

Chapitre 1 : Etat de l'art : Genres et discours scientifique

La notion de *genre* est de plus en plus présente, tant en linguistique, qu'en classification textuelle, ou qu'en didactique. Le concept a ainsi perduré de l'Antiquité à nos jours, et s'est vu approprié et redéfini par maintes disciplines des sciences humaines et sociales. S'il va de soi que nous n'énumérerons pas l'ensemble des perspectives qui ont un jour abordé la question des genres¹, nous dresserons un panorama raisonné et condensé de ces approches afin de présenter – et de légitimer – le cadre théorique d'analyse des genres qui fondera notre étude et nos analyses ultérieures. Ce n'est que dans un deuxième temps que nous introduirons et positionnerons notre objet, *i.e.* le genre de l'article scientifique dans son domaine linguistique.

1.1. Evolution/évaluation du concept de genre

Le terme « genres du discours » (*genera dicendi*) provient originellement de la tradition gréco-latine aristotélicienne. Les premiers modèles d'analyse discursive se limitent au modèle platonicien du discours (expression, réception et mimesis) dans lequel les facteurs sociaux, contextuels, culturels et dialectaux sont ignorés au profit d'un schéma rhétorique universel et écrit, dérivé des langues grecque et latine – ce qui semble surprenant dans un art qui prétend être concerné par une certaine efficacité interactionnelle. Les premières typologies génériques sont ainsi des systèmes de classements fermés : les genres oratoires, qui renvoyaient aux institutions de la cité, correspondaient à des stratégies argumentatives particulières, et étaient enseignés dans cette visée.

De même, les genres littéraires ont longtemps été présentés regroupés au sein d'une liste finie et normative, qui correspondait aux classements normatifs des collèges d'ancien régime :

Dans l'enseignement classique, elle permettait en effet d'attirer l'attention sur les co-occurrences de traits caractéristiques d'œuvres modèles et de les proposer comme des normes à intérioriser. Construite en vue de l'acquisition pratique des modèles, la notion de genre a été descriptivement adéquate tant qu'on s'est référé au corpus fermé des textes de la tradition. (Branca-Rosoff, 1999, p. 5)

En effet, les typologies des genres littéraires commencent à se périmer à partir du XVII^e siècle, les productions de l'époque ne correspondant plus au corpus fermé des œuvres enseignées. Le roman a ainsi eu du mal à trouver sa place. Les conventions génériques éclatent au XIX^e siècle, l'écriture cherchant désormais à transgresser les lois du genre :

Que l'œuvre « désobéisse » à son genre ne rend pas celui-ci inexistant ; on est tenté de dire : au contraire. Et ce pour une double raison. D'abord parce que la transgression, pour exister comme telle, a besoin d'une loi – qui sera précisément transgressée. On pourrait aller plus loin : la norme ne devient visible – ne vit – que grâce à ses transgressions (Todorov, p. 45, 46).

De manière générale, on prend progressivement conscience que le genre, loin d'être une catégorie figée dans une typologie immuable, est un objet en constante mutation, qui émerge, évolue et disparaît selon les époques et les contextes. Les genres deviennent des objets à décrire, plus que des unités de classement.

¹ D'autant que plusieurs études récentes s'y sont déjà consacrées, e.g Paltridge 1997, Rastier, 2001.

Parmi les travaux qui ont influencé nos conceptions du genre, mentionnons d'abord ceux de Bakhtine, qui ont connu un retentissement considérable en France de la fin des années 70 à nos jours, tant dans les théories littéraires ou linguistiques que dans d'autres branches scientifiques. C'est effectivement à cette date que l'analyse du discours, et des genres, attire l'attention de plusieurs branches, mettant en œuvre des approches différentes.

La particularité de Bakhtine est d'avoir relié les genres aux pratiques sociales en tant que sphères d'échange et de s'être intéressé au langage courant comme constructeur de genres :

Each separate utterance is individual, of course, but each sphere in which language is used developed its own *relatively stable types* of these utterances. These we may call *speech genres* (*The problem of speech genres*, p. 60)

Tous les domaines de l'activité humaine impliquant l'utilisation de la langue, la diversité des genres est infinie. Dans la mesure où ils permettent d'articuler l'étude de la langue à celle de la parole, la nature générale des genres doit être abordée, et leur description engagée. Or, on avait surtout privilégié l'étude et la description des genres littéraires, au détriment d'autres genres :

Dans une société, on institutionnalise la récurrence de certaines propriétés discursives, et les textes individuels sont produits et perçus par rapport à la norme que constitue cette codification. Un genre, littéraire ou non, n'est rien d'autre que cette codification de propriétés discursives. (Todorov, 1978)

Comme le souligne également Todorov (*ibid.*), les genres présentent cette particularité de pouvoir être décrits selon deux points de vue : le point de vue de l'utilisateur ordinaire, et celui de l'analyste. En ce sens, ils existent dans la compétence métalinguistique des sujets parlants : tout locuteur est capable d'identifier un genre et de restituer ses normes dans une situation donnée.

On comprend là pourquoi la question des genres a intéressé tant de disciplines et d'approches différentes : s'ils concernent bien entendu la linguistique, les genres peuvent être abordés selon des points de vue fort différents, tant sociologique, qu'anthropologique ou psychologique, et inspirer de nombreux axes d'analyse pluri- ou inter-disciplinaires.

Le genre est généralement perçu comme une entité à deux faces, l'une sociale et l'autre linguistique, ce qui rend son analyse complexe, ou simplifiée. Examinons d'abord les modalités de description de la dimension sociale des genres, et celles de son rattachement au linguistique.

1.1.1. L'inscription sociale des genres et les genres sociaux

On a d'abord souligné que l'étude des propriétés linguistiques des genres n'était pas envisageable sans prendre en compte leur pratique sociale d'appartenance. En effet, de nombreux travaux, et plus particulièrement des études de théorie littéraire (Lotman 1977, Sinclair 1982, Carter 1994), ont souligné qu'un genre ne pouvait être défini par ses propriétés strictement linguistiques ; par exemple, Selden et Widdowson ont montré qu'il n'y avait pas de raison linguistique intrinsèque à considérer les termes employés dans un ouvrage littéraire comme *littéraires* (1993) ; ce n'est que dans la mesure où l'ouvrage est considéré comme « littéraire » qu'il est possible de les lire ainsi.

La production et l'interprétation d'un genre semblent donc subordonnées à un ensemble de critères linguistiques et non linguistiques, imposés par une pratique sociale particulière. On a donc largement soutenu que l'étude d'un genre ne saurait se faire hors d'un cadre interdisciplinaire, qui instituerait un couplage des facteurs sociaux qui le déterminent aux propriétés linguistiques qui le caractérisent. L'étude du genre implique donc deux

dimensions : une dimension langagière et une dimension sociale, *champs homologues*, pour reprendre les termes bourdieusiens, dans la mesure où les formations discursives qui sont en concurrence dans le champ qu'elles constituent renvoient également à la compétition des groupes sociaux dans le champ social.

Toutefois, ce couplage est encore bien problématique, d'abord parce que les dimensions sociale et linguistique des genres se recouvrent largement :

La notion de genre est une notion biface qui fait correspondre une face interne (les fonctionnements linguistiques) avec une face externe (les pratiques socialement signifiantes). Mais les usagers de la langue utilisent également des termes de classification lorsqu'il y a non coïncidence entre les deux dimensions : tantôt ils privilégient la situation de communication, tantôt les marques formelles. L'instabilité de la relation entre formes et comportements sociaux institutionnalisés est une difficulté centrale pour toute définition *a priori* des genres. (Branca-Rosoff, 1999, p. 116)

C'est d'ailleurs pour cette raison que P. Achard (1995) distingue les genres et les registres², i.e. le linguistique et le social, et ce sont précisément les distorsions entre les deux niveaux qui intéressent sa sociologie. A l'instar des pratiques sociales, genres et registres ne sauraient être répertoriés au sein d'une typologie générale, le registre serait un fonctionnement social spécifique, sélectionné en fonction d'une hypothèse du sociologue, et analysé en genres.

Les genres sociaux ne semblent pas en effet être analysables au moyen de critères stables. Il est ainsi possible de prendre en compte différents paramètres, tant historiques, que sociologiques ou communicationnels, pour appréhender la dimension sociale des genres : le genre dépend en effet largement de sa représentation, et plus largement, de son interprétation et de sa production, variables selon les époques et les cultures. Il implique ainsi des conditions communicationnelles, ainsi que des conditions statutaires. Au nombre de celles-ci figurent par exemple la position de l'énonciateur, le statut de son co-énonciateur – pour reprendre le terme culiolien –, la représentation que le premier se fait du second, etc. L'importance de ces paramètres varie selon les genres, qui sont différemment normés. Par conséquent, des critères primordiaux pour décrire et caractériser un genre A se révéleront totalement inadéquats quand il s'agira d'analyser un genre B.

Cette instabilité a donné lieu à des définitions diverses des genres sociaux, selon les critères et les genres privilégiés : ainsi, l'ethnographie de la parole (Bauman et Sherzer, 1974), qui se concentre sur les moyens dont les locuteurs d'une communauté donnée disposent pour communiquer, les considère comme des *types d'événement langagier*, permettant l'emploi de catégories comme la conférence, la blague, le salut, etc.. Pour le courant de la nouvelle rhétorique, qui s'intéresse aux genres depuis une vingtaine d'année, le genre est un *moyen d'action sur le monde*, tandis que Bergmann et Luckmann, qui relie le constructionnisme social avec la sociologie de la connaissance (1995), parlent de *genre communicatif* : les genres sont perçus comme des modèles communicatifs socialement construits qui permettent de résoudre des problèmes de communication, estimés centraux pour la construction de la réalité sociale. La dialogique bakhtinienne a également influencé certaines conceptions du genre, et notamment le fait que les locuteurs ne construiraient pas leurs énoncés qu'en fonction des objets du discours et de leurs destinataires immédiats, mais également par rapport à la représentation particulière selon laquelle ils pensaient être interprétés, i.e. par rapport à un « super-destinataire » susceptible d'endosser diverses formes idéologiques (Dieu, la Vérité, l'histoire, la science, etc.). Le genre est ainsi défini en tant que *site d'action idéologique*,

² Opposition que l'on retrouve d'ailleurs dans la majorité des courants systémiques du genre.

imprégné des valeurs concrètes du jugement et exprimant les valeurs partagées de groupes particuliers.

Si les critères et les méthodologies varient selon les genres étudiés, l'étude d'un genre donné en tant qu'objet social *et* linguistique requerrait en amont la sélection d'un ensemble d'hypothèses et de descripteurs linguistiques et sociologiques, qui permettraient d'obtenir dans un premier temps une description des propriétés linguistiques et sociales du genre. Il conviendrait dans un dernier temps de faire émerger les lieux de corrélation entre les deux types de phénomènes observés. Une telle étude paraît toutefois difficilement envisageable hors d'un cadre collaboratif linguiste/sociologue, dans la mesure où elle nous semble nécessiter une véritable double compétence linguistique et sociologique.

Les approches les plus répandues, qui cherchent directement à corréler phénomènes linguistiques et sociologiques, nous semblent encore insatisfaisantes : on part généralement de la situation de communication ou d'hypothèses sociologiques spécifiques, et on cherche à leur associer des dispositifs linguistiques particuliers. La sélection des descripteurs linguistiques est donc orientée par leur interprétation/fonction sociale ultérieure, ce qui explique pourquoi on choisit généralement des marqueurs discursifs (pronoms personnels, modalisateurs³, etc.) et lexicaux (le plus souvent des phraséologies), catégories intuitives, et donc plus facilement interprétables.

Ce faisant, on exclut un grand nombre de catégories linguistiques pourtant constitutives de la langue et des discours, ce qui est discutable d'un point de vue linguistique. Il est toutefois à noter que ces restrictions ne sont pas l'apanage de la sociologie ; nombre d'études discursives, linguistiques ou non linguistiques, procèdent de même, et obtiennent d'ailleurs des résultats tout à fait satisfaisants par rapport aux objectifs descriptifs escomptés.

Notre objectif est en revanche autre, dans la mesure où nous visons d'abord à décrire la linguistique d'un genre. L'incidence du sociologique sur le linguistique nous semble donc relever d'une étape d'analyse ultérieure, que nous n'écarterons pas.

Si la dimension sociale des genres est difficilement paramétrable en sociologie, leur caractérisation linguistique n'est pas moins problématique, d'autant que la question des genres a longtemps été délaissée par la discipline.

1.1.2. La question des genres en linguistique

Force est de constater que la linguistique s'est longtemps cantonnée à l'étude de la phrase⁴ et que son domaine de prédilection est longtemps demeuré la morphosyntaxe. Cette limitation tient à l'histoire de la linguistique, et à la tradition grammaticale : la linguistique historique et comparée fondait en effet ses comparaisons sur l'analyse morphosyntaxique, et la dichotomie saussurienne langue-parole, qui a posé les jalons de la linguistique contemporaine, a entraîné une conception de la parole en tant que « matériau » permettant de retrouver le système de la langue : collecter et analyser des phrases ou des phonèmes pour en tirer des règles étant plus aisé que de recueillir des textes, certains domaines linguistiques ont progressé davantage que d'autres.

³ *Hedges* au sens de Lakoff (1972 : 195) : “words whose job is to make things fuzzier or less fuzzy”, de type *sort of, kind of*, etc.

⁴ Et à des unités d'analyse d'ordre inférieur bien entendu, comme le morphème ou le syntagme.

Cependant, on convient aujourd'hui que si la phrase représente l'unité d'analyse du domaine linguistique syntaxique, elle n'existe pas dans l'utilisation réelle que l'on fait du langage. Les exemples construits de linguistique, qui sont encore trop souvent des phrases et qui représentent de loin la majorité des exemples du genre de l'article (Poudat, 2003, et v. chapitre 5), ne renvoient ainsi à aucune réalité effective, ce qui laisse songeur quant à la pertinence de la description.

Diverses approches ont toutefois dépassé le cadre de la phrase : pour ne citer qu'elles, la perspective énonciativiste, introduite par Benveniste, s'est intéressé à l'énoncé et ses enchaînements par rapport à la situation d'énonciation, et la théorie pragmatique a démontré, avec Austin et la théorie des actes de langage, que chaque énoncé était étroitement lié à un ensemble de stratégies communicatives.

Même si la langue n'est plus perçue comme une succession de phrases structurellement uniformes et si on commence à prendre conscience que la parole est régie par un ensemble de stratégies communicatives qui font parties de notre comportement social, ces courants ne se prêtent pas à l'étude des genres, dans la mesure où ces derniers ne sauraient être abordés hors d'un cadre textuel : les théories de l'énonciation sont encore éloignées du texte, et se concentrent davantage sur des corpus phrastiques de taille réduite, tandis que la pragmatique a cherché à proposer un modèle communicationnel global à partir de l'étude de quelques genres oraux, et plus spécifiquement de la conversation, au détriment des genres écrits et des textes.

Genres et textes sont en effet intimement liés : si le genre affecte les différents niveaux d'analyse de la linguistique, l'étude de ses motifs linguistiques ne saurait se réaliser dans le cadre restreint du mot ou de la phrase. Le texte représente un palier de description adéquat, dans la mesure où c'est un objet attesté et produit dans une pratique sociale déterminée⁵, et non une unité de description artefactuelle, élaborée en vue d'une analyse particulière. Nous admettons donc, à la suite de François Rastier, que « les textes sont l'objet empirique de la linguistique » (Rastier, 1996).

Les approches textuelles sont pourtant relativement récentes : la *linguistique textuelle*, dont les objectifs descriptifs et méthodologiques sont décrits dans (Adam, 1990), a moins d'une cinquantaine d'années, et les recherches sur le texte et la textualité se sont longtemps heurtées – et se heurtent encore – aux questions suivantes :

Three such problems which are connected with the basic assumption of text-discourse theory concern the criteria we might use to define (a) the outer boundaries or conditions of a text, (b) the way the text 'hangs together' internally, and (c) the structural arrangement of the elements making up the text. (Hartmann, 1980)

Ces éléments relèvent précisément des normes instituées par le genre : si les textes possèdent des propriétés distinctes qui ne sauraient être décrites avec le même jeu de descripteurs, ils relèvent d'un genre particulier qui les détermine, et qui leur impose des contraintes, tant structurelles que cohésives ou stylistiques⁶. En ce sens, l'étude des textes ne saurait faire l'économie de celle des genres.

⁵ Par *texte*, nous entendons la définition proposée par Rastier, qui présente l'avantage d'intégrer la notion de *support* et d'intégrer les textes oraux :

Un texte est une suite linguistique empirique attestée, produite dans une pratique sociale déterminée, et fixée sur un support quelconque. (1996, 19)

⁶ Par exemple, l'article scientifique est un genre très structuré, qui contient plusieurs niveaux d'organisation numérotés, qui rendent compte d'une progression démonstrative, contrairement au genre du roman, divisé en chapitres articulés autour d'une progression narrative.

Pourtant, les textes ont souvent été abordés dans des cadres discursifs sans prise en compte de la notion de genre. Ainsi, l'*analyse de discours* issue des approches ethnographiques de l'anthropologie américaine et de la sociologie britannique⁷, s'est intéressée au problème du texte comme lieu d'intersection des phénomènes linguistiques et non linguistiques, *i.e.* sociologiques. On cherche à relier les propriétés linguistiques du texte à ses conditions sociales de production et d'interprétation, ce qui donne lieu à des typologies discursives pour le moins discutables, fondées sur des critères sociaux qui correspondent globalement aux vecteurs idéologiques et sociopolitiques. Cependant, comme le souligne Rastier :

[La] notion de *conditions de production* a été transposée du matérialisme historique et utilisée par l'Analyse du discours pour instaurer une transcendance du sens à l'égard du texte, car ces conditions sont considérées de fait comme des déterminations causales. (2001 : 245)

Ce qui explique pourquoi la notion de genre est souvent évacuée de tels modèles⁸ :

On comprend alors que pour l'Analyse du discours l'étude des genres textuels semble au mieux secondaire, car relevant de la dimension de l'énoncé ; au pire gênante, car elle suppose un rapport au social par des normes intralinguistiques dont relèvent les genres, et non par les « conditions de production », dans lesquelles résident les causes sociales externes auxquelles on réduit les textes. Comme la conception traditionnelle de la science entend remonter des effets aux causes, l'Analyse du discours entend remonter du texte aux déterminations sociales et idéologiques qui l'ont produit. (*Ibid.*)

Or, si les textes nécessitent un cadre générique pour être abordés, le discours pourrait difficilement se passer des genres, qui constituent un palier linguistique adéquat pour les relier aux textes.

Nous n'énumérerons pas toutes les acceptions de la notion de discours⁹, et admettrons à la suite de Rastier, que les discours permettent de relier les pratiques sociales aux genres, qui eux-mêmes assurent le lien entre textes et discours. Cette conception présente en effet l'intérêt de rendre les genres à l'étude linguistique tout en reconnaissant leur détermination sociale. *A fortiori* :

Puisque tout texte se rattache à la langue par un discours et à un discours par la médiation d'un genre (...), l'étude des genres doit devenir une tâche prioritaire pour la linguistique. (2001 : 230).

En effet, la linguistique, qui est de plus en plus confrontée au problème de l'attestation de ses corpus d'étude, et à la pertinence de ses descriptions, ne peut ignorer plus longtemps la question des genres et doit se consacrer à leur description : les genres sont en effet des lieux de régulation et de contraintes, des *espaces normatifs* qui permettent, avec les discours et les champs génériques, de réconcilier la linguistique de la langue avec celle de la parole :

Les deux linguistiques, celle de la langue et celle de la parole, que Saussure cherchait explicitement à articuler, restent unies par l'espace des normes. Les niveaux des genres, champs génériques et discours sont bien les niveaux stratégiques qui permettent de passer de la généralité de la langue aux particularités des textes, car les relations sémantiques entre textes s'établissent

⁷ Qu'il faut distinguer de l'école française d'analyse du discours, qui s'est essentiellement intéressée à l'étude du discours politique.

⁸ Notons que l'analyse du discours telle que la décrit Maingueneau intègre toutefois les genres, mais pour les relier à des lieux sociaux (Maingueneau 1995).

⁹ D'autant que l'étude du discours est un domaine d'analyse passablement mal défini, qui se positionne difficilement dans le champ linguistique, dans la mesure où son objet est supposé dépasser et intégrer les branches de la discipline, tout en nécessitant des compétences autres ; la notion de discours renvoie en outre à une extrême diversité d'acceptions ; si elle peut désigner la conversation en analyse conversationnelle, ou l'instance d'énonciation par opposition au récit selon Benveniste, les travaux de l'école française tendent à lui conférer une valeur sociosémantique, tandis qu'elle a longtemps été amalgamée à la parole saussurienne.

préférentiellement entre textes du même genre, du même champ générique et du même discours.
(...)

Bref, la linguistique peut prendre *de droit* pour objet de description l'espace des normes : au lieu de les édicter, comme elle le faisait naguère en frappant d'inacceptabilité des énoncés, alors même qu'ils sont attestés, elle doit les décrire et pour cela exploiter des corpus. (Rastier, 2002)

La tâche est fondamentalement complexe, dans la mesure où le genre est un palier déterminant pour le texte, mais déterminé par son discours et sa pratique sociale de rattachement. En outre, c'est un objet linguistique qui possède ses caractéristiques propres, et qui se prête par conséquent plus volontiers à certains types d'analyse. Il serait d'ailleurs plus approprié de dire que certaines théories, en privilégiant l'analyse de certains genres, ont finalement développé des modèles d'analyse générique – d'ailleurs particulièrement efficaces : il en va ainsi de la pragmatique anglo-saxonne, qui fonde la plupart de ses méthodes d'analyse sur le genre de la conversation.

Chaque énoncé étant d'abord individuel, il convient enfin de définir la question des *styles* (*personnels* ou *d'auteur*) par rapport à celle des genres. De nombreux travaux de stylistique, portant d'ailleurs essentiellement sur les discours littéraire et politique, ont en effet cherché à décrire le caractère individuel des énoncés par rapport à l'artefact de la « langue ordinaire ». Pourtant, si l'existence de « styles personnels » est indiscutable, il est également manifeste qu'ils varient selon les genres et les discours :

Any utterance – oral or written, primary or secondary, and in any sphere of communication – is individual and therefore can reflect the individuality of the speaker (or writer); that is, it possesses individual style. But not all genres are equally conducive to reflecting the individuality of the speaker in the language of the utterance, that is, to an individual style. The most conducive genres are those of artistic literature: here the individual style enters directly into the very task of the utterance, and this is one of its main goals (but even within artistic literature various genres offer different possibilities for expressing individuality in language and various aspects of individuality). (Bakhtine, p. 63)

En effet, les études de styles personnels portent plus volontiers sur les textes littéraires ou politiques que sur les décrets ou les notices de montage. L'étude des styles ne saurait faire l'économie d'un examen des genres :

Style enters as one element into the generic unity of the utterance. Of course, this does not mean that language style cannot be the subject of its own independent study. Such a study, that is, of language stylistics as an independent discipline, is both feasible and necessary. But this study will be correct and productive only if based on a constant awareness of the generic nature of language styles, and on a preliminary study of the subcategories of speech genres. (*Ibid.*)

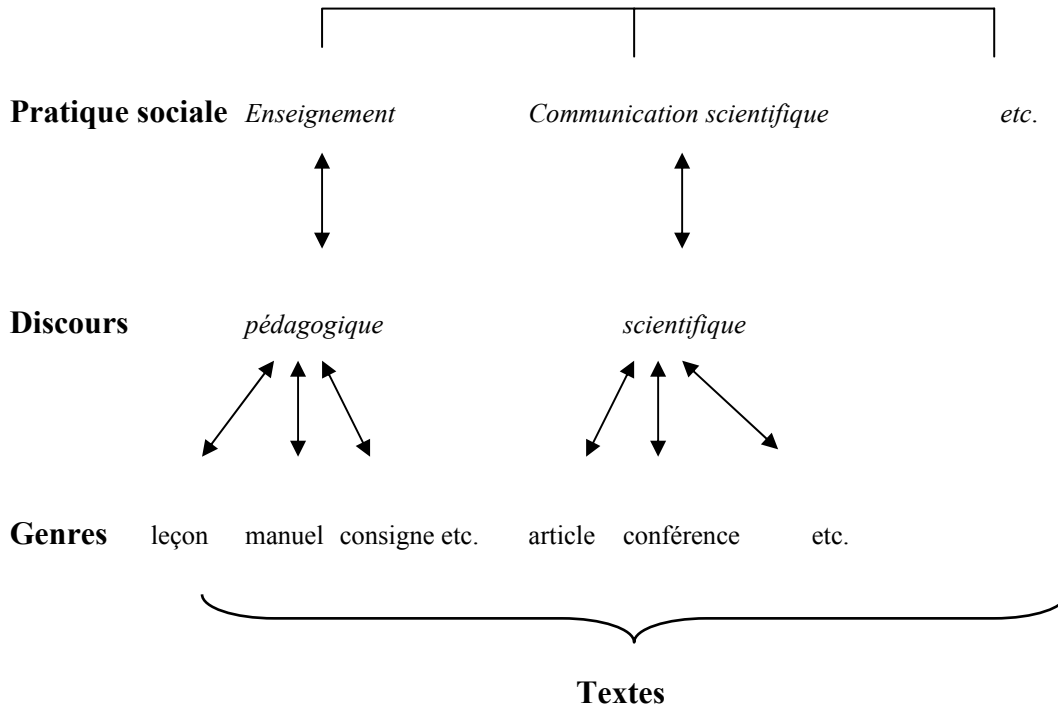
De fait, et comme le souligne Rastier (2001), les régularités génériques régissent les régularités stylistiques : les styles ne représentent que des « normes additionnelles », voire optionnelles, qui s'appuient sur les normes génériques, que ce soit pour s'en écarter ou y adhérer. Le style personnel est ainsi un *usage singulier du genre*, qui se manifeste plus ou moins selon les genres : l'étude des genres littéraires, et de l'ensemble des genres autorisant, voire préconisant l'usage d'un style individuel, requièrent un examen des styles d'auteur constitutifs du corpus, dans la mesure où la présence de styles trop singuliers pourrait entraver l'analyse du genre.

On voit bien que le genre est une notion centrale, un lieu de normes qui articule les linguistiques de la langue et de la parole et qui se doit d'être décrit par la linguistique. Il convient dès lors de s'interroger sur les modalités d'analyse des genres, et plus généralement, sur ce qui le définit.

1.1.3. Des genres et de leur description linguistique

1.1.3.1. Définition linguistique du genre

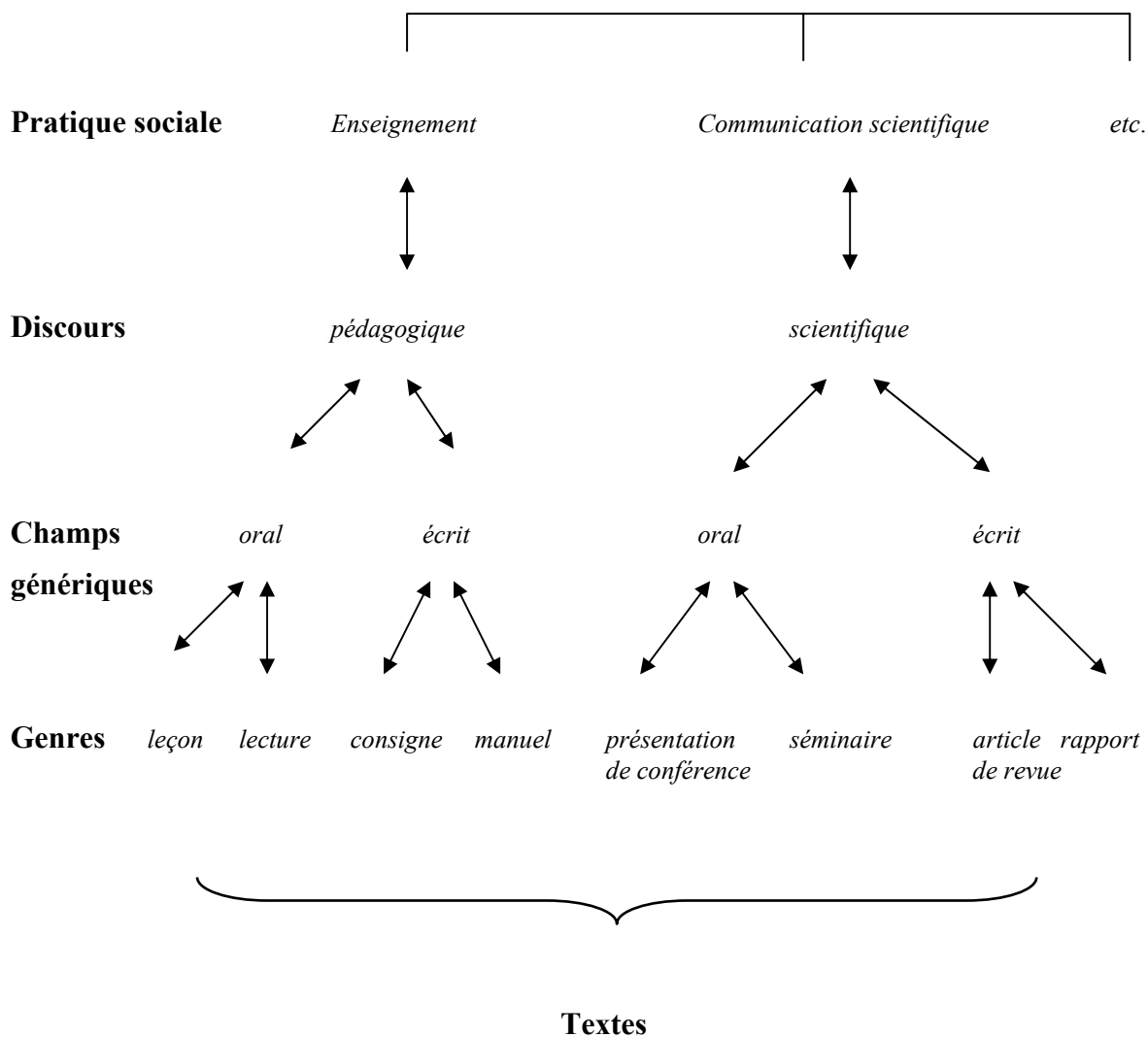
Si l'on considère que le genre est le palier permettant de relier un texte à son type de discours, qui lui-même assure la médiation entre pratique sociale et genre, on obtient la représentation non hiérarchique suivante :



On conviendra d'emblée et par principe qu'il n'est pas envisageable de lister les genres, ni les pratiques sociales auxquelles ils se rattachent, au sein d'une typologie exhaustive, d'autant qu'il y a parfois un écart entre le genre et sa désignation. Le fait que l'on hésite parfois entre plusieurs appellations pour qualifier un texte semble indiquer la présence d'une concurrence entre les genres : il est ainsi difficile de distinguer un article de recherche d'un article dans des actes en linguistique, d'autant que les deux genres sont publiés dans les mêmes revues et possèdent régulièrement des propriétés linguistiques identiques. François Rastier introduit donc une nouvelle médiation : *les champs génériques* :

Un champ générique est un 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). (2001)

Ce qui nous donne la représentation suivante :



Régulé par son type de discours, et par son champ générique, le genre institue des contraintes linguistiques sur le texte. En revanche, la nature et la rigidité de ces normes varient selon les genres, ce qui rend l'élaboration d'une définition générale du concept problématique et empêche l'adoption d'une grille d'analyse stable, convoquant un jeu de catégories linguistiques bien défini. C'est d'ailleurs pourquoi Bakhtine estimait qu'une analyse strictement linguistique des genres était compromise, en raison de la trop grande rigidité des unités linguistiques (ou grammaticales) par rapport au caractère évolutif et flexible des genres.

Il serait peu pertinent de privilégier un niveau d'analyse linguistique particulier pour caractériser un genre dans son intégrité : un genre n'est pas un objet syntaxique, ni morphologique, mais un lieu de contraintes, d'oppositions et de corrélations linguistiques. En ce sens, la phrase, ou le morphème n'ont aucun privilège à le définir, pas plus il nous semble que l'énoncé bakhtinien¹⁰ ou la séquence¹¹, telle que définie par Adam.

¹⁰ Le niveau de l'énoncé proposé par Bakhtine pour décrire les genres est en effet difficile à délimiter, les critères de « changement de sujet parlant » et de « finalisation » (le locuteur a dit tout ce qu'il avait à dire, et il n'est donc plus possible d'y répondre) nous semblant finalement peu applicables.

Plusieurs approches ont en effet proposé d'adopter une unité, ou plus largement, une modalité de description du texte : la notion de *fonction*, issue du Cercle de Prague, est ainsi fondamentale dans l'ensemble des approches systémiques du genre, dérivées des travaux de Halliday (1985/1994) et Halliday et Hasan (1989). On examine les énoncés en fonction de l'information qu'ils contiennent et on détermine la contribution sémantique de chacune de leurs parties par rapport à l'ensemble du texte. Les études de genres – essentiellement didactiques – menées dans cette perspective ont ainsi contrasté les genres *factuels* des genres *narratifs*, ce qui nous semble aussi discutable que la distinction bakhtinienne entre genres *premiers* et *seconds*¹².

Ainsi, l'étude d'un genre donné ne peut être menée sur la base de critères empruntés à d'autres entreprises typologiques, bien que l'on admette volontiers qu'il est plus aisé de s'attacher à l'une de ces contraintes en particulier, et d'élaborer une typologie textuelle en fonction du critère choisi. Pourtant, un genre n'est résolument pas un *type de texte*¹³, et étant donné qu'aucune typologie textuelle n'est encore parvenue à discriminer les genres, leur caractère multidimensionnel – certes abstrait – demeure encore définitoire de leur nature :

Les genres sont en effet définis par un *faisceau* de critères, et doivent d'ailleurs leur caractère d'objectivité à la multiplicité de ces critères. La cohésion du faisceau des critères, tant au plan du signifié qu'à celui du signifiant, structure la textualité et détermine la sémiologie textuelle. (...)

aucune typologie des textes fondée sur des critères définis indépendamment des genres (comme oral vs. écrit, public vs. privé, etc.) n'a permis d'isoler des genres (...). On objectera certes qu'un genre n'est qu'un type parmi d'autres, voire un "genre de type". La question sera tranchée quand on aura produit des faisceaux de critères stables et cohérents, indépendants de toute connaissance préalable des genres et des discours, qui ne recourent ni les genres ni les discours, et qui aient une pertinence théorique et pratique. Alors, — le défi est lancé —, la typologie des textes sera devenue indépendante de la théorie des genres. (Rastier, 2001 : 253, 254)

Les genres peuvent donc être appréhendés comme des objets linguistiques multidimensionnels, définis par des faisceaux de critères, qui permettent de relier les discours et les champs génériques aux textes. En ce sens, le genre est un lieu de régulation et de corrélation des phénomènes linguistiques, ce qu'admettent également les conceptions prototypiques du genre :

A prototypical theory of categorisation allows for the inclusion of such cases within the umbrella of the one single genre by basing deviations from the central prototypical core on how the world sees particular instances of a genre, rather than the much less flexible approach held in classical theories of categorisation where "entities are classified into sets with clear-cut boundaries and where an entity is seen as either belonging or not to a set. (Forbes 92: 378-9)" (Paltridge, 1997)

¹¹ En effet et comme le démontre Rastier, « la notion de séquence ne peut définir celle de genre mais la complète à un palier d'analyse inférieur ; mieux, le genre détermine les séquences et non l'inverse : c'est là un effet ordinaire du principe que le local détermine le global. Aussi, l'on ne peut tableer sur une compositionnalité générique : un genre n'est pas défini par ses séquences, au sens où il se réduirait simplement à un effet de leurs combinaisons. » (2001 : 265).

¹² Bakhtine distingue en effet les genres premiers, qui correspondraient aux interactions verbales spontanées, des genres seconds, ou complexes, qui renverraient aux genres culturellement plus « évolués », et donc écrits.

¹³ Adam (1999) propose d'ailleurs d'en finir avec les types de textes lorsqu'il énonce :

l'idée même de « types de textes » qui ne cesse de refaire surface, en particulier dans le champ de la didactique du français, doit être sérieusement interrogée. Le succès des typologies de textes a eu tendance à cantonner l'étude des genres dans le champ de la poétique littéraire. (1999 : 81)

La conception prototypique du genre présente l'intérêt de prendre en compte la variation des instances génériques ; malgré tout, le prototype d'un genre donné reste délicat à établir. Correspond-il à une moyenne des critères considérés ? Comment mesurer la notion d'*écart* ?

La question délicate de l'étude de ces « faisceaux d'indices », et de la mise en œuvre d'une critérisation adaptée à l'analyse des genres demeure de surcroît bien problématique.

1.1.3.2. Des corpus aux genres

Le développement du traitement automatique des langues, de la statistique et plus généralement de la linguistique de corpus a entraîné celui d'outils et de méthodes de traitement de grands corpus opérationnels qui permettent aujourd'hui de faire émerger des phénomènes et des régularités linguistiques qu'il n'était pas possible d'intercepter auparavant. Jusqu'aux années 60 – et encore aujourd'hui –, les intuitions des linguistes constituaient la seule source légitime d'information, tandis que l'intérêt pour les données authentiques et attestées était au plus bas.

De manière générale, un genre n'est saisissable qu'en corpus : si l'entreprise est descriptive, le corpus doit être génériquement homogène et représentatif du genre étudié. Etant donné leur constante évolutivité, on étudie généralement les genres en synchronie ; une étude diachronique nécessiterait l'établissement de plusieurs corpus, correspondant aux états $[1, n]$ du genre étudié¹⁴.

Les progrès de l'informatique et des possibilités de numérisation ont entraîné une – relative – disponibilité de corpus numériques¹⁵, ce qui est particulièrement favorable à l'analyse des genres. En effet, sans sombrer dans l'écueil du « Gros, c'est beau » (Habert, 2000), la caractérisation d'un genre requiert un corpus certes représentatif, mais de taille suffisamment importante pour autoriser et légitimer la description. On regroupe généralement les textes sous leur appellation générique d'origine, ce qui semble objectiver les études. Toutefois, et selon le genre observé, un corpus est fort susceptible d'être à la fois homogène en genre et extrêmement hétérogène, ce qui contraint souvent l'analyste à adopter des paramètres de constitution additionnels (auteurs, années de parution, domaine ou sujet, etc.).

Toute caractérisation étant différentielle, celle d'un genre donné s'effectue nécessairement par champs de contraste. La diversité des genres étant infinie et la linguistique de corpus encore peu développée, on manque globalement de grands corpus numérisés qui pourraient permettre de positionner un genre dans son discours ou dans son champ générique, afin d'en obtenir une description différentielle qu'il conviendrait ensuite d'explorer. Les grandes collections disponibles (corpus, bases de données)¹⁶ sont extrêmement hétérogènes : des discours et des genres pour le moins divers y sont représentés et les dates de parution des textes peuvent varier sur des empan de cinquante ans. La représentativité de ces ensembles est particulièrement discutable, dans la mesure où ils ne sont ni représentatifs de la *langue*, ni d'un discours ou d'un genre.

Dans cette perspective, la caractérisation d'un genre ne peut se fonder que sur son corpus. La démarche est donc nécessairement empirique, et ascendante : les variables de départ sont

¹⁴ A moins bien sûr de se concentrer sur un type d'évolution linguistique particulier, comme celle du lexique utilisé par exemple.

¹⁵ Ecrits bien sûr. Nous n'aborderons pas la question des genres oraux, dont le traitement demeure soumis à des méthodologies bien différentes. (cf. *Traitement automatique des corpus oraux*, TAL 45 : 2)

¹⁶ BNC, Brown corpus, Frantext, *Le Monde*, etc.

élaborées et sélectionnées en fonction des hypothèses de l'analyste, et ne sont donc pas directement caractéristiques du genre. Ce n'est que l'observation du corpus qui permettra de déterminer si ces éléments sont bien spécifiques au genre étudié ; par contre, ils n'auront qu'un statut d'indice par rapport à la caractérisation de leur discours d'affiliation. En effet, le discours, et le(s) champ(s) générique(s) de rattachement du genre ne sont jamais immédiatement saisissables : ce n'est qu'en comparant son corpus d'analyse à d'autres corpus génériques relevant du même discours, que l'analyste sera en mesure de caractériser les champs génériques, puis le discours auquel se rattache le genre étudié.

Il en va pour ainsi dire de même des hypothèses de variation qu'il est possible de formuler relativement à un genre donné : si les hypothèses de variation sociologiques peuvent être éprouvées par d'autres biais¹⁷, les hypothèses linguistiques sont encore une fois soumises au corpus, voire à la constitution d'un corpus distinct, élaboré dans cette perspective.

A. Qualitatif et quantitatif

On distingue fréquemment deux modalités d'analyse des genres et des corpus en général : les études dites « qualitatives » de celles dites « quantitatives ». La voie qualitative renvoie à l'étude d'un marqueur ou d'un phénomène linguistique en contexte¹⁸, tandis que la voie quantitative renvoie à l'utilisation de méthodes statistiques pour décrire un objet. La notion d'*analyse qualitative* est globalement apparue en réaction au développement des statistiques, qui introduisent en linguistique de corpus, et dans l'ensemble des sciences de l'homme et de la société, les exigences et les méthodes des sciences de la nature¹⁹.

Les méthodes quantitatives doivent bien entendu être manipulées avec précaution : les études statistiques sont parfois tentées de présenter leurs résultats chiffrés comme objectifs, et donc exempts d'interprétation, sans légitimer les mesures adoptées en amont par rapport à la définition de l'objet observé. Or, les données recueillies résultent nécessairement d'un travail de construction élaboré en fonction des hypothèses de l'analyste ; en ce sens, elles ne sauraient être immédiatement tenues pour *réelles* ou *objectives*. Comme le souligne Desrosières (1993/2000), ce n'est que par un travail d'*objectivation* qu'il est possible de sortir du débat entre objectivistes et relativistes :

La question de la consistance et de l'objectivité des mesures statistiques est souvent posée. La perspective proposée vise à sortir des dilemmes récurrents auxquels se heurte le constructeur du chiffre, s'il veut y répondre complètement. D'une part, il précisera que la mesure *dépend de conventions* portant sur la définition de l'objet et les procédures de codage. Mais, d'autre part, il ajoutera que sa mesure *reflète une réalité*. Le paradoxe est que ces deux affirmations sont incompatibles, et qu'il est pourtant impossible de répondre autrement. En remplaçant la question de *l'objectivité* par celle de *l'objectivation*, on se donne une possibilité de voir autrement cette

¹⁷ Il est par exemple possible d'élaborer un questionnaire destiné à évaluer les relations existant entre les représentations que les acteurs de la communauté de discours se font du genre étudié, et les observations linguistiques obtenues ; ou encore d'apprécier l'impact de facteurs sociologiques (e.g. le statut, l'âge ou le sexe) sur le genre analysé.

¹⁸ Ce dernier s'échelonnant généralement de la phrase, ou de l'énoncé, au paragraphe.

¹⁹ Qui sont d'ailleurs souvent mal perçues en linguistique : il nous semble d'ailleurs important de souligner que les linguistes français, contrairement aux psychologues ou aux sociologues, ne sont généralement pas formés aux statistiques : aussi les études linguistiques françaises ne recourent-elles généralement qu'aux statistiques descriptives (moyennes, écart-types, etc.), contrairement aux travaux anglo-saxons (e.g. le courant *Computers and the Humanities*), qui font régulièrement appel aux méthodes statistiques multidimensionnelles (analyses en composantes principales et méthodes factorielles en général, classifications, etc.) et explicatives (régression linéaire et multiple, etc.) pour décrire et contraster les textes, les genres ou les styles.

contradiction. La réalité apparaît comme le produit d'une série d'opérations matérielles d'inscriptions, produit d'autant plus réel que ces inscriptions sont plus générales, c'est-à-dire que les conventions d'équivalence qui les fondent sont plus solidement établies, à partir d'investissements de plus grande ampleur. Or ces investissements ne trouvent sens que dans une logique d'action englobant la logique apparemment cognitive de la mesure. Si une chose mesurée est vue comme relative à une telle logique, elle est tout à la fois réelle, puisque cette action peut prendre appui sur elle (ce qui est un bon critère de réalité), et construite, dans le cadre de cette logique. (Desrosières, p. 9)

L'étude des genres doit ainsi être objectivée, de telle sorte que le quantitatif puisse participer à la qualification des genres. En ce sens, il nous semble qu'il n'y a pas lieu d'opposer qualitatif et quantitatif en ces termes : les deux approches n'ont pas les mêmes objectifs « qualitatifs », ou descriptifs, et sont finalement très complémentaires : si l'analyse en contexte est pertinente pour mettre au jour un dispositif énonciatif ou le fonctionnement d'un marqueur en genre, la seconde permet d'aborder ce qu'il n'est justement pas envisageable d'étudier de la sorte, et contribue à la détermination des hypothèses à valider en contexte. En effet, c'est l'étude d'un genre qui doit pouvoir déterminer la pertinence des éléments de description adoptés, et ses caractéristiques. Etant donné que les méthodes quantitatives permettent de prendre en compte de nombreux critères de niveaux différents et d'éprouver leurs relations afin de faire émerger des phénomènes non, voire contre-intuitifs, il nous semble qu'elles doivent constituer un préalable à toute étude générique.

En rassemblant des éléments a priori singuliers, le travail statistique permet ainsi de donner réalité et consistance à des objets plus complexes et plus vastes, comme le sont les genres du discours.

B. Genres et méthodes inductives

Si les statistiques multivariées sont aujourd'hui très présentes dans les entreprises typologiques, il faut d'abord mentionner le travail de D. Biber (1988), qui a connu de larges répercussions. Dans un ouvrage-phare intitulé *Variation across Speech and Writing* (1988), Biber propose une méthodologie *ascendante* et *inductive* permettant de faire émerger des typologies textuelles à partir de critères linguistiques : un travail de comptage de 67 traits linguistiques, essentiellement morphosyntaxiques, est d'abord effectué sur les textes normalisés²⁰ de deux grands corpus de langue anglaise (le *Lancaster-Oslo-Bergen Corpus of British English*²¹ et le *London-Lund Corpus of Spoken English*²²), comptant au total 23 genres différents (17 écrits et 6 oraux). Le corpus est ensuite soumis à différentes techniques de statistique multivariée (et principalement à une analyse factorielle²³), permettant l'obtention de groupes de traits en cooccurrence (les facteurs), interprétés ensuite en termes de dimensions fonctionnelles.

Biber obtient finalement sept facteurs, et dégage six dimensions textuelles, à partir desquelles est proposée une description des types de textes du corpus. La prose académique serait ainsi hautement informative, non-narrative, abstraite et explicite, et moyennement persuasive et spontanée.

²⁰ Les textes étant de longueur différente, Biber choisit de les normaliser en séquences de 1 000 mots, correspondant au début des textes.

²¹ Connu sous le nom de LOB Corpus, comprend 61.500 textes de 2 000 mots, issus de 15 genres différents.

²² 87 textes d'anglais parlé de 5 000 mots.

²³ C'est-à-dire la réduction des variables à un petit ensemble de variables dérivées, les facteurs. Chaque facteur représente un groupement de traits linguistiques qui co-occurrent avec une fréquence élevée

Si les méthodes factorielles sont bien antérieures à l'étude de Biber, elles avaient surtout été appliquées à la psychologie (tests d'intelligence ou de mémoire, détermination de facteurs cachés, etc.). Lebart et Morineau (1982) rappellent en effet que l'*analyse factorielle classique*, fondée et perfectionnée par Spearman (1904) et Thurstone (1947), se proposait déjà d'aller au-delà des apparences en faisant émerger des variables échappant à l'observation directe. Quant à l'Analyse en Composantes Principales (ACP) qu'utilise Biber, elle est issue de la théorie antiréaliste de Pearson (1901) et de ses travaux sur la notion de corrélation, destinée à supplanter la notion de causalité, supposée insuffisante et infondée.

L'étude de Biber présente ainsi l'originalité d'appliquer des méthodes de statistique multivariée aux données linguistiques autres que lexicales²⁴, ce qui n'a été rendu possible qu'avec le développement de l'informatique – et plus précisément des corpus numérisés, des outils d'étiquetage automatique et des possibilités d'implémentation des méthodes statistiques.

Si Biber ne s'est pas intéressé à la question des genres sur le plan linguistique, les genres étant considérés comme des catégories sociales soumises à l'intuition du locuteur, la nature inductive de sa démarche, et les méthodes factorielles en général, ont particulièrement intéressé les entreprises typologiques et descriptives. Les outils de statistique multidimensionnelle ont ainsi montré leur efficacité en matière de comparaison de corpus (Rayson et Garside, 2000), ou pour faire émerger des typologies textuelles en fonction de critères spécifiques, comme par exemple le degré de technicité des textes (Copeck et al., 2000). Dans la lignée directe de Biber, diverses entreprises de profilage par induction (Karlgrén 1994, Karlgrén et Cutting 1994, Kessler et al. 1997, Habert 2000, etc.) ont cherché à dégager des typologies textuelles à partir d'un « bilan quantitatif fondé sur des indices linguistiques », les *types de textes* obtenus étant généralement considérés comme des « agglomérats fonctionnellement cohérents de traits linguistiques » (Habert, 1999).

Bien qu'ils n'accordent qu'une importance mineure aux genres, ces travaux illustrent les avancées encourageantes de la linguistique de corpus, et des méthodologies applicables à l'étude d'un genre donné : puisqu'on ne dispose pas de paramètres stables pour caractériser un genre *a priori* et qu'il n'existe pas de *niveau* général d'analyse des genres, les méthodes inductives, et les méthodes statistiques en général, intéressent particulièrement la description des genres. Plusieurs travaux ont souligné que cette dernière était bien plus complexe qu'on l'avait pensé, et de nombreux analystes se sont accordés sur le fait qu'elle devait être plus probabiliste que déterministe (Butler 1988, Nesbitt and Plum 1988, Couture, 1988, Halliday 1991, cités dans Paltridge 1997). S'il y a en effet nécessité de dépasser les niveaux d'analyse traditionnels, nombre des problèmes rencontrés sont dus à l'application de méthodes non adaptées ; on tend effectivement à transposer un type de description par ailleurs efficace dans d'autres cadres à des études qui en nécessiteraient d'autres.

En ce sens, les méthodes typologiques inductives et de statistique multidimensionnelle en général nous semblent particulièrement cruciales pour faire émerger ces (faisceaux de) critères qui définissent les genres, à condition bien entendu de disposer d'un ensemble de paramètres pertinents quant au genre étudié – ou du moins en éliminant de la sélection les éléments résolument non caractéristiques du genre : il serait ainsi incongru d'espérer décrire le genre juridique du décret au moyen de marqueurs de l'oral. Une fois les variables

²⁴ Les travaux de statistiques textuelles sur des données lexicales remontant au début des années 60, avec Charles Muller et Jean-Paul Benzécri.

sélectionnées, l'analyse des corrélations entre les variables et l'examen des axes factoriels obtenus²⁵ nous semble permettre d'obtenir une première approximation descriptive du genre.

C. Des critères éligibles pour décrire les genres

La description des genres s'articule globalement autour de deux plans corrélés : les plans du contenu et de l'expression, auxquels peuvent être associés de nombreux marqueurs potentiels qui seront (in)validés par le corpus.

Les corpus bruts et les chaînes de caractère ayant largement démontré leurs limites, l'étude des genres est généralement menée sur corpus annotés, d'autant que de nombreux standards d'annotation se sont développés ces dernières années. Ainsi, la Text Encoding Initiative (TEI) a élaboré un système de codage des textes fondé sur SGML qui permet de baliser la structure des textes, de même que certains de leurs éléments caractéristiques²⁶. Ce type d'annotation intéresse particulièrement l'étude des genres textuels et leur structure, d'autant que l'adoption d'un tel standard rend les corpus immédiatement comparables, ce qui est propice à la mise en contraste et à la description des genres et des discours.

De surcroît, plusieurs systèmes d'annotation linguistique – essentiellement morphosyntaxique et syntaxique – ont été élaborés (e.g. Penn TreeBank, Susanne, Eagles, etc.), et ont pu donner lieu au développement d'outils d'étiquetage automatique, aujourd'hui communément utilisés par les entreprises typologiques. L'étude pionnière de D. Biber (1988) a ainsi encouragé de nombreux travaux à privilégier le niveau morphosyntaxique pour contraster et caractériser les textes (Karlgrén 1994, Karlgrén et Cutting 1994, Kessler et al. 1997, Rayson et Garside, 2000, Habert 2000, etc.).

Toute annotation étant le produit d'une interprétation, ces catégories, qui s'avèrent particulièrement discriminantes, ne sauraient pour autant être tenues pour *objectives*. Aussi doivent-elles être manipulées avec prudence : on oublie trop souvent que la TEI renferme une conception particulière des textes²⁷, de même que les catégories morphosyntaxiques ou syntaxiques ne sont finalement que le reflet d'hypothèses linguistiques ou grammaticales, reproche qu'il est légitime de faire à Biber avec Branca-Rosoff (1999, p. 15) :

Le choix des paramètres est cependant posé comme naturel ; c'est-à-dire qu'il n'est pas explicité par l'analyste. Or, les traits retenus n'ont rien « d'objectif ». Leur sélection résulte des hypothèses du chercheur et de la tradition, plus précisément des bases typologiques qu'il juge intéressantes. De ce point de vue, il n'y a pas d'analyse purement immanente et on perçoit par exemple chez D. Biber l'influence des travaux sur la syntaxe de l'écrit ou des analyses énonciatives, sans que jamais ces sources soient mentionnées et discutées.

²⁵ Qui correspondent toutefois rarement à de véritables « dimensions textuelles », d'une part parce que les analyses factorielles sur corpus homogènes n'obtiennent généralement pas des seuils factoriels aussi élevés que dans l'étude de Biber, consacrée à la caractérisation de corpus très hétérogènes, et d'autre part parce que les axes sont souvent peu interprétables – ou facilement sur-interprétables.

²⁶ www.tei-c.org

²⁷ Et recommande d'ailleurs des balises parfois problématiques, qui correspondent à des hypothèses linguistiques spécifiques. La distinction faite entre <q> et <quote> est par exemple souvent difficile à trancher en contexte :

<q> contains a quotation or apparent quotation — a representation of speech or thought marked as being quoted from someone else (whether in fact quoted or not); in narrative, the words are usually those of a character or speaker; in dictionaries, q may be used to mark real or contrived examples of usage.

<quote> contains a phrase or passage attributed by the narrator or author to some agency external to the text. (TEI P4, web)

Ces réserves exprimées, une utilisation raisonnable des étiqueteurs morphosyntaxiques peut s'appliquer sinon à la définition immédiate des genres, du moins à leur exploration : les sorties ne peuvent être considérées que comme un moyen d'analyse, et non comme une finalité descriptive. Comme le soulignent Malrieu et Rastier, qui obtiennent des résultats très encourageants en exploitant les méthodologies de statistique multidimensionnelle pour retrouver – et valider – une typologie de 2 500 textes classés par genre à l'aide de 251 variables morphosyntaxiques (2001) :

(...) nous ne privilégions pas exclusivement une démarche « déductive » : la linguistique de corpus qui se développe actuellement doit permettre de refonder ou d'affiner les distinctions intuitives et empiriques entre genres. En effet, les variations morphosyntaxiques selon les genres sont notables. Par exemple, les textes littéraires contiennent significativement moins de passifs que les autres ; la position de l'adjectif, la nature des déterminants, des pronoms et des temps, l'usage du nombre varient aussi notablement. Ou encore, dans le domaine technique même, les variations sont importantes entre un manuel et une brochure commerciale : au premier, les acronymes, les impératifs, les ellipses de déterminants ; au second, les phrases longues, les pronoms nombreux, etc. (...)

Si l'approche empirique « inductive » peut aider à mettre en évidence des différences entre textes, nous n'attendons pas qu'elle *définisse* les genres textuels. Nous considérons en effet que les genres sont définis par l'interaction normée de composantes textuelles : nous jouons alors de difficulté, puisque les traits linguistiques dont nous disposons pour conduire l'étude présentée ici sont issus d'une analyse morphosyntaxique au palier de la phrase qui ne tient évidemment pas compte des contraintes des paliers supérieurs. (p. 5)

Le niveau morphosyntaxique permet ainsi de distinguer les genres et les discours entre eux, et d'un point de vue plus descriptif, de mettre l'accent sur certaines corrélations positives ou négatives dont on ne ferait pas naturellement l'hypothèse. Une telle méthodologie nous semble représenter un préalable crucial à l'étude d'un genre donné, tant pour appréhender le plan du contenu que celui de l'expression, d'une part parce que les systèmes d'étiquetage mis au point englobent des descripteurs intéressants des deux niveaux, et d'autre part parce que l'étude d'un genre est sur le plan pratique, difficilement envisageable sur chaînes de caractères.

Le genre assurant le lien entre les plans du contenu et de l'expression, les corrélations entre les deux niveaux sont souvent les plus intéressantes : comme le souligne Rastier (2002), les ponctèmes, qui relèvent du plan de l'expression, sont par exemple très souvent corrélés à certaines entités thématiques :

Alors que la ponctuation n'est pas considérée comme sémantique et qu'elle est tout simplement absente des grammaires formelles, l'étude en corpus permet de souligner les corrélations entre contenus lexicaux et ponctèmes. Par exemple, dans un corpus romanesque, Evelyne Bourion (2001) a ainsi pu confirmer la corrélation entre des noms de sentiments et les ponctuations dans les contextes où ces noms apparaissent. Ainsi les sentiments ponctuels, brusques, comme la colère ou la joie, sont significativement associés aux points de suspension.

Enfin, parmi les critères éligibles pour décrire un genre donné, restent les variables structurelles : les sections et autres configurations peuvent ainsi être balisées selon la TEI. Notons que l'étude des parties constitutives d'un genre, et de ses modalités de structuration, s'effectuera différemment en fonction des genres étudiés : s'il serait hasardeux de comparer les premiers chapitres des romans, les introductions des articles de recherche font sens en tant que *sections* constitutives du genre de l'article, ce qui les rend comparables. En ce sens, leur description est pertinente, et participe à celle du genre.

Si l'on admet que les genres ne sont saisissables qu'en corpus et qu'ils représentent un palier de description linguistique adéquat pour décrire les discours et la langue, il convient de poser la question *des langues*, qui demeure sinon l'enjeu, l'un des enjeux descriptifs majeurs

de la linguistique. Il nous faut donc nous interroger sur les modalités, et la potentialité d'une étude contrastive interlangue des genres, terrain encore largement inexploré.

1.1.4. Genres et langues

1.1.4.1. De la comparabilité des genres

Etant donné qu'il est plus aisé d'analyser un marqueur ou un concept spécifique, et de le contraster d'une langue à l'autre, l'analyse contrastive des langues se construit usuellement selon le point de vue auquel le linguiste se place : le point de vue onomasiologique, dans lequel on part des concepts pour arriver aux signes linguistiques, et le point de vue sémasiologique, dans lequel on part du signe linguistique pour atteindre les concepts. Une analyse *en genre*²⁸ peut être menée selon ces deux modalités : les réalisations d'un même objet peuvent ainsi être contrastées dans les langues étudiées, et les différences ou correspondances obtenues permettent, selon la perspective et les objectifs descriptifs de l'analyse, d'éclairer le fonctionnement des langues ou celui des institutions.

Il en va tout à fait différemment de la mise en contraste des genres, qui sont des objets autrement plus complexes : étant donné qu'on se situe au-delà des pratiques instituées au sein des communautés langagières, on ne saurait faire l'économie d'un examen des pratiques à l'œuvre dans les deux communautés, ne serait-ce que pour déterminer leur degré de *comparabilité*. En effet, si les réalités sociale et linguistique sont plus à même de coïncider dans une même communauté de discours ou dans un même pays, les pratiques diffèrent d'une communauté à l'autre : par conséquent, les genres, voire leurs pratiques de rattachement, ne renvoient pas nécessairement à la même réalité, même s'ils portent la même appellation.

Une étude contrastive des genres nécessite ainsi un examen attentif des champs de pratique génériques et discursifs, qui permettra d'une part de déterminer le statut de la langue dans le champ et d'apprécier la comparabilité des genres sur le plan social, et d'autre part, d'élaborer un ensemble d'hypothèses linguistiques quant aux genres observés.

1.1.4.2. Genres et corpus parallèles

S'il est nécessaire de mener une réflexion sur la comparabilité des genres étudiés, il n'est pas moins indispensable de poser le problème de la comparabilité des textes et des corpus : de manière générale, les contrastes interlinguistiques sont manifestes dans des textes de langues différentes, et ils le sont davantage dans la comparaison des traductions avec leurs originaux. Etant donné qu'on traduit d'abord *en genre*, et que la traduction d'un texte d'une langue à l'autre nécessite une connaissance des deux genres et de leurs conventions, il est d'abord naturel de s'orienter vers ce type de corpus parallèle. Surgissent alors plusieurs types de problèmes : force est d'abord de constater que la grande majorité des textes n'est pas traduite, ce qui restreint l'étude contrastive des genres à l'examen de quelques genres essentiellement littéraires, journalistiques et juridiques. De surcroît, la plupart des textes traduits le sont en anglais, ou dans *certaines langues*, ce qui constitue une nouvelle limite. L'étude contrastive des genres devient donc directement soumise à la pratique de la traduction et aux « textes traduits », ce qui est délicat à légitimer. Il va en outre de soi que la traduction constitue elle-même une pratique particulière, et que les textes sont différemment rendus selon les

²⁸ Et on sait bien que les travaux sur exemples phrastiques sont loin d'être l'apanage de la linguistique du français : les analyses contrastives sont bien souvent menées sur les mêmes types de corpus, sans prise en compte des genres.

traducteurs²⁹, comme le montrent de nombreuses études contrastives, qui intègrent d'ailleurs ce paramètre stylistique à leurs analyses. Eu égard à ces contraintes, l'étude des genres sur corpus de textes parallèles *textes-traductions* demeure encore marginale – bien qu'à explorer, pour les perspectives intéressantes qu'elle ouvre à l'analyse des genres et plus largement, aux études traductologiques.

Si l'on écarte les traductions, Srpová distingue deux autres types de corpus parallèles, en relation avec la notion de genre :

- textes de genres comparables, sans égard pour les contenus concrets qu'ils traitent,
 - textes de genres comparables traitant d'un « même » sujet concret,
- (Srpová, 1994 : 195)

La contrainte du « même sujet concret » est toutefois peu convaincante, dans la mesure où l'on ne contrastera finalement pas les genres, mais les modalités de rendu du thème choisi selon les genres : on pourra par exemple comparer le traitement de certaines entrées thématiques comme « la guerre en Irak » par l'article de quotidien, ou encore le thème de l'eau ou de la rupture dans les romans de X ou de Y.

Le corpus parallèle de type 1 est donc plus envisageable : dans cette perspective, l'étude de genres comparables est envisagée comme la *mise en texte* de genres comparables, ce qui privilégie le plan de l'expression sur celui du contenu. Chaque langue peut donc être étudiée séparément, avec un jeu de catégories équivalent. Il est ainsi possible de décrire des textes de genres comparables à partir d'un jeu de catégories de mise en page, de fonctions communicatives ou de marqueurs linguistiques, bref, d'un ensemble de catégories portant sur la forme du texte, en faisant abstraction de son contenu spécifique.

Différentes études ont été élaborées dans cette perspective, et ont pu mettre au jour des différences culturelles ou institutionnelles entre des communautés de langue différentes, qui intéressent particulièrement la description des genres. Par exemple, Srpová (*ibid.*) souligne que les textes des genres journalistique et scientifique français recourent davantage à la polémique que les textes tchèques : le sujet français se montre davantage, et procède à un examen critique des références bibliographiques qu'il utilise, contrairement au sujet tchèque, dont la présence est moins visible, et qui se livre à une évaluation positive de ses sources.

1.1.4.3. Méthodes inductives et genres en contraste

La démarche nous semble toutefois plus délicate à appliquer sur grand corpus et avec des méthodes inductives : si l'on reconsidère le niveau morphosyntaxique, qui a démontré son efficacité discriminante dans de nombreuses langues, la notion d'*équivalence* paraît problématique, eu égard aux régimes d'organisation spécifiques des langues, qui ne recourent pas aux mêmes systèmes de catégorisation linguistique. Comme le souligne Moirand (1992, pp. 30-31) :

Mais si les marques, fonctions et schématisations interviennent, quelle que soit la langue, y a-t-il correspondance catégorie à catégorie entre ces trois sortes de données, qu'il est pourtant nécessaire d'articuler ? Comment comparer les corpus en langues différentes si l'on ne peut pas distinguer entre ce qui relève des préconstruits culturels et ce qui relève des constructions socio-cognitives (scripts, praxéogrammes) de la situation de communication et du genre du discours impliqué dans une communauté donnée ?

²⁹ En outre, on se heurte au problème du *calque* et des traductions plus ou moins littérales, susceptibles d'effacer certaines différences linguistiques et génériques.

L'élaboration d'un ensemble de descripteurs équivalents est ainsi une tâche complexe, a fortiori lorsqu'il s'agit non pas d'un phénomène ou d'un marqueur spécifique, mais d'un faisceau d'indices. Les corrélations émergeant au sein d'un même corpus, homogène en genre ou non, sont souvent d'ordre grammatical (e.g. les noms communs pluriel sont bien entendu corrélés avec les adjectifs et les déterminants pluriel) et par conséquent attendues ; toutefois, de nombreuses liaisons moins intuitives mais tout aussi caractéristiques du genre se dégagent également, et requièrent de fait une très bonne connaissance des systèmes langagiers observés afin de mettre en évidence des contrastes pertinents quant aux genres observés. Dans cette perspective, l'étude des deux genres devra s'effectuer de manière disjointe, et ce sont les corrélations et les axes d'organisation observés qui permettront de faire émerger des pivots de comparaison.

Il est bien entendu possible, voire recommandable, de procéder par sous-systèmes d'organisation linguistique : on peut par exemple se focaliser par exemple sur les systèmes des temps, des personnes, ou des modes de détermination verbale et nominale : la distribution des catégories dans chaque système, de même que leur présence/absence permettra d'éclairer et de contraster les caractéristiques des deux genres, et de leurs langues de rattachement.

1.2. Genres scientifiques et genre de l'article

Si les genres littéraires ont été largement explorés, l'étude des genres scientifiques/académiques et professionnels n'est menée que depuis une vingtaine d'années dans le cadre d'un courant plus appliqué – pédagogique à l'origine, combinant trois approches principales fondées sur les textes écrits : les travaux de Hasan, Halliday et Martin, cherchant à construire une vision systémique des genres, l'ESP (*English for Specific Purpose*), courant rhétorique qui s'intéresse essentiellement aux descriptions macro-textuelles et à la description des genres d'un point de vue phrastique ou propositionnel, et un ensemble de travaux portant sur la communication professionnelle écrite aux Etats-Unis (Bazerman 1988, Berkenkotter et Huckin 1995, Bizell 1992, Miller 1992 et Trosborg 2000).

L'étude des genres académiques et professionnels démarre ainsi au début des années 80, suite au mécontentement causé par l'échec des approches centrées sur la notion de *processus* dans les écoles primaires australiennes. L'étude des textes narratifs avait alors été privilégiée à celle des textes 'factuels' comme les rapports, les procédures, les explications ou les présentations, qui vont dès lors être abordés en tant qu'événements communicatifs et culturels, et analysés de manière systémique fonctionnelle. On s'intéresse ainsi à l'article scientifique, la lettre, l'accusé de réception, etc. en insistant sur leur structure schématique (début, milieu et fin de texte) et sur leurs variations potentielles.

Les genres scientifiques, et notamment le genre de l'article, sont généralement abordés dans un cadre *English for Specific Purpose*, et plus précisément dans un sous-domaine dédié au discours académique³⁰.

1.2.1. Le courant English for Specific Purpose

Le courant anglo-saxon *English for Specific Purpose* (ESP) est un courant rhétorico-fonctionnel, qui tient ses origines de diverses approches, décrites dans (Swales, 1990) :

³⁰ L'ESP s'est effectivement considérablement spécialisé en différents sous-domaines revendiqués, comme l'EST (*English for Science and Technology*) ou encore l'EMP (*English for Medical Purposes*)

variety studies, situation-specific skill and strategy studies, notional-functional approaches to language program development, discourse analysis, sociolinguistics, composition studies and studies in the areas of cultural anthropology, and language and cognition.

L'ESP forme aujourd'hui une communauté internationale bien établie dans le monde scientifique anglo-saxon³¹. Les textes scientifiques constituent l'objet d'une sous-communauté sociologiquement bien délimitée, dans la mesure où elle rassemble essentiellement des enseignants d'anglais scientifique (*Academic English*), qui cherchent à décrire les différences sociologiques, culturelles et linguistiques observées dans les textes scientifiques. A l'instar des courants sociologiques dans leur ensemble, l'étude linguistique des genres est subordonnée à un ensemble d'hypothèses de variation d'ordre social : différences entre les cultures, les sexes, les domaines scientifiques, etc. On procède le plus souvent par champs de contrastes, en corrélant des jeux de descripteurs discursifs à des fonctionnements institutionnels et sociaux.

On distinguera ainsi deux types de variations communément abordées en ESP : les variations d'ordre sociologique (1) et les variations d'ordre culturel (2), généralement analysés en termes de mouvements rhétoriques et de marqueurs discursifs (3). On exposera enfin les limites posées par l'adoption d'une telle approche (4).

1.2.1.1. Variations sociales : Communauté de discours vs. communauté sociologiques

Le fonctionnement des institutions scientifiques est souvent placé au premier plan des analyses du discours scientifique : bien que ce dernier ait été longtemps exempt d'analyses rigoureuses, il faut en effet souligner qu'il représente le premier type de discours à propos duquel on s'est interrogé sur les communautés qu'il suppose. En outre, et comme le souligne Bouvier (1997) :

A la suite de Kuhn, des chercheurs comme Feyerabend ont développé des idées mettant au premier plan non plus la structure du milieu ou de la profession scientifique mais celle des procédures d'adhésion et de persuasion à une nouvelle théorie scientifique et, à travers elles, la construction de nouvelles communautés scientifiques.

Si la science est un facteur d'intégration très puissant dans les relations humaines et internationales, la production de discours savant ne serait de surcroît pas immédiatement liée à un désir de connaissances : elle « s'échangerait » contre de la notoriété à l'intérieur du milieu scientifique (Hagstrom, *The scientific community*, New York, Basic Books, 1965, cité par Maingueneau, 1987). Bourdieu propose une description similaire de la communauté scientifique. En termes bourdieusiens, la communauté scientifique est un champ, et le champ scientifique est un marché, le lieu d'une concurrence acharnée dont l'enjeu est « le monopole de *l'autorité scientifique* inséparablement définie comme capacité technique et comme pouvoir social », « la *compétence scientifique* [étant] entendue au sens de capacité de parler et d'agir légitimement (c'est-à-dire de manière autorisée et avec autorité) en matière de science, qui est socialement reconnue à un agent déterminé » (Bourdieu 1976).

Suite aux travaux de Bloor et Barnes, qui marquent une rupture avec le *programme fort* en sociologie de la connaissance :

A la manière de Durkheim dans ses pages les plus sociologiques, [on] considère que le contenu des sciences, y compris dans ses aspects les plus logiques, comme l'idée de nécessité, est une

³¹ L'anglais scientifique, ou académique, étant encore peu enseigné en France, l'ESP y demeure encore marginal et relativement peu connu.

projection du social (en l'occurrence de la *contrainte sociale*), idée totalement étrangère à Kuhn ou à Feyerabend. (Bouvier, 1997)

La science est ainsi complètement démythifiée, les notions d'*objectivité*, de *rationalité* et de *vérité* n'ayant plus qu'un statut de croyance sociale et d'hypothèses métaphysiques parmi tant d'autres³². On s'intéresse donc au champ scientifique, de la même manière que l'on s'intéresserait à tout autre champ social. Les relations sociales entre les membres de la communauté de discours scientifique, et les jeux de pouvoir à l'œuvre en son sein, de même que le statut de ses membres (*expert vs. novice*) offrent ainsi de nombreux axes d'analyse : il est ainsi possible de s'intéresser à l'expression des conflits scientifiques, à la polémique et aux relations de solidarité, de même qu'à l'influence du degré d'accréditation et du sexe de l'auteur, et de relier ces phénomènes sociaux à des phénomènes discursifs.

Les études ESP se démarquent des études sociologiques en considérant les communautés scientifiques comme des *communautés de discours* (Swales, 1990). Une communauté de discours est *socio-rhétorique*, et constituée par un groupe de personnes se rassemblant afin de poursuivre des objectifs supérieurs à ceux de socialisation et de solidarité³³. Fondée sur des valeurs et des objectifs communicatifs partagés, son fonctionnement serait donc déterminé par des considérations *fonctionnelles*. Dans cette perspective, les genres ont une importance primordiale, dans la mesure où ils représentent des instruments permettant à la communauté de poursuivre ses objectifs. L'article scientifique est ainsi conçu comme un instrument autorisant le partage des connaissances et l'expertise du discours.

Le courant dans son ensemble considère le genre comme fondateur de la communauté de discours. Swales le définit ainsi :

(...) a class of communicative events, the members of which share some set of communicative purposes. These purposes are recognised by the expert members of the parent discourse community, and thereby constitute the rationale for the genre. This rationale shapes the schematic structure of the discourse and influences and constrains choice of content and style. Communicative purpose is both a privileged criterion and one that operated to keep the scope of a genre as here conceived narrowly focused on comparable rhetorical action. In addition to purpose, exemplars of a genre exhibit various patterns of similarity in terms of structure, style, content and intended audience. If all high probability expectations are realised, the exemplar will be viewed as prototypical by the parent discourse community. The genre names inherited and produced by discourse communities and imported by others constitute valuable ethnographic communication, but typically need further validation (Swales, 1990: 58).

On insiste globalement sur la visée communicative du genre, ce qui entraîne parfois la segmentation des textes en différents (*sous-*)genres : le résumé (*abstract*), contenu dans le péri-texte de l'article est ainsi considéré comme un genre, et l'introduction comme un *sous-genre*. En règle générale, on cherche à associer les genres ou leurs sections à divers indices rhétoriques afin d'en proposer une description fonctionnelle : par exemple, Docherty et Smith (1999) soutiennent que les sections *discussions* des articles scientifiques, jugées trop « spéculatives »³⁴, ne seraient finalement qu'un moyen de promotion de l'auteur.

32 Ce que beaucoup de scientifiques reprochent d'ailleurs à la sociologie de la science, accusée d'avoir alimenté « ce relativisme généralisé selon lequel il y a autant de vérités, de valeurs et de points de vue que de groupes sociaux ou de communautés d'intérêts, de cultures diverses, de conceptions du monde. » (Busino, 1997 : 8). Latour, dont les travaux ont connu de larges retentissements, considère ainsi explicitement la science comme une croyance, et défend l'existence d'un *relativisme naturel* (Latour 1989, 1994).

33 Par opposition aux communautés de parole, qui seraient, elles, sociolinguistiques et qui reposent sur des besoins de socialisation et de solidarité ; ce qui est largement discutable, la constitution d'une communauté scientifique représentant une nouvelle forme de socialisation, créant/renforçant des liens de solidarité

34 Et même considérées comme « the weakest part of the paper ».

S'il est possible de procéder d'un genre à l'autre, on contraste également les *domaines scientifiques* (cf. *infra*) entre eux, ce qui est largement dû, nous semble-t-il, aux pratiques d'enseignement des chercheurs d'ESP, qui sont elles-mêmes spécialisées. Aussi la notion de communauté de discours est-elle soumise à la notion de *domaine*, ce qui offre un nombre de champs de contraste quasi-illimité, d'autant qu'à ces possibilités de comparaison s'ajoutent des contrastes d'ordre culturel, que la *mondialisation de la science* a rendu particulièrement visibles.

1.2.1.2. Variations culturelles et mondialisation de la science

Du XVII^e siècle à nos jours, la communauté scientifique a fonctionné sur un modèle transnational, quel qu'ait été le mode de communication. Les progrès réalisés dans la transmission des informations ont constamment été mis à profit pour renforcer la circulation des données et des résultats entre les chercheurs. Les canaux de circulation de l'information scientifique sont aujourd'hui sans commune mesure avec ceux des siècles précédents : Internet³⁵, revues (inter)nationales, (visio)conférences, etc. participent à un processus de *mondialisation* de la recherche scientifique. L'accroissement des coopérations internationales renforce la nécessité d'une langue commune. Tout au long de l'histoire, les scientifiques ont été amenés à utiliser plusieurs langues³⁶ ; après la seconde guerre mondiale, et jusqu'à nos jours, la situation hégémonique des Etats-Unis a consacré l'anglais comme langue scientifique universelle. La communauté de discours scientifique anglo-saxonne est aujourd'hui une communauté globalisante, les communautés française ou allemande représentant des communautés internationales locales. Dans ce contexte, il n'est pas surprenant que la langue et la culture anglo-saxonnes fassent souvent office de modèle qui tend à substituer à ses standards les normes régissant les autres traditions intellectuelles.

La mondialisation³⁷ de la science n'est toutefois pas un processus uniforme, et ne saurait se réduire à une absorption progressive du champ scientifique par la culture anglo-saxonne, d'une part parce que la *science* ne constitue pas un champ homogène, mais renvoie à un grand nombre de domaines scientifiques distincts dotés d'une culture spécifique, et d'autre part parce que les langues, reflets de traditions intellectuelles et culturelles, mettent en place des stratégies discursives qui leur sont propres.

Ainsi, bien que l'anglais tende à devenir la 'monoculture' du discours universitaire, certains pays et certaines communautés choisissent de conserver leurs standards rhétoriques, ce qui peut être sujet à confusion ; un article rédigé en anglais et soumis à un comité anglophone peut ainsi être considéré comme 'incompréhensible', en raison de sa non-conformité au style anglo-saxon.

L'efficacité conférée à l'auteur dans les textes anglo-saxons, de même que la nature empirique et pragmatique de la recherche, contraste en effet avec la conception européenne

³⁵ Faut-il rappeler que ARPANET, qui a mené à la création de l'Internet actuel, était à l'origine destiné aux chercheurs ?

³⁶ Selon la situation hégémonique des nations ou des langues (ce fut d'abord le latin, différentes langues européennes, et enfin, l'anglais, du moins dans l'aire occidentale).

³⁷ On parle en effet de *mondialisation* et non d'*internationalisation* de la science, phénomène qui n'a rien de nouveau, et qui renvoie aux échanges entre peuples ; la *mondialisation* de la science renvoie à une autre réalité, celle de l'extension de ces échanges et de ces relations à l'échelle du monde, conséquence des progrès toujours croissants de communication et de transport de la civilisation contemporaine. Un réseau scientifique complexe, fondé sur une pluralité d'acteurs et de partenaires est ainsi mis en place, entraînant des difficultés nouvelles qui ne se limitent pas aux Etats et aux économies.

idéalisée du discours scientifique, profondément ancrée dans la tradition rhétorique grecque, et plus précisément dans la tradition aristotélicienne : la science a longtemps été perçue comme le reflet de la vérité, et la langue, medium profane utilisé par un auteur faillible, comme un obstacle à l'expression de vérités scientifiques. Un langage simple et impersonnel, dépourvu de marques subjectives, a longtemps été recommandé par la tradition logique, ce qui explique le nombre impressionnant de formes impersonnelles constatées dans nombre d'articles scientifiques français, allemands ou tchèques. La recherche efficace de la vérité et l'image déshumanisée de la science est encore bien ancrée dans le discours européen, et offre ainsi de multiples champs de contraste avec le discours anglo-saxon. L'usage des pronoms personnels, et notamment celui du pronom personnel de première personne du singulier, plus présent dans les textes anglo-saxons, représente ainsi un objet d'étude très fréquent.

1.2.1.3. Etudes textuelles et mouvements rhétoriques

On distingue globalement deux types d'analyses linguistiques : les études de marqueurs, dans lesquelles on cherche à corréliser un phénomène rhétorico-linguistique à une fonction sociale, et les analyses de la structure des genres, fondée sur l'examen des *mouvements rhétoriques* (*moves*), *i.e.* des schémas d'organisation rhétorique. La structure universalisante Problème-Méthode-Résultats-Discussion ne résistant pas à l'épreuve des langues, la configuration des textes est généralement conçue comme une série de *mouvements rhétoriques* (Swales, 1990), chacun d'entre eux contenant au moins une étape (*step*). Ainsi, les introductions d'articles de recherche se découperaient en quatre *moves* (Swales, 1990).

1. Establishing field

- (a) Showing centrality
- (b) Stating current knowledge
- (c) Ascribing key characteristics

2. Summarizing previous research

- (a) Strong author-orientation
- (b) Weak author-orientation
- (c) Subject orientation

3. Preparing for present research

- (a) Indicating a gap
- (b) Question-raising
- (c) Extending a finding

4. Introducing present research

- (a) Giving the purpose
- (b) Describing present research

L'analyse des mouvements rhétoriques étant particulièrement délicate à mener sur grands corpus, voire sur textes entiers, on se limite généralement à certaines sections de l'article,

notamment les *introductions*, les *résultats* et les *résumés*³⁸. Si ces études portent essentiellement sur l'anglais, on recense plusieurs études de l'article scientifique en rhétorique contrastive : une étude des résumés américains et suédois (Melander, Swales et Fredrickson in Duszak, 1997) contraste ainsi des textes rédigés dans trois domaines différents (biologie, médecine et linguistique) ; les sciences de la nature étant soumises à des standards plus stricts, une plus grande homogénéité est constatée dans les articles de biologie et de médecine, les articles linguistiques manifestant un taux élevé de variation, bien qu'il ait été remarqué que les textes anglo-saxons étaient plus homogènes. Ce type d'analyse permet ainsi de mettre au jour des différences communautaires et de faire émerger certaines tendances culturelles : par exemple, les auteurs polonais n'annoncent pas leur objectif principal en premier, contrairement aux articles scientifiques anglo-saxons, dans lesquels de telles présentations sont typiques, voire obligées.

1.2.1.4. Des limites du courant ESP

Le courant ESP vise ainsi à faire émerger des contrastes culturels ou sociologiques en les corrélant au niveau linguistique, dans un but didactique et descriptif. Le genre étant conçu comme une entité à deux faces, l'une sociale et l'autre linguistique, les deux versants génériques doivent être articulés : la description linguistique des genres est donc fonctionnelle, et les critères sélectionnés sont directement corrélables à une ou à plusieurs fonctions sociales, le plus souvent rhétoriques.

Cette conception technique des genres scientifiques, symptomatique de l'utilitarisme et du pragmatisme anglo-saxon, s'oppose à la conception poétique européenne en restreignant le nombre de descripteurs potentiels des genres aux éléments les plus immédiatement efficaces et les plus propices à l'élaboration ultérieures de procédures rédactionnelles : on conseillera ainsi aux non natifs anglo-saxons de veiller aux structures divergentes décrites lorsqu'ils rédigent en anglais, et aux anglo-saxons de mettre en place une formation qui leur permettrait d'améliorer leur compréhension des papiers rédigés par des auteurs non natifs.

De surcroît, les marqueurs sont traités en tant que séquences isolées et non corrélables, dans la mesure où l'on postule qu'elles ne sont pas liées. L'ESP adopte en effet une perspective nominaliste (à l'instar de la philosophie analytique) en niant l'existence de catégories générales qui gouverneraient les individus, à l'inverse de la tradition aristotélicienne, dans laquelle les concepts généraux sont de prime importance. Par conséquent, les déterminations du global par le local ne sauraient être prises en considération, pas plus qu'une conception du genre en tant que *lieu de régulation des phénomènes linguistiques*, dans la mesure où la notion de *corrélation* au niveau textuel – et non au niveau phraséologique – semble trop contre-intuitive, et peut-être trop peu palpable, pour être immédiatement reliée à une fonction sociale.

Le courant dans son ensemble est de manière générale peu tourné vers les méthodes de statistique multidimensionnelle, et privilégie les méthodes plus basiques de statistique descriptive, qui sont certes plus réalisables et plus parlantes pour tout un chacun, a fortiori lorsqu'il s'agit de l'étude d'un marqueur spécifique et isolé. Les chercheurs travaillent donc exclusivement de manière qualitative (cf. *supra*), et le plus souvent sur corpus bruts et de taille restreinte, et procèdent par *hypothèses fonctionnalistes* quant à la présence/absence d'un

³⁸ Une étude de Myrna Gopnik (Gopnik, 1972) traite en outre de la structure des abstracts d'articles en vue d'une implémentation ultérieure

phénomène ou d'un marqueur particulier, ou par analyse fonctionnelle des mouvements rhétoriques des textes.

Si nous ne discutons pas l'intérêt pédagogique du courant ESP, et la pertinence de ses descriptions fonctionnelles, il nous semble cependant que les genres scientifiques manquent d'une description plus linguistique, plus formelle et plus quantitative, qui permettrait du reste aux chercheurs d'ESP d'orienter et d'objectiver leurs analyses qualitatives.

1.2.2. Genre de l'article et typologie du discours scientifique

Le discours scientifique s'inscrit dans la pratique sociale de la *communication scientifique*, qui pourrait elle-même se scinder en deux sous-pratiques, selon la nature du public visé. La *vulgarisation scientifique*, qui nous semble également s'apparenter au discours journalistique, vise à communiquer la recherche au grand public ; en ce sens, elle s'oppose à la pratique de communication scientifique académique, qui présente la particularité de faire coïncider le public de ses producteurs et celui de ses consommateurs. La présente étude porte exclusivement sur le genre de l'article de recherche, et n'abordera donc pas la question de la vulgarisation.

Le discours scientifique académique se réalise à travers des genres différents – comptes rendus, articles, conférences, posters, etc., qui varient dans leur forme et leur contenu selon la langue et le domaine considérés. De manière générale, les genres scientifiques présentent la particularité d'imposer des contraintes rigides, exprimées au moyen des feuilles de style et des consignes aux auteurs présentes dans les revues ou dans les modalités d'appel à communication des conférences ou des colloques. En ce sens et par rapport au cadre d'analyse dans lequel nous positionnons notre étude, les genres scientifiques présentent l'intérêt d'être plus régulés, et donc plus observables au niveau formel.

1.2.2.1. Le genre de l'article de revue

Si l'on s'intéresse à l'écrit, l'article de revue représente le genre scientifique le plus accrédité – du fait de son statut particulier dans le champ³⁹, le plus répandu et le plus observé. Plutôt analysé dans des travaux portant sur l'argumentation, ou dans la perspective rhétorico-fonctionnelle du courant ESP, l'article scientifique n'a jamais, à notre connaissance, été abordé comme lieu de régulation des phénomènes linguistiques, et dans une perspective formalisée de détermination du global sur le local, bien que l'on recense diverses tentatives de formalisation (e.g. Gopnik 1972, Paltridge 1997).

A. Contraintes éditoriales

Genre bureaucratique et fortement structuré, l'article de revue est soumis à des contraintes de forme et de contenu :

L'article scientifique obéit à des règles et à des codes particuliers au niveau du contenu (lourd appareillage théorique et méthodologique, présentation de résultats nouveaux ou de synthèses critiques de l'état des connaissances dans un domaine particulier du savoir) et de sa forme (importance du paratexte – notes, références bibliographiques, annexes, tableaux, schémas -, recours à un style impersonnel et utilisation d'un vocabulaire spécialisé). A travers son discours, le chercheur montre qu'il a intégré non seulement les connaissances de son domaine, mais aussi les savoir-faire, les codes, les valeurs, quand ce ne sont pas les tics. (Boure, 1998 : 107)

³⁹ Dans le champ (au sens bourdieusien du terme), la *valeur* du chercheur est souvent quantifiée au nombre d'articles qu'il a publiés (Merton, Latour et Woolgar).

La conformité scientifique et rédactionnelle des textes est généralement validée par un comité scientifique de lecture extérieur au comité de rédaction, bien que les modalités de sélection varient selon les revues (nombre de lecteurs, anonymat des textes, etc.) et selon leurs conceptions de la pratique scientifique :

Pour les tenants de la version « hard », calquée peu ou prou sur le modèle des sciences dites « dures », l'article scientifique est parfaitement codifié. D'abord dans son contenu : rigueur et rationalité du raisonnement (pas de place pour les explications idéologiques), recours à la méthode expérimentale, présentation de résultats empiriques et/ou méthodologiques nouveaux⁴⁰ (et non discours sur des discours) de telle sorte qu'ils puissent être discutés. Ensuite, mais plus accessoirement, dans sa forme, ce qui, dans les cas extrêmes, peut entraîner un conformisme d'exposition et d'écriture. Pour parler vite, disons que ce modèle est principalement celui des grandes revues disciplinaires, largement influencées par les revues anglo-saxonnes. (ibid., p. 111-112)

La fascination pour le modèle américain ne touche pas que les sciences dures, elle frappe également de nombreuses disciplines des SHS⁴¹, qui entérinent graduellement les critères internationaux instituant la distinction entre « bonnes » et « mauvaises » revues : adoption totale ou partielle de la langue anglaise et des standards de rédaction anglo-saxons (et donc du style *reader-friendly*⁴²), privilège accordé aux articles expérimentaux au détriment des textes plus théoriques, etc. Les articles considérés comme *majeurs* sont de plus en plus publiés dans les revues anglo-saxonnes, et on assiste à un désintérêt croissant des chercheurs à l'égard des revues françaises qui ne s'alignent pas sur le modèle anglo-saxon, et qui ne sont pas référencées par l'ISI (Institute for Scientific Information).

La plupart des revues sont aujourd'hui spécialisées, et correspondent aux sous-domaines scientifiques du champ. Les numéros sont souvent thématiques, ce qui renforce la cohésion des textes ; les articles hors thème éventuels sont alors regroupés au sein d'un *varia*. Dans ce cas, les textes ne sont pas soumis à l'évaluation du comité de lecture :

« Il faut savoir que la plupart des revues pratiquent la technique de la commande : elles s'adressent régulièrement à des spécialistes reconnus d'un domaine de recherche, souvent pour leur demander de faire le point soit sur leurs propres travaux, soit sur une partie de la littérature (*surveys*). En général, la rédaction et les auteurs sollicités se mettent d'accord sur les grandes lignes du contenu, ainsi que sur la forme (en particulier la taille), les ajustements entre les exigences des premiers et les propositions des seconds se réalisant à travers des navettes. » (ibid., p. 111)

Ces pratiques encouragent les phénomènes de cooptation, et les auteurs publiés au sein du numéro correspondent généralement au tissu social du responsable du numéro. On notera que ces modalités de publication sont plus fréquentes dans les communautés de taille restreinte, et relativement marginales au niveau international.

⁴⁰ Cette conception de la *nouveauté* est bien entendu relative, dans la mesure où les travaux qui démentissent véritablement les connaissances existantes sont pour le moins exceptionnels : la « nouveauté se ramène souvent à une simple différenciation au niveau des hypothèses, de la problématiques, ou des résultats, toujours compatible avec les paradigmes en cours, pour parler comme Kuhn » (ibid.)

⁴¹ La psychologie est aujourd'hui une discipline internationalisée, qui a complètement intégré le modèle de publication anglo-saxon.

⁴² Le style *reader-friendly* est largement demandé par les Anglo-Saxons ; très centré sur le lecteur, il se définit de la manière suivante (source : <http://ittimes.ucdavis.edu/tips.html>) : *Simplicité et transparence* (connaissance de son lectorat suffisante pour anticiper des questions éventuelles, pas de jargon, d'acronymes ou d'abréviations sans définition, phrases courtes et efficaces) ; *Utilisation de titres descriptifs et de sous-titres spécifiques et efficaces* (exemple : *Bad: Copyright Defined/Good: What is a copyright and who owns it?*) ; *Efficacité* (pas de digression, présentation des informations les plus importantes en premier) et *Utilisation d'un ton professionnel* (dans la mesure où le texte est destiné à ses pairs, adapter la terminologie et la syntaxe ; ne pas employer un style trop distant, autorisation du I, possibilité de recourir au YOU pour les conseils, les explications, etc.).

Ces considérations doivent bien entendu être prises en compte dans toute étude de l'article de revue, dans la mesure où elles influent sur son style et sa mise en texte générale. Un texte issu d'une revue imposant l'usage d'un style dialogique, ou *reader-friendly*, aura ainsi des caractéristiques linguistiques spécifiques (nombre plus important de phrases interrogatives, d'impératifs, de pronoms de première personne du singulier, etc.) qu'il serait impropre d'interpréter comme relevant d'un choix personnel de l'auteur.

B. Caractéristiques linguistiques du genre de l'article

Nous n'énumérerons pas l'ensemble des études dédiées à la description du discours scientifique et du genre de l'article, d'une part parce que les travaux recensés portent souvent sur l'examen d'un marqueur, ou d'un phénomène isolé, et d'autre part parce qu'ils sont fréquemment fondés sur des corpus de taille trop restreinte, ou trop spécifiques pour être généralisables⁴³.

On adopte habituellement en amont de l'étude d'un marqueur, ou de la caractérisation d'un genre, des éléments de description (très) généraux des textes scientifiques, comme par exemple ceux de Prince, repris par Prozorova (in *Cultures and styles of academic discours*, p. 305), qui s'intéresse au discours scientifique et à la notion d'inférence :

- Academic discourse is less explicit – it leaves more out, more inferencing is required.
- There is much metalinguistic inferencing (based on the assumption that the reader is familiar with various linguistic paradigms).
- Entities involved are highly abstract.
- There is an increased incidence of containing inferables, coupled with a blurring between unused and inferable entities (which is not the case with oral narrative).
- Cultural assumptions required for the inferencing are highly complex and abstract in nature.
- Entities are larger in academic discourse.
- No brand new entities are used at all, although there are some unused entities.
- Fifty percent of all subject NPs represent evoked entities, and one eighth of the non subjects do. More than 40 percent of the subjects are inferrables, mostly of the contained inferrable type.
- In academic discourse, as in oral narrative, subjects are more likely to be evoked than inferrable, and more likely to be inferrable than new (unused), which indicates a tendency toward a greater degree of givenness in subjects.

C. Structure du genre

L'article présente la particularité d'être un genre très structuré. Si la progression des sections de son développement est diversement soumise à la structure IMRAD – selon sa culture et son domaine scientifique de rattachement-, l'article se divise normalement en plusieurs éléments et sections facilement identifiables et encodables avec un standard comme la TEI :

- Le *titre*, placé en début d'article, dont la police varie en fonction des revues ; si le titre est marqué d'un astérisque, celui-ci renvoie aux *remerciements* éventuels de l'auteur, pratique plus fréquente dans la communauté anglo-saxonne que dans la communauté française ;
- Le(s) *nom(s)* de l'auteur (ou des co-auteurs), ainsi que leur(s) *affiliation(s)* ;

⁴³ Ce qui n'affecte bien entendu pas la qualité des études menées.

Le(s) *résumé(s)*⁴⁴ (*abstract*), placé en début ou en fin d'article (en fonction des normes imposées par la revue), et souvent suivi des *mots-clés* (*key-words*), correspondant aux différents thèmes abordés par l'article. On notera que la pratique du résumé est plus généralisée dans les disciplines des sciences de la nature, et qu'elle est moins systématique dans les sciences humaines : elle nous semble en effet largement dépendre de l'importance et des pratiques de l'éditeur de la revue. Les grands éditeurs scientifiques internationaux, comme Kluwer, Elsevier ou MIT Press, ont développé des sites web référençant les périodiques publiés, et proposent à la vente les numéros, voire les articles des revues, en version numérisée. Le résumé a donc une importance primordiale, dans la mesure où il représente la vitrine scientifique et commerciale de l'article. Il en va bien entendu différemment des revues universitaires françaises, qui disposent de moyens plus restreints : le résumé n'a donc pas la même importance, et il a souvent une vocation plus introductive que récapitulative.

- La *bibliographie*, généralement placée à la suite de la conclusion et facilement identifiable par sa structure ;
- Le *corps* de l'article, qui se découpe lui-même en plusieurs séquences :
 - *L'introduction* : placée au début du corps de l'article, elle peut être signalée par un titre (*introduction*), numérotée ou non (selon le cas, elle épouse plus ou moins le corps de l'article) ; certains articles ne comportent pas d'introduction explicite, l'article commençant directement par une première section constitutive du corps de l'article (mais remplissant en partie la fonction d'introduction) ;
 - Le *développement* (ou *analyse* ou *discussion*), divisé en différentes parties, souvent numérotées, organisées ou non selon une structure IMRAD ;
 - La *conclusion*, placée à la suite du corps de l'article, généralement indiquée par un titre variable (*conclusion(s)*, *perspectives*, *conclusions et perspectives*, *bilan*, etc.) ;
 - Les *notes de bas de page*.

D. Diversité sémiotique du genre de l'article

A l'instar de l'ensemble du discours scientifiques, le genre de l'article se caractérise par la présence de sémiotiques distinctes, le plus souvent délimitées dans le texte. La science étant généralement associée à un processus de formalisation, on recense de nombreux objets non linguistiques à visée d'objectivation : *tableaux*, *graphiques*, *formules*, *équations*, etc.

On pourrait penser que ces éléments sont l'apanage des sciences exactes, qui privilégient les modèles mathématisés hypothético-déductifs, considérés comme étant la forme *achevée* d'une théorie scientifique, mais il n'en est rien : les sciences humaines se formalisent de plus en plus. Ce phénomène est lié d'une part à l'importation de modèles mathématiques – le plus issus des mathématiques appliquées, i.e. informatique, statistiques -, et d'autre part à une certaine imitation des processus d'objectivation des sciences exactes, parfois bien caricaturale :

(2) l'équation écrit = oral surmarqué, c'est-à-dire écrit = oral + effet de spontanéité + expressivité + subjectivité, soit la formule :

É = O + [/S/+Ex+Sub] [SCL091] (exemple emprunté à F. Rinck, 2004, extrait d'un article de critique littéraire)

⁴⁴ Le nombre de résumés dépend des normes imposées par la revue ; certaines revues exigent ainsi la rédaction de plusieurs abstracts, rédigés dans des langues différentes – dont l'anglais la plupart du temps

Dans la mesure où ils sont caractéristiques du genre, et étant donné qu'ils participent à l'effet d'objectivation des textes scientifiques, l'examen de ces hors-textes nous semble particulièrement intéressant. Pourtant, comme le souligne Rastier (*Pour une sémantique des textes théoriques, à paraître*) :

Faute d'une sémiotique adaptée, nous manquons cependant d'une théorie suffisamment élaborée des effets de réel théorique ; elle serait d'autant plus nécessaire que les textes scientifiques multimédia se multiplient et se complexifient.

Si les éléments de description mentionnés *supra* doivent être pris en compte comme un préalable à l'étude du genre de l'article, on ne saurait aborder ce dernier sans l'inscrire dans la typologie du discours scientifique, ou en d'autres termes, sans aborder ce qui participe à sa détermination, c'est-à-dire les *champs génériques* dans lesquels il rivalise, les domaines scientifiques, l'axe théorique/appliqué et les langues auxquels il se rattache et les styles personnels et nationaux par lesquels il s'instancie.

1.2.2.2. *Champs génériques*

Les genres scientifiques oraux nous semblent d'abord rivaliser avec les genres scientifiques écrits : la communication scientifique s'effectue en effet à travers ces deux canaux, et la même recherche peut être exposée dans le cadre d'un article de revue ou d'une conférence, bien que l'article ait en principe plus de valeur que la présentation. La présentation de conférence donnant généralement lieu à publication (dans le cadre du genre écrit – dérivé de l'oral – de l'*article dans des actes*), l'examen des deux genres concurrents est réalisable sur corpus écrits, ce qui facilite l'émergence de lieux de contrastes⁴⁵. La distinction écrit/oral ne constitue bien évidemment pas le seul terrain possible de contraste et de rivalité des genres scientifiques : il est par exemple concevable de contraster la communication scientifique *interne* aux laboratoires et aux groupes de recherche avec la communication *externe*, destinée à un public plus large de pairs, mais cette rivalité concerne peu le genre qui nous intéresse.

En outre, il nous semble qu'il existe une certaine concurrence entre les genres au sein de la revue, dans la mesure où certains genres tentent de s'imposer face à l'article, qui commence d'ailleurs à se diversifier en plusieurs genres distincts dans le monde anglo-saxon : le genre de l'*article* s'opposerait ainsi au genre plus polémique de la *discussion*, au genre plus court de la *communication courte* (*short communication*), au genre présentatif de la *présentation de revue*, ou encore au genre post-discussion de la *réponse* (*response*).

Ces champs de concurrence, et cette multiplication des genres scientifiques, posent problème à l'analyste contrastive, dans la mesure où les genres peuvent être comparables d'une communauté à l'autre, même s'ils ne portent pas la même appellation : par exemple, le genre de la *discussion*, qui est largement répandu dans les communautés scientifiques anglo-saxonnes, ne l'est pas (encore ?) en France, ce qui ne signifie certes pas que certains articles scientifiques français ne soient pas également dédiés à la controverse. L'écart entre le genre et sa désignation est ainsi manifeste selon les langues, et il semble difficile à l'analyse contrastive des genres de s'en tenir systématiquement au principe d'objectivité émis par Todorov, et préconisé par nombre d'analystes, de convenir « d'appeler genres les seules classes de textes qui ont été perçues comme telles au cours de l'histoire ».

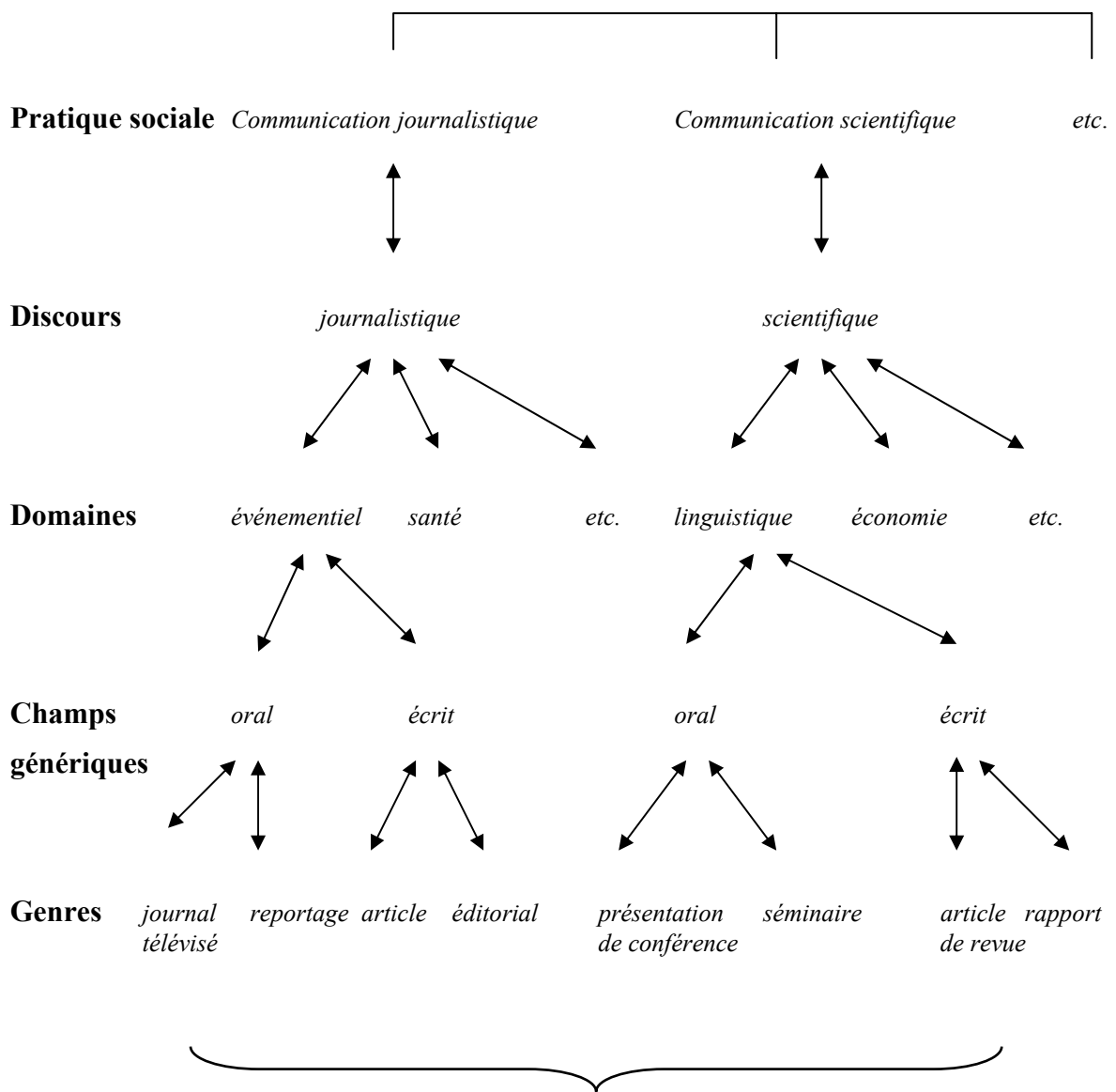
⁴⁵ Notons qu'il est parfaitement envisageable d'analyser les genres scientifiques oraux (e.g. Rowley-Jolivet, 2002) et de les contraster aux genres écrits à l'aide de paramètres tels que : intervention/non intervention du public avec l'auteur, représentation du public par l'auteur possible/impossible, etc.

1.2.2.3. Genres et domaines

Genres scientifiques écrits et oraux contrastent ainsi à l'intérieur du discours scientifique, qui se scinde lui-même en différents *domaines*. Notons d'emblée que cette notion de *domaine*, qui semble cruciale pour décrire les textes spécialisés, est fondamentalement ambiguë, dans la mesure où elle renvoie à la fois à l'idée de *sous-pratique* – ou *sous-spécialisation* – et à la notion de *langue de spécialité*, plus linguistique – et plus précisément terminologique dans la plupart des acceptions.

A. Domaines et disciplines scientifiques

Les genres scientifiques ne sauraient d'abord être décrits sans recourir à la notion de *domaine de spécialisation* : la mondialisation de la science, évoquée *supra*, a entraîné une compartimentation et une hyper-spécialisation de la recherche scientifique qui a donné lieu à autant de *domaines scientifiques* que de spécialisations. Ce phénomène n'est d'ailleurs pas l'apanage du discours scientifique, et concerne également de nombreux genres professionnels. Ainsi, les domaines scientifiques et journalistiques renvoient respectivement aux *disciplines scientifiques* et aux *sujets d'information*. Le discours scientifique se subdivise d'abord traditionnellement en deux branches : les sciences de la nature et les sciences de l'homme et de la société, qui se scindent elles-mêmes en autant de branches que de disciplines scientifiques institutionnalisées. La presse journalistique se subdivise elle-même en *domaines d'information*, qui englobent des sujets aussi distincts que l'informatique, la santé ou la décoration intérieure. Les domaines nous semblent donc renvoyer aux sous-spécialisations d'une pratique, qui les rassemble toutefois au sein d'un même type de discours :



A première vue, on pourrait considérer les domaines comme des *sous-pratiques sociales et discursives*, qui détermineraient leurs champs génériques et leurs genres propres. Pourtant, il faut reconnaître que les genres leur correspondant transcendent souvent la notion de domaine, dans la mesure où ils portent la même désignation, et renvoient généralement à la même réalité sociale et communicative : la *présentation de conférence*, ou l'*éditorial* sont ainsi parfaitement identifiables, quel que soit le domaine scientifique ou journalistique. Il semble donc que l'ensemble des domaines soit fondé sur un modèle discursif prototypique, et probablement historique, qui déterminerait les champs génériques et les genres : les domaines se différencieraient selon leurs modalités d'appropriation du modèle et des genres, ce qui explique par exemple les différences notoires observées entre les domaines en ce qui concerne le genre de l'article (cf. projet KIAP).

Si la notion de domaine, de même que ses relations avec le genre, reste encore à explorer, l'étude d'un genre donné doit prendre en compte ce facteur de variation : à moins de chercher

à l'explorer comme les chercheurs du projet KIAP, un corpus génériquement et domanialement homogène nous semble plus propice à la qualification d'un genre, malgré les limitations que la démarche implique : les conclusions obtenues ne pourront effectivement pas être généralisables à l'ensemble des sous-pratiques.

B. Domaines et langues de spécialité

La notion de *domaine* est généralement associée, sur le plan linguistique, aux niveaux lexical et terminologique⁴⁶ : l'ensemble des genres scientifiques d'une discipline appartiendrait ainsi à un domaine terminologique particulier, qui établirait ses conventions sémantiques propres. Cette conception s'est longtemps appliquée à l'étude du discours scientifique⁴⁷, qui bénéficie encore aujourd'hui d'une validation universelle du plan de son contenu, qui l'affranchirait des langues et des conventions textuelles. En effet et comme le souligne Rastier, « la tradition scientifique occidentale fait de la science une affaire de concepts et de termes, non de textes » (*Pour une sémantique des textes théoriques, à paraître*).

Certaines théories franchissent pourtant le niveau conceptuel en considérant les domaines comme des *langues de spécialité*, voire des *sous-langages* :

La pertinence d'une hypothétique opposition entre « langue usuelle » et « langue spécialisée » - à laquelle adhère un grand nombre de locuteurs francophones confrontés à l'hermétisme de certains textes techniques - ne peut être envisagée qu'à partir d'une définition de la langue. Si nous nous référons à la conception saussurienne, l'antagonisme envisagé précédemment entre *usuelle* et *spécialisée* convierait à appréhender le problème sous la forme de deux systèmes linguistiques (lexicaux et syntaxiques) formellement indépendants et à considérer que les règles qui régissent par exemple le discours technique sont étrangères au système français. (Cusin-Berche, p. 105)

Les langues de spécialités seraient ainsi des sous-systèmes linguistiques, bien qu'on insiste encore largement sur le niveau lexical, en s'accordant à reconnaître les règles syntaxiques employées comme identiques à un « système linguistique général », eu égard aux similarités observées entre les discours. Etant donné qu'il n'existe pas d'activités humaines entièrement cloisonnées et que les discours et les domaines s'influencent entre eux, Lerat (1995) parle d'ailleurs de *langues spécialisées*, qui renverraient à des « usages socialement normés de plurisystèmes » : outre une priorité accordée à l'écrit, ces usages se caractériseraient par :

- une syntaxe identique à celle des langues de référence, mais avec des prédilections énonciatives (tournures impersonnelles, phraséologies professionnelles, etc.) ;
- une morphologie composite accueillante aux alternances de morphèmes, aux emprunts dénommatifs et aux formules complexes, incluant du lexical général et du lexical spécifique ;
- la présence de signes non linguistiques au sein des énoncés, selon les notations adoptées par les disciplines ;
- une conception sémantique non ethnocentrique, du fait de l'universalité potentielle des notions scientifiques et techniques et de leur propagation d'une culture à l'autre.

⁴⁶ Ils renverraient dans ce cas à un « fonctionnement sémantique original des unités lexicales constitutives des vocabulaires spécialisés » (Cusin-Berche, p. 94) qui participerait à la définition des textes spécialisés.

⁴⁷ On a par exemple contrasté les terminologies plus « dures » des sciences de la nature avec les terminologies plus « molles » des sciences de l'homme et de la société.

Ces spécificités ne peuvent être (in)validées qu'à partir d'une étude linguistique des usages et des genres, que les études domaniales tendent à négliger. Il nous semble en outre qu'un corpus homogène en genre et en domaine pourrait autoriser l'émergence et la formalisation d'un sous-système linguistique, à partir de l'analyse des régularités et des corrélations observées au niveau textuel : l'acceptabilité des structures syntaxiques et de l'ensemble des phénomènes linguistiques varie en effet selon les genres, et nous semble devoir, avec la linguistique de corpus, être remplacée par celle d'*usage*. Enfin, la thématique des textes scientifiques, qui renvoie aux concepts et aux notions, doit certes être abordée par domaine, mais également par genre, dans la mesure où c'est au niveau du genre que les plans de l'expression et du contenu sont par définition corrélés.

1.2.2.4. Recherche théorique/appliquée

De même que l'on oppose les sciences de la nature à celles de la culture, on a coutume de distinguer les recherches *théoriques* des recherches *appliquées*, de même que l'on différencie le *technique* du *scientifique*. Si la première distinction est relativement bien tranchée, la seconde semble moins nette, dans la mesure où elle est plus ou moins caractérisée selon les disciplines, les cultures et les époques. Ainsi, elle était avant tout sociale au XIX^e siècle, et instaurait une hiérarchie sociale et intellectuelle.

En règle générale, on aime à penser à la recherche scientifique comme poursuite de ce que l'on appelle le 'savoir pur', et à la recherche appliquée comme à une activité moins noble : il semble ainsi établi que la science poursuit des objectifs visant à élaborer des connaissances, tandis que la technique a pour objectif la réalisation d'outils et d'objets permettant à l'homme d'améliorer son existence et sa survie.

Cependant, la distinction est, aujourd'hui plus qu'hier, discutable, dans la mesure où la frontière entre science et technique tend à s'estomper. En effet, l'avènement de la *Big Science*⁴⁸, issue de la physique nucléaire et de l'espace, a contribué à une réorganisation de la distinction. Certains domaines scientifiques, comme la physique des particules élémentaires ou l'astrophysique nécessitent des investissements si coûteux qu'ils ne peuvent être pris en charge que par les états. La science est de fait matériellement subordonnée à la technique, dans la mesure où les applications technologiques de découvertes scientifiques peuvent entraîner un développement inattendu de la discipline.

Les sciences de l'information et du codage se sont développées à cause de cette socialisation d'une technologie et de nouveaux paradigmes scientifiques sont en train de naître dans un processus dialectique d'aller et retour entre la science fondatrice et les technologies qui, d'abord associées et dérivées, conquièrent leur autonomie pour questionner et interpeller à leur tour la science originelle. (Pignon, 97, p. 40)

De fait, l'institutionnalisation du financement de la recherche scientifique révèle l'impossibilité d'une science 'pure' de toute contingence ; avec l'utilitarisme de la société moderne la science est devenue une « fonction publique » qui doit justifier ses objectifs.

Au sein du domaine linguistique, le succès actuel de la linguistique appliquée illustre ces changements ; le développement massif des industries de la langue et des applications en didactique a considérablement accru le nombre d'articles *appliqués* dans le champ.

En outre, la séparation cartésienne en deux branches, fondamentale et appliquée, ne résiste pas aux différences culturelles ; ainsi, en polonais, le style du discours scientifique n'est

⁴⁸ Expression anglo-saxonne qui n'a pas d'équivalent satisfaisant en français – il est toutefois possible de le traduire par *Grands Projets* ; renvoie à une « science dont les nécessités internes réclament des équipements et des équipes qui ne sont plus à la mesure artisanale de l'homme. » (Pignon, 1997, p. 34)

utilisé que dans le cadre d'un échange entre 'égaux', et sur des sujets considérés comme relevant de la 'vraie science'. Ce style ne sera pas employé dans des échanges pourtant universitaires dans lesquels le savoir est transmis dans des perspectives différentes, comme la didactique. La langue polonaise ne dispose de fait pas de termes comme l'adjectif 'academic' qui qualifie la transmission du savoir dans des contextes tant scientifiques qu'éducatifs. De manière générale, la culture européenne privilégie la théorie alors que la culture anglo-saxonne donne le pas aux recherches empiriques. Par exemple, dans l'étude du discours, les recherches linguistiques américaines s'orientent plus délibérément vers des applications tandis que les recherches européennes cherchent à construire des modèles du discours et se désintéressent généralement de leur applicabilité potentielle (Galtung, 1985).

Le tableau de F. Gadet (cité dans Maingueneau, 1987) illustre nettement cette différence :

	AD française	AD anglo-saxonne
Type de discours	Ecrit Cadre institutionnel doctrinaire	Oral Conversation quotidienne ordinaire
Buts assignés	Visées textuelles Explication-forme Construction de l'objet	Visées communicationnelles Description-usage Immanence de l'objet
Méthode	« structuralisme » linguistique et histoire	Interactionisme Psychologie et sociologie
Origine	linguistique	anthropologie

Tableau : AD française et AD anglo-saxonne

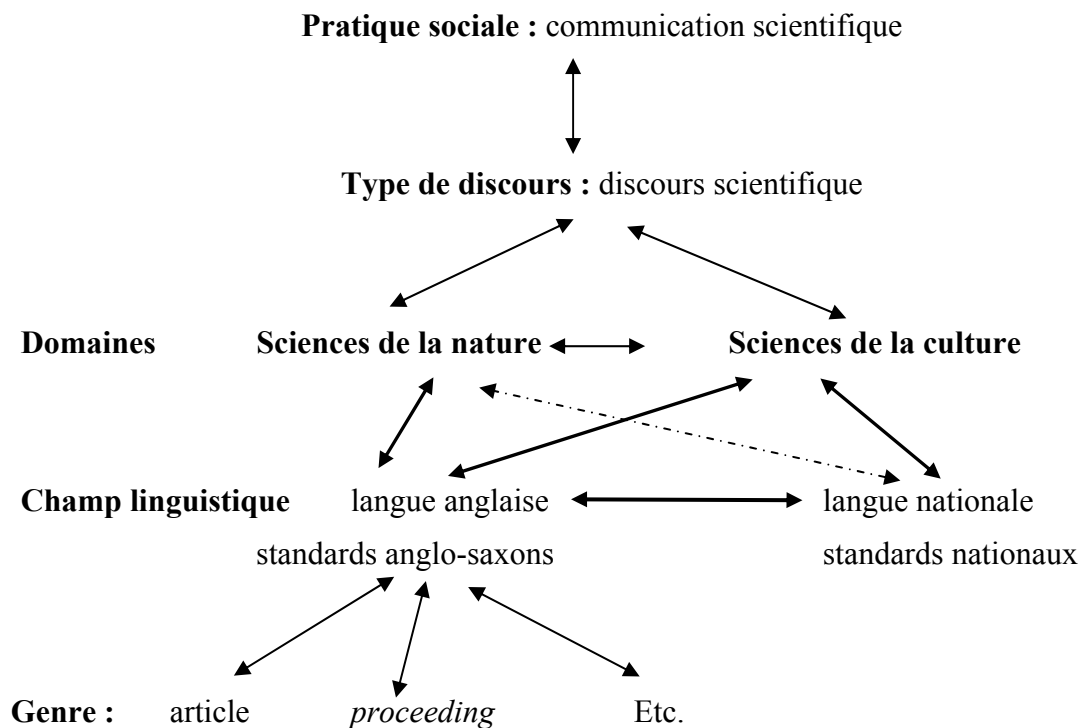
1.2.2.5. Genres scientifiques et langues en contraste

La mondialisation de la science a conféré à la langue anglaise un statut de *lingua franca* scientifique et l'ensemble des sciences de la nature est aujourd'hui *internationalisé*, i.e. les articles majeurs sont publiés en anglais et dans des revues internationales anglophones. L'analyse des genres, et *a fortiori*, l'analyse contrastive des genres, ne saurait ignorer ce phénomène, qui entraîne des conséquences importantes sur la typologie linguistique du discours scientifique.

Si la langue anglaise – et le style rhétorique anglo-saxon – est la norme rédactionnelle imposée par l'ensemble des sciences de la nature, il en va différemment des sciences humaines et sociales : certaines disciplines, comme la psychologie par exemple, se sont internationalisées et ont adopté un mode de communication anglo-saxon (rédaction en anglais, revues à points, etc.), alors que de nombreuses disciplines interprétatives ont conservé leurs standards et leur langue nationale de publication. Il serait pourtant inexact de soutenir que ces communautés scientifiques fonctionnent à circuit fermé et sans interaction avec la communauté internationale anglo-saxonne : un nombre important de revues nationales

acceptent, ou commencent à accepter des textes rédigés en anglais, ce qui leur confère d'ailleurs un statut de *revue internationale*.

On relève ainsi une certaine rivalité entre les langues et les standards de publication, qui induirait l'existence d'un *champ linguistique* de concurrence entre la langue hégémonique anglaise et les langues nationales : si ce *champ linguistique* est aujourd'hui révolu dans l'ensemble des sciences de la nature, étant donné que l'anglais l'a emporté sur les autres langues, il est encore bien actuel dans de nombreuses disciplines interprétatives :



Il va de soi que l'anglais et les standards anglo-saxons partent avec un sérieux avantage bien que la langue anglaise ait elle-même à subir certaines modifications, liées à son appropriation par des non natifs : par exemple, l'emploi de la structure *to allow*, initialement incorrecte, est aujourd'hui autorisée, et de nombreux non natifs transposent les structures impersonnelles d'usage dans leur langue à l'anglais (*one might think, one can consider, etc.*).

La normalisation de la langue anglaise influe considérablement sur les genres, et sur le discours scientifique en général, dans la mesure où certains genres n'existent que dans certaines langues, ce qui restreint considérablement les champs de comparaison. En outre, les sciences de la nature accordent plus de valeur aux textes publiés en anglais, dans des revues internationales anglophones. Le genre de l'article scientifique de biologie en français aura par conséquent un statut bien différent de celui de l'article rédigé en anglais : les rares spécimens d'articles scientifiques rédigés en français sont des pièces mineures ou des redites de communications déjà publiées en langue anglaise, et sont peu, voire non accrédités par la communauté scientifique internationale, ni même par la communauté française. Ces textes ne sont donc pas produits, reçus et interprétés de la même manière que les textes publiés en

anglais : on pourrait éventuellement considérer qu'ils représentent deux genres différents⁴⁹. Ils sont pourtant particulièrement intéressants dans une perspective contrastive interlangue, dans la mesure où la pratique d'écriture scientifique anglo-saxonne influence très nettement la pratique française, ce qui offre de nombreux pivots de comparaison, contrairement à la mise en contraste des productions génériques de deux communautés scientifiques plus distinctes, dans la mesure où les lieux de contrastes linguistiques sont moins évidents.

1.2.2.6. Styles personnels et styles nationaux

Si l'on a pu penser que les textes scientifiques étaient exempts de styles personnels, de la même manière que l'on a pu concevoir le discours scientifique comme le discours impersonnel de la Vérité, de nombreux travaux, sociologiques, linguistiques ou de la veine ESP, ont depuis exploré les variations des genres scientifiques et fait émerger maintes différences stylistiques entre les textes, tant personnelles (ou d'auteurs) que nationales (ou culturelles).

Les variations des styles personnels observées notamment dans le genre de l'article transcendent les positions sociales et statutaires des auteurs, dans la mesure où le discours scientifique en général se caractérise par une relation entre pairs, dans laquelle les acteurs ont un statut sociologique en principe analogue, sinon identique. Chaque pratique instituant son système de valeurs et ses jeux de pouvoir susceptibles d'affecter son discours, la communication scientifique met en scène des luttes de pouvoir symbolique qui s'expriment souvent à travers le degré d'accréditation et la reconnaissance de l'auteur dans son champ. La notion d'*expert* est ainsi fondamentale, dans la mesure où les standards sont discutables entre membres compétents d'une communauté de discours. Si un expert devient une autorité, il pourra développer un style idiosyncrasique qui ne sera pas imité – mais qui pourra être critiqué du fait qu'il viole des standards académiques.

Le caractère plus ou moins marqué des styles personnels varie bien entendu selon la taille et les standards d'une communauté scientifique donnée. Les communautés de taille restreinte fonctionnent ainsi sur des modes plus relationnels, ce qui encourage les procédures de sélection non anonymes des textes⁵⁰, ainsi que certains phénomènes de cooptation : de tels cadres sont propices au développement de styles personnels plus prononcés, et au détournement des canons du genre, dans la mesure où ces libertés ne sauraient remettre en question l'accréditation et le statut du chercheur, bien (re)connu de ses pairs. Il en va bien sûr différemment des communautés scientifiques plus importantes, qui normalisent l'usage d'un style « neutre » par le biais de procédures de sélection anonymes : dans un tel cadre, seuls les experts internationaux, qui échappent à de telles contraintes en étant invités à communiquer, sont autorisés à cultiver un style bien à eux, tandis que les chercheurs moins distingués sont encouragés à se maintenir dans un rôle stylistique plus conventionnel et fixé *a priori*, comme l'exige l'ensemble des genres plus normés.

Les styles personnels varient de surcroît avec les cultures, ce que le courant ESP a bien mis en valeur ; en d'autres termes, les styles personnels dépendent eux-mêmes des *styles nationaux* de leurs communautés scientifiques d'appartenance, les genres reflétant des styles d'écriture normés par des pratiques. Ainsi, les articles anglo-saxons sont généralement orientés vers le lecteur ; l'emploi du style *dialogique* (ou *reader-friendly*) présente l'article

⁴⁹ Ce qu'ils sont d'ailleurs amenés à devenir, dans la mesure où le genre de l'article français dans les sciences de la nature se fait de plus en plus rare.

⁵⁰ Légitimé par le fait que les chercheurs se connaissent et se reconnaissent aisément.

scientifique comme un contrat à négocier, dans lequel on donne au lecteur le sentiment de participer ; contrairement à la tradition intellectuelle germanique⁵¹, qui favorise un style monologique (exposition/contemplation) et une part importante de créativité. Clyne (1998) compare par exemple les différences organisationnelles des textes académiques anglo-saxons et allemands ; si les premiers privilégient les développements linéaires, l'explicitation préalable des termes employés ou le style *reader-friendly*, les seconds préfèrent les développements digressifs et la non-explicitation des termes, liée à l'idéalisation du savoir. Chez les Anglo-Saxons, c'est effectivement l'auteur qui doit s'efforcer de rendre son texte lisible, tandis que ce sont les lecteurs qui doivent faire l'effort d'adaptation dans les communautés germanique et française. Diverses études (e.g. Saviile-Troike, 1982) décrivent par ailleurs la tendance des cultures de l'ouest, plus individualistes, à employer des styles plus directs et explicites, tandis que les sociétés orientales insisteraient davantage sur les valeurs collectives et l'harmonie du groupe, employant un style interactif plus défensif et plus imprécis. La culture de l'auteur se reflète ainsi dans la langue, à travers l'emploi de structures affirmatives ou modales, de *hedges*, etc.

Ainsi, l'emploi de la référence directe (deixis) est variable : dans certains cultures, l'emploi du pronom personnel de première personne du singulier peut avoir des effets non désirés.

Perceptions of politeness are likely to differ across cultures, sometimes quite alarmingly, and it is of crucial practical significance that we should know how to predict the kinds of communication breakdown that are likely to arise from such conflicting perceptions. (Watts *et al.* 1992 : 15).

Sauf les cas où l'auteur est accrédité par l'ensemble de la communauté (comme Chomsky) et reconnu en tant qu'*expert*⁵², l'usage de la première personne du singulier peut être considéré comme trop informel, voire impropre, ou au contraire comme autoritaire, voire arrogant. S'il semble plus répandu dans les articles anglo-saxons, le français ou les langues de l'Est ne l'emploient guère. Certaines occurrences commencent pourtant à être observées en français, en raison de l'influence de la norme imposée par la communauté anglo-saxonne. Le style employé dans des articles rédigés en anglais par des non-anglophones sera ainsi plus ou moins influencé par des procédés linguistiques usités dans la langue première ; cette dernière subira toutefois l'autorité du style requis en anglais.

La question des *styles (personnels ou nationaux)* en tant qu'*usages singuliers du genre*, intéresse ainsi la description des genres scientifiques, et *a fortiori* le cadre méthodologique dans lequel nous nous inscrivons. En effet, le style étant plus souvent abordé en termes morphosyntaxiques ou phonologiques, selon la nature écrite ou orale du genre, genres et styles varient ainsi sensiblement au niveau morphosyntaxique : si l'on sait que les genres l'emportent sur les styles en matière de classification automatique, et qu'un corpus de taille significative neutralise normalement les styles au profit du genre, la rigidité stylistique du genre peut néanmoins être éprouvée au moyen d'une classification menée sur corpus génériquement homogène : un classement du corpus par auteur indiquerait en effet la normalisation d'une plus grande amplitude stylistique, et la présence de styles personnels scientifiques.

⁵¹ Allemands, polonais et tchèques

⁵² L'auteur souligne dans ce cas sa contribution.

1.2.3. Genre de l'article et domaine linguistique

Comme nous l'avons évoqué *supra*, analyser l'article scientifique ne saurait se faire hors du choix d'une discipline. Notre formation et nos compétences nous conduisent à adopter, à l'intérieur des sciences humaines, le domaine linguistique. La sélection d'un domaine inconnu de l'analyste pose en effet des problèmes d'expertise et de connaissance des standards régissant le genre. Nous admettons que l'ignorance d'un champ et de ses enjeux peut constituer un principe d'objectivation des observations effectuées, à l'instar de la démarche adoptée par Latour et Woolgar (1979). Cependant, si les études sociologiques doivent dissocier les faits sociaux de leurs représentations sociales, il nous semble que l'analyse linguistique n'est pas soumise aux mêmes contraintes, dans la mesure où la caractérisation d'un genre s'objectivise d'abord par son corpus. La méconnaissance d'un domaine scientifique entrave non seulement la sélection des descripteurs linguistiques – hypothèses qui seront (in)validées par le corpus, mais encore l'interprétation des résultats, qui doivent dans ce cas être soumis à la validation d'un ou plusieurs experts – non linguistes – qui confronteront leurs représentations aux résultats linguistiques obtenus. Cette remarque ne signifie bien évidemment pas que le domaine scientifique linguistique doit être privilégié par le linguiste⁵³, mais que son adoption ne remet nullement en cause la scientificité ou l'objectivité de l'étude.

Le domaine scientifique linguistique présente la particularité d'être à la croisée des sciences humaines et sociales et des sciences de la nature. La discipline ne s'est institutionnalisée en France que récemment et nombre de ses chercheurs ont une formation annexe de lettres modernes ou de philosophie du langage, qui transparaît dans le style et les références des textes publiés. Le développement de l'informatique et du traitement automatique des langues, et le succès des formalismes linguistiques a en outre orienté un versant de la discipline vers les sciences dites 'dures' : on recense ainsi des revues de linguistique logique et mathématique dans le monde linguistique anglo-saxon et plusieurs revues de linguistique informatique dans les deux communautés française et anglo-saxonne.

Outre diverses études issues du courant ESP qui ont relevé un taux important de variation dans les articles de linguistique, qui traduirait une absence de standard multinational dans le champ (Duszak, 1997, Evangelisti Allori, 1994), peu d'études ont à notre connaissance cherché à caractériser le discours et les genres scientifiques linguistiques en corpus : la plupart des travaux recensés contrastent le genre de l'article de linguistique à d'autres domaines de recherche et s'inscrivent dans des perspectives scientifiques et méthodologiques distinctes des nôtres (le projet KIAP, cf. Breivega et al. 2002, Rinck Fanny – thèse en cours portant sur les aspects linguistiques et socio-discursifs de l'écriture d'articles de recherche et de l'identité disciplinaire et de la singularité des scripteurs doctorants).

1.2.3.1. Linguistique théorique et appliquée

La distinction entre linguistique théorique et appliquée remonterait aux années 40, date à laquelle on relève le premier usage de l'expression *applied linguistics* (*linguistique appliquée*), qui faisait alors référence aux pratiques d'enseignement :

It is worth recalling that this term, now recognized by many as a misnomer, arose by historical accident, its first systematic use having been with reference to foreign-language teaching methodology in the US of the 1940's. (Burgarski, 1987:13)

Comme le souligne Miseska (1987), le champ de la linguistique appliquée se serait élargi avec l'avènement du générativisme :

⁵³ Les linguistes recourent d'ailleurs fréquemment à des exemples extraits de textes de linguistique.

In an endeavor to secure for linguistics membership in the distinguished society of exact sciences, generative grammarians devoted their energies to the study of the linguistic competence of the ideal language user and relegated to applied linguistics all issues suspected of being theoretically "impure". As a consequence, the scope of applied linguistics widened, growing into a complex of activities largely concerned with language acquisition and language use. (Miseska, 1987:XV)

Si la distinction est aujourd'hui communément admise, elle demeure pourtant mal délimitée, tant sur le plan pragmatique que théorique :

Pragmatically, it is all right to talk of applied linguistics in the sense of a complex of various research activities of considerable social relevance. Theoretically, however, the scope and content of the notion are open to dispute as on this level applied linguistics does not seem to exist as a distinct and coherent field of study. Few if any convincing criteria can be found which might justify the separation of "applied linguistics" from "linguistics" in terms of either subject matter or methodology, the study of language as a whole having of late become increasingly interdisciplinary and socially conscious. A distinction may perhaps be drawn only in terms of goals, the general aim of linguistics being to understand human language, and that of applied linguistics, to contribute to the solution of problems arising in the use of language in human societies – though even this appears questionable in light of contemporary developments. (Burgarski, *ibid.*)

Il serait ainsi délicat d'élaborer une cartographie précise des champs scientifiques linguistiques français et anglo-saxons au regard de cette opposition⁵⁴, d'autant que les sous-domaines linguistiques procèdent parfois à la même distinction en leur sein, en « théorisant la pratique » par exemple. Il n'est en outre pas certain que les articles plus appliqués (comme ceux de didactique linguistique, ou de traitement automatique) se distinguent des textes théoriques à partir de critères linguistiques stables, dans la mesure où la distinction n'est pas institutionnalisée comme elle l'est en Pologne par exemple :

It is also arguable that alienation of academic from general language supports alienation of theoretical from practical knowledge. In Polish, scientificity implies a closed-circuit transmission of knowledge among the knowledgeable. The term "scientific style" can only be applied to exchange information among "equals", and only on matters concerning "real" science. It is not used in contexts where knowledge is transmitted for other academic purposes (e.g., education); in such cases its qualification as "popular" scientific style becomes mandatory. As a result, Polish has no single equivalent of the English term "academic" for describing communication styles in both research and educational settings. (Duszak, 1997)

Ce n'est donc qu'à partir de l'examen du corpus qu'il sera éventuellement possible de préciser la distinction.

1.2.3.2. Les langues de la linguistique

Un examen du champ scientifique révèle rapidement que les articles faisant autorité dans le domaine ne sont pas nécessairement rédigés en anglais⁵⁵. Certaines théories ont de fait connu un succès considérable dans certaines collectivités, mais s'avèrent peu représentées dans d'autres⁵⁶ ; en outre, certaines branches de la linguistique, comme la linguistique romane, ne s'intéressent pas à la langue anglaise, ce qui n'encourage pas les publications en anglais.

⁵⁴ Il n'existe d'ailleurs pas, à notre connaissance, de cartographie précise et délimitée des branches de la linguistique française, les revues et les articles n'étant pas indexés par mots-clés, ni même numérisées dans certains cas, ce qui nous semble d'ailleurs dommageable.

⁵⁵ Sans doute est-ce dû à l'objet même de la linguistique – la langue et *les* langues...

⁵⁶ La théorie culiolienne a ainsi connu un succès considérable auprès des anglicistes français et des applications quasi-immédiates dans l'enseignement de la langue anglaise, alors qu'elle connaissait un succès moindre en linguistique française.

Le domaine linguistique est, à l'instar de l'ensemble des sciences humaines, plus marqué par les *styles nationaux*, contrairement aux sciences de la nature, tenues, dans leur effort d'objectivation, de respecter le principe d'interchangeabilité de l'observateur et de fait de limiter, voire d'effacer les différences culturelles. Au contraire, les sciences humaines privilégient une posture critique que reflète l'interprétation des textes, et portent la marque de leurs conditions socio-historiques de production.

L'influence de la communauté anglo-saxonne sur l'ensemble du domaine linguistique est cependant indiscutable. La multiplication des revues et des conférences internationales exige régulièrement la rédaction d'articles ou de présentations en anglais par des non natifs ; le taux important de variation constaté dans ces articles traduit une absence de standard multinational dans le champ linguistique, d'ailleurs confirmée par plusieurs études contrastives portant sur des mouvements rhétoriques⁵⁷ différents de l'article scientifique – résumés, introduction (Duszak, 1997, Evangelisti Allori, 1994). La communauté anglo-saxonne impose toutefois une norme rédactionnelle, en témoignent les refus *liés au style* qu'essuient régulièrement certains chercheurs non anglophones lorsqu'ils soumettent leurs articles dans des revues internationales. De fait sont dispensés aux Etats-Unis des cours de rédaction d'articles scientifiques destinés aux jeunes chercheurs, ce qui est encore peu répandu en France.

⁵⁷ Au sens des *moves* de Swales (Swales 1990)

1.2.3.3. Typologie du discours scientifique linguistique

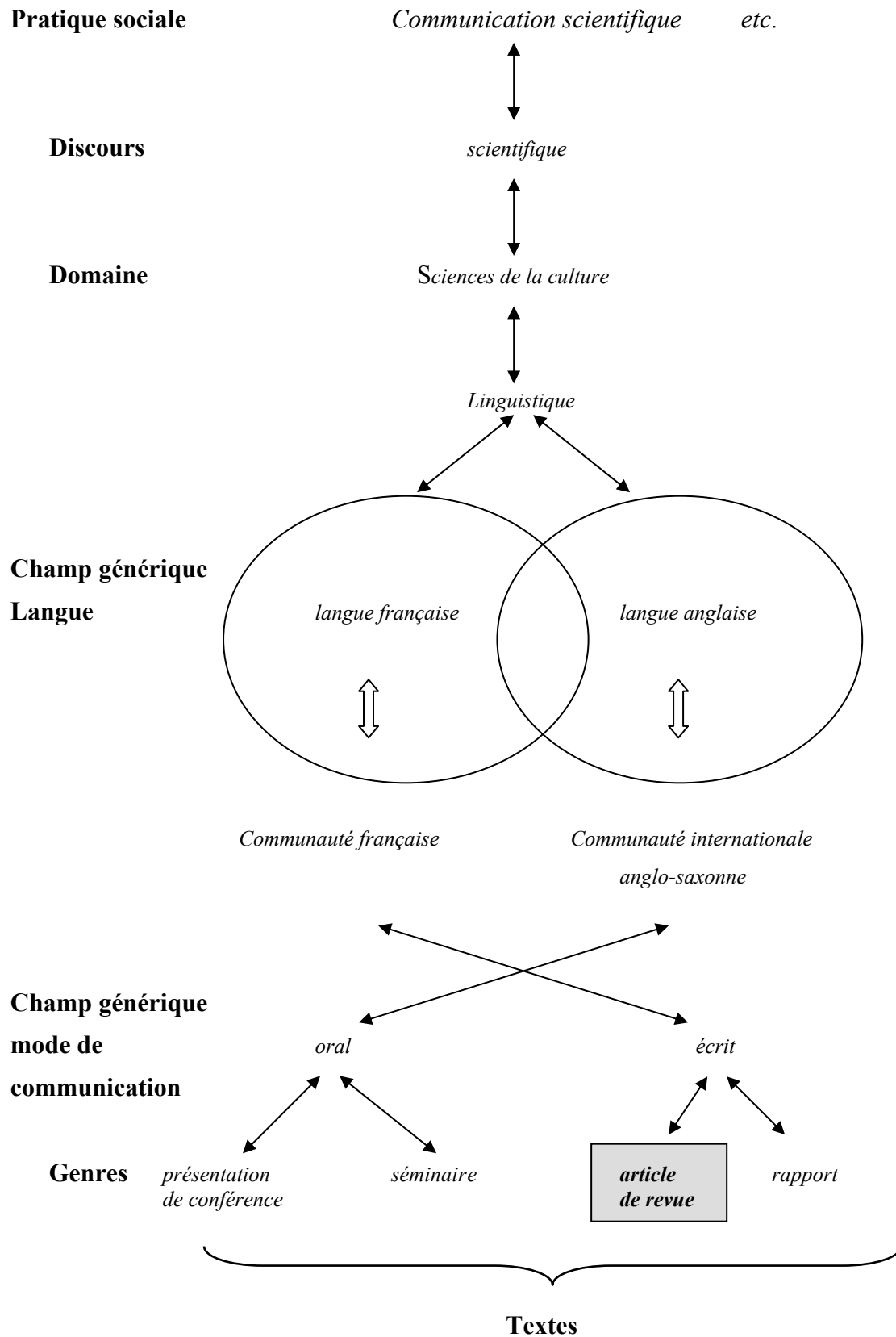


Figure : Typologie du discours scientifique linguistique

1.3. Conclusion

Dans ce chapitre, nous avons d'abord montré l'intérêt et les enjeux que soulèvent l'étude des genres en linguistique : objet empirique et attesté, le texte représente un niveau de description linguistique pertinent qui ne saurait être abordé sans prise en compte de la notion de genre, palier global déterminant les phénomènes locaux qui l'affectent. Le genre est ainsi une notion centrale, un espace normatif permettant avec les discours de réconcilier les linguistiques de la langue et de la parole.

Si les genres littéraires ont été largement explorés, l'étude des genres scientifiques, et celle du genre de l'article, est plus récente, et demeure encore essentiellement menée dans le cadre du courant ESP, qui autonomise les descripteurs linguistiques et empêche ainsi l'observation globale des genres.

Dans la mesure où les genres sont des objets fondamentalement multidimensionnels, ce qui explique leur appropriation par de nombreuses disciplines, ils ne sauraient être décrit à partir d'un niveau ou d'un critère linguistique unique, mais à partir de faisceaux d'observations. Le développement du traitement automatique des langues, des statistiques textuelles et de la linguistique de corpus permet aujourd'hui de saisir les genres et leurs régulations linguistiques, en corpus et au moyen de méthodes d'exploration quantitative.

Dans cette perspective, nous avons choisi d'explorer les lieux de stabilité du genre de l'article en partant d'un corpus homogène en genre et de quatre points d'entrée : le niveau morphosyntaxique, qui a montré son efficacité descriptive et discriminante dans diverses études de classification et de validation de typologies textuelles, et qui nous permettra d'obtenir une première approximation descriptive du genre et de ses axes d'organisation interne (chapitre 3), le niveau structurel des sections de l'article (chapitre 4), le niveau configuratif des composantes génériques à travers l'analyse des exemples des textes (chapitre 5) et le niveau lexico-thématique du genre (chapitre 6).

Le corpus sera ensuite mis à l'épreuve avec d'autres collections recoupant d'autres typologies possibles (auteur, domaine, genre et langue), ce qui nous permettra de le saisir de manière plus précise.

Le chapitre qui suit décrit les corpus mobilisés par notre démarche, ainsi que le jeu de descripteurs morphosyntaxiques qui servira d'entrée à l'élaboration du profil de l'article (chapitre 3).