

MONDES LEXICAUX ET TOPOÏ DANS L'APPROCHE "ALCESTE"

Max Reinert

*UMR 5610 Equipe de Recherche en Syntaxe et Sémantique (CNRS,
Université de Toulouse-Le Mirail - reinert@cict.fr)*

L'objectif est de montrer que si tout discours met en jeu un système de "topoï", il organise du même coup la distribution du lexique dans ses énoncés et cela peut être mis en évidence à l'aide d'une analyse statistique.

La première partie est consacrée au développement et l'argumentation de cette proposition. Dans la seconde partie, l'auteur présente rapidement deux exemples d'application dans lesquels, selon lui, les différents « mondes lexicaux » mis en évidence par l'analyse statistique reflète un même système de topoï fondant l'activité énonciative des locuteurs.

Analyse de discours, Lexicométrie, Sémantique, Alceste, Corpus, Statistique, Analyse des Données Textuelles, Mondes lexicaux, Topoï, Catégories.

0. INTRODUCTION

La méthodologie Alceste (1986), au niveau de ses techniques, est une méthodologie de statistique textuelle selon la terminologie de L. Lebart et A. Salem (1994) et se situe en continuité avec l'approche Benzécriste (1973). Son objectif se rapproche davantage de celui de l'analyse de discours que de celui de la lexicométrie. Plus précisément, l'objectif est surtout de rendre compte de l'organisation distributionnelle d'un discours et non de comparer différents textes. De ce point de vue, l'objet visé ici est différent de celui visé par E. Brunet (1993) dont plusieurs de ses études (1988, 1992) sont relatives à la propagation d'un thème particulier dans différentes œuvres d'une époque ou d'un corpus particulier - thème appréhendé à travers les distributions d'un champ lexical particulier. Cela montre, s'il en était besoin, que la statistique textuelle peut permettre d'appréhender de bien des manières ce que signifient les textes... d'où la nécessité de préciser, dans un premier temps, l'objet visé par l'analyse statistique proposée.

Quel objet pour une analyse de discours de type statistique ?

Selon l'Encyclopédie Universalis, le mot *discours* viendrait d'un mot latin *discurrere*, signifiant « courir ça et là ». Il évoque le « chemin hasardeux de la conversation et de l'entretien, avant de renvoyer à toute mise en forme, parlée ou écrite de la pensée ».

Cette évolution du sens du mot *discours* traduit bien l'évolution de tout discours écrit qui commence avec les aléas des intentions premières plus ou moins contradictoires pour se terminer, au moins idéalement, par une composition ordonnée. La problématique des *lieux du discours* ou *topoï* se situe à ce niveau-là et notre principale thèse est que tout discours met en

jeu un système de *topoi* ou *lieux communs* à travers lequel une rationalité s'organise en même temps que quelque chose s'objective, se construit selon différents éclairages.

En tant que notion première (précatégorielle), la notion de *topoi* telle que nous l'utilisons n'est pas définie à partir d'un trait opératoire abstrait permettant de définir avec précision une catégorie de pensée. Nous proposons justement d'approcher cette notion de manière pragmatique à partir de l'analyse statistique des usages discursifs et plus précisément des usages lexicaux. C'est l'objet de la méthodologie "Alceste" (pour *Analyse des Lexèmes Cooccurents dans les Enoncés Simples d'un Texte*).

Notre objectif est donc de montrer que si tout discours met en jeu un système de "topoi", il organise du même coup la distribution du lexique dans ses énoncés et cela peut être mis en évidence à l'aide d'une analyse statistique.

On développera d'abord cette notion de *topoi* en partant de la notion de distribution.

On présentera ensuite de manière succincte deux exemples d'analyse mettant en évidence, selon nous, un même jeu de *topoi*.

1. DE LA NOTION DE DISTRIBUTION À CELLE DE TOPOÏ

Si Harris a nettement orienté l'approche distributionnelle vers l'analyse de discours, sa notion d'équivalence distributionnelle reste dépendante d'une approche syntaxique et transformationnelle. Dès 1962, Benzécri propose une autre interprétation de cette notion, une interprétation statistique. Sa métrique du c2 est définie de manière à satisfaire au principe « *d'équivalence distributionnelle* » spécifiquement pour l'étude de « *données linguistiques* ».

Ce courant Benzécriste s'est mêlé à celui de la lexicométrie et du traitement des enquêtes et les enjeux de la « statistique textuelle » (Lebart, Salem) sont encore très ouverts et discutés. Nos recherches sur l'approche statistique "Alceste" renouent d'une certaine manière avec les premiers objectifs de Harris en analyse de discours en ce sens que nous cherchons à analyser les distributions dans les limites d'un seul discours mais avec une opérationnalisation statistique particulière de cette notion de distribution.

Au lieu de considérer la distribution comme une combinaison ordonnée d'éléments, nous la considérons comme un simple indicateur des « lieux d'énonciation ». Par exemple, dans ces vers de Reverdy (Extrait de "Courte vie" du recueil "Les sources du vent") « *L'heure pleine est passée sur une autre qui sonne. Les pas des voyageurs courrent déjà loin* » la simple présence de mots comme *heure, passée, pas, voyageurs, courrent, loin*, dessinent dans ce cas une isotopie du passage indépendamment de la manière dont sont reliés ces mots au niveau syntaxique. Mais antérieurement à cet acte de lecture notre hypothèse consiste à dire que l'énoncé porte en lui, par le choix même du vocabulaire, cette possibilité isotopique d'un lieu qui ne doit pas être interprété comme un thème (il pourra le devenir pour un lecteur particulier).

Ce choix agit sur le lecteur comme une force perlocutoire et le place dans une réceptivité particulière. Elle le prédispose à certains lieux. Ce sont ces prédispositions de lieux que nous avons cherché à étudier à travers les « *mondes lexicaux* ». Cela dit, un énoncé ne définit pas un lieu en lui-même et cette notion de lieu n'a de sens que comme pôle d'un système de tensions à partir duquel le discours se construit. En effet ces lieux n'ayant au départ aucune forme propre, ne peuvent émerger que comme pure différence. Chez Bakhtine, par exemple, un énoncé n'est pas défini par sa forme propositionnelle propre mais comme réplique d'un dialogue social. Pour nous, ce processus dialogique correspond à une première approche du sens, celui où le locuteur ou bien l'allocutaire cherche ses marques, celui où justement ce fouillis d'impressions commence à s'organiser comme système dynamique de marques, mettant en jeu d'abord des oppositions identitaires entre le "je" et le "tu" pour ensuite éventuellement se stabiliser et s'objectiver dans une forme qui deviendra alors elle-même l'expression d'un lieu d'échange et de négociation, c'est à dire un signe.

Au niveau méthodologique, notre notion de distribution se résume donc à l'étude des présences de lexèmes dans des unités de contexte dont les frontières ne peuvent être que floues.

Sans entrer dans le détail, le choix des lexèmes ou de manière moins précise, des mots dits "pleins" se justifie par le fait qu'ils sont surtout remplis de nos usages du monde et sont donc plus aptes à connoter ce fond isotopique des énoncés. Bien sûr, nous ne disons pas que les mots "outils" ne peuvent pas être porteur de trait isotopique ici ou là mais ce rôle revient davantage aux mots pleins par définition même de leur "plénitude".

Cela dit, si la notion de lexème n'est pas complètement opérationnelle que dire de celle d'énoncé ? Ce n'est pas l'énoncé en soi qui nous intéresse mais l'énoncé en tant qu'il fonde globalement un lieu d'énonciation comme *fondement topique*, ce qu'il faut distinguer de son contenu représentationnel dépendant de l'organisation syntaxico-sémantique de cet énoncé. C'est ce changement de perspective qui justifie notre méthodologie. Celle-ci consiste à prendre plusieurs découpages du texte en unités de contexte et à chercher les distributions stables, c'est à dire, indépendantes de l'arbitraire de chaque découpage : ceci permet de ne définir les unités de contexte que par leur ordre de grandeur (souvent de l'ordre du paragraphe).

L'analyse statistique, en distinguant des classes d'unités de contexte de même distribution, distingue du même coup des lieux d'énonciation plus habituels où le locuteur se situe pour coordonner les objets auxquels il s'intéresse à un moment donné de son énonciation.

Nous appelons "*mondes lexicaux*" l'empreinte lexicale de ces lieux dans l'énonciation, mondes qui sont visualisés, techniquement, par le vocabulaire spécifique des classes. Cette hypothèse d'un système de lieux permet de différencier le « *locuteur* » de « *l'énonciateur* » selon une distinction proche de celle de Ducrot : dans un discours, le locuteur gère les changements de lieux et prend en charge la totalité de l'énonciation alors que l'énonciateur ne prend en charge que la cohérence d'un lieu particulier, d'un point de vue muni de sa logique particulière. Pour nous, ces deux entités sont distinguables au niveau sémiotique du fait qu'elles sont inscrites dans le jeu même des distributions.

Le jeu entre énonciateurs (au pluriel) et locuteur (au singulier) rend compte de l'ambiguïté d'un lieu qui se définit autant par sa cohérence logique interne que par le fait qu'il s'oppose à d'autres lieux. Cet aspect rapproche cette notion de lieux de la notion classique de "topoi" en ce sens qu'on ne construit pas un discours à l'aide d'un lieu unique mais par un agencement de "lieux communs", ce qui constitue justement l'art du rhéteur.

La nature de ces lieux dépend, bien sûr, de la nature du discours étudié, allant du discours rationnel imposant une complémentarité logique des lieux, aux actes de langage de la vie quotidienne dépendant d'une dynamique de construction plus dialogique voire contradictoire ou conflictuelle.

Cela dit, quel statut donner à ces "lieux" qui ne sont pas des "thèmes" ; quel statut par rapport à la notion de "référent" ?

Cette dernière notion est elle-même ambiguë : chez Benveniste, par exemple, elle renvoie parfois à la situation réelle d'énonciation (sens 1), parfois aux usages du locuteur (sens 2). C'est dans cette opposition que semble s'inscrire l'opposition référence (d'une phrase) / référent (d'un mot). Pour cet auteur, « *si le «sens» de la phrase est l'idée qu'elle exprime, la «référence» de la phrase est l'état de choses qui la provoque, la situation de discours ou de fait à laquelle elle se rapporte et que nous ne pouvons jamais ni prévoir, ni deviner* ». Alors que le référent d'un mot est « *l'objet particulier auquel le mot correspond dans le concret de la circonstance ou de l'usage* ». Cette dernière formulation est elle-même ambiguë. Si le référent renvoie au concret de la circonstance, il est défini au sens 1 ; si le référent renvoie à l'usage, il évoque davantage le sens 2.

L'utilisation du mot "référent" au sens 2 ne peut se justifier que si l'on abandonne la notion de réalité du référent en soi. Elle implique des « *habitus* » c'est à dire des formes d'activité consacrées par l'usage qui reflètent autant un "monde extérieur" (plutôt dans le sens de l'*Umwelt* de Uexküll) qu'une prédisposition à investir certains types de lieux. Pour cette raison, au niveau de la terminologie, nous utiliserons plutôt le terme de *topoï* que celui de référent du fait qu'il évoque une pluralité construite, même s'il est bien évident que ce construit doit tenir compte des contraintes des existants impliqués dans chaque situation d'énonciation.

Ce point de vue, bien que statistique, rencontre, nous semble-t-il, celui d'Anscombe et Ducrot (1995) pour qui, «*d'une façon générale, se représenter comme discourant sur un état de choses, c'est avant tout lui faire correspondre des topoï. En d'autres termes, utiliser des mots c'est convoquer des topoï. D'où l'hypothèse que le sens des mots n'est fondamentalement pas la donation d'un référent, mais la mise à disposition d'un faisceau de topoï.*» (Théorie des topoï, p51)

En résumé, de même que, selon Bourdieu, les activités humaines sont structurées par des «*habitus*», les discours, qui ne sont que les traces langagières de ces activités (Achard), sont structurés par des systèmes de lieux ou topoï agissant comme des attracteurs pour le locuteur. Un "lieu" ne peut véritablement être défini en soi. Un lieu se définit par l'existence même d'autres lieux auquel il s'oppose et sans lesquels il perd son identité. Dans le cadre de son discours, un locuteur doit gérer au fil du processus énonciatif ce passage d'un lieu à l'autre qui est aussi le passage d'une identité à l'autre. Cela est le propre même de l'argumentation rhétorique et dialectique. En conséquence, la mise en place d'un système de lieux ou «*topoï*» n'a pas qu'une visée référentielle, elle vise également à gérer l'activité d'un «*locuteur* » dans la manière de définir ses positions d'énonciation. Dans un corpus particulier, les "lieux habituels" peuvent être indexés par leurs «*mondes lexicaux* ». Cela, une analyse statistique peut le montrer de manière relativement automatique. C'est l'objectif même de la méthodologie "Alceste" (pour « Analyse des Lexèmes Cooccurrents dans les Enoncés Simples d'un Texte »).

2. DEUX EXEMPLES D'APPLICATION CONNOTANT LE MÊME SYSTÈME DE TOPOÏ

Afin de préciser la relation entre "mondes lexicaux" et "topoï", présentons rapidement les résultats de l'analyse de deux corpus qui semblent bien différents dans leur condition de production : Le premier est un texte littéraire d'un auteur - «*Aurélia* » de G. de Nerval - et le second est le corpus des six numéros de la revue «*le Surréalisme au Service de la Révolution* » (noté SaSDLR) qui est l'oeuvre d'un groupe, le groupe surréaliste animé par André Breton entre les années 1930 et 1933.

Dans ces deux analyses, les "topoï" reflètent, selon notre interprétation, un même processus de démarquage de type catégorique. Les résultats de ces analyses ont été publiés séparément (1994, 1997). Nous nous contenterons ici de suggérer la ressemblance existant entre les partitions en trois classes d'énoncés obtenues par l'analyse statistique.

L'ordre des bases lexicales présentées dans le tableau ci-joint est fourni par le logiciel et ne dépend que d'un coefficient de spécificité du mot à la classe (les mots sont triés par spécificité décroissante ; entre parenthèses, la valeur indique le nombre d'u.c.e. de la classe contenant le mot). Par contre, les intitulés "*imaginaire, réel, symbolique*" correspondent à notre appréciation. Nous avons renumérotés les classes afin de les apparier.

AURELIA		SaSdIR	
Nombre de formes distinctes	4580	Nombre de formes distinctes	21487
Nombre de hapax	2820	Nombre de hapax	10858
Nombre d'occurrences	21880	Nombre d'occurrences	203477
Nombre d'occur. analysées	5474	Nombre d'occur. analysées	52109
Nombre d'u.c.e.	739	Nombre d'u.c.e.	9756
Nombre d'u.c.i.	19	Nombre d'u.c.i.	179
Dendrogramme des classes :		Dendrogramme des classes :	
Cl. 1 (343uce) -----+ + Cl. 2 (89uce) ---+ -+ Cl. 3 (166uce) ---+		Cl. 1 (3399uce) -----+ + Cl. 2 (1938uce) ---+ -+ Cl. 3 (3110uce) ---+	
598 u.c.e classées sur 739 soit 80.92 % des u.c.e		8447 u.c.e classées sur 9756 soit 86.58 % des u.c.e	
C vis(23), fleur+(16), yeux(19), I figure+(24), lumiere+(13), soleil+(19), a terre+(33), etoile+(15), mur+(10), s vaste+(12), corps(9), couleur+(9), s rayon+(10), brill+er(9), devenir.(12), e elev+er(12), entour+er(12), sembl+er(40), i nuag+e(9), vieill<(9), cru+(11), l eclaire+(7), grand+(24), profond+(7), teinte+(8), blan+14(7), air+(11), feu+(8), lune+(8), nuit+(23), regard+(8), sang+(7), fix+er(7), march+er(8), sortir.(16), I cote+(7), montagn+e(7), ombr+e(11), M douce+(6), doux(6), vetu+(6), arbre+(9), A astre+(9), ciel+(11), demeure+(6), G porte+(11), sein+(6), distingu+er(6), T eclair+er(6), entr+er(12), penetr+er(6), N enf+ant(9), treill<(6), aile+(4), A antique+(4), clair+(4), couvert+(5), T divin+(13), longue+(7), naturel+(5), R ouvrier<(4), particulier+(4), partie+(8), E peint+(5), souterrain+(4), taille+(5), rouge<(4), vert+(4), angle+(5), anima+l(5), bras(5), clarite+(5), colline+(5), compagnon+(4), costume+(4), couchant+(4), eau+(8), epreuve+(4), escalier+(4), fleuve+(5), germe+(4);	C noir+(69), rouge<(53), eau+(76), femme+(152), main+(118), nuit+(75), pied+(62), soleil+(38), statue+(58), terre+(50), tete+(78), nu+(40), blan+14(58), air+(47), cheminee+(29), fille+(45), yeux(76), vill+23(40), mur+(38), rose+(28), bleu+(26), vert+(29), arbre+(31), bouche+(40), boule+(34), bras(33), cheveu+(27), couleur+(55), dent+(23), droite+(35), feu+(44), soir+(33), ventre+(22), voiture+(22), petit+(116), belle+(41), favorable+(32), gauche+(40), nocturne+(24), se+14(20), legerement(17), arcade+(20), bois(22), cheva+l(24), ciel+(35), coeur+(45), corps(44), doigt+(24), escalier+(19), fenetre+(20), levre+(23), mer+(23), metre+(19), oiseau+(26), pain+(24), poitrine+(19), sable+(23), sexe+(23), vent+(19), visage+(28), asseoir.(19), chant+er(25), dechir+er(17), descendre.(20), regard+er(30), roul+er(21), feuill+23(37), ombr+e(34), bete+(14), chaud+(17), doux(13), etoile+(17), matin+(24), mort+(73), pale+(14), mere+(26), amour+(65), batiment+(18), boite+(20);		

C l a s s e 2 R É E L	paris(4), ami+(13), amour+(8), cimetiere+(7), faute+(6), lettre+(6), papier+(4), pus(6), tombe+(7), donn+er(11), laiss+er(4), regard+er(4), rendre.(12), suivre.(6), visit+er(5), aller.(14), bague+(3), larme+(4), os+er(4), retrou+er(6), vierge+(3), eprouv+er(4), poe+19(3), malade+(3), maison+(8), parent+(3), veille+(4), chant+er(4), cherch+er(6), pass+er(7), rappel+er(4), pauvre+(3), fou+(2), heur+eux(4), suivant+(3), argent(3), campagne+(2), coeur+(4), doigt+(2), femme+(7), mot+(3), voyage+(2), pri+er(2), fievre+(2), tendre+(2), grave+(2), mauvais+(2), anneau+(2), douceur+(2), hasard+(3), memoire+(2), nom+(3), apprendre.(2), attendr+ir(2), consult+er(2), dirig+er(5), entendre.(6), parcourir.(2), parl+er(5), racont+er(2), rencontr+er(3), retent+ir(2), retour+er(2), reveill+er(2), inconnu+(3), prendre.(6), sentir.(6), vill+23(4), cher+(1), confus+(2), content+(1), long+(2), seul+(3), singulier+(1), mere+(1), annee+(2), centre+(1), certitude+(1), choeur+(1);	allemand+(31), francais+(92), milita+ire(35), ouvrier<(42), decembre+(19), article+(29), camarade+(27), commissariat+(15), communis+me(40), ecole+(25), ecrivain+(34), fascis+me(17), genera+l(48), guerre+(61), journa+l(31), lettre+(89), livre+(59), paix(21), police+(31), revue+(29), sport+(15), titre+(30), adress+er(17), ecrire.(40), publi+er(31), bourgeois<(69), collaborat+ion(18), edit+ion(39), elev+e(34), justic<(23), litterair<(25), polit+16(24), traduct+ion(18), cyr(28), affaire+(23), directeur+(15), publi+14(26), seri+eux(18), sportif(14), novembre+(16), septembre+(14), etudiant+(14), film+(14), pretre+(23), succes(14), menac+er(33), cathol+16(21), revolution<(71), soviet<(15), sadoul(15), arme+(32), saint+(23), fevrier+(15), ami+(21), auteur+(40), gens(28), minist+12(11), occasion+(24), officier+(17), ouvrage+(20), page+(30), peuple+(22), texte+(27), theatre+(15), admir+er(18), apprendre.(24), declar+er(20), demand+er(32), excus+er(16), faire.(152);
C l a s s e 3 S Y M B O L I Q U E	dieu+(19), esprit+(30), vie+(18), monde+(23), double+(6), evenement+(6), humain+(9), vrai+(5), etude+(5), idee+(15), science+(7), exist+er(7), amer+(4), chretien+(6), vague+(7), maladie+(4), rapport+(4), souvenir+(9), comprendre.(8), occup+er(5), vivre.(5), entier+(6), mort+(15), perdu+(7), oncle+(4), desespoir+(3), fois(7), proie+(3), raison+(5), sommeil(7), tradition+(3), appel+er(7), frem+ir(6), pardonn+er(3), reun+ir(5), song+er(5), act+ion(3), croy+ant(4), immort+el(6), religi<(7), revolution<(4), fata+l(7), circonference+(2), douleur+(6), existence+(7), recit+(2), philo<(2), celeste+(3), recu+(3), ame+(12), compte+(3), ignorance+(3), moment+(9), moyen+(3), songe+(4), coincid+er(3), esper+er(3), expliqu+er(4), mourir.(6), sauv+er(3), maudi<(3), pres+ent(3), chose+(4), image+(7), vision+(4), pens+er(12), etern+el(5), age+(3), certain+(2), cheri+(2), dernier+(5), distinct+(2), divers+(7), ennemi+(2), exterieur+(2), force+(7), froide+(1), meilleur+(2), precis+(3);	concret+(47), evoluti+f(38), expressi+f(80), humain+(63), mora+l(90), particulier+(43), pratique+(42), connaissance+(63), contenu+(42), developpement+(43), devenir+(54), dialectique(91), domaine+(41), element+(43), esprit+(175), homme+(175), humour+(40), idee+(81), image+(89), materialis+me(81), methode+(42), monde+(132), mot+(107), moyen+(74), nature+(57), objet+(156), phenomene+(39), processus(33), rapport+(70), realite+(97), sens(95), societe+(55), surrealism+me(150), valeur+(49), ag+ir(81), ni+er(32), pens+er(233), act+ion(163), log+16(49), objectiv<(40), philo<(71), poe+19(227), possi+ble(101), psych+16(33), rev+e(126), re+el(76), theori<(40), exterieur+(42), propre+(55), desir+(49), dire+(109), experience+(32), limite+(36), notion+(41), principe+(39), science+(35), tendance+(31), vie+(103), tendre.(41), imaginat+ion(30), paranoia<(26), problem<(43), techn+16(33), abstrait+(24), contraire+(26), inconscient+(36), necessaire+(32), negati+f(28), plan+(37);

Dans l'analyse du *SaSdIR*, nous avions dénommé les trois grandes classes (voir tableau) : *discours poétique*, *discours polémique* et *discours philosophique*. Des informations

supplémentaires sur l'origine des énoncés regroupés dans ces classes permettent de situer, par exemple, les énoncés d'Eluard plutôt dans la première classe, ceux d'Aragon plutôt dans la deuxième et ceux de Breton plutôt dans la troisième.

Si nous comparons avec l'analyse d'*Aurélia* de G. de Nerval, nous sommes frappés à la fois par la ressemblance globale des "lieux" et la différence avec laquelle ils sont investis par les énonciateurs avec des contenus propres liés à leur expérience singulière.

Le premier monde lexical d'*Aurélia* regroupe plutôt les énoncés relatifs à la description des rêves notamment les grands rêves évoqués dans la première partie de l'oeuvre et qui sont à l'origine de l'oeuvre elle-même. C'est de cette expérience des rêves, expérience vécue lors d'un premier internement, que l'auteur veut justement rendre compte.

Le second monde lexical regroupe des énoncés plus narratifs où Nerval évoque ses errances principalement dans Paris, évoque aussi des sentiments envers des amis, des proches, des disparus.

Le troisième monde lexical est plus cognitif et spirituel. Il recouvre les énoncés relatifs aux interrogations sur la religion, les sciences, sur la possibilité une fois entrevue dans ses rêves de retrouver, dans un autre monde, les êtres chers qu'il a perdus.

La relation de ces trois mondes avec la topique de Lacan - *Imaginaire, Réel, Symbolique* - nous semble suffisamment intuitive pour l'évoquer rapidement.

Il est de même, relativement aisément de relier (toujours intuitivement) le discours "poétique" de l'analyse de la revue surréaliste au "monde imaginaire", le discours "polémique" au monde "réel" et le discours "philosophique" au monde "symbolique".

Ce lien entre ces deux structures lexicales peut être objectivé directement par une certaine communauté de mots entre les classes correspondantes : 16 bases lexicales sont communes au monde imaginaire : *rouge, eau+, nuit+, soleil+, terre+, blan+, air+, bras, couleur+, feu+, ciel+, corps, escalier+, ombre+, doux, étoile+* ; 14 au monde symbolique : *esprit+, vie+, monde+, humain+, idée+, science+, rapport+, act+ion, (r)évolution<, philo<, moyen+, image+, pens+er, extérieur+* ; quant à ce qui correspond au monde réel, les contenus sont très différents ; seules trois bases sont communes : *lettre+, ami+, apprendre* ou curieusement les lettres d'amour chez Nerval font écho dans l'autre corpus à une lettre d'injure de Sadoul au premier de la promotion de Saint-Cyr. Cela pour dire que les contenus associés à ces mondes lexicaux dans les deux œuvres sont bien différents : c'est dans le rapport des mondes entre eux qu'une topique commune reste lisible.

Le seul point commun des conditions de production de ces œuvres que nous percevons est dans leur relation d'une expérience vécue originale. C'est justement l'originalité de l'expérience qui pose un problème sémiotique de mise en forme : comment retranscrire un tel vécu hors - norme dans les canons littéraires du moment ? (du moins pour Nerval... pour le groupe "Breton", il revendique sa différence par la création même de la revue. Mais celle-ci ne pouvait pas fonctionner de manière "canonique", notamment dans la distribution des rôles et des rubriques. Aussi cette revue qui est la trace de l'aventure intellectuelle de ce groupe durant trois années, peut bien être considérée, par sa création même, comme la mise en forme d'une expérience de vie, hors - norme).

Dans ce cas, si nous nous risquons à développer notre point de vue, la forme prise par le développement de ces œuvres en tant que trace d'une telle expérience dépendrait d'un processus plus archaïque, moins dépendant des idéologies et catégories d'une époque mais dépendant de caractéristiques endogènes à la sémiotique même... du moins, telles sont nos hypothèses façonnées par les travaux de Peirce. Le rapport intuitif des mondes lexicaux avec la topique de Lacan en est d'ailleurs un indice. Ceci va en tout cas dans le sens des travaux de Everaert-Desmedt. Cette sémioticienne propose en effet une correspondance entre la topique de Lacan et les catégories de Peirce. Cela en tout cas ouvre une voie pour comprendre cette possibilité de faire apparaître la topique lacanienne à travers le prisme des usages langagiers.

Cela dit, d'un point de vue "naïf", il peut sembler évident que l'expression d'une expérience nouvelle (ou d'un traumatisme, par exemple) passe par différentes phases oscillantes :

1) le sujet cherche à en rendre compte d'abord du contenu même de son expérience : ce qu'il a perçu ou ressenti (mais de ce premier moment, on ne peut généralement pas parler... sauf par la poésie ou le rêve).

2) ensuite, il cherche à re-situer cette expérience dans le contexte d'une situation, d'un temps où elle s'inscrit comme passage ou comme conflit (dans le cas d'un trauma, c'est souvent ce moment qui est d'abord exprimé : voir par exemple les témoignages de guerre)

3) enfin par une interrogation sur le sens de tout cela : comment construire un nouveau point de vue qui intègre le vécu de l'expérience dans un nouveau prédicat. En quoi celle-ci est fondatrice d'une nouvelle conception sur les choses, sur la manière de voir la réalité ? En quoi, elle est capable de rendre compte de ce dont on n'a pu parler ?

Mais en posant cela nous ne faisons que reproduire le même processus sémiotique tel que le développe Peirce. Pour cet auteur les trois moments de ce processus sémiotique se retrouvent dès la construction d'une proposition : sujet, prédicat, proposition. Le sujet s'impose de l'extérieur (Réel), le prédicat s'impose de l'intérieur (Imaginaire) et la proposition unifie ces deux points de vue en une même unité de jugement, en une même forme (Symbolique), en un même signe (Pour une étude de la notion de proposition chez Peirce, se reporter à l'article de Christiane Chauviré : « Peirce, le langage et l'action »).

Quoiqu'il en soit, notre objectif sera atteint ici si nous avons pu montrer la possibilité d'étudier les traces d'un tel processus à l'aide d'outils statistiques.

RÉFÉRENCES

- Achard, Pierre (1993) *La sociologie du Langage*, P.U.F.
- Anscombe, Jean-Claude (1995) *Théorie des Topoi*, Kimé.
- Bakhtine, Mikhaïl (Tr. fr. 1977), *Le marxisme et la philosophie du langage*, Éditions de Minuit.
- Benveniste, Emile (1966) *Problèmes de linguistique générale*, 1, Gallimard
- Benzécri, Jean-Paul (1973) *l'Analyse des Données*, DUNOD.
- Benzécri, Jean-Paul (1981) *Pratique de l'Analyse des Données : linguistique et lexicologie*, DUNOD.
- Bourdieu, Pierre (1982) *Ce que parler veut dire*, Fayard.
- Brunet, Etienne (1988) Le vocabulaire religieux dans trois siècles de littérature française, *Bible et Informatique : méthodes, outils résultats*, Jérusalem, 9-13 juin 1988, p147-165
- Brunet, Etienne (1992) Voyages autour des mots, in Robert Martin (ed.), *Dictionnaire et lexicographie*, Didier Érudition, p167-184.
- Brunet, Etienne (1993) Un hypertexte statistique : Hyperbase, *Actes des secondes journées internationales d'analyse statistique de données textuelles*, Telecom Paris 93 S 003, p1-16
- Chauviré, Christiane (1979) Peirce, le langage et l'action, *Les Etudes philosophiques*, n° 1.

Ducrot, Oswald (1989) *Topoi et sens*, in George Morand, *Le Texte et l'Image*, C.A.L.S. Toulouse.

Grize, Jean-Blaise (1982) *De la logique à l'argumentation*, DROZ, Genève.

Harris, Zellig Sabbetai (1952) Discourse analysis, *Language*, 28, 1-30 [traduction de Dubois-Chalier dans Langage n° 13 (1969)]

Lalande, André (1926, 3ème ed; 1993) *Vocabulaire technique et critique de la philosophie*, P.U.F.

Lebart, Ludovic, Salem, André (1994) *statistique textuelle*, Dunod.

Pécheux, Michel (1969) *L'analyse automatique du discours*, Dunod.

Peirce, Charles Sanders (tr. fr. 1987) «*Textes fondamentaux de sémiotique*», Klincksieck.

Reinert, Max (1986) « Un logiciel d'analyse lexicale : ALCESTE », *Les cahiers de l'Analyse des Données*, 4, 471-484.

Reinert, Max (1993) « Les "mondes lexicaux" et leur "logique" à travers l'analyse statistique d'un corpus de récits de cauchemars », *Langage et Société*, 66, 5-39.

Reinert, Max (1994) « L'approche des mondes lexicaux dans Aurélia de G. de Nerval », in Evelyne Martin, *Les Textes et l'informatique*, Didier Érudition, 145-175.

Reinert, Max (1997) «Les Mondes lexicaux des six numéros de la revue "Le Surréalisme au Service de la Révolution" », *Cahiers du Centre de Recherche sur le Surréalisme (Mélusine)*, L'Age d'Homme, XVI, 270-302