Les routines de topicalisation dans les écrits scientifiques en français

Ji Yujing, Agnès Tutin

LIDILEM, Université Grenoble Alpes, Bâtiment Stendhal, 38058 Grenoble cedex 9 ;
yujing.ji@univ-grenoble-alpes.fr ; agnes.tutin@univ-grenoble-alpes.fr

Résumé : Le discours des écrits scientifiques est dialogique. L'auteur des écrits scientifiques présente non seulement des contenus informatifs, mais utilise aussi des éléments métadiscursifs pour évaluer les contenus et interagir avec ses lecteurs. Dans notre étude, nous nous focalisons sur l'aspect interlocutif des écrits scientifiques réalisé à travers des éléments métadiscursifs complexes, les routines de guidage du lecteur comme *L'objectif de cet article est d'étudier...* Plus précisément, nous examinons un petit ensemble des routines que nous nommons « routines de topicalisation » afin de comprendre comment l'auteur met en relief les idées ou les informations importantes pour ses lecteurs. À partir d'un corpus d'articles de recherches en sciences humaines et sociales, nous appliquons la méthode des « arbres lexico-syntaxiques récurrents » pour extraire des routines. Nous modélisons et analysons ensuite les fonctions rhétoriques et argumentatives des routines de topicalisation.

Mots-clés : écrit scientifique, dialogisme, métadiscours, routines, topicalisation.

1 Introduction

Comme dans de nombreuses formes de communication, l'écrit scientifique, en tant qu'écrit argumentatif, a une fonction dialogique (Bakhtine, 1981). L'auteur de l'écrit scientifique ne présente pas seulement des contenus informatifs, mais utilise aussi des éléments l'aidant à évaluer les contenus et à dialoguer avec ses lecteurs (Hyland, 2001). Ces éléments relèvent du « métadiscours », notion qui renvoie à la façon dont l'auteur interagit avec le lecteur à travers l'usage du langage (Hyland, 2005, p. 3). En observant ces éléments métadiscursifs, nous remarquons des expressions récurrentes qui servent à faciliter la lecture du texte que nous appelons les routines de guidage du lecteur. À titre d'exemple, *L'objectif de cet article est d'étudier...* est une routine qui sert à annoncer au lecteur l'objectif de la recherche dès le début de l'étude présentée. Parmi ces routines, nous nous focalisons dans cet article sur les routines de topicalisation dans l'écrit scientifique (par exemple : *il est important de souligner*), qui jouent un rôle de mise en relief.

Dans les études en linguistique, la notion de topicalisation est souvent appréhendée à travers les termes de « thème » ou de « topique » dans une perspective grammaticale ou lexicale pour évoquer un sujet dont il a déjà été question dans le contexte (Combettes, 2003a). De ce fait, les marqueurs de topicalisation étudiés sont souvent des locutions prépositionnelles comme *quant à* ou *en ce qui concerne* ou des structures détachées comme *c'est qui... est de...* (Combettes, 1998, 1999, 2003 b, 2007; Combettes & Prévost, 2001 ; Crismore, 1983 ; Prévost, 2003). Notre article propose de traiter cette notion dans une perspective plus rhétorique en nous appuyant sur certains travaux autour de la notion de métadiscours. Nous envisageons la topicalisation comme une fonction rhétorique qui vise à mettre en relief des informations essentielles du texte scientifique pour le lecteur, ce qui permet à l'auteur de l'écrit scientifique de mieux orienter la lecture.

L'article comporte trois parties principales. La première présentera le cadre théorique afin de préciser notre définition des routines de topicalisation dans l'optique rhétorique. La deuxième partie expose le choix des corpus et la méthode de recherche. La troisième partie présente une analyse à la fois syntaxique, sémantique et rhétorique des routines de topicalisation pour faire comprendre comment elles contribuent à la prise en charge du guidage du lecteur. Nous finirons par une réflexion sur les limites de l'analyse réalisée et proposerons quelques perspectives.

2 L'écrit scientifique : un discours dialogique

Le discours des écrits scientifiques est fondamentalement dialogique (Grossmann & Tutin, 2010). Élaborée par Bakhtine (1981), la notion de dialogisme « réfère aux relations que tout énoncé entretient avec les énoncés produits antérieurement ainsi qu'avec les énoncés à venir que pourraient produire ses destinataires » (Charaudeau & Maingueneau, 2002, p. 175). Elle met en avant le fait que tout discours se produit et se développe entre des sujets et que l'étude linguistique du discours doit intégrer les sujets participants du discours. Dans le cas des écrits scientifiques, ce seront l'auteur et le lecteur qui doivent être pris en compte.

2.1 Une perspective dialogique : la prise en compte du lecteur dans l'écrit scientifique

Deux formes de dialogisme interviennent principalement dans l'écrit scientifique — interdiscursive et interlocutive (Grossmann & Tutin, 2010) :

— La forme interdiscursive concerne l'interaction entre les énoncés de l'auteur et ceux de ses pairs précédemment formulés souvent sur un même objet. (Brès, 2005, 2017).

Dans l'écrit scientifique, l'auteur se constitue une identité scripturale scientifique fiable à partir des discours préexistants de ses pairs. Il peut, par exemple, synthétiser, citer, comparer, ou même remettre en question les propos des pairs afin de se frayer une route pour accéder à sa propre recherche.

— La forme interlocutive se rapporte avec « la réponse de l'allocutaire qu'il sollicite et sur laquelle il anticipe » (Brès, 2017, p. 3).

Published in : M. Mangeot et A. Tutin (dir.), *Lexique(s) et genre(s) textuel(s) : approches sur corpus*, Editions des archives contemporaines, ISBN : 9782813003454, doi : 10.17184/eac.9782813003454

Cette dimension repose sur une prémisse importante : la prise en compte du lecteur. Il est à noter que le discours est avant tout un acte de communication qui dépend non seulement du sujet parlant, mais aussi de l'interlocuteur. Comme le constate Benveniste (1974, p. 241), « Toute énonciation supposant un locuteur et un auditeur, et chez le premier l'intention d'influencer l'autre en quelque manière ».

Nous pouvons voir que dans l'optique interlocutive, l'auteur et le lecteur constituent deux paramètres centraux pour l'analyse linguistique et textuelle des écrits scientifiques. Le rapport tissé entre l'auteur et le lecteur forge un espace intersubjectif où les deux participants se rencontrent pour se comprendre. Pour l'écrit scientifique, malgré l'absence du lecteur lors de la rédaction, le lecteur se manifeste comme un agent très important qui exerce un impact sur les marques énonciatives, les choix lexicaux et les stratégies adoptées par l'auteur. Dans cette optique, l'auteur doit toujours avoir une bonne connaissance du public cible afin de pouvoir se faire d'une image représentative du lecteur type. Comme le souligne Todorov (1981, p. 70), « S'il n'y a pas d'interlocuteur réel, on le présuppose en la personne du représentant normal, pour ainsi dire, du groupe social auquel appartient le locuteur ». Pour l'auteur de l'écrit scientifique, l'image du lecteur sera construite à partir de l'image représentative de ses homologues dans son domaine scientifique. Cette image virtuelle du lecteur aide l'auteur à mieux anticiper les attentes du lecteur, parce qu'une bonne anticipation du lecteur peut rendre plus facile la communication et garantir davantage la réussite de l'objectif d'écriture (Hyland, 2005, p. 4).

2.2 La nécessité d'anticiper les attentes du lecteur de l'écrit scientifique : bien guider la lecture

Une bonne anticipation des attentes du lecteur peut faciliter la « réussite » d'un écrit. Les marques de guidage du lecteur constituent un paramètre essentiel pour atteindre ce but. D'une part, le guidage du lecteur peut assurer un bon suivi du développement logique des idées de l'auteur ; d'autre part, il facilite largement la compréhension des informations et des points de vue adoptés par l'auteur. Par exemple, l'auteur peut utiliser des marqueurs de reformulation pour expliciter des points éventuellement difficiles à comprendre pour ses lecteurs :

> [EX] La statistique descriptive a pour objet, comme son nom l'indique, de décrire des données. En d'autres termes, il s'agit de décrire la distribution d'une variable. (Corpus Scientext, Linguistique, thèse)

Les articles de recherche, qui constituent un sous-genre emblématique des écrits scientifiques (Bazerman, 1988), ont pour objectif de transmettre les découvertes dans un domaine scientifique donné. Il est essentiel que l'écrit soit facilement compréhensible pour le lecteur, afin d'obtenir la reconnaissance et la validation de la communauté scientifique (Swales, 1990). C'est pourquoi l'auteur expert a tendance à prêter une attention particulière à l'usage du langage, par exemple aux marqueurs discursifs ou métadiscursifs, dont les routines de topicalisation sont un type spécifique.

Published in : M. Mangeot et A. Tutin (dir.), *Lexique(s) et genre(s) textuel(s) : approches sur corpus*, Editions des archives contemporaines, ISBN : 9782813003454, doi : 10.17184/eac.9782813003454

2.3 Les routines de topicalisation : des expressions récurrentes qui ont une fonction de « mise en relief »

L'écrit scientifique utilise de nombreux moyens linguistiques pour assurer une bonne compréhension du texte, en particulier, les marqueurs linguistiques regroupés sous le terme de *métadiscours*. Le métadiscours constitue un moyen essentiel pour faciliter la communication, consolider les arguments, améliorer la lisibilité du texte et construire une relation intersubjective entre l'auteur et le lecteur (Hyland, 2005, p. 5). Par conséquent, il exerce une grande influence sur la perception du texte pour les destinataires d'un discours. Pour les écrits scientifiques, l'emploi des marqueurs métadiscursifs permet à l'auteur de s'exprimer avec efficacité et d'établir un lien avec leur lecteur.

2.4 Le métadiscours comme entrée pour l'étude des routines de topicalisation

Williams, le premier à utiliser le terme « métadiscours », l'a identifié dans une perspective didactique des écrits, comme l'ensemble des phénomènes qui aident à construire le discours. Il le définit de la façon suivante : « writing about writing, whatever does not refer to the subject matter being addressed » (Williams, 1981, p. 211-212). Par la suite, certains héritiers du modèle fonctionnaliste de Halliday (1973; 2014) ont développé le concept du métadiscours dans une perspective fonctionnelle. Vande Kopple (1985) et Crismore *et al.* (1993) envisagent le métadiscours comme le matériel linguistique qui n'ajoute pas d'informations propositionnelles, mais signale la présence de l'auteur. Hyland (2005, p. 37) l'envisage plutôt comme un outil ou un moyen que l'auteur adopte pour intégrer ses points de vue et attitudes, pour négocier les sens interactionnels dans un texte, et pour engager les lecteurs en tant que membres d'une communauté particulière. La nuance entre ces deux définitions donne lieu à deux catégorisations lexicales du métadiscours. Vande Kopple (1985) et Crismore et *al.* (1993) ont proposé une typologie en deux grandes catégories :

1. une catégorie textuelle qui désigne des éléments principaux liant de différents niveaux du texte, comportant des marqueurs textuels comme des connecteurs, des marqueurs de rappel, etc. et
2. la catégorie interpersonnelle qui comporte des éléments du lexique permettant de refléter la relation auteur-lecteur, comme les marqueurs d'attitude *I wish that*, les marqueurs de commentaire *you may not agree that*, etc.

De son côté, Hyland refuse de séparer les deux dimensions textuelle et interpersonnelle du métadiscours. Il estime que tout métadiscours doit être de nature interpersonnelle et propose donc une catégorisation en niveaux interactionnel et interactif. La dimension interactionnelle implique davantage les moyens que l'auteur utilise pour commenter le contenu, exprimer ses attitudes ou ses sentiments comme des marqueurs d'engagement *you can see that* (Hyland, 2005, p. 49) classés par ses prédécesseurs dans la classe interpersonnelle. La dimension interactive reflète la prise en compte du lecteur chez l'auteur, les moyens linguistiques que l'auteur utilise pour anticiper les attentes du lecteur, dans laquelle se classent des marqueurs semblables à la catégorie textuelle citée plus haut, comme ceux de transition, de structuration, etc.

Published in : M. Mangeot et A. Tutin (dir.), *Lexique(s) et genre(s) textuel(s) : approches sur corpus*, Editions des archives contemporaines, ISBN : 9782813003454, doi : 10.17184/eac.9782813003454

Notre étude adopte la conception de Hyland, car le dialogisme interlocutif de l'écrit scientifique nous permet d'envisager l'écrit scientifique comme interpersonnel, ce qui correspond à la compréhension de Hyland sur le métadiscours. Cependant, nous ne nous limitons pas à l'observation de la typologie du métadiscours de Hyland dans notre étude, car les autres typologies nous fournissent également des données linguistiques pertinentes pour accéder à la question des routines de topicalisation.

2.4.1 Les routines de guidage du lecteur dans les écrits scientifiques en français

Les routines, objet d'étude émergent du domaine phraséologique (Legallois & Tutin, 2013, p. 3), proposent une entrée intéressante et riche pour comprendre la réalisation du guidage du lecteur. Pour nous, un candidat au statut de routine doit répondre parallèlement à deux critères :

— du point de vue linguistique, il est une configuration lexico-syntaxique complexe, mais récurrente située à la frontière entre le lexique et la phrase (Kraif & Tutin, 2017) ;
— du point de vue rhétorique, il remplit des fonctions rhétoriques et argumentatives dans les écrits scientifiques (*Ibid).*

Il faut préciser qu'à première vue, les routines ressemblent aux locutions figées comme *en ce qui concerne*, mais elles présentent quelques caractéristiques différentes de ces dernières (Kraif & Tutin, 2017; Tutin & Kraif, 2016). Au lieu d'avoir une structure linéaire figée, les routines correspondent à des configurations lexico-syntaxiques dont les éléments peuvent être des paradigmes lexicaux (par exemple, *comme on/nous l'a/avons vu/observé/constaté*). Elles correspondent également généralement à des phrases entières, dont certains constituants sont actualisés dans le contexte (et renvoient à l'auteur, au texte, à la problématique). Elles ont enfin une fonction rhétorique spécifique au genre (*ibid.*). Les routines de guidage du lecteur sont ces configurations lexico-syntaxiques complexes, regroupant diverses fonctions rhétorique et argumentative pour faciliter la lecture et de guider l'interprétation de l'écrit scientifique. Par exemple, *l'objectif de cet article est d'étudier...* est une routine qui aide à annoncer au lecteur l'objectif de la recherche. Les routines de topicalisation constituent un sous-ensemble de routines de guidage du lecteur dont nous souhaitons observer le fonctionnement.

2.5 Les routines de topicalisation : définition et fonctions rhétoriques

Dans la tradition française, le concept de topicalisation est plutôt abordé dans une perspective grammaticale. Les recherches les plus détaillées et complètes sur ce sujet sont les études de Combettes (1999, 2003 a, 2003 b, 2006, 2007), de Grobet (2002) et de Prévost (2003). Pour eux, les marqueurs de topicalisation sont liés au thème ou topique pour évoquer le sujet d'un propos. Ils peuvent être des dispositifs grammaticaux comme des éléments déictiques ou anaphoriques, ou des éléments lexicaux ou phraséologiques comme des locutions prépositionnelles *quant à, à propos de,* des syntagmes *en ce que concerne, en ce qui est de*, ou des structures détachées *ce qui... est de, c'est... qui/que...* (Combettes & Prévost, 2001 ; Grobet, 2002 ; Prévost, 2003).

Published in : M. Mangeot et A. Tutin (dir.), *Lexique(s) et genre(s) textuel(s) : approches sur corpus*,
Editions des archives contemporaines, ISBN : 9782813003454, doi : 10.17184/eac.9782813003454

Dans les recherches sur le métadiscours en anglais, Meyer (1975, cité par Crismore, 1983, p. 9-10) identifie des marqueurs de pointage (« *Pointer words* ») qui ont pour fonction d'informer explicitement le lecteur sur une idée particulière dans la perspective de l'auteur, par exemple, *My first important point is that ERA is not a dead issue.* Ces marqueurs peuvent être employés afin d'attirer l'attention sur l'importance d'une idée ou sur l'opinion d'un auteur. Williams (1981) range quant à lui les marqueurs de topicalisation (« *Topicalizers* ») en tant que sous-catégorie du métadiscours. Ces marqueurs attirent l'attention sur un sujet particulier d'une phrase ou d'un paragraphe ou d'une section entière (Crismore, 1983, p. 8). Les *topicalizers* attirent non seulement l'attention sur le thème, mais aussi sur le changement de thème, et regroupent donc des expressions comme *in regard to, in the matter of, turning now to*, ou des phrases clivées comme *there are three reasons why we...* Inspiré par cette description, Vande Kopple (1985) reprend les éléments de cette sous-catégorie dans la classe des « Connecteurs » de sa propre typologie, en les définissant en tant qu'éléments métadiscursifs qui aident à prêter attention sur le thème d'un segment textuel, avec des marqueurs comme *with regard to/en ce qui concerne.* Crismore *et al.* (1993) les envisagent aussi comme des connecteurs qui ont pour fonction de signaler le changement du sujet : *now I will discuss / maintenant, je vais discuter.*

Le concept de topicalisation apparaît donc sous plusieurs définitions et présente un certain flou. Dans cette étude, nous proposons une définition plus restreinte de la topicalisation qui se rapproche davantage du sens de la mise en relief d'une idée ou d'un élément dans un texte scientifique, et excluons le sens du changement de thème. Au niveau formel, nous ne nous intéressons qu'aux expressions récurrentes plus larges sur le plan syntaxique dans la perspective d'une phraséologie scientifique transdisciplinaire, sans prendre en compte des marqueurs de nature grammaticale comme des éléments déictiques ou anaphoriques, ni les phrases clivées non plus.

Nous donnons ici une définition aux routines de topicalisation : ce sont des énoncés typiques et récurrents, ayant une construction lexicale et syntaxique complexe, qui aident l'auteur à mettre en relief une idée ou un élément qu'il juge important ou essentiel. Avec la mise en place des routines de topicalisation, l'auteur peut non seulement assurer un bon guidage du lecteur afin d'exercer une bonne influence sur l'interprétation du texte pour le lecteur, mais aussi valoriser les résultats de ses travaux.

3 Méthodologie et analyse des routines

3.1 Méthodologie — Corpus Scientext SHS et méthode d'extraction

Pour observer les routines de topicalisation dans les textes, nous utilisons les méthodes de la linguistique de corpus. Nous effectuons une extraction de configurations lexico-syntaxiques, à partir d'un corpus diversifié de 500 articles de sciences humaines (Hatier 2016). La méthode d'extraction utilisée utilise la notion d'arbre lexico-syntaxique récurrent (désormais ALR) que l'on peut considérer comme un observable « textométrique » extrait à partir des relations syntaxiques et des mesures d'association (Kraif 2016 ; Tutin & Kraif 2016 ; Kraif & Tutin 2017). Cette méthode permet d'extraire des arbres de dépendance correspondant à des ensembles d'expressions récurrentes.

Published in : M. Mangeot et A. Tutin (dir.), *Lexique(s) et genre(s) textuel(s) : approches sur corpus*, Editions des archives contemporaines, ISBN : 9782813003454, doi : 10.17184/eac.9782813003454

Pour élargir les résultats, nous effectuons cette extraction des ALR à partir d'un ensemble de classes sémantiques, définies dans le cadre du projet TermITH (Hatier 2016 ; Jacques & Tutin 2018), qui paraissent pertinentes pour la question de la topicalisation. À cette fin, 15 classes sémantiques (6 classes de verbes, 3 classes d'adjectifs, 3 classes de noms et 3 classes d'adverbes) sont employées. Par exemple, nous exploitons la classe sémantique des « verbes d'analyse » (*constater, noter, remarquer...*) ou des « adjectifs d'importance » (*central, crucial, déterminant, essentiel...*). Les classes sémantiques utilisées sont présentées dans le tableau suivant :

Tableau 1 : 15 classes sémantiques comme les pivots de recherche

Classes sémantiques	**Mots classés**
Verbes d'analyse=VANALYSE	*constater, noter, remarquer, signaler, observer, relever, repérer, préciser, spécifier, montrer.*
Verbes de communications=VCOMM	*insister, souligner, citer, dire, indiquer, mentionner.*
Verbes de processus non humain = VPROCES	*affirmer, indiquer, marquer, révéler, signaler, spécifier.*
Verbes d'ancrage =VANCRAGE	*centrer, focaliser, privilégier.*
Verbes de modalité=VMOD	*exiger, importer, nécessiter, révéler, falloir, devoir.*
Verbes de l'état et de relation=VETATRELAT	*imposer, intéresser.*
Adjectifs d'importance positive =ADJIMPORT	*central, crucial, déterminant, essentiel, important, majeur, prépondérant, primordial, principal, significatif, fondamental, fatal.*
Adjectifs de modalité=ADJMOD	*certain, évident, incontournable, indispensable, nécessaire, obligatoire, imposé, inévitable, immanquable, impératif.*
Adjectifs d'évaluation axiologique=ADJEVALUAXIO	*décisif, intéressant, remarquable, notable.*
Noms d'état =NETAT	*importance, poids, portée, priorité, nécessité, exigence, relief, besoin, évidence.*
Noms de processus=NPROCES	*affirmation, intérêt, attention.*
Noms d'objet=NOBJ	*marque, question, problème, force.*
Adverbes d'intensité forte=ADVINTEN	*particulièrement, fondamentalement, significativement, singulièrement.*
Adverbes de manière et de modalité=ADVMANMOD	*essentiellement, nécessairement, principalement.*
Adverbes discursifs=ADVDIS :	*notamment, spécialement, surtout.*

Cette technique des ALR associée aux classes sémantiques permet ainsi de repérer des configurations comme celle qui apparaît dans la figure 1 et qui correspondent aux phrases *il est vraiment important de noter, il crucial ici de remarquer, etc.* comme on le verra plus bas.

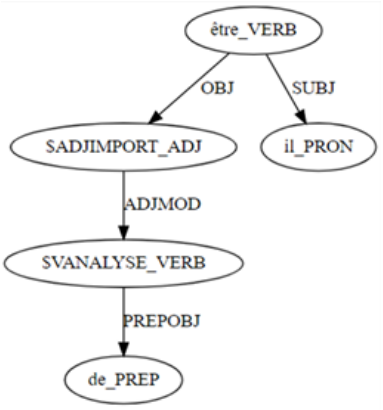


Figure 1: Un exemple d'ALR extrait avec des classes sémantiques

Published in : M. Mangeot et A. Tutin (dir.), *Lexique(s) et genre(s) textuel(s) : approches sur corpus*, Editions des archives contemporaines, ISBN : 9782813003454, doi : 10.17184/eac.9782813003454

Un ensemble d'ALR a été extrait et nous avons sélectionné les données qui apparaissent pertinentes pour notre étude.

3.2 Analyses des routines de topicalisation

L'extraction des routines a fourni des données riches, à partir desquelles nous avons sélectionné quelques routines liées à notre thème. Nous observons maintenant leurs caractéristiques lexicales, syntaxiques, sémantiques et rhétoriques. Après analyse des ALR extraits que nous regroupons, nous obtenons cinq types de routines de topicalisation. Rappelons que notre étude n'est que préliminaire et que ces classes peuvent donc être élargies après une extraction plus détaillée. Le tableau 1 résume les principales routines.

Tableau 2 : classement des routines après leurs modélisations

	Les routines de topicalisation	Exemples
1	*il est* ADJIMPORT *de* VCOMM	*Il est important de souligner/dire que…*
2	*il est* ADJIMPORT *de* VANALYSE	*Il est important de noter/remarquer que…* *Il est fondamental de noter que…*
3	*il* VMOD-déontique VANALYSE/ VCOMM	*Il nous faut préciser que…* *Enfin, il faut remarquer/préciser/noter/signaler que…* *Il importe enfin de remarquer que.* *Il faut souligner ici que…* *Et pourtant il faut le dire…* *Il faut le souligner…*
4	Pronom_auteur + (*vouloir/pouvoir*) + VCOMM/VANALYSE	*Je voudrais néanmoins souligner…* *Nous pouvons remarquer que*
5	*l'enjeu/le problème* ADJIMPORT *est…*	*Le principal problème est…* *L'enjeu principal est de…*

La colonne « les routines de topicalisation » regroupe six types de routines. Nous classons dans un type de routine les verbes appartenant à un même paradigme sémantique et à une structure syntaxique identique. Prenons l'exemple de la routine *il est* ADJIMPORT *de* VANALYSE comme illustration : la classe sémantique ADJIMPORT regroupe des adjectifs d'importance positive, dans lesquels se trouvent *central*, *crucial*, *déterminant*, *essentiel*, *important*, *majeur*, *prépondérant*, *primordial*, *principal*, *significatif*, *fondamental*, *fatal*. La classe VANALYSE rassemble des verbes d'analyse : *constater*, *noter*, *remarquer*, *signaler*, *observer*, *relever*, *repérer, préciser, spécifier, montrer*. Ces deux classes permettent de formuler des routines comme *il est important de noter*, *il est essentiel de remarquer*, etc., illustrées dans la colonne « exemples ».

Nous invitons à observer ces routines sous trois aspects :

— la construction syntaxique avec les arguments, les compléments ou les marqueurs discursifs
— la structure sémantique
— les fonctions rhétorique et argumentative

La structure **il est ADJIMPORT *de* VANALYSE/VCOMM + complétive *que*/argument nominal** regroupe deux classes de routines construites respectivement à l'aide des verbes d'analyse et des verbes de communication. Les adjectifs d'importance comme *important*, *essentiel*, *fondamental*, etc. dans ces routines sont

Published in : M. Mangeot et A. Tutin (dir.), *Lexique(s) et genre(s) textuel(s) : approches sur corpus*, Editions des archives contemporaines, ISBN : 9782813003454, doi : 10.17184/eac.9782813003454

primordiaux, car c'est avec eux que le sens de la mise en relief peut être véhiculé et renforcé.

1. il est ADIIMPORT DE VCOMM :

Cette routine intègre des verbes de communication, plus précisément, des verbes d'insistance (*souligner, insister*) et des verbes de diffusion (*citer*, *dire*, *indiquer*, *mentionner*). Elles permettent de rendre manifeste l'importance d'une idée ou d'une information parmi des idées ou des informations transmises. Nous recensons trois constructions syntaxiques :

a) Routine +argument nominal :

(1) Enfin, **il est important de souligner le poids** que pèse, dans le processus adoptif, la germanité prise, ici, dans toutes ses expressions de sexe, d'âge, de biologique et de social. (anthropologie.xml-s5332).

Cette structure est fréquente parmi les occurrences. Elle vise à mettre en évidence la valeur d'importance des idées ou des informations. Dans l'exemple [1], la routine *il est important de souligner* indique un élément important dans le processus adoptif parmi plusieurs éléments caractérisant la parenté : l'influence de la germanité exerce aussi une influence importante dans les expressions de sexes, d'âge, etc. En conséquence, avec cette routine, l'auteur peut sélectionner par souci de clarté une idée fondamentale pour ses lecteurs.

b) Routine + complétive que

(2) Finalement, **il est important de souligner que** le nombre de participants à cette étude n'était pas très conséquent (n=181). (psycho.xml-s222).

Cette construction syntaxique est également fréquente dans notre corpus. Le fait qu'elle soit suivie d'une phrase complétive facilite la construction d'une idée ou information plus complexe. Dans l'exemple [2], l'auteur met en relief l'information que le nombre de participants ne constitue pas un facteur décisif pour influencer le résultat de son étude, cette précision dissipe le doute éventuellement généré chez le lecteur, ce qui rend la lecture plus claire.

c) Routine avec des connecteurs ou des marqueurs discursifs

(3) **Néanmoins, il est important de souligner que** cette validation a été réalisée avec des lycéens âgés de 12 à 18 ans. (psycho.xml-s7165).

(4) Ce que nous avons pu montrer à travers l'étude du texte de Jolicoeur repose sur une situation **qu'il est important de souligner, en conclusion**, pour indiquer en quoi ce texte garde une résonance moderne. (histoire.xml-s3126).

Published in : M. Mangeot et A. Tutin (dir.), *Lexique(s) et genre(s) textuel(s) : approches sur corpus*, Editions des archives contemporaines, ISBN : 9782813003454, doi : 10.17184/eac.9782813003454

Il faut signaler que, les routines de topicalisation apparaissent souvent avec des marqueurs discursifs, éléments indispensables pour la cohésion et la cohérence du texte (Tran, 2014, p. 28). La présence des marqueurs discursifs permet par conséquent à une routine de topicalisation de remplir certaines fonctions rhétoriques et argumentatives au-delà de la fonction de mise en relief. Dans l'exemple [3], l'auteur utilise *néanmoins, il est important de souligner que* pour introduire une nouvelle information qui mérite aussi une attention dans le texte. Cette information constitue à la fois un argument complémentaire (compléter l'information de l'âge) et un contre-argument (exclure la généralisation de cette validation pour les autres tranches d'âge). *Enfin* et *finalement* dans l'exemple [1] et l'exemple [2] servent à introduire un élément comme le dernier élément de l'argumentation, et signale en même temps le statut important de cet élément, tandis que *en conclusion* dans l'exemple [4] ajoute une fonction de récapitulation à la routine de topicalisation.

2. il est ADJIMPORT de VANALYSE :

Les verbes d'analyse sont des verbes qui visent à aider l'auteur de l'écrit scientifique à « présenter les démarches scientifiques et le processus de construction des connaissances dans la discipline » (Yan, 2017, p. 126). Les verbes d'analyse retenus dans la routine sont plutôt des verbes de constat comme *noter*, *remarquer*, qui constituent une catégorie sémantique spécifique à l'intérieur de la classe des verbes d'analyse. Ces verbes montrent que, en dehors de la présentation des démarches ou des processus, l'auteur de l'écrit scientifique doit aussi « mettre en scène une opération intellectuelle d'analyse de données plus que de simple perception ou d'activité graphique. » (Grossmann, 2014, p. 760). Ce type de routines a souvent pour fonction de mettre en évidence une information ou un constat après des réflexions intellectuelles sur des activités scientifiques. De plus, il y a ici une opération de co-constat exprimée par l'auteur : celui-ci prend en compte la présence de ses lecteurs en invitant le lecteur à faire un co-constat avec lui, ce qui montre son intention d'intégrer son lecteur (exemple [6]). Ce modèle de routines a les mêmes constructions syntaxiques que les routines de modèle *il est ADJIMPORT de VCOMM*. Le connecteur *Toutefois* (exemple [8]) fournit au constat mis en avant par la routine *il est fondamental de noter que* un sens d'opposition : malgré les considérations sur des observations ou des circonstances, l'auteur décide finalement de faire un constat qu'il juge important après la réflexion intellectuelle et invite au lecteur de le partager.

(5) **Il est important de noter** la forte insistance – nous y reviendrons – mise par les parents adoptifs à intégrer les trois sœurs dans leur propre univers affectif et social, sans s'immiscer dans le leur. (antrop. xml-s5460)

(6) **Il est important de remarquer que** nous comparons une solution de premier rang avec une situation où toutes les décisions sont décentralisées dans un processus de marché. (eco.xml-s20251)

(7) **Il est important de noter ici que** cette hypothèse n'est pas restrictive dans la mesure où la prise en compte d'un tel coût fixe strictement

Published in : M. Mangeot et A. Tutin (dir.), *Lexique(s) et genre(s) textuel(s) : approches sur corpus*, Editions des archives contemporaines, ISBN : 9782813003454, doi : 10.17184/eac.9782813003454

positif consisterait, dans cette analyse, juste à diminuer le profit de cet opérateur d'une constante. (eco.xml-s20335).

(8) **Toutefois, il est fondamental de noter que** la partie fractale existe dans la partie non fractale, mais elle ne « s'exprime » pas ; de même, la partie non fractale existe dans la partie fractale, mais elle n'est pas apparente d'où les pointillés sur la figure 1. (geo.xml-s6303).

Ces deux modèles de routines présentent une grande similitude au niveau syntaxique, mais leurs sens et leurs fonctions rhétoriques diffèrent. La routine *il est* ADJIMPORT *de* VANALYSE prend souvent une idée ou une information comme complément, alors que la routine il est ADJIMPORT de VCOMM prend un fait ou un constat en tant que complément.

3. il + VMOD-déontique + VANALYSE/VCOMM

Les deux routines **il + VMOD-déontique + VANALYSE/VCOMM**, en mobilisant les verbes déontiques *falloir* ou *importer* avec les verbes d'analyse et de communication, ont principalement quatre constructions syntaxiques :

a) Routine + arguments nominaux

(9) Concernant les régressions multiples, **il faut remarquer** les variances relativement faibles. (psycho.xml-s16318).

Cette construction syntaxique avec les verbes d'analyse vise à mettre en évidence la valeur importante d'une entité scientifique. Dans notre corpus, elle apparaît souvent avec un connecteur comme *enfin* pour structurer les arguments. Par exemple :

(10) **Enfin, il faut signaler** le cas de l'expression dans les + quantifieur. (ling.xml-s11179).

L'auteur utilise *Enfin, il faut signaler* pour introduire un nouvel élément comme le dernier élément dans une énumération, et signale en même temps son importance.

b) Routine + complétive que

(11) **Il faut noter que** la structure commune ne comprend alors que deux cycles. (sciencespolitiques.xml-s12225).

Avec la phrase complétive introduite par *que,* cette structure indique plutôt la valeur d'un fait ou d'un constat. Dans l'exemple [11], c'est le constat que la structure commune ne comprend que deux cycles qui sont mis en avant.

c) Routine en structure clivée

(12) D'autre part, aux racines de l'art moderne, **c'est** bien une incessante manipulation sur le national et l'international **qu'il faut remarquer**, rendant impossible toute approche de l'histoire de l'art en termes d'art national comme international. (histoire.xml-s2599).

Published in : M. Mangeot et A. Tutin (dir.), *Lexique(s) et genre(s) textuel(s) : approches sur corpus*, Editions des archives contemporaines, ISBN : 9782813003454, doi : 10.17184/eac.9782813003454

En français, le clivage d'une phrase est un moyen syntaxique souvent utilisé pour accentuer les informations ou les éléments essentiels. Cet élément extrait est nommé « focus » ou « foyer », traité ensuite par un processus de focalisation (Riegel, Pellat, & Rioul, 2018, p. 726). La focalisation de l'élément extrait permet de spécifier cet élément et laisse comprendre que cet élément constitue la seule information nouvelle introduite par le clivage. En conséquence, quand cette routine se forme en structure clivée, elle va renforcer l'important de l'élément extrait.

d) Routine avec des connecteurs ou d'autres marqueurs discursifs

(13) **Mais il faut remarquer d'emblée que** les variables de contrôle introduites dans le modèle ne sont pas postulées dans cette théorie et ne sont par conséquent pas justiciables d'une interprétation unique et cohérente dans ce cadre. (scienceseducation.xml-s8381)

Parfois, on introduit *mais* et *cependant* dans la routine pour introduire un élément différent qui mérite aussi une attention dans le texte. En dehors de ces routines de structure impersonnelle, nous recensons aussi des routines faisant appel aux pronoms personnels.

4. Pronom + (vouloir/pouvoir) + VCOMM/VANALYSE + complétif que/ argument nominal

Ce modèle recourt particulièrement aux verbes de communication et d'insistance comme *souligner, insister* pour manifester l'accent mis par l'auteur sur les informations, les éléments, les entités scientifiques.

(14) Avant de les évoquer brièvement, **je voudrais souligner que** l'immense majorité des emplois prépositionnels manifestent les propriétés évoquées en 1. (ling.xml-s11223).

(15) Au-delà du caractère anecdotique de cette agression, **nous voudrions souligner** les injonctions contradictoires auxquels peuvent être confrontés les agents de sécurité. (sociologie.xml-s1049).

(16) Par substance sémantique, **nous voulons dire que** nous ne formons que des représentations digitales ou propositionnelles de la scène visuelle. (psyco.xml-s6625).

(17) **On peut remarquer que** le taux de taxation total est identique dans les deux pays. (eco.xml-s16783).

(18) À partir des graphiques 1 et 2, **nous pouvons remarquer que** le niveau d'indemnisation de départ joue de manière différente dans les deux réglementations. (eco.xml-s9630).

Ce qui nous intéresse ici, c'est l'utilisation des pronoms personnels renvoyant à l'auteur dans la routine. Dans les écrits scientifiques, *je* renvoie seulement à l'auteur tandis que *nous* et *on* peuvent renvoyer à plusieurs référents. Ils peuvent être inclusifs en incluant l'auteur et le lecteur, mais aussi exclusifs qui revoient seulement à l'auteur.

Published in : M. Mangeot et A. Tutin (dir.), *Lexique(s) et genre(s) textuel(s) : approches sur corpus*, Editions des archives contemporaines, ISBN : 9782813003454, doi : 10.17184/eac.9782813003454

Dans nos exemples, *nous* dans les routines constituées par verbes de communication renvoie plutôt à l'auteur lui-même ((14), (15), (16)). Cela correspond à la tentative de sélectionner les informations importantes de l'auteur pour ses lecteurs. Dans les routines à verbes d'analyse et de constat, *on* ou *nous* sont inclusifs. L'auteur invite le lecteur de faire un co-constat sur un point essentiel avec lui. L'ajout parfois des verbes de modalité comme *vouloir* ou *pouvoir* aide à nuancer les formulations de l'auteur : *vouloir* renforce l'intention de la mise en relief d'un élément ((14), (15), (16)) et *pouvoir* adoucit l'affirmation ((17), (18)).

Sur le plan syntaxique, les routines apparaissent sous trois formes :

a) routine +argument nominal : exemple (15)

Cette structure attire l'attention souvent sur une entité scientifique. La routine *nous voudrions souligner* fait ressortir une entité scientifique qui est un autre caractère de l'agression — les injonctions contradictoires.

b) routine + complétive que : exemples (14), (16), (17), (18)

La subordonnée introduite par *que* renvoie à une information importante.

c) routine des connecteurs ou d'autres marqueurs discursifs logiques

(19) **On peut cependant remarquer** sur le tableau 6 a que lorsqu'on compare les mêmes points de vue issus des deux situations, on n'obtient alors pas de différence significative. (psycho.xml-s12233).

À l'instar des routines modélisant en *il est* ADJIMPORT *de* VANALYSE/VCOMM, les routines *Pronom + (vouloir/pouvoir) +* VCOMM/VANALYSE peuvent également apparaître avec d'autres marqueurs afin de participer à l'argumentation du texte. Le marqueur *aussi* fournit une autre interprétation d'un fait pour faciliter la compréhension du lecteur. La locution *par exemple* donne un autre fait essentiel par une opération d'exemplification.

5. L'enjeu/le problème ADJIMORT est (de)

Le dernier modèle de routines mobilise des noms scientifiques comme *enjeu*, *problème*, *question*, etc. qui sont des « shells nouns » largement utilisés dans les écrits scientifiques, ces noms abstraits qui servent à synthétiser ou présenter les informations ou des concepts (Schmid, 2012).

(20) **L'enjeu fondamental est donc** de savoir si la question des identités collectives, en particulier culturelles, doit être renvoyée à la seule sphère privée, comme ce fut le cas pour la question religieuse au XVII e siècle, ou si au contraire la notion de communauté politique doit être repensée de manière à relativiser la valeur des droits universels en fonction du pluralisme culturel. (sociologie.xml-S872).

Published in : M. Mangeot et A. Tutin (dir.), *Lexique(s) et genre(s) textuel(s) : approches sur corpus*,
Editions des archives contemporaines, ISBN : 9782813003454, doi : 10.17184/eac.9782813003454

(21) **Le problème essentiel est** sans doute celui des prépositions souvent dites « vides » ou, du moins, dont certains emplois peuvent être dits « vides ». (ling.xml-s7509).

Avec cette routine, l'auteur fait aussi un constat pour indiquer le problème central au lecteur qui pourra ainsi comprendre la complexité et la difficulté de la recherche.

En résumé :

1. Au niveau syntaxique, les routines s'actualisent avec plusieurs structures syntaxiques ;
2. Au niveau sémantique, leur sens est compositionnel. Les routines admettent au moins une classe lexicale d'un paradigme sémantique identique ;
3. Au niveau fonctionnel, en dehors la fonction de mise en relief des éléments importants, les routines de topicalisation semblent remplir aussi d'autres fonctions rhétoriques comme l'exemplification, la récapitulation, etc. quand elles travaillent avec d'autres marqueurs.

4 Conclusion

Pour conclure, nous ferons trois remarques. Premièrement, l'écrit scientifique en français présente une abondance de routines productives qui permettent la mise en relief des faits, des informations, ou des entités scientifiques. Deuxièmement, la méthode des Arbres lexico-syntaxiques récurrents, couplée avec des classes lexicales sémantiques, peut produire des résultats prometteurs qui restent à évaluer à plus grande échelle. Troisièmement, les routines de topicalisation s'actualisent sous la forme de plusieurs structures syntaxiques en fonction de contextes précis. Le fait qu'elles apparaissent souvent en cooccurrence avec des marqueurs discursifs de nature logique est aussi à prendre en compte. Nous faisons l'hypothèse que ces cooccurrences ont des rôles argumentatifs spécifiques. Notre étude n'est ici qu'un travail préliminaire, qui présente bien évidemment des limites : il faut affiner les classes et créer de nouvelles classes pour une meilleure modélisation, par exemple, pour certains mots qui ne sont pas inclus dans la liste actuelle du lexique scientifique transdisciplinaire (LST). Nous remarquons ainsi que certaines expressions ne sont pas extraites, car il y a quelques erreurs dans l'analyse syntaxique. À titre d'exemple, *force est de constater* ou *il est à noter/souligner* ne sont pas présentes dans nos résultats. Enfin, à plus long terme, nous envisageons un prolongement didactique à notre étude en mettant en place des outils informatiques qui permettront d'illustrer dans les corpus les différentes routines de guidage du lecteur dans les corpus pour les apprenants du français académique.

Bibliographie

Bakhtin, M. M., & Holquist, M. (1981). *The dialogic imagination : four essays*. Austin : University of Texas Press.

Brès, J. (2005). 3. Savoir de quoi on parle : dialogue, dialogal, dialogique ; dialogisme, polyphonie... In *Dialogisme et polyphonie* (p. 47-61). De Boeck Supérieur. Consulté à l'adresse http://www.cairn.info/dialogisme-et-polyphonie-approches-linguistiques--9782801113646-p-47.htm

Published in : M. Mangeot et A. Tutin (dir.), *Lexique(s) et genre(s) textuel(s) : approches sur corpus*, Editions des archives contemporaines, ISBN : 9782813003454, doi : 10.17184/eac.9782813003454

Brès, J. (2017). Dialogisme, éléments pour l'analyse. *Recherches en didactique des langues et des cultures. Les cahiers de l'Acedle*, *14*(14-2). https ://doi.org/10.4000/rdlc.1842

Charaudeau, P., & Maingueneau, D. (2002). *Dictionnaire d'analyse du discours*. Paris : Éd. du Seuil.

Combettes, B. (1998). Thématisation, topicalisation, et éléments non-référentiels : le cas de l'adjectif détaché. *Cahiers de praxématique* (30), 133-159.

Combettes, B. (1999). Les marqueurs de topicalisation dans l'« Histoire d'un voyage en terre de Brésil » de Jean de Léry, 25-40.

Combettes, B. (2003a). Le topique comme constituant périphérique : aspects diachroniques, Abstract. *Travaux de linguistique*, *no47*(2), 137-161. https ://doi.org/10.3917/tl.047.0137

Combettes, B. (2003 b). Les marqueurs de topicalisation en français : un cas de grammaticalisation. In *Ordre et distinction dans la langue et le discours* (p. 149-167). Honoré Champion. Consulté à l'adresse https ://halshs.archives-ouvertes.fr/halshs-00149703

Combettes, B. (2006). Du niveau textuel au niveau énonciatif dans la grammaticalisation : le rôle du contexte, Abstract. *Langue française* (149), 48-60. https ://doi.org/10.3917/lf.149.0048

Combettes, B. (2007). Grammaticalisation des marqueurs de topicalisation en français : les expressions du type pour ce qui regarde. *Langue française* (156), 93-107. https ://doi.org/10.3917/lf.156.0093

Combettes, B., & Prévost, S. (2001). Évolution des marqueurs de topicalisation. *Cahiers de praxématique* (37), 103-124.

Crismore, A. (1983). *Metadiscourse : What it is and how it is Used in School and Non-school Social Science Texts*. University of Illinois at Urbana-Champaign.

Crismore, A., Markkanen, R., & Steffensen, M. S. (1993). Metadiscourse in Persuasive Writing : A Study of Texts Written by American and Finnish University Students. *Written Communication*, *10*(1), 39-71. https ://doi.org/10.1177/0741088393010001002

Fløttum, K., & Vold, E. T. (2010). L'éthos auto-attribué d'auteurs-doctorants dans le discours scientifique. *Lidil. Revue de linguistique et de didactique des langues* (41), 41-58.

Grobet, A. (2002). *L'identification des topiques dans les dialogues*. De Boeck Supérieur. https://doi.org/10.3917/dbu.grobe.2002.01

Grossmann, F. (2014). De quelques routines phraséologiques liées aux verbes parenthétiques dans les genres scientifiques. In *Congrès Mondial de Linguistique Français* (vol. 8, p. 759-770). Berlin, Germany. https ://doi.org/10.1051/shsconf/20140801306

Grossmann, F., & Tutin, A. (2010). Les marqueurs verbaux de constat : un lieu de dialogisme dans l'écrit scientifique. In *Actes du colloque Dialogisme : langue, discours*.Hatier, S. (2016). *Identification et analyse linguistique du lexique scientifique transdisciplinaire. Approche outillée sur un corpus d'articles de recherche en SHS* (Doctorat, Université Grenoble Alpes).

Hyland, K. (2001). Bringing in the Reader : Addressee Features in Academic Articles. *Written Communication*, *18*(4), 549-574. https ://doi.org/10.1177/0741088301018004005

Hyland, K. (2005). *Metadiscourse : Exploring Interaction in Writing*. Bloomsbury Publishing.

Kraif, O., & Tutin, A. (2017). Des motifs séquentiels aux motifs hiérarchiques : l'apport des arbres lexico-syntaxiques récurrents pour le repérage des routines discursives. *Corpus* (17).

Legallois, D., & Tutin, A. (2013). Présentation : Vers une extension du domaine de la phraséologie. *Langages* (189), 3-25. https ://doi.org/10.3917/lang.189.0003

Prévost, S. (2003). Détachement et topicalisation : des niveaux d'analyse différents. *Cahiers de praxématique* (40), 97-126.

Riegel, M., Pellat, J.-C., & Rioul, R. (2018). *Grammaire méthodique du français*. PUF.

Schmid, H.-J. (2012). *English Abstract Nouns as Conceptual Shells : From Corpus to Cognition*. Walter de Gruyter.

Todorov, T. (1981). *Mikhaïl Bakhtine : le principe dialogique*. Paris : Éditions du Seuil. DL 1981.

Tran, T. T. H. (2014). *Description de la phraséologie transdisciplinaire des écrits scientifiques et réflexions didactiques pour l'enseignement à des étudiants non-natifs : application aux marqueurs discursifs* (PhD Thesis). Université de Grenoble.

Published in : M. Mangeot et A. Tutin (dir.), *Lexique(s) et genre(s) textuel(s) : approches sur corpus*, Editions des archives contemporaines, ISBN : 9782813003454, doi : 10.17184/eac.9782813003454

Tutin, A. & Kraif, O. (2016). Routines sémantico-rhétoriques dans l'écrit scientifique de sciences humaines : l'apport des arbres lexico-syntaxiques récurrents. *Lidil. Revue de linguistique et de didactique des langues* (53), 119-141.

William J. Vande Kopple. (1985). Some Exploratory Discourse on Metadiscourse. *College Composition and Communication*, *36*(1), 82-93. https ://doi.org/10.2307/357609

Williams, J. M. (1981). *Style : Ten lessons in clarity & grace*. Glenview, Ill : Scott, Foresman.

Yan, R. (2017). Étude des constructions verbales scientifiques dans une perspective didactique : utilisation des corpus dans le diagnostic des besoins langagiers du FLE à l'aide des techniques de TAL. Consulté 19 novembre 2017, à l'adresse

Published in : M. Mangeot et A. Tutin (dir.), *Lexique(s) et genre(s) textuel(s) : approches sur corpus*, Editions des archives contemporaines, ISBN : 9782813003454, doi : 10.17184/eac.9782813003454