

LA MACROSÉMANTIQUE

François RASTIER
C.N.R.S.

(Texte issu du chapitre VII de *Sémantique pour l'analyse*, Paris, Masson, 1994 ; révisé pour la traduction anglaise *Semantics for Descriptions*, Chicago UP, 2002)

SOMMAIRE

1. Le paradigme du texte

1.1. Trois définitions négatives

1.2. Définition positive

1.3. Diversité des textes

2. Texte empirique et textualité

3. Un modèle cognitiviste de la compréhension de texte

4. L'entreprise typologique

4.1. Problématique

4.2. La thématique

4.3. La dialectique

4.4. La dialogique

4.5. La tactique

4.6. L'interaction entre composantes sémantiques et les genres

5. La conception morphosémantique du texte

6. La lancinante question des dinosaures

6.1. Les anaphores et la construction des acteurs et des fonctions

6.2. Les isotopies temporelles et tonales

7. Analyse de textes d'expert

7.1. Deux stratégies

7.2. L'exemple de l'exemple

7.3. L'énonciation représentée

1. Le paradigme du texte □

Pour le caractériser, il convient d'abord de définir ce qu'est un texte.

1.1. Trois définitions négatives □

a) Considérer un texte comme une *chaîne de caractères* serait le réduire à sa seule substance graphique, encourager son traitement séquentiel ou plus exactement déterministe (au moyen d'une fenêtre de lecture déplacée linéairement, comme c'est souvent le cas dans les analyseurs syntaxiques) ; enfin le couper de ses entours local (sa situation) et global (la culture dont il procède). Ce serait aussi réduire le langage au seul plan du signifiant ; encore ne s'agirait-il que des mots, à quoi correspondent dans le meilleur des cas les chaînes de caractères.

b) Un texte n'est pas non plus une *suite d'instructions*, algorithmique ou non, comme le voudrait la sémantique dite *procédurale* et la psychologie qu'elle a influencée (chez Johnson-Laird, par exemple). En fait, cette conception assimile le texte à un programme informatique (qui est bien une suite d'instructions) et la compréhension à sa mise en œuvre par l'esprit, implicitement comparé à un ordinateur. Cependant, une instruction informatique est nécessairement exécutée par l'ordinateur, tandis qu'un interprétant, tel qu'on le définit en sémantique interprétative, n'est qu'un indice qui doit être reconnu comme tel par le lecteur. Par ailleurs, si un programme consiste en une suite d'instructions, elles sont collectivement nécessaires et suffisantes à son exécution. En revanche, les interprétants d'un texte peuvent lui être extrinsèques, et appartenir à un autre système de signes (des illustrations, par exemple) ; en outre, le recours à des interprétants est réglé par des facteurs comme le pacte générique que son lecteur ou son auditeur scelle en lui donnant sens [1].

La notion d'instruction suppose, semble-t-il, une conception normative de l'interprétation. Or, il ne s'agit pas d'imposer des normes, mais simplement de les décrire. Dans un cadre général de prohibitions et de prescriptions, l'activité interprétative ne se déploie pas selon un système, mais dans le cadre de ce qu'on pourrait appeler une *grammaire permissive*.

c) Enfin, un texte n'est pas une suite de *schémas cognitifs* (comme les propositions mentales, modèles mentaux, scripts, plans etc.). Sa lecture suscite certes des corrélats mentaux, mais sa structure ne consiste pas en de tels corrélats. Par rapport à la psychologie, un texte apparaît non pas comme un ensemble de représentations, mais comme (i) un ensemble structuré de contraintes sur la formation des représentations et (ii) un réseau de récepteurs sémantiques, qui ne sont pas des marques, mais des lieux d'accomplissement des attentes, ou du moins d'épreuve des présomptions.

Pour ces raisons, nous préconisons une approche non-cognitive — sans préjuger des mérites des approches cognitives.

1.2. Définition positive □

Résumons-nous en une définition préliminaire. *Un texte est une suite linguistique empirique attestée, produite dans une pratique sociale déterminée, et fixée sur un support quelconque.* Un texte peut être écrit ou oral, voire présenté par d'autres codes conventionnels (Morse, Ascii, etc.), et en interaction avec d'autres sémiotiques (film, etc.). Ces trois conditions s'entendent ainsi :

1. Le texte est attesté : il n'est pas une création théorique comme l'exemple de linguistique, même considéré comme texte. Cette première condition énonce un principe d'objectivité.

2. Il est produit dans une pratique sociale déterminée : c'est là un principe d'écologie. La connaissance ou la restitution hypothétique de cette pratique est nécessaire, bien que non suffisante. La délimitation du texte est assurée par la pratique sociale dont il procède [2]

3. Il est fixé sur un support : c'est la condition de son étude critique, supposant débat des conjectures. Cette condition empirique rompt avec le privilège exclusif de l'écrit et rappelle que la substance de l'expression n'est pas définitoire du texte.

L'unité empirique du texte ne préjuge pas (i) de la fixité de sa signification, (ii) de l'intention de son ou ses auteurs, (iii) de ses références non ostensives, (iv) ni de l'interprétation qui en est donnée par ses destinataires. Au sein de la sémiotique des cultures, c'est à la linguistique considérée comme science des textes qu'il revient de caractériser, par des conjectures rationnelles, sinon formelles, ces quatre facteurs. On convient alors que l'analyse du sens permet des conjectures rationnelles sur le rapport du texte au monde physique, (tel qu'il trouverait des corrélats dans ses références non ostensives), comme sur son rapport au monde des représentations (tel qu'il trouve ses corrélats dans les intentions de son auteur et dans les interprétations de ses destinataires).

Nous n'évoquons pas pour l'instant de traits structurels qui définiraient la textualité, car il nous semble que *le* texte en soi n'existe pas, et que la textualité est une abstraction (comme d'ailleurs *le* langage). Nous ne postulons pas d'universaux textuels. Certains genres peuvent prescrire des textes composés d'une phrase, d'un mot, d'une énumération. Par ailleurs, de nouvelles pratiques sociales peuvent demain susciter des genres aujourd'hui imprévisibles.

S'il existe des règles de bonne formation [3], elles sont relatives aux genres, non à la textualité. Les universaux en la matière ne sont que des catégories descriptives, en d'autres termes des universaux de méthode qui paraissent utiles pour décrire les discours, les genres, et les textes. Nous proposerons donc des critères propres à décrire la textualité, dans le cadre méthodologique d'une typologie des textes.

1.3. Diversité des textes □

Si nous proposons une sémantique *des* textes, et non *du* texte, c'est parce que le texte en soi n'est

qu'une abstraction temporairement utile. Une théorie du texte est utile si elle élabore et hiérarchise des concepts descriptifs pour affronter la diversité. De la même façon que la diversité des langues est le problème fondateur de la linguistique, la diversité des textes fonde la sémantique des textes.

Dans les domaines techniques, on a pu sous-estimer cette diversité, car on assiste à présent à une normalisation des textes industriels, qui unifie leurs règles de rédaction, d'indexation et de stockage, et facilite leur traitement informatique, du moins dans la phase de documentation. La norme la plus connue est la norme SGML (*Standard Generalized Markup Language*, ou langage normalisé de balisage généralisé). Cependant les normes usuelles formatent le plan de l'expression mais non celui du contenu, et laissent ouverte la question du rapport entre ces deux plans.

En outre, trois facteurs de diversité doivent être rappelés : la diversité des genres, la diversité corrélative des formats et présentations des textes, enfin celle des descriptions elles-mêmes.

1) Les besoins de description sémantique concernent aussi des textes non normalisés, ou dont les normes ne sont pas explicites. Par exemple, tel assureur veut analyser 9.000 constats d'accidents mortels qui comprennent des déclarations sous forme libre, tel service officieux chargé de faire de la prospective cherche à traiter 120.000 pages de textes d'histoire ou d'extraits de presse sur la politique balkanique, rédigés à diverses époques et selon des perspectives variées.

Enfin, plus généralement, si la description sémantique doit satisfaire des demandes techniques, elles ne concernent pas nécessairement des textes techniques. Beaucoup de demandes intéressent des textes fortement évaluatifs, idéologiques : études d'opinion, analyses d'image institutionnelle, d'impact, etc. Aussi faut-il éviter d'ériger les textes techniques en parangons de toute textualité. Une sémantique des textes doit pour répondre aux besoins pouvoir s'adapter à tous les genres et discours.

2) De plus en plus de textes sont aujourd'hui présentés sous une forme délinéarisable : la structure arborescente traditionnelle qui les découpait en chapitres, paragraphes, etc., se trouve redoublée voire dépassée par une structure en réseau qui exploite les possibilités ouvertes par l'indexation et la co-indexation. On les nomme *hypertextes*. Les nouveaux supports numériques permettent de nouveaux moyens d'édition et de consultation, et sans doute facilitent certaines stratégies de production et de lecture.

En outre, les textes sont de plus en plus souvent inclus dans des documents plus complexes, hétérogènes, qui juxtaposent plusieurs systèmes de signes, codes, ou moyens de représentation. Ces *hyperdocuments* se multiplient avec l'essor de la communication multimédia. Les problèmes qu'ils posent relèvent de la sémiotique, qui permet d'articuler la linguistique avec les disciplines de l'image, de la musique, etc. Ils ne sont pas nouveaux dans leur principe, car une grand-messe ou un opéra sont à leur manière des documents multimédia ; mais ils ne concernent plus seulement l'ethnolinguistique ou l'esthétique. Aussi la description sémantique doit pouvoir s'articuler avec la description d'autres systèmes de signes, au sein d'une sémiotique générale.

3) Enfin les principes de description doivent permettre de s'adapter à la diversité des besoins. Elle est plus grande qu'on ne le pense, et prend parfois un tour imprévisible. Par exemple, pour les besoins d'un procès, telle compagnie pourra demander si elle emploie tel mot dans le même sens depuis vingt ans... Les demandes les plus courantes vont de l'indexation automatique à l'extraction de connaissances.

Disposer d'un modèle cohérent peut séduire le linguiste, mais il doit s'apprêter à taire son souci d'orthodoxie pour mettre en application son modèle. Dans une recherche terminologique, par exemple, les contextes lointains passent au second plan. Il convient de tenir compte des domaines d'application et de leurs spécificités, mais encore des objectifs et des contraintes de l'application : ils ont un rôle constituant, car l'espace du problème détermine alors la pertinence sémantique.

La sémantique interprétative a été testée sur des textes littéraires comme sur des textes techniques,

ce qui a demandé de détailler certains de ses secteurs, sans toutefois imposer de remaniements majeurs. À ceux qui demandent comment traiter les textes neutres ou ordinaires, qui ne seraient ni littéraires ni techniques, nous répondons qu'il n'en existe pas. Cette question est inspirée sans doute par la philosophie du langage ordinaire, et plus généralement par l'idée qu'il existe un emploi neutre du langage, littéral, à la fois d'usage général et simplement dénotatif. En fait, tous les usages linguistiques sont normés : ils relèvent d'un genre et d'une pratique sociale, et même ceux qui donnent l'impression de liberté, notamment les usages privés, n'échappent pas à ces déterminations.

2. Texte empirique et textualité □

Le palier du texte est le palier de complexité supérieur de l'usage linguistique. On peut provisoirement appeler *textualité* ce qui rend le texte irréductible à une suite de phrases [4]. L'étude des structures textuelles permet d'aborder le problème de la textualité. Elle a cependant été négligée, tant pour des raisons théoriques que pratiques.

1) Elle s'est heurtée tout d'abord à des obstacles épistémologiques : le traductionnisme en sémantique logique, le positivisme en linguistique, et le propositionnalisme en psycholinguistique.

(a) Dans le domaine théorique, le paradigme logique se trouve devant un dilemme. Ou bien la textualité n'existe pas, car le sens du texte se réduit à celui de ses phrases, que l'on fait généralement correspondre à des propositions susceptibles de valeurs de vérité. Ou bien elle existe, mais on ne peut calculer strictement le sens du texte, car il n'existe pas de règles syntaxiques qui permettraient de le faire. C'est précisément l'impossibilité de définir des règles au sens fort qui a conduit à l'abandon des grammaires de textes fondées sur le paradigme formel (comme la montré le débat dans *Cognitive Science* à la fin des années quatre-vingts), et la faiblesse descriptive aujourd'hui reconnue du modèle de représentations discursives de Kamp n'a pas permis de tracer une direction de recherche plus encourageante.

(b) L'extension au texte des théories et des méthodes d'analyse syntaxique n'a jusqu'ici pas conduit à des progrès décisifs [5]. La notion de *texture* proposée par Halliday et Hasan montre bien les limites d'une approche linguistique discrètement positiviste. La texture, qui définit la spécificité du texte, est " assurée par la relation cohésive " (1976, p. 4), décrite en termes de présuppositions, et dont l'exemple le plus simple est l'anaphore pronominale. En fait, la textualité se réduit ici à la cohésion locale, marquée par des liens explicites (*cohésive ties*) comme par exemple des adverbes, ou ce que l'on a appelé des *connecteurs pragmatiques*. Cependant, les relations sémantiques ne sont pas nécessairement explicites, au sens où chacune aurait une marque isolable. D'autre part, les relations locales peuvent rendre compte de la période, en tant qu'unité syntaxique élargie, mais non de textes plus étendus qu'un paragraphe. Le problème des relations sémantiques à longue distance ne peut être résolu ni même posé dans le cadre théorique d'une syntaxe phrastique étendue, que certains nomment aujourd'hui *macrosyntaxe*. La macrosyntaxe est nécessaire, mais non suffisante pour rendre compte de la textualité. L'étude des pronoms et des connecteurs met en évidence des fonctionnements sémantiques beaucoup plus généraux et globaux que la recherche des "marques" ne le laisse supposer.

C'est au palier du texte que la conception commune de la compositionnalité laisse apparaître le plus clairement ses lacunes : en effet, le global y détermine le local et le *recompose*. C'est pourquoi une phrase et *a fortiori* un mot peuvent changer de sens quand se modifie leur contexte immédiat et lointain.

(c) Les théories psycholinguistiques de la compréhension de texte considérées comme classiques (cf. van Dijk et Kintsch, 1983) ont abordé l'espace du texte avec plus de décision que les grammaires de texte dont elles ont en quelque sorte pallié les lacunes ; mais elles sont demeurées comme elles tributaires d'un paradigme logiciste qui a limité leur adéquation.

2) En pratique, dans certaines applications, on peut traiter du texte sans tenir compte de la textualité. Dans l'objet empirique qu'est le texte, on peut ne retenir que des unités de petite taille (pour

l'indexation, l'étude de co-occurrences, etc.) sans les rapporter à la textualité, ni même à la totalité du texte. Les méthodes statistiques occupent une situation intermédiaire entre analyse "atomiste" et analyse "globale" : elles retiennent de petites unités du signifiant, mais évaluent leur récurrence par rapport à l'ensemble du corpus (fréquences relatives, écarts réduits). En outre, beaucoup de corpus sont composés de plusieurs textes, parfois en très grand nombre, et les tâches de traitement excluent souvent de rendre compte de la spécificité de chacun. Il nous paraît cependant nécessaire de préciser la nature de la textualité : même si l'on est amené à la négliger, que ce soit du moins en connaissance de cause.

3) L'enjeu pour la sémantique est crucial. En effet, comme les structures textuelles sont essentiellement sémantiques, elles relèvent plutôt de normes et de régularités que de règles — et échappent à une linguistique restreinte qui concevrait les règles conformément à la théorie des langages formels. Le caractère culturel de ces normes dissuade de considérer la textualité comme un invariant. Du moins, s'il existe au plan sémantique des formes générales voire universelles de la textualité, c'est à une sémantique comparée de les caractériser.

3. Un modèle cognitiviste de la compréhension de texte □

Sans le présenter dans son détail, ce qui a été fait maintes fois, nous allons formuler quelques observations sur le modèle de van Dijk et Kintsch (1983), car pendant vingt-cinq ans il a servi de référence obligée et, bien qu'il ait été abandonné par Kintsch, il n'a pas été vraiment remplacé. Nous précisons à cette occasion notre point de vue.

Pour van Dijk et Kintsch, la première phase du traitement consiste à coder chaque phrase du texte en une ou plusieurs formules du calcul des propositions ou des prédicats. L'ensemble de ces formules constitue la *microstructure* du texte. Les propositions que l'on peut construire en généralisant par abstraction plusieurs propositions de la microstructure sont dites *macropropositions*. L'ensemble des macropropositions constitue la *macrostructure*. La microstructure et la macrostructure constituent ensemble la *base de texte*. Le thème, sujet, ou sens global du texte est une macroproposition, dont je donnerais volontiers en exemple, pour *Guerre et Paix*, le chef-d'œuvre de Tolstoï, cette traduction linguistique : " Une femme aime un homme et en épouse un autre ". Toute l'analyse repose sur un certain nombre de postulats dont nous soulignons les conséquences :

(i) Le sens d'un texte se compose de propositions logiques, et van Dijk souligne même que toute structure sémantique est propositionnelle (1981, p. 74). Cette thèse reprend celle des grammaires générales du XVIII^{ème} siècle, héritières de la scolastique : elle préexistait à la constitution de la linguistique en discipline.

(ii) Le sens d'un texte est de nature conceptuelle ; et Kintsch rappelle que la microstructure comme la macrostructure sont des niveaux conceptuels (1991, p. 107). Le linguistique et le sémantique sont ainsi séparés, le linguistique étant réduit à la morphosyntaxe. Comment alors concevoir la spécificité sémantique des langues ?

(iii) On peut passer sans reste du linguistique (phrases du texte) au logique (propositions). Or les expériences séculaires pour enrégimenter (*regiment*, selon le mot de Quine) les langues dans un calcul logique n'ont guère été concluantes, car les signifiés ne sont ni des constantes ni des variables. Ici, non seulement les procédures de transcription restent vagues, mais leurs résultats montrent que cet enrôlement ne se fait pas sans désertions massives : les modalités, les tons, les différences de niveau de langue, locales, historiques, les évaluations, les focalisations, les mises en relief et tous les effets contextuels d'afférence disparaissent tout simplement.

Ce montage théorique obéit à une stratégie de réduction, d'une part dans le codage des microstructures, puisque le langage de codage est drastiquement restreint, d'autre part dans le passage des microstructures aux macrostructures, puisqu'une règle de suppression permet d'éliminer toutes les propositions qui ne conditionnent pas l'interprétation d'autres propositions. Le montage repose en outre sur une hypothèse forte touchant la compréhension effective des textes par tout un

chacun : elle est identique au dogme du cognitivisme orthodoxe qui affirme le format propositionnel des états mentaux (cf. Fodor, 1975). En l'occurrence, les résultats expérimentaux avancés pour affirmer la validité psychologique de ce modèle cachent un artefact : partant du principe que comprendre c'est résumer, on demande à des étudiants de résumer des textes, et l'on s'aperçoit que leurs résumés ne diffèrent pas trop de la macrostructure, résumé au format propositionnel élaboré par le chercheur. On en conclut à la validité psychologique du modèle. Or le résumé est en lui-même un genre qui sélectionne conventionnellement ce que l'on appelle les informations importantes. Le résumé ne représente pas le texte, c'est une réécriture dans un autre genre textuel, qui s'accorde avec le postulat que tout texte est un support d'information. Dans un contexte universitaire, les sujets d'expérience ont naturellement à cœur de le confirmer.

Le cognitivisme orthodoxe a reformulé la conception informationnelle du texte en la liant au computationnalisme. Cette conception repose sur l'idée traditionnelle que le langage est un simple instrument de la pensée. Dès lors l'analyse d'un texte consiste à restituer la pensée dont il procède, ce qui conduit en général à réduire sa spécificité linguistique pour lui substituer une structure cognitive que Kintsch nomme un *modèle de situation*, et dont il précise qu'à ce niveau " le texte lui-même a perdu son individualité " (1991, p. 107).

Corrélativement, les problèmes herméneutiques que soulève le codage en propositions ne sont pas abordés. En d'autres termes, la compréhension, définie comme constitution d'un modèle de situation, oblitère l'interprétation proprement dite. Prenons pour exemple le syntagme *the old men and women*, qui se transcrit *and [old [men], old [women]]*. Kintsch est pris d'un doute justifié : "peut-être les femmes n'étaient pas du tout vieilles" (1991, p. 108), et en conclut que "la connaissance générale sur toute chose contraint la construction des représentations de discours à tous les niveaux" (ibid.). Cette conclusion est des plus vagues, car c'est la connaissance particulière de certaines choses qui contraint la construction du sens[6], et non la connaissance générale de toutes. Il faut, bien entendu, réduire la généralité encyclopédique par des normes de pertinence, et montrer leur incidence sur l'actualisation sémantique : c'est précisément là un objectif de la sémantique interprétative.

4. L'entreprise typologique □

4.1. Problématique □

a) Alors que le problème fondateur de la linguistique est celui de la diversité des langues, il faut poser aussi celui de la diversité des types de textes. D'une part, il n'existe pas de texte sans *genre* ; en outre, tout genre relève d'un *discours* (ex. politique, juridique, religieux, etc.). Par exemple, dans le discours médical, un professeur des hôpitaux dispose dans sa pratique professionnelle de trois genres écrits : l'observation clinique, l'article scientifique et la lettre au collègue. Même les échanges linguistiques qui paraissent les plus spontanés sont réglés par les pratiques sociales dans lesquelles ils prennent place, et relèvent donc d'un discours et d'un genre. Malgré certains théoriciens de l'analyse conversationnelle, la conversation, par exemple, n'est pas un genre ni un discours. Nous disposons tous de plusieurs genres conversationnels, liés à des pratiques différentes, de l'entretien à la conversation de cantine, et dont chacun a ses spécificités.

Comme les textes ne sont pas gouvernés par des règles (au sens trop fort du terme qui est d'usage en linguistique), mais par des normes, la description des genres est un objectif pour la typologie des textes, comme aussi un enjeu important pour la linguistique. Elle permet en effet d'articuler les connaissances sur la structure de la langue avec les observations sur la structure des textes particuliers. Les linguistiques textuelles dont nous disposons ont sous-estimé ces problèmes : la plupart sont universelles, et se révèlent trop puissantes en théorie pour être utiles en pratique. Il importe donc de discuter des critères de typologie. La méthodologie adoptée dépend elle-même des genres. C'est en effet dans le cadre de genres déterminés que doivent être définis les parcours de production et d'interprétation, d'où le caractère primordial de leur typologie.

En outre, la méthode doit s'adapter aux objectifs visés. De la même façon qu'il n'y a pas d'usage neutre de la langue, on peut écarter l'idée d'une méthodologie neutre et exhaustive.

b) Les critères d'une typologie des textes se divisent en trois groupes dont on ne peut dire qu'ils soient étanches : les critères anthropologiques, les critères philosophiques et les critères sémantiques. Il y a deux moyens de les définir, soit déductivement en instituant des classes *a priori*, soit inductivement par une méthode comparative qui permettrait de justifier ou du moins de refléter les classifications culturelles.

La plupart des typologies dont nous disposons transcendent la division en discours et en genres. Chacune de leurs classes peut inclure plusieurs discours, ou du moins des genres relevant de plusieurs discours. La plupart relèvent implicitement de l'anthropologie philosophique, et énumèrent les genres d'usage que l'homme peut faire du langage. Ainsi, Longacre distingue quatre types de discours (narratif, procédural, expositif et exhortatif) ; de Beaugrande et Dressler trois types (descriptif, narratif et argumentatif), ces fonctions dominant inégalement dans les trois types de textes qu'il distingue par ailleurs (littéraire, poétique, et scientifique) ; Greimas quatre types, en croisant une distinction anthropologique entre le pratique et le mythique, et une distinction philosophique entre qualités et actions ; Adam deux types, en reprenant la distinction précédente sous le nom du *faire* et de l' *être* qui définissent les pôles narratif et descriptif (à quoi il ajoute par exemple l'injonctif) ; van Dijk deux, narratif et argumentatif (qu'il définit comme des schémas textuels ou *superstructures*, cf. 1980a). Soit ce tableau :

	<i>Narratif</i>	<i>Descriptif</i>	<i>Procédural</i>	<i>Argumentatif</i>
Longacre	+	+	+	+ exhortatif
Greimas	+ idéologique	+ scientifique	+ technologique	+ axiologique
B & D	+	+		+
van Dijk	+	+		+
Adam	+			+ injonctif

Ces typologies pourraient être détaillées et bien d'autres citées (Werlich, Garavelli). Soulignons seulement leur stabilité, et le caractère traditionnel des types toujours reformulés : par exemple, la distinction entre récit et description a en rhétorique une histoire millénaire. Il semble que les quatre colonnes de notre tableau fassent système : le narratif et le descriptif sont deux modes de la représentation, définis comme tels depuis la *Poétique* d'Aristote. Ils s'opposent par là au procédural et à l'argumentatif qui sont deux modes d'action, sur les choses et sur les esprits respectivement, et qui s'opposent en quelque sorte comme le faire au faire-faire. Nous retrouvons ici les articulations majeures des deux grands paradigmes de la signification, représentationnaliste et intentionnaliste (cf. *supra*, ch. I). Leur usage est sans doute simplement ethnocentrique, ce qui n'enlève rien à leur valeur descriptive quand il s'agit de textes issus de notre culture, mais ne les qualifie pas cependant pour fonder une typologie à prétention universelle.

Greimas fait ici preuve d'originalité avec l'opposition pratique / mythique, d'origine lévi-straussienne, qui justifie la distinction entre les textes scientifiques et les textes axiologiques (dont il donne pour exemple les descriptions littéraires). Cette opposition reste une catégorie à fonder, car il existe des textes qui décrivent des pratiques mythiques (les rituels, les recettes alchimiques, par exemple), et qui par leur précision, leur organisation séquentielle, n'ont rien à envier à nos textes techniques. Dans bien des sociétés, la distinction entre le pratique et le mythique semble difficile à établir. Elle ne reflète peut-être qu'une exigence de notre civilisation laïcisée, où elle est utilisée pour distinguer le "véridique" du "fictionnel".

A part celle-là, les autres catégories sont fonctionnelles, et par là-même difficiles à appliquer, car tout texte résulterait d'un "dosage" de ces fonctions : une notice d'entretien, par exemple, est un texte tout aussi procédural qu'injonctif. La problématique fonctionnelle spécifie certes la conception instrumentale du langage, mais reste nécessairement générale, si l'on estime que le langage n'a qu'un petit nombre de fonctions (comme l'atteste le tableau ci-dessus). Dès lors, les types

fonctionnels définissent des classes trop accueillantes et conduisent à regrouper des textes hétéroclites : si par exemple on range ensemble les textes à dominante argumentative, on pourra être conduit à juxtaposer la thèse du philosophe et la plaidoirie de l'avocat. Aussi, plutôt que des fonctions supposées du langage nous préférons partir des fonctions attestées des textes : elles varient avec les pratiques sociales, comme avec les cultures où elles prennent place, et leur nombre n'est pas fixé *a priori*. Il se peut bien entendu que le langage ait des fonctions, mais elles sont à établir par une typologie interculturelle, et ne la fondent pas : l'on peut douter que le langage ait les mêmes fonctions en tout temps et en tout lieu, alors même que les cultures codifient si diversement son usage. En outre, les fonctions spécifiques de chaque genre n'ont d'intérêt pour notre propos que si l'on peut les rapporter à des modes de structuration sémantique des textes. Nous préférons donc proposer des catégories sémantiques propres à la caractérisation des genres.

b) La typologie des genres paraît indispensable pour les traitements automatiques. Soit en général, car l'analyse des corpus en situation montre que le lexique, la morphosyntaxe, la manière dont se posent les problèmes sémantiques de l'ambiguïté et de l'implicite, tout cela varie avec les genres. Les systèmes d'analyse et de génération doivent tenir compte de ces spécificités, et les projets de systèmes universels sont ainsi irréalistes, linguistiquement parlant [7]. Soit en particulier, car les genres sont déterminés par des pratiques sociales spécifiques, dans lesquelles les applications informatiques prennent place. Elles doivent donc tenir compte des contraintes propres aux pratiques où elles s'insèrent.

c) Pour établir le cadre conceptuel d'une typologie des textes, on peut concevoir la production et l'interprétation des textes comme une interaction non-séquentielle des composantes autonomes que nous avons introduites au deuxième chapitre : thématique, dialectique, dialogique et tactique.

4.2. La thématique □

a) La notion passablement vague de *thème* a divers usages, en critique thématique, en linguistique d'inspiration pragoise (par opposition à *rhème*) en analyse du discours (*topic* par opposition à *focus*) [8]. La microsémantique descriptive peut clarifier cette notion en définissant le thème comme un ensemble structuré de sèmes. Leur statut et leur nombre, les relations établies entre eux, tout cela varie selon les genres. Selon le statut de ces sèmes, il convient de distinguer les thèmes génériques et les thèmes spécifiques.

b) Un *thème générique* est défini par un sème ou une structure de sèmes génériques récurrents. Cette récurrence définit une isotopie ou un faisceau d'isotopies génériques (c'est-à-dire un groupement de sèmes génériques co-récurrents). Les isotopies génériques, et particulièrement domaniales, déterminent le "sujet" (*topic*) du texte, ou plus précisément induisent les impressions référentielles dominantes.

Etendons ici au texte les critères proposés au chapitre précédent pour la typologie des énoncés. Par exemple, les textes techniques étant contraints par un domaine d'application, ils ne manifestent qu'un domaine sémantique, alors que les textes littéraires peuvent en superposer plusieurs, en poésie notamment, ou n'en comporter aucun (dans le cas des soties ou de certains écrits dadaïstes). Selon les types de classes sémantiques qu'ils manifestent, on peut distinguer quatre sortes.

(i) Les thèmes taxémiques sont portés par les membres d'un même *taxème*, comme le //tabac// dans *Madame Bovary*, qui se décline en 'pipe', 'cigarette' et 'cigare', corrélés respectivement au mari, à la femme infidèle, et au premier amant.

(ii) Les thèmes domaniaux sont portés par les membres d'un même *domaine* sémantique (ex. //alimentation// dans *Cuisine pour toi et moi*, de Ginette Mathiot, qui prodigue outre des recettes des conseils touchant la diététique, l'approvisionnement, etc.). Les "langues" de spécialité décrivent et structurent un domaine sémantique.

(iii) Les thèmes dimensionnels sont portés par les membres d'une même *dimension* sémantique. Les

dimensions évaluatives ou thymiques, les tons, les espaces modaux, les plans temporels ou *chronotopes* définissent dans les textes autant de sortes d'isotopies dimensionnelles. Nous en donnerons des exemples dans les analyses de textes à la fin de ce chapitre.

(iv) Les thèmes liés à un champ sémantique sont plus variables, puisque les champs ne sont pas des classes de langue. Les plus faciles à identifier sont ceux qui sont codifiés (comme en topique littéraire) ou définitoires d'un corpus : par exemple, un corpus de notices d'entretien destiné au Falcon 900 se situe évidemment dans le domaine aéronautique (et les acceptions lexicales référeront toutes à ce domaine) mais elles en définissent un champ restreint.

c) Les *thèmes spécifiques* sont des groupements récurrents de sèmes spécifiques. Nous avons proposé de nommer *molécules sémiques* ces groupements. Ils peuvent être représentés par des graphes microsémantiques, mais ils ne sont pas nécessairement dépendants d'une lexicalisation particulière. Certains n'ont, pour ainsi dire, de nom en aucune langue : nous avons décrit par exemple, dans *L'Assommoir* de Zola, une molécule sémique qui regroupe les sèmes /jaune/, /chaud/, /visqueux/ et /néfaste/, et qui est lexicalisée par *alcool*, *sauce*, *morve*, *huile*, *pipi*, mais aussi dans toutes sortes de syntagmes voire de lexicalisations partielles dispersées (cf. 1989 b, II, ch. 2 ; traduction anglaise 1998). En règle générale, dans les textes techniques, les molécules sémiques ont cependant une lexicalisation privilégiée, voire exclusive (le terme), car les disciplines techniques répugnent à l'équivocité.

Dans la suite, nous conserverons les expressions *isotopie générique* pour désigner les thèmes génériques, et *molécule sémique* pour désigner les thèmes spécifiques. Nous utiliserons la notion de thème uniquement pour les formes sémantiques (génériques ou spécifiques) récurrentes de texte à texte, voire entre les genres d'un même discours.

4.3. La dialectique □

Comme elle traite des intervalles de temps représenté, la dialectique rencontre les théories du récit. Elle est définie à deux niveaux, dont nous présentons simplement les principaux concepts. Le premier niveau, dit *événementiel*, apparaît dans tous les textes structurés par une composante dialectique. Ses unités de base sont les *acteurs*, les *rôles* et les *fonctions* (au sens de Propp, soit un type d'action représentée)[9].

a) Un *acteur* peut être défini comme une classe d'actants : il est constitué par totalisation d'actants anaphoriques des périodes (au palier inférieur du texte, mésosémantique). Dans la période, les actants peuvent être nommés ou recevoir diverses descriptions (définies ou non). Chaque dénomination ou description lexicalise un ou plusieurs sèmes de l'acteur. L'acteur se compose de trois sortes de structures sémiques :

(i) Sa *molécule sémique* est constituée des sèmes spécifiques de ces actants. Par exemple, dans un même article de journal espagnol, la même personne est nommée : *la guapa senorita*, *la gentil novia*, *la amable anfitriona* (cf. Pottier, 1974, p. 95). Les trois adjectifs, comme les trois substantifs, contiennent des sèmes communs, inhérents dans les adjectifs, afférents dans les substantifs : /évaluation positive/, et /féminin/. Ce sont les sèmes récurrents de la molécule. Chaque syntagme caractérise en outre un champ différent (respectivement : qualités physiques, matrimoniales, et de socialité) et cette aimable personne s'y voit affecter les sèmes /jeune/, /fiancée/, et /hôtesse/.

(ii) Outre leur molécule sémique, les acteurs comportent par ailleurs des sèmes génériques, ici /humain/ ; mais dans certains genres les acteurs ne comprennent pas nécessairement le trait /animé/[10]

(iii) Enfin un acteur comprend des sèmes afférents, qui sont des *rôles*. Ce sont des cas sémantiques associés aux actants qu'il subsume. Par exemple, si A, B et C sont des acteurs, [A]->(ERGATIF)->[B] est un rôle, [A]->(DATIF)->[C] un autre. L'ensemble des rôles d'un acteur définit sa sphère interactionnelle. Elle évolue avec les intervalles temporels. On peut classer les acteurs selon le

nombre et la nature des rôles qui les caractérisent. Par exemple, dans le *Dom Juan* de Molière, le personnage éponyme dispose de huit types de rôles, mais son valet La Violette d'un seul. Les rôles connexes se groupent en espaces d'interaction, tous les acteurs n'étant pas nécessairement en interaction réciproque.

Selon la stratégie descriptive adoptée, on peut garder en mémoire tous les composants attribués à un acteur, ou simplement les sèmes invariants, communs à toutes les occurrences de tous ses actants. Mais il faut tenir compte du fait que dans certains genres textuels l'acquisition des sèmes n'est pas simplement cumulative : selon les intervalles temporels, l'acteur gagne ou perd des sèmes et peut même perdre tous les sèmes initiaux, voire être remplacé par un ou plusieurs autres acteurs. Que reste-t-il d'un blanc d'œuf devenu meringue au café ? Ce problème inquiétant pour la philosophie du langage peut être appelé celui de la *rigidité* de l'acteur.

b) Les *fonctions* sont des interactions typiques entre acteurs : ce sont des classes de processus. Comme les acteurs, elles sont définies par une molécule sémique et des sèmes génériques : par exemple, le *don* est une fonction irénique, le *défi* une fonction polémique. Les fonctions correspondent à des *valences* actorielles. Si [A] et [B] sont des acteurs, une fonction [DÉFI] peut s'écrire :

$$[A] \leftarrow (\text{ERG}) \leftarrow [\text{DÉFI}] \rightarrow (\text{DAT}) \rightarrow [B].$$

Les affinités sémiques entre acteurs et fonctions rendent compte des relations de concordance ou de rection qui s'établissent entre eux.

Les grammaires universelles du récit ont rapporté les fonctions à des actions primitives (Schank), à des chaînes causales abstraites (Rumelhart), à une praxéologie spéculative (Greimas). Il paraît plus utile de construire l'inventaire des fonctions à partir des corpus. Par exemple, la théorie de Greimas prévoit une fonction *contrat* : cela ne suffit évidemment pas pour décrire un corpus d'analyses financières. Tout comme au palier mésosémantique il faut faire proliférer les cas, au palier macrosémantique on peut multiplier les fonctions. Bref, l'inventaire des fonctions doit s'adapter aux discours (juridique, politique, etc.) et même aux genres ; il en va de même pour l'inventaire des acteurs. En outre, leur pertinence dépend du corpus choisi, comme des objectifs de la description.

Les fonctions peuvent se grouper en syntagmes fonctionnels (ex : *échange* se compose de deux transmissions, *affrontement* d'une attaque et d'une contre-attaque). Les *scripts* dont on fait grand cas en Intelligence Artificielle sont des syntagmes fonctionnels typés par une isotopie générique de champ ou de domaine.

c) Le niveau *agonistique* est hiérarchiquement supérieur au niveau événementiel. Comme il n'apparaît que dans les textes mythiques, et que les applications concernent pour l'essentiel des textes pratiques (notamment technologiques), sa présentation sera ici succincte. Ses unités de base sont les *agonistes* et les *séquences*.

(i) Un *agoniste* est un type constitutif d'une classe d'acteurs. Dans les textes mythiques au moins, il est fréquent que les acteurs relevant d'un même agoniste soient indexés sur des isotopies génériques différentes, voire dans des univers et des mondes divers, mais se trouvent cependant en relation métaphorique. Par exemple, dans *Toine* de Maupassant, la vieille est un acteur sur l'isotopie humaine, qui se trouve en relation métaphorique avec le coq sur l'isotopie animale, le vent du large sur l'isotopie météorologique, et la mort sur l'isotopie métaphysique. Bien qu'apparemment hétéroclites, ces acteurs reçoivent dans le texte des traits spécifiques communs, comme /agression/, /intensité/ (cf. Rastier, 1989 b, p. 220 ; traduction anglaise, 1998, p. 165).

Les agonistes sont définis d'une part par les composants invariants ou équivalents des molécules sémiques de leurs acteurs, et aussi par les rôles identiques ou équivalents dans les sphères interactionnelles de ces acteurs. À la différence des acteurs, ils ne comportent pas nécessairement de sèmes génériques qui les indexent sur des isotopies. La description des agonistes permet une

économie descriptive : par exemple, dans le *Dom Juan* de Molière, aux centaines d'occurrences d'actants correspondent 24 acteurs, et seulement 9 agonistes. L'inventaire des agonistes ne nous paraît pas devoir être fixé a priori [11], mais doit être construit en fonction des cultures, des discours et des genres, de façon à ménager une étude comparative.

(ii) Les *séquences* sont définies par homologation de syntagmes fonctionnels isomorphes. Par exemple, dans le *Dom Juan* de Molière, on relève 118 fonctions, que l'on peut regrouper en onze séquences seulement (cf. Rastier, 1971, p. 307). Comme les syntagmes qu'elles homologuent occupent des positions différentes dans le temps dialectique, les séquences sont ordonnées par des relations de logique narrative (présuppositions) mais non chronologiques.

La distinction entre niveau événementiel et niveau agonistique permet une typologie dialectique et conduit à détailler la notion de récit (cf. Rastier, 1989 b, I, ch. V ; traduction anglaise 1998). On appelle volontiers *récit* toute relation d'événements mettant en jeu des personnes ou des personnages, mais cette notion ne permet pas de distinguer l'épopée de l'anecdote (comme la *Margie* de Schank). La notion de récit ne suffit pas à définir un genre, car elle ne relève que d'une des quatre composantes de la textualité. En général, la composante dialectique des textes pratiques ne comporte que le niveau événementiel alors que les textes de fiction le redoublent par un niveau agonistique.

4.4. La dialogique □

La dialogique rend compte de la modalisation des unités sémantiques à tous les paliers de complexité du texte.

a) Un *univers* est l'ensemble des unités textuelles associées à un acteur ou à un foyer énonciatif : toute modalité est relative à un site (un univers) et un repère (un acteur). Par exemple, quand le narrateur de *la Cousine Bette* parle d'une *mauvaise bonne action*, 'bonne' renvoie à l'univers de deux acteurs, et 'mauvaise' à son propre univers.

Chaque univers est susceptible de se diviser en trois *mondes* : (i) Le *monde factuel* est composé des unités comportant la modalité assertorique ; (ii) Le *monde contrefactuel*, des unités comportant les modalités de l'impossible ou de l'irréel ; (iii) Le *monde du possible*, des unités comportant la modalité correspondante [12]. Mondes et univers peuvent évoluer selon les intervalles de temps dialectique. L'accessibilité entre mondes d'un même univers pose des problèmes délicats. En effet, dans chacun de ses mondes, un univers peut contenir des unités qui sont soit des images voire des répliques d'unités situées dans d'autres univers.

b) Les unités considérées comme vraies dans tous les mondes factuels des univers associés aux acteurs constituent ce qu'on peut appeler l' *univers de référence*. Il est associé, par défaut, à l'énonciateur représenté (ex. le narrateur d'un récit). C'est relativement à l'univers de référence que l'on peut attribuer en dernière analyse une valeur de vérité aux unités de tous les autres univers [13].

c) Dans le cadre ainsi esquissé, le récit relève de la dialectique et la narration de la dialogique. Ces deux composantes sont évidemment liées. Par exemple, une fonction *contrat* consiste en un échange de processus de transmission situés dans les mondes du possible associés aux acteurs contractants : en tant qu'échange "effectif" de promesses, le contrat relève de la dialectique, mais ces promesses elles-mêmes, en tant qu'unités modalisées, relèvent de la dialogique.

La dialogique fonde la typologie des énonciateurs représentés. Par exemple, en général, les textes d'instructions techniques ne comportent qu'un foyer énonciatif et un foyer interprétatif fixes et non nommés, et se présentent souvent comme une suite de phrases à l'impératif. Les articles scientifiques en revanche multiplient les énonciateurs délégués par le biais de citations ou d'allusions, et précisent les foyers interprétatifs par de discrets clins d'œil aux initiés. Nous verrons aussi que dans les textes de vulgarisation ou les entretiens techniques les structures dialogiques peuvent devenir complexes, la complexité maximale restant toutefois l'apanage de la littérature

quand elle multiplie les foyers énonciatifs et interprétatifs sans les hiérarchiser ni même les signaler (cf. le “style indirect libre”).

4.5. La tactique □

Cette dernière composante rend compte de la disposition linéaire des unités sémantiques à tous les paliers. La linéarité du signifié entretient certes des rapports étroits avec celle du signifiant, mais ne se confond cependant avec elle à aucun des paliers : rappelons les morphèmes à signifiant zéro, les ellipses, les unités macrosémantiques comme les isotopies ou les séquences, dont les manifestations sont discontinues. En outre, des unités sémantiques de toute dimension sont restituées par des inférences [14] : elles appartiennent alors au sens du texte, sans qu'on puisse pour autant leur assigner une position dans sa linéarité.

La position relative des unités du signifiant peut être évidemment utilisée comme indice des relations distributionnelles entre unités du signifié. Prenons l'exemple des unités dialectiques. La maxime *post hoc ergo propter hoc* est un des principes de l'intelligence narrative : elle conclut que ce qui advient après un événement en est la conséquence. Dans les textes procéduraux, la succession des positions dans le texte est en général homologue de la succession des intervalles du temps représenté : le texte commence par ce qui advient d'abord et finit par ce qui advient en dernier lieu. Mais les positions tactiques des unités sémantiques ne sont pas nécessairement homologues aux positions dans le temps textuel. Par exemple, dans le roman policier, c'est toujours à la fin du texte que l'on connaît les événements qui se sont déroulés en premier lieu : cette inversion fait précisément partie des règles du genre. Bref, même quand le texte a une structure dialectique, et quand le temps représenté est linéaire, les positions des unités sémantiques ne s'ordonnent pas nécessairement selon la linéarité du signifiant, ni selon celle du temps représenté.

4.6. L'interaction entre composantes sémantiques et les genres □

1. Nous venons de définir des unités et des relations dans les quatre composantes : elles sont dérivées, par diverses opérations méthodologiques, des unités et des relations du palier mésosémantique, et par là du palier microsémantique. Si complexes ou si abstraites soient-elles, elles restent analysables en sèmes. Nous avons procédé dans une perspective de production ou d'interprétation, qui en reste à la sémantique : les objets décrits sont tous constitués d'unités et de relations sémantiques organisant des fonds et des formes. Nous n'avons pas eu recours cependant aux hypothèses cognitives que suscite aujourd'hui l'étude de l'énonciation effective et de la compréhension au sens fort.

2. Chacune des composantes peut être la source de critères typologiques. On peut classer les textes selon les critères les plus divers, nombre et nature des isotopies génériques, représentation de l'énonciation, etc. Un jeune homme, par exemple, ne lisait que les romans en *il*, et repoussait les romans en *je*. Ce critère respectable, sans nul doute productif, permet de distinguer un romanesque objectiviste, à prétention sociologique, d'un romanesque subjectiviste et psychologisant. Mais nos critères doivent ici se régler sur l'objectif d'une typologie, et l'on ne gagnerait rien à créer une catégorie des textes en *je*, qui regrouperait par exemple Proust, les serments, et les lettres, quand les textes en *il* comprendraient Balzac et les rapports de police. Voilà pourquoi en somme, et toutes proportions gardées, nous préférons ne pas reprendre les classifications fonctionnelles comme la distinction narratif / descriptif ou procédural / argumentatif.

3. Aussi proposons-nous cette hypothèse : sur le plan sémantique, les *genres* sont définis par des *interactions normées* entre les composantes que nous venons de décrire. Par exemple, une notice d'entretien comme celle que nous reproduisons en annexe se caractérise comme suit.

(i) Une thématique fixe pour trois raisons : l'isotopie générique de champ est unique (et lexicalisée par le titre) ; les molécules sémiques correspondent à des lexicalisations à fort recouvrement (ex. *la cuve du filtre / la cuve*) ; ces molécules sont invariables.

(ii) Dans la dialectique, l'inventaire des acteurs est fixe (il n'y a pas création ni disparition d'acteurs). Ils sont présentés en trois groupes (les outillages et ingrédients, énumérés d'abord ; les parties de l'avion ; les agents présumés par l'impératif), chaque groupe correspondant à un type de rôles (adjuvants, objets, agents). Les fonctions sont réversibles (ex. *déposer / reposer, défreiner / freiner*). Leur succession est ordonnée fixement, comme le montre pour les indexer l'usage de chiffres et de lettres successifs.

(iii) La dialogique se caractérise par une prépondérance du factuel, le contrefactuel étant réservé aux notes (ex. : *ne jamais employer de produits chlorés*), le possible renvoyant à une procédure intercalaire facultative (*effectuer éventuellement [...]*). Les foyers énonciatif et interprétatif sont fixes.

(iv) La tactique ne correspond à aucune succession temporelle dans la première partie (*outillage et ingrédients*), mais à une succession stricte dans la seconde. On relève des inversions en forme de chiasme (ex. : 2. *Accéder à la soute arrière*. 3. *Déposer les panneaux d'accès aux racks hydrauliques*, inversé dans 9. *Reposer les panneaux d'accès aux racks hydrauliques*. 10. *Refermer la porte de la soute arrière*). Ces inversions sont liées aux phases initiales et finales des processus prescrits.

Les différentes formes de fixité que nous avons soulignées doivent être rapportées à des contraintes issues du domaine d'objectivité (la structure de l'avion impose un ordre des procédures), mais aussi à des formes rhétoriques de l'objectivation (ex : l'effacement de l'énonciateur).

Subsidiairement, les caractères particuliers de chaque texte sont définis par une interaction propre de ses composantes sémantiques, qui spécifie l'interaction définitoire du genre [15]. Seules la thématique et la tactique sont nécessaires dans tout texte : c'est le cas limite de l'énumération (encore que même une énumération de lexèmes puisse être organisée par des structures dialectiques ou dialogiques)[16]. Les autres interactions binaires attestées relient thématique et dialogique ou thématique et dialectique.

(i) La thématique d'un texte peut d'abord se décrire par son étendue, c'est-à-dire la part de l'univers sémantique mise en jeu, comme par ses restrictions de fait. Rapportée à la tactique, cette étendue est linéarisée en isotopies ou spécifiée en molécules sémiques, les unes comme les autres pouvant être caractérisées par leur position relative dans le texte.

L'étendue thématique peut varier au cours du texte. Par exemple, dans les résumés d'observation médicaux, la rubrique initiale qui rapporte les antécédents du patient comporte une thématique ouverte (souvent sans rapport avec l'affection, comme la profession, le lieu de naissance, ou l'âge des enfants) ; cette ouverture est sans doute un reliquat de la médecine hippocratique, qui considérait à bon droit le patient comme une personne digne d'intérêt pour elle-même. Dès le paragraphe suivant, l'étendue thématique se restreint drastiquement.

Rapportées à la dialectique, les molécules sémiques deviennent, par adjonction de traits casuels afférents, des acteurs ou des fonctions, voire — après homologation — des agonistes ou des séquences. Rapportées à la dialogique, les isotopies et molécules sémiques se placent dans des espaces modaux. Pour les textes narratifs, cela ouvre la possibilité de décrire des sphères possibles ou irréelles du récit. Mais encore, pour les textes descriptifs, cela permet de rendre compte des "points de vue" et des évaluations qui en sont inséparables.

(ii) La dialectique d'un texte peut être spécifiée par le type de processus qu'elle met en œuvre, notamment selon qu'ils sont réversibles ou non. Parmi les processus irréversibles, certains constituent des molécules sémiques nouvelles. Ainsi, alors qu'on classe les notices d'entretien et les recettes de cuisine parmi les textes procéduraux, les recettes seules créent par détermination et substitution de sèmes des "objets" nouveaux. L'évolution interne de l'univers sémantique décrit dépend ainsi des structures dialectiques du texte.

Les textes pratiques, du moins ceux qui décrivent des procédures, n'utilisent pas l'ellipse des

fonctions, puisqu'ils ont des objectifs didactiques. En revanche, les textes mythiques utilisent ces ellipses, dans la mesure où ils renforcent l'appartenance à une culture en présupposant la connaissance de ses règles. La succession stéréotypée des fonctions dans les syntagmes fonctionnels sert ici d'interprétant et permet de suppléer les fonctions attendues mais manquantes.

Sans en être définitoire, la structure dialectique peut être spécifique d'un genre. Par exemple, remarquait Zumthor, dans les chansons de toile médiévales, la structure narrative aurait pu être reprise dans des contes, mais il n'en a rien été.

Le couplage de la structure dialectique avec la thématique définit l'orientation dialectique : selon que les unités thématiques des premiers intervalles temporels sont mieux ou moins bien valorisées que celles des derniers, la morale des récits, et le sens de l'histoire pour la société considérée, sont optimistes ou pessimistes.

(iii) La dialogique a été bien étudiée pour les textes littéraires, mais fort peu pour les autres. En effet, la conception représentationnelle du langage conduit à considérer les textes scientifiques et techniques comme objectifs, d'autant plus qu'ils se règlent eux-mêmes sur cette conception et multiplient les rites d'objectivation. Rapportées à la thématique et à la dialectique, les variations dialogiques introduisent des dénivellations entre mondes et des ramifications du temps.

(iv) Rapportée aux autres composantes, la composante tactique permet de définir des rythmes thématiques (utilisés notamment en poésie), des rythmes dialectiques (mis à profit dans les arts du récit), des rythmes dialogiques (cf. les romans "polyphoniques", le théâtre). Hors de leur utilisation esthétique, ces rythmes jouent vraisemblablement un rôle dans le traitement sémantique, en favorisant d'une part la mémorisation, et d'autre part la reconnaissance de formes.

Ici se pose toutefois le problème herméneutique de la significativité des formes. Les chiasmes que l'on peut relever dans une notice d'entretien ou dans un manuel de sondage pétrolier sont simplement dus au fait que dans le processus décrit les opérations de remontage inversent celles du démontage. Leur symétrie est un gage de bon ordre : si une inversion finale venait à manquer, on oublierait par exemple des outils sur le chantier. Ces chiasmes sont par exemple formellement analogues à ceux des *Géorgiques* de Virgile, où ils soulignent les oppositions entre les saisons, et induisent en outre un effet de complétude et de clôture du texte. C'est donc évidemment le type de discours dont relève le texte qui détermine, par les pactes génériques de production et d'interprétation, la significativité des formes sémantiques.

Bref, des formes sémantiques comparables peuvent avoir des corrélats extralinguistiques différents. Si l'on prend l'exemple d'une corrélation entre un thème spécifique et une position tactique, elle peut être liée à une prescription esthétique, mais aussi à une contrainte factuelle : par exemple, dans les résumés d'observation sur le cancer de la thyroïde, le lévothyrox est toujours mentionné dans le dernier ou dans l'avant-dernier paragraphe, tout simplement parce que c'est le médicament qu'on donne après intervention.

5. Les interactions ternaires et quaternaires entre composantes pourraient à leur tour être explorées. Prenons l'exemple de la technique narrative du suspens. Un récit est constitué par l'interaction entre une structure thématique et une structure dialectique. Leur découplage avec la structure tactique se fait de deux manières codifiées : par prolepse, quand le narrateur dévoile au début quelle sera la fin du récit ; par analepse, quand il tait, souvent jusqu'à sa fin, le début du récit [17]. C'est en la matière ce qui oppose la tragédie au roman policier (où l'analepse correspond à deux stratégies dialogiques, celle du secret pour le narrateur, celle du dévoilement pour l'interprète). En revanche, un couplage avec la tactique qui ferait correspondre la succession des intervalles narratifs au déroulement du texte aurait déterminé d'autres stratégies de production et d'interprétation.

Il nous semble que les interactions des composantes sémantiques n'ont pas à être explorées *in abstracto*. Elles sont codifiées par les discours et les genres (dont chacun peut être définis comme des types d'interactions) ; en cela, elles relèvent de normes, évidemment culturelles. Aussi

n'entendons-nous pas formuler une typologie, mais en définir les critères.

Les genres mêmes doivent être étudiés au sein des discours et des pratiques sociales où ils prennent place. On peut trouver un intérêt théorique à comparer la structure dialectique de notices de montage et de recettes de cuisine, mais on ne peut négliger qu'elles ne relèvent pas du même discours, et ne sont pas interprétées ni appliquées de la même façon (les bons cuisiniers suivent leur inspiration). En outre, les discours mettent en jeu plusieurs genres, et il faut restituer la systématique de ces genres pour comprendre les spécificités de chacun. Comment par exemple, dans le discours juridique, étudier les réquisitoires sans les distinguer des plaidoiries ?

5. La conception morphosémantique du texte □

Les invariants de la description sont les fonds et les formes sémantiques qu'il faut répertorier en les reliant les uns aux autres : liens entre fonds (dans le cas par exemple des genres qui comportent plusieurs isotopies génériques), liens entre formes (dont il faut faire une description différentielle), et surtout liens des formes aux fonds (cruciaux pour l'étude de la perception sémantique). Ces morphologies sémantiques peuvent faire l'objet de diverses descriptions. Par exemple, rapportées aux quatre composantes, une molécule sémique peut être décrite comme thème, comme acteur, comme but ou source d'un point de vue modal, comme place dans la linéarité du texte. En outre, à chaque composante correspondent des types d'opérations productives et interprétatives.

La conception *morphosémantique* du texte a pour objectif de décrire les dynamiques de ces fonds et de ces formes ; par exemple, la construction des molécules sémiques, leur évolution, et leur dissolution éventuelle. Ces dynamiques et leurs optimisations sont paramétrées différemment selon les genres et les discours. D'une part, les formes et les fonds sont constitués et reconnus par rapport à des présomptions, et rapportés à des stéréotypes différents selon les pratiques sociales. En outre, les contrats de production et d'interprétation qui sont associés aux genres et aux discours codent le traitement de ces morphologies sémantiques.

□ Aller au chapitre 6. La lancinante question des dinosaures

NOTES

[1] La notion d'instruction est souvent utilisée pour l'analyse des textes littéraires, par exemple chez Eco, voire par les chercheurs qui se réclament de l'esthétique de la réception. La notion parfois avancée de *résolution de problèmes* (Adam, 1990, p. 114) appartient au même paradigme computationnel ou calculatoire.

[2] La définition empirique de Weinrich : " Un texte [...] peut être défini comme une suite signifiante de signes entre deux interruptions manifestes de la communication " (1974, p. 198) rompt certes avec l'écrit, mais repose sur la notion de communication. La notion d'interruption manifeste appelle des précisions ; par exemple, un dialogue doit être considéré comme un seul et même texte, car sa clôture dépend de normes d'interlocution propres au genre conversationnel dont il relève.

[3] Selon Slakta, le texte se définit ainsi : " Séquence bien formée de phrases liées qui progressent vers une fin " (1985, p. 138). Cette définition nous paraît susciter plusieurs questions ou objections : (i) Un texte ne se réduit pas à une séquence de phrases, fussent-elles liées. (ii) Qu'est-ce qu'une séquence bien formée ? La notion d'*expression bien formée* suppose des règles syntaxiques strictes, au sens logico-mathématique du terme. On se souvient du débat dans *Cognitive Science* au début des années 80 : il a montré, pour qui en doutait encore, que les grammaires de textes ne produisaient pas de règles de bonne formation au sens technique. (iii) Tendre vers une fin est, dans une perspective aristotélicienne, le propre des genres narratifs. On peut douter que ce soit une propriété de tout texte.

[4] Les études sur les enchaînements de phrases, quelle que soit leur inspiration (Ch. Bally, I. Bellert, O. Ducrot, L. Danlos) gardent un intérêt local, mais n'ont pas conduit à élaborer de théories du texte.

[5] La *discourse analysis* de Harris et de ses successeurs, dont l'école française d'analyse du discours, ont cherché à étendre au texte les méthodes de l'analyse distributionnelle, sans succès notable puisque les structures

textuelles sont essentiellement sémantiques.

[6] Ainsi, selon que l'on se trouve à une table de jeu du casino de Monaco, ou dans un car de retraités éblouis qui passe devant, les différences d'âge entre hommes et femmes se réduisent.

[7] Ils reposent en effet sur le préjugé que la langue est homogène et identique à elle-même dans tous les textes et dans toutes les situations de communication.

[8] Le thème est parfois défini comme l'élément le plus à gauche de la phrase (cf. Brown et Yule, 1983, p. 126) et souvent assimilé au "sujet grammatical" (p. 127). Cet ordre serait naturel : Levelt ressuscite à ce propos la notion d'*ordo naturalis* (1981), et van Dijk prétend que "l'ordre normal" irait du général au particulier, du tout aux parties, de l'ensemble aux éléments, du grand au petit, du dehors au dedans, du possesseur au possédé (cf. 1977, p. 106). Pour Katz, le *discourse topic* sera tout simplement le thème commun à une série de phrases, c'est-à-dire "le sujet de leur prédication" (1980, p. 26). Pour Perfetti et Goldman, le thème est le référent qui sert de "sujet central" à un discours" (1974, p. 71), c'est-à-dire généralement le personnage principal (*main character*). On voit que la confusion n'est pas mince, si l'on peut passer de l'élément de gauche de la phrase au héros de l'histoire.

[9] Utilisant la méthode comparative de la linguistique historique, Propp a relevé 31 fonctions dans un corpus de contes merveilleux. Elles se succèdent de manière fixe, même quand elles ne sont pas toutes présentes.

[10] Le concept d'*acteur* ainsi défini n'a pas de rapport privilégié avec la notion de personnage (pour une discussion, cf. Rastier, 1973, III, 2). Les personnages d'une épopée sont certes des acteurs, mais au même titre que les pièces et les outils d'une notice de montage.

[11] Quand Greimas étend à tous les textes la liste des personnages du conte populaire russe selon Propp, tout en l'assimilant à la liste des actants selon Tesnière, il tente là deux assimilations discutables. Les personnages du conte populaire russe sont sans doute spécifiques à une partie de l'aire indo-européenne. Et si Tesnière présente la phrase comme un petit récit, c'est par une métaphore didactique qui ne permet pas d'assimiler des niveaux de complexité différents. Les hypothèses universalistes sont certes stimulantes, mais l'universalité ne se décrète pas.

[12] Nous ne prétendons pas pour autant que la théorie logique des mondes possibles puisse s'appliquer à la sémantique des textes.

[13] Faute de pouvoir construire un univers de référence, apparaissent des situations ambiguës comme dans *À chacun sa vérité* (Pirandello).

[14] La validité de ces inférences dépend du discours et du genre. Elles sont permises par des énoncés normatifs implicites, comme l'a établi la théorie aristotélicienne des enthymèmes ou syllogismes incomplets.

[15] Nous rencontrons naturellement ici Adam (1993), quand il adapte à la définition du texte nos propositions pour la définition du genre.

[16] Cf. *Veni, vidi, vici* ; ou ce grognard de Napoléon qui avait traversé les guerres avec ces trois mots : *Brot, Kuss, Ehre!* (*pain, baiser, honneur!*, en évidente progression évaluative, du matériel au moral).

[17] Sur ces questions, voir notamment les travaux classiques de Gérard Genette, 1972, pp. 78 sq.

Vous pouvez adresser vos commentaires et suggestions à : Lpe2@ext.jussieu.fr

©  juin 2002 pour l'édition électronique.

Référence bibliographique : RASTIER, François. La macrosémantique. *Texto !* juin 2002 [en ligne]. Disponible sur : <http://www.revue-texto.net/Inedits/Rastier/Rastier_Macrosemantique1.html>. (Consultée le ...).