

Autour du concept de *texte*. Pour un dialogue des disciplines de l'analyse des données textuelles

Jean-Michel Adam

Pôle de recherche et d'enseignement interdisciplinaires
en Sciences des textes et analyse comparée des discours
de l'Université de Lausanne (www.unil.ch/nss-iris4)

1. Les conditions d'un dialogue interdisciplinaire

Je suis très honoré d'avoir été convié à prononcer une des trois conférences plénières de ces huitièmes Journées internationales d'Analyse des Données Textuelles. Je n'ai eu l'audace d'accepter ce défi que par amitié pour nos collègues bisontins et en me disant égoïstement que je serais ainsi poussé à me mettre à niveau en préparant cette conférence et en assistant à vos débats. Autour de l'analyse de discours et des disciplines et sciences des textes, nous avons entrepris, avec Jean-Marie VIPREY, une réflexion épistémologique qui croise l'analyse textuelle et comparée des discours que nous développons à Lausanne et l'analyse informatique des données textuelles du pôle *Archives, Bases, Corpus* de la Maison des Sciences de l'Homme de Franche-Comté. Nous avons en commun une interrogation qui porte sur les concepts de *texte*, de *discours*, d'*archive* et de *corpus* et le pôle 3 de la M.S.H. place au centre de ses travaux une question qui nous concerne tous : « *Qu'est-ce qu'un texte à l'ère numérique ?* » A question, je ne répondrai pas en informaticien, mais en linguiste spécialiste du texte et de l'analyse des discours. Les conditions d'un réel dialogue interdisciplinaire, qui caractérisent l'esprit des JADT, comme l'a rappelé en introduction Jean-Marie VIPREY, ont un prix : nous devons faire l'effort de nous

lire, de nous écouter et d'interroger les propositions des uns et des autres depuis nos propres positions disciplinaires. C'est à cet exercice que je me suis livré et c'est pourquoi mon intervention prendra assez largement appui sur un article de Jean-Marie VIPREY, dans le dernier numéro 21 de *Semen* : « Quelle place pour les sciences des textes dans l'analyse de discours ? » (2006 : 167-182). Je compléterai mon exploration des 3000 pages des JADT de Saint-Malo, de Louvain-la-Neuve et de Besançon par le premier numéro, paru en 2002, de l'excellente revue *Corpus* de l'équipe niçoise « Bases, Corpus et Langage ».

Par sciences et disciplines des textes, je fais allusion à quelques-unes de celles que vos travaux et les miens croisent à l'occasion : la **stylistique** et la **théorie des genres** sont souvent croisées, par exemple par Denise MALRIEU dans les JADT 2002 (483-494), par Margareta KASTBERG SJÖBLOM, dans les JADT 2004 (672-691) et dans les JADT 2006, aussi bien par la stylistique générique des récits de voyage de Véronique MAGRI-MOURGUES (655-666) que par l'étude des « contrastes internes et variations stylistiques du genre de l'article scientifique de linguistique » de Céline POU DAT & Fanny RINCK (785-795). En fait, cette question des styles et des genres traverse visiblement un très grand nombre d'autres interventions. Je fais également allusion à un autre domaine que je connais bien, celui de l'**analyse des discours politiques et médiatiques**. Je citerai seulement les études, par Carmen PINEIRA-TRESMONTANT, des discours du 24 décembre (Navidad) et du 6 janvier (Pascua Militar) du roi Juan-Carlos d'Espagne de 1975 à 1999, dans les JADT 2004 (874-884), par Damon MAYAFFRE, des discours de Thorez, Blum, Tardieu et Flandin entre 1930 et 1939, dans les JADT 2002 (517-528), des discours de la cohabitation Jospin-Chirac entre 1997 et 2002, dans les JADT 2004 (785-792) et des discours présidentiels sous la Vème République (2006 : 677-685), par Pierre FIALA et Jean-Marc LEBLANC du « *Je* présidentiel » dans les messages de vœux aux français sous la Cinquième République : entre 1959 et 2001, dans les JADT 2004 (716-725).

J'ajouterai l'analyse contrastive de discours des présidents de la Vème République proposée par Jean-Marc LEBLANC et William MARTINEZ dans les JADT 2006 (601-614) ou encore l'analyse de l'influence des « plumes de l'ombre » sur les discours des premiers ministres québécois proposée par Denis **MONIERE** et Dominique LABBE (2006 : 687-696), la « méthodologie pour l'analyse comparée » des discours des Verts français et italiens de Béatrice FRACCHIOLLA (2006 : 415-425) ou encore le travail d'Alpha Ousmane BARRY sur la rhétorique oratoire de Sékou Touré (2006 : 107-120).

J'ai constaté qu'un certain nombre de travaux se revendiquent de l'**analyse textuelle**. Une analyse textuelle qui est utilisée en sémiologie clinique par Jean-Marie VIDAL et une équipe de Rennes, dans les JADT 2002 (767-777), par Pierre FIALA et Jean-Marc LEBLANC qui parlent de « **comparaison en analyse textuelle** » dans leur étude des messages de vœux aux français. J'ai également trouvé dans vos travaux la confirmation d'ouvertures en direction de la **génétique textuelle**, avec la présentation du logiciel MEDITE de comparaison de versions par Jean-Gabriel GANASCIA, Irène FENOGLIO et Jean-Louis LEBRAVE de l'ITEM, dans les JADT 2004 (468-478) et de la **philologie**, avec la présentation du projet d'encodage des chartes rédigées en français de Wallonie en 1300, présenté par Nicolas MAZZIOTTA, dans les JADT 2004 (793-803). Jean-Gabriel GANASCIA et Julien BOURDAILLET lient même génétique textuelle et philologie dans leur article des JADT 2006 (427-437).

J'ai surtout été frappé – et c'est ce qui a motivé la mise au point qui suit – par le fait que nous nous référons à plusieurs de ces concepts que Jean-Marie VIPREY appelle fort justement des concepts « régulateurs » ou « transactionnels [...] indispensables au jeu des disciplines humanistes entre elles, même si certains ont [...] une ou plusieurs acceptions dures dans ces mêmes disciplines » (2006 : 168). Parmi les concepts que nous nous empruntons et que nous allons

chercher dans d'autres disciplines, je vais retenir ceux de *genres de discours*, d'*intertextualité*, de *co-textualité*, d'*unités de l'analyse textuelle* et de *corpus* autour desquels je concentrerai mes propositions.

Si les conditions d'un dialogue de nos disciplines semblent réunies, c'est en raison de la double évolution identifiée par Jean-Marie VIPREY : d'une part, leur recentrage sur des données résolument textuelles et, d'autre part, le fait que travaillant sur des textes de plus en plus nombreux, différents, complexes et ouverts, nous avons besoin de textes très solidement établis et de réunir les conditions d'un traitement informatisé et d'une circulation scientifiquement contrôlée de cette masse de nouvelles données textuelles. A cela j'ajouterai le fait que : « La controverse entre linguistiques qualitative et quantitative est aujourd'hui quelque peu désuète », comme le disent Jean-Marie VIDAL et ses collègues, dans les JADT 2002 (767). Prenant acte du dépassement de ce faux débat, je dirai que le traitement informatisé des données textuelles implique la création de réseaux de collaborations entre des équipes de **recherche** travaillant à une véritable refonte des sciences et disciplines académiques des textes. Nous avons, de toute évidence, besoin les uns des autres : tandis que nous linguistes nous mettons l'accent sur la définition des unités élémentaires, sur le traitement de la linéarité des textes, sur les enchaînements transphrastique et sur la combinatoire d'unités de rangs de complexité supérieurs à la phrase, vos travaux insistent sur la structure non-séquentielle et réticulaire des textes. C'est ce que font, dans deux articles d'un prochain numéro de *Langages*, Dominique LEGALLOIS, du CRISCO de Caen, qui a dirigé ce numéro consacré aux « Unité(s) du texte », et Jean-Marie VIPREY. Je vais partir de ma lecture de leur travail pour montrer l'utilité de notre dialogue.

Jean-Marie VIPREY aborde **l'organisation non-séquentielle des textes** en portant son attention sur une composante lexicale de la cohésion textuelle indé-

niablement négligée par les linguistiques du texte, même quand elles intègrent le concept encore trop rudimentaire d'isotopie. J.-M. VIPREY montre que la **cooccurrence** ou la **collocation** (chez M. A. K. HALLIDAY & R. HASAN ou chez Z. S. HARRIS) est un facteur primordial de textualité : « La collocation modifie et constitue les vocables en tant que tels, non tellement comme individus que comme constituants d'un tout engrené, influant ainsi sur la cohésion par l'intermédiaire du palier lexical intégralement considéré (et non disséminé phrase par phrase) » (à paraître 2006). Dominique LEGALLOIS utilise quant à lui la répétition lexicale au sein d'unités syntaxiques pour établir des réseaux de phrases semblables qui dessinent l'organisation globale, non-séquentielle des textes. Il montre que la répétition de trois lexèmes ou plus dans certaines phrases d'un texte est responsable d'une organisation réticulaire « en parfaite congruence avec l'étymologie du mot *texte*. C'est bien un *tissu* de phrases enchevêtrées, une *trame*, une *texture*, toute une constitution et une conception d'un objet complexe que met en évidence l'analyse de la répétition lexicale dans les discours » (à paraître 2006). Comme Dominique LEGALLOIS le note en passant, cette méthode d'identification de phrases appariées au sein d'un texte pourrait s'appliquer à l'identification de phrases intertextuellement liées. Les travaux de D. LEGALLOIS et de J.-M. VIPREY ne remettent pas en cause le caractère structurant de la linéarité textuelle, mais ils prouvent que **la textualité doit résolument être pensée comme la combinaison de parcours linéaires et réticulaires**. C'est ce que je vais examiner à partir d'un exemple dans un moment, mais je tiens à revenir auparavant sur le traitement de la discontinuité textuelle.

Les travaux linguistiques traitent la discontinuité textuelle à la lumière des phénomènes de portée des connecteurs (ouverture à droite en attente d'une fermeture qui déborde souvent très largement la phrase pour s'étendre de la période au paragraphe). C'est également ce que j'ai travaillé dans mes publications relatives à la complexité de l'organisation séquentielle des textes et en particulier à

l'identification de zones textuelles narratives, descriptives, explicatives, argumentatives et dialogales (Adam 2001). Les zones dialogales sont particulièrement étudiées dans tout ce qui s'écrit actuellement sur le discours rapporté et plus largement sur le discours et la pensée représentés. Ces préoccupations linguistiques sont en phase avec les propositions de lemmatisation qui tiennent compte de la linéarité des textes et de la localisation d'unités comme les temps verbaux. C'est ce que développent, à propos des parties et de la structure de textes d'historiens latins, Dominique LONGREE, Xuan LUONG et Sylvie MELLET dans les JADT 2004 (743-752) et 2006 (643-654). Leur travail porte à la fois sur des différences génériques et sur la structure des textes. Dans sa contribution aux JADT 2002, Denise MALRIEU (483-494) montre également qu'il est nécessaire de développer des outils de balisage des parties des textes graphiquement marquées ou non « pour ne plus confondre des lexies qui peuvent être identiques quant à leur forme et leur fonction syntaxique mais non homogènes sur la composante dialogique » (2002 : 493). Son étude de corrélations entre variables a l'intérêt de se présenter comme une « analyse contrastive de fréquences de distributions respectant les relations syntaxiques à l'intérieur d'unités typées de longueurs variables » (id.). Denise MALRIEU ajoute fort justement qu'il faut absolument tenir compte de deux ordres de variables : non seulement les « genres » mais « les séquences textuelles » (2002 : 493). Denise MALRIEU fait une remarque toute simple et pertinente à propos de l'impératif : l'impératif est certes un indice d'interlocution dans un dialogue ou dans une forme d'adresse, mais dans un texte du type conseil-consigne, dans les notices de montage et même dans les proverbes, il n'a pas du tout le même sens.

Un autre aspect de la discontinuité est moins pris en compte : on a l'habitude de distinguer des textes de genres discursifs différents, mais on oublie souvent qu'un texte n'est pas monolithiquement d'un seul genre. Il est fréquent de voir se mêler au sein d'un texte plusieurs catégories génériques signalées par

des configurations de marques différentes. Ainsi *La Barbe bleue* de Perrault n'est un conte que parce qu'il commence par la formule devenue canonique : « Il était une fois... » et parce que si la clé tombée dans le sang des victimes de ce tueur en série ne peut pas être nettoyée, c'est parce que « la clé était fée », comme le dit le narrateur dans une incise explicative. Hors ces deux énoncés, le genre du conte merveilleux est totalement absent du reste du texte. La barbe du personnage n'est qu'un signe distinctif, pas un élément aussi merveilleux que la « petite pantoufle de verre » de Cendrillon. *Barbe bleue* relève plutôt du genre des histoires sanglantes très prisé au XVII^{ème} siècle. On pourrait aussi s'intéresser à la critique sociale ironique que développe le rapport entre le récit et les deux Moralités... Bref, la belle homogénéité du genre n'est pas la loi des textes. C'est pourquoi, avec Ute HEIDMANN, nous préférons parler de généricité (Adam & Heidmann 2004).

2. La phrase, le texte et l'intertexte : d'André Breton à Lautréamont et Thiers

Pour ne pas rester dans la pure réflexion théorique, je vous propose d'examiner un exemple sur lequel j'ai travaillé pour cette conférence. Que nous apporte la proposition d'identification de phrases appariées dans un texte et lorsqu'on dépasse les limites d'un texte ? Je partirai, pour cela, d'une phrase d'André Breton : la clause de *Nadja* (1928).

(1) **La beauté sera CONVULSIVE ou ne sera pas.** (*Nadja*)

On se trouve très à l'étroit si l'on reste dans le seul co-texte gauche de cet énoncé de *Nadja*. Le texte qui précède ne dit pas grand chose qui permette de comprendre cet énoncé. Ainsi renvoyés prioritairement à l'œuvre de Breton, on tombe alors, par recherche de cooccurrence de la collocation rare du substantif *beauté* et de l'adjectif *convulsive*, sur le début de *L'Amour fou* (1937) :

(2) **Il ne peut, selon moi, y avoir beauté – beauté convulsive – qu'au prix de l'affirmation du rapport réciproque qui lie l'objet considéré dans son mouvement**

et dans son repos.

(*L'Amour fou*)

Nous avons affaire à un intertexte auctorial auquel la fin de (2) ajoute l'écho d'un domaine de mémoire externe : l'interdiscours médical dans lequel l'énoncé prend en effet sens. Il suffit de lire la définition médicale de la convulsion donnée par le *Larousse du XXe siècle*, dans son édition de 1929, contemporaine du texte de Breton :

- (3) **[Convulsion] Méd. Contractions musculaires, involontaires et instantanées, locales et intéressant un ou plusieurs groupes musculaires, ou généralisées à tout le corps.**
(*Larousse du XXe siècle* 1929)

Le texte de la fin de *Nadja* confirme cette la provenance médicale de l'épithète :

- (4) **[...] ni dynamique ni statique, la beauté je la vois comme je t'ai vue. [...] Elle est comme un train qui bondit sans cesse dans la gare de Lyon et dont je sais qu'il ne va jamais partir, qu'il n'est pas parti. Elle est faite de saccades [...] La beauté, ni dynamique ni statique. Le cœur humain, beau comme un sismographe.** (*Nadja*)

La phrase nominale « La beauté, ni dynamique ni statique » apparaît ainsi comme une annonce paraphrastique de l'énigmatique « beauté convulsive » de la clause. Par ailleurs, la phrase finale « Le cœur humain, beau comme un sismographe » ouvre sur l'intertexte des *Chants de Maldoror* d'ailleurs explicitement signalé au début de *L'Amour fou* dans une phrase (5) qui fait coexister la formule de Lautréamont et l'épithète de Breton :

- (5) **Les "beau comme" de Lautréamont constituent le manifeste même de la poésie convulsive.**
(*L'Amour fou*)

Une recherche entreprise rapidement à partir du corpus **numérisé** du Centre de Recherche Hubert de Phalèse, de Paris III, actuellement dirigé par Michel Bernard, et auquel collaborent Jean-Pierre Goldenstein et Pascal Mougin, nous permet de constater l'existence de tout un réseau de segments textuels déterminés par les collocations des lexèmes « beauté » et du connecteur comparatif « comme ». Je ne donne ici que la première occurrence en en soulignant le rythme périodique carré (à 4 membres) :

- (6) Ta grandeur morale, image de l'infini, est immense **comme** la réflexion du philosophe, **comme** l'amour de la femme, **comme la beauté** divine de l'oiseau, **comme** les méditations du poète. (*Chant I* ; je souligne ici et dans les exemples qui suivent)

La seconde occurrence (7) est une déclaration majeure de la poétique d'Isidore Ducasse, une glose de la complexité surréaliste de tous ces « beau comme », la clé du dispositif qui séduira tant les surréalistes.

- (7) Chacun a le bon sens de confesser sans difficulté (quoique avec un peu de mauvaise grâce) **qu'il ne s'aperçoit pas, au premier abord, du rapport, si lointain qu'il soit, que je signale entre la beauté** du vol du milan royal, **et celle de** la figure de l'enfant, s'élevant doucement, au-dessus du cercueil découvert, **comme** un nénuphar qui perce la surface des eaux ; **et voilà précisément en quoi consiste l'impardonnable faute qu'entraîne l'inamovible situation d'un manque de repentir, touchant l'ignorance volontaire dans laquelle on croupit.** (*Chant V*)

De la suite, je n'ai retenu que les trois segments textuels qui constituent un réseau à structure périodique presque identique syntaxiquement :

- | | |
|-----------------------------------|------------------------------------|
| (11) (a) — beau comme — ; | (15) & (16) (a) — B/beau comme — ; |
| (b) <i>ou plutôt</i> , comme — ; | (c) <i>ou encore</i> , comme — ; |
| (c) <i>ou encore</i> , comme — ; | (b) <i>ou plutôt</i> , comme — ; |
| (d) <i>et surtout</i> , comme — ! | (d) <i>et surtout</i> comme — ! |
- (11) Avec quelle satisfaction de n'être pas tout à fait ignorant sur les secrets de son double organisme, et quelle avidité d'en savoir davantage, je le contemplais dans sa métamorphose durable ! Quoiqu'il ne possédât pas un visage humain, il me paraissait [a] **beau comme** les deux longs filaments tentaculiformes d'un insecte ; [b] *ou plutôt*, **comme** une inhumation précipitée ; [c] *ou encore*, **comme** la loi de la reconstitution des organes mutilés ; [d] *et surtout*, **comme** un liquide éminemment putrescible ! Mais ne prêtant aucune attention à ce qui se passait aux alentours, l'étranger regardait toujours devant lui, avec sa tête de pélican ! Un autre jour, je reprendrai la fin de cette histoire. (*Chant V*)
- (15) Je me connais à lire l'âge dans les lignes physiognomoniques du front : il a seize ans et quatre mois ! [a] Il est **beau comme** la rétractilité des serres des oiseaux rapaces ; [c] *ou encore*, **comme** l'incertitude des mouvements musculaires dans les plaies des parties molles de la région cervicale postérieure ; [b] *ou plutôt*, **comme** ce piège à rats perpétuel, toujours retendu par l'animal pris, qui peut prendre seul des rongeurs indéfiniment, et fonctionner même caché sous la paille ; [d] *et surtout* **comme** la rencontre fortuite sur une table de dissection d'une machine à coudre et d'un parapluie ! Mervyn, ce fils de la blonde Angleterre, vient de prendre chez son professeur une leçon d'escrime, et, enveloppé dans son tartan écossais, il retourne chez ses parents. (*Chant VI*)
- (16) Et je me trouve **beau** ! [a] **Beau comme** le vice de conformation congénital des organes sexuels de l'homme, consistant dans la brièveté relative du canal de l'urètre et la

division ou l'absence de sa paroi inférieure, de telle sorte que ce canal s'ouvre à une distance variable du gland et au-dessous du pénis ; [c] ou encore, **comme** la caroncule charnue, de forme conique, sillonnée par des rides transversales assez profondes, qui s'élève sur la base supérieure du bec du dindon ; [b] ou plutôt, **comme** la vérité qui suit : « le système des gammes, des modes et de leur enchaînement harmonique ne repose pas sur des lois naturelles invariables, mais il est, au contraire, la conséquence de principes esthétiques qui ont varié avec le développement progressif de l'humanité, et qui varieront encore » ; [d] et surtout **comme** une corvette cuirassée à tourelles ! Oui, je maintiens l'exactitude de mon assertion. (*Chant VI*)

Cette identité de structure et de fonctionnement sémantique des trois segments périodiques nous met sur la piste d'une structure réticulaire en lien avec la progression du récit. La phrase périodique (11), du *Chant V*, se termine par un énoncé métatextuel : « **Un autre jour, je reprendrai la fin de cette histoire** », qui semble bien renvoyer à (15) et à (16) du *Chant VI*. Alors que le (15) a été largement cité et retenu depuis les surréalistes, la fin métatextuelle de (16) me semble capitale : « **Oui, je maintiens l'exactitude de mon assertion** ».

Pour cerner minimalement la constitution du lexème « convulsive » en **vocabulaire** d'au moins deux textes de Breton, il me faut encore dire un mot d'un autre réseau intertextuel de contextualisation de la phrase-clausule que j'ai prise comme point de départ de mon propos. Le *Larousse du XXe siècle* ajoute au sens médical de l'entrée « Convulsion » que je signalais tout à l'heure un sens figuré :

(17) *Fig. : Les CONVULSIONS du désespoir. Les CONVULSIONS politiques.*
(*Larousse du XXe siècle* 1929)

Notons d'abord la même graphie en majuscules dans le dictionnaire et dans la phrase de Breton. Mais cette définition nous guide vers un autre domaine de l'interdiscours : une phrase présente dans la mémoire intertextuelle du discours politique français. Élu chef du pouvoir exécutif de la République française par l'Assemblée nationale, en février 1871, Louis-Adolphe Thiers, dans un message adressé à l'Assemblée nationale du 13 novembre 1872, a résumé sa conception politique par une phrase célèbre :

(18) **La République sera conservatrice ou elle ne sera pas.** (Thiers, 13.11.1871)

Si l'on fait de cet énoncé politique un possible intertexte de la clause de *Nadja*, en tenant compte du fait que la phrase de Breton vient après un autre collage (un fait-divers tragique en provenance de l'interdiscours journalistique), on peut dire qu'on est en présence d'un collage-détournement de nature à la fois littéraire et politique. Dans le champ littéraire, cette phrase se présente comme une application de la poétique de Lautréamont-Isidore Ducasse, et de son jeu favori avec le plagiat-détournement. Je rappelle sa déclaration célèbre des *Poésies II* :

(19) **Le plagiat est nécessaire. Le progrès l'implique. Il serre de près la phrase d'un auteur, se sert de ses expressions, efface une idée fausse, la remplace par l'idée juste.** (*Poésies II*, édition établie par J.-P. Goldenstein, Presses-Pocket, 1992 : 275).

Dans le champ politique et révolutionnaire de l'engagement surréaliste, le détournement de la phrase-« idée fausse » de celui qui réprima l'insurrection de la Commune devient particulièrement pertinent. Le procédé du collage est patent dans les exemples (16b) et (16c), dont on connaît aujourd'hui très exactement la provenance : la 6^{ème} partie, consacrée aux oiseaux, de l'*Encyclopédie d'histoire naturelle*, du Docteur Chenu (Paris, Marescq & Cie, 1854 : 100) pour (16b) et le traité de *Théorie physiologique de la musique fondée sur l'étude des sensations auditives* de H. Helmholtz, traduit de l'allemand par M. G. Guérault (Paris, Masson, 1868 : 306) pour la section entre guillemets de (16c).

Bien sûr, le problème de l'identification automatique du dialogue intertextuel des phrases de Breton et de Thiers est assez problématique. Est-ce que la loi des trois éléments cooccurrents de LEGALLOIS s'applique ? Pour ma part, je soulignerai la presque identité de structure syntaxique qui se prolonge dans les échos phoniques des signifiants des lexèmes *CONSerVatRICE* et *CONVulSIVE* : même syllabe d'attaque et redoublement du phonème /s/ dans le premier, /v/ dans le second et, pour finir le mot, appui sur la même voyelle /i/ suivie d'une des deux consonnes /s/ ou /v/ et d'une finale muette identique : /is-/ et /iv-/.

(18)	La République sera	CONSerVatRICE	ou elle ne sera pas.
(1)	La beauté sera	CONVulSIVE	ou ne sera pas.

Si j'ai traité un peu longuement ces exemples, c'est pour en tirer deux conséquences pour la définition du texte :

A). La linguistique textuelle a bien décrit la répétition-progression qui fait d'un texte un objet complexe structuré par une linéarité textuelle constitutive, qui est celle de la linéarité de la langue. Mais un texte est aussi une structure réticulaire d'énoncés. La prise en compte de l'organisation des vocables en isotopies et des réseaux phrastiques qui permettent de mettre en relation des énoncés éloignés, ces deux approches non-linéaires mettent un peu de clarté dans des textes aussi complexes que *Nadja* ou *Les Chants de Maldoror*.

B). Ce qui me paraît également intéressant du point de vue théorique, c'est que la phrase-clausule de Breton n'a **pas un** intertexte mais **plusieurs**. Le sens de la phrase-clausule de Breton est, à la fois, dans le co-texte de *Nadja*, dans l'intertexte auctorial du début de *L'Amour fou*, dans l'intertexte politique de la phrase de Thiers, dans l'interdiscours médical de la neuropsychiatrie, dans l'intertexte littéraire du Lautréamont de *Chants de Maldoror*. Nous retrouvons là une question posée par les travaux de Damon MAYAFFRE (2002), de Ludovic TANGUY & Théodore THLIVITIS (1999) ainsi que de Ioannis KANELLOS (1999).

Il me semble que se pose une question essentielle pour nos sciences du texte : celle des limites de l'unité texte et des limites de l'interprétation des énoncés. Nous avons besoin d'une **redéfinition de la co-textualité des énoncés**, (chez RASTIER, le « contexte intra-textuel » 1998 : 106). Les travaux que je viens de citer vont dans la même direction que celle que dessine Jacques GUILHAUMOU (1993 & 2002 : 32). Pour ce dernier – qui prolonge là des remarques de Michel PECHEUX et de Denise MALDIDIER –, l'analyse de discours doit être considérée comme une « **discipline herméneutique à part entière** ». Ce tour-

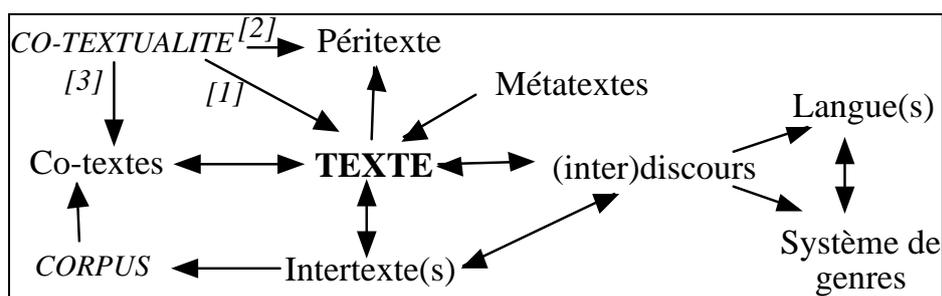
nant herméneutique et plus largement d'ouverture de la linguistique à l'interprétation est polémiquement prôné par RASTIER (2001) et raisonné de façon un peu plus sereine par Frédéric COSSUTTA (2004). Ce **tournant herméneutique de l'analyse de discours** n'est possible, selon Guilhaumou, que si l'on « situ[e] les sources interprétatives des textes en leur sein » (2002 : 32). À la différence de l'oral, la contextualité de l'écrit, comme le dit RASTIER, « va de texte à texte » (1998 : 106). En d'autres termes, à l'écrit, c'est le co-texte qui domine la situation, et même qui la supplée.

Tout ceci nous ramène à la question des ressources ou données textuelles nécessaires à la construction du sens des énoncés. Il faut, à ce niveau de la réflexion, se pencher sur une redéfinition du concept de texte que je ferai reposer sur une distinction du *texte* et de la *textualité* d'où découlera une distinction du *texte* et de sa *textualisation*.

3. Distinguer texte et textualité

Nous avons parlé de l'ouverture intertextuelle du sens des énoncés comme d'un phénomène de relation de texte à texte.

Schéma 1 La textualité comme système de concepts



A ce type de relation mémorielle *in absentia*, il faut ajouter un type plus évident de relation *in praesentia* de texte à texte, généralement négligé, je veux dire les relations co-textuelles non pas entre les **énoncés en relation co-textuelle interne à un texte [1]**, mais les **relations co-textuelles entre plusieurs textes [3]**, réunis matériellement au sein d'une page ou d'une double

page de journal, dans ce qu'on appelle les hyperstructures journalistiques, qui regroupent plusieurs articles sur un même sujet (Adam & Lugrin 2000, Lugrin 2000 & 2001). Ces co-textes peuvent également se trouver matériellement réunis dans un recueil de poèmes, de contes ou de nouvelles. Les co-textes présents à gauche et/ou à droite d'un texte jouent un rôle dans la détermination de son sens comme le prouvent par exemple le passage du manuscrit de 1695 à l'édition des *Histoires ou contes du temps passé. Avec des moralités* de Perrault en 1697 :

Manuscrit 1695		Première édition (Barbin, 1697)
<i>Contes de ma Mère Loye</i>		<i>Histoires ou contes du temps passé. Avec des Moralités.</i>
La Belle au Bois Dormant.	<i>Conte</i>	La Belle au Bois Dormant. <i>Conte</i>
Le Petit Chaperon Rouge.	<i>Conte</i>	Le Petit Chaperon Rouge. <i>Conte</i>
La Barbe Bleue.	<i>Conte</i>	La Barbe Bleue.
Le Maître Chat ou le Chat botté.	<i>Conte</i>	Le Maître Chat ou le Chat botté. <i>Conte</i>
Les Fées.	<i>Conte</i>	Les Fées. <i>Conte</i>
		Cendrillon ou la petite pantoufle de verre <i>Conte</i>
		Riquet à la houppe. <i>Conte</i>
		Petit Poucet. <i>Conte</i>

La co-textualité des *Fées* change du manuscrit à la première édition, avec l'ajout de *Cendrillon* qui en est l'écho amplifié.

Un troisième type de **relations co-textuelles [2]** doit être considéré : les relations du texte aux **énoncés péri-textuels** qui en délimitent les frontières (titre, sous-titre, noms d'auteur et d'éditeur, couverture et page de titre, dédicace, épigraphe, préface et postface, annexes, sommaire, etc.). Le recueil de Perrault change de titre, un récit comme *La Barbe bleue* n'a plus le sous-titre *conte* conformément à ce que je disais tout à l'heure de ses caractéristiques génériques bien plus complexes. L'« **Epître** à Mademoiselle » qui ouvre le recueil est péri-textuellement à la fois une *dédicace* à la nièce de Louis XIV et une *préface* destinée à tous les lecteurs.

Un texte n'est donc pas une entité autonome et fermée, mais un système

complexe de relations que je désigne par des termes partiellement empruntés à la poétique de Gérard GENETTE. Dire que **la textualité est plus complexe que le texte**, c'est remettre en cause l'idée trop schématique d'une division stricte entre espace interne et externe au *texte*. Tout texte est ainsi souvent déterminé par les **commentaires métatextuels** (gloses autoriales d'un texte ou gloses critiques, journalistiques) qui l'entourent matériellement à l'occasion de certaines éditions savantes ; parfois plus lointaines et diffuses, les interprétations passées pèsent d'une façon ou d'une autre au point d'occulter les sens que le texte propose.

Le **système de genres** et **la langue** sont les deux principales composantes de l'**interdiscursivité**, définie comme *possibilité de formes de discours* disponibles (tant à la production qu'à l'interprétation) dans la communauté socio-discursive des auteurs, des éditeurs et des lecteurs. Selon une très juste formule de François RASTIER : « Aucun texte n'est écrit seulement "dans une langue" : il est écrit dans un genre en tenant compte des contraintes d'une langue » (2004 : 126). La généricité place un texte donné dans une société de textes qui change avec la culture des lecteurs et dans le temps historique. Elle est corrélée à la convocation d'intertextes aussi indispensables eux-mêmes à l'interprétation du sens des énoncés que les énoncés co-textuellement liés, co-présents matériellement. Comme le dit Damon MAYAFFRE, le traitement de ressources co-textuelles ne saurait être envisagé de façon différente de celui de ressources intratextuelles (2002 : 63). On peut, en effet, définir les **co-textes** par leur co-présence matérielle au sein d'un ensemble **éditorial** au sens classique de l'ère Gutenberg ou **numérique** dans le cadre de la constitution d'une édition numérisée et d'un corpus. La décision de placer un intertexte à côté d'un texte (dans le corpus co-textuel ainsi constitué), c'est rendre cet intertexte aussi lisible que les autres énoncés co-textuels et le faire glisser vers un statut de co-texte. Une glose métatextuelle peut passer dans le péri-texte sous forme de préface ou de postface.

Je ne crois pas utile de conserver le concept-calembour d'architexte de

RASTIER, même si son « principe d'architecte-Archimède » est très juste : « tout texte placé dans un corpus en reçoit des déterminations sémantiques, et modifie potentiellement le sens de chacun des textes qui le composent » (2001 : 92). Je crois préférable de concentrer notre intérêt sur l'idée de « corpus réflexifs » dont MAYAFFRE dit qu'ils « devront être organisés techniquement comme des *hypertextes* : chaque texte constituant devra être relié aux textes considérés comme parents. [...] Le but étant [...] que chaque texte puisse être replacé dans son réseau co-textuel utile et nécessaire pour sa compréhension et son analyse » (2002 : 64-65). MAYAFFRE distingue deux types de corpus : les **corpus centripètes**, qui convergent (comme je l'ai fait tout à l'heure en voulant élucider le sens possible de la phrase de Breton) ; dans ce cas le site hypertextuel est orienté clairement vers la compréhension du texte placé en son centre. La création d'un **corpus réflexif** « sans direction pré-établie, [...] non focalisé *a priori* sur un de ses constituants » (Mayaffre 2002 : 65) est d'une autre nature et offre des possibilités herméneutiques nouvelles. C'est ce que nous travaillons actuellement à Lausanne, sans ressources hypertextuelles numérisées, à propos des contes de Perrault et d'Andersen. Nous nous appuyons pour cela sur la théorie de la comparaison d'Ute HEIDMANN, comparaison différentielle qu'elle théorise comme non-hiérarchique (2005a & b).

4. Distinguer texte et textualisations

Les problèmes liés à l'établissement des textes sont communs à l'analyse de discours et au « scriptorium numérique », pour reprendre une expression de Jean-Marie VIPREY. Nous ne devons pas reconduire le leurre de l'évidence de l'objet texte qui a été assez largement dénoncé dès le début des années 1980. Pierre KUENTZ déplorait alors l'oubli, par l'analyse de discours, des conditions de distribution et de consommation des textes, ainsi que du rôle de l'histoire du livre et de l'imprimerie (1981 : 43). Cet important reproche a été relancé, dans les an-

nées 1990, aussi bien par le littéraire Michel CHARLES que par l'historien Roger CHARTIER. Le premier considère la croyance en l'existence du texte comme le premier des préjugés critiques :

Soit un texte, je vais l'étudier. Tout se passe comme si le texte existait hors du regard que je porte sur lui, hors de l'expérience que j'en ai, hors des opérations que je lui fais subir pour que précisément il devienne texte. (Charles 1995 : 40)

Quant à Roger CHARTIER, il décrit ainsi ce qu'il considère comme la première limite des diverses sortes de critiques textuelles :

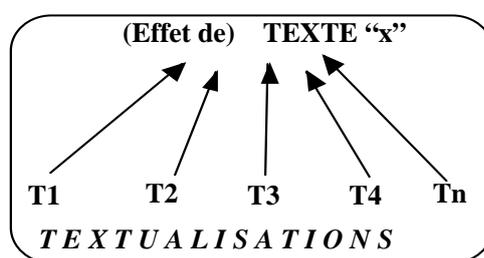
Leur première limite tient au fait qu'elles considèrent (le plus souvent) les textes comme s'ils existaient en eux-mêmes, en dehors des matérialités (quelles qu'elles soient) qui en sont les supports et les véhicules. Contre cette « abstraction » des textes, il faut rappeler que les formes qui les donnent à lire, à entendre ou à voir participent, elles aussi, à la construction de leur signification. Le « même » texte, fixe dans sa lettre, n'est pas le « même » si changent les dispositifs de son inscription ou de sa communication. (Chartier 1998 : 270)

La circulation matérielle des textes et de leur(s) édition(s) philologique, commerciale et/ou numérique, nous oblige à questionner le concept de texte sous l'angle de sa textualisation. En complétant ce que je développe dans Adam 2005b, je dirai seulement que l'on ne doit pas négliger l'importance des médiations éditoriales. Je suis entièrement d'accord avec François RASTIER : « La philologie rappelle que les textes ne sont pas des données, mais des constructions problématiques issues de diverses procédures » (2001 : 82).

Ainsi, pour traiter un discours politique oralisé du type de ceux du général de Gaulle sur lesquels j'ai travaillé, il est, d'un point de vue linguistique, nécessaire d'établir une transcription de travail qui tienne compte du fait qu'on a affaire à un écrit oralisé. J'ai montré l'utilité d'une transcription qui rende compte des interruptions de la foule et du rythme de la parole des discours du 24 juillet 1967, au balcon de l'Hôtel de ville de Montréal (Adam 2005 : 55-57 ; de façon plus détaillée dans Adam 2004) ou du 4 juin 1958 sur le forum d'Alger (Adam 1999 : 74-76 & 132-136). Ces transcriptions de travail sont des objets construits pour l'analyse. Un texte comme le discours de Montréal du général De Gaulle

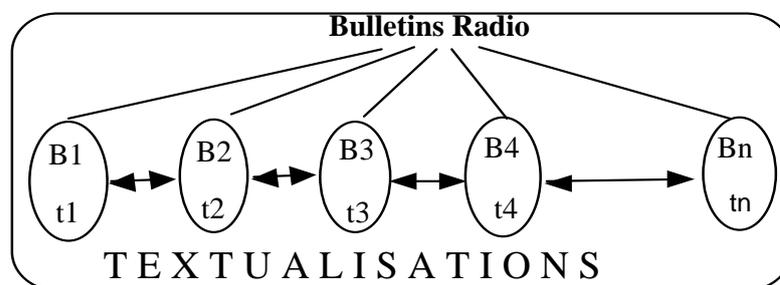
n'existe pas autrement que sous forme de textualisations diverses : une version écrite disponible sur Internet et qui semble la version « officielle » du côté des responsables de l'héritage politique du général De Gaulle (appelons-là T1), une version des archives nationales du Canada (T2), une version audio-visuelle disponible sur la toile dans les archives de Radio Canada (T3) dont j'ai proposé une transcription de travail T4 dans Adam 2005a (56-57). Le schéma 2 résume cette complexité du texte que l'on pourrait appeler « discours de Montréal » :

Schéma 2



De la même manière, les bulletins d'une station de radio, sur une matinée ou une journée, à propos de la même information, sont en relation de co/inter-textualité tellement forte qu'on ne peut pas limiter le « texte » de cette information au seul bulletin B1 ou à B2 ou à Bn. Le texte est littéralement dans la série en variation des informations successives, reprises et modifiées dans le temps (t1, t2, t3, etc.). On est, là aussi, en face de textualisations successives, variationnelles, dont le sens est précisé dans la suite des n-textes considérés :

Schéma 3 : Reformulations ou re-textualisations d'une information radiophonique



Au concept de texte conçu comme version définitive, on substitue, dans cette perspective proche de la « sémiotique de l'altération » de Jean PEYTARD

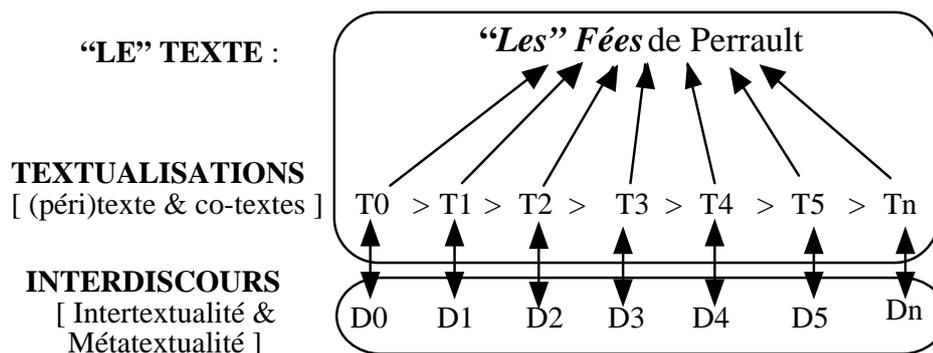
(1993), un concept de texte en variation et considéré comme la somme différentielle de ses états (reformulations, reprises) successifs

Il n'est de version définitive d'un texte que celle sur laquelle on s'arrête, comme *ne varietur*, pour en montrer la constante stabilité acquise certes après quelques variations. N'est-ce pas dans le cours de cette mouvance, de ces variations, qu'il convient de se situer pour apercevoir le sens, qui n'est jamais, semble-t-il, une fois pour toutes donné, mais constamment effet et source d'altérations ? (Peytard 1993 : 176)

Le problème qui se pose ici est celui de l'établissement des données textuelles. Dans sa contribution aux JADT 2004 : « Le texte dans tous ses états. Philosophie d'encodage du projet Khartès » (793-803), Nicolas MAZZIOTTA se pose, à propos de l'encodage de chartes rédigées en français de Wallonie de 1300, les mêmes problèmes que nous à propos des contes de Perrault. Comment respecter les majuscules, l'orthographe, la ponctuation d'origine et traiter les énoncés ? La solution est dans le double dispositif dont parle également Jean-Marie VIPREY dans son article *de Semen* 21, à propos des corpus de presse : « Le *texte brut* n'est la *surface graphique* que par approximation, et il a été parfois indûment sacralisé au détriment de celle-ci. L'objectif le plus ambitieux, mais qui est programmatiquement nécessaire, est de pouvoir donner accès au PDF. La non-destructivité se formule ainsi : à partir de toute étape du traitement des ressources, on doit pouvoir remonter sans perte à l'une des étapes antérieures » (2006 : 176)

Le schéma 3 représente les différents états d'un conte comme *Les Fées* de Perrault tels qu'ils pourraient constituer un corpus attentif à l'instabilité même de tout texte :

Schéma 4 : Le texte comme variation ou effet de textualisations successives



Les schémas 2, 3 & 4 veulent seulement signaler que les textualisations successives auxquelles nous avons (ou donnons) accès matériellement, contribuent à la construction d'un effet de texte dit « Les Fées de Perrault », « La Chanson de Roland », « Le discours du 24 juillet 1967 », « L'appel du 18 juin », « Les chants de Maldoror », etc. Et chacune de ces textualisations est prise dans un interdiscours changeant qui transforme, chaque nouvelle textualisation en un fait de discours singulier.

Cette réflexion sur le caractère variationnel des textes est importante au regard de l'illusion que provoque leur évidence matérielle. Considéré dans son épaisseur variationnelle, un texte acquiert cette dimension temporelle et matérielle d'**archive** qui est à proprement parler son historicité comme fait de discours. D'un point de vue épistémologique, il faut encore en tirer deux conséquences majeures décrites par Ute HEIDMANN (2005) : les textualisations T1, T2, Tn ne sont pas des variantes altérées d'un « vrai » texte, qui serait, dans le cas du discours de Montréal de De Gaulle, T3 ou, chez Perrault, l'édition Barbin originale. Ces **textualisations** sont à lire dans ce qui fait leurs différences (Heidmann 2005b) en passant d'une approche simplement contrastive (phase descriptive d'établissement des faits) à une méthode comparative (phase interprétative des faits observés). C'est la méthode que nous avons appliquée à la traduction d'un petit texte de Franz Kafka (Adam & Heidmann 2003b) et à la comparaison de deux contes des Grimm et d'Andersen (Adam & Heidmann 2002). C'est également la méthode que je mets en œuvre à propos des appels du

17 et du 18 juin du maréchal Pétain et du général De Gaulle ou d'un poème de Cendrars et du fait-divers qu'il **plagie** (Adam 1999 : 139-155 & 175-188). Je ne développe pas et souligne seulement, pour conclure, que l'édition numérique ouvre des perspectives assez considérables à une conception du texte comme « série ouverte dans l'ordre de ses matérialités », comme dit Jean-Marie VIPREY ou, comme le dit programmatiquement Jean PEYTARD :

C'est « là où ça varie », Là « où ça se différencie » que se construit le sens (variable et variant) à inter-échanger ; que ce n'est pas une analyse du « cohésif/cohérent » qui permet de fonder une sémantique. Mais peut-être, plutôt, la prise en compte du « différentiel » en variance, où s'enracinerait une « sémiotique de l'altération », une « sémiotique différentielle ». A élaborer. (1993 : 147)

REFERENCES

- ADAM Jean-Michel 1999 : *Linguistique textuelle. Des genres de discours aux textes*, Paris, Nathan, coll. FAC.
- 2001 (1997) : *Les textes : types et prototypes*, Paris, Nathan, coll. FAC.
- 2004 : « Quand dire “Vive le Québec libre !” c'est faire l'histoire avec des mots », *Discours et constructions identitaires*, D. Deshaies & D. Vincent édés., Presses de l'Université Laval, 13-38.
- 2005a : *La linguistique textuelle. Introduction à l'analyse textuelle des discours*, Paris, A. Colin, coll. Cursus.
- 2005b : « Les sciences de l'établissement des textes et la question de la variation », in *Sciences du texte et analyse de discours. Enjeux d'une interdisciplinarité*, J.-M. Adam & U. Heidmann édés., Genève, Slatkine, 69-96.
- ADAM Jean-Michel & HEIDMANN Ute 2002 : « Réarranger les motifs, c'est changer le sens. Princesses et petits pois chez Andersen et les Grimm », in *Contes : l'universel et le singulier*, A. Petitat éd., Lausanne, Payot, 155-174.
- 2003a : « Discursivité et transtextualité : la comparaison pour méthode. L'exemple du conte », in R. Amossy & D. Maingueneau édés. *L'analyse du discours dans les études littéraires*, Toulouse, Presses Universitaires du Mirail, 29-49.
- 2003b : « Du récit au rocher : Prométhée d'après Kafka », in U. Heidmann éd. *Poétiques comparées des mythes*, Lausanne, Payot, 187-212.
- 2004 : « Des genres à la genericité. L'exemple des contes (Perrault et les Grimm) », *Langages* 153, Paris, Larousse, 62-72.
- ADAM, Jean-Michel & LUGRIN, Gilles 2000 : « L'hyperstructure : un mode privilégié de présentation des événements scientifiques », in Fabienne Cusin-Berche (dir.) : *Rencontres discursives entre science et politique. Spécificités linguistiques et constructions sémiotiques*, Carnets du CEDISCOR, n° 6, Paris, Presses de la Sorbonne Nouvelle, 133-149.
- CHARLES Michel 1995 : *Introduction à l'étude des textes*, Paris, Seuil.
- CHARTIER Roger 1998 : « Histoire et littérature », in *Au bord de la falaise*, Paris, Albin Michel, 269-287.

- COSSUTTA Frédéric 2004 : « Catégories descriptives et catégories interprétatives en analyse du discours », in J.-M. Adam, J.-B. Grize et M. Ali Bouacha éd. *Texte et discours : catégories pour l'analyse*, Dijon, Editions Universitaires de Dijon, 189-213.
- GUILHAUMOU Jacques 1993 : « À propos de l'analyse de discours : les historiens et le "tournant linguistique" (l'exemple du porte-parole pendant la Révolution française) », *Langage & Société* 65, Paris, Maison des sciences de l'homme, 5-39.
- 2002 : « Le corpus en analyse de discours : perspective historique », *Corpus* 1, Nice, 21-49.
- HEIDMANN Ute 2005a : « Comparatisme et analyse de discours. La comparaison différentielle comme méthode », in *Sciences du texte et analyse de discours. Enjeux d'une interdisciplinarité*, J.-M. Adam & U. Heidmann éd., Genève, Slatkine, 99-118.
- 2005b : « Epistémologie et pratique de la comparaison différentielle », in *Comparer les comparatismes*, M. Burger & C. Calame éd., Lausanne, *Etudes de Lettres*, 141-159.
- KANELLOS Ioannis 1999 : « De la vie sociale du texte. L'intertexte comme facteur de la coopération interprétative », *Cahiers de Praxématique* n° 33, Montpellier, Université Paul-Valéry, 41-82.
- KLEIBER Georges 1994 : « Contexte, interprétation et mémoire : approche standard VS approche cognitive », *Langue Française* 103, Paris, Larousse, 9-22.
- KUENTZ Pierre 1981 : « Les "oublis" de la nouvelle rhétorique », in *Matérialités discursives*, CONEIN et al. éd., Presses Universitaires de Lille, 35-43.
- LUGRIN Gilles 2000 : « Les ensembles rédactionnels : multitexte et hyperstructure », *Médiatiques*, Louvain-la-Neuve, Observatoire du récit médiatique, 34-36.
- 2001 : « Le mélange des genres dans l'hyperstructure », *Semen* n° 13, Besançon, 65-96.
- MAINGUENEAU Dominique 2004 : « Retour sur une catégorie : le genre », in J.-M. Adam, J.-B. Grize et M. Ali Bouacha éd. *Texte et discours : catégories pour l'analyse*, Dijon, Editions Universitaires de Dijon, 107-118.
- MALRIEU Denise 2002 : « Stylistique et statistique textuelle », in *6èmes JADT-Saint-Malo*, Rennes, IRISA-INRIA, 483-494.
- MAYAFFRE Damon 2002 : « Les corpus réflexifs : entre architextualité et hypertextualité », *Corpus* 1, Nice, 51-69.
- MOIRAND Sophie 1999 : « Les indices dialogiques de contextualisation de la presse ordinaire », *Cahiers de Praxématique* n° 33, Montpellier, Université Paul-Valéry, 145-184.
- 2004 : « Le Texte et ses contextes », in J.-M. Adam, J.-B. Grize et M. Ali Bouacha éd. *Texte et discours : catégories pour l'analyse*, Dijon, Editions Universitaires de Dijon, 129-143.
- PEYTARD Jean 1971 : « La rature de Dazet, ou la métamorphose du sens », *Littérature* 4, Paris, Larousse, 68-89.
- 1982 : « Les variantes de ponctuation dans le chant premier des *Chants de Maldoror* », in *La genèse du texte : les modèles linguistiques*, C. Fuchs et al. éd., Paris, Editions du CNRS, 13-71.
- 1993 : « D'une sémiotique de l'altération », *Semen* 8, Besançon, Presses Universitaires de Franche-Comté, 145-177.
- RASTIER François 1998 : « Le problème épistémologique du contexte et le statut de l'interprétation dans les sciences du langage », *Langages* 129, Paris, Larousse, 97-111.
- 2001 : *Arts et sciences du texte*, Paris, P.U.F.
- 2004 : « Poétique et textualité », *Langages* 153, Paris, Larousse, 120-126.
- TANGUY Ludovic & THLIVITIS Théodore 1999 : « Parcours interprétatifs (inter)textuels : vers une assistance informatique », *Cahiers de Praxématique* 33, Montpellier, Université Paul-Valéry, 185-215.

- VAN DIJK Teun A. 1997 : « Cognitive Contexte Models and Discourse », in M. Stamenow éd., *Language Structure, Discourse and the Access to Consciousness*, Amsterdam, Benjamins, 189-226.
- 1999 : « Context Models in Discours Processing », in H. van Oostendorp & S. Goldman éd., *The Construction of Mental Representations During Reading*, London, Lawrence Erlbaum, 123-148.