

Les sphères de contextualisation. Réflexion méthodologique sur les passages de texte à texte (s) et la constitution des corpus

Vincent Capt, Jérôme Jacquin et Raphaël Micheli



Édition électronique

URL : <http://journals.openedition.org/corpus/1685>

ISSN : 1765-3126

Éditeur

Bases ; corpus et langage - UMR 6039

Édition imprimée

Date de publication : 15 novembre 2009

Pagination : 129-147

ISSN : 1638-9808

Référence électronique

Vincent Capt, Jérôme Jacquin et Raphaël Micheli, « Les sphères de contextualisation. Réflexion méthodologique sur les passages de texte à texte(s) et la constitution des corpus », *Corpus* [En ligne], 8 | 2009, mis en ligne le 01 juillet 2010, consulté le 01 mai 2019. URL : <http://journals.openedition.org/corpus/1685>

**Les sphères de contextualisation.
Réflexion méthodologique sur les passages
de texte à texte(s) et la constitution des corpus**

Vincent CAPT
Jérôme JACQUIN
Raphaël MICHELI
Université de Lausanne – Faculté des Lettres

1. Introduction

Cette contribution propose d'opérer un retour réflexif sur une pratique d'analyse, menée au sein d'un groupe de recherches en analyse textuelle des discours. L'ancrage pratique de la réflexion épistémologique rend moins compte de la phase terminale d'un projet que d'une articulation globale de problèmes méthodologiques communs à un paradigme, touchant à la systématisation des rapports texte / corpus. La réflexion est menée à partir d'objets diversifiés¹, dont émerge un questionnement commun, qui porte sur la *contextualisation* – comprise ici comme le *processus par lequel le chercheur tente d'établir la pertinence d'une mise en relation entre un texte et un autre (ou plusieurs autres) et de leur regroupement au sein d'un corpus*. Il s'agit ainsi de s'intéresser aux *passages de texte à texte(s)* impliqués dans la constitution d'un corpus, et de se demander à quelles *conditions méthodologiques* ces passages peuvent être effectués. Toute constitution de corpus est, on le

¹ Il s'agit de trois corpus sur lesquels nous travaillons et que nous aurons l'occasion d'évoquer à plusieurs reprises dans cet article : (i) lettres d'artistes placés en institution psychiatrique et archivés à la Collection de l'Art Brut à Lausanne, (ii) comptes-rendus écrits et officiels de débats parlementaires français relatifs à l'abolition de la peine de mort, (iii) débats politico-médiatiques organisés par des associations d'étudiants sur un campus universitaire, filmés et retranscrits par le chercheur lui-même.

sait, inséparable d'un objectif d'analyse² et implique de mobiliser des catégories propres à conduire ladite analyse. Dans cet article, nous nous concentrerons davantage sur les principes justifiant le regroupement de textes que sur la nature des catégories sollicitées pour l'analyse de ceux-ci. L'objectif n'est pas d'établir une marche à suivre : nous souhaitons plutôt, à partir d'exemples concrets, faire émerger quelques problèmes susceptibles d'intéresser l'analyste des discours.

Avant d'aborder frontalement la question de ces passages de texte à texte(s), il convient, au préalable, de préciser le sens que nous accordons aux notions fondamentales de texte, de cotexte et de contexte, et de corpus. On ne prétend bien évidemment pas aborder ici ces notions dans toute leur complexité, ni restituer au lecteur les multiples ramifications des débats auxquels elles donnent lieu en sciences du langage. Il s'agit, plus modestement, de formuler les postulats théoriques qui guident notre pratique.

A un niveau très général, nous dirons, pour commencer, que les *textes* constituent l'*objet propre* d'une approche comme la nôtre. Ce postulat se justifie à un niveau *empirique* : l'activité langagière des locuteurs prenant la forme de *textes* (non de morphèmes, de syntagmes ou de phrases isolés), ceux-ci constituent de plein droit les objets, certes non exclusifs, d'une science du langage. Nous considérons ainsi, avec J.-P. Bronckart, qu'« une démarche de sciences du langage se doit d'aborder d'abord les *textes*, en tant qu'ils constituent les réalisations empiriques premières de l'ordre langagier » (2008 : 39)³.

Il est impossible d'entrer ici de plain-pied dans la problématique de la définition du texte et des critères souvent hétérogènes qui la sous-tendent. Nous nous contenterons d'affirmer trois points très généraux, mais incontournables. (i) Les textes sont des productions *en langue naturelle* (Rastier

2 Dans notre cas, l'objectif d'analyse concerne des phénomènes comme la subjectivation du sujet de l'énonciation épistolaire, la construction discursive de l'émotion en contexte argumentatif ou encore la négociation interactionnelle des identités lors de débats.

3 D'autres linguistes spécialistes du texte formulent un argument similaire, au premier rang desquels Jean-Michel Adam (1999 : 40) et François Rastier (2005 : 31).

parle de « *suite[s] linguistique[s]* », 2001 : 21), quand bien même ils articulent très fréquemment le verbal à d'autres systèmes sémiotiques, problème qui ne nous retiendra pas ici. (ii) Conformément à l'exigence d'empiricité énoncée plus haut, on ne parle de « textes » qu'au sujet de productions verbales *attestées* (c'est ce que Rastier appelle le « principe d'objectivité », 2001 : 21). (iii) Il n'est pas possible, selon nous, de fournir une définition *essentialiste* au moyen d'une liste de propriétés structurelles que tout texte exhiberait minimalement. En revanche, il existe un critère essentiel de textualité, qui a trait à la possibilité, pour une production verbale, de donner lieu à un *jugement de complétude*⁴. Il importe de souligner qu'un tel jugement n'est pas l'apanage du point de vue savant et du travail *a posteriori* de l'analyste. La complétude fait l'objet d'un travail de *marquage* qui est le fait des participants eux-mêmes, et qui se manifeste de façon explicite ou implicite dans leurs productions verbales. La problématique de l'« unité » du texte résulte donc nécessairement d'une dialectique entre l'activité du locuteur et de l'allocutaire, d'une part, et celle de l'analyste, d'autre part. Ensuite – et peut être en conséquence de ce qui précède – les jugements de complétude présentent toujours un caractère de *relativité* : face à une production verbale un tant soit peu étendue et complexe, l'analyste pourra être amené à opérer plusieurs segmentations et à considérer comme « texte » des zones de taille variable. Ainsi, pour prendre l'exemple du débat parlementaire, on pourra considérer comme « texte » l'intervention d'un député particulier lors de la discussion d'un projet de loi : ce jugement de complétude s'appuiera sur un marquage *péritextuel* (le président donne la parole au député et signale la fin de l'intervention en donnant la parole à un autre député ou en levant la séance) et sur des *indices de structuration textuelle* (par exemple une division en

4 Nous rangeons sous ce terme de complétude deux types d'effets que les textes sont réputés produire, et qui ont été souvent décrits en linguistique textuelle : la cohésion, qui repose sur des moyens linguistiques assurant les liens intra- et inter-phrastiques, et, dans une optique sémantico-pragmatique, la cohérence, qui concerne l'ajustement d'une production verbale à une situation de communication. Voir, pour une synthèse, Adam *in* Charaudeau et Maingueneau 2002 : 99-100.

parties héritées de la rhétorique classique avec exorde, exposé des arguments et péroraison). Mais le chercheur pourra aussi considérer comme « texte » non pas l'intervention d'un seul député, mais, de façon plus globale, la discussion d'un amendement : il arguera, à nouveau, d'un marquage péritextuel (le président ouvre la discussion et clôt celle-ci en faisant voter l'assemblée sur l'amendement) et d'une structuration textuelle marquée par d'autres indices de complétude (les députés se succèdent pour soutenir ou, au contraire, critiquer ledit amendement).

Si le texte est l'objet propre de notre approche, il n'est cependant pas son horizon ultime. A la suite notamment de Rastier, nous considérons le texte comme « l'unité *minimale*, et le corpus l'ensemble dans lequel cette unité prend son sens » (2005 : 31). Nous réclamant d'une analyse textuelle des discours à orientation herméneutique, nous concevons et construisons nos corpus comme systèmes interprétatifs de mise en relation textuelle, donnant ainsi une importance fondamentale aux rapports privilégiés que les textes – unités minimales de construction, mais non d'analyse, du corpus – sont susceptibles d'entretenir entre eux : la généricité, sur laquelle nous reviendrons abondamment, mais aussi l'intertextualité (un texte peut pointer de façon explicite ou implicite vers un autre texte qui semble être « présent » en son sein) et la métatextualité (un texte peut faire l'objet d'une glose dans un autre texte).

Si les modes de traitement des textes et l'identification des différentes unités constitutives de la textualité peuvent être communs, notre conception de la construction du corpus se distingue d'une autre, celle fonctionnant par accumulation non génériquement déterminée d'occurrences d'un phénomène linguistique. Cette approche privilégie d'autres objectifs, la plupart du temps la description des faits langagiers (un connecteur, un pronom, un type de construction syntaxique, ...) dans ce qu'ils ont de plus systématique, ce qui implique forcément une part d'arrachement aux spécificités – notamment génériques – des textes.

La constitution de nos corpus dépend donc crucialement d'un processus de contextualisation. Dans le présent article,

Les sphères de contextualisation

nous parlerons de *contextualisation* plutôt que de *contexte*, car le terme permet de désigner à la fois un *processus* (le fait de contextualiser, pour les participants à l'interaction *et* pour le chercheur) et un *résultat* (l'ensemble de textes qui résultent du processus et qui forment le corpus). La contextualisation a fait l'objet de nombreux travaux marquants en ethnographie de la communication et en analyse conversationnelle et interactionnelle (voir le volume collectif Duranti et Goodwin 1992), ainsi qu'en analyse critique du discours (notamment les travaux de Van Dijk). Ces travaux se signalent, au-delà de leurs différences, par une idée : le contexte est avant tout une construction des participants et, suivant cette perspective émique⁵, l'étude de la contextualisation est l'étude des activités par lesquelles les participants sélectionnent les traits pertinents de la situation dans laquelle ils sont en train d'interagir. Notre propos se focalisera sur la contextualisation *savante* : nous nous intéresserons au processus par lequel le chercheur décide de mettre en relation un texte avec un ou plusieurs autres textes et de regrouper ceux-ci au sein d'un corpus, ainsi qu'aux types de justification qui sont susceptibles de fonder une telle décision. Nous verrons que certaines de ces justifications sont en lien avec les activités de contextualisation prises en charge par les participants eux-mêmes : il s'agira donc d'articuler au mieux le point de vue savant et le point de vue émique.

Notre pratique de recherche privilégie trois sphères de contextualisation – générique, auctoriale et thématique. Dans un premier temps (§ 2.), ces sphères seront présentées séparément : nous nous attacherons à les définir et à envisager leurs modes de saisie par l'analyste. Dans un second temps (§ 3.), nous tenterons de voir comment elles peuvent être hiérarchisées et articulées.

⁵ Depuis le linguiste Kenneth Pike, on distingue traditionnellement la perspective émique, consistant à approcher l'objet d'étude à partir des représentations subjectives des participants à la communication, de la perspective étique qui vise à établir les causes ou les facteurs systématiques et globaux surplombant ces mêmes participants.

2. Identification et saisie des sphères de contextualisation

2.1 La contextualisation générique

Nous admettons que tout texte est « susceptible d'être qualifié en terme de *genre* » (Bronckart 1997 : 75) ou, selon les termes d'Adam & Heidmann, que « dès qu'il y a texte [...], il y a *effet de généricité* – c'est-à-dire inscription de cette suite d'énoncés dans une classe de discours » (2006 : 23).

Si l'importance du palier générique fait aujourd'hui largement consensus en analyse des discours (c'est-à-dire au delà de sa restriction au seul champ littéraire), il n'est pas évident de s'y retrouver dans la nébuleuse de définitions, de grilles de reconnaissance et de typologies. Notre but étant de considérer le genre comme *un des facteurs privilégiés⁶ de contextualisation de texte à texte(s)*, nous avons tout d'abord besoin d'une définition – la plus opératoire et consensuelle possible – du genre. Profitant du fait qu'Adam & Heidmann (2006) s'en revendiquent, nous choisissons celle proposée par Bronckart dans *L'acquisition des discours* :

formes communicatives historiquement construites par diverses formations sociales, en fonction de leurs intérêts et de leurs objectifs propres; genres de la sorte socialement « indexés » et qui sont plus largement [...] à la fois producteurs et produits de modalités spécifiques d'élaboration des connaissances (1996 : 56).

Néanmoins, au-delà d'une définition consensuelle des genres, nous avons surtout besoin de *critères* permettant de rattacher une occurrence textuelle à un prototype générique, ceci afin de

⁶ Le qualificatif « privilégié » n'est pas à prendre à la légère. Que ce soit Bakhtine (1984) ou Bronckart (par ex. 2008), les auteurs soumettent nos autres sphères de contextualisation (identité et place des participants ainsi que thèmes abordés) à l'aune de la généricité. Rastier établit même cette dernière – sa position se fait encore plus catégorique dans d'autres articles – comme premier facteur de contextualisation pertinent au sein d'un corpus : « Par le biais notamment de son genre, le texte pointe sur son corpus et sélectionne en quelque sorte les autres textes qui permettent de l'interpréter [...]. Le premier cercle du corpus, corpus nécessaire mais pas toujours suffisant, est ainsi constitué des textes du même genre » (Rastier 2004 : 125, cf. aussi 2001 : 83 & 232).

Les sphères de contextualisation

justifier la mise en corpus d'au moins deux textes dont on postule et, à la fois, tente d'établir une généricité identique (genre épistolaire, genres du débat public et du débat parlementaire). Du fait que notre corpus de travail propose (ou plutôt tend à proposer) un parcours interprétatif (parmi les autres possibles) allant de texte à texte(s) au sein d'une même sphère de généricité, se pose donc de manière urgente la question de la *reconnaissance* d'un genre. Or nous sommes déjà contraints d'abandonner toute velléité formaliste ou systématique (si ce n'est dans la rigueur de l'approche même du phénomène) :

L'identification d'un genre n'est pas un raisonnement abstrait, fondé sur le repérage d'ensembles de propriétés définies. Il s'agit plutôt de regroupements par « airs de famille ». Les genres sont des catégories prototypiques définissables par des tendances ou des gradients de typicalité, par des faisceaux de régularités et des phénomènes de dominante. (Adam & Heidmann 2006 : 30)

Dans le souci d'armer méthodologiquement ce « jugement de participation » (*ibid.*) d'un texte à un genre et de rendre compte de ce dernier sous forme de « *faisceau* de critères » (Rastier 2001 : 253), nous proposons ici quelques jalons d'un parcours méthodique de reconnaissance de la généricité d'un texte. Nous nous inspirons d'un article de Jean-Claude Beacco (2004) qui propose *trois approches complémentaires* pour la reconnaissance et l'analyse de la généricité en sciences du langage, les trois approches relevant en fin de compte d'une démarche progressant par paliers de complexité croissante et de valeur interprétative ajoutée. L'objectif de notre contribution restant méthodologique, nous restreignons et reformulons la présentation de ces trois dimensions de manière à les faire entrer en résonance avec nos propres pratiques de recherche.

a) Le genre par les « noms des genres »

Partant du constat que le genre est avant tout une *catégorie ordinaire*, quotidienne, de traitement de la communication verbale, il est possible de reconnaître des genres – et surtout des

textes du même genre – par les catégorisations que les locuteurs eux-mêmes proposent. Ces derniers font explicitement référence à des genres, en proposant des *formulations* à toutes fins utiles (*i.e.* afin de se rendre mutuellement reconnaissable ce qu'ils sont en train de faire), donnant ainsi aux mentions génériques le statut de base contractuelle pour l'interaction en cours, que ces mentions soient situées au niveau transtextuel⁷ ou textuel. Dans nos corpus, des labels tels que « débat » (voire « débat public » ou « débat parlementaire ») ou « lettre » servent aux locuteurs à s'orienter vers la reconnaissance et l'accomplissement de l'activité proposée.

b) Le genre par les régularités observées au niveau de la textualité et de la transtextualité

Adam & Heidmann (2004 puis 2006 : 31 ss), prolongeant les réflexions de Bakhtine (1984 : 265) qui identifiait trois dimensions textuelles (contenu thématique, style et construction compositionnelle), proposent six niveaux textuels susceptibles d'être « engagés » par la généricité et donc d'en constituer les indicateurs potentiels.

Au niveau textuel, tant le genre épistolaire que le genre du débat public ou du débat parlementaire impliquent de manière singulière les niveaux *sémantique*, *énonciatif*, *pragmatico-argumentatif*, *stylistique*, *compositionnel* et *médiatique* (dans le sens du support exploité). Ainsi, le genre épistolaire et le genre du débat public ou parlementaire s'opposent essentiellement au vu de différences de composition (en termes de plans de texte régulièrement constatés), de support médiatique (écrit pour l'épistolaire, oral pour les débats

⁷ Dans le cas d'un débat public, on trouve régulièrement des « mentions de genre » (« débat », « débat public », « débat contradictoire », etc.) dans les cotextes participant à la médiatisation de l'événement (affiches, annonces par courrier électronique, etc.). Toujours au niveau transtextuel, on s'intéressera également aux cas d'intertextualité, lorsqu'un texte « indépendant » (dans le sens de « non génériquement déterminé » à l'inverse par exemple des textes médiatiques cités ci-dessus qui entretiennent un rapport générique aux débats publics) pointe (par exemple par citation) sur le texte cible en lui conférant, par mention, un genre (par ex. « lundi soir, avait lieu un débat public contradictoire [...] » lisible dans un compte-rendu journalistique).

Les sphères de contextualisation

public et parlementaire) ou encore de régime énonciatif (la réciprocité et l'alternance *je-tu* est de mise pour l'épistolaire alors que les débats fonctionnent de manière trilogale).

Adam & Heidmann définissent en outre quatre plans transtextuels parmi lesquels deux sont susceptibles de témoigner de régularités propices à indexer un genre. Au niveau *péritextuel*, les frontières du texte, outre qu'elles contiennent parfois une mention de genre, peuvent être génériquement déterminées (normes régissant les titres par exemple). Le niveau *cotextuel* implique quant à lui les textes composant l'entourage « direct » du texte. Le genre du commentaire journalistique, par exemple, peut s'identifier *via* son inscription dans une « aire scripturale » plus large (la page de journal).

c) Le genre par l'inscription dans les sphères sociales d'utilisation

Au delà de l'identification du genre par le repérage des régularités textuelles, il subsiste encore une piste d'analyse, quand bien même Beacco la considère comme une approche plus interprétative qu'identifiante de la généricité. En se basant sur les situations de production / réception des textes, il est possible de prendre en compte la généricité de ces derniers en termes de *distribution*, de *valeur* et de *buts* dans la sphère d'activités sociales analysée⁸. Les genres seraient ainsi identifiés au sein d'un système de genres rendu pertinent par leur institution d'accueil et de circulation. Si cette approche ne tend pas à proposer des critères d'identification à prétention auto-suffisante, elle apporte néanmoins un complément bienvenu aux deux perspectives précédentes.

Dans notre cas, les sphères d'activités du politique (institution parlementaire, espace public démocratique) ou du médical (institution asilaire) sont pertinentes. Dans le cas des débats parlementaires, nous voyons que tant ces derniers que les comptes-rendus qui les rendent accessibles sont des genres indexés à une institution particulière.

⁸ Il s'agit notamment de l'approche des genres en Critical Discourse Analysis, telle qu'elle est pratiquée entre autres par l'équipe de Ruth Wodak.

Le poids donné à la description linguistique des genres, c'est-à-dire à la reconnaissance des effets de généricité produits par un texte à travers ses régularités formelles, ne doit pourtant pas nous faire oublier l'écueil pointé par Beacco lorsqu'il s'agit de considérer le lien du genre au corpus : le genre participe à la construction du corpus, tandis que ce dernier participe à la définition du genre. Il y a là un *risque de circularité*, et Beacco propose comme solution l'entretien d'une complémentarité et d'une articulation fortes entre la description formelle des régularités (b) et la prise en compte des labels génériques ordinaires (a) :

En matière de description linguistique des genres, la prudence méthodologique commande que l'on s'en tienne à des ensembles de textes dont on examine la consistance par rapport à des noms de genres attestant du caractère partagé des représentations sociales qui s'y attachent. (2004 : 113)

Partant de l'idée (i) que les sphères de contextualisation favorisent des parcours intertextuels propices à l'élaboration du sens et (ii) que la généricité constitue, avant l'auteur et la thématique que nous abordons ci-dessous, le principe majeur de ces ajouts textuels en corpus, il nous faut ainsi admettre que postuler et établir une généricité identique entre deux textes est un acte aux conséquences herméneutiques importantes, qu'il s'agit d'outiller le plus possible.

2.2 La contextualisation auctoriale

La deuxième sphère de contextualisation de texte à texte(s) concerne l'instance en charge du texte, que nous nommons, après Bronckart (1997), l'auteur. Ici, cette instance vaut par l'articulation de deux principes, celui de l'origine et celui de la responsabilité :

S'agissant de la production langagière ou textuelle, la notion d'auteur semble [...] bien correspondre à la définition générale qu'en proposent les dictionnaires : « celui qui est à l'origine » et « celui qui est responsable ». (1997 : 318)

Les sphères de contextualisation

L'origine relève de l'acte matériel d'une production langagière. Le plus souvent, elle coïncide avec le principe de responsabilité, qui concerne l'assomption du texte par l'auteur. La responsabilité a ainsi trait à une instance qui performe le texte. Dans certains genres de discours, il peut y avoir disjonction entre ces deux principes : les interventions officielles d'un politicien, par exemple, sont le plus souvent élaborées en amont par divers conseillers en communication. Pour autant, ce qui prime, *pour la reconnaissance spécifique de l'auteur*, c'est la responsabilité de ce dernier, en tant qu'il est amené *in fine* à porter le texte *comme l'ayant initialement produit*. L'auteur est ainsi présenté comme l'instance à laquelle on impute la responsabilité d'un texte.

Bien qu'il soit peu consensuel en sciences du langage, nous choisissons ce concept car il permet de désigner *une instance renvoyant à un seul et même organisme*. Un tel organisme ne se restreint toutefois pas nécessairement à une personne humaine unique. Les cas d'institutionnalisation de l'instance auctoriale, telles la signature-logo d'un parti dans un communiqué politique ou la marque d'un produit dans le discours publicitaire, sont en effet légion. Il en va de même pour les cas de pluralisation de l'auteur, comme la co-signature du présent article⁹. Dans tous ces cas, ce qui importe, *in fine*, c'est que la contextualisation de texte à texte(s) s'établisse selon un rapport d'*identité auctoriale*, où la même instance se retrouve d'un texte à l'autre.

Si l'enjeu est de regrouper plusieurs textes selon une même instance auctoriale, se pose alors le problème de son *attestation*. À partir des textes, il s'agit d'identifier les critères permettant un jugement d'auctorialité. Plus que de simplement réguler le passage intertextuel, la sphère auctoriale est à considérer comme une instance *discursive* de contextualisation, c'est-à-dire empirique, particularisée et historicisée. Deux critères permettent de lui attribuer des attestations fortes : soit

⁹ Les cas de pluralisation d'auteurs sont complexes. Dans cette configuration, ce qui prime globalement, c'est que les différents auteurs soient réunis au sein d'une seule et même instance, collective, qui, pour ainsi dire, les fédère.

les attestations correspondent à des désignations par le nom de l'auteur, soit elles sont des indices qui le désignent par un statut social (comme le Président de la République) ou un rôle interlocutif (par exemple le conférencier). Les indices de personne nécessitant une connaissance de la situation de communication pour référer et ne livrant leur référence que relativement à l'instance auctoriale attestée, ils ne sauraient suffire. Dans le même ordre d'idée, les régularités idiolectales *esquissent* l'identification d'un auteur, mais ne suffisent pas à le référer pleinement. Dans le genre épistolaire, le cas où manquent la signature (en fin du plan de texte de la lettre) et la mention de l'expéditeur (sur le péricaractère de l'enveloppe) est significatif de conditions insuffisantes d'attestation.

Ajoutons encore qu'il est utile d'identifier deux possibilités d'attestation : l'*auto-attestation* et l'*hétéro-attestation*. Dans le premier cas – par ex. la signature au bas d'un tract –, l'individu revendique lui-même l'auctorialité de la production verbale. Dans le second cas, l'auctorialité d'une production verbale est attestée dans le texte d'un autre auteur : on prendra pour exemple, au Parlement, les mentions des participants par le président de séance lors de l'allocation des tours de parole.

Lorsqu'elle est mobilisée *seule*, la contextualisation auctoriale fait l'objet de suspicions en sciences du langage, alors qu'elle peut recouper les habitudes d'autres champs d'étude, où l'on peut regrouper des textes en œuvres complètes d'un auteur, et qui peuvent donner lieu à des études monographiques. Ce même constat s'applique à la contextualisation thématique, que nous abordons à présent. Nous défendrons plus loin (§ 3.) l'hypothèse que ces deux sphères gagnent à être mobilisées dans le cadre d'*activations conjointes*.

2.3 La contextualisation thématique

Pour ce qui est de la contextualisation *thématique*, la question qui se pose est la suivante : dans quelle mesure est-il pertinent d'activer le critère du thème¹⁰ lors des passages de texte à

10 On aurait pu attendre ici, en lieu et place de la notion certes problématique de « thème », celle d'« objet de discours », développée notamment dans le cadre de la logique naturelle de Jean-Blaise Grize. Nous préférons toutefois parler de contextualisation thématique que de contextualisation

Les sphères de contextualisation

texte(s) qui sont inhérents à la constitution d'un corpus ? Nous ne prétendons bien sûr pas pouvoir traiter de ce problème dans toute sa complexité : il s'agit plutôt d'identifier quelques questions méthodologiques auxquelles l'analyste des discours est susceptible d'être confronté.

Dans une perspective textuelle, le thème n'est pas une notion aisée à circonscrire. La tentative la plus récente et la plus développée est celle de François Rastier (voir principalement 2001 : 189-226). Notre propos n'est pas ici d'entrer de plain-pied dans une discussion de nature définitionnelle sur cette notion. Rastier montre bien qu'une (re)définition textuelle du thème doit se distinguer d'une perspective littéraire (en particulier de la critique dite « thématique »), et d'une analyse communicative de la phrase qui s'intéresse à l'articulation thème / rhème. Nous prendrons appui sur la définition générale¹¹ que propose Rastier : « Structure stable de traits sémantiques (ou *sèmes*), récurrente dans un corpus, et susceptible de lexicalisations diverses » (2001 : 197). Cette définition a l'avantage de distinguer deux *modes d'inscription* du thème dans le matériau textuel et, de façon corollaire, deux *modes de saisie* par l'analyste. Ces deux modes vont souvent – mais pas nécessairement¹² – de pair. L'idée de « lexicalisation » indique qu'un thème peut se *condenser* en une unité qui, pour ainsi dire, le dénomme et qui va potentiellement du morphème au syntagme. L'idée de « récurrent[ce] » de « traits sémantiques » indique pour sa part que, s'il peut être lexicalisé par un seul mot, le thème n'est pas pour autant une unité lexicale : il se *diffuse*, parfois à l'échelle du texte entier, parfois à l'échelle

par l'objet du discours, et cela pour la raison suivante : tel qu'il est défini par Grize, l'objet du discours pose le problème de « l'articulation [...] à la matérialité langagière – tout particulièrement à travers la question des marques permettant [son] repérage » (Sitri 2003 : 36). La notion de thème dans la théorie du texte de Rastier a l'avantage d'éviter cet écueil par le marquage lexical et sémique.

11 Une distinction est ensuite opérée entre thème générique et thème spécifique, que nous ne pouvons commenter ici.

12 Rastier rappelle à juste titre qu'il est des thèmes sans lexicalisation privilégiée.

intermédiaire de la « période ou du paragraphe » (Rastier 2001 : 194).

Dans le cadre de notre réflexion sur la contextualisation et les passages de texte à texte(s), nous nous intéressons surtout aux *processus émiques de « lexicalisation » du thème*, c'est-à-dire à la *façon dont les locuteurs pointent le thème* (du texte qu'ils produisent eux-mêmes et d'autres textes auxquels ils répondent). Dans le genre fortement ritualisé du débat parlementaire, on sera par exemple attentif aux formules d'ouverture prononcées par le Président (« L'ordre du jour appelle la discussion du projet de loi portant abolition de la peine de mort ») et aux exordes des interventions des députés. L'étude du corpus de débats publics montre quant à elle que le médiateur a entre autres pour tâche de circonscrire l'objet du débat en annonçant lexicalement les thèmes (« On va commencer par le thème du financement dans l'enseignement »).

Dans les développements qui précèdent, nous voulions avant tout nous demander de quelle manière on peut *attester* les différentes sphères de contextualisation – en quoi celles-ci sont *reconnaissables*, si l'on peut dire. Il faut maintenant voir comment ces sphères peuvent être hiérarchisées et articulées, et les problèmes concrets que cela pose dans la pratique de l'analyse des discours.

3. Hiérarchisation et articulation des sphères de contextualisation

De façon générale, la réflexion en sciences du langage quant à la constitution et au traitement des corpus insiste sur les régularités *génériques*. Le genre apparaît, chez la plupart des auteurs, comme la *première sphère de contextualisation* et comme la voie privilégiée vers la création d'un corpus. Rastier a soutenu une telle position avec force en affirmant que « le corpus de textes du même genre s'impose en général : on juge [...] d'un poème par rapport à d'autres poèmes, et non par rapport à des romans » (2001 : 232)¹³. La primauté du genre se

¹³ Dans le domaine de la lexicologie quantitative, la réflexion de Damon Mayaffre fait de l'« homogénéité » une caractéristique fondamentale de

Les sphères de contextualisation

retrouve dans nos pratiques de recherche, où les sphères de l'auteur et de la thématique sont pensées *relativement* à la contextualisation générique. Dans notre approche, une telle prégnance du genre sur l'auteur et/ou la thématique signifie qu'il y a un *ordre* des sphères de contextualisation : générique d'abord et, partant, auctoriale et/ou thématique ensuite. Cela rend aussi compte du fait qu'il est possible d'activer *conjointement* les différentes sphères de contextualisation. A l'aide du diagramme ci-dessous, nous envisagerons trois¹⁴ cas d'intersection entre les sphères :

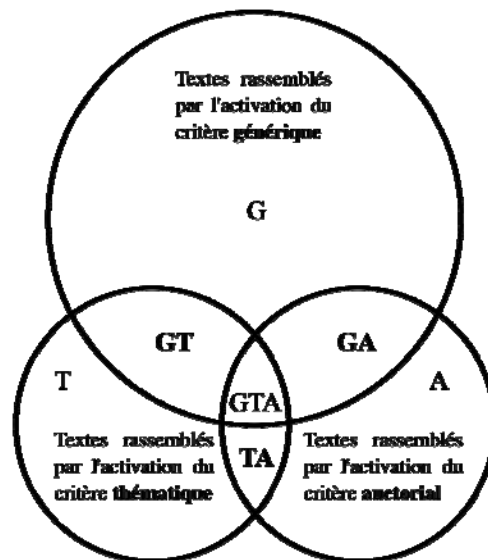


Fig. 1 : Les sphères de contextualisation.

tout corpus, en envisageant des textes « produits dans des conditions identiques » – on retrouve ici la contextualisation générique – ou « traitant d'un même sujet » (2000 : 39).

14 Le quatrième cas est représenté par la zone centrale du diagramme (GTA) et consiste en une réunion de textes relevant d'une même genericité, auctorialité et thématique. Ce cas ne sera pas abordé ici dans la mesure où il est moins attesté dans nos pratiques. Nous n'illustrerons pas non plus les activations non conjointes de sphères de contextualisation (zones G, T et A). Retenons ici qu'elles reçoivent, mise à part la sphère générique au bénéfice d'une autonomie plus importante, une valeur interprétative par leur mise en contraste avec les zones d'intersection.

Le premier cas envisageable est celui de l'intersection entre les sphères de contextualisation *générique* et *thématique* (zone GT du diagramme). C'est le parti que nous avons pris dans notre étude diachronique de la construction des émotions dans les débats parlementaires français sur l'abolition de la peine de mort¹⁵. Dans cette recherche, la pertinence des passages de texte à texte(s) et la cohérence du corpus étaient justifiées par le fait que les différents textes traitaient d'un thème similaire au sein d'un genre similaire. Lorsqu'on exploite un tel corpus, dont l'homogénéité se veut à la fois thématique et générique, on cherche à saisir les spécificités qu'exercent les normes d'un genre sur le traitement d'un thème.

Dans le cas d'un corpus génériquement déterminé comme celui constitué de débats publics – notre intérêt se portant sur les modalités identitaires de participation et sur l'argumentabilité des identités – la contextualisation générique-thématique (par exemple sous la forme de deux débats publics consacrés à la question du financement des études supérieures) permet de contraster des auteurs différents et leurs productions verbales respectives dans des contextes comparables.

Dans le même ordre d'idée, les sphères *générique* et *auctoriale* peuvent également se croiser (zone GA du diagramme). Telle est l'une des conjonctions exploitées lors d'un travail autour de la subjectivation du discours et du sujet de l'énonciation épistolaire, où un même auteur a été retenu en fonction des seules lettres qu'il a écrites. Dans le cas de l'épistolier et auteur d'Art Brut Samuel Daiber, cette intersection a permis d'enrichir le corpus de lettres contenues dans les archives de la Collection de l'Art Brut en y ajoutant de nouvelles lettres signées par le même auteur et contenues dans les archives médicales d'un hôpital psychiatrique¹⁶. La conjonction de ces deux sphères a permis de remarquer dans

15 Voir R. Micheli : « L'analyse argumentative en diachronie : le pathos dans les débats parlementaires sur l'abolition de la peine de mort », *Argumentation et Analyse du Discours*, n° 1, 2008, [en ligne], mis en ligne le 18 septembre 2008. URL : <http://aad.revues.org/index482.html>.

16 On peut préciser que dans ce cas le genre subit un déplacement subtil, en ce que les institutions initiales d'accueil et de circulation des textes diffèrent.

une diachronie la constance de la valeur illocutoire de ces lettres, l'homogénéité des destinataires auxquels l'auteur s'adressait, ainsi que la forte régularité de certaines formules idiolectales.

On peut encore, de façon certes davantage problématique, activer conjointement les sphères *thématique* et *auctoriale* (zone TA du diagramme). Dans le corpus de débats parlementaires sur l'abolition de la peine de mort, la contextualisation thématique-auctoriale apparaît incontournable dans les cas de Victor Hugo et de Robert Badinter. Les deux auteurs ont participé, respectivement en 1848 et en 1981, aux assemblées parlementaires, mais ont également investi le thème de l'abolition dans d'autres genres de discours (pour Badinter, il s'agit principalement de la plaidoirie, de l'article de journal et du récit autobiographique ; pour Hugo, la liste est plus longue et comprend par exemple la lettre ouverte et la préface de roman). Dans ces cas, l'intérêt de la contextualisation *thématique-auctoriale* était de faire apparaître, par comparaison, les spécificités *génériques* de l'argumentation que déploient les auteurs en faveur de l'abolition et, plus précisément, leur recours très variable à un *pathos* centré sur les souffrances endurées par le condamné durant l'exécution (recours très faible, voire nul dans le genre du débat parlementaire, et abondant dans les autres genres).

Dans le corpus rassemblant des occurrences de débats publics et où la question générique est au centre de l'intérêt, la contextualisation thématique-auctoriale n'a une fois encore de sens que dans une optique contrastive : il s'agit d'analyser les productions verbales d'un même auteur sur un même thème, mais dans des genres différents (tract, interview, lettre ouverte, pétition, conférence, etc.) afin de saisir les spécificités du genre ciblé quant aux modes de construction de l'argumentation, de production d'identités discursives, d'attestation de l'auteur, d'organisation thématique, etc.

4. Conclusion

L'objectif de cet article était de réfléchir à la contextualisation, que nous avons définie comme un processus par lequel le chercheur tente d'établir la pertinence d'une mise en relation

entre un texte et un autre (ou plusieurs autres) et de leur regroupement au sein d'un corpus. Nous avons choisi de mettre en évidence trois principes qui nous semblaient incontournables dans notre pratique d'analyse des discours, ces trois principes autorisant des passages de texte à texte(s) et dessinant ce que nous avons proposé d'appeler des *sphères de contextualisation*. Il s'agissait, dans un premier temps, d'identifier ces sphères et de s'interroger brièvement sur ce qui, dans les textes, permet de les attester. Dans un deuxième temps, il s'agissait d'envisager une hiérarchisation de ces sphères et leurs articulations possibles. Le diagramme proposé visait à montrer de façon synthétique ce que nos recherches empiriques illustrent : la contextualisation repose souvent sur l'activation conjointe de deux sphères au moins parmi les trois envisagées (à l'exception de la sphère générique qui peut être activée seule). La constitution d'un corpus implique que le chercheur privilégie des zones d'intersection entre les sphères, qu'il les fasse contraster et, de façon corollaire, qu'il en délaisse d'autres si elles s'avèrent moins productives en termes d'ajouts textuels ou si elles ne sont simplement pas pertinentes en regard de l'hypothèse de recherche. En opérant un retour réflexif sur une pratique d'analyse des discours et en interrogeant les conditions méthodologiques des passages de texte à texte(s), la modélisation esquissée ici visait modestement à contribuer à une « théorie des parcours interprétatifs » qui, selon Rastier (2001 : 111), reste à élaborer.

Références bibliographiques

- Adam J.-M. (1999). *Linguistique textuelle : des genres de discours aux textes*. Paris : Nathan, 208 p.
- Adam J.-M. (2008). *La linguistique textuelle. Introduction à l'analyse textuelle des discours*. Paris : Armand Colin [1^{ère} éd. 2005], 239 p.
- Adam J.-M. & Heidmann U. (2004). « Des genres à la généricité : l'exemple des contes (Perrault et les Grimm) », *Langages* 153 : 62-72.
- Adam J.-M. & Heidmann U. (2006). « Six propositions pour l'étude de la généricité », *La Licorne* 79 : 21-34.

Les sphères de contextualisation

- Bakhtine M. (1984). « Les genres de discours », in *Esthétique de la création verbale*. Paris : Gallimard [1^{ère} éd. 1952], pp. 263-308.
- Beacco J.-C. (2004). « Trois perspectives sur la notion de genre discursif », *Langages* 153 : 109-119.
- Bronckart J.-P. (1996). « L'acquisition des discours. Le point de vue de l'interactionnisme socio-discursif », *Le français dans le monde. Le discours: enjeux et perspectives*, N° spécial : 55-64.
- Bronckart J.-P. (1997). *Activité langagière, textes et discours : pour un interactionnisme socio-discursif*. Lausanne / Paris : Delachaux et Niestlé.
- Bronckart J.-P. (2008). « Genres de textes, types de discours et "degrés" de langue : hommage à François Rastier », *Texto!*, Vol. XIII.
- Charaudeau P. & Maingueneau D. (2002). *Dictionnaire d'analyse du discours*. Paris : Seuil.
- Duranti A. & Goodwin C. (1992). *Rethinking context : language as an interactive phenomenon*. Cambridge : Cambridge University Press.
- Mayaffre D. (2000). *Le poids des mots. Le discours de gauche et de droite dans l'entre-deux-guerres*. Paris : Champion.
- Rastier F. (1996). « Pour une sémantique des textes : questions d'épistémologie », *Texto !*
www.revue-texto.net/Inedits/Rastier/Rastier_PourSdT.html
- Rastier F. (1998). « Le problème épistémologique du contexte et le statut de l'interprétation dans les sciences du langage », *Langages* 128 : 97-111.
- Rastier F. (2001). *Arts et sciences du texte*. Paris : PUF, coll. « Formes sémiotiques ».
- Rastier F. (2004). « Poétique et textualité », *Langages* 153 : 120-126.
- Rastier F. (2005). « Enjeux épistémologiques de la linguistique de corpus », in Williams G. (éd.), *La linguistique de corpus*. Rennes : PUR, pp. 31-45.
- Sitri F. (2003). *L'objet du débat*. Paris : PSN.