l'analyse linguistique des textes, et réélaborées dans ce cadre.

En linguistique, le thème pourrait se définir par diverses voies, selon qu'on privilégie le signe ou le texte, et dans le signe, le signifiant ou le signifié.

a) La voie lexicographique définit le thème comme un mot-vedette, généralement un substantif, auquel sont rapportés divers paronymes ou équivalents partiels. Un dictionnaire de thèmes sera donc un sous-ensemble d'un dictionnaire. Cette voie est liée à une linguistique du signe.

b) La voie sémantique ne confère pas de prééminence à un mot-vedette identifié par son signifiant. Elle met en évidence des réseaux de récurrences sémantiques. Elle relève plutôt de la linguistique du texte que de la linguistique du signe. Explorons donc cette voie.

#### a) La voie sémantique

Nous nommerons *thème* une structure stable de traits sémantiques (ou *sèmes*), récurrente dans un corpus, et susceptible de lexicalisations diverses. Cette définition provisoire va être spécifiée, en caractérisant ces traits, ce corpus, et ces lexicalisations ; puis en posant la question de l'interrelation des thèmes entre eux et des paradigmes thématiques.

(i) La notion de catégorie sémantique est trop vague<sup>3</sup> et l'on convient de distinguer entre sèmes génériques et spécifiques (*cf.* Portier, 1974) : les uns indexent les sèmes dans des classes (taxâmes, domaines et dimensions), les autres les opposent aux membres de leur classe de définition.

La récurrence d'un sème générique induit une isotopie générique. Et parfois, dans son acception générale, le mot *thème* est employé pour désigner le « sujet » d'un texte, c'est-à-dire son isotopie générique dominante, ordinairement un domaine sémantique<sup>4</sup>. Par exemple, on pourra caractériser (très faiblement) *La Princesse de Clèves* comme un roman d'amour. Nous avons parlé à ce propos de *thème générique*, ce qui est quelque peu abusif, car une isotopie n'est pas une structure<sup>5</sup>.

En revanche, un thème spécifique peut se définir comme une molécule sémique, c'est-à-dire un groupement structuré de sèmes spécifiques. Une molécule se représente par un graphe sémantique, dont les nœuds sont étiquetés par des sèmes, et les liens, par des primitives sémantiques (cas et relations structurales). En voici un exemple<sup>6</sup>.

Excursus 1 : De l'ennui

On hésite à revenir sur l'ennui après la thèse de Sagnes (1969), les études de Bouchez (1973), et d'Éveline Martin (1993, pp. 146-202) et de David Erlich ici même. Mais les données déjà recueillies, publiées et élaborées nous permettront d'abrégier notre illustration.

Un sentiment est une structure actantielle où un actant humain se trouve affecté d'évaluations. Cette structure peut être représentée par un graphe sémantique<sup>7</sup>, qui décrit sa molécule sémique. Définissons d'abord le contenu du graphe ci-dessous en soulignant de façon différenciée les étiquettes de ses nœuds et de ses liens. Le graphe se divise en deux sections : l'une niée (désormais N) décrit la possession d'un objet de valeur, l'autre affirmée (désormais A) l'action du sentiment sur le sujet.

Voici les sèmes constituants de la molécule sémique suivis de quelques expressions typiques.

L'actant Ego n'est pas spécifié autrement que par le sème /humain/. Les autres sèmes sont propagés par le contexte. Il est Ergatif dans N, et Accusatif dans A. L'agent est défini alors par le irait /sentiment/ ; les autres spécifications de l'ennui sont définies par les relations entre l'agent et le reste du graphe.

L'actant Accusatif dans N est spécifié par une évaluation positive /mélioratif/.

Le procès dans N est nié. On trouvera donc dans les contextes le sème /privation/, exprimé par des tours négatifs (*aucun moyen de comprendre, aucune occupation, ne sachant que faire, il ne rencontre rien de ferme, sans que le désir s'y mêlât*) ; des préfixes (inaction, impossibilité). D'autres expressions ne sont négatives que par inférence (ex. : *vide, mains ballantes, province*).

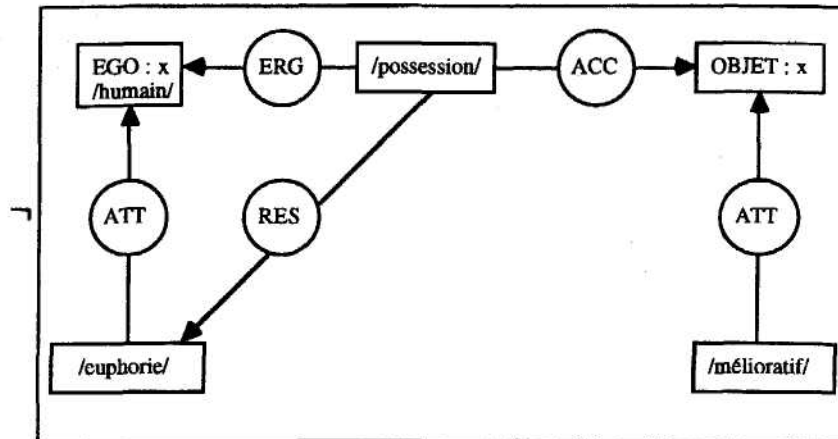
Dans A, le procès est caractérisé par deux attributs sémiques :

/imperfectivité/ : *indéfiniment, hébétude méditative, langueurs de moribond, abattement sans bornes* ; mais aussi l'emploi fréquent de temps verbaux imperfectifs, notamment l'imparfait ; ou encore *araignée, dimanche*.

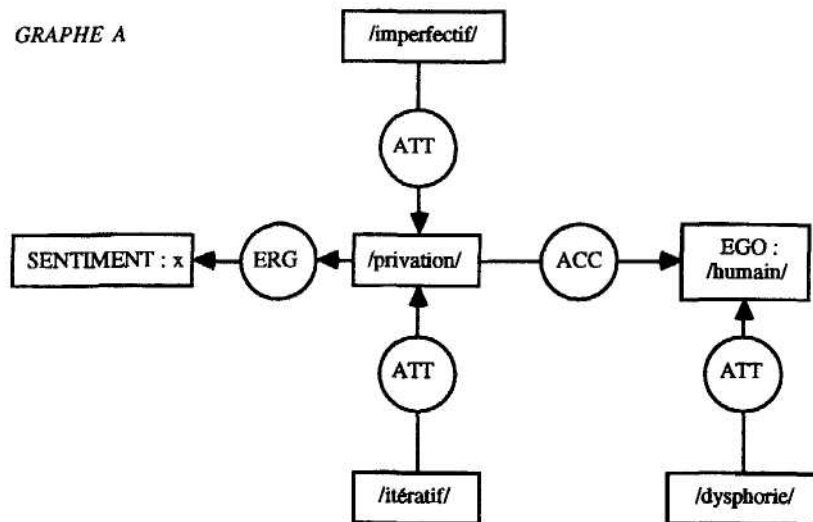
/itérativité/ : *désagrège, dévore, pétrit, rongé, lèpre, habitudes sans cesse recommençantes, les jours se succédaient, une régularité monotone, il retrouvait [...] le même ennui*. Dans ces expressions, les deux sèmes sont fréquemment lexicalisés ensemble\* : ex. *errer*.

Le résultat du procès est un attribut /dysphorie/affecté à Ego, et qui se lexicalisé par des expressions comme : *vide, hébétude, endort, alanguir*.

GRAPHE N



GRAPHE A



N.B. : ERG abrège ergatif, ACC accusatif, ATT attributif, RES résultatif ; - symbolise la négation. Ces graphes sont des types, car certains de leurs nœuds, étiquetés en capitales, peuvent ne pas être instanciés (comme le rappelle le symbole x).

La capacité descriptive de cette représentation peut être mise à l'épreuve sur des exemples, de manière à souligner aussi comment les écrivains s'approprient le thème. Quand Prévert écrit : « ceux qui crévent d'ennui le dimanche après-midi parce qu'ils voient venir le lundi » (*Paroles*, 1946, pp. 19-20), il remotive le dimanche, fréquemment associé à l'ennui, par le sème /imperfectif. Par la menace du lundi, topos du populisme de l'époque<sup>9</sup>, il fait du travail ce qui gâche même l'inaction.

Ou encore, Barthes note : « L'ennui n'est pas loin de la jouissance : il est la jouissance vue des rives du plaisir » (1973, p. 43), il instancie par la jouissance l'objet du désir dont Ego est séparé (graphe N)- La séparation est figurée par l'image du fleuve, qui reprend les traits /imperfectivité/ et /itérativité/. Et classiquement, le plaisir reste associé à l'ennui (É. Martin, *op. cit.*, p. 108, relève dans des contextes antérieurs : *l'air d'ennui dans les étreintes ardentes, l'ennui de cet accouplement*), et il se trouve dans le graphe N à sa place. Le remaniement lacanien du thème aura consisté à mettre la jouissance là où les poètes fin de siècle plaçaient l'azur.

(ii) Un thème, défini comme molécule sémique, peut recevoir des expressions diverses, par des unités qui vont du morphème au syntagme. Nous les nommerons, pour simplifier, *lexicalisations*. On peut distinguer des lexicalisations synthétiques qui manifestent au moins deux sèmes, et des lexicalisations analytiques, qui n'en manifestent qu'un. Ainsi, un thème peut être manifesté de manière diffuse, par exemple dans un paragraphe où divers sèmes seront lexicalisés tour à tour. La lexicalisation la plus synthétique ne jouit d'aucune prééminence théorique par rapport aux autres lexicalisations : elle n'est pas le « mot juste » dont toutes les autres expressions ne seraient que d'imparfaits avatars<sup>10</sup>.

Selon les discours et les genres, les normes de lexicalisation des thèmes varient : la poésie lyrique (la seule qui nous reste) cultive les lexicalisations analytiques, alors que dans les discours techniques, les synthétiques sont de rigueur.

Même si leur lexicalisation la plus synthétique relève d'une classe bien identifiée, les thèmes sont indépendants d'une classe sémantique, ou plus exactement peuvent se manifester sur diverses isotopies génériques. Par exemple, si l'on nomme Ennui la molécule sémique qui comprend les traits /privation/ (notamment : /inactivité/), /imperfectif/, /itératif/ (souvent combinés en /monotonie/), ce thème peut se manifester par *araignée*, par *dimanche*, ou par *monotone*. Il ne s'agit pas de métaphores, car un thème spécifique est par principe indépendant de toute isotopie générique ; en d'autres termes, il n'y pas de mot propre, même si le mot *ennui* reste une dénomination commode. Et alors que ce mot se rencontre seulement quatre fois dans *Madame Bovary*, les composants du thème apparaissent souvent, notamment à propos de Charles. Par exemple, dans cette phrase célèbre : « La conversation de Charles était [/imperfectif/] plate [/imperfectif/, /monotonie/] comme un trottoir de rue [/monotonie/], et les idées de tout le monde [/itératif/, /monotonie/] y défilaient [/imperfectif/, /itératif/], dans leur costume ordinaire [/itératif/, /monotonie/], sans [/privation/] exciter d'émotion [/euphorie/], de rire [/euphorie/] ou de rêverie [/euphorie/] » (I, vn)<sup>11</sup>.

Cela rappelle le principe — souligné par Hjelmslev — qu'il n'y a pas d'isomorphisme entre les plans de l'expression et du contenu, malgré les vœux d'une certaine linguistique du signe.

Les molécules sémiques sont des formes sémantiques simples, alors que les isotopies génériques sont des fonds sémantiques sur lesquels elles se présentent à la perception. Dans certains discours, le rapport entre forme et fond est univoque ; mais dans le discours littéraire, des formes analogues peuvent se présenter sur des fonds différents, qui sont entre eux dans des relations de hiérarchie et de dominance. Quand deux occurrences de la même molécule sont relevées sur deux isotopies différentes, on peut les dire en relation métaphorique. Mais l'orientation de cette relation dépend de la dominance entre isotopies et du parcours interprétatif qui l'établit. En d'autres termes, *frisson* n'est pas une métaphore de *peur* ; ces deux mots lexicalisent sur des isotopies génériques différentes le même thème. L'opposition entre les dimensions //physique// et //moral// est alors secondaire pour l'analyse thématique.

Il reste bien sûr à préciser l'incidence des fonds sur les formes, et les remaniements que peut apporter à une molécule sémique l'adjonction d'un trait générique<sup>12</sup>. Mais on voit pourquoi une théorie des classes sémantiques reste utile, mais ne suffit pas à constituer une thématique, car elle ne concerne que les fonds. En prenant pour exemple les sentiments, nous entendons inventorier et décrire les formes sémantiques reliées à un même fond de référence par le déploiement d'une classe sémantique.

Le rapport entre forme et fond souligne la dépendance de la perception sémantique à l'égard du contexte (*cf.* l'auteur, 1991, ch. VIII). Ce rapport n'est pas statique, et ne doit pas être conçu à l'image d'une figure géométrique sur un plan ; la comparaison avec la composition musicale ne serait pas mauvaise. Les formes elles-mêmes ont en effet des modes de diffusion divers, et peuvent passer à l'arrière-plan en constituant des faisceaux d'isotopies spécifiques. Ainsi un thème peut être latent ou saillant, selon que des constituants sont épars ou regroupés.

#### *b) Le palier de l'analyse thématique*

La notion d'unité textuelle est quelque peu ambiguë. Une analyse de texte peut en rester au palier du mot, comme le fait par exemple la statistique lexicale. Rappelons donc qu'un texte peut être analysé à trois paliers principaux : micro-, méso-, et macrosémantique, qui correspondent au sémème, au contenu de la période, et à la structure textuelle.

Ces trois paliers correspondent d'une part à des degrés de systématicité dominants : système fonctionnel de la langue, normes socialisées, normes idiolectales. D'autre part, à des zones de localité qui intéressent la propagation des traits sémantiques : elle est maximale au sein du syntagme ; bonne entre les syntagmes d'une même période ; entre périodes, cette propagation demande une prise en charge par des structures macrosémantiques.

Chaque trait sémantique a un potentiel d'activation, qui se diffuse localement, en fonction des inhibitions et facilitations régulées par les structures morphosyntaxiques. Selon la nature des traits, et le mode de propagation de l'activation, on peut distinguer trois cas remarquables.

L'actualisation d'un trait favorise sa répétition. En ce cas, et selon le statut de ce trait, cela constitue une isotopie générique ou spécifique. La production des antonymes, massivement attestée par les associationnistes du siècle dernier, est un exemple d'activation au sein d'un même taxème, par la constitution d'une isotopie générique minimale.

L'actualisation d'un trait favorise aussi la répétition des traits voisins dans la même molécule sémique : c'est pourquoi des lexicalisations partielles d'un même thème sont fréquemment cooccurentes dans la même période, voire dans le même syntagme. Ce phénomène pourrait être appelé *paratopie*. Il est à l'œuvre dans ce que l'on nomme les anaphores associatives<sup>13</sup>. Ces diffusions d'activation sont le corrélat sémantique des phénomènes que la *Gestalt* nommait *lois de bonne continuité*, et que la psychologie cognitive étudie sous le nom général d'amorçage (*priming*). Elles justifient sémantiquement l'étude statistique des cooccurrences lexicales pour l'analyse thématique.

Le thème, par ses récurrences, intéresse la macrosémantique. Mais en tant qu'unité, il relève de la mésosémantique. D'après nos relevés, ses lexicalisations diverses apparaissent généralement dans un espace inférieur à trois cents mots. Un espace de cinquante mots environ suffit pour identifier quatre occurrences d'un thème sur cinq.

Présenter la thématique relativement aux différents domaines de l'analyse textuelle excéderait l'objectif de cette contribution. Notons simplement les relations du thème par rapport aux autres composantes textuelles (c/. l'auteur, 1989,1). Relativement à la tactique, le thème a des positions identifiables — dans ses manifestations denses — on peut relever des rythmes thématiques. Relativement à la dialectique, les occurrences groupées des corrélats d'un thème ont une même position dans un intervalle dialectique ; et relativement à la dialogique, elles sont situées dans un même monde et dans un même univers. Cela spécifie les thèmes par rapport aux unités des autres composantes sémantiques, qui mettent en jeu des différences dialectiques, dialogiques, ou tactiques. Par exemple, un thème récurrent dans plusieurs intervalles dialectiques et intégré à des structures dialectiques se définit alors comme un acteur<sup>14</sup>.

### c) Les structures thématiques et topiques

Aux conditions statistiques de définition du thème s'ajoutent des conditions structurales.

Le problème de l'organisation paradigmatique des thèmes se pose d'une façon différente de celle des paradigmes lexicaux proprement dits, car ils relèvent d'un palier de complexité supérieur. Les classes de thèmes ne relèvent donc pas de la lexicologie, encore moins de la lexicographie<sup>15</sup>, mais de la thématique ; ou de la *topique*, si l'on convient de nommer ainsi l'étude des formes sémantiques stéréotypées au palier mésosémantique<sup>16</sup>. Pour tracer la limite entre thème et topos, admettons qu'un thème est récurrent au moins une fois dans le même texte ; un topos au moins une fois chez deux auteurs différents.

S'il existe des paradigmes thématiques, la méthodologie de leur (re)constitution diffère sans doute de celle des paradigmes lexicaux, car ils n'ont pas le même statut. Par rapport aux thèmes, les sémèmes sont des unités primaires, et qui sont réputées appartenir au système de la langue, alors que les thèmes relèvent de normes socialisées.

La relation structurale la plus simple est l'antonymie. Et les études de cooccurrence lexicale, quand elles sont menées avec soin, montrent qu'un nom de sentiment est fréquemment associé à son antonyme<sup>17</sup>. Mais si l'antonymie est fréquente au palier lexical, il n'est pas certain qu'elle soit généralisée au palier thématique. Parmi les lexicalisations de thèmes, certaines peuvent se trouver en relation d'antonymie, d'autres non.

Quand elle est attestée, elle se manifeste par des séries d'oppositions sémiques (alors que les sémèmes antonymes ne diffèrent ordinairement que par un sème). Par exemple, le topos complexe de la fleur au bord de l'abîme, fort récurrent à l'époque romantique, comprend deux thèmes, Fleur et Abîme, qui peuvent être lexicalisés par *rosé*, *plante*, *gouffre*, *précipice*, *vertige*, *profondeur*, etc. Ils s'opposent par les catégories sémiques /saillant/ vs /creux/, /fragile/ vs /puissant/, /attirant/ vs /repoussant/, /vivant/ vs /mortel/, /coloré/ vs /sombre/ (cf. l'auteur, 1989, p. 63). Mais on ne saurait généraliser cet exemple, car l'antonymie lexicale entre sémèmes ne se traduit pas au palier supérieur par une antonymie entre les molécules sémiques qui constituent les thèmes<sup>18</sup>.

Retenons que thèmes et topoï sont susceptibles de divers types de groupements. Une seule molécule sémique comme celle de l'Épée peut devenir partie d'un groupement qui met en relation deux molécules (ex. la Plume et l'Épée ; les Armes et les Amours ; la Maman et la Putain<sup>19</sup>). Les groupements ternaires sont légion dans la tradition indo-européenne, en liaison de ce que Dumézil a appelé l'idéologie trifonctionnelle ; les groupements quinaires abondent en Chine (comme l'a souligné notamment Gernet) ; les septénaires dans l'antiquité moyen-orientale (sans doute sous l'influence du culte astral babylonien). Bref, ces groupements sont l'homologue des taxèmes au palier lexical, mais leurs structures reflètent des normes d'un autre ordre. Par ailleurs,



les couples antithétiques, comme les Armes et les Amours opposent et relie des domaines et dimensions topiques.

On ne peut cependant transposer directement ce que l'on sait des structures lexicales aux structures thématiques. C'est précisément un des buts de la thématique de repérer les regroupements de thèmes, de faire la part entre topoï et thèmes personnels<sup>20</sup> ; enfin, montrer les traitements personnels des topoï, ou plus précisément souligner comment les topoï sont spécifiés, voire remaniés par leur contexte.

#### d) *L'étude des motifs*

Après le topos, un mot du motif, car thème et motif sont fréquemment associés, dans une tradition qui remonte à la folkloristique du siècle dernier<sup>21</sup>. Cette notion libérale appelle des clarifications : malgré les propositions théoriques des formalistes russes, les listes de motifs ne sont pas distinguées des inventaires de thèmes, et selon les bibliographies, on y trouve aussi bien le Juif errant, le féminisme dans l'*Enéide*, le mesmérisme et le tabac. Trousson a naguère proposé une distinction : un motif est « une toile de fond, un concept large, désignant soit une certaine attitude — par exemple la révolte — soit une situation de base, impersonnelle, dont les acteurs n'ont pas encore été individualisés », alors qu'un thème serait « l'expression particulière d'un motif, son individualisation [...] On dira que le motif de la séduction s'incarne, s'individualise et se concrétise dans le personnage de Dom Juan » (1981, pp. 21-22). Il reste une confusion des relations type/instance et fonction/acteur qui se conjoignent dans le rapport *séduction / Dom Juan* ; en outre, ces relations ne sont pas distinguées des relations fond/forme (le motif est comparé à une toile de fond, comme chez Sauer).

Les motifs peuvent être (re) définis comme des structures textuelles complexes de rang supérieur (macrosémantique) qui comportent des éléments thématiques, mais aussi dialectiques (par changement d'intervalle temporel), et dialogiques (par changement de modalité). Par exemple, le motif du *mort reconnaissant*, relevé par le *Motif-Index of Folk-Literature* de Thompson, est une structure thématique et dialectique complexe, qui met en jeu des fonctions décès, bienfait, et gratitude, ainsi que des acteurs humains. En somme, le motif est un syntagme narratif stéréotypé, et partiellement instancié par des topoï, alors que le thème est une unité du palier inférieur, non nécessairement stéréotypée, et qui se trouve dans toutes les sortes de textes. Bref, le thème est au syntagme narratif ce que le topos est au motif.

Soit, en bref :

	Unités thématiques	Unités dialectiques
D'un discours ou d'un genre	<i>Topos</i>	<i>Motif</i>
D'un texte	<i>Thème</i>	<i>Fonction, syntagme narratif</i>

Ces formes sont susceptibles de degrés de typicité, selon que l'on considère — dans une représentation par graphes — seulement les étiquettes des liens, certaines étiquettes des nœuds, ou toutes les étiquettes des nœuds.

N.B. : Le *type*, tel qu'il est défini par Panofsky en iconologie (1967, p. 17), correspond explicitement au *topos* en histoire de la littérature chez Curtius. Proposons une précision, d'après l'exemple de Panofsky : le *topos* de la femme à l'épée devient un *type* dès lors par exemple qu'on l'identifie comme Judith. En ce cas un *type* est une sorte de *topos*, dont la molécule sémique contient un nom propre.

*Topoi* et motifs sont deux sortes de formes sémantiques. Comme tels, ils ont une histoire, et sont tributaires de corpus. Leur étude relève de la sémantique historique et comparée, partie intégrante de la sémiotique des cultures. En étudiant les sentiments nous avons voulu contribuer à cette entreprise. Cela impose une déontologie philologique dans la délimitation et l'exploitation des corpus ; cela suscite parallèlement des questions herméneutiques.

*e) Le problème de la pertinence et l'assise herméneutique de la thématique*

Comme toutes les unités sémantiques, un thème est un construit, non un donné : en cela, la thématique dépend de conditions herméneutiques.

Les dictionnaires et index thématiques reflètent l'état moyen des conceptions contemporaines du thème. Si par exemple, comme l'a relevé plus haut F. Surdel, le thème de la pitié en est absent, c'est que les sentiments, *a fortiori* les sentiments charitables, sont de fait exclus du champ de la thématique à la française — qui, influencée notamment par Bachelard et Merleau-Ponty, s'est attachée à décrire plutôt les qualités sensibles que les qualités morales. A notre connaissance, aucun dictionnaire de thèmes n'est encore construit à partir d'un corpus, aucun ne part de considérations linguistiques. La plupart se bornent à compiler des listes sans principe de définition commun, et dont des classements de bon sens cachent le caractère hétéroclite. L'intuition, indispensable, doit être soumise à un contrôle méthodologique. De prétendus thèmes comme l'agressivité sont des catégories descriptives générales qui peuvent correspondre à des hypothèses anthropologiques, non à des thèmes dans un corpus littéraire classique. Le caractère anachronique des catégories descriptives doit être problématisé : si l'agressivité est vieille comme Caïn, le concept d'agressivité est éminemment moderne. Rendre justice aux textes du passé, pouvoir leur donner sens, c'est aussi apprécier ce qui nous en sépare.

En outre, l'interprétation des données textuelles se place dans un cercle méthodologique — dépendant du cercle herméneutique. L'analyse lexicale, dont la statistique est un auxiliaire, ne propose pas d'elle-même des indices à l'analyse thématique. Toute sélection de corpus, tout prélèvement dans un corpus, tout recueil de données est tributaire de choix qu'il importe de rendre explicites. En d'autres termes, pour atteindre ses objectifs, la thématique doit guider l'analyse lexicale, puis interpréter ses résultats qui sans cela resteraient inutilisables pour une sémantique textuelle. Les logiciels d'interrogation imposent certaines démarches, mais ne proposent rien. Ils servent à confirmer ou infirmer des hypothèses, qui dépendent de la stratégie d'interprétation.

*Remarque* : Plus généralement, la notion de *donnée* invite à la prudence. En premier lieu, les informations non interprétables — comme les fréquences absolues — ne méritent pas d'être considérées comme des données. D'autre part, les données textuelles n'ont rien de sémantique, ce sont des chaînes de caractères localisées par quelques précisions philologiques minimales : l'auteur, l'œuvre, la date de l'œuvre et la place dans l'œuvre. Les chaînes de caractères ont le même type d'objectivité que les suites de chiffres. Elles sont également dépourvues de sens : on peut les identifier, non les interpréter.

En somme, une donnée est « donnée » non à l'observateur, mais *par* l'observateur, du seul fait qu'une hypothèse a présidé à son recueil. D'une part, toute acception est liée à un genre et à un discours, il importe donc, pour la rendre interprétable, de restituer cet entour. De la même façon qu'en archéologie, mêler des vestiges d'un même champ de fouilles conduit à l'échec, mêler des textes hétérogènes par le type de discours, le genre et l'époque, interdirait de restituer les normes sémantiques — qui sont notamment l'objet de la thématique.

Le choix d'un corpus répond à une attente globale. En outre, tout sous-ensemble du corpus convoqué par une requête répond à une attente plus précise. L'art de l'interprétation — j'adapte ici l'expression de Schleiermacher — consiste alors à croiser les requêtes pour objectiver les attentes.

## II. De l'analyse lexicale à l'analyse thématique

### 1. Difficultés de principe

Détaillons les difficultés théoriques rencontrées pour passer de l'analyse lexicale à l'analyse thématique. À la différence des lexèmes, les thèmes ne sont pas des signes, ni, corrélativement, des unités du français : ils dépendent en effet d'autres normes que la langue. Si le lexème et le thème diffèrent aussi bien par le niveau que par le palier d'analyse, le premier étant un signe, et relevant de la morphologie et de la microsémantique, le second une unité du contenu au palier mésosémantique, il est clair que tout lexème n'est pas un thème.

Une analyse thématique qui en resterait au palier lexical compterait potentiellement autant de thèmes que de mots de la langue. Sauf bien sûr à restreindre cet inventaire, comme le font les dictionnaires de thématique, de façon normative et non critique. On objectera que les thèmes sont ordinairement dénommés par un lexème. Mais ce lexème est simplement une lexicalisation privilégiée du thème. Et l'on pourrait fort bien rencontrer des thèmes sans lexicalisation privilégiée (comme celui que nous avons décrit chez Zola, 1989).

Comment savoir si un lexème donne ou non accès à un thème ? Nous retrouvons ici le problème de la pertinence. Pour l'éclaircir, posons-le à propos d'un mot, *nombril* chez Flaubert. Après enquête, nous pouvons faire état de quelques modestes découvertes.

Excursus 2 : *Nombril dans l'œuvre de Flaubert*

Le choix de ce « thème d'étude » nous a été suggéré par une affirmation vertigineuse de Frédéric Nef : « Madame Bovary n'a pas une infinité de nombrils possibles ; tout simplement, elle n'a pas de nombril. Dans tous les mondes accessibles à partir du nôtre c'est un individu incomplet à qui manque cette propriété » (1992, p. 10).

Le mot *nombril* revient seize fois dans l'œuvre de Flaubert, récits de voyage et correspondance comprise (éd. Conard). Il ne figure pas dans *Madame Bovary*.

Dans la première version de l' *Education sentimentale* (1845) on trouve une commère « d'environ quarante-sept ans, assez fraîche encore, bien nippée et bien nourrie, un peu haute en couleur, l'œil vif et le caquet prompt, très fournie de gorge, puisqu'on entend par là ce qui s'étend depuis le menton jusqu'au nombril » (p. 39). La *Tentation* de 1849 reprend : « à l'ombre de la vigne, couchée sur le gazon, elle avance les lèvres pour saisir le raisin mûr : un grain tombe, il glisse sur sa joue, et, roulant entre ses seins, la chatouille tout entière, depuis le menton jusqu'au nombril » (p. 360).

La correspondance atteste aussi cette acception originale de *nombril* comme limite inférieure d'une gorge abondante, à propos des femmes grecques, qu'elles soient modernes : «Elles prennent leur vêtement, et pour se cacher la mine, se découvrent ce qu'il est convenu d'appeler la gorge, c'est-à-dire l'espace compris entre le menton et le nombril » (*Corr.*, 1850, p. 136) —ou bien antiques comme l'atteste cette description d'un bas relief sur l'Acropole : « Il ne reste plus que les deux seins, depuis la naissance du cou jusqu'au dessus du nombril » (*Corr.*, 1852, p. 298).

Dans la *Tentation* de 1874, le nombril féminin s'érotise encore jusqu'à annoncer le sexe : « J'ai d'abord cherché une femme comme il convient : de race militaire, épouse d'un roi, très bonne, extrêmement belle, le nombril profond, le corps ferme comme du diamant : et au temps de pleine lune, sans l'auxiliaire d'aucun mâle, je suis entré dans son ventre » (p. 123).

Dans *Salammbô* et la *Tentation*, *nombril* prend une autre acception : il est alors lié au divin ou au magique, comme origine d'un ou plusieurs cordons ombilicaux, voire de chaînes qui devraient enchanter les nombreux psychanalystes de Gustave : « Son corps occupait le plafond tout entier. De son nombril pendait à un fil un œuf énorme » (*Salammbô*, p. 81). Ou encore : « plusieurs rues partaient devant lui. Dans chacune d'elles, un triple rang de chaînes en bronze, fixées au nombril des Dieux Pateques, s'étendait d'un bout à l'autre, parallèlement » (p. 162). Dans la *Tentation* de 1849 des courtisanes « ont acheté à des mages des plaques de métal qui se portent au nombril » (p. 219).

L'ombilic géniteur se retrouve dans : « des fœtus quadruplés se tenant par le nombril » (*Tentation*, 1849, p. 408) ; puis en 1856 : « des fœtus quadruplés se tenant par le nombril et valsant comme des toupies » (p. 599, repris littéralement en 1874, p. 197).

Mais il devient omphalopsychique<sup>22</sup> dans : « je contemplais en souriant s'éleva de mon nombril la tige verte d'où devait éclore le dieu nouveau » (p. 448). En 1874 : « sur le nombril du dieu, une tige de lotus a poussé : et, dans son calice, paraît un autre dieu à trois visages » (p. 123).

Une dernière acception se trouve dans la correspondance. Pour les hommes, le nombril divise, platoniquement, le corps en deux : « nous ne devons entrer dans la vie réelle que jusqu'au nombril. Laissons le mouvement dans la région des jambes » (*Corr.*, 1857, p. 46).

Ces trois acceptions correspondent à trois formes textuelles différentes. La première lexicalisé fort obliquement un thème propre à la correspondance, c'est le thème de la poitrine féminine<sup>23</sup>. La seconde est une lexicalisation partielle du motif de la Genèse monstrueuse, récurrent dans les romans orientalisants (la *Tentation* et *Salammbô*). La troisième acception ne correspond pas non plus à un thème, si l'on convient qu'il faut au moins une récurrence pour faire thème<sup>24</sup>. En somme, le mot *nombril* est une lexicalisation partielle d'un thème dans la correspondance, d'un motif dans deux romans, et d'un topos isolé dans la correspondance.

Résumons. Un lexème peut ne lexicaliser aucun thème ; par exemple le mot *thème* ne correspond à aucun thème dans le corpus romanesque que nous avons étudié. Mais il peut aussi en lexicaliser plusieurs. Enfin, son lien avec le palier thématique est relatif à un discours (littéraire, médical, etc.), un genre, et un corpus.

Par suite, le rapport au corpus n'est pas le même pour l'analyse lexicale et l'analyse thématique. Le corpus doit être maximal pour une étude lexicographique qui entend mettre en évidence toutes les possibilités de la langue. Dans l'analyse thématique, il doit être restreint à bon escient pour pouvoir caractériser la spécificité des discours et des genres : les thèmes du roman ne sont pas ceux de l'essai ni du poème. Ainsi, en dépouillant un corpus trop étendu qui mêlait des romans et des essais dans la période 1830-1870, nous nous sommes aperçu que les sentiments du roman n'étaient pas ceux de l'essai. Par exemple, le sentiment *de fraternité*, récurrent dans les ouvrages de Leroux, et celui *d'équité* chez Proudhon, n'ont pas été relevés dans les romans, à l'exception confirmatrice des *Misérables*, qui alterne des chapitres romanesques et d'autres qui relèvent du genre de l'essai.

À supposer même que le même mot se rencontre dans des genres différents, rien n'assure qu'il se rapporte aux mêmes thèmes : *amour* se rencontrera certes en poésie et dans le roman, mais le thème de l'Amour diffère pourtant avec ces genres. Il n'a pas la même molécule sémique, ni les mêmes lexicalisations, ni les mêmes antonymes — et l'on remarque que l'Ambition n'est pas un thème en poésie, alors que dans le roman elle est fréquemment couplée avec l'Amour (*cf. supra* la contribution de Erlich).

Tout cela éloigne la thématique de la lexicologie, et *a fortiori* de la lexicographie. En particulier, les méthodes lexicographiques visant à l'élargissement du corpus sont impropres pour l'analyse thématique. Les éléments encyclopédiques qui se mêlent souvent aux définitions seraient alors oiseux. La structure de l'article de dictionnaire ne conviendrait pas à un répertoire thématique : la distinction des acceptions et leur hiérarchisation reflète une ontologie ; en outre, les dictionnaires, par leur fonction même, ont un rôle normatif, alors que la thématique ne peut être que descriptive.

Pour progresser, la thématique doit donc dépasser l'analyse lexicale. Elle en est cependant tributaire, notamment pour des raisons de fait qui tiennent à l'état de l'art. D'une part, seule l'interrogation de banques informatisées permet de vérifier certaines hypothèses, et de recueillir des faits là où personne n'aurait l'idée d'en chercher<sup>25</sup>. D'autre part, les logiciels d'interrogation acceptent pour unité la chaîne de caractères, simple signifiant, et ne constatent que des cooccurrences de signifiants qui peuvent être soumises à traitement statistique.

L'enjeu consiste à passer de ce *zéro meaning* à l'analyse thématique, à pallier l'absence de « données sémantiques » en tirant profit de la théorie sémantique.

Cette situation est sans doute temporaire, et l'évolution des architectures connexionnistes laisse prévoir des logiciels qui permettront de faire émerger, par apprentissage et discrétisation sur de grands corpus, des faisceaux de cooccurrences spécifiques. Cela permettra de guider la formulation des hypothèses, mais sans modifier pour autant les conditions herméneutiques que nous avons soulignées.

Si l'on veut préciser encore le rapport de l'analyse lexicale à l'analyse thématique, il faut préciser que le mot à partir duquel commence la recherche n'en est pas l'objet, à la différence d'un mot-vedette qui ferait l'objet d'une recherche lexicographique. On va certes chercher, en utilisant les moyens d'assistance informatisés, d'autres mots et expressions qui sont cooccurrents. Une fois interprétés, les cooccurrents pour lesquels on aura identifié une relation sémantique seront considérés comme des *corrélats*, c'est-à-dire comme des lexicalisations complémentaires de la même molécule sémique.

Le réseau des corrélats relie les manifestations lexicales du thème. Mais il faut pouvoir discerner le(s) meilleur(s) point d'entrée(s) dans ce réseau. La « vedette » n'est alors qu'un mot d'entrée, choisi pour sa fréquence, et dans l'hypothèse qu'il présente une lexicalisation synthétique du thème que l'on cherche à décrire.

## 2. *Qu'est-ce qu'un sentiment ?*

En analysant le champ lexical des sentiments, nous ne postulons pas que ce champ soit uniforme, ni qu'il constitue une unité de langue. Il contient sans doute plusieurs taxèmes. Il ne constitue pas un domaine délimité par l'incidence d'une pratique sociale. Il s'agit donc d'un regroupement *ad hoc*, convoqué par la pratique descriptive en cours.

En dénombrant les noms de sentiments, nous ne prétendons pas compter autant de thèmes. Certains peuvent lexicaliser le même thème (*pitié, compassion*), d'autres des thèmes différents : par exemple, le mot *sentiment* signifie par défaut 'amour' (cf. *Le Rouge et le Noir*, p. 406 ; *Eugénie Grandet*, p. 57) et prend pour antonymes : *mariage* (*Nucingen*, p. 627) ; *malheur* (*Le Rouge et le Noir*, p. 407) ; *intérêt* (*La Duchesse de Langeais*, p. 249).

Pour mettre en évidence les difficultés, retraçons rapidement les étapes d'une recherche lexicale préparatoire à certaines contributions de ce volume. Il s'agissait de faire manuellement — c'est-à-dire intellectuellement — un inventaire des sentiments dans un sous-corpus constitué de 138 romans, soit un par an dans l'intervalle chronologique choisi (à l'exception de deux années où la

banque ne comptait aucun roman). Nous avons pour cela sélectionné les contextes de *sentiment* et de *sentiment de* ou *d'*, dans l'étendue d'une phrase.

La liste ainsi obtenue, détaillée ci-dessous, a été ensuite croisée sur elle-même : en appelant tous les passages qui contenaient au moins deux membres de cette liste, nous avons pu sélectionner un sous-corpus à haute teneur sentimentale et éviter pour l'essentiel les ambiguïtés qui peuvent s'attacher aux occurrences isolées<sup>26</sup>.

#### a) Réserves d'inventaire

Au cours de cette enrichissante lecture, nous avons rencontré deux problèmes classiques. Celui de la polysémie : *sentiment*, qui désignait d'abord la faculté de sentir, qu'il s'agisse de la sensibilité physique ou de la conscience, est venu à désigner aussi, au cours du XVIII<sup>e</sup> siècle, les affections et passions de l'âme. Cette évolution a sans doute accompagné la conquête de l'intériorité. Mais dans notre corpus, la première acception perdure, et l'on relève par exemple *le sentiment de son costume*. Elle se rencontre jusqu'au premier tiers de ce siècle, notamment chez des auteurs conservateurs en politique (Barrés) et qui se plaisent à quelques archaïsmes, innocents moyens de retrouver le bon vieux temps.

Certains sentiments n'ont pas de lexicalisation synthétique, et sont manifestés par des lexies complexes. Dans le meilleur des cas, elles sont récurrentes, mais parfois plusieurs sont employées de façon apparemment équivalente. On pourrait alors parler de sentiments sans nom, comme le sentiment du beau, que Balzac nomme *sentiment inexprimable* (*Le Colonel Chabert*, p. 48) ; ou le sentiment maternel, dit aussi *de la maternité* (*le plus saint de tous*, Dumas, *Monte Cristo*, H, p. 254 ; *tendresse des tripes*, Groult, p. 883), *d'être mère* ; *de mère* (Rops, p. 597).

Il s'agit, plus techniquement, de thèmes sans lexicalisation privilégiée. Ces sentiments sans nom mériteraient à eux seuls une étude particulière, dont la monographie de Danielle Bouverot (*cf. supra*) permet de mesurer la difficulté et l'intérêt. Ils engagent à souligner encore les limites d'une approche purement lexicale. D'autant plus que les méthodes d'interrogation et les calculs statistiques diffèrent selon qu'il s'agit de lexicalisations synthétiques, de syntagmes intégrés, ou de syntagmes non intégrés. Or il faudrait, pour l'analyse sémantique, intégrer les résultats de ces traitements divers dans un protocole commun.

Faut-il considérer comme équivalents les parasyonymes comme *espoir*, *espérance* ? Pour les homologuer, il convient de chercher s'ils se trouvent en relation de collocation. Si ce n'est pas le cas, ils sont sémantiquement assez proches pour être substituables en contexte. Cela n'entraîne cependant pas qu'ils aient les mêmes corrélats ; *espérance*, par exemple, convient mieux à des

contextes religieux. En somme, les parasynonymes ont des distributions complémentaires. Cela explique que leurs collocations soient rares ou nulles.

#### *b) Anatomie des corps étrangers*

Les relevés font apparaître des contenus qui ne sont pas ordinairement classés comme des sentiments. Par exemple, 'dignité' avait été chassé d'une première liste par un « sentiment linguistique » majoritaire dans un groupe de travail. Cependant, on trouve « un sentiment de dignité personnelle et d'orgueil » (Maupassant, p. 177). On peut conclure que localement 'dignité personnelle' entre dans la classe des sentiments.

Mais comment interpréter des coordinations comme « de fureur et de vengeance » (France, p. 333), « d'autre sentiment que l'orgueil et la force » (Tharaud, p. 324), « de malaise et d'aversion » (Aymé, p. 478), « de colère et de revendication » (Rops, p. 537), « de sérénité, de puissance, de liberté » (Martin du Gard, p. 691), « de regret, de mortification » (id., p. 693). Il est vraisemblable que coexistent ici des deux acceptions de *sentiment* que nous relevions plus haut, par une sorte d'antanaclase diachronique qui apparaît notamment dans : « sentiment de joie inconnue et d'élection merveilleuse » (Gracq, p. 812), ou « sentiment de permanent dégoût, de permanente impuissance et de permanente décomposition » (Simon, p. 929). Dans ces cas là, plutôt que de considérer le sentiment inattendu comme le corrélat d'un autre, mieux vaut admettre que l'inventaire des sentiments n'est aucunement arrêté, et que les variations contextuelles reflètent des normes individuelles propres à des auteurs particuliers.

Ici encore, il faut tenir compte du fait que le syntagme est la zone de localité qui facilite le plus la propagation des traits sémantiques. La parataxe notamment instaure des relations d'équivalence (cf. « l'amour, la tendresse, la fidélité », Abellio, p. 746), de même que la comparaison (cf. « sentiment d'étonnement comparable à l'angoisse », Duhamel, p. 704), la gradation (ex. : « sentiment de responsabilité, voire de culpabilité », Martin du Gard, p. 674 ; « sentiment d'admiration, presque d'envie », Green, p. 786).

On ne saurait oublier non plus que le mélange des sentiments fait partie de la psychologie romanesque, et l'on trouve des énumérations comme : « amitié, sentiment, tentation, oubli, silence, érotisme » (Sabatier, p. 915), qui semblent, plutôt que des coq-à-l'âne poétiques, des romans miniature.

#### *c) Classement proposé*

Les sentiments ont été classés selon deux critères. D'une part, leur structure actantielle : ego passif, réflexif, actif sur un ou plusieurs congénères ou objets ; d'autre part, le taxème où ils sont indexés. Le premier critère domine le premier :



les rubriques finales comprennent des sentiments relationnels. La présentation ci-dessous est grossière, car il y manque une classification croisée :

I. *Sentiments antiques* :

1. existentiels : *angoisse, anxiété, de culpabilité, égarement, ennui, espoir, étonnement, de la fatalité, impuissance, incertitude, inquiétude, insouciance, mélancolie, nostalgie, résignation, de sécurité, de (sa, la) solitude, stupeur, trouble*. Les aspectualisés sont anticipateurs : *anxiété, espoir, inquiétude, insouciance* ; ou rétrospectifs : *nostalgie, résignation*.

2. thymiques : *aise, allégresse, amertume, ardeur, bonheur, colère, dégoût, dépit, désespoir, détresse, enthousiasme, épouvante, exaltation, extase, félicité, fureur, gaieté, horreur, indifférence, indignation, joie, lassitude, malaise, malheur, mécontentement, panique, peur, plaisir, plénitude, répugnance, répulsion, satisfaction, sérénité, surprise, tristesse*. On peut les diviser en euphoriques *aise, satisfaction, gaieté, joie, allégresse, bonheur, félicité, extase* ; et dysphoriques : *amertume, malaise, mécontentement, malheur, désespoir, détresse*.

II. 3. *Sentiments relationnels* :

*admiration, affection, amitié, amour, amour-propre, anti-sympathie, antipathie, aversion, bienveillance, camaraderie, complicité, confiance, crainte, cruauté, curiosité, défiance, délicatesse, effroi, embarras, envie, estime, exaspération, frustration, gêne, gratitude, haine, honte, humiliation, sentiment d'infériorité, indulgence, ingratitude, ironie, orgueil, pudeur, rancune, reconnaissance, respect, sentiment (par défaut = amour), soumission, de (la, sa) supériorité, sympathie, tact, timidité, vanité, vénération*. Ds se divisent à leur tour en passifs et actifs, euphoriques et dysphoriques.

III. 3.1. *Sentiments de filiation* : *amour filial, amour maternel, de la famille, sentiment filial, sentiment maternel, de la maternité, de parenté, paternel, de la paternité, de la piété filiale, tendresse*.

3.2. *d'alliance* : *fidélité, jalousie, amour*.

3.3. *sociaux* : *ambition, sentiment colonial, dignité, du devoir, de ses devoirs, de ses droits, de discipline, fraternité, héroïsme, sentiment hiérarchique, honneur, humanité, intérêt, sentiment national, d'oppression, patriotisme, sentiment patriotique, de la propriété, de rébellion, de responsabilité, révolte, de solidarité*.

IV. 3.4. *religieux* : *charité, compassion, sentiment de faute, espérance, ferveur, foi, humilité, piété, sentiment religieux, de religion, remords*.

3.5. *esthétiques* : *de l'art, du beau, de la beauté, de la vraie beauté, sentiment inexplicable, poétique*.

Soit 98 désignations de sentiments, avant réduction des parasyonymes, pour les rubriques I et II ; et 34 pour les rubriques III et IV. Soit 128 en tout. Ce chiffre dépend sans doute pour une part du corpus traité, mais nous avons relevé un effet de saturation : au fil de la lecture, les sentiments nouveaux deviennent de plus en plus rares.

La distinction déterminante est certainement celle de l'évaluation positive ou négative. Ces dimensions sémantiques divisent le champ des sentiments, ce que confirme par d'autres moyens l'analyse factorielle conduite par Brunet (*cf.* la fig. 5 de son étude, dont la partie gauche, dysphorique, s'oppose à la partie droite). On vérifie subsidiairement la distinction entre les sentiments on tiques (dans le bas du tableau) et les sentiments relationnels (dans le haut). Plus généralement, il est remarquable que les résultats de notre analyse sémantique et ceux de son analyse statistique concordent globalement, bien que ces études aient été conduites séparément. Cela confirme que les proximités statistiques sont bien fondées sur des proximités sémantiques.

### 3. Cooccurrence et corrélation

Le passage de l'analyse lexicale à l'analyse thématique conduit de signes non interprétés à des unités sémantiques qui résultent d'un parcours interprétatif. Il se concrétise par le passage des cooccurents aux corrélats.

#### a) La cooccurrence intéresse les signifiants

Les cooccurents ne sont que des signifiants — certains disent aussi *formes*, ou *chaînes de caractères* — associées statistiquement par la méthode de l'écart réduit ou de l'écart hypergéométrique<sup>27</sup>. Il convient de distinguer plusieurs zones de cooccurrence : immédiate, proche, et large.

Contrairement à l'intuition, les cooccurents immédiats ne sont pas toujours d'un grand intérêt pour l'étude thématique. Si l'on interroge la banque textuelle à partir de mots isolés, le bruit le plus sensible est dû aux constituants des expressions phraséologiques reviennent avec insistance. On peut distinguer ici entre les expressions complètement intégrées comme *avoir à cœur*, ou *savoir par cœur*, qui fausseront par exemple les résultats d'une interrogation sur les parties du corps ; et les phénomènes de cooccurrence restreinte, propres à des syntagmes moins intégrés (comme *éprouver un sentiment*), et que Mel'cuk a étudié sous le nom de *fonctions lexicales*. Par exemple, dans son analyse de *joie* (1981, p. 25) et de *désespoir* (p. 22), il relève la fonction *Oper 1*, qui prend la même valeur *éprouver*. Il est clair que ce cooccurrent ne permet pas de différencier la joie du désespoir. En règle générale, on relève dans les syntagmes en voie de figement une désémantisation des constituants, dont les lexies se confondent en une seule. Aussi la cooccurrence de ces constituants n'a-t-elle qu'une pertinence faible ou nulle.

C'est dans le contexte indépendant des figements phraséologiques que l'on trouve les cooccurents les plus pertinents. Par extrapolation, peut-on faire l'hypothèse que le contexte proche soit plus sensible aux nonnes sociolectales que le contexte large, de l'ordre du paragraphe ? La question est délicate, car l'on trouve des usages individuels des topoï aussi bien au palier du syntagme qu'au palier du paragraphe.

Retenons cependant que le paragraphe est l'homologue (non l'analogue) sur le plan de l'expression de la période sur le plan du contenu. C'est ainsi dans l'espace du paragraphe que l'on trouve l'essentiel des cooccurents correspondant au thème recherché.

Selon l'étendue du corpus et les objectifs de la description, on peut faire varier les seuils de pertinence quantitative retenus. Les listes de cooccurents ainsi établies contiennent de bons candidats à la qualification sémantique.

*b) La corrélation intéresse les signifiés*

Les unités de cooccurrence sont des chaînes de caractères, et ne correspondent pas ternie à terme aux unités de corrélation qui sont des lexies. Ces dernières connaissent deux zones de localité qui favorisent de façon décroissante la propagation des traits sémantiques : le syntagme et la période (*cf.* l'auteur, 1994, ch. V).

La qualification des cooccurrents est cruciale, car elle permet le passage du quantitatif (les cooccurrents) au qualitatif (les corrélatés). Elle se règle sur les principes herméneutiques que le global détermine le local, et que l'hypothèse gouverne l'objectivation. Ainsi, les cooccurrents ne sont élevés à la dignité de corrélatés que s'il est possible d'établir une relation d'isotopie ou de paratopie avec d'autres cooccurrents. Par exemple, parmi les cooccurrents à *ennui*, *dimanche* et *araignée* se sélectionnent mutuellement, dans le contexte d'*inaction*. Ils lexicalisent un des composants du thème recherché, et c'est à ce titre qu'ils sont qualifiés. Comme les corrélatés sont des signifiés, les classes morphologiques qui les manifestent n'ont pas d'importance à ce niveau d'analyse, et l'on sait que le même sème peut être lexicalisé par un préfixe, un nom, un adjectif, un verbe, etc.<sup>28</sup>.

L'hypothèse qui fonde la transformation de la cooccurrence en corrélation est celle-ci : le contexte proche est structuré par des isotopies qui marquent l'appartenance à un même fond sémantique, ou des paratopies — qui marquent l'appartenance à la même forme sémantique.

Tout corrélat peut être relié par une relation casuelle avec un autre corrélat, ou partager au moins un sème avec au moins un autre corrélat. Ainsi, un corrélat X peut partager le sème *a* avec le corrélat Y, et le sème *b*, avec le corrélat Z, etc. ; cependant Z partage le sème *c* avec X ou W, etc. Le réseau des relations casuelles ou d'équivalence partielle ainsi dessiné constitue le thème. Si ces relations ne peuvent être établies, l'hypothèse initiale doit être révisée, car l'on n'a pas choisi un bon point d'entrée, et l'on a conduit l'interrogation à partir d'un sémème faiblement corrélaté au thème recherché ; soit encore on n'a pas affaire à un thème — ou du moins à un thème stabilisé dans le corpus choisi.

La recherche des cooccurrents et la promotion de certains au rang de corrélat reprend, à un palier inférieur, mais à une échelle quantitative nouvelle, l'antique technique herméneutique des passages parallèles, déjà théorisée avant notre ère par Hillel l'ancien.

Bref, résumons les principales étapes d'une recherche thématique assistée : 1. Choix des hypothèses, en fonction de l'objectif général de la recherche. Une préanalyse statistique peut guider la recherche d'hypothèse, mais la fréquentation préalable du corpus reste indispensable pour guider les intuitions.

2. Recherche de cooccurrents par la méthode statistique des écarts réduit ou hypergéométrique.

3. Transformation interprétative des cooccurrents en corrélats, et constitution des réseaux thématiques. Cette étape est facilitée si l'on a pratiqué une interrogation simultanée sur plusieurs cooccurrents (*cf. supra* Bourion, I. 2)<sup>29</sup>.

4. Validation des résultats : par croisement de l'analyse thématique avec l'analyse d'autres composantes du même corpus ; par test sur un corpus de contrôle ; par confrontation avec d'autres recherches thématiques.

### **m. Problèmes épistémologiques : de la philologie à l'herméneutique**

Tout protocole méthodologique se règle sur des principes épistémologiques, explicites ou non. Aussi, plutôt que de détailler ces protocoles eux-mêmes, qui dépendent de l'état de l'art, et doivent se régler sur le type de corpus choisi, comme sur le type d'hypothèse à valider, il nous paraît plus utile de préciser leurs principes déontologiques.

La philologie et l'herméneutique sont les deux disciplines concernées au premier chef. Une sémantique du texte doit les concilier, car elle ne peut se satisfaire de la séparation entre la lettre et l'esprit, quand bien même la philologie a connu naguère une dérive positiviste, et l'herméneutique une involution spéculative.

#### *1. Les choix méthodologiques et la philologie*

Même si l'on convient de dépasser par l'étude des corrélations le stade de la cooccurrence lexicale, comment formuler l'hypothèse initiale, comment choisir le ou les mots qui permettent d'entrer dans le réseau thématique ? Une recommandation générale suggère de choisir des mots du corpus. J'emprunte par exemple les résultats préliminaires d'une étude restée inédite de R. Choné. Voulant étudier le sentiment d'infériorité dans le roman du XIX<sup>e</sup> siècle, elle n'en a pas trouvé d'attestation ; en revanche, le sentiment *d'écrasement* n'est pas rare, bien qu'il ne figure pas dans les dictionnaires thématiques, et soit étranger à la doxa contemporaine en matière de sentiments. Si bien que l'hypothèse qui a présidé à l'interrogation doit être reformulée, ou abandonnée. Le sentiment d'écrasement affecte des personnages situés au bas de l'échelle sociale, et résulte de l'oppression et de l'injustice qui s'exerce sur eux ; il est décrit pour ainsi dire de l'extérieur. En revanche le sentiment d'infériorité est une création contemporaine, qui affecte des personnages vus de l'intérieur, et n'a pas de cause sociale bien identifiée.

Faut-il, peut-on rechercher les thèmes contemporains dans les textes des siècles passés ? Cette question dépasse la distinction entre l'approche onomasiologique qui part de signifiés ou de concepts pour en rechercher l'expression, et l'approche sémasiologique qui prend le chemin inverse. Elle engage aussi bien la philologie que l'herméneutique, et nous allons l'illustrer par un parallèle.

Les deux contributions de Brunet et Béhar & Bernard se placent explicitement dans le cadre d'une *histoire des mentalités*.

On a vu comment Béhar et Bernard compilent neuf listes de sentiments « à la manière d'un sondage d'opinion »<sup>30</sup>. La liste résultante, qui ne retient que les termes figurant déjà dans le thésaurus de la BDHL, soit 106, reste hétérogène, puisqu'on y trouve l'alcoolisme, le culte, la dépression nerveuse, le mysticisme, la soif, la toxicomanie. La garantie d'une attestation majoritaire dans les listes de départ réduit ce nombre à une liste homogène de vingt. Cela entraîne l'élimination de quatre-vingts noms de sentiments attestés par ailleurs dans le corpus romanesque de langue française de 1830 à 1970. Ici apparaît le caractère normatif d'un thésaurus, qui, pour une étude de thème conduit à ne sélectionner que ce qui peut répondre à une indexation préalable. Les auteurs soulignent bien dans leur conclusion que l'étude de la littérature « a affaire à des opinions », mais « en se plaçant du côté du lecteur, se range du côté de l'histoire des mentalités ». Il s'agit bien entendu du lecteur d'aujourd'hui. La question du lecteur d'hier et des mentalités du siècle dernier ne se pose pas dans ce cadre.

Pour sa part, Brunet ne prend en considération que le corpus Frantext et part de notre relevé indicatif. Il fixe comme seuil de pertinence un seuil de fréquence<sup>31</sup>. Il trouve statistiquement 124 mots-cooccurrents du mot *sentiment*. Ils ne sont pas nécessairement des noms de sentiment, mais des formes comme *éprouvait*, *éprouvais* ou *visage*, *cri*, *vie*. Dans l'environnement de ces 124 entrées, il trouve 164 noms de sentiment, et le mot *sentiment* lui-même. Il accroît ainsi d'un quart la liste indicative initiale, établie à partir du tiers du corpus. Et à partir de cette liste étendue, il construit des données nouvelles sur la structure du champ des sentiments, son histoire et son évolution.

Si ces deux études partageaient les mêmes objectifs, on pourrait par généralisation opposer deux stratégies. La première semble accueillante, car elle multiplie les sources. Mais elle reste normative, dans la mesure où elle ne retient que ce qu'elles ont de commun, et conduit à une restriction des hypothèses. La seconde, simplement descriptive, conduit à une extension des hypothèses et à la constitution de données riches.

Cette symétrie reste fautive, car la sélection des œuvres par indexation et leur exploration assistée n'obéissent pas aux mêmes objectifs. Une question reste posée : comment choisir les thèmes qui servent à indexer les œuvres ? Ou bien

l'on reflète la doxa contemporaine en la matière, ou bien l'on cherche à restituer les thèmes importants à l'époque — tâche qu'un logiciel d'analyse thématique pourrait précisément assister.

La question philologique reste posée. L'exploration d'un corpus exige des précautions — à l'image de tout travail historique. La première est le respect des limites du corpus. Si l'on convient que tout relevé, pour être pertinent, doit correspondre au corpus, c'est-à-dire à des coordonnées précises dans le temps et dans l'espace textuel, les sources annexes sont par là-même oiseuses. Enfin, un corpus inanalysé ne peut en éclairer un autre que si l'on démontre que l'on peut les confondre.

Nous n'avons pas l'illusion pour autant de nous transporter par la thématique dans la mentalité de naguère. L'entreprise même d'une sémantique des textes et les moyens d'assistance informatique sont bien de notre temps. Du moins, pour éviter le narcissisme des recherches qui ne retiennent que ce qui conforte leurs hypothèses, pouvons-nous soumettre à révision les nôtres, pour accueillir par exemples des sentiments sans nom, mêlés ou disparus. Sans quoi la thématique, comme la plus belle fille du monde selon Proust, ne donnerait que ce que l'on croit recevoir.

#### IV. Directions de recherche

##### 1. Ouvertures méthodologiques

Ce n'est pas le lieu ici de revenir sur les problèmes classiques de l'analyse statistique et de la consultation de banques de données, comme la polysémie, le bruit, et le silence.

Il faut cependant souligner que l'amélioration des outils logiciels et l'enrichissement des corpus conditionnent en pratique l'analyse sémantique. Les résultats magistraux présentés ici-même par Brunet n'auraient pas été obtenus sans l'usage des logiciels qu'il a élaborés. Plus généralement, retenons que le quantitatif et le qualitatif ne s'opposent aucunement : seule une analyse qualitative peut rendre significatifs des phénomènes quantitatifs remarquables, et interpréter leurs variations. La sémantique des textes est d'ailleurs la partie de la linguistique la plus utile pour cela, dans la mesure aussi où elle permet de relativiser les phénomènes quantitatifs. Par exemple, Erlich ici-même a figuré les *rafales* thématiques de l'Ambition et de l'Amour dans *Le Père Goriot*. Mais les silences lexicaux, comme il le souligne, ne sont pas pour autant des silences thématiques : ces thèmes, pour n'être pas nommés, sont intégrés ailleurs et d'une autre manière à des structures dialectiques. L'on pourrait par paradoxe soutenir que dans ce roman l'Ambition et l'Amour sont partout à l'œuvre quand ils ne sont pas

nommés<sup>32</sup>, et culminent à la mort du personnage éponyme, où leurs lexicalisations privilégiées font pourtant défaut.

### 2. Pour une thématique historique et comparée

Les thèmes, en tant que formes sémantiques, sont liés à l'histoire culturelle, dont ils sont une expression privilégiée, surtout quand leur récurrence les promeut au rang de topoï. L'exemple des sentiments est révélateur.

Les dates d'apparition de nouveaux sentiments ne manquent pas d'intérêt, qu'ils soient liés à l'histoire politique, comme le *sentiment patriotique* qui apparaît en 1912 et reparaît en 1940, ou à l'histoire de la métapsychologie, comme ce *sentiment de frustration* qui apparaît en 1960.

L'évolution diachronique des sentiments, et de leur évaluation trouve un excellent exemple dans la contribution de F. Surdel : à partir des années 1870, la pitié devient un sentiment suspect. Nous ne nous chargerons pas d'accuser les Communards ou leurs adversaires. Mais il faut souligner que les attaques venues des écrivains catholiques, de Bloy à Mauriac, ont accompagné sa fin, si bien que ce sentiment, perdant ses derniers soutiens, s'est absenté du roman, et sans doute de la vie. L'humanitaire a pris sa place, dans de tout autres contextes.

Plus généralement, une histoire globale du monde sentimental devient possible. Par exemple, le tableau que présente Brunet (figure 9) est éloquent : on assiste à un assombrissement continu des sentiments dans le roman depuis 1830.

Enfin, et là encore les résultats de Brunet sont pleins d'enseignement, une thématique comparée selon les genres permettrait de spécifier les particularités de chacun, et de restituer l'articulation générale du discours littéraire.

### 3. Des formes culturelles au vécu : les sentiments et les affects

Nous avons défini les thèmes comme des unités sémantiques, mais ils correspondent sans doute à des représentations socialisées. Comment sont-elles liées à leur expression linguistique ?

Gardons l'exemple des sentiments. Une première voie de réponse consiste à rechercher comment les sentiments éprouvés trouvent leur expression linguistique. Mais le monde sentimental n'est-il pas lié à sa représentation ? Pour les écrivains peut-être, et pour leurs lecteurs assidus. Hors même de la culture littéraire, les sentiments ne sont-ils pas inséparables de leur verbalisation ?

La demi-douzaine d'émotions de base répertoriées par les éthologistes expliquent peut-être l'existence des affects, mais non la variété des sentiments, et encore moins le caractère historique et culturel de leur expression. L'acédie érémitique et la dépression nerveuse ont peut-être les mêmes substrats biochimiques, mais non la même définition. Les sentiments, en tant que formes

culturelles, varient selon les époques et les sociétés ; ils ne sont d'ailleurs pas nécessairement individualisés comme tels, car ils supposent un sujet autonome et doué d'une vie intérieure. Et chacun sait combien d'une langue à l'autre les sentiments sont difficiles à traduire, même dans des sociétés très proches culturellement.

Bref, une culture transforme les affects en sentiments. Un homme normal, dit Sacks, est un homme qui peut se raconter son histoire. Mais ce récit demande une verbalisation, et qui sait si notre vécu, du moins ce qui en demeure hors de l'instant, n'en dépend pas ?

#### NOTES

1. Trousson (1981, p. 12 sq.) relève par exemple ces définitions : « point de rencontre d'un esprit créateur et d'une matière littéraire ou simplement humaine » (Cl. Pichois et A.-M. Rousseau) ; « un événement ou une situation infantiles » (J.-P. Weber) ; « réseau organisé d'obsessions » (Barthes). A quoi Collot (1988, p. 80) ajoute : « le thème [...] n'est rien d'autre que la coloration affective de toute expérience humaine » (Dobrovsky) et renchérit en soulignant que le thème « exprime la relation d'un sujet au monde sensible » (p. 81).
2. La définition du TLF témoigne du poids de l'anthropologie en matière de thématique : « Unité du contenu (d'un discours, d'un texte ou d'une œuvre littéraire) correspondant à une constante du symbolique ou de l'imaginaire. » Or un thème peut être spécifique à un texte ou à une groupe de textes, et comment faire fond sur ces constantes éminemment discutables ? La critique thématique a d'ailleurs mis en évidence des « thèmes personnels », et l'étude de Richard sur Mallarmé est l'exemple éminent d'une analyse idiographique.
3. Todorov définissait le thème comme « une catégorie sémantique qui peut être présente tout au long du texte ou même dans l'ensemble de la littérature (le thème de la mort, par exemple) » (in Ducrot et Todorov, 1972, p. 283).
4. Nous nous écartons évidemment de la problématique réaliste ordinaire en philosophie du langage comme de la linguistique qu'elle influence (cf. Reinhardt et Kuno), pour laquelle le thème est le réfèrent de l'énoncé.
5. Pour une typologie des thèmes génériques, cf. l'auteur, 1989, ch. IV.
6. Nous avons proposé des exemples de thèmes sans lexicalisation privilégiée chez Zola (1989, H, ch. II), et chez Eluard (1991, ch. IV").
7. Nous appelons ainsi l'adaptation des graphes conceptuels définis par Sowa (1984) aux besoins de la description linguistique ; sur les avantages et les limites de cette sorte de représentation, cf. l'auteur, 1989, ch. V.
8. Ces deux sèmes aspectuels ont une grande importance dans la représentation du vécu, dans la mesure où les formes du vécu correspondent à des types de temporalisation.
9. On dit encore dans les ateliers *ça va comme un lundi*, pour signifier *ça ira bien si on n'avait pas encore une semaine à tirer*.
10. Un thème peut avoir une lexicalisation privilégiée (ex. *ambition*), ou plusieurs (*pitié*, *commisération*, *compassion*). Il peut s'agir d'une lexie (*amour paternel*) ou n'avoir pas de nom retenu par l'usage (*sentiment du beau*, *amour de l'art*). La parasynonymie des lexicalisations privilégiées se manifeste par le fait qu'ils n'apparaissent pas dans les mêmes contextes (ex. *spleen* apparaît rarement dans le contexte de *ennui*, non parce qu'il en est trop éloigné, mais vraisemblablement parce qu'il est trop proche (cf. *supra* Gorcy).



11. Le foyer énonciatif qui préside à la description est alors celui d'Emma. On sait que Flaubert passe sans doute à bon droit pour le créateur du style indirect libre, précisément dans *Madame Bovary*.  
Cet exemple vaut pour les graphes A et N. Le contenu du graphe N se trouve détaillé dans les phrases qui suivent, et que résume cette interrogation : « Un homme, au contraire, ne devait-il pas tout connaître, exceller en des activités multiples, vous initier aux énergies de la passion, aux raffinements de la vie, à tous les mystères ? »
12. Une molécule sémique a le statut de type. A ses nœuds peuvent être attachés des sèmes ou des axes sémantiques (*attributs* dans la terminologie de l'intelligence artificielle) susceptibles de prendre diverses valeurs selon les occurrences.
13. Par exemple, dans *Nous montâmes dans un fiacre. La promenade fut enivrante, promenade* lexicalisé un sème de *fiacre*.  
Nous ne postulons pas de traitement déterministe, tel que la propagation d'activation irait du mot vers ceux qui le suivent immédiatement.
14. Le rapport du thème aux lexicalisations de ses attributs est homologue du rapport de l'acteur aux lexicalisations de ses actants. Nous emploierons d'ailleurs la même notation pour le thème et l'acteur : une majuscule à l'initiale.
15. Encore que les « dictionnaires de thèmes » puissent rendre des services, en pédagogie notamment.
16. On peut distinguer trois acceptions du mot *topos*. La plus traditionnelle (depuis la rhétorique d'Aristote) est une forme argumentative stéréotypée (elle a été reprise, avec une extension moindre, par certains pragmaticiens). La seconde, que nous avons utilisée (1987) est un axiome normatif socialisé (comme *Les gascons sont vantards*) qui permet une afférence — en termes logiques, cette proposition complète un enthymème. La troisième, que nous employons ici, désigne une structure thématique stéréotypée, familière en histoire de la littérature : ainsi, le *topos* du *locus amoenus*.
17. Cf. *supra* l'élude de Brunet, et les syntagmes contradictoires relevés par Bourion, comme *épouvané de joie, terroriser d'amour, désir panique*.
18. C'est peut-être là une des raisons des difficultés relevées plus haut par Gérard Gorcy : les noms de sentiments peuvent se trouver en relation antonymique, mais cela ne devient certain que si l'on relève des cooccurrences. Cependant, si les noms de sentiments étudiés relèvent d'un thème, fût-ce en qualité de corrélats privilégiés, on ne peut conclure sans médiation à une antonymie entre les thèmes, car d'autres corrélats des mêmes thèmes pourraient ne pas être opposés.
19. A la différence des précédents, ce complexe topique reste omniprésent. Cheb Khaled expliquait récemment les contradictions de l'Algérie en opposant deux acrostiches : FIS signifierait *Fatma interdiction de sortir*, et FLN *Femme libre la nuit...*
20. Personnel signifie ici *idiolectal*, relevant de nonnes individuelles. Certains auteurs, comme Trousson, estiment que les thèmes personnels ne sont pas à proprement parler des thèmes, réservant le nom de *thèmes* aux *topoi*. C'est là un effet des objectifs de la littérature comparée, qui privilégie la topique. Mais il reste à préciser, au sein d'un même texte, les rapports entre thèmes et *topoi*, ne serait-ce que pour montrer comment sa globalité particulière intègre des éléments généraux.
21. Bédier employait le mot *élément*, et *motif* nous vient de Vessélovski ; il a été repris par Chklovski : « Par motif, j'entends l'unité narrative la plus simple, qui, en forme d'image, répondait aux diverses exigences de l'esprit primitif et de l'observation quotidienne » (d'après Segre, 1988, p. 10). Cette définition conduira Tomachevski et Propp à voir dans les motifs des fonctions narratives.
22. Le nombril eut un rôle pour ainsi dire central dans la controverse sur l'hésychasme — qui dans la Byzance du XVe opposa notamment Barlaam de Seminara à Grégoire Palamas. En effet, les mystiques athonites pratiquaient la prière du cœur pour accéder à la vision de la lumière divine ; or, ils considéraient le nombril comme le « lieu du cœur » sur lequel il fallait se concentrer, d'où le sobriquet d'\**omphalopsychiques* (qui ont l'âme au nombril) dont les affubla le caustique Barlaam.
23. Cf. e.g. cette description de Kuchuk-Hanem (1850) *Corr.*, t. II p. 174 : « C'est une impériale bougresse, tétonneuse, viande. »

24. Cette division platonicienne, encore vivace dans des expressions comme *au-dessous de la ceinture*, constitue bien un topos, mais son occurrence isolée dans notre corpus ne permet pas de l'intégrer à une analyse thématique.
25. Cf. *supra* Béhar et Bernard sur l'abondance des sentiments chez Raymond Roussel, alors que toute la tradition critique insiste sur sa méthode d'écriture qui ne leur laisse aucune place.
26. Quand par exemple une requête sur *amour* sélectionne des occurrences dans des expressions comme *pour l'amour de dieu* ou *les enfants de l'amour*.
27. Prendre en considération l'écart absolu interdit en pratique d'apprécier les contrastes, et en théorie éloigne irrémédiablement la recherche des principes de la linguistique comparée.
28. Même les signes de ponctuation peuvent se trouver resémantisés : par exemple, à certains sentiments sont associés des points d'exclamation, à d'autres des points de suspension...
29. La requête simultanée sur plusieurs cooccurrents (relevant de diverses classes morphologiques) permet de sélectionner un sous-corpus dense et patinant.
30. Cinq dictionnaires dont quatre électroniques, un analogique, un thésaurus d'indexation (BDHL), un relevé dans les *Passions de l'âme* de Descartes, notre relevé dans le corpus de romans de Frantext ; enfin une liste de mots pouvant figurer après *éprouver*.
31. Les seuils de fréquence engagent l'analyse statistique sur la voie de la topique plutôt que la thématique individuelle. Cela est nécessaire pour que ses résultats soient significatifs, mais il faut y reconnaître une contrainte quantitative sur le qualitatif. Les hapax cependant ne signifient pas moins que les autres occurrences, même s'ils signifient autrement.
32. De même plus haut, à propos de l'Ennui dans *Madame Bovary*.

## BIBLIOGRAPHIE

- BOUCHEZ (M.), *L'Ennui de Sénèque à Moravia*, Paris, Bordas, 1973.
- BOURION (É.), *L'Expérience d'extraction des contenus thématiques à l'aide du test de l'écart réduit*, ms, 1994, 5 p. BRUNET (É.), *THIEF*, in *Traitements informatisés de corpus textuels*, éd. par É. Martin, Paris, CNRS-INaLP, Didier Érudition, 1994, pp. 239-261.
- COLLOT (M.), Le Thème selon la critique thématique, *Communications*, n° 47, 1988, pp. 79-92.
- DUCROT (O.), TODOROV (T.), *Dictionnaire encyclopédique des sciences du langage*, Paris, Seuil, 1972.
- LEBART (L.), SALEM (A.), *L'Analyse statistique des données textuelles*, Paris, Dunod, 1988.
- MARTIN (É.), *Reconnaissance de contextes thématiques dans un corpus textuel. Éléments de lexico-sémantique*, Paris, CNRS-INaLF ; Didier Érudition, 1993 (Études de sémantique lexicale).
- MEL'CHUK (L.), IORDANSKAJA (L.), ARBATCHEWSKY-JUMARIE (N.), Un nouveau type de dictionnaire : le dictionnaire explicatif et combinatoire du français contemporain, *Cahiers de lexicologie*, 38-1, 3-34, 1981.
- MEUNIER (J.-G.), Le Traitement et l'analyse informatique des textes, *Actes de VICO Québec*, 1990, pp. 9-18.
- NEF (F.), Sémantique et ontologie II, *Sémiotiques*, n° 2, 1992, pp. 7-18.
- PANOFSKY (E.), *Essais d'icnologie*, Paris, Gallimard, 1967.
- RASTIER (F.), *Sémantique interprétative*, Paris, PUF, 1987.
- RASTIER (F.), *Sens et textualité*, Paris, Hachette, 1989.
- RASTIER (F.), *Sémantique et recherches cognitives*, Paris, PUF, 1991.
- RASTIER (F.), Thématique et génétique, *Poétique*, n° 90, 1992a, pp. 205-228.
- RASTIER (F.), Réalisme sémantique et réalisme esthétique, *TLE*, D° 10, 1992b, pp. 81-119.
- RASTIER (F.), *Sémantique pour l'analyse*, avec la collab. de A. Abeille et M. Cavazza, Paris, Masson, 1994.
- RICHARD (J.-P.), *L'Univers imaginaire de Stéphane Mallarmé*, Paris, Seuil, 1961.

- ROBERT (R.), Jalons pour une étude du traitement des *Topoi* romanesques dans le roman sentimental, in *Le Roman sentimental*, Limoges, Pulim, 1990, pp. 13-27.
- SAGNES (G.), *L'Ennui dans la littérature française de Flaubert à Laforgue*, Paris, Armand Colin, 1969.
- SEGRE (C), Du motif à la fonction, et vice versa, *Communications*, n° 47, 1988, pp. 9-22.
- SOWA (J.), *Conceptual Structures*, New York, Addison-Wesley, 1984.
- TROUSSON (R.), *Thèmes et mythes : questions de méthodes*, Bruxelles, Éditions de l'Université de Bruxelles, 1981.

*NB* : J'ai plaisir à remercier de leurs observations Henri BÉHAR, Danielle BOUVEROT, Pierre LAFON.



