

GROUPEMENTS DE TEXTES ET CORPUS : POINT DE VUE DE LINGUISTE

Carine DUTEIL-MOUGEL
ATILF, Nancy-Université, CNRS

SOMMAIRE

1. Un groupement de textes *nommé corpus*...
 - 1.1. Textualité, intertextualité et architextualité
 - 1.2. Parcours plurisémiotiques
 - 1.3. L'inspiration linguistique
2. Corpus et textualité
 - 2.1. Corpus
 - 2.2. Typologies textuelles
- Conclusion

Résumé : *Trois situations de lecture des textes littéraires sont proposées aux élèves dans l'enseignement secondaire : au sein d'anthologies littéraires, dans le cadre de groupements de textes et dans le cadre d'une œuvre intégrale.*

Les groupements de textes rassemblent des extraits (scènes de pièces de théâtre, passages de romans ou de nouvelles, extraits d'essais, extraits de poèmes ; textes tirés d'un recueil : fables, lettres, nouvelles, poèmes) choisis et étudiés selon une cohérence thématique ou problématique. Leur étude prend place dans le cadre d'une séquence didactique, et doit permettre d'aborder un objet d'étude au programme.

L'une des épreuves écrites du Capes interne (depuis 2001) consiste justement à analyser un groupement de textes, accompagné très souvent d'un document iconographique, et à proposer une exploitation didactique de ce groupement sous la forme d'un projet de séquence. Notre réflexion portera sur ces « groupements », appelés de plus en plus souvent « corpus ». Nous nous interrogerons sur leur statut : s'agit-il de véritables CORPUS ? Quels critères président au choix de leur constitution ? Comment leur pertinence est-elle établie ?

L'étude d'exemples attestés (groupements au sein de séquences didactiques, groupements proposés lors de l'épreuve anticipée du baccalauréat de français, groupements proposés lors de l'« épreuve de didactique » au Capes interne de Lettres Modernes) nous conduira à analyser les relations qui s'établissent entre les éléments ainsi rassemblés au sein des groupements (liens thématiques, rapports d'interprétation, ...).

On s'interrogera sur les effets de la décontextualisation des extraits (incidence sur la textualité, sur les conditions d'interprétation), et on réfléchira aux moyens nécessaires à la constitution de corpus pleinement utilisables pour la caractérisation des textes.

Cette réflexion sur les « corpus » dans l'enseignement du français nous conduira à examiner plus en détail cette notion dans le cadre de la linguistique, et à présenter nos perspectives de recherche dans ce domaine (profilage et typologie des textes).

1. Un groupement de textes *nommé corpus*...

Trois situations de lecture des textes littéraires sont proposées aux élèves dans l'enseignement secondaire : au sein d'anthologies littéraires¹, dans le cadre de groupements de textes et dans le cadre d'une œuvre intégrale².

La notion de « groupement de textes » est récente dans l'enseignement du français ; elle apparaît pour la première fois dans le *Bulletin Officiel* du 7 juillet 1983. Son adoption est alors un moyen de renouveler les modalités de l'épreuve anticipée de français au baccalauréat. Comme le précise Michel Descotes (formateur IUFM) : « il s'agissait d'éviter l'évaluation des capacités de l'élève à

¹ Une anthologie présente des extraits d'œuvres « significatives » (textes canoniques, extraits exemplaires...). Elle s'apparente à l'encyclopédie, définie comme *archive de passages de textes décontextualisés* (cf. Rastier F., 2001, « Sémiotique et sciences de la culture », *Linx*, n°44-45, pp. 149-168).

² On entend par œuvre intégrale un roman, un essai, des mémoires, un recueil de poèmes, un recueil de nouvelles, une pièce de théâtre, etc. Il s'agit donc d'un texte intégral - et non d'une œuvre complète (*i.e.* l'ensemble des écrits d'un même auteur).

On soulignera que les élèves étudient beaucoup plus de groupements de textes que d'œuvres intégrales.

partir de la lecture d'un texte "à l'état de fragment isolé". Pour cela, il est préconisé de constituer la liste d'oral avec : - des oeuvres qui ont été lues dans leur intégralité ; - des groupements de textes choisis et étudiés selon une cohérence thématique ou problématique clairement formulée : par exemple, un groupe de poèmes permettant d'étudier la fuite du temps chez les poètes romantiques ou l'automne dans la poésie d'Apollinaire ; un choix de pages groupées pour une étude de la condition humaine selon Pascal et Voltaire ou de la critique de la société chez les philosophes du XVIIIème siècle, etc. ».

À l'heure actuelle la pratique du groupement s'est développée, et le groupement apparaît de plus en plus clairement comme *un instrument didactique* permettant de construire des savoirs et des savoir-faire.

« Réunissant un nombre limité mais suffisant de textes (quatre à cinq paraît une bonne moyenne), le groupement de textes constitue une unité d'apprentissage autour d'une notion littéraire, culturelle ou fonctionnelle. »

Programmes et accompagnement (enseigner au collège, Français), 2004 (rééd.), Paris, Centre National de Documentation Pédagogique.

Le groupement de textes rassemble des extraits (scènes de pièces de théâtre, passages de romans ou de nouvelles, extraits d'essais, extraits de poèmes ; textes tirés d'un recueil : fables, lettres, nouvelles, poèmes) choisis et étudiés selon une cohérence thématique ou problématique. Son étude prend place dans le cadre d'une séquence didactique¹, et doit permettre d'aborder un *objet d'étude* au programme.

Par exemple, au collège, les groupements de textes permettent d'étudier les différentes *formes de discours* au programme² – l'objectif majeur des nouveaux programmes pour l'enseignement du français au collège étant la maîtrise des discours³.

▼ RACONTER À LA PREMIÈRE PERSONNE		
R. MERLE, <i>Madrapour</i> ; A. DUMAS, <i>La Dame aux camélias</i> ; R. WRIGHT, <i>Black Boy</i>		
▼ IMPLIQUER LE LECTEUR		
I. CALVINO, <i>Si par une nuit d'hiver...</i> ; M. BUTOR, <i>La Modification</i>		
▼ CHOISIR ET VARIER LES POINTS DE VUE	▼ GROUPEMENT DE TEXTES : La violation des Droits de l'Homme	
A. TCHEKOV, <i>La Dame au petit chien</i> ; STENDHAL, <i>Vanina Vanini</i> ; A. ROBBE-GRILLET, <i>Dans le labyrinthe</i>	VOLTAIRE, <i>Candide ou l'Optimiste</i> ; V. HUGO, <i>Bug-Jargal</i> ; V. SCHELCHER, article « <i>Esclavage, Esclavage</i> » ; B. HOPQUIN, « Les enfants de la balle au Pakistan » (<i>Le Monde</i> , 1998) ; Débat rapporté par A. COJEAN (<i>Le Monde</i> , 1998)	Rédiger une argumentation Lire au CDI
G. FLAUBERT, <i>Madame Bovary</i> ; É. ZOLA, <i>Une page d'amour</i>		

Fig. 1 : Extraits d'un manuel de Français pour la classe de troisième, Hatier, 1999, pp. 6-7
à Gauche : « Les formes de discours, 1. Raconter, décrire »
à Droite : « Les formes de discours, 2. Argumenter convaincre »

Notons que lesdites *formes de discours* renvoient aux *séquences textuelles* définies par la linguistique textuelle⁴.

¹ Les nouveaux programmes favorisent le travail par séquences didactiques. On désigne par là « un mode d'organisation des activités qui rassemble des contenus d'ordre différent autour d'un même objectif, sur un ensemble de plusieurs séances ».

² « **Choix théorique : le discours**

Le premier objectif de l'enseignement du français est de "donner aux élèves la maîtrise des principales formes de discours" (*BO* n° 10, 15 octobre 1998, p. 4). », *Programmes et accompagnement* (enseigner au collège, Français), 2004.

³ « Le français se fixe comme objectif central au collège la maîtrise des discours. », *ibid.*

⁴ Cf. notamment Adam J.-M., 1992 et 1999.

« **Formes de discours.** Suivant les finalités de l'énonciation, les discours adoptent des dominantes différentes : c'est en ce sens qu'on parle ici de discours narratif, descriptif, explicatif, argumentatif, pour les formes les plus évidentes.

– Le discours narratif rapporte un ou des événements et les situe dans le temps.

– Le discours descriptif vise à nommer, caractériser, qualifier.

– Le discours explicatif cherche à faire comprendre.

– Le discours argumentatif valorise un ou plusieurs points de vue, une ou plusieurs thèses.

On peut trouver diverses formes de discours dans un même texte. »

Glossaire, *Programmes et accompagnement* (enseigner au collège, Français), 2004 (rééd.), Paris, Centre National de Documentation Pédagogique

La progression d'ensemble¹, de la 6^{ème} à la 3^{ème}, s'organise ainsi autour de deux pôles : le pôle narratif et le pôle argumentatif, que nous représentons comme suit :

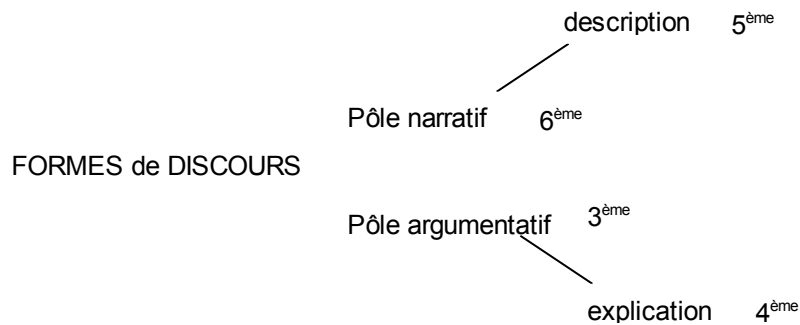


Fig. 2 : Etude des formes de discours au collège

L'appellation « corpus » pour de tels groupements apparaît à plusieurs reprises dans les nouveaux programmes pour le secondaire. Cette appellation figure également dans les modalités et les libellés des sujets de l'épreuve anticipée (écrite) de français au baccalauréat :

« Épreuve écrite

Les sujets prennent appui sur un ensemble de textes (corpus) distribués au candidat, éventuellement accompagnés par un document iconographique si celui-ci contribue à la compréhension ou enrichit la signification de l'ensemble. »

Note de service n° 2001-117 du 20 juin 2001. BO n° 26 du 28 juin 2001

Modifiée par la note de service n° 2003-002 du 8 janvier 2003. BO n° 3 du 16 janvier 2003

Les sujets invitent les candidats à s'interroger sur l'unité du corpus et à réfléchir au rapprochement des extraits choisis (cf. l'exemple ci-dessous). Il peut s'agir d'un thème commun, d'un genre commun, d'un registre commun, d'un dispositif énonciatif commun, de l'appartenance à un même mouvement littéraire, etc.

***SUJET : Série ES, S – Juin 2003**

Objet d'étude : La poésie

CORPUS

Texte A. Paul Verlaine, « Mon rêve familial », *Poèmes saturniens*, 1866.

Texte B. Robert Desnos, « J'ai tant rêvé de toi », in « À la mystérieuse », *Corps et Biens*, 1930.

Texte C. Paul Eluard, « La dame de carreau », *Les Dessous d'une vie*, 1926.

Texte D. Claude Roy, « Tant », *Le Voyage d'automne*, 1987.

I. Question

Justifiez le rapprochement des quatre poèmes.

II. Ecriture : au choix

1. Commentaire : Vous ferez le commentaire du poème de Robert Desnos (texte B.)

¹ Les formes de discours sont progressivement plus complexes (dominantes, insertions, textes mêlant narratif et argumentatif, etc.). L'étude du récit est privilégiée en 6^{ème}. En 5^{ème} l'accent est mis sur la description ; en 4^{ème}, sur l'explication. Les principales formes d'argumentation sont étudiées en 3^{ème}.

2. Dissertation : Pour Eluard, le poète « aimant l'amour » n'est pas tant amoureux d'une femme que de l'amour lui-même.

La vocation de la poésie est-elle, selon-vous, de célébrer l'amour ou privilégiez-vous d'autres fonctions ?

Vous vous appuyerez pour répondre à cette question, sur les textes du corpus et les poèmes que vous avez lus et étudiés.

3. Invention : Dans la préface d'une anthologie de poèmes d'amour que vous avez réunis, vous démontrez comment l'inspiration poétique et l'amour sont à vos yeux liés.

Rédigez la préface.

Vous devez nourrir votre texte de citations de poèmes et de références à des auteurs.

La nouvelle épreuve écrite du Capes interne de Lettres Modernes (depuis 2001) – l'« épreuve de didactique » – consiste également à analyser un corpus, accompagné d'un **document iconographique** (cf. les deux sujets ci-dessous), et à proposer une exploitation didactique de ce corpus sous la forme d'un projet de séquence.

« Epreuve de didactique de la discipline. Un corpus de textes éventuellement accompagné de documents iconographiques est proposé aux candidats.

Ceux-ci, dans un devoir rédigé et argumenté :

- analysent les textes, en fonction d'une problématique indiquée par le sujet ;

- proposent une exploitation didactique de ces textes, sous la forme d'un projet de séquence destinée à la classe de collègue ou de lycée indiquée par le sujet. Il appartient au candidat de déterminer l'objectif qu'il fixe à sa séquence. Une séance d'étude de la langue est obligatoirement comprise dans cette séquence.

Durée de l'épreuve : Six heures ; coefficient : 1.

Le programme des épreuves est celui des lycées d'enseignement général et technologique et des collèges. »

Extrait de l'Arrêté du 2 mars 2000

***Exemple 1 : session 2001**

SUJET

Vous analyserez les documents joints dans la perspective de l'étude du récit autobiographique dans une classe de 3^e, en vous attachant à la diversité des approches et des projets de chaque auteur. Vous en proposerez une exploitation didactique sous forme d'un projet de séquence, incluant une séance d'étude de la langue et des prolongements par des lectures cursives.

Documents extraits de¹ :

- Chateaubriand, *Mémoires d'Outre-Tombe* [Livre premier, de « Il y a quatre ans qu'à mon retour de terre sainte, j'achetai près du hameau d'Aulnay... » à « ... Et enfin l'Histoire des grands officiers de la couronne du P. Anselme. »]

- M. Yourcenar, *Archives du Nord* [éd. Gallimard, 1977, de « Dans un volume destiné à former avec celui-ci les deux panneaux d'un diptyque... » à « Survolons-le à une époque où il était encore sans habitants et sans nom. »]

- P. Néruda, *J'avoue que j'ai vécu* [éd. Gallimard, traduction de Claude Couffon, 1975, de « Je dirai pour commencer cette évocation des jours et des années de mon enfance que le seul personnage... » à « La vie était dure pour les petits agriculteurs du centre du pays. »]

- A. Nothomb, *Métaphysique des tubes* [éd. Albin Michel, 2000, de « En me donnant une identité, le chocolat blanc... » à « On commença à m'appeler par mon prénom. »]

- R. Queneau, *Chêne et chien* [1937, De « Je naquis au Havre un vingt et un février » à « je n'eus jamais de sœur. »]

Iconographie : P. Gauguin, Autoportrait au christ

***Exemple 2 : session 2006**

SUJET

Dans le cadre de l'étude de la poésie en classe de troisième, vous analyserez le corpus ci-joint en vous attachant à la diversité des démarches et des écritures.

¹ Notons qu'en 2001, il n'était pas encore question de CORPUS.

En vous fondant sur cette analyse, vous proposerez une exploitation didactique clairement problématisée tenant compte des objectifs d'étude fixés pour la classe de troisième. Votre séquence comportera obligatoirement une séance d'étude de la langue. Vous pouvez enrichir votre projet de textes ou de références complémentaires.

CORPUS :

Victor Hugo, "Fonction du poète", *Les rayons et les ombres* (extrait), 1840.

Tristan Corbière, "Le crapaud", *Les Amours jaunes*, 1873.

Léon-Paul Fargue, *Haute solitude*, extrait d'"Azazel", 1941.

Max Jacob, *Derniers poèmes en vers et en prose*, "Amour du prochain", 1944.

Yves Bonnefoy, *Ce qui fut sans lumière*, "La tâche d'espérance", 1987.

Philippe Jaccottet, *Leçons*, "Autrefois...", in *Poésies 1946-1967*.

Jacques Réda, *Amen*, "Espère et tremble", 1968.

Pablo Picasso, *Femme en pleurs*, 1937.

[Reproduction tirée de Carsten-Peter Warncke, *Picasso*, Taschen, 1988]

1.1. Textualité, intertextualité et architextualité

Mais ces groupements de textes peuvent-ils prétendre au statut de corpus ?

La question du statut des groupements se pose dans la mesure où les critères qui président au choix de leur constitution ne sont pas explicités, et dans la mesure où la pertinence du regroupement des textes n'est pas véritablement discutée. Tout regroupement de textes ne mérite pas le nom de corpus, et on pourrait s'interroger notamment sur l'hétérogénéité des groupements - auteurs, œuvres, époques...

De plus, les textes s'apparentent davantage à des *morceaux choisis* (extraits, éléments d'un recueil, etc.). Se pose alors le problème de leur décontextualisation (texte - *le contexte, c'est tout le texte* - et corpus d'origine - *tout texte est situé dans son intertexte*) et de leur recontextualisation au sein du groupement.

Lorsqu'un texte change de corpus, il change inévitablement de sens, selon le principe d'architextualité : *tout texte placé dans un corpus en reçoit des déterminations sémantiques et modifie potentiellement le sens de chacun des textes qui le composent*¹. L'interprétation² étant conditionnée par le *corpus*, l'intertexte nouveau crée ainsi des relations d'interprétance mutuelle et invite à de nouveaux parcours interprétatifs (liens thématiques, isotopies, corrélations) qu'il s'agit de pouvoir légitimer.

1.2. Parcours plurisémiotiques

L'intégration du document iconographique au corpus pose également des difficultés ; son exploitation en lien avec le corpus textuel relève parfois de l'artefact. Les programmes scolaires mettent pourtant l'accent sur l'étude de l'image³, qui s'inscrit dans le cadre de l'apprentissage des formes de discours⁴.

« Fixe ou mobile, l'image est envisagée comme discours : elle raconte et décrit, mais elle a aussi un rôle explicatif ou argumentatif. »

Programmes et accompagnement (enseigner au collège, Français), 2004 (rééd.), Paris, Centre National de Documentation Pédagogique.

Il s'agit par exemple de conduire les élèves :

¹ Cf. Rastier F. 2001, p. 92.

² L'interprétation est toutefois limitée puisqu'il s'agit de privilégier la construction d'une séquence didactique (en fonction du niveau de classe, et en suivant les indications et l'orientation du sujet) ; et que le grand nombre d'extraits condamne au butinage textuel.

³ « Les enseignants de français, bien que non professionnels de l'image, sont amenés à intégrer dans leur enseignement la dimension visuelle, qui imprègne de plus en plus profondément la formation culturelle et les pratiques quotidiennes de leurs élèves. Afin de les conduire progressivement à une approche raisonnée de l'image, les professeurs travaillent en particulier sur les relations entre le langage verbal et le langage visuel, dans la perspective du discours. » *Programmes et accompagnement* (enseigner au collège, Français), 2004.

⁴ « l'étude de l'image peut très facilement se développer à l'intérieur d'un groupement de textes et de documents, en particulier si l'on envisage les différentes formes d'argumentation rendues possibles par les langages non verbaux. » *ibid.*

« à reconnaître, à identifier et à analyser les différentes fonctions discursives de l'image : esthétique, informative, descriptive, narrative, explicative, persuasive, critique et plus largement argumentative. » (*ibid.*)

Les manuels s'efforcent ainsi d'associer une image aux textes rassemblés. Mais la mise en relation texte(s)-image et la création de parcours interprétatifs plurisémiotiques est parfois problématique, comme l'illustre l'extrait de manuel qui figure en annexe.

L'image revêt alors très souvent une fonction illustrative, ce à quoi la réduit d'ailleurs le descriptif des modalités de l'épreuve anticipée (écrite) de français au baccalauréat :

« Le document iconographique, s'il est joint au corpus, ne peut servir que de support. En aucun cas il ne sera demandé d'en faire une étude pour lui-même »

Note de service n° 2001-117 du 20 juin 2001. BO n° 26 du 28 juin 2001

Modifiée par la note de service n° 2003-002 du 8 janvier 2003. BO n° 3 du 16 janvier 2003

1.3. L'inspiration linguistique

Les programmes du français dans le secondaire sont sur de nombreux points sous-tendus par des apports de la linguistique¹ et il n'est pas exclu – cela est même fort probable – que l'utilisation récente du mot « corpus » dans les programmes soit liée au regain d'intérêt des linguistes pour cette notion² et à l'émergence depuis les dix dernières années de la *linguistique de corpus*.

Après nous être intéressée aux « corpus » de l'enseignement du français, nous examinerons plus en détail cette notion dans le cadre de la linguistique et nous présenterons nos perspectives de recherche dans ce domaine (profilage et typologie des textes).

2. Corpus et textualité

2.1. Corpus

La notion de *corpus* reçoit en linguistique différentes acceptions. Le corpus peut être défini comme *une collection de données langagières* ou comme *un échantillon de langage* ; il peut également être conçu comme *un ensemble de mots*, ou comme *un ensemble d'énoncés*, ou encore, dans notre perspective, comme *un ensemble structuré de textes*.

Mais tout ensemble de textes n'est pas un corpus. Pour B. Bommier-Pincemin (1999 : 416), le corpus doit vérifier trois types de conditions :

« • *Des conditions de signifiante* : Un corpus est constitué en vue d'une étude déterminée (*pertinence*), portant sur un objet particulier, une réalité telle qu'elle est perçue sous un certain angle de vue (et non sur plusieurs thèmes ou facettes indépendants, simultanément) (*cohérence*).

• *Des conditions d'acceptabilité* : Le corpus doit apporter une représentation fidèle (*représentativité*), sans être parasité par des contraintes externes (*régularité*). Il doit avoir une

¹ Pour ne citer que quelques exemples :

« Un futur professeur ne saurait ignorer que l'ensemble des programmes et instructions officielles s'inscrit dans un cadre théorique de référence, très précisément défini et illustré par les documents d'accompagnement des programmes du collège et du lycée. », *Rapport officiel du jury Capes interne et Caer de Lettres Modernes*, 2005 ;

« Le programme fixe comme objectif de fin de 3^e la maîtrise des principales formes du discours.

On entend par **discours** toutes les mises en pratique du langage à l'oral et à l'écrit (voir *Glossaire*). La conception générale du programme se fonde donc sur l'énonciation, définie comme l'actualisation de la langue dans des situations concrètes d'utilisation. peu à peu à l'analyse des discours à dominante argumentative. » *Programmes et accompagnement* (enseigner au collège, Français), 2004 ; « Au collège, l'étude de la langue n'est pas une fin en soi, mais elle est subordonnée à l'objectif de la maîtrise des discours. Elle se fonde donc sur la prise en compte des situations de communication. » *ibid.* ;

« Pour ce qui concerne les catégories grammaticales, il convient de procéder à une révision et à un approfondissement systématique des notions fondamentales en grammaire de phrase, en grammaire de texte et en grammaire du discours.

Les développements contemporains de la pragmatique du discours (énoncé, énonciateur, situation d'énonciation, implicite et sous-entendu, etc.) nécessitent un apport de formation tout particulier. », *Rapport officiel du jury Capes interne et Caer de Lettres Modernes*, 2003.

² En témoigne également le colloque qui nous réunit : « Corpus en Lettres et Sciences sociales - Des documents numériques à l'interprétation ».

ampleur et un niveau de détail adaptés au degré de finesse et à la richesse attendue en résultat de l'analyse (*complétude*).

• *Des conditions d'exploitabilité* : Les textes qui forment le corpus doivent être commensurables (*homogénéité*). Le corpus doit apporter suffisamment d'éléments pour pouvoir repérer des comportements significatifs (au sens statistique du terme) (*volume*). »

Nous suivons F. Rastier (2005 : 32), qui propose la définition suivante du *corpus* (c'est nous qui soulignons) :

« Un corpus est un regroupement **structuré de textes intégraux, documentés, éventuellement enrichis par des étiquetages, et rassemblés** : (i) de manière théorique réflexive **en tenant compte des discours et des genres**, et (ii) de manière pratique en vue d'une gamme d'applications. »

Nos travaux s'inscrivent dans la perspective d'une linguistique textuelle capable de penser la diversité des textes, et privilégiant l'étude du sens textuel. Nous adoptons une conception praxéologique (praxéologie entendue comme *théorie de l'action dans et par le langage*) selon laquelle tout texte procède d'un genre, et tout genre est relatif à un discours (politique, littéraire, scientifique, etc.), attaché à une pratique sociale¹.

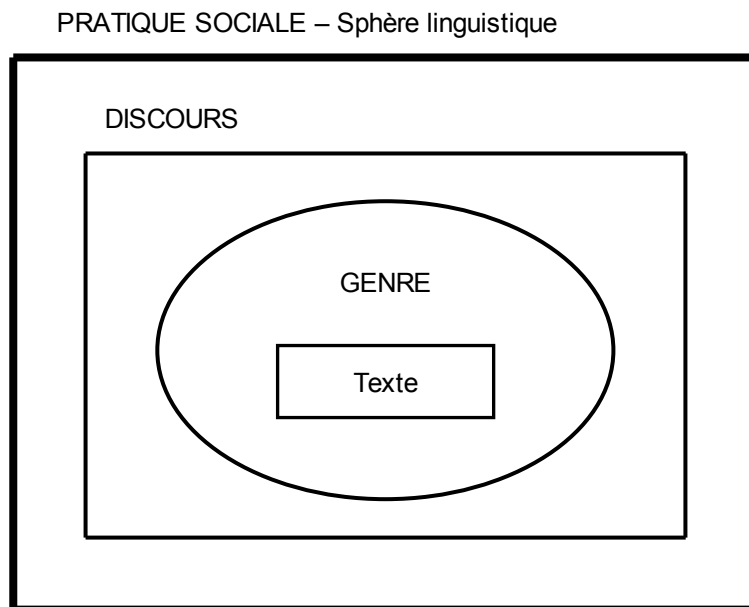


Fig. 3 : Conception praxéologique

Les textes sont ainsi étudiés dans leur production et leur interprétation, et sont reliés à leur entour social et historique.

Il s'agit notamment de déterminer l'incidence du genre textuel, du discours d'appartenance ainsi que de l'intertexte (corpus de référence) sur les unités linguistiques locales. Déterminations du global sur le local que l'on peut représenter ainsi :

¹ Cf. notamment Rastier F., 2001.

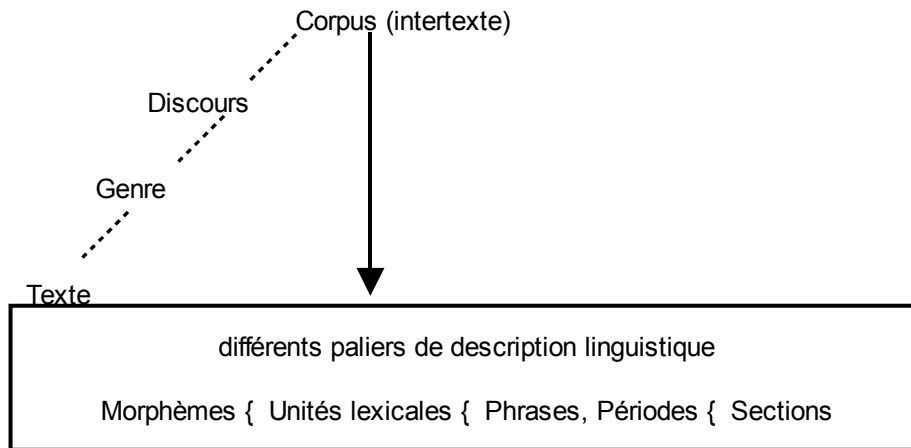


Fig. 4 : Déterminations du global sur le local (principe herméneutique)

2.2. Typologies textuelles

Comme les Traitements automatiques du langage ont affaire à des textes, non à des phrases, leur typologie est une condition de leur analyse. La diversité des genres et des discours amène à s'interroger sur l'importance de ces normes. Il apparaît ainsi nécessaire, dans la perspective d'établir des typologies textuelles¹, de tenir compte des genres et des discours.

L'étude des corpus montre par exemple que le lexique, la morphosyntaxe varient avec les genres et les discours. Denise Malrieu et François Rastier (2001) ont montré leur incidence sur l'ensemble des catégories morphosyntaxiques, ainsi que sur des variables comme la longueur des mots et des phrases. Aussi, disposer de critères et d'outils pour la classification des genres et des discours représente un enjeu pour la linguistique de corpus et la constitution de corpus pleinement utilisables (structuration des bases textuelles) pour des tâches de description linguistique.

Nous proposons de définir le genre par la cohésion d'un faisceau de critères linguistiques. Étiqueter le corpus avec des catégories morpho-syntaxiques ne suffit pas à faire émerger des genres. Les genres sont définis par des interactions normées entre diverses composantes et par des critères de corrélation entre ces composantes et leurs constituants :

lexico-syntaxiques, sémantiques, sémantico-rhétoriques

Les genres ne sont pas isolés mais contrastent au sein de champs génériques². On ne caractérise donc pas les genres en soi mais par rapport aux autres genres, à commencer par les plus proches, à savoir ceux relevant d'un même champ générique³.

Le schéma suivant illustre les relations intergenres au sein d'un même champ générique. Le sous-corpus 1 rassemble des textes appartenant au 'genre X'. Le sous-corpus 2 rassemble des textes appartenant au 'genre Y'. 'Genre X' et 'Genre Y' contrastent au sein d'un même //champ générique//, le champ générique A.

¹ Il ne s'agit pas de classer des types (fonctionnels) de textes ou d'isoler des séquences textuelles (narrative, descriptive, explicative, argumentative, etc.). Un texte ne se réduit pas à une étendue de plusieurs phrases ou propositions, et les unités textuelles ne sont pas nécessairement discrètes ; elles sont davantage à envisager comme des lieux et moments de parcours énonciatifs et interprétatifs. Cf. Rastier F., 2003.

² Cf. Rastier F., 2001, Glossaire p. 297.

« *Champ générique* : groupe de genres qui contrastent voire rivalisent dans un champ pratique : par exemple, au sein du discours littéraire, le champ générique du théâtre se divisait en comédie et tragédie ; au sein du discours juridique, les genres oraux constituent un champ générique propre (réquisitoire, plaidoirie, sentence). ».

³ Le regroupement des genres en champs génériques fait appel à des critères situationnels (comme le contexte de production/de réception, les objectifs et visées, le support médiatique, le mode de diffusion, etc.). Chaque champ générique correspond à un champ pratique ; une connaissance de la pratique sociale paraît nécessaire pour effectuer les regroupements.

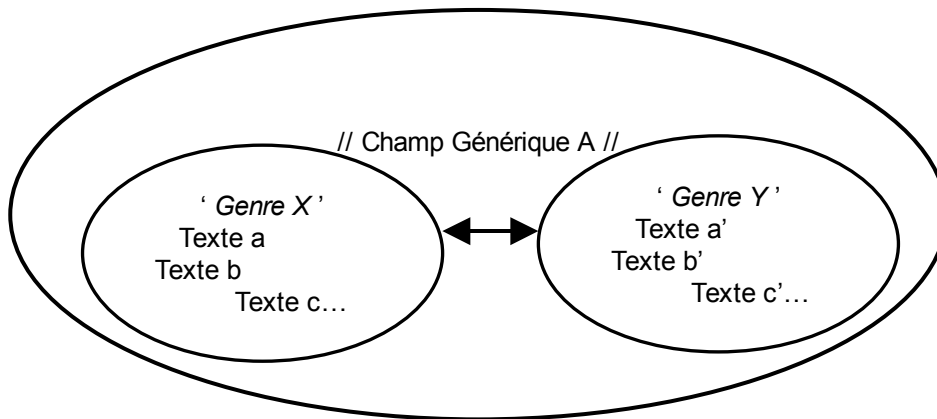


Fig. 5 : Relations contrastives entre genres au sein d'un même champ générique

Les champs génériques sont ainsi envisagés comme des classes d'interdéfinition des genres. Il s'agit alors de proposer une caractérisation différentielle des genres et d'établir les traits spécifiques de chacun d'eux.

Nous faisons l'hypothèse que les procédés sémantico-rhétoriques¹ jouent un rôle important dans la caractérisation des genres. Pour permettre leur traitement, nous cherchons actuellement à identifier des marqueurs de procédés sémantico-rhétoriques et à automatiser leur balisage.

Conclusion

L'étude du statut des « groupements de textes » et de la place qu'ils occupent dans l'enseignement du français (dans le secondaire) nous a conduit à problématiser la notion de « corpus » en linguistique et à proposer des critères pour la caractérisation des genres et la typologie des textes.

En lien avec cette problématique, une étude sur l'exploitation des corpus numériques en classe mériterait d'être menée. Il s'agirait d'analyser les apports du travail sur corpus et de s'interroger sur les applications qui peuvent être proposées : recherche par mots-clés ou mots-vedettes ? ; étude de corrélations ? ; avec quels outils et suivant quels objectifs ?

On pourra notamment s'appuyer sur les nombreuses analyses de corpus réalisées par Thierry Mézaille, dans la perspective d'une *sémantique textuelle, littéraire et didactique*².

BIBLIOGRAPHIE

- ADAM, J.-M. 1992. *Les textes : types et prototypes*, Paris, Nathan.
- ADAM, J.-M. 1999. *Linguistique textuelle. Des genres de discours aux textes*, Paris, Nathan.
- ARMAND, A., DESCOTES, M. et al. 1992. *La séquence didactique en français : classes de lycée, CAPES, agrégation*, Paris, Bertrand Lacoste.
- BIARD, J. & DENIS, F. 1993. *Didactique du texte littéraire. Progressions et séquences*, Paris, Nathan.
- BIBER, D. 1988. *Variations across Speech and Writing*, Cambridge, Cambridge University Press.
- BIBER, D. 1995. *Dimensions of register variation : a cross-linguistic comparison*. Cambridge University Press, Cambridge.
- BOMMIER-PINCEMIN, B. 1999. *Diffusion ciblée automatique d'informations : conception et mise en œuvre d'une linguistique textuelle pour la caractérisation des destinataires et des documents*, Thèse de l'Université Paris IV, 805 p.
- BRONCKART, J.-P. 1985. *Le fonctionnement des discours*, Neuchâtel ; Paris, Delachaux et Niestlé.
- COMBETTES, B. 2002. Analyse linguistique des textes et stylistique, *Langue française*, 135, pp. 95-113.

¹ Les procédés sémantico-rhétoriques renvoient notamment aux composantes de l'*invention-élocution* (composantes éthique, argumentative, pathétique). Pour une présentation, cf. Duteil-Mougel C., 2005.

² Cf. le site personnel de l'auteur : <http://www.chez.com/mezaille/>.

- CONDAMINES, A. (dir.) 2005. *Sémantique et corpus*, Paris, Hermès.
- DUTEIL-MOUGEL C. 2005. Les mécanismes persuasifs des textes politiques. Propositions théoriques pour l'analyse de corpus, *Corpus*, 4, pp. 129-155.
- HABERT, B. 2006. Portrait de linguiste(s) à l'instrument, *Texte ! Textes et cultures*, Vol. X, n°4.
- HABERT, B., NAZARENKO, A. & SALEM, A. 1997. *Les linguistiques de corpus*, Paris, Armand Colin - Masson.
- HABERT, B. et coll. 2000. Profilage de textes : cadre de travail et expérience, in *Actes des 5èmes JADT*.
- IDE, N. et VÉRONIS, J. 1995. *Text Encoding Initiative - Background and Context*, Dordrecht (The Netherlands), Kluwer Academic Publishers.
- ILLOUZ, G., HABERT, B. et al. 1999. Maîtriser les déluges de données hétérogènes, in A. Condamines et al. (éds.), *Corpus et traitement automatique des langues : pour une réflexion méthodologique*, Actes de l'atelier thématique TALN, Cargèse, pp. 37-46.
- JORDY, J. 1995. Groupement de textes poétiques – Une séquence autour du blason, in M. Descotes et al., *Lire méthodiquement des textes*, Paris, Bertrand Lacoste.
- JORDY, J. 1991. *Le groupement de textes*, Toulouse, CRDP.
- Le Français aujourd'hui*, « Le groupement de textes », n°97, Mars 1992.
- MALRIEU, D. & RASTIER, F. 2001, Genres et variations morphosyntaxiques, *Traitement Automatique des langues*, vol. 42, 2, pp. 548-577.
- MAYAFFRE, D. 2002. Les corpus réflexifs : entre architextualité et intertextualité, *Corpus*, I, 1, pp. 51-70.
- MEZAILLE, T. 2003. *Thématiques littéraires - enseignement des textes numériques pour un accès sémantique et didactique aux banques textuelles*, HDR.
- RASTIER, F. 2001. *Arts et sciences du texte*, Paris, PUF.
- RASTIER, F. 2003. Parcours de production et d'interprétation, in A. Ouattara (éd.), *Parcours énonciatifs et parcours interprétatifs*, Ophrys, pp. 221-242.
- RASTIER, F. 2005. Enjeux épistémologiques de la linguistique de corpus, in G. Williams (éd.), *La linguistique de corpus*, Rennes, PU Rennes, pp. 31-45.
- RASTIER, F. à paraître. La traduction : interprétation et genèse du sens.
- SUEUR, J.-P. 1982. Pour une grammaire du discours, *Mots*, 5, pp. 145-185.
- VECK, B. 1995. *Groupements de texte et projet de lecture*, (2 vol.), Paris, Bertrand Lacoste.

Documents officiels :

***Programmes et accompagnement* (enseigner au collège, Français), 2004 (rééd.), Paris, Centre National de Documentation Pédagogique.

*Organisation des enseignements dans les classes de 6^{ème} des collèges : Arrêté du 29 mai 1996 – (BO n° 25 du 20 juin 1996), modifié par l'arrêté du 14 janvier 2002 (BO n° 8 du 21 février 2002) ; Arrêté du 22 novembre 1995 relatif aux programmes de la classe de 6^{ème} des collèges.

*Programme du cycle central des collèges : Arrêté du 10 janvier 1997. JO du 21 janvier 1997 – (BO Hors série n° 1 du 13 février 1997).

*Programme de la classe de 3^{ème}, BOHS n° 10 du 15 octobre 1998.

***Français (classe de seconde)*, 2003 (rééd.), Paris, Centre National de Documentation Pédagogique.

*Programme de la classe de seconde générale et technologique.

BO n° 41 du 7 novembre 2002.

***Français (classe de première), Littérature (classe terminale série littéraire)*, 2004 (rééd.), Paris, Centre National de Documentation Pédagogique.

*Programme de français des classes de première des séries générales et technologiques.

BO n° 28 du 12 juillet 2001.

**Epreuves anticipées de français

Note de service n° 2001-117 du 20 juin 2001. BO n° 26 du 28 juin 2001.

Modifiée par la note de service n° 2003-002 du 8 janvier 2003. BO n° 3 du 16 janvier 2003.

**Epreuve d'admissibilité au Capes interne de Lettres Modernes

Arrêté du 2 mars 2000 modifiant l'arrêté du 30 avril 1991 modifié fixant les sections et les modalités d'organisation des concours du certificat d'aptitude au professorat de l'enseignement du second degré.

**Rapport officiel du jury Capes interne et Caer de Lettres Modernes, 2005, présenté par Madame Catherine BECCHETTI-BIZOT, Inspectrice générale de l'éducation nationale, Présidente du jury, Paris, Centre National de Documentation Pédagogique.

**Rapport officiel du jury Capes interne et Caer de Lettres Modernes, 2004, présenté par Madame Catherine BECCHETTI-BIZOT, Inspectrice générale de l'éducation nationale, Présidente du jury, Paris, Centre National de Documentation Pédagogique.

**Rapport officiel du jury Capes interne et Caer de Lettres Modernes, 2003, présenté par Madame Catherine BIZOT, Inspectrice générale de l'éducation nationale, Présidente du jury, Paris, Centre National de Documentation Pédagogique.

ANNEXE

Extrait d'un manuel de Français pour la classe de troisième, Hatier, 1999, p.20



Vieira Da Silva (1908-1992),
Composition, 1951 (huile sur
toile; Bâle, Kuntmuseum,
Offentliche Kunstsammlung).

RACONTER À LA PREMIÈRE PERSONNE

Leçon p. 54

J'écris cette histoire

TEXTE 1

Robert Merle,
écrivain français,
né en 1908.

Madrapour,
roman, 1976.

Il s'agit ici du début du roman.

13 novembre.

J'écris cette histoire en même temps que je la vis. De jour en jour. Ou plutôt – ne soyons donc pas si ambitieux – d'heure en heure. Nous serions bien avisés, d'ailleurs, de faire tenir un monde dans chaque minute qui passe. Nous n'en avons pas tant à notre disposition. La vie la plus longue peut se chiffrer en secondes. Faites le calcul : c'est un chiffre qui n'a rien d'astronomique – ni de particulièrement rassurant.

Tandis que j'écris ceci, je suis bien incapable de prévoir la fin de mon aventure. Je ne pénètre pas non plus sa signification. Ce n'est pourtant pas faute de hasarder des hypothèses.

Mon histoire aura sûrement un terme, ne serait-ce que le plus évident. Mais il n'est pas certain qu'elle ait un sens ou – ce qui revient au même – il n'est pas sûr que je sois capable de lui en trouver un : « Un moucheron qui naît à l'aube et meurt au coucher du soleil ne peut pas comprendre le mot nuit. »

Quand le taxi me dépose à l'aéroport de Roissy-en-France, une surprise m'attend. Tout est vide. Ni voyageurs, ni employés, ni hôtesses. Je suis seul, rigoureusement seul, dans ce monument de métal et de glace, où règne un silence de crypte. Comparaison absurde : Roissy ressemblerait plutôt, avec ses immenses vitres, à une serre démesurée.

Je pose mes valises sur un chariot et, dans ce désert lumineux, je pousse le chariot devant moi. En même temps, je sens tout le ridicule de convoier ainsi mes biens terrestres, alors qu'aucun préposé ne peut les prendre en charge.

Question posée aux élèves :

« Lire l'image

Observez l'ensemble des lignes, des figures, les tonalités dominantes dans le tableau de la page 20. Quel lieu ce tableau peut-il évoquer ? Citez une phrase du texte qui pourrait servir de légende à cette peinture. »