



HAL
open science

L'évaluation dans l'écrit scientifique en français : étude sur des collocations évaluatives du modèle Nscientifique + Adjévaluatif

Yujing Ji

► **To cite this version:**

Yujing Ji. L'évaluation dans l'écrit scientifique en français : étude sur des collocations évaluatives du modèle Nscientifique + Adjévaluatif. Linguistique. 2017. dumas-01718722

HAL Id: dumas-01718722

<https://dumas.ccsd.cnrs.fr/dumas-01718722>

Submitted on 27 Feb 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License



L'évaluation dans l'écrit scientifique en français : étude sur des collocations évaluatives du modèle

N_{scientifique} + Adj_{évaluatif}

Nom Prénom : JI YUJING

UFR LLD – Département de Didactique du FLE

Mémoire de Master 2 – 28 ECTS - Mention

Spécialité : Didactique du français langue étrangère et seconde dans le monde

Sous la direction de Mme Cavalla

Année universitaire 2016-2017

Remerciements

Tout d'abord, je voudrais remercier tout particulièrement ma directrice Madame Cavalla. Sans ses soutien et guide tout au long de ma rédaction, je ne serais pas arrivée à finir ce travail. Ses commentaires, nos entretiens réguliers m'ont beaucoup aidé pendant la conceptualisation de mes pensées. De plus, son aide sur la langue, la forme m'ont été également très profitable.

Ensuite, je tiens à remercier à l'Université Sorbonne Nouvelle et tous les professeurs de l'université. Grâce à leur dévouement, j'ai découvert des cours très intéressants et inspirants qui m'ont conduite à connaître le monde de la didactique du français langue étrangère, à l'étude de laquelle je vais me consacrer durablement.

Encore, je suis reconnaissante de leur soutien à mes amies en France et en Chine, surtout Liyuan, Yiheng, Menglu, qui ont fait ma vie en couleur à Paris, mes parents qui m'ont toujours soutenue.

À la fin, je voudrais dire merci à M. Jean de Bannes qui m'a offert son aide pour améliorer mon expression écrite.

DÉCLARATION

1. Ce travail est le fruit d'un travail personnel et constitue un document original.
2. Je sais que prétendre être l'auteur d'un travail écrit par une autre personne est une pratique sévèrement sanctionnée par la loi.
3. Personne d'autre que moi n'a le droit de faire valoir ce travail, en totalité ou en partie, comme le sien.
4. Les propos repris mot à mot à d'autres auteurs figurent en citations et les auteurs sont mentionnés.
5. Les écrits sur lesquels je m'appuie dans ce mémoire sont systématiquement référencés selon un système de renvoi bibliographique clair et précis.

NOM : PRENOM :

DATE : SIGNATURE :

Tableau des matières

Remerciements.....	2
Introduction.....	8
Partie I - Cadre théorique.....	12
Chapitre 1. Le discours scientifique	13
1.1. La notion de discours	13
1.2. Le type du discours et le genre du discours : précision terminologique.....	15
1.3. Le discours scientifique et les genres du discours scientifique	16
1.4. L'écrit scientifique universitaire	17
1.5. Les difficultés des étudiants non natifs face à l'écrit scientifique	17
1.6. L'exigence spéciale de l'écrit scientifique : savoir évaluer	19
Bilan.....	20
Chapitre 2. L'évaluation dans le discours : les études antérieures	21
2.1. La genèse des recherches sur l'évaluation du texte	22
2.2. La théorie de l'évaluation de Hunston et Tompson (2000).....	23
2.3. Appraisal theory (Martin et White : 2005).....	25
2.4. L'évaluation dans le discours scientifique en français.....	29
Bilan.....	30
Chapitre 3. Les collocations évaluatives transdisciplinaires du genre N _{scientifique} + Adj _{évaluatif}	32
3.1. Lexique scientifique transdisciplinaire	32
3.2. La définition des collocations	34
3.2.1. L'acceptation statistique de la notion de collocations	34
3.2.2. L'acceptation sémantique de la notion des collocations	35
Bilan.....	37
3.3. Les collocations transdisciplinaires évaluatives du genre N _{scientifique} + Adj _{évaluatif} : des caractéristiques principales	38
3.4. Le traitement sémantique des constituants des collocations N _{scientifique} + Adj _{évaluatif}	40
3.4.1. Des noms scientifiques dans les collocations en question.....	40
3.4.2. Des adjectifs évaluatifs dans les collocations en question.....	42
Bilan.....	44

Chapitre 4. La linguistique de corpus pour extraire les collocations.....	45
4.1. Sur l'aspect qualitatif : l'importance du contexte.....	45
4.2. Sur l'aspect quantitatif : l'importance de la statistique.....	47
4.3 Constitution des corpus.....	48
4.3.1. La notion de corpus.....	48
4.3.2. Les critères de la sélection des données pour construire le corpus	49
4.3.3. La taille du corpus : grande ou restreinte?.....	50
4.4. La question de la généralité et de la représentativité du corpus spécialisé.....	50
Bilan.....	51
Partie II - Matériels et Méthodologie.....	52
Chapitre 5. Matériels et méthodologie	53
5.1 Matériels :	53
5.1.1 Corpus Scientext.....	53
5.1.2. Le mode d'emploi du corpus Scientext : ScienQuest pour la requête des éléments linguistiques - l'exemple du corpus de l'écrit scientifique en français.....	54
5.1.3. Le traitement statistique des données du site ScienQuest.....	56
5.2. La constitution du corpus de notre recherche	58
5.2.1. Le corpus de base.....	58
5.2.2. Tentative de sélection des collocations	59
5.2.2.1. La forme ou le lemme.....	60
5.2.2.2. Le seuil de spécificité pour sélectionner les collocations à traiter	60
5.2.3. Le traitement manuel pour supprimer les bruits dans la liste de collocations obtenues par traitement informatique.....	68
5.3. Les principes de classement des collocations $N_{\text{scientifique}} + \text{Adj}_{\text{évaluatif}}$: le modèle et la méthode	69
5.4. Les aspects d'analyses.	72
5.4.1. Analyse quantitative.....	73
5.4.2. Analyse qualitative sur la sémantique des collocations et leurs fonctions rhétoriques	73
Bilan.....	73
Partie III - Analyses	75

Chapitre 6. Analyses et résultats.....	76
6.1. Analyses quantitative	76
6.1.1. Analyse de la fréquence des collocations évaluatives dans les huit disciplines	76
6.1.2. Analyse des collocations axiologiques et non axiologiques dans les huit disciplines.....	81
6.1.3. Mise en commun des huit disciplines	88
6.1.4. Analyse de la distribution des noms scientifiques dans les collocations évaluatives	91
Bilan.....	93
6.2. Analyses qualitatives.....	94
6.2.1. Analyses sémantiques des collocations évaluatives.....	94
6.2.1.1. Analyses sémantiques sur des collocations axiologiques	95
6.2.1.2. Analyses sémantiques sur des collocations non axiologiques ..	100
6.2.1.3. Mise en commun des analyses sur la combinaison entre les constituants des collocations	106
6.2.2. Analyses sur les patrons syntaxiques et syntagmatiques des collocations évaluatives	107
6.3. Transdisciplinarité des collocations évaluatives	111
Bilan.....	112
Limite des résultats	113
Partie IV - Propositions didactique	114
Chapitre 7. Propositions didactiques à l'aide du corpus Scientext.....	115
7. 1. Corpus Scientext comme outil d'enseignement.....	115
7.1.1. Corpus comme outil d'enseignement : avantages.....	116
7.1.2. Scientext comme outil d'enseignement pour les collocations N _{scientifique} + Adj _{évaluatif}	116
7.2. Exploitation du corpus Scientext : principes et processus.....	117
7.2.1. Principes d'exploitation du Corpus Scientext.....	117
7.2.2. Processus d'enseignement dans l'approche indirecte	118
7.2.2.1. Constitution des petits corpus à enseigner	118
7.2.2.2. Processus d'enseignement des collocations évaluatives.....	119
Bilan.....	126

7.3. Vers l'autonomie de l'apprenant non natif.....	126
Bilan.....	128
Conclusion et perspective	129
Références bibliographiques :.....	132
Annexe 01 : Tableaux des collocations évaluatives N _{scientifique} + Adj _{évaluatif} en question.....	143
Annexe 02 : Tableaux de classement des collocations évaluatives N _{scientifique} + Adj _{évaluatif}	152

Introduction

Ce mémoire s'inscrit dans le domaine du français sur objectifs universitaires (FOU). Comme la filière du français sur objectifs spécifiques (FOS), le FOU a été créé pour répondre aux besoins linguistiques de l'intégration universitaire des étudiants non natifs voulant poursuivre les études dans des pays francophones (Mangiante et Parpette, 2011. 5). Le français enseigné ici n'est plus le français général, mais le français scientifique qui leur permet de réaliser des tâches académiques dans le cadre institutionnel, y compris l'écrit scientifique. Ce dernier, en tant qu'outil de communication indispensable et essentiel dans le monde de la recherche, doit satisfaire aux exigences de la rhétorique scientifique propre à la communauté scientifique de la langue cible.

Des étudiants non natifs du français, souvent accablés simultanément par le besoin de s'adapter au contexte académique inhabituel, le poids de l'appréhension des connaissances de leur discipline et la difficulté liée à la langue française (Cavalla, 2007 : 38), éprouvent plus de difficultés que les natifs face à l'exigence de l'écrit scientifique, tant sur la stratégie rhétorique que sur le lexique à utiliser pour composer leur propre discours à la française. L'expression *à la française* signifie ici faire quelque chose «comme font les Français»¹. En conséquence, rédiger un écrit scientifique à la française consiste à élaborer un discours dont la langue et la rhétorique impliquées répondent aux habitudes langagières inhérentes à la communauté de recherche des Français.

Parmi ces habitudes langagières, le lexique scientifique transdisciplinaire s'avère très important, car il représente, selon Tutin (2007 : 6), la manière habituellement utilisée par la communauté scientifique pour décrire ou présenter les activités scientifiques. Il ne s'agit pas ici des termes scientifiques appartenant à une discipline précise, mais des lexiques qui peuvent être employés dans les activités rédactionnelles pour construire des discours visant à décrire les objets scientifiques ou bien les procédures scientifiques.

Notre étude ne se limite pas purement aux lexiques transdisciplinaires dans le

¹ à la française — Wiktionnaire. (s. d.). Consulté le 3 février, 2017, à l'adresse https://fr.wiktionary.org/wiki/%C3%A0_la_fran%C3%A7aise

discours scientifique, mais se concentre plus précisément sur des combinaisons phraséologiques qui relèvent typiquement du discours scientifique. Plus précisément, notre étude concerne un genre typique des phénomènes phraséologiques : les collocations évaluatives du genre $N_{\text{scientifique}} + \text{Adj}_{\text{évaluatif}}$. Nous nous intéressons à ce genre de collocations pour quatre raisons :

1. Les collocations du genre $N_{\text{scientifique}} + \text{Adj}_{\text{évaluatif}}$ dans le discours scientifique en français sont rarement étudiées : nous retenons seulement le travail de Tutin et Cavalla en 2008 qui traite ce genre de travail.

2. Les adjectifs dans les collocations évaluatives du genre $N_{\text{scientifique}} + \text{Adj}_{\text{évaluatif}}$ jouent un rôle important dans l'argumentation et remplissent une grande fonction interpersonnelle dans le discours scientifique (Tutin, 2010 : 1).

3. Les adjectifs montrent l'attitude (ou le positionnement) de l'auteur envers les activités scientifiques et dévoilent sa stratégie rhétorique pour valoriser et promouvoir son étude.

4. Nous croyons que l'enseignement de ce genre de collocations évaluatives peut aider l'apprenant non natif à mieux évaluer dans l'écrit scientifique.

Dans le but de découvrir l'utilisation des collocations évaluatives dans l'écrit scientifique en français, notre travail invite à entrer également dans le domaine de la linguistique de corpus, et c'est sur les textes scientifiques en français regroupés dans le corpus Scientext² que nous envisageons de mener notre recherche. Le corpus Scientext est un corpus conçu principalement par le Laboratoire de Linguistique et de Didactique du Français Langue Étrangère et Maternelle (LIDILEM) de l'Université Grenoble 3-Stendhal. Il a pour objectif d'étudier les phénomènes phraséologiques ainsi le positionnement et le raisonnement dans l'écrit scientifique. Les collocations évaluatives seront extraites du

² Le projet ANR Scientext, développé par le LIDILEM de l'Université Grenoble 3-Stendhal, a pour objectif de mener des recherches sur « l'étude des marques du positionnement et du raisonnement dans les écrits scientifiques » (Tutin et Cavalla, 2008). Nous allons l'aborder minutieusement dans le chapitre suivant qui est dédié à la méthodologie. Le site d'internet : <http://scientext.msh-alpes.fr/scientext-site/spip.php?article39>

corpus Scientext et classées en fonction des valeurs sémantiques des noms scientifiques et des adjectifs évaluatifs dans les collocations évaluatives du patron $N_{\text{scientifique}} + \text{Adj}_{\text{évaluatif}}$. Par la suite, nous effectuerons des observations et des analyses sur les classes des collocations évaluatives afin de démontrer leurs propriétés sémantiques, syntaxiques et rhétoriques.

De ce fait, ma problématique est ainsi la suivante : comment les collocations évaluatives manifestent-elles l'évaluation de l'auteur dans les différentes disciplines et dans les différentes parties textuelles de l'écrit scientifique en français. Cette problématique peut être mieux précisée par les questions suivantes : quelles valeurs sémantiques les collocations évaluatives du genre $N_{\text{scientifique}} + \text{Adj}_{\text{évaluatif}}$ interprètent-elles? Quelles fonctions rhétoriques sont remplies par les collocations évaluatives du patron $N_{\text{scientifique}} + \text{Adj}_{\text{évaluatif}}$ dans le discours scientifique? Existe-t-il des différences ou des similarités d'utilisation dans des disciplines différentes? Dans des parties textuelles différentes? Quelles propriétés les combinaison entre des noms scientifiques et des adjectifs manifestent-elles? Comment enseigner ces collocations évaluatives ?

Pour traiter cette problématique, il nous faut avancer les hypothèses suivantes :

1. L'utilisation des collocations évaluatives est différente d'une discipline à l'autre.
2. L'utilisation des collocations évaluatives varie en fonction des parties textuelles de l'écrit scientifique.
3. Chaque classe des collocations évaluatives représente une valeur sémantique spécifique et dispose d'une fonction rhétorique spécifique qui est liée à un aspect spécifique de la dimension interpersonnelle.
4. Le corpus Scientext peut servir d'outil d'enseignement et d'apprentissage des collocations évaluatives.

À partir de ces hypothèses, notre étude comportera quatre parties :

La première partie comporte quatre chapitres et s'articulera autour du cadre théorique pour encadrer notre étude dans le champ théorique approprié. Le premier

chapitre traitera la question du discours scientifique et les difficultés rencontrées par les étudiants non natifs de français. Le deuxième chapitre discutera les études de nos devanciers sur l'évaluation dans le discours pour s'orienter vers une définition de l'évaluation adaptée à notre étude. Le troisième chapitre se concentrera d'abord sur la distinction entre lexique transdisciplinaire et lexique général, et puis mettra en lumière la notion de collocation et son évolution dans le domaine phraséologique, et enfin présentera les traits sémantiques des noms scientifiques et des adjectifs évaluatifs, ce qui constituera les fondements du classement des collocations évaluatives. Le quatrième chapitre expliquera la linguistique de corpus et les théories liées à la constitution du corpus.

La deuxième partie aura pour ambition de présenter le corpus Scientext, son mode d'emploi, et la constitution de notre corpus, ainsi que les principes d'extraction des collocations dans notre corpus, la manière d'analyses.

La troisième partie abordera les analyses quantitatives et qualitatives sur le classement des collocations évaluatives afin de répondre aux hypothèses émises dans l'introduction.

La quatrième partie exposera nos propositions didactiques des collocations évaluatives dans le but d'aider l'apprenant à mieux pratiquer l'évaluation dans le cadre de l'écrit scientifique.

Nous concluons notre étude en rappelant dans un premier temps les résultats de nos observations ; dans un deuxième temps nous préciserons les limites de notre étude; et dans un dernier temps nous ouvrirons d'autres pistes de recherches pour de futures études.

Partie I - Cadre théorique

Chapitre 1. Le discours scientifique

1.1. La notion de discours

Il est difficile de donner une définition à la notion de discours et maints chercheurs ont tenté d'en trouver une. Aujourd'hui, la notion de discours a connu une extension, ce qui nous met face à plusieurs définitions possibles. En les interrogeant, nous n'avons pas manqué de trouver que cette notion allait souvent de pair avec quelques notions proches telles que la langue, l'énoncé, le texte ou la phrase qui sont sujettes à causer des confusions.

Au tout début, Saussure a opposé la parole à la langue. Pour lui, la langue est un « système de signes acceptés par une communauté linguistique » alors que la parole représente « l'acte de d'individu qui utilise la langue pour émettre ou pour comprendre des énoncés » (Saussure, 1979 : 17). Ici, le terme *parole* saussurien est traité comme le synonyme du terme *discours*. Maingueneau (1987) a précisé cette opposition en allant plus loin et expliqué que le discours, enveloppe « toute production langagière », « n'a affaire au langage que par là où il fait sens pour des sujets inscrits dans des stratégies d'interlocution, des positions sociales ou des conjonctures historiques » (Maingueneau, 1987 : 7). La distinction entre la langue et le discours nous conduit à voir que le discours se différencie par son orientation vers la dimension sociale ou mentale. A.H.Gardiner considère le discours dans la dimension sociale, il confirme que le discours constitue un outil pour que les hommes communiquent entre eux à travers des signifiants en cas d'exprimer leur opinion. La dimension mentale du discours s'explique par le fait que le discours se réalise en fonction de la condition psychique (Charaudeau & Maingueneau, 2002 : 185-186).

Quant à Benveniste (1966), dans sa définition, le terme *discours* est proche du terme *énoncé*. D'après lui, le discours est « tout énonciation supposant un locuteur et un auditeur et chez le premier l'intention d'influencer l'autre en quelque manière » (Benveniste, 1966 : 242). Cependant, il faut signaler que les deux termes ne sont pas les mêmes bien qu'ils soient proches. L'énoncé, le produit de l'acte de l'énonciation, est « la suite des phrases émises entre deux blancs sémantiques, deux arrêts de la communication »

tandis que le discours pourrait être vu comme « l'énoncé considéré du point de vue du mécanisme discursif qui le conditionne ». (Guespin, 1971 : 10). Dans cette citation, nous comprenons le fait qu'il existe trois niveaux hiérarchiques entre la phrase, l'énoncé et le discours : les phrases constituent l'énoncé qui constitue enfin le discours en ajoutant la prise en compte des conditions de la production de cet énoncé. Même si on dit que la phrase est l'un des constituant du discours, ceci n'implique pas qu'un discours doive toujours se manifester plus long qu'une phrase car une seule phrase pourrait aussi devenir un discours, c'est par exemple le cas d'un proverbe. La différence réside justement dans l'organisation des phrases. Le discours est obligé de répondre aux exigences d'organisation des phrases d'une communauté ou d'un genre de discours (un résumé ne s'organise pas de même façon qu'une nouvelle).

En parallèle, nous trouvons la définition de Maingueneau (1976) qui préfère voir le discours en tant que « résultat d'une construction », « résultat de l'articulation d'une pluralité plus ou moins grande de structurations transphrastiques, en fonction des conditions de production ». (Maingueneau, 1976 : 16). Cet essai de définir le discours reprend les éléments importants dans la distinction entre l'énoncé et le discours que nous trouvons chez Guespin : les phrases articulées (l'énoncé) et les conditions. Mais il complète ces deux éléments par le résultat de constitution ou de production. Le discours doit être compris en même temps comme résultat de constitution et les conditions dans lesquelles l'on obtient ce résultat. Ici, ce n'est plus le processus qui compte mais plutôt le fruit du processus de production et la situation où se passe cette production qui s'avèrent plus importants. C'est aussi pourquoi Jaubert (1990) pense que le discours est « du langage en situation » (Jaubert, 1990 : 22). La situation ou le contexte constitue donc une condition indispensable pour le discours. Maingueneau a déjà mentionné l'importance de la situation en 1976 en formulant le fameux schéma :

«ÉNONCÉ + situation de communication = DISCOURS» (Maingueneau, 1976 : 12)

C'est aussi l'élément qui oppose le discours au texte : le premier est « conçu

comme l'inclusion » du dernier dans son contexte (la situation de communication) (Charaudeau & Maingueneau, 2002 : 186). Le texte, selon Charaudeau (1988), est « un objet qui représente la matérialisation de la mise en scène de l'autre langage. Il est un résultat toujours singulier d'un processus qui dépend d'un sujet parlant particulier et de circonstances de production particulières » (Charaudeau 1988 : 69 cité par AOUADI Lemya, 2015, p43). Et « une étude linguistique des conditions de production de ce texte en fera un discours » (Guespin, 1971 :10).

En interrogeant ces définitions variées, nous pouvons voir combien il est complexe de définir le discours. Mais nous pouvons quand même dégager trois dimensions du discours : la dimension linguistique (énoncé en forme de texte écrit ou oral), la dimension communicationnelle et la dimension sociologique (le contexte de production) (AOUADI Lemya, 2015, p44). La dimension linguistique discute le discours en examinant la langue, la phrase, l'organisation des phrases et l'acte de langage là-dedans. Au niveau de l'organisation, un discours sera orienté vers une finalité du locuteur. La dimension communicationnelle traite l'interaction entre le locuteur et l'interlocuteur. La dimension sociologique est le contexte de production des deux premières dimensions.

1.2 Le type du discours et le genre du discours : précision terminologique

Les idées en matière de l'emploi terminologique pour désigner des variétés discursives sont divergentes. Nous confrontons par exemple le « genre de discours » et le « type de textes » (J.-M.Adam, 1999) , le « type de discours » et le « genre de texte » distingués par J.-P.Bronckart (1996), etc. Dans notre travail, nous voudrions prendre la distinction de Maingueneau : « type de discours », « genre de discours » (Maingueneau, 1991 : 213). Les deux termes sont précisés ainsi :

- a. **Type de discours** : il s'agit d'une catégorisation élémentaire des discours en fonction des secteurs de l'activité discursive, de la cohérence textuelle. Comme le discours littéraire, politique, scientifique...
- b. **Genre de discours** : il sous-catégorise les types de discours en prisant compte « des lieux d'énonciation déterminés » (Ibid). Par exemple, pour le discours littéraire,

nous trouvons le théâtre, la poésie, le roman, etc.

1.3. Le discours scientifique et les genres du discours scientifique

Le discours scientifique est un type de discours produit au cours de l'activité scientifique. Il est

les pratiques, formes orales, écrites, pluricodiques, dont le but est la construction et /ou la diffusion de connaissances et de concepts appartenant aux champs qu'on s'accorde socialement à considérer comme scientifiques. (Astofli, 1995 : 57, cité par Mroue, 2014 : 9).

Cette définition permet de distinguer en premier abord le discours scientifique selon deux principes élémentaires : la forme et le champ où s'inscrit le contenu. Si l'on distingue le discours scientifique du point de vue formel, il peut se diviser en discours scientifique écrit dont le support des connaissances est un matériel, et discours scientifique oral. En fonction du contenu, le discours scientifique peut se classer selon les domaines des connaissances qu'il traite, par exemple la linguistique, les mathématiques et d'autres domaines. Ces deux catégorisations du discours scientifique ne sont qu'élémentaires. Pour aboutir à l'écrit scientifique universitaire, il nous faut catégoriser le discours scientifique d'un autre point de vue. Nous nous inspirons de la catégorisation de Loffler-Laurian (1983). Ce linguiste catégorise le discours scientifique à l'aune de la situation de communication, de l'identité de l'émetteur, du récepteur et de la nature du message (Laurian, 1983 : 9-12) :

1. Discours scientifique spécialisé, ex : l'article dans la revue de recherche
2. Discours de semi-vulgarisation scientifique, ex : des revues pour des formants universitaires
3. Discours de vulgarisation scientifique, ex : domaines spécialisés dans la presse générale
4. Discours de type mémoire, thèse, etc.
5. Discours scientifique pédagogique, ex : un manuel pour l'étudiant
6. Discours scientifiques officiels, ex : rapport destiné à la politique

Selon la situation de communication, le discours de type mémoire, thèses et le discours de type pédagogiques sont les deux discours scientifiques auxquels est confronté souvent l'étudiant universitaire. En particulier pour le discours de type mémoire, thèses etc., l'étudiant joue le rôle d'émetteur de message pour montrer ses acquis des connaissances d'un champ très précis à l'attention de l'enseignant, ou de jury, etc. ce qui construit un discours scientifique très spécial : le discours scientifique universitaire.

1.4. L'écrit scientifique universitaire

L'écrit scientifique universitaire s'inscrit dans le discours scientifique universitaire qui est réalisé via l'écrit des textes, non à travers l'oral. Nous proposons ici de subdiviser encore ce genre de discours scientifique en adoptant le classement de Cavalla et Bou (2010, cité par Mroue, 2014). Ces deux chercheurs ont mené une enquête auprès des responsables de formations de Master à l'université Grenoble Alpes dans différentes disciplines pour connaître les modalités d'évaluation scientifique mises en place dans ces disciplines et ont proposé ensuite trois ensembles (Mroue, 2014 : 13) :

1. Les productions initiales, ex : les recherches bibliographiques, les prises de notes
2. Les productions de processus, ex : rapports de stages, les résumés, les synthèses...
3. Les productions finales, ex : écrit longs comme les mémoires, les thèses.

Dans notre travail, nous fixons notre objectif comme l'aide à la rédaction de l'écrit scientifique du genre des productions finales, en particulier le mémoire et la thèse.

1.5. Les difficultés des étudiants non natifs face à l'écrit scientifique

Pour valider ses études en France, tout étudiant, qu'il soit natif de langue française ou non, doit obligatoirement passer les épreuves des universités ou établissements d'enseignement supérieurs en français (sauf quelques exceptions). En dépit de la diversité des épreuves en fonction des disciplines et surtout des enseignants, le mémoire ou la thèse, fruit de la conceptualisation et de la présentation de l'activité scientifique à la fin des

études supérieures, reste un travail incontournable pour des étudiants voulant obtenir le diplôme master ou doctorat. Les étudiants manifestent par conséquent un besoin immédiat de bien rédiger leur écrit scientifique.

Sachant que pour les étudiants français la rédaction de l'écrit scientifique comme le mémoire ou la thèse pose déjà beaucoup de problèmes, pour les étudiants étrangers dont la langue maternelle n'est guère le français, la situation s'avère plus grave, en particulier sur la langue française (Boche et Firer, 2015 : 86). Cavalla (2007) constate que les étudiants non natifs du français sont confrontés principalement à trois types de difficultés face à l'écrit scientifique :

- acquérir les connaissances dans la discipline dans laquelle ils doivent s'intégrer
- s'adapter à la culture de la communauté scientifique.
- maîtriser les connaissances linguistiques pour bien rédiger leur écrit scientifique (Cavalla, 2007 : 38)

La première difficulté appartient aux activités cognitives propres à l'appréhension des savoirs ou savoirs-faire disciplinaires qui ne fait pas l'objet de notre recherche.

Ce qui nous intéresse le plus, ce sont les deux dernières difficultés qui sont en lien avec l'écrit scientifique du point de vue sociologique et linguistique. Sociologiquement, l'une des caractéristiques de l'écrit est d'être étroitement liée à une communauté plus restreinte qui « à travers leur rites langagiers partagent un même territoire » (Maingueneau, 1992 : 117). On appelle habituellement communauté un ensemble de personnes qui partagent les mêmes rites. Par exemple du point de vue linguistique, les gens qui partagent la même langue constituent une communauté linguistique mais de façon hétérogène (les Québécois et les Français). Pour les étudiants non natifs, ils possèdent déjà un bagage culturel et langagier qui est probablement étranger par rapport à la culture cible. La deuxième difficulté des étudiants non natifs consiste donc à l'acculturation à la culture de la communauté scientifique. Plus précisément, la communauté scientifique s'avère spécifique, car elle fonctionne avec ses propres critères discursifs qui s'incarnent par des

choix linguistiques spécifiques tels que les verbes de démarcation (Chavez 2008, Tran 2011), les adjectifs d'évaluation (Tutin, Cavalla, 2008, 2010), les pronoms personnels (Hyland 2002).

1.6. L'exigence spéciale de l'écrit scientifique : savoir évaluer

Depuis longtemps, les étudiants non natifs se sentent démunis face au positionnement dans l'écrit scientifique (Hyland, 2002 : 1093). Pour eux, évaluer signifie de prendre une posture subjective dans l'écrit scientifique qui a l'air "objectif". Néanmoins, par opposition à ce que les étudiants croient, l'écrit scientifique comporte lui aussi un aspect bien subjectif.

Quand on parle de l'écrit scientifique, on pense souvent à un genre objectif. Certes, de nombreux outils pour la rédaction de l'écrit scientifique (tel que celui de Cislaru et Claudel) recommandent de garder une attitude objective. Prenons comme exemple Cislaru, dans *L'écrit universitaire en pratique* :

Un travail de recherche répond à un principal d'objectivité [...] Cette spécificité du discours de recherches est représentée, dans les textes, par un effacement des marqueurs de subjectivité tels que les évaluatifs (bon, meilleur, bien, intéressant, etc.) et par un usage des modalisations opérant une mise à distance (emploi du conditionnel ou d'adverbes comme assez, plutôt, etc.) ((Cislaru, Claudel, & Vlad, 2009 : 111)

Ce petit exemple nous permet de voir combien l'objectivité compte dans l'écrit scientifique, et que les éléments évaluatifs sont souvent déconseillés. Cependant, d'autres voix réagissent et contestent le caractère purement objectif de l'écrit scientifique. Production de l'être humain, l'écrit scientifique ne peut pas rester tout à fait indifférent à l'attitude de son producteur. Il ne s'agit pas seulement d'un texte scientifique où l'auteur explique un fait en entassant des connaissances scientifiques en vigueur, mais aussi et inévitablement d'un lieu où l'auteur donne son propre point de vue vis-à-vis des recherches faites par les prédécesseurs, éventuellement établit ses propres recherches innovantes et les évalue.

Les étudiants non natifs ne sont pas à l'aise pour se positionner ou évaluer. Selon

les enquêtes de Chavez (2008) sur l'utilisation des verbes de démarcations dans les corpus des écrits scientifiques de natifs et de non natifs, les verbes de démarcation recensés dans le corpus de natifs sont totalement introuvables dans le corpus de non natifs, ce qui démontre que, contrairement aux natifs, les non natifs ne savent pas bien se positionner. L'évaluation, en tant qu'un aspect du positionnement, s'avère aussi indispensable dans l'écrit scientifique et mérite une attention particulière lors de l'aide à l'écrit scientifique des apprenants non natifs.

Bilan

Le discours scientifique comporte plusieurs genres selon la typologie de Lorrain, mais nous nous intéressons à l'écrit scientifique universitaire, en particulier le mémoire et la thèse qui représentent le travail final pour obtenir un diplôme dans le système français d'éducation. Les étudiants non natifs éprouvent des difficultés face à ce genre de l'écrit scientifique. L'acculturation à la communauté scientifique leur impose de respecter non seulement la manière dont se positionnent les auteurs dans celle-ci, mais aussi les ressources linguistiques pour la réaliser. L'évaluation, manifeste le positionnement de l'auteur dans l'écrit scientifique, sa maîtrise exige une attention particulière de la part des étudiants non natifs.

Chapitre 2. L'évaluation dans le discours : les études antérieures

L'évaluation est un substantif dérivé du verbe *évaluer*. Dans le dictionnaire étymologique, "évaluer" a pris son origine dans le verbe « esvaluer » qui signifie « déterminer la valeur, le prix de quelque chose » (Dictionnaire étymologie CNRTL³ en ligne, article *évaluer*). En d'autres termes, *évaluer* signifie « extraire de la valeur » ou « faire ressortir la valeur de » d'un objet (Pelpel, 2003). L'évaluation est donc l'action d'évaluer l'objet. Dans le domaine de l'analyse du discours, on entend par évaluation les outils linguistiques que la langue mobilise pour évaluer les objets.

Les recherches sur le phénomène d'évaluation dans le discours font preuve d'éclectisme en matière de perspectives, ce qui contribue à la profusion de termes. En fin de compte, différents termes sont dus aux différents penchants pour des théories linguistiques, des centres, des méthodes et des modalités de recherches. Dans le monde anglophone, nous trouvons « connotation » (Lyons, 1977), « intensity » (Labov, 1984), « attitude » (Halliday, 1985), « evidentiality » (Chafe, 1986), « stance » (Barton, 1993), « evaluation » (Hunston, 2000), « appraisal » (Martin et White, 2005). Dans le monde francophone, et pour ce travail, nous nous référons aux études de Tutin et Cavalla (2008) qui ont pris le terme *évaluation* dans le cadre du positionnement. Les ressources linguistiques à étudier pour représenter l'évaluation dans le discours par des pionniers ne font pas l'unanimité. À titre d'exemple, Ochs (1989 : 1) a indiqué que l'affect est un vaste domaine sémantique que les langues codent richement à travers les structures lexicales, grammaticales et discursives, tels que la prosodie, la variation phonologique, les particules morphologiques, les affixes, le système pronominal, les quantificateurs, l'emphatie, l'adverbe, les structures parallèles, les répétitions etc. Chaque approche se singularise par le recours à une ou plusieurs ressources linguistiques, ce qui rend la compréhension de la question de l'évaluation plus difficile. Comme nous jugeons que la diversité terminologique ne permet pas une compréhension claire du phénomène de l'évaluation

³ CNRTL : Centre National de Ressources Textuelles et Lexicales. Site d'internet : <http://www.cnrtl.fr/etymologie/>

dans le discours, il vaut mieux débiter par un éclaircissement sur cette situation complexe en abordant quelques théories et travaux sur l'évaluation afin de préciser quelle recherche sera le mieux adaptée à notre travail.

2.1. La genèse des recherches sur l'évaluation du texte

Halliday est le pionnier de la recherche sur l'évaluation du discours. À partir des années cinquante, il commence à contribuer à l'établissement et au développement de la Linguistique Systémique Fonctionnelle qui est devenue aujourd'hui une « approche sémantique et sociale » comme outil visant à analyser le discours très répandu dans le monde anglophone (Banks, 2011). Nous disons que l'approche est à la fois sémantique et sociale parce qu'elle regarde autant sémantiquement sur l'origine du sens que socialement sur la relation interpersonnelle dans le discours. Halliday (1985) distingue trois méta-fonctions sémantiques de la langue anglaise : idéationnelle, interpersonnelle, et textuelle. La méta-fonction idéationnelle concerne les représentations des processus et des éléments dans le monde, la méta-fonction interpersonnelle observe les relations entre des membres d'une communauté sociale ainsi que celles entre le locuteur et les messages. La méta-fonction textuelle concerne la constitution des messages dans un contexte donné qui leur permet de posséder un sens contextuel (Matthiessen & Halliday, 2009 : 12-13). Dans le cadre de l'analyse de la méta-fonction interpersonnelle, Halliday insiste sur l'interaction entre le mode et la modalité au niveau de la phrase. Il souligne le fait que la réalisation de la fonction interpersonnelle se fait par le biais des ressources linguistiques telles que le mode, la modalité et l'intonation qui interagissent de façon rythmique dans le texte. Parmi ces ressources linguistiques, le système de modalité autour du jugement et de l'évaluation constitue le moyen d'exprimer l'évaluation dans le texte. Inspirés par sa théorie, les linguistes successifs, dans l'espoir d'aller plus loin, ont développé l'étude de l'évaluation du texte et proposé d'autres théories plus complètes. Parmi lesquelles, il est indispensable de mentionner la théorie de l'évaluation de Hunston et Tompson et la théorie de l'Appraisal⁴ de Martin et White.

⁴ La théorie de l'Appraisal est la traduction de l'anglais 'appraisal theory' par Pierre Gardin en 2009

2.2. La théorie de l'évaluation de Hunston et Tompson (2000)

Dans leur ouvrage *Evaluation in text : authorial stance and the construction of discourse*, les deux auteurs ont essayé de donner une définition à l'évaluation :

Evaluation is the broad cover term for the expression of the speaker or writer's attitude or stance towards, viewpoint on, or feelings about the entities or propositions that he or she is talking about. (2000 : 5)⁵

Selon cette définition, l'évaluation concerne l'attitude, le positionnement et le sentiment de l'auteur envers l'objet à évaluer. Il est clair que les deux auteurs différencient l'objet de l'évaluation en entité et proposition. Pour les distinguer d'une manière plus approfondie, ils divisent l'évaluation en deux opinions : affective et épistémique. La première attribue aux entités (le groupe nominal) une opinion "positive" ou "négative" (bon/mauvais, certain/incertain), la dernière se rapporte à l'opinion sur les propositions qui s'exprime souvent par des paraphrases. De plus, les deux chercheurs ont tenté de définir l'évaluation à partir de trois fonctions :

1. Exprimer le point de vue de l'auteur ou de l'orateur, pour refléter le système de valeur de cette personne et de sa communauté.
2. Construire et maintenir les relations entre l'orateur ou l'auteur et l'auditeur ou le lecteur.
3. Organiser le discours.

La première fonction occupe un statut central, elle permet de repérer le point de vue de l'auteur. Cet acte d'évaluation révèle la valeur de la communauté (Hunston et Tompson,

dans son article « Application de la théorie de l'Appraisal à l'analyse d'opinion ». Source en ligne : http://majecstic2009.univ-avignon.fr/Actes_MajecSTIC_RJCP/MajecSTIC/articles/1272.pdf

⁵ Traduction personnelle /de l'auteur: L'évaluation est le terme générique pour l'expression de l'attitude ou du positionnement de l'auteur, ses points de vue ou ses sentiments à propos des entités ou des propositions dont il parle.

2000 : 6). La deuxième fonction interroge la relation interpersonnelle, elle est représentée par trois aspects : la manipulation, la protection et la politesse. La manipulation consiste à manipuler le lecteur, plus l'évaluation est discrète, mieux l'auteur peut persuader le destinataire avec succès. La protection et la politesse ont l'intention de modérer la certitude des connaissances de manière à maintenir la relation entre l'auteur et le lecteur. La relation interpersonnelle n'est pas révélée seulement par les informations connotées dans le texte, mais existe aussi en termes du texte lui-même (Ibid. 6-13). La prise en compte des trois fonctions de l'évaluation amène les deux auteurs à proposer de décrire les ressources linguistiques du phénomène d'évaluation à travers des prismes plus variés : l'évaluation dans le texte fait appel aux ressources linguistiques s'étalant sur trois niveaux : lexical, grammatical et textuel.

Le niveau lexical : quatre genres lexicaux, à savoir l'adjectif, l'adverbe, le nom et le verbe.

Le niveau textuel ouvre l'étude de l'évaluation à l'échelle plus large qui traverse l'ensemble du texte.

Du point de vue grammatical, la description de l'évaluation du discours prend en compte les structures syntaxiques et les fonctions sémantiques remplies par les catégories grammaticales. Nous nous intéressons ici au travail de Hunston et Sinclair (2000 : 74-101) qui analyse les structures syntaxiques exprimant l'évaluation à l'aide du corpus numérique. Plus précisément, nous nous référons au contenu qui traite des structures syntaxiques incluant l'adjectif. Selon les deux auteurs, l'identification du caractère évaluatif d'un adjectif peut se réaliser par trois principes :

1. Si l'adjectif se produit dans des motifs évaluatifs : tels que le motif "it + link verb + adjective group + clause" (il + verbe + groupe adjectif + proposition subordonnée, ex. : It was certain that he was much to blame / il était certain qu'il avait beaucoup de reproches à faire);

2. Lorsque l'adjectif dispose des formes comparatives ou superlatives ou en occurrence avec l'adverbe de classement, comme le motif " noun group + link verb +

superlatif adjectif group + prepositional phrase” (groupe nominal + verbe + groupe adjectif superlatif + phrase prépositionnelle, ex. : The race was one of the greatest in modern times / ce fut l'une des plus grandes courses des temps modernes);

En fonction de la position de l'adjectif : préposé ou postposé par rapport au nom.

① adjectif + noun : interesting statistic (adjectif + nom : statistique intéressante), l'adjectif remplit la fonction d'épithète.

② noun group + link verb + adjectif group : they're interesting (groupe nominal + verbe + groupe adjectif : ils sont intéressants). L'adjectif ici joue le rôle d'un attribut.

Cette structure pourrait être plus complexe :

③ noun group + link verb + adjectif group + to noun group : Colonel Lewis is interesting to historians (groupe nominal + verbe + groupe adjectif + pour groupe nominal : Le colonel Lewis est intéressant pour les historiens).

Parfois, la copule peut être remplacée par d'autres verbes.

④ noun groupe + link verb + adjectif groupe : She felt miserable (groupe nominal + verbe + groupe adjectif : elle se sentait malheureuse).

Le résumé de la théorie de l'évaluation et l'explication sur des ressources linguistiques en particulier en ce qui concerne les structures syntaxiques comportant l'adjectif illustrent de façon générale la caractéristique de la théorie de l'évaluation de Hunston : cette théorie prête attention aux structures syntaxiques (soit la grammaire locale) notamment en s'arrêtant sur la combinatoire lexico-syntaxique.

2.3. Appraisal theory (Martin et White : 2005)

Cette théorie est un système interpersonnel au niveau du discours sémantique qui a été développé au cours des études dans le cadre du SFL (Systemic Functional Linguistics / Linguistique systémique fonctionnelle). Avant la théorie de l'appraisal, les études sur l'évaluation du texte dans l'approche de la linguistique systémique fonctionnelle se focalisent davantage au niveau structural et fonctionnel, comme chez Hunston et Sinclair

qui mettent l'accent sur la morphologie et la structure syntaxique de l'évaluation dans le texte. Martin juge donc que l'approche SFL manque d'une interprétation sémantique de l'évaluation :

What has tended to be elided in SFL approaches to data of this kind is the semantics of evaluation - how the interlocutors are feeling, the judgements they make, and the value they place on the various phenomena of their experience (Martin, 2000 : 144)⁶.

Dans l'intention de combler cette lacune, Martin et ses collègues suggèrent la théorie de l'Appraisal. En mettant l'accent sur le plan lexical, la théorie constitue une méthodologie destinée à étudier, décrire et interpréter les fonctions visant à l'évaluation, à construire le positionnement, le style et la relation interpersonnelle du langage dans le discours (Martin et White, 2005 : 1). En conséquence, elle est interprétative et herméneutique. Son cadre se construit généralement autour de trois sous-systèmes : l'attitude, l'engagement et la gradation, dont chacun comporte à son tour des sous-genres.

L'attitude : elle regroupe l'affect, le jugement et l'appréciation. L'affect traite les ressources pour construire les émotions, ces émotions se divisent en trois classes dichotomiques :

① Un/happiness (mé/contentement) regarde le sentiment de l'humain comme "content, mécontent, heureux, malheureux" etc.

② In/security (in/sécurité) explique la situation interpersonnelle, tel que "la peur, la confiance" ;

③ Dis/satisfaction (in/satisfaction) est liée à la poursuite d'accomplissement de l'objectif, "monotone, intéressant".

Le jugement s'occupe des ressources pour évaluer les comportements en fonction des principes normatifs et se répartit en « social esteem » (estime sociale) qui estime le comportement en fonction de l'aspect moral ou éthique et « social sanction » (sanction

⁶ La traduction personnelle en français : Ce qui a tendance à être éliminé dans les données de SFL, c'est la sémantique de l'évaluation - comment les interlocuteurs ressentent, les jugements qu'ils font et la valeur qu'ils mettent sur les différents phénomènes de leur expérience.

sociale) qui l'évalue en vertu des règles légales ou sociales. L'appréciation se consacre aux ressources pour construire la valeur des entités, y compris le phénomène naturel et sémiotique (le produit ou le processus). Elle se divise en trois sous-genres : "reaction" (réaction) signifie notre réaction face aux objets, "composition" (composition) étudie la balance et la complexité, "valuation" (évaluation) évalue les objets. Ce sous-système se divise en deux polarités : positive et négative, chaque polarité se caractérise par les intensités variées telles que basse, moyenne et haute. (Martin et White, 2005 : 35-36, 42-61)

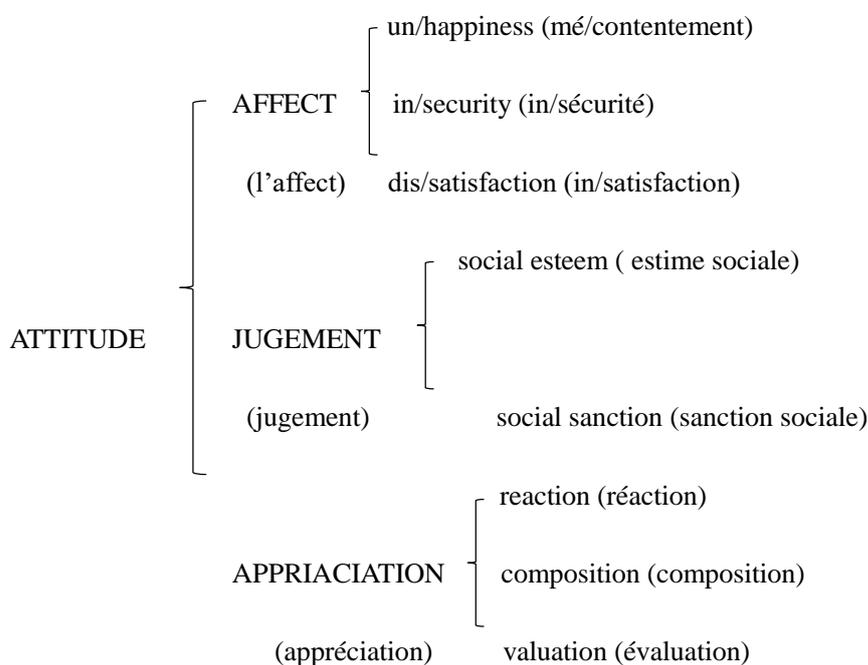


Figure 1: Sous-système Attitude du système de la théorie de l'Appraisal

a. L'engagement : ayant affaire à l'origine de l'attitude et au ton de la voix, il englobe " monogloss" langue de l'auteur et "heterogloss" langue de l'autre. Ce sous-système règle la responsabilité et le droit du contenu du texte et traite concrètement les sources des phrases dans le texte (Martin et White, 2005 : 36). La réalisation de l'engagement fait appel à la projection, l'atténuation (hedging), la modalité, etc.

b. La gradation : ce système travaille de façon complémentaire avec le système de l'attitude et l'engagement. En se combinant avec l'attitude, l'affect, le jugement et

l'appréciation peuvent être nuancés par intensification ou affaiblissement. Quant à l'engagement, l'intervention de la gradation permet à l'auteur d'exprimer par degré son point de vue sur les objets ou les travaux effectués par lui-même ou par d'autrui. La gradation se représente par deux sous genres : « force » et « focus ». Le premier discute la quantité et l'intensité des ressources linguistiques qui peuvent être graduées par degré, tandis que le deuxième vise à clarifier ou estomper la limite entre des catégories pour la raison que les ressources traitées ici ne permettent pas une division par degré, en d'autres termes, « focus » essaie à « modéliser le degré de prototypicalité » (Gardin, P. 2009 : 2).

En ce qui concerne les ressources linguistiques à étudier, comme ce système se concentre davantage sur le plan lexical, il est évident que la réalisation de l'évaluation se fait habituellement par quatre genres de ressources linguistiques qui sont semblables à celles mobilisées par la théorie de l'évaluation de Hunston et Tompson au niveau lexical : le verbe, le nom, l'adjectif et l'adverbe.

Ce système fournit un très bon cadre qui permet de raffiner l'évaluation dans le discours et de souvent limiter les ressources linguistiques à quelques valeurs définies principalement par leur polarité : positive, négative ou neutre (Gardin, P. 2009 : 2). Malgré son avantage en matière d'intégralité, une autre difficulté émerge lors de son application : ce système paraît très complexe et il est donc difficile de saisir le langage spécifique de chaque sous-système de la théorie de l'Appraisal, car la délimitation entre les sous-systèmes est souvent subjective et varie à la guise des usagers. C'est aussi la raison pour laquelle que nous jugeons que cette théorie n'est pas très appropriée à notre objet : les adjectifs d'évaluation. Ceux-ci surgissent dans plusieurs sous-systèmes de cette théorie (l'attitude et la gradation) avec les autres ressources linguistiques telles que le verbe, l'adverbe ou le nom, ce qui empêche une recherche spécifiquement dédiée à l'adjectif. De plus, à cause de la difficulté liée à la délimitation des sous-systèmes, cette théorie rencontre des difficultés et des lacunes lors de son application à une analyse du corpus à grande échelle (Song, 2010).

2.4. L'évaluation dans le discours scientifique en français

Si l'évaluation dans le discours a connu un développement à grande échelle dans le monde anglophone pour la langue anglaise, son application au discours en français s'avère plus restreinte. Tutin et Cavalla s'y sont intéressées, en particulier autour du positionnement dans l'écrit scientifique. Devant le choix du terme pour décrire ce phénomène dans l'écrit scientifique, les deux chercheuses préfèrent le terme « positionnement » qui est une dénomination déjà existante et reprise du projet ANR Scientext. Dans le cadre du projet du corpus Scientext, le « positionnement » ne signifie pas simplement « le point de vue », mais implique « la façon dont l'auteur s'inscrit dans une communauté de discours, comment il s'évalue et évalue ses pairs, et quelles propositions propres il met en avant » (Tutin, Grossmann, Falaise & Kraif, 2009). L'évaluation s'avère donc très liée au positionnement. Nous pouvons dire que c'est par s'évaluer et évaluer que l'auteur constitue son propre positionnement dans ses écrits. Tout comme la théorie de l'Appraisal de Martin et la théorie de l'évaluation de Hunston et Tompson pour l'anglais, l'étude sur le positionnement dans l'écrit scientifique en français s'effectue par le biais de l'étude sur les ressources linguistiques ou en autre terme les « marques énonciatives, lexicales et syntaxiques spécifiques » (Ibid).

Parmi les recherches déjà accomplies, le travail de Tutin et Cavalla en 2008 sur les collocations évaluatives, en particulier les collocations $N_{\text{scientifique}} + \text{Adj}_{\text{évaluatif}}$, nous intéresse le plus compte tenu de notre objectif d'étude. Dans leur travail, les deux chercheuses académiques ont limité la définition de l'évaluation de Hunston et Tompson par une mise en précision sur quatre points : « la comparaison, le jugement (par rapport à une norme), le jugement de valeur (axiologique), certains marqueurs de modalité » (2008). En conséquence, pour elles, l'évaluation des adjectifs répond aux trois catégories primaires en fonction de la sémantique : jugement axiologique qui évalue la valeur, jugement non axiologique qui se répartit en plusieurs sous catégories telles que l'importance, la nouveauté, la temporalité, le degré, etc., et les adjectifs de modalité qui évaluent souvent un énoncé.

Très originale, cette définition de l'évaluation est différente de deux théories que nous venons d'aborder. Par rapport au système très complexe de la théorie de l'Appraisal qui est capable de s'appliquer à tout genre discursif, l'évaluation de Tutin et Cavalla a l'air plus légère mais plus précise pour se concentrer sur l'écrit scientifique. Ce qui est différent dans la définition de l'évaluation de Tutin et Cavalla par rapport à celle de Hunston et Tompson, c'est qu'elle ne prend pas en compte les marqueurs d'affectivité. Cet abandon est probablement dû à la rareté de la présence des marqueurs d'affectivité dans l'écrit scientifique.

La précision de leur notion de l'évaluation n'est pas la seule raison pour laquelle nous décidons de faire allusion à leur travail, une autre dimension ouverte à l'échelle phraséologique pour étudier l'évaluation dans l'écrit scientifique nous paraît très pertinente dans la présente étude. Les deux auteurs associent les adjectifs d'évaluation avec des noms scientifiques transdisciplinaires (comme *résultat*, *méthode*, *rôle...*) pour trouver des collocations évaluatives qui incarnent l'évaluation dans l'écrit scientifique. En s'inspirant de la grammaire locale sur les structures syntaxiques existant entre l'adjectif et le nom scientifique, les deux auteurs ont pris généralement deux structures lors de l'extraction des collocations évaluatives du genre $N_{\text{scientifique}} + \text{Adj}_{\text{évaluatif}}$ (Tutin, 2010) :

1. Nom + adjectif, correspond à la première structure de Hunston et Sinclair (①)
2. Nom + verbe + (adverbe) + adjectif, correspond à la deuxième structure de Hunston et Sinclair (②).

À partir de ces deux grammaires locales, les deux auteurs ont extrait les collocations évaluatives en question et les ont classifiées selon le critère de leur redéfinition de l'évaluation. En comparaison avec les deux théories présentées auparavant, il est facile de constater que le travail de Tutin et Cavalla associe les avantages de chacune d'elles : il prend en compte les grammaires locales des collocations évaluatives de Hunston et Sinclair et la dimension sémantique de l'évaluation de Martin.

Bilan

Cette partie a parcouru les théories les plus célèbres concernant l'évaluation dans

les textes. Nous constatons que l'étude sur l'évaluation débute par linguistique systémique fonctionnelle de Halliday. Hunston s'en inspire et oriente ses études sur les grammaires locales tandis que Martin propose la théorie de l'Appraisal pour y ajouter la dimension sémantique de l'évaluation. Mais les deux théories prennent toutes les deux en compte l'affectivité qui est rarement présente dans l'écrit scientifique. Quant à Tutin et Cavalla, elles combinent les deux avantages des leurs deux devanciers en effaçant l'affectivité pour que l'évaluation s'adapte mieux à l'écrit scientifique. De plus, elles appliquent l'évaluation dans le domaine phraséologique pour extraire et analyser des collocations évaluatives dans l'écrit scientifique.

En conséquence, nous aimerions donner à l'évaluation une définition en prenant en compte la définition de Tutin et Cavalla (2008) sur la base de celle de Hunston et Tompson : l'évaluation est un large terme qui couvre généralement l'expression de l'attitude ou de la position de l'auteur ou de l'auteur sur les mots, les points de vue à propos des entités ou des propositions dont il parle, elle s'incarne en linguistique par la comparaison, le jugement non axiologique, le jugement de valeur axiologique et certains marqueurs de modalité.

Chapitre 3. Les collocations évaluatives transdisciplinaires du genre

$N_{\text{scientifique}} + \text{Adj}_{\text{évaluatif}}$

Dans la partie précédente, nous avons parcouru certaines théories et études sur l'évaluation dans le discours. Il est le temps de regarder de plus près des ressources linguistiques que mobilise l'écrit scientifique pour exprimer l'évaluation. Dans le cadre de notre travail, nous prenons comme objet d'étude les collocations évaluatives transdisciplinaires du genre $N_{\text{scientifique}} + \text{Adj}_{\text{évaluatif}}$ et essayons de découvrir comment l'écrit scientifique en français réalise l'évaluation à travers elles.

Au commencement de cette partie, nous invitons à approfondir la distinction entre le lexique général et le lexique transdisciplinaire afin de mettre en lumière l'identité des noms scientifiques transdisciplinaires dans les collocations en question. Notre réflexion va ensuite se diriger vers la précision du terme « collocation », ce qui nous autorisera à mieux analyser des propriétés combinatoires des collocations évaluatives tels que leurs aspects formel et sémantique. Sur la base de ces deux aspects, nous envisagerons de les classer en fonction d'une double considération : le traitement sémantique des noms scientifiques et des adjectifs d'évaluation dans les collocations évaluatives transdisciplinaires du genre $N_{\text{scientifique}} + \text{Adj}_{\text{évaluatif}}$.

3.1. Lexique scientifique transdisciplinaire

Les collocations que nous allons traiter dans ce mémoire comportent deux éléments : les noms scientifiques transdisciplinaires et les adjectifs évaluatifs.

Qu'est-ce que c'est un nom scientifique transdisciplinaire? Les noms scientifiques transdisciplinaires appartiennent au lexique scientifique transdisciplinaire (LST) qui relève de la langue scientifique générale (LSG) mais en constitue une zone plus restreinte. Donc, nous pouvons proposer une figure inclusive des deux termes.

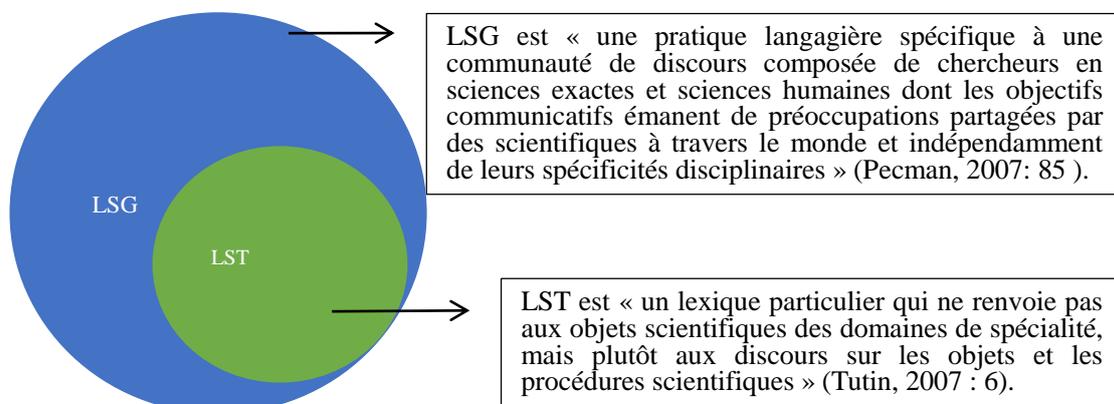


Figure 3 : les relations inclusives entre LSG, LST.

La langue scientifique transdisciplinaire est constituée de lexiques scientifiques transdisciplinaires. La définition du lexique scientifique transdisciplinaire de l'écrit scientifique de Tutin citée en haut dans la figure 3 nous permet de comprendre que le LST, ayant pour fonction de décrire ou de présenter l'activité scientifique, est indépendant des disciplines. S'inscrivant dans le LTS, les noms scientifiques transdisciplinaires sont ainsi un genre de noms qui ne concernent pas la terminologie des domaines spécialisés, mais qui en revanche, renvoient aux discours sur les objets et les procédures scientifiques.

À l'instar de Tutin, Drouin (2007) indique également la transdisciplinarité du LST et souligne en même temps son importance dans l'activité scientifique :

On considère que le LST transcende les domaines de spécialité et présente un noyau lexical commun significatif entre les disciplines. Le lexique scientifique transdisciplinaire n'est pas saillant dans les textes scientifiques dans la mesure où, contrairement à la terminologie, il se rencontre également dans la langue commune. Par contre, il est au cœur même de l'argumentation et de la structuration du discours et de la pensée scientifique. (Drouin, 2007 : 45)

Si le LST joue déjà un rôle indispensable dans la présentation et la description de l'activité scientifique, les collocations transdisciplinaires au sein desquelles se combinent des noms scientifiques transdisciplinaires avec des adjectifs évaluatifs s'avèrent plus

importantes pour garantir la cohésion de la présentation et de la description sur le plan linguistique, parce que la combinaison entre les éléments d'une collocation ne fonctionne pas de façon aléatoire, mais au contraire, d'une façon restreinte qui reflète la stratégie rhétorique de l'évaluation de l'auteur dans l'écrit scientifique.

3.2. La définition des collocations

Pour la définition des collocations, deux acceptations majeures sont marquantes dans le domaine phraséologique. L'acceptation statistique essaie de fournir une explication sur la notion des collocations sous l'angle de la fréquence de cooccurrence des unités lexicales dans le corpus. Les linguistes représentatifs sont Firth, Halliday et Sinclair. L'acceptation sémantique s'attache à décrire les propriétés des collocations en faisant des observations sur la distribution sémantique et syntaxique entre leurs constituants. Dans cette lignée, nous trouvons les études de Mel'čuk (1983, 2003, 2013), de Haussmann (1998, 2006) et de Tutin et Grossmann (2002), de Tutin (2013).

3.2.1. L'acceptation statistique de la notion de collocations

Il nous semble logique de parler tout d'abord d'acceptation statistique de la notion de collocations d'autant plus que le phénomène collocatif des lexèmes a été baptisé sous le nom de « collocation » par le linguiste anglais Firth (Tutin et Grossmann, 2002 : 8). Aujourd'hui, ce terme est largement accepté dans le domaine phraséologique. La définition statistique des collocations a un lien très étroit avec le développement de la linguistique de corpus qui fournit à l'école anglo-saxonne un outil parfait pour faire dégager l'association entre lexèmes qui aide à découvrir le style du discours.

Firth (1957), dans son œuvre *Papers of Linguistics*, introduit la notion de collocation pour compléter sa théorie sur le sens. Il indique l'importance des collocations dans la construction du sens et de la cohésion du texte. Halliday hérite de la théorie de son maître Firth en développant la description statistique des collocations :

collocation is a syntagmatic association of lexical items, quantifiable, textually, as the probability that there will occur at n removes (a distance of n lexical items) from a item x, the item a, b,c... (Halliday, 1961 : 276, cité par Gledhill, 2000 : 8).⁷

Une collocation est considérée ainsi comme une unité lexicale qui co-apparaît avec un mot central. Les unités lexicales en cooccurrence ne s'attachent pas totalement étroitement sans intervalle, au contraire, il est possible qu'une distance se produise entre elles. À sa suite, Sinclair (1970) a renforcé la définition statistique des collocations :

La cooccurrence de deux unités dans un contexte à l'intérieur d'un environnement spécifié. UNE COLLOCATION SIGNIFICATIVE est une collocation habituelle entre deux unités, telles qu'elles se trouvent ensemble plus souvent que leurs fréquences respectives et la longueur du texte dans lequel elles apparaissent ne peuvent le prédire (Sinclair, 1970, p.150, cité et traduit par Williams, 2003, p.37).

Outre la caractéristique liée à la fréquence des collocations, cette définition met l'accent sur la binarité, et la posture côte à côte des constituants collocationnels. Mais au fur et à mesure, Sinclair étend le nombre de constituants d'une collocation de deux aux plus de deux mots, même l'intervalle entre les constituants a été admis (Gonzalez, 2015 : 73).

Même si les apports de la perspective statistique de la définition de collocation permettent aujourd'hui aux linguistes qui travaillent sur les corpus numériques de repérer les collocations, il manque toujours une explication sémantico-syntaxique des collocations.

3.2.2. L'acceptation sémantique de la notion des collocations

La perspective sémantique de la notion de collocation complète la perspective statistique. Elle met un peu de côté la fréquence et l'occurrence des collocations, et prend en même temps position sur la compositionnalité sémantique et syntaxique des unités lexicales.

⁷ Traduction personnelle : La collocation est une association syntagmatique des unités lexicales, textuellement quantifiable selon la probabilité qu'il se produise à n se remuant (une distance d'unité lexicale n) à partir d'une unité x, l'unité a, b, c ...

La définition la plus classique et répandue est sans doute celle de Mel'čuk (1989, 2003, 2013). Sa définition est élaborée en s'inscrivant dans la théorie Sens-Texte qui a pour dessein de décrire « le fonctionnement de la langue par un modèle fonctionnel (=système de règles) qui commence son travail en construisant le sens linguistique à exprimer, duquel il passe aux textes correspondants» (Mel'čuk, 2013 : 1). Pour elle, une collocation est un demi-phrasème⁸ se définissant ainsi :

Collocation (= locution semi-figée)

L'expression **AB** ayant le sens (**S**) est appelée une collocation si et seulement si les trois conditions suivantes sont simultanément remplies :

1. 'S' \supset 'A' ;
2. **A** est sélectionné par le locuteur de façon régulière et non contrainte ;
3. **B** n'est pas sélectionné de façon régulière et non contrainte, mais en fonction de **A** et du sens (**S**) à exprimer. (Mel'čuk, 2003 : 23)

Ici, **A** = la base de collocation, **B** = le collocatif. Dans la collocation *peur bleue*, le lexème *peur* est la base de collocation, et *bleue* fait office de collocatif qui est sélectionné selon la base *peur* pour intensifier la peur. Cette définition met l'accent sur le caractère binaire des constituants des collocations. Ce caractère ne manifeste non seulement dans la forme des collocations (la binarité catégorielle), mais aussi dans le fonctionnement des constituants des collocations (la binarité fonctionnelle).

La binarité catégorielle indique le fait que la collocation regroupe souvent deux unités lexicales de catégories différentes (Tutin, 2013, 5). Par exemple, nos collocations du type $N_{\text{scientifique}} + \text{Adj}_{\text{évaluatif}}$ mobilisent deux types de constituants : les noms scientifiques et les adjectifs évaluatifs.

⁸ Un phrasème est un « énoncé multilexémique non libre » (Mel'čuk, 2013:1), c'est à dire « une configuration de deux ou plus lexèmes syntaxiquement liés» (Ibid) de façon non libre. Ce terme est privilégié par Mel'čuk, mais dans le domaine phraséologique, nous confrontons souvent à plusieurs appellations pour la collocation : collocation (largement acceptée par des linguistes), le demi-phrasème, locution semi-figée etc.

La binarité fonctionnelle est constituée par la dissymétrie au niveau du fonctionnement des constituants de la collocation. Un constituant fonctionne comme base qui garde son sens habituel et l'autre comme le collocatif qui est mobilisé souvent par le premier (Ibid). Dans notre cas, nous préférons considérer les noms scientifiques comme la base, car ils sont choisis précisément par le locuteur lors de sa rédaction pour bien indiquer l'objet ou le processus de l'activité scientifique qu'il veut décrire. En tant que bases, les noms scientifiques gardent leur sens habituel. Si les noms scientifiques remplissent la fonction de la base, les adjectifs évaluatifs font office de collocatif, qui est sélectionné par le locuteur en fonction de la base et du sens à exprimer. La binarité fonctionnelle connote en réalité une autre dissymétrie sémantique tout comme ce qu'explique la définition des collocations de Hausmann et Blumenthal (2006) :

La base est un mot (plus précisément l'acception d'un mot, appelée aussi « lexie ») que le locuteur choisit librement parce qu'il est définissable, traduisible et apprenable sans le collocatif. Le collocatif est un mot (ou l'acception d'un mot) que le locuteur sélectionne en fonction de la base parce qu'il n'est pas définissable, traduisible ou apprenable sans la base (Hausmann et Blumenthal, 2006 : 4).

Cette précision sémantique des constituants des collocations met en accent sur la prédictibilité du sens du collocatif. Comme indiqué plus haut, le collocatif est souvent considéré comme difficile à prédire. Néanmoins, dans notre cas, les collocations évaluatives en question ne respectent pas cette règle à cent pour cent. Nous pouvons voir que pour la plupart des adjectifs évaluatifs, leur sens est bien prédictible car qu'ils gardent généralement leur sens habituel. Ce n'est que pour certains adjectifs polysémiques que la spécification de leur sens doit s'analyser à l'aide d'un contexte précis.

Bilan

Cette partie examine les différentes définitions des collocations. La définition statistique de l'école fonctionnelle systémique met en avant la cooccurrence des constituantes des collocations dans le corpus. Les deux approches statistique et sémantique se complètent surtout pour notre objectif didactique. En outre, l'aspect statistique fournit un accès informatique très intéressant pour l'extraction des collocations dans les corpus

numériques. Les définitions sémantiques ont pris leur essor dès l'instant où Mel'čuk et Hausmann ont précisé les critères de reconnaissance. En conséquence, pour ce travail, nous conservons les deux aspects que sont la fréquence et la sémantique pour la définition de la collocation telle que nous voulons l'aborder dans notre projet à visée didactique. Nous jugeons qu'une série d'unités lexicales peut constituer une collocation si elle peut répondre à l'un des critères suivants :

1. On la trouve régulièrement associée dans le corpus.

2. Elle est la combinaison des deux lexèmes, dont l'un est la base et l'autre est le collocatif. La base est choisie selon le locuteur et garde son sens habituel lorsque le collocatif est sélectionné en fonction de la base. Son sens est parfois prédictible à partir de son sens habituel ou parfois difficilement prédictible.

Quant à nos collocations $N_{\text{scientifique}} + \text{adj}_{\text{évaluatif}}$, elles sont statistiquement présentes dans le corpus : nous pouvons constater les deux constituants en cooccurrence dans des corpus. Plus précisément, nous considérons les noms scientifiques comme bases qui gardent leurs sens habituels, les adjectifs évaluatifs comme collocatifs dont le sens est prédictible du fait qu'ils gardent également leur sens habituel avec précision sémantique en cas de polysémie.

3.3. Les collocations transdisciplinaires évaluatives du genre $N_{\text{scientifique}} + \text{Adj}_{\text{évaluatif}}$: des caractéristiques principales

Les collocations transdisciplinaires évaluatives ajoutent aux collocations transdisciplinaires une autre dimension concernant la fonction rhétorique et pragmatique. Nous voudrions rappeler ici la définition de l'évaluation dans le discours scientifique de Tutin et Cavalla (2008) afin de mieux cerner les collocations en question. L'évaluation a pour objectif d'exprimer l'attitude ou la position de l'auteur ou de l'auteur sur les mots, les points de vue à propos des objets, des processus scientifiques, ou des énoncés. Les collocations transdisciplinaires évaluatives sont donc des collocations qui franchissent la limite des disciplines pour exprimer l'attitude ou le point de vue de l'auteur sur l'objet, le processus de l'activité scientifique, ou des énoncés.

Du point de vue sémantique, ce genre de collocations transdisciplinaires s'inscrivent partiellement dans les définitions sémantiques de Mel'čuk et Hausmann : en ce qui concerne la binarité, elles sont composées du nom scientifique en tant que base et de l'adjectif d'évaluation servant de collocatif. La base porte sur l'objet ou le processus scientifique que le locuteur veut évaluer, le collocatif est choisi en fonction de la base en prenant compte de la sémantique évaluative que le locuteur veut exprimer. Dans notre cas, l'adjectif évaluatif, en tant que collocatif, prend souvent son sens habituel tout comme la base. À cet effet, le sens global est prédictible pour la plupart des collocations évaluatives.

Du point de vue syntaxique, les constituants d'une collocation en question établissent deux patrons avec des ajouts variés :

① Patron canonique / syntagme nominal où l'adjectif fait fonction d'épithète : Nom scientifique + adjectif d'évaluation; les constituants (en gras) sont en occurrence côte à côte dans la phrase. Ici, l'adjectif d'évaluation (le collocatif) fonctionne en tant qu'épithète pour déterminer le nom scientifique (la base).

Ex : Le **rôle principal** des CPP est d'examiner le protocole d'étude, les modalités d'information des patients afin de recueillir leur consentement éclairé, et les modalités de suivi pour s'assurer des précautions prévues pour veiller à la sécurité des malades durant le déroulement de l'essai. (Corpus Scientext, Médecine)

② Patron verbal où l'adjectif fait fonction d'attribut : nom scientifique + verbe support + adjectif d'évaluation; les constituants (en gras) sont séparés cette fois par un verbe support (en italique) qui modifie le rôle sémantique/fonction syntaxique des constituants mais pas leur sens, de plus, le verbe support met en avant l'adjectif. L'adjectif d'évaluation (le collocatif) travaille comme l'attribut lorsque la base devient le sujet.

Ex : Les **résultats sont satisfaisants**, le taux d'économie de saisies est d'environ 55 % pour une liste de 9 propositions et d'environ 50 % pour une liste de 5 propositions. (Corpus Scientext, TAL)

③ Des ajouts variés dans les deux patrons syntaxiques tels que l'adverbe : nom scientifique + adverbe + adjectif évaluatif, les constituants (en gras) ne se sont pas accolés, mais séparés par un adverbe (souligné). L'adjectif d'évaluation (le collocatif) remplit toujours la fonction d'épithète. L'adverbe renforce ou affaiblit la sémantique évaluative du

collocatif, mais ne modifie pas son sens. Comme l'exemple suivant :

Ex : Dans la perspective des développements futurs des moteurs à combustion offrir un modèle et une méthode expérimentale fiable de caractérisation de gaz résiduels est un **résultat très important**. (Corpus Scientext, Mécanique)

Ou soit nom scientifique + verbe support + adverbe + adjectif évaluatif. Dans ce cas, la distance entre les deux constituants (en gras) semble plus grande à cause de l'insertion simultanément des verbes (en italique) et adverbes (souligné). L'adjectif d'évaluation fait office toujours de l'attribut, l'adverbe contribue également à nuancer le degré d'évaluation lors que le verbe support dirige l'attention de lecteur vers l'adjectif d'évaluation.

Ex : Pour les catégories intermédiaires dont la **fonction** sémantique *est moins importante*, comme les catégories Pronom et Adverbe, la saisie d' environ 1 caractères en moyenne suffit à le faire apparaître dans la liste des propositions. (Corpus Scientext, TAL)

L'observation sur les variations syntagmatiques des collocations $N_{\text{scientifique}} + \text{Adj}_{\text{évaluatif}}$ nous amène à faire attention aux éléments périphériques autour des collocations lors de son application réelle dans l'écrit scientifique; ce qui enrichit sans doute la didactique des collocations transdisciplinaires évaluatives. En parallèle, sa caractéristique sémantique en matière de compositionnalité et de prédictibilité donne un accès au sens des collocations, ce qui rend possible le classement sémantique des collocations.

3.4. Le traitement sémantique des constituants des collocations $N_{\text{scientifique}} + \text{Adj}_{\text{évaluatif}}$

3.4.1. Des noms scientifiques dans les collocations en question

Les noms scientifiques sont les bases de nos collocations en question, ils représentent le choix de l'auteur en matière d'objet ou de processus scientifique à évaluer. Souvent, dans le but de mieux favoriser l'application de ce genre de lexique à l'aide à la rédaction, il vaut mieux proposer une analyse profonde des noms scientifiques. Dans le domaine TAL (Traitement automatique des langues), le traitement sémantique à l'aune de l'analyse distributionnelle du lexique a beaucoup contribué à l'analyse et la catégorisation du lexique. Cette analyse distributionnelle est une méthode d'exploitation du corpus qui vise à catégoriser des lexiques sémantiquement proches en fonction des relations de dépendance syntaxiques similaire (Baneyx, 2005). Au début, cette méthode de

catégorisation s'applique principalement au traitement sémantique des lexiques terminologiques, tels que ceux du droit effectué par Bourigault, ou de l'immunologie de Harris etc. Puis Tutin (2007) découvre l'intérêt de l'appliquer au traitement sémantique des lexiques scientifiques transdisciplinaires des écrits scientifiques dans le but d'aider « à la rédaction en Français Langue étrangère » (Tutin, 2007 : 283). Dans le but d'évaluer la performance en matière de traitement sémantique de l'application des méthodes distributionnelles machinales aux lexiques transdisciplinaires, l'auteur propose dans un premier temps un classement manuel, applique dans un deuxième temps respectivement deux méthodes distributionnelles machinales pour obtenir les classements des noms scientifiques transdisciplinaires et dans un dernier temps, les comparer avec le classement manuel. Étant donné que les méthodes distributionnelles machinales donnent à la fin des résultats partiellement concluants, nous préférons prendre ici le classement manuel qui paraît plus fiable. De ce fait, nous aurons sept classes majeures :

1. Les noms de processus de l'activité scientifique : communication application, choix...
2. Les noms d'objets construits par l'activité scientifique : approche, argument, concept, conception, démarche...
3. Les noms d'observables de l'activité scientifique : cas, donnée, échantillons, exemple, facteur...
4. Les noms de supports de la rédaction scientifique : article, chapitre, conclusion, document, figure...
5. Les noms de caractérisation : caractère, caractéristique, différence, difficulté, fonction...
6. Les noms d'acteurs de l'activité scientifique : auteur, chercheur...
7. Les noms de relation logique : but, cause, conséquence, corrélation, effet, influence, liaison, lien, rapport, relation ...(Tutin, 2007 : 285)

Les sept classes manifestent des différences en matière de nature, d'association préférée des verbes. Les noms de processus de l'activité scientifique sont des seuls noms extensifs qui préfèrent se produire en cooccurrence avec des locutions prépositionnelles indiquant le temps telles que *lors de, durant, pendant, au cours de...*; des verbes phasiques⁹ comme *faire, commencer, débiter*, etc. Les noms d'objets construits par l'activité scientifique et d'observables de l'activité scientifique sont tous non extensifs, les premiers se combinent souvent avec des verbes *élaborer, construire* alors que les derniers ont un faible pour le verbe auxiliaire *être* et d'autres verbes comme *analyser, examiner*.

⁹ Les verbes phasiques sont un genre de verbes pour décrire les phases de l'action, allant du début jusqu'à l'achèvement.

Les noms de supports de la rédaction scientifique et ceux de relation logique sont également des noms non extensifs, leur différence réside sur le fait que les premiers comprennent des noms concrets et abstraits tandis que les derniers ne sont qu'abstrait. Quant aux noms de caractérisation, ils montrent une propriété intensive à se combiner avec l'auxiliaire *avoir* et des adjectifs. Les noms d'acteurs de l'activité scientifique sont des noms humains qui dirigent l'activité scientifique. Ils se combinent généralement avec des verbes d'activité scientifique tels qu'*examiner, observer...*

Il faut signaler que certains noms scientifiques peuvent entrer en même temps dans plusieurs classes. Le nom scientifique polysémique *question* peut être vu comme le nom de l'objet construit ou de l'observable de l'activité scientifique. La distinction se fait à l'aide d'un contexte bien concret.

3.4.2. Des adjectifs évaluatifs dans les collocations en question

Les adjectifs d'évaluation sont des collocatifs des collocations $N_{\text{scientifique}} + \text{Adj}_{\text{évaluatif}}$, ils accordent aux bases des collocations une sémantique évaluative afin de faire véhiculer le point de vue de l'auteur sur la base. Les premières études portant sur les adjectifs d'évaluation ont été réalisées par Kerbrat-Orecchioni (1980) en abordant la subjectivité dans l'énonciation. Suite à sa théorie, les adjectifs d'évaluation, se subdivisant en adjectifs axiologiques et non axiologiques, rejoignent avec les adjectifs affectifs la catégorie des adjectifs subjectifs qui se dissocient des adjectifs objectifs.

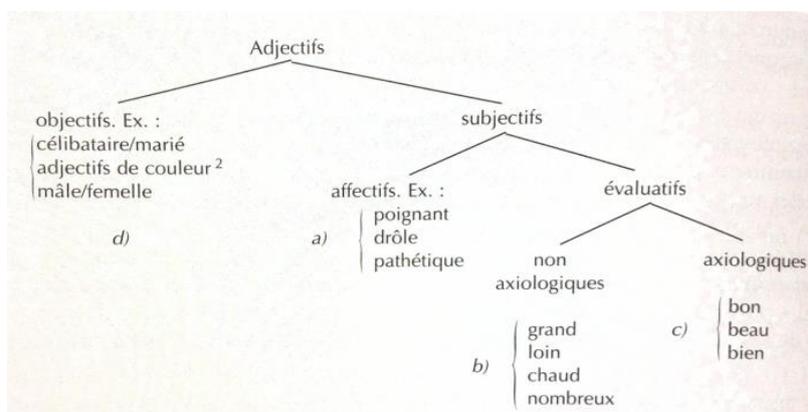


Figure 4 : typologie des adjectifs de Kerbrat-Orecchioni (1980 : 84)

Kerbrat nous fait comprendre que les catégories des adjectifs subjectifs en les

définissant ainsi :

- a) Les adjectifs affectifs décrivent les émotions du locuteur envers un objet.
- b) Les adjectifs évaluatifs non axiologiques n'impliquent pas de jugement de valeur, excluent l'engagement émotionnel, et mettent l'accent sur la quantité et la qualité de l'objet. Ils renvoient à la qualité de l'objet.
- c) Les adjectifs axiologiques s'engagent à évaluer la valeur de l'objet avec la polarité positive ou négative. Ils renvoient aux propriétés de l'objet.

Tutin et Cavalla (2008) s'inspirent de cette typologie pour en élaborer une nouvelle. En premier lieu, les deux auteurs ont exclu la classe des adjectifs affectifs en considération du simple fait que l'écrit scientifique ne met pas en jeu des affections. Contrairement à Kerbrat qui met de côté les adjectifs modaux, Tutin et Cavalla jugent nécessaire de les inclure dans le champ des adjectifs évaluatifs, eu égard à la fonction rhétorique remplie par des adjectifs modaux pour exprimer l'attitude de l'auteur à propos de l'énoncé. La modification ne se limite pas uniquement au niveau d'inclusion ou d'exclusion des catégories primaires, elle touche pareillement aux sous-catégories au sein des adjectifs évaluatifs: à part des adjectifs axiologiques qui apportent un jugement sur la valeur, les adjectifs non axiologiques connaissent une opération d'extension. La comparaison, le jugement au sens large sur une norme comme *grand*, *important*, *récent* ont été incorporés dans cette sous-catégorie. En résumé, nous avons un tableau récapitulatif :

Adjectifs d'évaluation		
Adjectifs axiologiques	Adjectifs non axiologiques	Adjectifs modaux
Jugement sur la valeur Ex : satisfaisant, pertinent, intéressant, bon, mauvais... Polarité : + /-	La comparaison, le jugement au sens large en relation à une norme. Ils se discriminent encore en sept sous-classes : 1.Degré (quantité) et intensité : forte/faible... 2.Comparaison :différente/similaire... 3.Importance : central/fondamental/essentielle... 4.Complexité : complexe... 5.Nouveauté (+/-): classique/nouveau...	Évaluer un énoncé : possible, certain...

	6.Temporalité : antérieur/récent...	
	7.Autre : spécifique, inverse...	

Tableau 4: Nouvelle typologie des adjectifs d'évaluation de Tutin et Cavalla (2008)

Bilan

Les collocations évaluatives s'inscrivent dans la famille des collocations scientifiques transdisciplinaires tout en manifestant une fonction rhétorique et pragmatique spécifique : évaluer l'objet, le processus de l'activité scientifique ou bien un énoncé. Plus précises, les collocations du genre spécifique $N_{\text{scientifique}} + \text{Adj}_{\text{évaluatif}}$ constituent un moyen particulier sur le plan linguistique pour exprimer l'évaluation dans l'écrit scientifique. Les recherches sur leurs constituants (des noms scientifiques et des adjectifs évaluatifs) permettent de les classer raisonnablement en vue d'analyse et de leur enseignement. Pour les extraire des corpus larges, nous avons besoin d'introduire la notion de linguistique de corpus qui sera développé dans le chapitre suivant.

Chapitre 4. La linguistique de corpus pour extraire les collocations

En se situant au carrefour de la linguistique appliquée, de l'informatique et de l'enseignement des langues, la linguistique de corpus – souvent classée dans le TAL : Traitement Automatique des Langues – représente une approche d'analyses des données linguistiques fondées sur des corpus numériques dans le but de reconstruire les spécificités linguistiques d'une langue en faveur de l'enseignement de cette langue ou de la conception de dictionnaires à l'aide de logiciels informatiques. Dans le champ de la phraséologie, la linguistique de corpus constitue une méthode pour révéler et étudier des phénomènes phraséologiques comme les collocations, les expressions figées, les routines sémantico-rhétoriques dans le discours. Dans notre cas, il s'agit bien des collocations. Cette méthode permet deux aspects complémentaires de l'analyse des collocations : qualitatif et quantitatif. L'analyse qualitative consiste à analyser les phénomènes collocationnels dans un contexte naturel ou réel, l'analyse quantitative fournit les étayages statistiques concernant les occurrences des éléments étudiés à partir du corpus afin de permettre de dégager des régularités et de comparer des différences. Maintenant, nous allons développer les deux aspects et leur importance.

4.1. Sur l'aspect qualitatif : l'importance du contexte

Le développement de la linguistique de corpus a une relation étroite avec le concept du contexte. Dans les années 20, Malinowski, professeur à l'université de Londres, a constaté lors de ses recherches concernant les langues primitives du Pacifique Sud que :

utterance and situation are bound up inextricably with each other and the context of situation is indispensable for the understanding of those words. Exactly as in the reality of spoken or written languages, a word without linguistic context is a mere figment and stands for nothing by itself, so in the reality of a spoken living tongue, the utterance has no meaning except in the context of situation¹⁰. (Malinowski, 1923 : 296-336, cité par Williams, 1999)

¹⁰ Traduction personnelle : L'énoncé et la situation sont liés inextricablement l'un avec l'autre

Nous pouvons dégager de cette citation l'un des éléments très importants pour l'analyse de la langue, qui est devenue par la suite la caractéristique essentielle de la linguistique de corpus : le contexte. Ce dernier joue un rôle crucial dans la constitution du sens des mots: un mot isolé perdrait son sens en dehors du contexte. Influencés par lui, les linguistes de l'école de Londres ont repris son idée et mis en évidence l'importance du contexte dans les études des langues surtout pour le lexique. Firth associe le contexte surtout avec les collocations : il pense que connaître un mot c'est connaître la « company words it keeps » (Firth, 1935 cité par Williams, 2006 : 153), soit les mots en cooccurrence. Ses disciples Halliday et Sinclair ont repris ses idées sur le contextualisme et ont travaillé ensuite sur l'étude contextualisée de la grammaire et du lexique.

Dans le premier ouvrage dédié à la discipline de la linguistique de corpus, *English Collocation Studies*, nous trouvons des idées qui rejoignent celles de l'école de Londres sur le rôle du contexte : le dictionnaire ne permet pas une compréhension sans difficulté du sens contextuel à partir des significations dans le dictionnaire (Teubert, 2009). La signification d'un mot se représente par les définitions données par le dictionnaire, et celles - ci sont à la fois nombreuses et manifestent une propriété polysémique, tandis que le sens contextuel s'appuie largement sur le contexte et il n'est quasiment pas polysémique sauf quand le contexte est ambigu.

Mais qu'est-ce que le contexte? En linguistique, le contexte est « les réalités, linguistiques ou non, qui entourent le texte et sa production » (Costa, 2006 : 2). En linguistique, le contexte est indispensable pour l'interprétation des textes, des phrases, et même des lexies. En d'autres termes, il peut servir à limiter le sens d'une information à une réinterprétation plus restreinte imposée par une situation donnée ou à conditionner la construction du sens par la prise en compte de la relation entre la langue et son extériorité (Rastier 1998 : 99). Par exemple, au niveau de la lexie et de la phrase, les deux peuvent

et le contexte de la situation est indispensable pour la compréhension de ces mots. Comme dans les réalités où les langues parlées ou écrites, un mot sans contexte linguistique est un simple produit et ne représente pas rien par lui-même, donc, dans la réalité d'une langue vivante parlée, l'énoncé n'a d'autre sens que dans le contexte de la situation.

toujours être interprétées sans contexte, mais si elles trouvent dans un contexte donné, le sens de la lexie et l'information portée par la phrase auraient besoin d'une réinterprétation contextuelle. Au niveau du texte, le contexte agit sur la relation entre la langue et les éléments extérieurs de la langue. Dans ce cas, ce n'est plus la lexie ou la phrase qui est en question, mais le passage du texte qui mérite l'étude contextualisée. On compte davantage sur le global que le local.

En résumé, les textes écrits ou oraux et leur situation de production travaillent comme des contextes plus précis pour décrypter les sens de leurs phrases internes et celles-ci à leur tour servent de fondements dans l'interprétation des sens des mots qui les constituent en combinant avec « leurs collocats, des unités pertinentes, grammaticales ou lexicales, qui les entourent » (Teubert, 2009).

Cette petite mise en lumière du concept de contexte et de son importance dans l'interprétation du sens nous permet de comprendre l'idée de base de la linguistique de corpus sur l'analyse qualitative : apporter une nouvelle méthodologie sur l'étude et l'enseignement de la langue. Des textes authentiques permettent de traiter des langues en contexte, ce qui peut « révéler les choix linguistiques opérés par des locuteurs dans des contextes réels » (Williams, 1999).

4.2. Sur l'aspect quantitatif : l'importance de la statistique

Si nous disons que le contexte assume une analyse qualitative interrogeant les mots qui mobilisent avec un lexique donné dans un contexte et désambiguïsant la polysémie, la statistique sur laquelle s'appuie l'analyse quantitative fournit alors les étayages statistiques du phénomène linguistique à étudier pour justifier sa spécificité ou sa généralité dans un grand nombre de textes - les corpus.

Auparavant, les études étaient souvent basées sur des corpus de très petite taille faute de capacité informatique, aujourd'hui, grâce à l'essor des techniques informatiques, l'analyse quantitative recourt davantage à l'ordinateur qui peut regrouper une quantité non exhaustive des textes. En même temps, l'accès aux données électroniques est beaucoup plus facile, ce qui nous conduit à porter notre attention sur la constitution des corpus.

4.3 Constitution des corpus

Dans la linguistique de corpus, la constitution des corpus s'avère très importante si l'on veut mener une recherche stricte et relativement plus fructueuse. Comme le développement de l'informatique donne accès à une quantité énorme de textes écrits et oraux, et que les capacités de stockage des données s'élèvent de façon impressionnante, la constitution d'un corpus pertinent est devenue un enjeu crucial. Avant de parler des critères conditionnant la sélection des données, nous voudrions nous attarder sur la notion de corpus qui nous paraît inévitable pour aborder sa constitution.

4.3.1. La notion de corpus

Nous allons reprendre la définition du corpus de Sinclair : du point de vue linguistique, le corpus est

a collection of pieces of languages that are selected and ordered according to explicit linguistic in order to be used as a sample of the language¹¹ (Sinclair, 1996, p4).

Dans ce propos, trois points essentiels pour la constitution d'un corpus émergent :

- Un but linguistique précis
- La sélection en fonction des critères linguistiques explicites
- La collection des données (Williams, 1999)

Le but linguistique dont se sert ce corpus peut aussi s'entendre comme les hypothèses de recherches que veut vérifier le concepteur du corpus. Il constitue la prémisse majeure pour élaborer des critères de sélection. Sans objectif précis, nous ne pouvons pas savoir quel type et quel genre de données recueillir car différents corpus peuvent avoir des utilisations très variées en fonction de leur contenu. Le corpus PFC regroupe des textes oraux de divers lieux géographiques pour permettre de mener des recherches sur les variétés du français oral. Le corpus Frantext s'intéressent plutôt à la littérature française, les textes collectés relèvent donc des œuvres littéraires. Tous ces exemples indiquent la nécessité de « créer un échantillon raisonné de la langue » (Ibid). Pour ce faire, au lieu d'archiver simplement des données, la sélection doit se faire en fonction des critères plus

¹¹ Traduction : une collection de données langagières qui sont sélectionnées et organisées selon des critères linguistiques explicites pour servir d'échantillon au langage.

précis qui répondent aux fins de recherches. Il est évident que la lexicographie du français général et l'étude sur un but plus précis telle que l'analyse des éléments phraséologiques dans les écrits scientifiques ne requièrent pas toujours la même collection de données.

4.3.2. Les critères de la sélection des données pour construire le corpus

En vue d'aller plus loin, nous nous attardons maintenant sur le deuxième point - la sélection des données en fonction de critères précis. Lors de la constitution du corpus, il est nécessaire de prendre en compte le type, le genre, la quantité des textes et le nombre des mots. Principalement, deux genres de corpus sont fréquents : le mégacorpus et le corpus spécialisé. Le premier, nommé aussi le corpus de référence, fonctionne souvent comme une base pour la lexicographie qui envisage de fabriquer des dictionnaires fondés sur les langues authentiques ou pour mener des recherches sur les généralités d'une ou plusieurs langues (corpus monolingue ou parallèle). Il est constitué d'une très grande variété de textes qui couvre presque tous les domaines, tous les genres, allant de la littérature jusqu'aux sciences naturelles, de l'écrit scientifique à la poésie, de manière à répondre aux fins plus générales de la lexicographie. Au contraire du mégacorpus, le corpus spécialisé est créé par les linguistes pour un objectif plus précis. Le critère de la sélection des données devient ainsi plus strict afin de garder une homogénéité du corpus qui garantit la précision de la recherche menée. Martin (1997) a signalé deux erreurs à éviter pour construire un corpus :

- « Le syncrétisme de domaines hétéroclites »
- « La confusion des finalités discursives ».

La première erreur consiste à mettre en commun les données provenant des domaines divers. Même pour une famille de sciences, les disciplines diffèrent l'une de l'autre. Un bon corpus doit savoir différencier les domaines. La deuxième erreur se manifeste principalement dans l'objectif des données: pour garantir l'homogénéité du corpus, l'objectif des données doit aussi être homogène. Prenons l'exemple des corpus Scientext et Emolex. Le corpus Scientext ne regroupe que des textes scientifiques en plusieurs disciplines avec plusieurs genres de textes, ce qui permet donc de découvrir les

choix linguistiques de la communauté scientifique, tandis qu'Emolex a choisi des textes de la presse et de la littérature dans l'intention de favoriser l'étude et l'enseignement des éléments phraséologiques concernant l'expression de l'émotion.

Une fois l'homogénéité obtenue, il faut voir la taille du corpus qui joue un rôle crucial dans sa construction puis dans les résultats de l'interrogation du corpus.

4.3.3. La taille du corpus : grande ou restreinte?

La question de la taille s'avère très importante pour la constitution du corpus. Différentes opinions sur la taille pertinente d'un corpus retenues par des chercheurs ont fait débat dans le champ de recherche en linguistique de corpus. Certains chercheurs comme Geens (1984 : 211) jugent qu'une grande taille semble indispensable pour obtenir des conclusions convaincantes. Mais est-ce qu'une grande taille est vraiment toujours adéquate pour tout objectif de recherche ? Peut-être pour la lexicographie, oui, mais pour un corpus spécialisé, certains linguistes affirment que la taille de 500 000 mots provenant de plusieurs textes serait satisfaisante (Williams, 1999). Bien que ce chiffre soit précis, nous ne pouvons pas oublier que ce seuil n'est pas obligatoire, car il s'agit d'une situation idéale. De plus, la réalité ne permet pas toujours d'avoir une situation aussi idéale concernant la quantité de mots dans le corpus à étudier. À ce moment-là, il ne faut pas être entravé par la quantité mais considérer la requête des besoins. N'oublions pas que tant pour les méga-corpus que pour les corpus spécialisés, leur taille est principalement bornée par deux facteurs : le droit de l'accès aux données (les droits d'auteur) et la finalité de recherche du créateur du corpus. Le premier facteur, en tant que force majeure, empêche d'accéder à l'intégralité des données. Le créateur du corpus peut, quant à lui, juger que sa recherche ne nécessite qu'un corpus limité. C'est aussi les raisons pour lesquelles nous ne devons pas nous limiter uniquement à ce seuil de 500 000 mots pour le corpus. Le fait est que, selon nos objectifs de recherche, nous pouvons travailler sur la totalité du corpus ou seulement sur la partie concernée.

4.4. La question de la généralité et de la représentativité du corpus spécialisé

La taille d'un corpus peut varier en fonction des fins de recherches, surtout pour le

corpus spécialisé qui est établi pour la plupart du temps avec un but très précis. Le corpus spécialisé est donc susceptible d'avoir une taille plus restreinte par rapport au seuil de nombre de mots (500 000) que nous avons indiqué dans la partie précédente. Cela pose un problème sur la question de généralité et aussi celle de la représentativité.

La généralité d'un phénomène linguistique en linguistique s'explique par l'apparition de ce phénomène dans des textes de nature différente. Dans notre recherche, ce sont des collocations $N_{\text{scientifique}} + \text{adj}_{\text{évaluatif}}$ qui s'emploient dans l'écrit scientifique de différentes disciplines. La représentativité des collocations en question peut aussi être conditionnée par la taille de corpus car plus la taille est grande, plus les occurrences des collocations sont censées être plus présentes. Pourtant, au cas où le corpus serait réduit, si les critères de sélections des données sont bien spécifiques, les conclusions tirées peuvent avoir une transportabilité relativisée. (William, 1999)

Bilan

Dans cette partie, nous avons abordé la linguistique de corpus et l'importance de la constitution du corpus, surtout sur la question de sa taille, de sa généralité et celle de sa représentativité. En guise de conclusion, nous rappelons que d'une part, le besoin exprimé par le courant linguistique de l'école de Londres qui considère nécessaire d'effectuer les analyses du langage dans le contexte, et d'autre part le développement continu des moyens offerts par l'informatique en matière de stockage des données et de puissance de traitement des outils d'interrogation ou d'annotation, ont conduit à la constitution de corpus de taille considérable. Mais cela n'a pas rien tranché sur les idées de la pertinence de la taille du corpus. En conséquence, nous croyons que la constitution du corpus spécialisé doit mettre l'attention sur la sélection des données de manière à garder l'homogénéité du corpus compte tenu des fins de recherches. Si la taille du corpus spécialisé est plus restreinte, il est impératif de comprendre que la généralité et la représentativité ne peuvent pas affirmées sans nuance mais de manière plus relativisée.

Partie II - Matériels et Méthodologie

Chapitre 5. Matériels et méthodologie

5.1 Matériels :

5.1.1 Corpus Scientext

Développé dans le cadre du projet ANR Scientext¹² conçu par le Laboratoire de Linguistique et de Didactique du Français Langue Étrangère et Maternelle (LIDILEM) de l'Université Grenoble Alpes (UGA) en collaboration avec le laboratoire Littérature Langue Société (LLS) de l'université de Chambéry et l'équipe LiCorn de l'université de Bretagne Sud, Scientext est un corpus en deux langues (l'anglais et le français) avec un accès libre pour tout usager. Autant les linguistes qui veulent mener des recherches sur les caractéristiques de l'écrit scientifique en français ou en anglais, que les apprenants de ces deux langues qui ont besoin de vérifier leurs hypothèses en matière d'utilisation de la langue lors de leurs rédactions académiques, peuvent l'utiliser. Dans le but de diversifier les corpus académiques et de proposer plusieurs pistes de recherches, les concepteurs ont collecté quatre sous-corpus : les deux premiers sont respectivement l'écrit scientifique en langue française et anglaise. Le troisième consiste en un corpus d'apprenants de l'écrit scientifique en anglais rédigé par des apprenants français de l'anglais et, le dernier, porte sur des évaluations de propositions de communications d'un colloque. Ces quatre sous-corpus contiennent quatre genres de l'écrit scientifique : des articles de recherche, des communications écrites, des thèses de doctorat et des mémoires d'habilitation à diriger les recherches (HDR).

Au-delà des quatre sous-corpus, différentes requêtes ont également été conçues en fonction de deux thèmes principaux : « le positionnement » (Tutin, 2014 : 11) et « le raisonnement » (Ibid). Le premier permet d'étudier la façon dont l'auteur s'inscrit en tant que sujet de l'écrit scientifique afin d'évaluer des recherches antérieures et de former son propre système de connaissances dans un domaine donné. La deuxième entrée d'analyse sur le raisonnement permet de « retracer son cheminement intellectuel et ce sur quoi il

¹² Le site d'internet : <http://scientext.msh-alpes.fr/scientext-site/spip.php?article1>

s'appuie et les déductions qu'il opère » (Ibid). Les deux thèmes fonctionnent de façon complémentaire dans l'écrit scientifique et leur étude permet de cerner les caractéristiques rhétoriques de l'écrit scientifique. Par ailleurs, le projet Scientext a permis la constitution d'un corpus qui regroupe « des écrits scientifiques variés permettant à la fois de comparer des ensembles de disciplines et des sous-genres scientifiques, comme la thèse ou l'article de recherche » (Tutin et Grossmann, 2009 : 1).

5.1.2. Le mode d'emploi du corpus Scientext : ScienQuest pour la requête des éléments linguistiques - l'exemple du corpus de l'écrit scientifique en français

Dans cette partie, nous allons résumer avec concision le mode d'emploi du site Scientext. Accessible gratuitement en ligne, le site Scientext permet de personnaliser les corpus lors de la requête et propose plusieurs dimensions de recherches grâce à l'interface ScienQuest. La première étape consiste à choisir le corpus. Nous prenons ici l'écrit scientifique en français comme exemple. Le site propose la sélection des textes en fonction de la combinaison de trois aspects. Dans la page dédiée aux textes de l'écrit scientifique en français, trois rubriques nommées respectivement *discipline*, *genre*, *parties textuelles* sont proposées en regroupant chacun des sous-thèmes à détailler au gré du besoin de l'utilisateur. Une fois les textes sélectionnés, le site fournit la possibilité de les mémoriser pour des sessions ultérieures. Voici la capture de l'écran de la page des choix à opérer pour constituer son propre corpus.

The screenshot shows the ScienQuest interface with a navigation bar at the top containing 'Corpus', 'Textes', 'Recherche', and 'Historique'. Below the navigation bar, there are tabs for 'Critères' and 'Liste'. A status message indicates: 'Actuellement votre sélection comporte 595 632 mots (205 textes), sur les 5 063 315 mots (205 textes) du corpus. Cliquez pour voir les détails de la sélection...'. The main content area is divided into three panels:

- Discipline:**
 - Sciences expérimentales:
 - Biologie (744 534 mots, 28 textes)
 - Médecine (97 139 mots, 8 textes)
 - Sciences humaines:
 - Linguistique (1 633 579 mots, 67 textes)
 - Psychologie (528 554 mots, 17 textes)
 - Sciences de l'éducation (1 280 515 mots, 60 textes)
 - TAL (753 025 mots, 17 textes)
 - Sciences appliquées:
 - Mécanique (139 912 mots, 3 textes)
 - Électronique (419 234 mots, 5 textes)
- Genre:**
 - HDR (517 025 mots, 7 textes)
 - Article (348 168 mots, 45 textes)
 - Communication (532 333 mots, 112 textes)
 - Thèse (4 198 966 mots, 41 textes)
- Parties textuelles:**
 - Parties principales:
 - Développement (4 310 279 mots)
 - Introduction (154 732 mots)
 - Conclusion (202 371 mots)
 - Autres parties:
 - Résumé (42 330 mots)
 - Notes de bas de page (249 156 mots)
 - Remerciements (22 090 mots)
 - Annexes (71 878 mots)
 - Avant-propos (4 145 mots)
 - Mots-clés (6 334 mots)

Figure 5 : sélection des textes

En ce qui concerne la requête, la page de recherche offre trois modalités de requêtes dans les textes déjà choisis : les requêtes libre, guidée et avancée. La modalité la plus classique est sans doute la requête libre. En interrogeant les lexiques lemmatisés, l'utilisateur peut trouver le mot dans le corpus simplement en saisissant le mot dans la barre de recherche en choisissant la forme, le lemme ou la catégorie. Pour pouvoir chercher simultanément plus d'un mot, ScienQuest propose une relation syntaxique entre les mots de manière à mieux les circonscrire.

The screenshot shows the ScienQuest search interface. At the top, there is a navigation bar with 'Corpus', 'Textes', 'Recherche', and 'Résultats'. Under 'Recherche', three options are available: 'Sémantique', 'Libre', and 'Avancée', with 'Libre' being the selected one. Below this, there are two input boxes for 'Mot 1' and 'Mot 2'. Each box contains three fields: 'Forme', 'Lemme', and 'Catégorie'. To the right of these boxes is a plus sign button. Below the input boxes is a section titled 'Relations syntaxiques' with a dropdown menu and a plus sign button. At the bottom of the interface, there is a 'Chercher' button, a search count of '0 occurrences', and two limiters: 'Limiter la recherche à 1 000 000 mots' and 'Limiter le contexte à 10 mots'.

Figure 6 : requête libre

La deuxième modalité, qui se fait à l'aide de la grammaire locale, consiste à chercher des données conformées aux sémantiques spécifiques telles que les expressions autour de la formulation d'une hypothèse. Elle permet d'entrer directement dans les thèmes liés au positionnement et au raisonnement de l'écrit scientifique. Cinq fonctions sémantiques essentielles pour construire un écrit scientifique sont fournies : *Auteurs cités* trie toutes les citations dans les textes, ce qui permet de faire comprendre la forme de citation employée par la communauté scientifique française. *Autour des hypothèses* regroupe toutes les formulations et validations d'une hypothèse. *Dénomination* donne une vue d'ensemble sur la manière de dénommer les objets scientifiques. *Propositions propres de l'auteur* fait apparaître les verbes exprimant les choix, l'intention et montrant les résultats et apports scientifiques. Quant à *Évaluation et opinion*, il s'agit d'une option qui aide l'utilisateur à

identifier des verbes ou des adjectifs, voire des adverbes, qui pourraient exprimer l'évaluation ou l'opinion de l'auteur dans l'écrit scientifique. Par exemple, si on choisit *adjectifs d'évaluation* et qu'on lance l'extraction, on peut obtenir tous les adjectifs évaluatifs associés aux noms scientifiques.

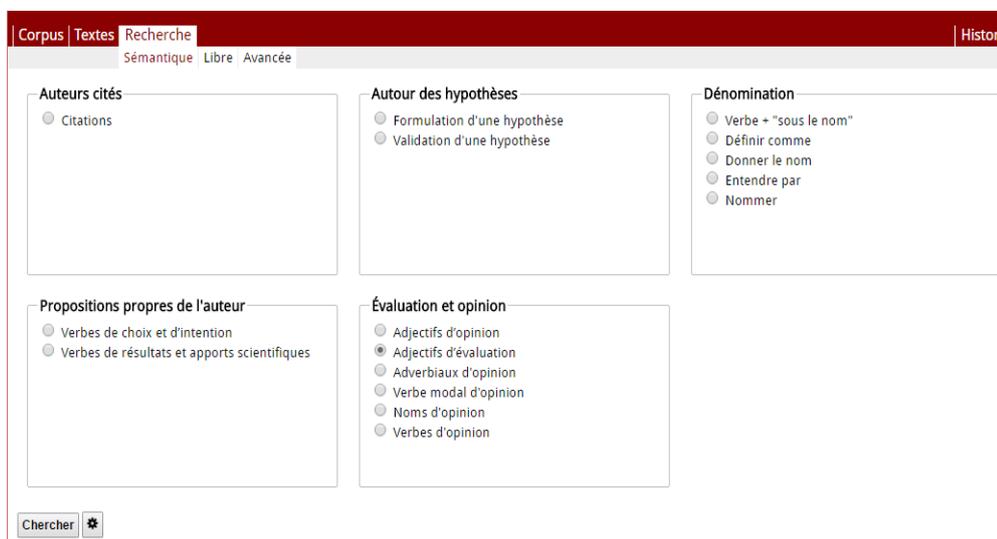


Figure 7 : requête guidée (sémantique)

La dernière modalité, à savoir la requête avancée, répond aux besoins des usagers expérimentés qui voudraient élaborer leur propre grammaire locale dans le but de mieux préciser la sémantique du lexique ou l'expression à chercher.

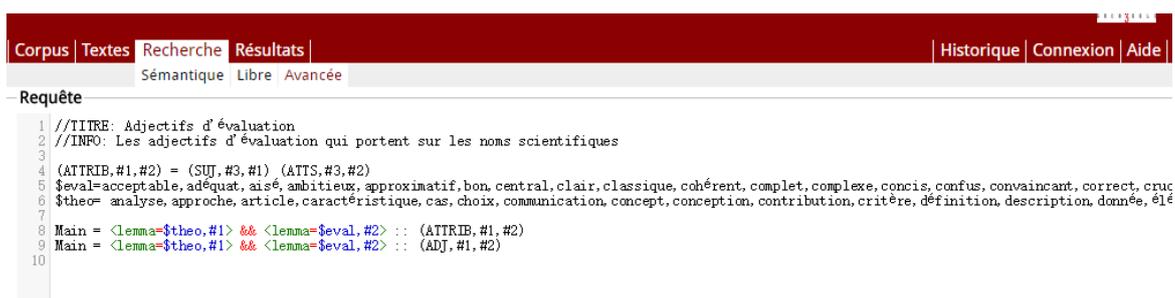


Figure 8 : la requête avancée

5.1.3. Le traitement statistique des données du site ScienQuest

Le site Scienquest ne se contente pas de s'arrêter juste sur le développement des modalités de requêtes, il intègre en parallèle un système de traitement statistique des données. Arrêtons-nous sur la page consacrée aux résultats, nous pouvons trouver deux

genres de résultats.

1. Le choix *Concordances* liste toutes les occurrences des mots à trier avec leur contexte proche et permet de l'élargir à un contexte d'ordre de 200 mots, comme ce que montre l'image ci-dessous :

#	Contexte gauche	Occurrences	Contexte droit	Texte
1	observé . Aussi peut -on dire que par principe , un	modèle est erroné	: on pourra toujours montrer qu' il y a un	#1 - HDR - Sciences de l'éducation - Introduction
2	cela n' a donc aucun intérêt de montrer qu' un	modèle est erroné	: on sait a priori qu' il l' est dans une certaine mesure .	#1 - HDR - Sciences de l'éducation - Introduction
3	les élèves vont pouvoir calculer un certain nombre de termes ; mais le	phénomène qu' on observe à la calculatrice est moins facile	à interpréter que dans le cas de Pn : en effet la suite semble	#1 - HDR - Sciences de l'éducation - Développement

Afficher

Contexte
 Arbre syntaxique

qu' à partir d' un rang à trouver , la valeur de Pn dépasse la capacité de la calculatrice . Les conjectures qui pourront être faites à partir de ces constatations sont au nombre de deux : soit Pn dépasse toute capacité supposée de n' importe quelle calculatrice , c'est-à-dire qu' elle n' a pas de maximum ; soit Pn a un maximum , à déterminer .
 b) Pour An
 Pour l' aire les élèves vont pouvoir calculer un certain nombre de termes ; mais le phénomène qu' on observe à la calculatrice est moins facile à interpréter que dans le cas de Pn : en effet la suite semble se stabiliser et devenir constante . Cette affirmation peut cependant être contestée grâce au milieu matériel , puisqu' il est clair qu' à chaque étape , on rajoute des triangles , même s' ils sont de plus en plus petits : l' aire est donc strictement croissante . Le premier calcul de An doit d' ailleurs confirmer cette propriété , puisque ce calcul fait apparaître une somme de termes tous positifs , un terme étant ajouté à la somme à chaque étape . Le milieu matériel et objectif fournit donc une rétroaction qui permet de mettre en doute l' apparente constance de la suite des

Isabelle BLOCH
 QUELQUES APPORTS DE LA
 THEORIE DES SITUATIONS A LA
 DIDACTIQUE DES
 MATHEMATIQUES DANS

Figure 9 : Les concordances

2. Le choix *Statistiques* consiste à fournir des données, des paramètres et à faire des comptabilisations sous sept aspects différents : la forme, le lemme, la catégorie, les traits morpho-syntaxiques, la partie textuelle, le genre et la discipline. Le tableau est téléchargeable pour un traitement éventuel plus précis.

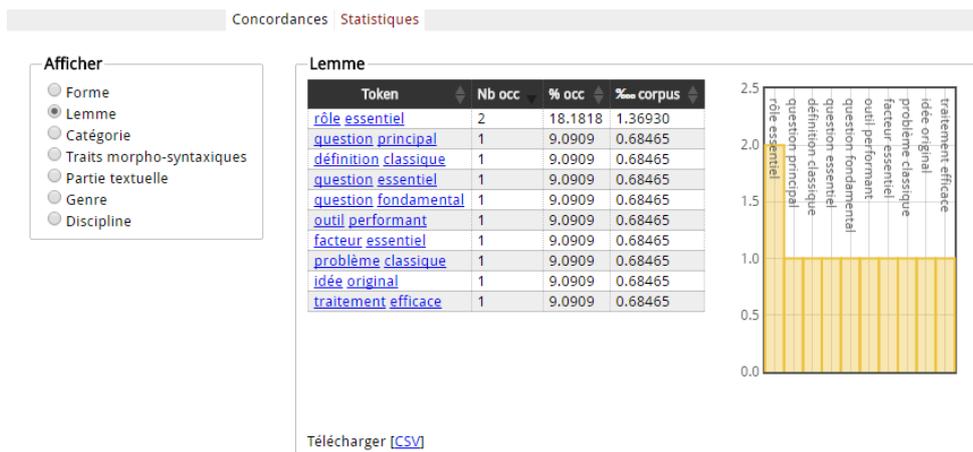


Figure 10 : le résultat statistique du ScienQuest

En résumé, le site ScienQuest offre trois modalités de requêtes qui essaient de satisfaire la diversité des usagers tout en proposant la possibilité de personnaliser le corpus selon les besoins. Simultanément, les usagers disposent d'un accès aux concordances, permettant d'étudier la phrase, et un autre accès aux statistiques, fournissant une vision globale des occurrences de l'objet à analyser.

5.2. La constitution du corpus de notre recherche

5.2.1. Le corpus de base

Notre travail s'inscrit dans l'étude linguistique de corpus en français afin de dégager les spécificités de l'écrit scientifique en matière de préférence des collocations évaluatives ($N_{\text{scientifique}} + \text{Adj}_{\text{évaluatif}}$) des Français natifs. Avec cette considération, le corpus qui nous intéresse ici est le corpus de l'écrit scientifique en français établi par LIDILEM. Ce corpus se divise en trois familles de disciplines :

- a. Les sciences expérimentales rassemblent la biologie et la médecine
- b. Les sciences humaines regroupent quatre disciplines : la linguistique, la psychologie, les sciences de l'éducation et le traitement automatique des langues (TAL) ;
- c. Les sciences appliquées réunissent la mécanique et l'électronique.

Vu que notre objectif sera l'étude des collocations évaluatives transdisciplinaires qui

aidera l'apprenant de n'importe quelle spécialité à marquer son évaluation à travers l'emploi des collocations évaluatives, nous allons étudier toutes les disciplines proposées par le corpus Scientext afin d'identifier les adjectifs les plus utilisés dans toutes les disciplines. Notre corpus de base est présenté dans le tableau 5 :

Disciplines	Articles	Communications écrites	Thèses de doctorat	Mémoire d'Habilitation à Diriger des Recherches	Nombre de textes au total	Nombre de mots
Biologie	17	0	11	0	28	744534
Médecine	6	0	2	0	8	97139
Total Sciences expérimentales	23	0	13	0	36	681439
Linguistique	8	47	8	4	67	1633579
Psychologie	1	11	4	1	17	528554
Sciences de l'éducation	0	54	5	1	60	1280515
TAL	13	0	4	0	17	753025
Total Sciences humaines	22	112	21	6	161	4195673
Mécanique	0	0	2	1	3	139912
Électronique	0	0	5	0	5	419234
Total sciences appliquées	0	0	7	1	8	510554

Tableau 5 : corpus de base sur l'écrit scientifique en français

Ce corpus de base constitue notre terrain de travail, nous allons discuter par la suite la sélection des collocations extraites dans ce corpus.

5.2.2. Tentative de sélection des collocations

La sélection des collocations a un lien très étroit avec la constitution du corpus à étudier surtout dans le cas où on doit restreindre la liste des collocations et la taille du corpus afin de faciliter des analyses dans un temps limité. Dans cette partie, nous tentons

de présenter nos principes de sélection face au grand nombre de collocations rencontrées dans le corpus original proposé par la requête sémantique dans Scienquest.

5.2.2.1. La forme ou le lemme

Une requête permet une distinction entre la forme et le lemme d'une collocation. Dans notre travail, nous ne recensons que des collocations en lemme, sans considérer le paradigme morphologique des collocations, et leur forme neutre sera comprise dans la liste dans un premier temps. Par exemple, la collocation *résultat intéressant*, est le lemme que nous allons garder dans notre liste et toutes les occurrences des différentes formes telles que *résultats intéressants* seront comptées dans le nombre d'occurrences de sa forme neutre. Les listes des collocations en lemme peuvent fournir le nombre d'occurrences des collocations recensées. Vu que les listes en lemme fournies par ScienQuest ne prennent que l'ordre de $N_{\text{scientifique}} + \text{Adj}_{\text{évaluatif}}$ pour toutes les collocations, même si certains adjectifs se trouvent devant le nom lors de leur utilisation, tel est le cas de la collocation *meilleur traitement*, et que les adjectifs dans les collocations sont tous masculins (forme lemmatisée) en cooccurrence avec des noms féminins comme *fonction essentiel*, nous allons observer de plus près des collocations lemmatisées afin de corriger manuellement certaines erreurs morphologiques.

5.2.2.2. Le seuil de spécificité pour sélectionner les collocations à traiter

Dans un premier temps, nous voulons savoir si nous pouvons prendre en compte toutes les collocations évaluatives $N_{\text{scientifique}} + \text{adj}_{\text{évaluatif}}$ dans l'ensemble du corpus de l'écrit scientifique français. Ce corpus comporte 5 063 315 mots (205 textes). Nous lançons la requête sémantique en choisissant "Adjectifs d'évaluation" et obtenons huit tableaux des collocations. Les chiffres des collocations évaluatives, avec la construction $N_{\text{scientifique}} + \text{Adj}_{\text{évaluatif}}$ dans les huit disciplines, sont récapitulés dans le tableau suivant :

Disciplines	Nombre des textes	Nombre des mots	Nombre d'occurrences des collocations $N_{\text{scientifique}} + \text{adj}_{\text{évaluatif}}$	Nombre des collocations $N_{\text{scientifique}} + \text{adj}_{\text{évaluatif}}$
Biologie	28	595 632	296	132

Médecine	8	97 139	95	49
Linguistique	67	1633 579	993	342
Psychologie	17	528 554	238	130
Sciences de l'éducation	60	1 280 515	712	294
TAL	17	753 025	716	326
Mécanique	3	139 912	441	75
Électronique	5	419 234	322	151
Au total	205	5 063 315	3813	745

Tableau 6 : Récapitulation des nombres des collocations dans les huit disciplines

Néanmoins, ceci ne signifie pas que nous allons prendre en compte toutes les collocations évaluatives extraites, et ce pour deux raisons : le grand nombre de disciplines et de collocations à étudier en raison de la grande taille du corpus et le manque de temps pour notre étude. Nous devons par conséquent élaborer des principes selon lesquels la sélection des collocations sera réalisée afin de limiter le corpus de recherche.

Pour limiter les collocations à étudier, nous faisons appel à la notion du « seuil de spécialité » qui « calculera la spécificité des expressions dans le corpus analysé » (Tutin et Kraif, 2016). Tutin a pris deux seuils de spécificité pour extraire des collocations $N_{\text{scientifique}} + \text{adj}_{\text{évaluatif}}$: en 2008, elle a présenté avec son homologue Cavalla pour les collocations évaluatives $N_{\text{scientifique}} + \text{adj}_{\text{évaluatif}}$ dans l'écrit scientifique linguistique et économique un seuil de spécificité à partir de 15 occurrences pour chaque collocation. Et en 2010, dans son article *Evaluative adjectives in academic writing in the humanities and social sciences*, elle a pris les 15 collocations les plus fréquentes par discipline. Donc, nous avons deux seuils de spécificité à prendre en compte :

* les collocations qui ont un nombre d'occurrences ≥ 15 dans le corpus de chaque discipline

* les 15 collocations les plus fréquentes dans le corpus de chaque discipline

Il n'est pas convenable d'appliquer les deux seuils sans regarder la différence entre la taille du corpus du Tutin et celle du nôtre. Les corpus linguistique et économique de Tutin comportent respectivement 640 693 mots et 661 169 mots.

	Linguistique	Économique
Articles scientifique (Corpus KIAP)	285 881 mots	374 516 mots
Thèses, rapports, manuels	364 812 mots	286 653 mots
Total	650 693 mots	661 169 mots

Tableau 7 : Constitution des corpus de Tutin (2010 : 4).

Pourtant, notre corpus n'a pas la même taille que celui de Tutin. C'est aussi à cause de la taille différente des corpus appartenant aux disciplines que nous ne pouvons pas utiliser ces seuils de la sorte. Il est par conséquent nécessaire d'examiner la faisabilité de l'application de ces deux seuils de spécialité employés par Tutin à nos corpus.

Nous commençons par le premier seuil de spécialité. Généralement, deux méthodologies ont été mobilisées pour vérifier si nous pouvions l'appliquer. La première consiste à utiliser directement ce seuil de spécialité (occurrences ≥ 15), La deuxième méthode permet de convertir ce seuil en ratio afin d'être plus précis. Après l'observation des collocations dans les huit disciplines, nous constatons qu'il y a très peu de collocations qui ont un nombre d'occurrences supérieur à 15.

Disciplines	Nombres des collocations $N_{\text{scientifique}} + \text{adj}_{\text{évaluatif}}$	Nombre des collocations dont l'occurrence ≥ 15
Biologie	132	2
Médecine	49	0
Linguistique	342	7
Psychologie	130	0
Sciences de l'éducation	294	3
TAL	326	2
Mécanique	75	4

Électronique	151	0
En total	745	16

Tableau 8 : nombre de collocations dont l'occurrence ≥ 15

Appliquer le premier seuil de spécialité a donc échoué. Nous voulons voir si le ratio converti à partir du seuil fonctionne. Le ratio est calculé ainsi : 15 (nombre minimal d'occurrences) est divisé par le nombre de mots dans le corpus de Tutin (650 693 pour la linguistique, 661 169 pour l'économique). Ainsi, nous aurons :

$$15 \div 650693 = 0.00002341214 \dots\dots\dots\text{ratio (1)}$$

$$15 \div 661169 = 0.00002268708 \dots\dots\dots\text{ratio (2)}$$

Pour nous assurer que nous pouvons obtenir plus de collocations, nous prenons le taux le moins élevé des deux, soit le ratio (2). Puis, nous allons voir combien de collocations dont le ratio d'occurrences par rapport au nombre de mots dans le corpus est supérieur au ratio (2) et nous obtenons le tableau suivant :

Disciplines	Nombres des collocations $N_{\text{scientifique}} + \text{adj}_{\text{évaluatif}}$	Nombre des collocations dont le ratio d'occurrences \geq 0.00002268708
Biologie	132	2
Médecine	49	16
Linguistique	342	0
Psychologie	130	0
Sciences de l'éducation	294	0
TAL	326	2
Mécanique	75	16
Électronique	151	5
En total	745	41

Tableau 9 : Nombres des collocations dont le ratio d'occurrence ≥ 0.00002268708

Avec cette vérification, nous constatons pour la deuxième fois que le ratio d'occurrences par rapport au nombre de mots dans le corpus n'est pas efficient. En effet,

nous constatons, d'après le tableau 9, que certaines disciplines ne contiennent aucune collocation dont le ratio d'occurrences répond à notre critère.

Dans cette perspective, nous devons renoncer au premier seuil. En ce qui concerne le deuxième seuil qui prend les 15 collocations les plus fréquentes dans le corpus, nous avons aussi détecté des soucis après examen. Nous proposons de prendre l'exemple du corpus de la linguistique et celui de la médecine comme exemples afin d'illustrer nos soucis :

N°	Médecin		Linguistique	
	Collocations	Nombre d'occurrences	Collocations	Nombre d'occurrences
1	facteur principal	9	rôle important	28
2	résultat meilleur	7	proposition principal	23
3	facteur classique	6	structure complexe	21
4	cas nouveau	5	résultat meilleur	18
5	critère majeur	5	problème crucial	16
6	rôle principal	4	objectif principal	16
7	traitement efficace	4	problème classique	15
8	rôle important	4	élément nouveau	14
9	traitement nouveau	3	problème complexe	14
10	rôle central	3	rôle central	13
11	donnée pertinent	2	résultat excellent	13
12	rôle essentiel	2	résultat principal	13
13	donnée insuffisante	2	problème épineux	13
14	résultat significatif	2	problème majeur	12
15	méthode prometteuse	2	élément central	12
16	rôle majeur	2	rôle essentiel	11
17	méthode satisfaisante	1	problème principal	11

Tableau 10 : les 15 collocations les plus fréquentes dans linguistique et la médecine

Le premier souci rencontré porte sur l'inégalité du nombre d'occurrences entre les 15 collocations les plus fréquentes dans chaque discipline. Nous pouvons remarquer que la quinzième collocation dans la médecine (*méthode prometteuse*) n'a que 2 occurrences alors qu'au même rang, dans la linguistique, *élément central* apparaît 12 fois. Il semble que nous pouvons tolérer les collocations qui apparaissent très peu dans certaines disciplines mais pas dans d'autres. Cela ne répond pas à l'exigence de la rigueur des recherches scientifiques.

Le deuxième souci se rencontre au sein de la discipline. Prenons encore la quinzième collocation dans la médecine *méthode prometteuse* comme exemple, nous remarquons dans la liste des collocations de la médecine, la seizième collocation *rôle majeur* a le même nombre d'occurrences que la quinzième, mais nous gardons la quinzième en éliminant la seizième. Cela ne paraît pas logique. Maintenant, on comprend clairement la difficulté de limiter la liste des collocations à analyser, surtout après avoir constaté que les deux seuils ne peuvent pas être appliqués à notre cas. Mais la contrainte liée au grand nombre d'occurrences et au manque de temps nous oblige à restreindre notre nombre de collocations ainsi que le corpus à étudier.

Enfin, nous proposons de regarder la question de la limitation de la liste de collocations dans une autre perspective. Si l'on ne peut pas amoindrir directement le nombre des collocations, nous pouvons en revanche réduire notre corpus pour arriver à l'obtention d'une liste non exhaustive et moins importante.

ScienQuest propose une division du texte en deux grandes parties. La partie principale se divise en trois sous-parties indispensables pour les textes : l'introduction, la conclusion et le développement (le corps principal du texte). Les autres parties concernent davantage le résumé, l'annexe, etc. Nous pouvons choisir une ou deux parties à étudier. Nous décidons de prendre l'introduction et la conclusion comme corpus de recherche pour deux raisons :

1. Tout écrit scientifique accorde une grande importance à ces deux parties du texte. Notons que leur structure est souvent identique alors que, en fonction des

disciplines, le développement peut varier fortement.

2. Après avoir calculé les proportions d'occurrences des collocations dans l'introduction, la conclusion et le développement des huit disciplines, nous trouvons que les collocations évaluatives sont proportionnellement et relativement plus présentes dans l'introduction et la conclusion des textes¹³.

Disciplines	Proportion des collocations évaluatives dans l'introduction	Proportion des collocations évaluatives dans le développement	Proportion des collocations évaluatives dans la conclusion
Biologie	0.00075312	0.00049804	0.00040105
Médecine	0.00323939	0.00107432	0.002290170
Linguistique	0.0086116	0.00067024	0.00093585
Psychologie	0.00049034	0.000500081	0.00108963
Sciences de l'éducation	0.00092807	0.00060319	0.00096321
TAL	0.00098213	0.00108005	0.00185698
Mécanique	0.00601974	0.00293295	0.00956603
Électronique	0.00173461	0.00076837	0.00204180

Tableau 11 : Proportions des collocations évaluatives dans les trois parties textuelles des huit disciplines

En conséquence, nous fixons donc notre corpus de recherche comme ce qui est montré dans le tableau 12 :

Familles de disciplines	Disciplines	Nombre de textes	Nombre de mots dans l'introduction	Nombre de mots dans la conclusion	Nombre de mots au total
Sciences	Biologie	28	14606	57350	71956

¹³ Même si une supériorité très légère du développement de la biologie par rapport à la conclusion et une autre de la psychologie par rapport à son introduction se manifestent, nous pouvons toujours dire que les collocations évaluatives sont relativement plus nombreuses dans l'introduction et la conclusion que dans le développement.

expérimentales	Médecine	8	3087	2757	5844
Sciences humaines	Linguistique	67	31353	44879	76232
	Psychologie	17	30591	21108	51699
	Sciences de l'éducation	60	32325	26993	59318
	TAL	17	30546	26387	56933
Sciences appliquées	Mécanique	3	4153	4286	8439
	Électronique	5	8071	18611	26682
Au total		205	154732	202371	357103

Tableau 12 : Constitution de notre corpus à étudier

Par la suite, nous obtenons nos listes de collocations (16 listes pour huit disciplines), le tableau 13 récapitule le nombre d'occurrences et celui des collocations dans les deux parties des huit disciplines.

Discipline	Partie textuelle	Nombre des collocations	Nombre d'occurrences
Biologie	Introduction	10	11
	Conclusion	20	23
Médecine	Introduction	7	10
	Conclusion	7	8
Linguistique	Introduction	24	27
	Conclusion	36	42
Psychologique	Introduction	13	15
	Conclusion	20	23
TAL	Introduction	25	30
	Conclusion	40	49
Sciences de l'éducation	Introduction	25	30
	Conclusion	22	26
Électronique	Introduction	11	38

	Conclusion	28	14
Mécanique	Introduction	8	25
	Conclusion	17	41
Au total		205	412

Tableau 13 : Nombre des collocations obtenues machinalement

5.2.3. Le traitement manuel pour supprimer les bruits dans la liste de collocations obtenues par traitement informatique

Dans le but d'éradiquer des bruits éventuels, produits par l'extraction automatique, il est nécessaire de vérifier les collocations une par une avec leur contexte. Pour des collocations d'origine anglaise comme *important role* qui est retenue en tant que collocation en français *rôle important*, nous les excluons de nos listes de collocations. En plus des collocations anglaises, nous remarquons aussi quelques fausses collocations qui ont été prises à tort comme de vraies collocations. Dans l'exemple suivant :

Dans le présent texte, il a été question de trois concepts principaux : le travail en équipe cycle, en équipe de collègues ; la dimension affective ; le bien-être professionnel. (Corpus Scientext, Sciences de l'éducation)

Question et *principal* ont été traités comme une collocation, même si l'adjectif *principal* ici a pour fonction de déterminer le nom *concept*. Dans ce cas, nous pensons correct d'accepter la collocation *concept principal* au lieu de *question principale*.

Ces deux exemples expliquent clairement la manière de mettre au point des collocations obtenues automatiquement. Enfin, nous obtenons des listes correctes sans bruit (annexe 01).

Discipline	Partie textuelle	Nombre des collocations	Nombre d'occurrences	Nombre d'occurrences au total
Biologie	Introduction	10	11	32
	Conclusion	19	21	
Médecine	Introduction	7	10	18
	Conclusion	7	8	

Linguistique	Introduction	24	27	69
	Conclusion	36	42	
Psychologique	Introduction	13	15	38
	Conclusion	20	23	
TAL	Introduction	25	30	79
	Conclusion	40	49	
Sciences de l'éducation	Introduction	25	30	56
	Conclusion	22	26	
Électronique	Introduction	11	38	52
	Conclusion	28	14	
Mécanique	Introduction	8	24	65
	Conclusion	17	41	
Au total		312	409	409

Tableau 14 : Nombre des collocations sans bruit

5.3. Les principes de classement des collocations $N_{\text{scientifique}}$ + $Adj_{\text{évaluatif}}$: le modèle et la méthode

Dans cette partie, nous allons présenter les principes que nous retenons pour classer les collocations évaluatives $N_{\text{scientifique}}$ + $Adj_{\text{évaluatif}}$ selon les caractéristiques sémantiques des constituants analysées dans la partie théorique.

Étant donné que notre but final consiste à réfléchir à la didactisation de ces éléments linguistiques, afin d'aider les apprenants non-natifs à la rédaction de leur écrit scientifique, spécialement le mémoire et la thèse, il est nécessaire d'adopter une approche didactique pour classer les collocations à traiter. L'approche onomasiologique s'occupe de l'aspect sémantique ou analogique des collocations, qui favorise leur mémorisation et leur utilisation à bon escient dans le contexte car elle est « pour point de départ un ensemble de concepts plutôt qu'un ensemble de mots » (Picoche, 1992 cité par Cavalla, 2009 : 8). Le principe est de classer des collocations en fonction de la sémantique de leurs constituants. Deux étapes sont donc mobilisées : la première étape consiste à classer les collocations en

fonction des bases (les noms scientifiques) et la deuxième étape à diviser les collocations déjà classées dans la première étape à partir de la sémantique évaluative des collocatifs (les adjectifs évaluatifs).

Nous avons donc deux principes en résumé :

a. À partir de la base (des noms scientifiques transdisciplinaires). Les noms scientifiques peuvent être catégorisés en sept classes selon Tutin (2007 : 285). Mais en regardant de plus près les sept classes, nous trouvons que les noms d'acteur de l'activité scientifique sont des noms humains qui n'entrent pas dans nos collocations. À cet effet, nous proposons de l'exclure de notre catégorie. Ce qui nous conduit à classer des collocations dans six classes :

1) Processus de l'activité scientifique, nous avons des noms scientifiques comme *choix, communication, contribution, application...*

2) Objets construits par l'activité scientifique, nous recensons les noms scientifiques comme *outil, idée, modèle, solution, principe, structure, système...*

3) Observables de l'activité scientifique, les noms scientifiques comme *résultat, exemple, élément, facteur, paramètre...*

4) Supports de la rédaction scientifique, *article, document...*

5) Caractérisation : *rôle, caractéristique, fonction...*

6) Relation logique : *but, cause...*

Il faut signaler ici que certains noms scientifiques sont polysémiques. Dans ce cas, une analyse sémantique avec la mise en contexte sera indispensable. Prenons le nom scientifique *question* comme exemple. Nous avons deux situations :

① *Question* relève des noms d'objets construits par l'activité scientifique qui ont un agent humain.

En nous interrogeant sur la pertinence de ces notions (marquage conscient d' une identité, et la manipulation lexicale) pour le scanto, nous poserons deux **questions (Objet construit)**

principales : quels processus (stratégies) de manipulation lexicale sont les plus importants dans le cas du scamto ? Jusqu' à quel point la " manipulation lexicale " peut -elle révéler des éléments identitaires des locuteurs ? (Corpus Scientext, Linguistique)

② *Question* appartient aux noms d'observables de l'activité scientifique. C'est après l'observation des phénomènes que les questions sur certain aspect sont posées.

En effet, ce traitement chirurgical soulève des **questions (Observable) fondamentales** qui impliquent la compréhension des interactions des réseaux de neurones appartenant aux différentes voies des ganglions de la base, à la fois dans une situation normale mais aussi pathologique. L'une des **principales questions (Observable)** vise à tenter d' élucider l' impact de la SHF du NST sur l' activité des neurones du NST eux -m mais aussi de ses cibles directes et/ou indirectes au sein des ganglions de la base et du cortex .(Corpus Scientext, Biologie)

b. À partir des valeurs sémantiques des adjectifs évaluatifs : dans cette étape, les collocations classées selon la distribution sémantique des noms scientifiques vont être reclassées en fonction des valeurs sémantiques des adjectifs évaluatifs discriminés en trois catégories : modaux, axiologiques et non axiologiques.

1) Les adjectifs modaux qui n'évaluent ni la valeur ni un autre sens comme la complexité, la temporalité ; leur présence a pour objectif de nuancer la tournure ou d'évaluer un énoncé : *développement possible, certain principe...*

2) Les adjectifs axiologiques qui évaluent les valeurs des noms scientifiques et souvent expriment une polarité de valeur, soit positive ou négative : *meilleure solution, étude pertinente, résultat intéressant...*

3) Les adjectifs non axiologiques qui ne prétendent pas à évaluer les valeurs mais plutôt de montrer l'intensité, la nouveauté et les caractéristiques similaires. Ils regroupent plusieurs sous catégories (Tutin, 2010 : 4-5) :

* Degré (quantité) et intensité tels que *hypothèse forte, influence faible...*

* Comparaison : *analyse différente, méthode similaire...*

* Importance : *objectif central, principe fondamental, question essentielle...*

* Complexité : *idée de base, outil complexe...*

* Nouveauté (polarité +/-) : *traitement traditionnel, démarche classique, nouveau concept...*

* Temporalité : *étude antérieure, travail récent...*

* Autre : *objectif spécifique, conclusion inverse...*

Pour des adjectifs polysémiques, nous reprenons également leur contexte pour vérifier leur sens contextuel pour garantir l'exactitude du classement.

En suivant ces deux étapes, nous pouvons obtenir un tableau de classement des collocations pour chaque discipline et à total huit tableaux (annexe 02). Dans chaque tableau, les collocations appartenant à l'introduction seront indiquées en bleu lorsque celles de la conclusion seront indiquées en violet. De plus, les collocations dont l'occurrence dépasse 1 seront suivies par le nombre d'occurrences indiqué dans les parenthèses. Pour des collocations axiologiques, la polarité est aussi indiquée grâce au signe « + » placé après la collocation. Un exemple pour la biologie :

		Processus de l'activité scientifique	Objets construits par l'activité scientifique	Observables de l'activité scientifique	Supports de la rédaction scientifique	Caractérisation	Relation logique
Adjectifs modaux							
Adjectifs axiologique Polarité +/-		Communication riche + (être plus)	Traitement efficace + Outil performant + Idée originale + Outil performant + Modèle cohérent + Traitement efficace +	Résultat intéressant + Bon exemple +			
Adjectifs non axiologiques	Degré(quantité) et intensité		Analyse fine				
	Comparaison						
	Importance	choix crucial	Structure principale Méthode importante (être) Outil essentiel	Question fondamentale Question essentielle Question principale Facteur essentiel Elément majeur Elément important		Rôle essentiel (2) Rôle important (2) Rôle crucial Fonction essentielle	
	Complexité		Système complexe(très)(dynamique)	Résultat difficile (être)			
	Nouveauté Polarité +/-		Définition classique Problème classique Outil nouveau (2)				
	Temporalité						
	Autre						

Tableau 15 : Classement des collocations dans l'introduction (bleu) et la conclusion (violet)

de la biologie

5.4. Les aspects d'analyses.

Le classement des collocations évaluatives du genre $N_{\text{scientifique}} + \text{adj}_{\text{évaluatif}}$ nous permet par la suite de les analyser en détail. Nous inspirant de Tutin (2010) sur les

collocations évaluatives dans l'économie et la linguistique, nous prétendons mener notre analyse selon les deux dimensions quantitative et qualitative.

5.4.1. Analyse quantitative

L'analyse quantitative vise à découvrir les similarités ou les différences entre les huit disciplines, ainsi qu'entre les deux parties textuelles - l'introduction et la conclusion. En premier lieu, nous allons comparer les pourcentages des collocations par rapport au nombre de mots du corpus dans l'introduction et la conclusion de chaque discipline afin de faire émerger la différence de présence des collocations entre les deux parties textuelles ainsi qu'entre les disciplines. En deuxième lieu, en utilisant les huit tableaux de classement des collocations, nous allons observer de plus près la proportion des différentes catégories de collocations dans l'introduction et la conclusion des huit disciplines prises une par une et ce d'une manière plus détaillée, pour que les différences entre disciplines soient plus affinées. Une mise en commun sera établie en dernier lieu.

5.4.2. Analyse qualitative sur la sémantique des collocations et leurs fonctions rhétoriques

Basée sur l'analyse quantitative déjà réalisée, l'analyse qualitative va entrer dans le champ sémantique des collocations qui trahit la stratégie rhétorique des auteurs natifs du français en matière d'évaluation. Plus précisément, nous comptons analyser les sémantiques évaluatives des collocations à travers des adjectifs évaluatifs tout en prenant en compte leur contexte qui nous permet de mieux comprendre le sens de la collocation cible et de rechercher le lien grammatical entre les deux éléments de la collocation. Nous menons, à la fin, une petite enquête sur la transdisciplinarité des collocations en question.

Bilan

Ce chapitre a présenté le corpus Scientext et l'emploi du site ScienQuest avec des études sur le positionnement et le raisonnement dans l'écrit scientifique en français. Grâce aux possibilités de choix libre des corpus, nous sélectionnons les parties introductive et conclusive des huit disciplines comme corpus, et enfin, ScienQuest nous fournit les listes originales des collocations évaluatives par la requête sémantique (en choisissant les

adjectifs évaluatifs). Pour éviter les bruits, les listes obtenues automatiquement ont été vérifiées manuellement. Par la suite, les collocations évaluatives ont été classées en deux étapes selon la base (nom scientifique) et le collocatif (adjectif évaluatif) pour enfin permettre d'accéder aux analyses quantitatives et qualitatives qui sont censées faire découvrir des similarités et des différences disciplinaires, ainsi que les stratégies rhétoriques de l'évaluation derrière l'emploi des collocations évaluatives.

Partie III - Analyses

Chapitre 6. Analyses et résultats

Dans ce chapitre, nous allons mener des analyses quantitative et qualitative ainsi que la comparaison avec les études précédentes.

6.1. Analyses quantitative

6.1.1. Analyse de la fréquence des collocations évaluatives dans les huit disciplines

Il faut rappeler ici que les tailles des sous-corpus (l'introduction et la conclusion des disciplines) ayant de grandes différences, une comparaison limitée au seul niveau du nombre d'occurrences des collocations évaluatives n'aurait donc pas de sens. Nous proposons par conséquent de regarder le rapport entre le nombre d'occurrences et celui de mots.

Dans la première étape, nous regardons les différences en matière de la fréquence des collocations par familles de disciplines :

	Sciences expérimentales	Sciences humaines	Sciences appliquées
Rapport	0.0643%	0.0991%	0.3331%

Tableau 16 : rapport de la fréquence des collocations évaluatives dans les trois disciplines

Les chiffres dans le tableau 16 montrent que l'emploi des collocations évaluatives s'avère différent entre les trois familles : les sciences appliquées emploient plus fréquemment les collocations évaluatives que les deux autres familles de disciplines. Les sciences humaines se trouvent au milieu alors que les sciences expérimentales paraissent les plus restreintes. Dans la deuxième étape, nous invitons à observer la fréquence des collocations évaluatives dans chaque discipline (en étudiant seulement l'introduction et la conclusion).

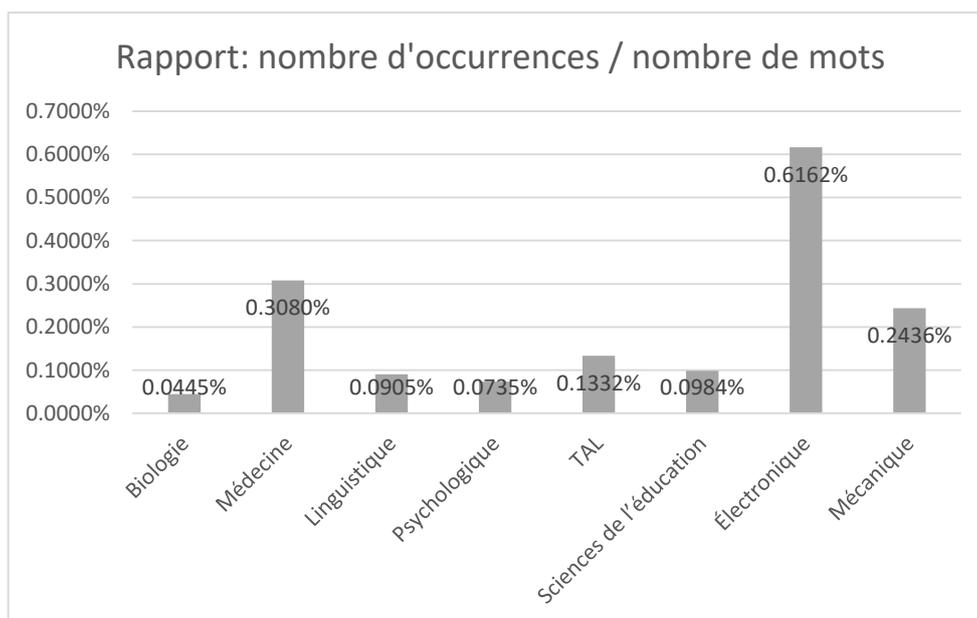


Figure 11 : Fréquence des collocations évaluatives dans les huit disciplines

La figure 11 permet de dégager les constats suivants:

1. La médecine, la mécanique et l'électronique se distinguent des autres disciplines par un nombre important d'occurrences des collocations évaluatives.

2. Les disciplines de la famille des sciences expérimentales manifestent une bipolarité : la biologie représente la discipline utilisant le moins de collocations évaluatives parmi toutes les disciplines étudiées, alors que la médecine occupe la deuxième place juste après l'électronique.

3. La famille des sciences appliquées semble préférer utiliser plus de collocations évaluatives dans l'introduction et la conclusion que les autres disciplines.

4. Excepté la biologie, les autres disciplines des sciences expérimentales et appliquées mobilisent plus fréquemment les collocations évaluatives que l'ensemble des disciplines de sciences humaines. Ce constat va à l'encontre des résultats de Fløttum et al. (2006) sur l'emploi des pronoms personnels dans l'économie et la linguistique et ceux de Stotesbury (2003) sur les résumés des articles de recherches dans trois familles de discipline¹⁴. La recherche de Fløttum montre que l'économie, en tant que discipline

¹⁴ La taxonomie des disciplines chez Stotesbury (2003) n'est pas identique à la nôtre. Les trois familles

appartenant aux sciences sociales, qui est proche des sciences naturelles, garde une objectivité plus évidente que la linguistique. Les études sur les attributs évaluatifs dans les sciences naturelles, sociales et humaines permettent à Stotesbury d'affirmer que l'abondance des attributs évaluatifs est plus importante dans les sciences sociales et humaines que dans les sciences naturelles. Mais dans notre étude, outre la biologie qui possède moins de collocations évaluatives par rapport à la taille du corpus, la médecine est souvent proche des sciences naturelles, la mécanique et l'électronique, comme toutes les branches de la physique, proches elles aussi des sciences naturelles, ont tendance à recourir à plus de collocations évaluatives dans les parties introductive et conclusive.

5. Les sciences humaines semblent plus homogènes que les autres familles de disciplines en termes d'utilisation des collocations évaluatives. Si l'on compare en détail les quatre disciplines de SHS selon la taxonomie de Stotesbury, la psychologie et les sciences de l'éducation (sciences sociales) contiennent moins d'adjectifs évaluatifs que la linguistique et le TAL (sciences humaines). Dans son étude, Stotesbury constatait pourtant que les sciences sociales mobilisaient davantage des lexiques évaluatifs que les sciences humaines (17 occurrences par 1000 mots dans les sciences sociales contre 15 occurrences par 1000 mots dans les sciences humaines).

La troisième étape vise à observer de plus près la répartition des collocations évaluatives entre les parties introductives et conclusives des disciplines.

que Stotesbury a étudié sont les sciences humaines qui regroupent la linguistique générale et appliquée, la littérature, l'anthropologie, la culture, l'histoire; les sciences sociales comprenant l'éducation, la psychologie, la sociologie, la géographie humaine, l'économie, les affaires, l'administration ; les sciences naturelles qui rassemblent l'écologie, la géologie, la botanique, la biologie animale, la limnologie, l'océanographie, la mécanique des sols, la sylviculture, la physique, la chimie, les mathématiques, les statistiques mathématiques.

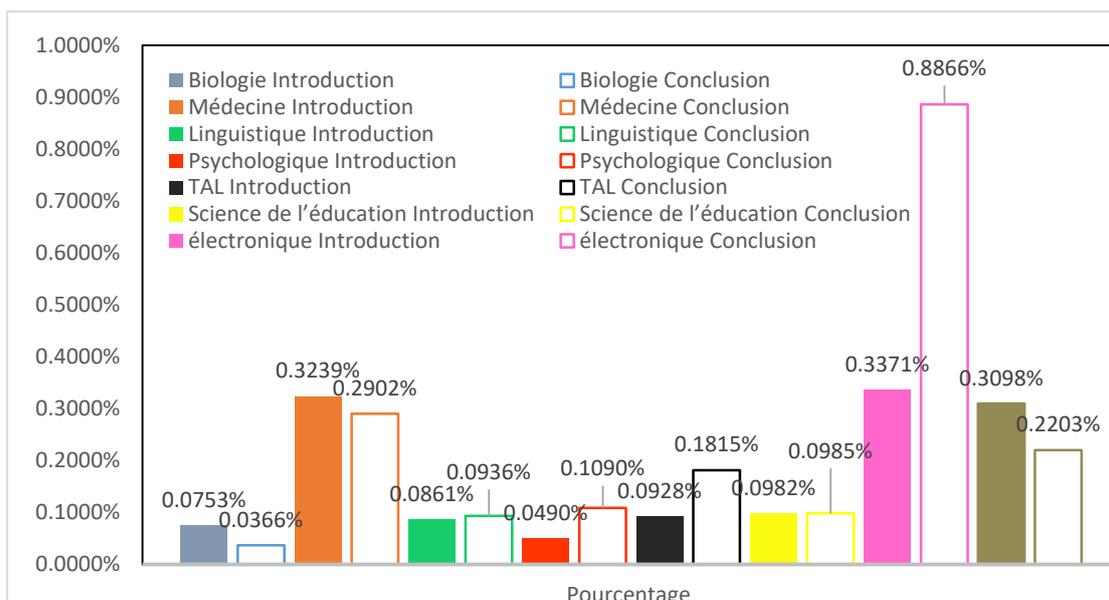


Figure 12 : Pourcentages des collocations évaluatives dans l'introduction et la conclusion des huit disciplines

Il est intéressant de voir qu'il y a seulement trois disciplines qui utilisent moins de collocations évaluatives dans la conclusion :

- la biologie : 0.0753 % pour l'introduction, > 0.0366 % pour la conclusion,
- la médecine : 0.3239 % pour l'introduction, > 0.2902 % pour la conclusion,
- la mécanique : 0.3266 % pour l'introduction, > 0.2203 % pour la conclusion.

Les cinq autres disciplines mobilisent plus de collocations évaluatives dans la conclusion que dans l'introduction :

- la linguistique : 0.0861 % pour l'introduction, < 0.0936 % pour la conclusion,
- la psychologie : 0.0490 % pour l'introduction, < 0.1090 % pour la conclusion,
- les sciences de l'éducation : 0.0982 % pour l'introduction, < 0.0985 %

pour la conclusion,

- le TAL : 0.0928% pour l'introduction, < 0.1815%,
- l'électronique : 0.3371 % pour l'introduction, < 0.8866 % pour la conclusion.

En résumé, nous constatons que cinq disciplines mobilisent davantage de collocations évaluatives dans la partie conclusive que dans la partie introductive. Néanmoins, l'analyse de Tutin et Cavalla (2008) faite sur les collocations évaluatives dans l'introduction et la conclusion de l'économie et de la linguistique aboutit à un constat inverse : l'introduction se sert plus de collocations évaluatives que la conclusion. Comme nous observons aussi trois disciplines qui possèdent plus de collocation évaluatives dans l'introduction que dans la conclusion, nous pourrions affirmer que la proportion des collocations évaluatives dans les deux parties textuelles est tout à fait variable selon les disciplines. Nous émettons aussi l'hypothèse que, au sein d'une discipline, la proportion des collocations évaluatives peut aussi varier suivant la taille et le genre des écrits scientifiques du corpus. C'est le cas de la linguistique : Tutin et Cavalla ont observé la surreprésentation des collocations évaluatives dans l'introduction alors que nous avons obtenu un résultat tout à fait inverse. Deux différences entre les deux études ont été remarquées : la taille du corpus, les genres des écrits scientifiques. Pour la taille, nous rappelons ici le critère de la constitution de corpus discuté dans la partie théorique : pour un corpus de spécialité, il est préférable d'avoir une taille d'au moins 500 000 mots provenant de plusieurs textes, la représentativité va être plus sûre suite à la croissance de la taille de corpus (Williams, 1999). Notre corpus de linguistique comporte 1 633 579 mots au total (67 textes) dont 31353 pour l'introduction et la conclusion, le corpus de Tutin et Cavalla regroupe 50 textes avec 650 693 mots en total. Nous jugeons donc que la taille des introductions et conclusions dans le corpus de Tutin et Cavalla seraient probablement plus petite par rapport à la nôtre. Entre les deux corpus, nous observons une augmentation de la proportion des collocations évaluatives dans la conclusion : elle s'élève suite à l'agrandissement de la taille de corpus et enfin dépasse celle que nous trouvons dans l'introduction dans notre corpus. Quant au genre des écrits, notre corpus de linguistique

regroupe des articles (8), des communications écrites (47), des thèses de doctorat (8) et des mémoires d’Habilitation à Diriger des Recherches (4), tandis que le corpus de Tutin et Cavalla comporte des articles, des thèses, des rapports ainsi que des manuels. La différence de genre de textes pourrait aussi avoir un impact sur les résultats.

En parallèle, quatre disciplines, à savoir la biologie, la psychologie, le TAL et l’électronique manifestent une grande différence sur le nombre d’occurrences des collocations évaluatives dans chaque partie, leur rapport allant du simple au double.

Contrairement à l’hétérogénéité en matière de distribution des collocations dans les deux parties textuelles dans les sciences expérimentales et appliquées (la fréquence des collocations évaluatives n’a pas la même tendance au sein de la même famille de disciplines), les sciences humaines sont sujettes à se servir davantage de collocations évaluatives dans la conclusion que dans l’introduction.

6.1.2. Analyse des collocations axiologiques et non axiologiques dans les huit disciplines

Après le constat général sur la globalité des disciplines, nous prétendons enquêter à l’intérieur de chaque discipline. Ces analyses se font sur la base des tableaux de classement des collocations dans l’annexe 2. Nous proposons de faire afficher la distinction quantitative des collocations évaluatives axiologiques et non-axiologiques, ainsi que les sous-catégories au sein des collocations non axiologiques pour chaque discipline. Nous mettrons en commun à la fin des résultats et des analyses obtenues pour chaque discipline afin de dégager des régularités et des singularités au niveau quantitatif de l’emploi des collocations évaluatives dans les huit disciplines.

① La biologie :

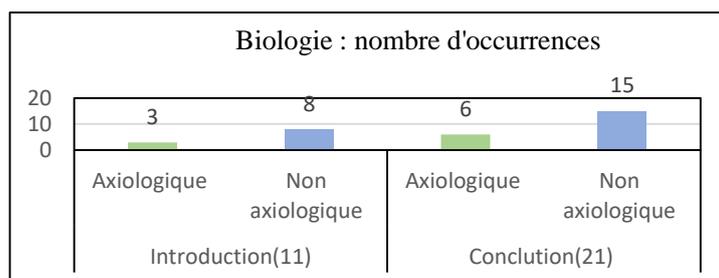


Figure 13 : Distinction des collocations axiologiques et non axiologiques dans la biologie

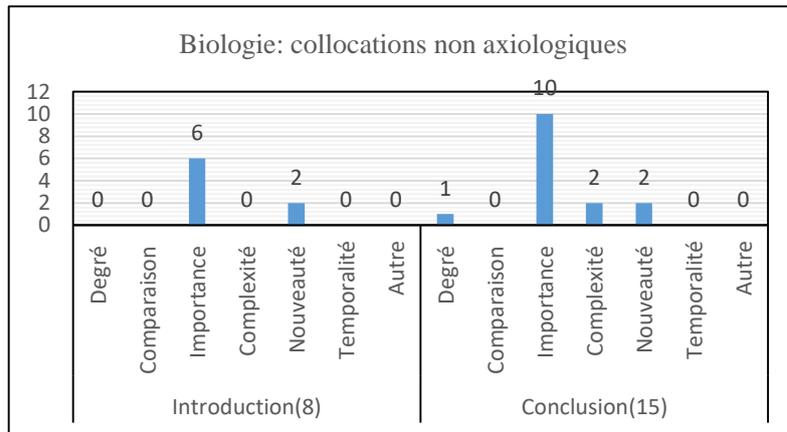


Figure 14 : Sous-catégories des collocations non axiologiques dans la biologie

La figure 13 montre que les collocations axiologiques sont moins fréquentes que les collocations non axiologiques dans l'introduction et la conclusion des textes en biologie. Cependant, la conclusion a tendance à mobiliser plus de collocations axiologiques que l'introduction ($3/8 = 0.375$ dans l'introduction, $< 6/9 = 0.4$ dans la conclusion).

Parmi les collocations non axiologiques, les sous-catégories Importance et Nouveauté sont toutes représentées alors que Degré et Complexité sont seulement présentées dans la conclusion. On ne recense aucune collocation impliquant les dimensions Comparaison, Temporalité et Autre. De plus, la sous-catégorie Importance représente plus de la moitié des autres sous-catégories de collocations (employée dans 75% des cas pour l'introduction et dans 66.7% pour la conclusion).

② La médecine :

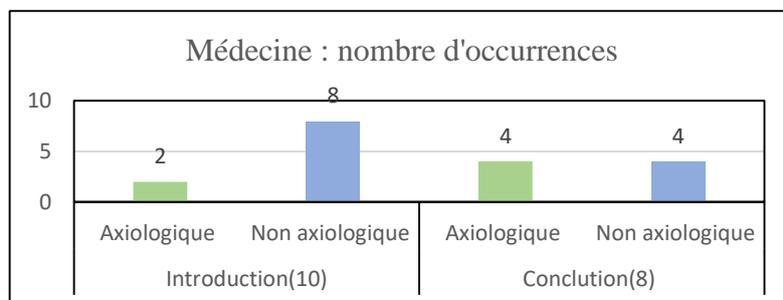


Figure 15 : Distinction des collocations axiologiques et non axiologiques dans la médecine

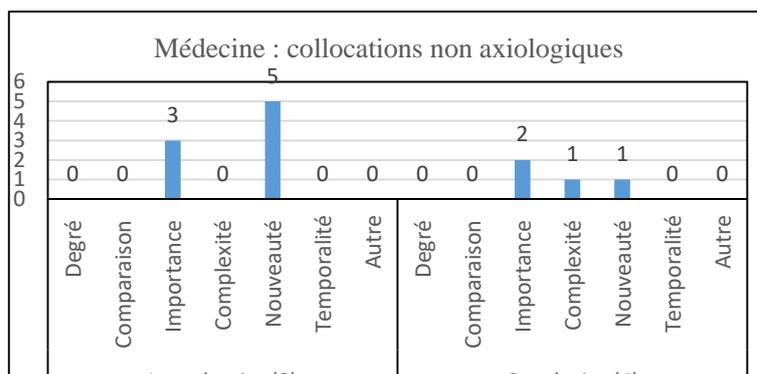


Figure 16 : Sous-catégories des collocations non axiologiques dans la médecine

Dans l'introduction des textes en médecine, les collocations axiologiques ne sont pas nombreuses et représentent 20% des cas d'occurrences de $N_{\text{scientifique}} + \text{Adj}_{\text{évaluatif}}$ générales; cette statistique atteint 50% dans les conclusions. Quant au classement des collocations non axiologiques, c'est encore Importance et Nouveauté qui dominent dans les deux parties textuelles, comme c'est le cas en biologie. Les sous-catégories Degré et Complexité demeurent moins représentées. Nous témoignons encore une fois l'absence des collocations s'inscrivant dans les dimensions Temporalité et Autre.

③ La linguistique :

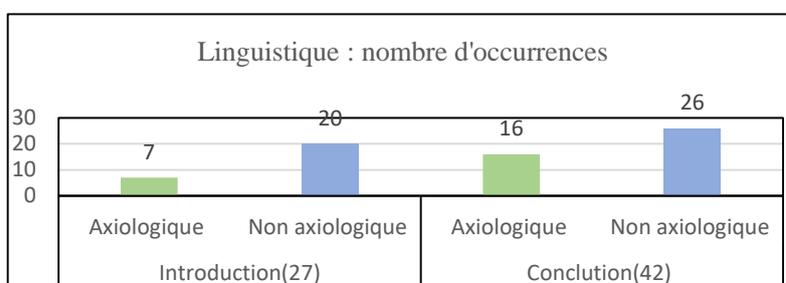


Figure 17 : Distinction des collocations axiologiques et non axiologiques dans la linguistique

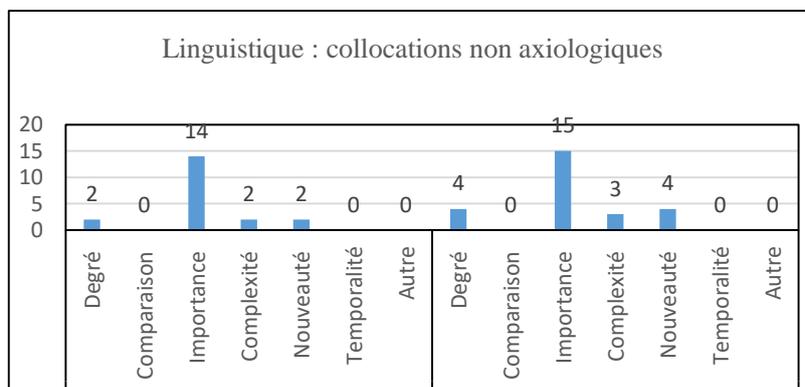


Figure 18 : Sous-catégories des collocations non axiologiques dans la linguistique

Tout comme la biologie et la médecine, la linguistique a tendance à employer plus fréquemment les collocations non axiologiques que les collocations axiologiques dans les deux parties textuelles. Ces dernières sont plus fréquentes dans la conclusion que dans l'introduction. À la différence de la médecine et la biologie, où les sous-catégories des collocations non axiologiques ne sont pas assez comparables, la linguistique manifeste une très forte cohérence. En effet, les collocations non axiologiques dans l'introduction et la conclusion se répartissent dans quatre classes : le degré, l'importance, la complexité et la nouveauté. Parmi elles, l'importance prend un avantage absolu. En ce qui concerne les trois autres sous-catégories, nous n'avons rien recueilli.

④ La psychologie :

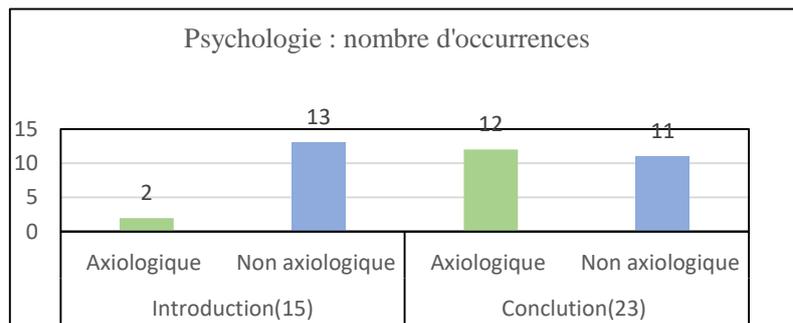


Figure 19 : Distinction des collocations axiologiques et non axiologiques dans la psychologie

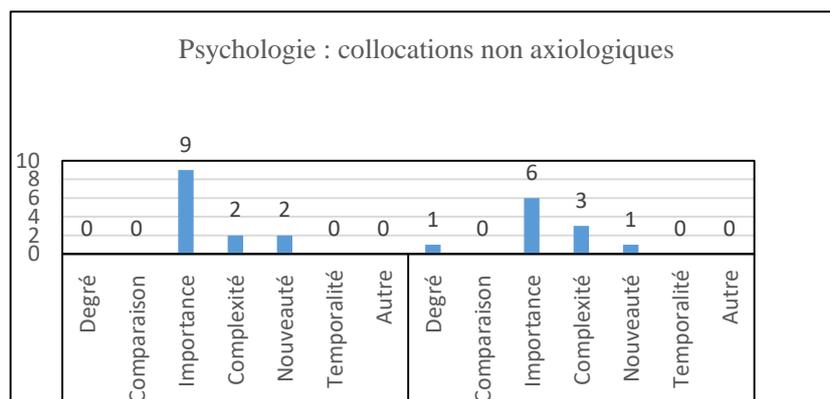


Figure 20 : Sous-catégories des collocations non axiologiques dans la Psychologie

D'abord, les collocations axiologiques restent moins présentes dans l'introduction que les collocations non axiologiques comme dans les disciplines analysées précédemment.

Mais dans la conclusion, le nombre d'occurrence des collocations axiologiques dépasse pour la première fois légèrement celui des collocations non axiologiques (12 > 11). La figure 20 illustre pourtant que la sous-catégorie Importance se trouve toujours en tête, les autres catégories comme Degré, Complexité, Nouveauté étant présentes mais dans une proportion moindre par rapport à la dimension Importance. Les dimensions Comparaison, Temporalité et Autre sont toujours absentes.

⑤ Les sciences de l'éducation :

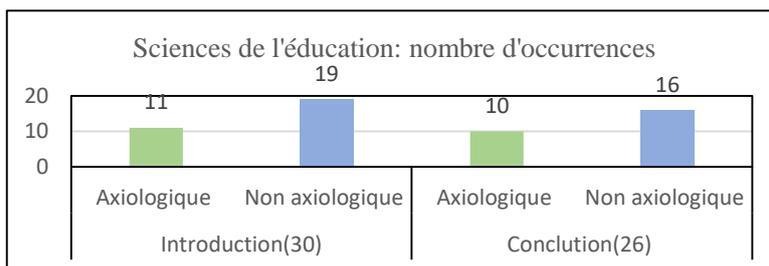


Figure 21 : Distinction des collocations axiologiques et non axiologiques dans les sciences de l'éducation

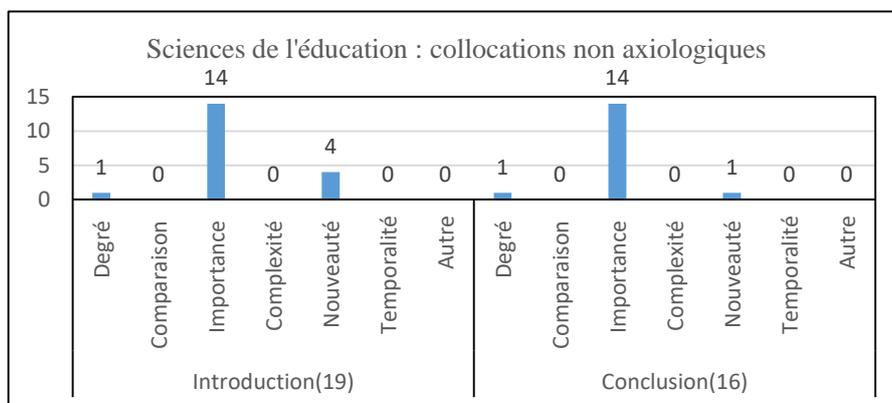


Figure 22 : Sous-catégories des collocations non axiologiques dans les sciences de l'éducation

Les collocations axiologiques occupent toujours une place inférieure par rapport aux collocations non axiologiques, tant dans l'introduction que dans la conclusion, où leur proportion augmente légèrement ($11/30 = 0.367 < 10/26=0.385$). En ce qui concerne les collocations non axiologiques, elles se répartissent en trois sous-catégories : Degré, Importance, Nouveauté, dont Importance s'attribue la plupart des collocations non axiologiques. Les sous-catégories comme Comparaison, Complexité, Temporalité, Autre

n'apparaissent pas.

⑥ Le TAL :

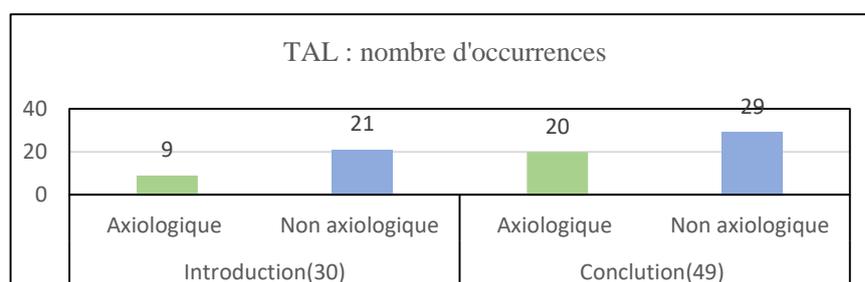


Figure 23 : Distinction des collocations axiologiques et non axiologiques dans le TAL

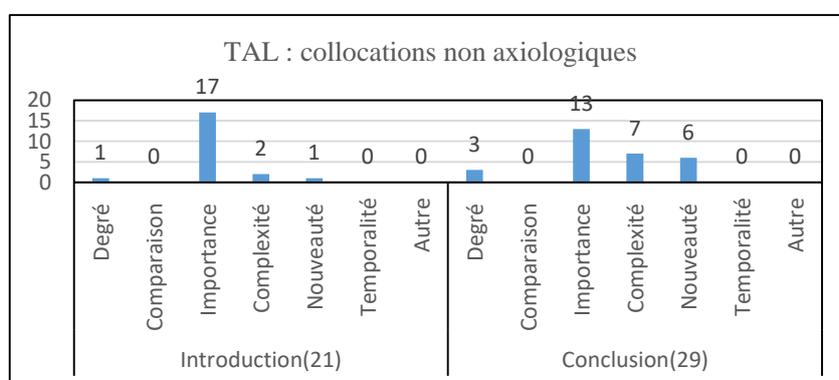


Figure 24 : Sous-catégories des collocations non axiologiques dans le TAL

L'emploi des collocations axiologiques et leur répartition entre les deux parties textuelles sont tout à fait identiques à ce qui est observé dans les sciences de l'éducation : les collocations axiologiques sont moins importantes que les collocations non axiologiques, l'introduction mobilise moins de collocations axiologiques que la conclusion.

Les collocations non axiologiques dans le TAL se répartissent de la même manière que celles dans la linguistique : quatre sous-catégories Degré, Importance, Complexité et Nouveauté se partagent toutes les collocations non axiologiques, et ne laissent aucune collocation non axiologique aux sous-catégories restantes. Les collocations portant le sens d'importance sont toujours les collocations les plus nombreuses dans les deux parties textuelles.

⑦ La mécanique :

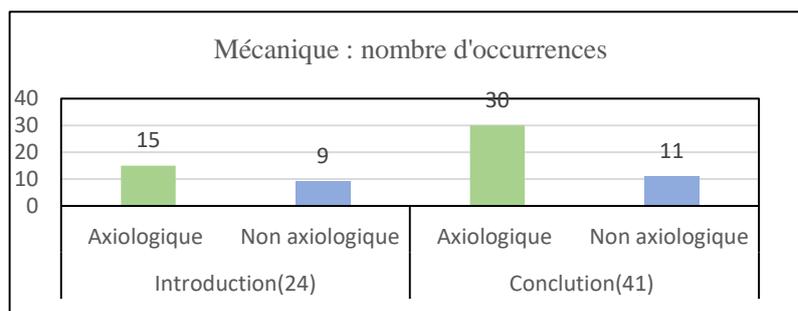


Figure 25 : Distinction des collocations axiologiques et non axiologiques dans la mécanique

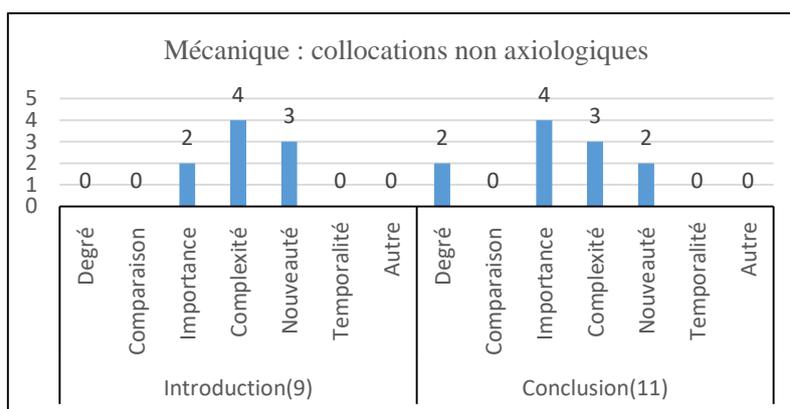


Figure 26 : Sous-catégories des collocations non axiologiques dans la mécanique

Pour la première fois, les collocations axiologiques sont simultanément plus nombreuses que les collocations non axiologiques dans les deux parties textuelles. Mais leurs proportions de fréquence suivent le même développement que dans les disciplines précédentes : de 62.5% dans l'introduction à 73.2% dans la conclusion, ce qui explique une augmentation des collocations évaluatives dans la conclusion par rapport à l'introduction.

Les collocations non axiologiques manifestent la même propriété que d'autres disciplines : quatre sous-catégories (Degré, Importance, Complexité, Nouveauté) prédominent avec des préférences différentes et les autres sont absentes. Dans l'introduction, ce sont les collocations de la sous-catégorie Complexité qui deviennent les plus nombreuses. Les collocations signifiant l'importance restent les plus fréquentes dans la conclusion.

⑧ L'électronique :

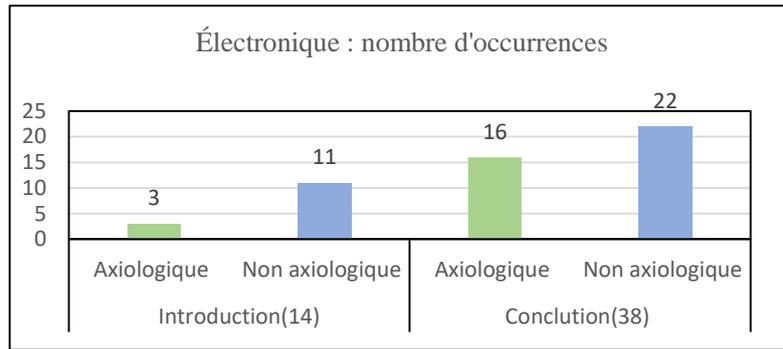


Figure 27 : Distinction des collocations axiologiques et non axiologiques dans l'électronique

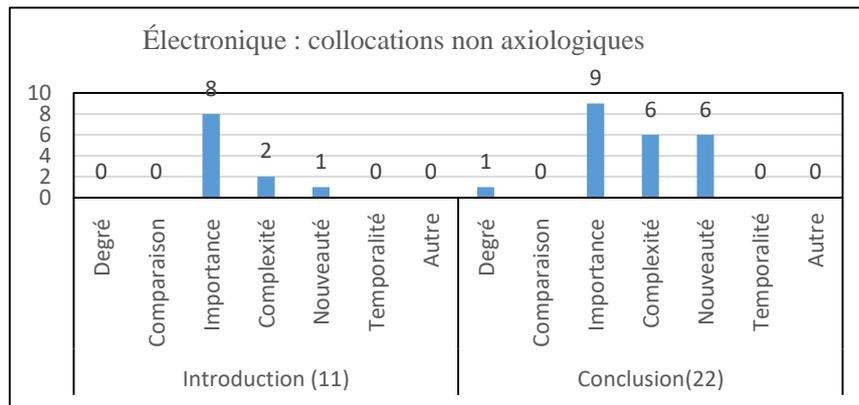


Figure 28 : Sous-catégories des collocations non axiologiques dans l'électronique

Les collocations non axiologiques sont plus nombreuses que les collocations axiologiques. Dans les conclusions, les auteurs préfèrent utiliser plus de collocations axiologiques que dans leurs introductions ($16/38 = 42\% > 3/14 = 21.4\%$). Il est intéressant de constater que les collocations non axiologiques se divisent toujours principalement en quatre sous-catégories : Degré, Importance, Complexité, Nouveauté, dont l'importance reste la plus significative.

6.1.3. Mise en commun des huit disciplines

Les analyses réalisées par disciplines nous permettent de dégager en premier lieu les régularités sur l'emploi des collocations évaluatives :

1. Aucune collocation modale n'a été repérée. Nous suggérons donc deux explications possibles :

a) L'insuffisance de la taille des corpus

b) Les adjectifs modaux qualifient souvent l'énoncé, il est rare qu'ils soient associés aux noms.

2. A l'exception des textes de mécanique, où les collocations axiologiques semblent l'emporter sur les collocations non axiologiques dans les deux parties textuelles, et des textes de psychologie, où le nombre des collocations axiologiques est légèrement supérieur à celui des collocations non axiologiques dans la conclusion, les autres disciplines mobilisent plus de collocations non axiologiques que les collocations axiologiques dans les deux parties textuelles.

Pour la surreprésentation des collocations axiologiques dans la conclusion des textes psychologiques, nous pensons qu'il s'agit plutôt de la question de l'idiolecte, car parmi les 11 textes contenant des collocations évaluatives dans la conclusion, deux de ces textes ont mobilisé 11 des 12 collocations axiologiques recensées.

Quant à la mécanique, la supériorité des collocations axiologiques dans les deux parties textuelles pourrait s'expliquer par l'insuffisance du nombre de textes et aussi par l'idiolecte. La mécanique n'a que trois textes, 15 sur 15 collocations axiologiques dans l'introduction et 26 sur 30 collocations axiologiques dans la conclusion sont dues à un même auteur (nous recensons 13 occurrences de la collocation *structure cohérente*, 2 occurrences de *phénomène cohérent* dans l'introduction, 21 occurrences de *structure cohérente*, 2 occurrences de *résultat cohérent*, 2 *phénomènes cohérents*, 1 *caractéristiques cohérentes* dans la conclusion).

En conséquence, nous pouvons dire que les collocations non axiologiques sont généralement plus représentées que les collocations axiologiques dans l'introduction et la conclusion. La figure 29 illustre la situation de l'emploi des collocations axiologiques dans les deux parties textuelles respectivement des huit disciplines :

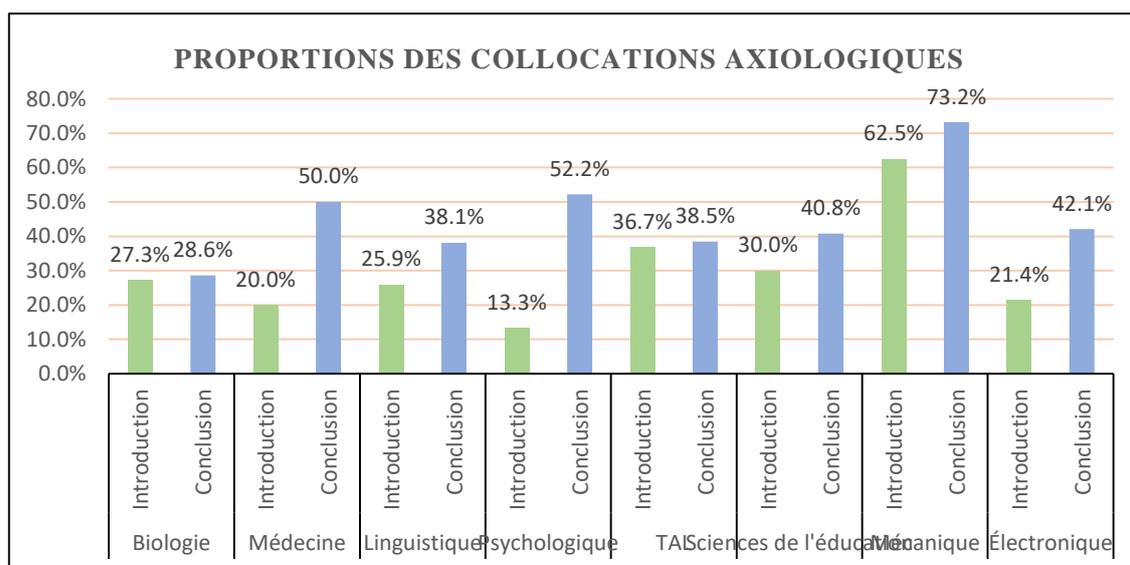


Figure 29 : Proportion des collocations axiologiques dans les deux parties textuelles des huit disciplines

Nos résultats font écho au constat du travail de Tutin (2010) sur les adjectifs évaluatifs dans l'économie et la linguistique, qui explique que les adjectifs axiologiques sont relativement moins nombreux par rapport aux adjectifs non axiologiques dans l'ensemble des deux disciplines.

3. La figure 29 montre que la tendance à l'emploi des collocations évaluatives entre l'introduction et la conclusion dans toutes les disciplines est identique : la conclusion est susceptible de mobiliser davantage de collocations axiologiques que l'introduction. C'est probablement lié à la stratégie rhétorique requise par la conclusion : c'est souvent dans la conclusion que l'auteur ressent la nécessité d'évaluer la qualité de ses travaux et d'émettre des jugements sur les résultats obtenus afin de promouvoir son travail.

4. Les collocations non axiologiques se répartissent principalement en quatre sous-catégories : Degré, Importance, Complexité, Nouveauté. Les dimensions Importance et Nouveauté sont présentes dans les deux parties textuelles des huit disciplines. Les dimensions Degré et Complexité sont moins fréquentes, leur occurrence n'est pas régulière. L'abondance des dimensions Importance et Nouveauté peut également s'expliquer par la stratégie rhétorique que l'auteur a besoin d'adopter dans l'introduction et la conclusion. L'auteur doit tenter de souligner dans les deux parties l'intérêt et l'originalité

de ses études. Pour ce faire, il aurait besoin de recourir aux collocations mettant en valeur l'importance et la nouveauté. La situation est présentée par la figure 30 :

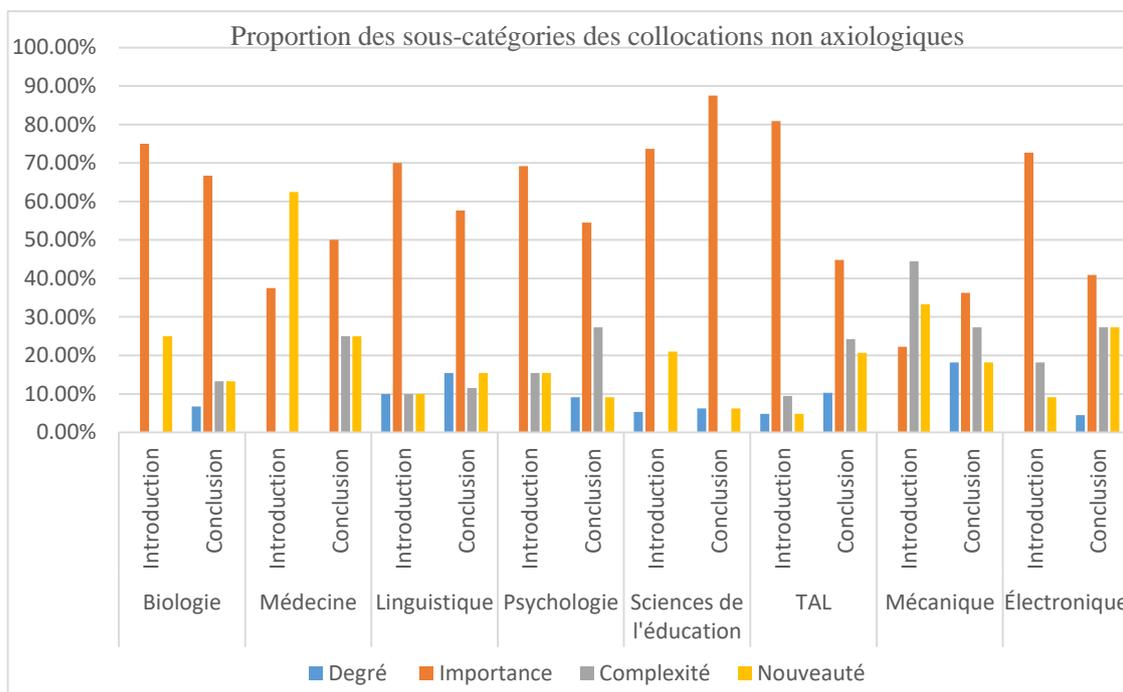


Figure 30 : Proportion des sous catégories des collocations non axiologiques dans l'introduction et la conclusion des huit disciplines

Encore une fois, notre résultat diffère du travail de Tutin et Cavalla (2008). Dans ce dernier, les sous-catégories Importance, Nouveauté et Temporalité contiennent la plupart des collocations dans l'économie et les sous-catégories Nouveauté, Temporalité et Comparaison contiennent la plupart des collocations dans la linguistique. Nous admettons que cette différence est probablement due à la différence de corpus. (Corpus KIAP vs Corpus Scientext).

Mais l'absence de collocation relative à la dimension *Temporalité* dans notre corpus nous paraît étrange. De même pour les dimensions *Comparaison*, et *Autre*.

6.1.4. Analyse de la distribution des noms scientifiques dans les collocations évaluatives

Nos collocations évaluatives comportent deux constituants, la base est le nom scientifique et le collocatif est l'adjectif évaluatif. Dans la partie précédente, nous avons

recensé quatre dimensions pour les adjectifs évaluatifs : Degré, Importance, Complexité et Nouveauté.

Arrêtons-nous sur la distribution des noms scientifiques dans les huit disciplines. Il faut rappeler que dans la partie méthodologique, nous avons retenu six classes : les noms de processus de l'activité scientifique, les noms d'objets construits par l'activité scientifique, les noms d'observables de l'activité scientifique, les noms de support de la rédaction scientifique, les noms de caractérisation et les noms de relation logique. Après le recensement, nous constatons que les noms de support de la rédaction scientifique et ceux de la relation logique sont introuvables dans nos listes de collocations évaluatives. Les noms scientifiques recensés se partagent en quatre classes : Processus, Objet construit, Observable et Caractérisation. Au sein des quatre classes, nous remarquons une inégalité de distribution. Voici la figure suivante qui illustre notre constat :

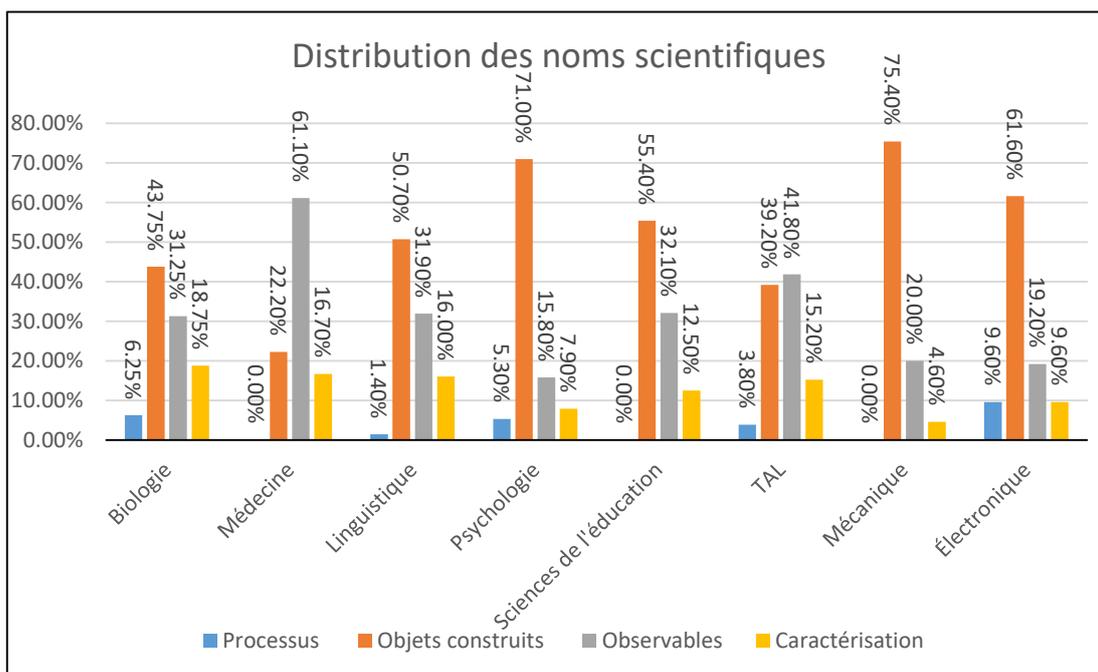


Figure 31 : Distribution des noms scientifiques dans les huit disciplines

Si nous observons plus attentivement la figure 31, certains phénomènes intéressants pourraient être dégagés :

1. La figure met en lumière la prédominance de deux classes dans toutes les

disciplines : les noms d'objet construit par l'activité scientifique et ceux d'observable de l'activité scientifique. Les classes Processus et Caractérisation sont moins représentées. Parmi huit disciplines, six ont plus de combinaisons entre les adjectifs évaluatifs et les noms d'objet construit (la biologie, la linguistique, la psychologie, les sciences de l'éducation, la mécanique et l'électronique). La médecine et le TAL disposent de plus de combinaisons liées aux noms d'observables de l'activité scientifique. Cela est peut-être dû à un besoin accru d'analyser des données, des facteurs, et aussi les résultats dans les deux disciplines.

2. La classe Caractérisation est plus représentée que la classe Processus, nous trouvons que la classe Processus n'est pas toujours présentée, tandis que Caractérisation est commune à toutes les disciplines.

Bilan

Les analyses statistiques dans cette partie nous ont permis de comprendre d'une façon globale les constantes et les singularités des huit disciplines en matière d'utilisation des collocations évaluatives dans l'introduction et la conclusion. Nous proposons de résumer ici les constats principaux :

1) La médecine, la mécanique et l'électronique mobilisent plus de collocations évaluatives dans l'introduction et la conclusion.

2) Cinq parmi huit disciplines mobilisent plus de collocations évaluatives dans la conclusion que dans l'introduction.

3) Les huit disciplines ont tendance à employer plus de collocations non axiologiques que de collocations axiologiques dans les deux parties textuelles, sauf en cas d'influence de l'idiolecte. Les collocations modales sont introuvables dans les huit disciplines.

4) La conclusion semble plus sujette à utiliser plus de collocations axiologiques que l'introduction.

5) Pour les collocations non axiologiques, les dimensions Importance et

Nouveauté sont omniprésentes dans toutes les parties textuelles en transcendant les huit disciplines, alors que les dimensions Degré et Complexité s'y manifestent mais dans une proportion plus faible. Les dimensions Comparaison, Temporalité, et Autre sont tout à fait invisibles pour les huit disciplines.

6) En ce qui concerne les noms scientifiques dans les collocations évaluatives, les classes présentées sont : Processus, Objets construits, Observables et Caractéristique, dont Objets construits et Observables ont un statut prédominant par rapport aux deux autres classes.

6.2. Analyses qualitatives

Les analyses qualitatives se focalisent sur la sémantique des adjectifs évaluatifs et leur combinaison avec des noms scientifiques, ainsi que sur les caractéristiques en matière de fonctions rhétoriques traduites par les sémantiques des collocations évaluatives au sein de chaque discipline. Nos objectifs sont de faire ressortir les caractéristique sémantiques des collocations évaluatives et leurs fonctions rhétoriques dans un premier temps, de découvrir les patrons syntaxiques des collocations évaluatives et leur préférence dans les huit disciplines dans un deuxième temps, et d'interroger la présence de la transdisciplinarité des collocations évaluatives dans le dernier temps. Tous les travaux vont contribuer enfin à l'élaboration des propositions didactiques des collocations évaluatives dans notre dernier chapitre.

6.2.1. Analyses sémantiques des collocations évaluatives

Le sens global des collocations en question sont tout à fait compositionnel : le nom scientifique comme la base garde le sens habituel lorsque l'adjectif évaluatif comme le collocatif prend le sens habituel avec la prise en compte le contexte pour donner à la base une dimension évaluative. En conséquence, la sémantique évaluative d'une collocation évaluative est forcément traduite par le collocatif - l'adjectif évaluatif. Dans ce cas-là, il convient d'opérer des observations au sein des dimensions que nous avons discutées dans les analyses quantitatives.

Nous suivons toujours la même logique : les collocations axiologiques seront

analysées dans la première étape, ensuite des sous-catégories au sein des collocations non axiologiques ainsi que leur préférence en matière de combinaison avec des noms scientifiques et une mise en commun à la fin.

6.2.1.1. Analyses sémantiques sur des collocations axiologiques

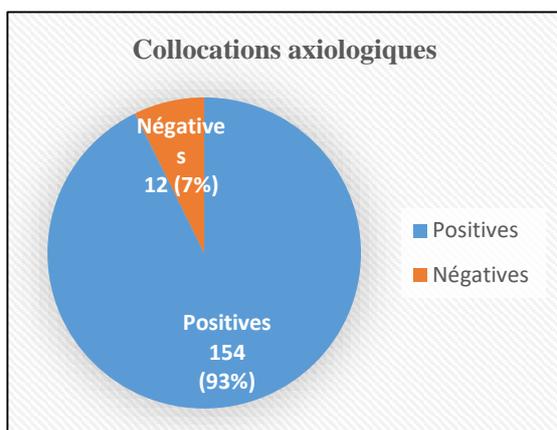


Figure 32 : Polarités des collocations axiologiques

Même si les collocations axiologiques ne sont pas autant représentées que les collocations non axiologiques, elles occupent quand même une place indispensable dans les textes scientifiques. Les adjectifs axiologiques se divisent en fonction de leur

polarité en deux classes : positive et négative.

Les collocations positives représentent la majorité des collocations axiologiques ; ce qui s'explique par le fait que l'auteur de l'écrit scientifique préfère valoriser son travail au lieu de le déprécier.

Parmi les collocations axiologiques négatives, les adjectifs mobilisés sont : *insuffisant* (2 occurrences), *épineux* (1 occurrence), *mauvais* (3 occurrences), *erroné* (6 occurrences). Ils préfèrent s'associer avec les noms d'objets construits (*modèle*, *conception*), d'observables (*problème*, *donnée*), de caractérisation (*qualité*).

Pour des collocations évaluatives positives, les adjectifs les plus fréquents (nombre d'occurrences ≥ 5) sont : *cohérent*, *bon/meilleur*, *pertinent*, *intéressant*, *satisfaisant*, *encourageant*. Ils se combinent avec les noms d'objets construits, d'observations, de caractérisation et de processus.

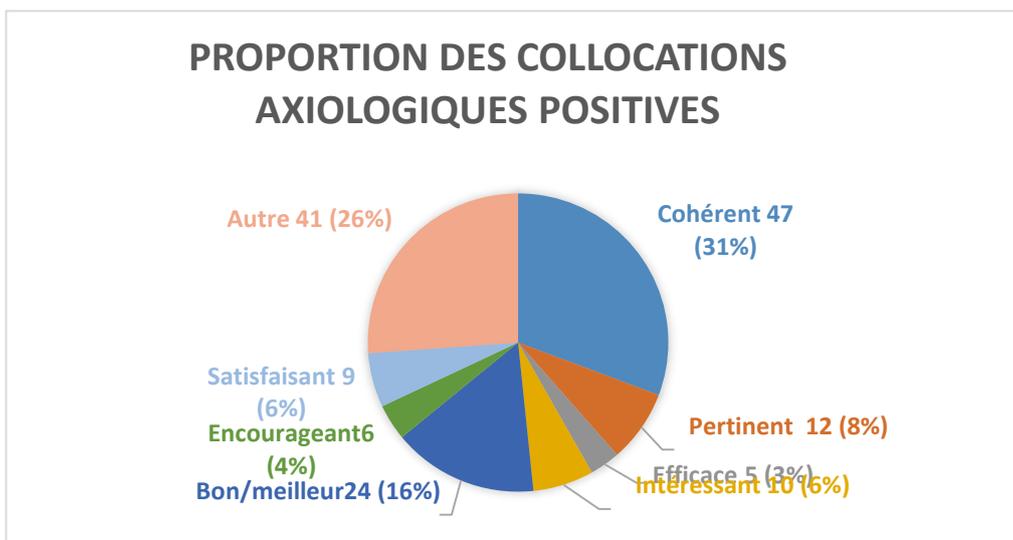


Figure 33 : distribution des adjectifs axiologiques

Cohérent : il décrit un objet « qui se compose de parties liées et harmonisées entre elles » (Dictionnaire le Robert illustré, 2017 : 402 - Article *cohérent*). Cet adjectif est prédominant dans nos corpus du fait qu'un seul auteur d'une thèse en mécanique a utilisé 33 fois la même collocation - *structure cohérente*. En dehors de ces 33 occurrences, il se combine souvent avec les noms d'objets construits comme *modèle*, *idée*, *description*, et les noms d'observables *résultat*, *phénomène* qui peuvent manifester une cohérence soit en structure, soit en logique.

- (1) Ce **modèle (Objet construit)** est **cohérent** avec le mode de fixation des métallos-régulateurs en général : l'espacement entre les deux hélices de reconnaissance doit être compatible avec l'espacement imposé par la séquence de reconnaissance. (Corpus Scientext, Biologie)
- (2) Cette seconde **idée (Objet construit)** est **cohérente** avec certaines approches pédagogiques actuelles, consistant à éviter autant que faire se peut les feed-back négatifs, les indications d'échecs, et les conflits, c'est-à-dire tout ce qui signifie à l'élève ou l'étudiant qu'il peut se tromper, ou s'être trompé. (Corpus Scientext, Psychologie)

Cet adjectif se combine différemment avec les noms scientifiques selon les disciplines. Pour les disciplines des sciences appliquées et expérimentales, les noms scientifiques en cooccurrence avec *cohérent* sont plutôt des noms appartenant aux « appareils théoriques » comme *modèle* (l'exemple (1)), *structure* ou des noms d'observables comme *phénomène*, *résultat*, tandis que les sciences humaines ont tendance à combiner l'adjectif cohérent davantage avec les noms relatifs à la conceptualisation scientifiques

comme *idée* (l'exemple (2)), *conception*, *définition*.

Bon/meilleur : ces adjectifs émettent un jugement très positif envers l'entité. Ils se combinent avec les noms d'objets construits comme *traitement*, *méthode*, *modèle*, etc, surtout les Observables comme *résultat* (12 occurrences parmi 24), *exemple*, et le nom de caractérisation comme *qualité* (5 parmi 24). La combinaison entre ces deux adjectifs avec les noms susmentionnés peut donner une impression favorable au lecteur sur les études présentées dans l'écrit scientifique.

(3) Si comme je le crois le dialogue entre apprenants est une mine d'or pour les recherches en sciences cognitives 64, une logique cognitive et langagière est le **meilleur outil (Objet construit)** pour analyser les processus ayant aboutis à la géologie actuelle, et les technologies de la communication et de la connaissance constituent les **meilleurs outils** pour extraire et travailler la matière précieuse. (Corpus Scientext, psychologique).

(4) Ainsi , la version COb a apporté de **meilleurs** résultats (Observable), tant au plan du rappel que de la précision . (Corpus Scientext, TAL)

(5) Les valeurs limites des trois paramètres mis en jeu (SNR & gt ; 6 dB , SIR & gt ; - 25 dB et IIP3 & gt ; - 3 dBm) ont été déterminées afin d' obtenir une **bonne qualité (Caractérisation)** de transmission qui a été définie par une valeur de BER inférieure à 10⁻⁵. (Corpus Scientext, Électronique).

Les noms scientifiques se trouvant à côté de l'adjectif *bon/meilleur* ne semblent pas très différents d'une discipline à l'autre. Le choix des noms à évaluer ici est plus restreint par l'adjectif que par les disciplines.

Pertinent : Qui « convient exactement à l'objet dont il s'agit, qui dénote du bon sens » ou « doué d'une fonction dans un système, un ensemble » (Dictionnaire le Robert illustré, 2017 : 1470 - Article *pertinent*). Il peut se combiner avec les classes Objets construits, Observables, Caractérisation, sans préférence pour une classe particulière des noms.

(6) Nous dégagerons aussi de ces contextes les marques formelles qui constituent des **éléments (Observable) pertinents** pour un repérage automatique, second objectif que nous poursuivons. (Corpus Scientext, Linguistique)

Il est très intéressant pour nous de remarquer que les occurrences de l'adjectif *pertinent* se situent presque exclusivement dans les sciences humaines (11/12), aucune

occurrence n'apparaissant dans la biologie, la médecine, et l'électronique, et une seule dans la mécanique (*solution pertinente*).

Intéressant : « Qui retient l'attention, captive l'esprit » (Dictionnaire le Robert illustré, 2017 : 1010 - Article *intéressant*), se combine avec Objets construits, Observables, Caractérisation, et a un faible pour le nom d'observable *résultat*.

(7) Il serait nécessaire de confirmer les **résultats (Observable) intéressants** que nous avons obtenus par la lecture spectrophotométrique ainsi que d'évaluer l'utilisation d'indicateurs colorés. (Corpus Scientext, Biologie)

Satisfaisant : « Qui satisfait, est conforme à ce qu'on peut attendre » (Dictionnaire le Robert illustré, 2017 : 1753 - Article *satisfaisant*). Il s'associe avec les noms d'objets construits et d'observables, en particulier avec *traitement, résultat*.

(8) Le passage en langue cible à partir d'une représentation du sens apporte des **résultats (Observable) plus satisfaisants** qu'une traduction de type syntactico-lexicale, car elle considère la compréhension et la représentation du sens comme un préalable à une traduction de qualité. (Corpus Scientext, Linguistique)

Encourageant : « Qui encourage, est propre à encourager » (Dictionnaire le Robert illustré, 2017 : 649 - Article *encourageant*). Dans notre corpus, il se combine seulement avec le nom d'observables *résultat*.

(9) Les **résultats(Observable)** sur les voiles souples sont très **encourageants**. (Corpus Scientext, Mécanique)

Efficace : c'est un adjectif polysémique, il peut signifier soit quelque chose « qui produit l'effet qu'on en attend », soit une personne « dont la volonté, l'activité produisent leur effet, opposé à inefficace » (Dictionnaire le Robert illustré, 2017 : 625 - Article *efficace*). Il s'associe avec le nom de processus et d'objets construits, surtout avec *traitement*.

(10) Pour rendre la **communication efficace**, le plus souvent, ce sont des séquences sonores différentes qui nomment des objets différents. C2 pénalise les vocabulaires avec des mots identiques. (Corpus Scientext, Linguistique)

(11) La production ne paraît toutefois pas nécessaire à la réalisation de **traitements efficaces** si ceux -ci sont opérationnalisés par d'autres tâches, comme en témoignent les résultats de l'expérience 8 Cela confirme l'hypothèse d'un déficit de production du traitement sémantique

lié au vieillissement, auquel nous avons conclu à la fin du chapitre précédent. (Corpus Scientext, Psychologie)

Pour les autres adjectifs, nous ne pouvons pas les traiter un par un faute de temps. Pourtant, nous aimerions attirer l'attention sur certains adjectifs polysémiques qui peuvent entrer dans la dimension axiologique ou non axiologique simultanément, et dont la distinction du sens polysémique s'effectuera à l'aide d'un contexte. Prenons ici *véritable* comme exemple :

Véritable : il peut signifier : a) conforme à l'apparence, qui n'est pas imité, qui est réel, ce sens n'est pas axiologique; b) conforme à l'idée qu'on s'en fait, qui mérite son nom, ce sens renvoie au jugement des propriétés de l'entité, il est donc axiologique.

(12) Dans le contexte scientifique de l'UMR GRIC, et également dans la communauté de recherche internationale travaillant sur les interactions inter-humaines, j'ai été frappé par les difficultés d'ordre épistémologique qui freinent l'émergence d'une **véritable théorie (Objet construit)** cognitive de l'interaction communicative (j'en reviens dans le dernier chapitre de cette synthèse). (Corpus Scientext, Psychologie)

(13) Le fait que le **véritable critère (Objet construit)** qui sous-tend l'assignation des articles (manifestant le caractère typiquement nominal) soit en réalité la saillance ou la PROÉMINENCE dans le discours est fortement suggéré par des faits concernant la relation entre la morphologie nominale et le discours dans [diverses] langues. (Corpus Scientext, TAL)

(14) L'une des difficultés les plus simples à résoudre est sans doute la prise de décisions menant à un échec futur. Il suffit en effet de remettre en cause le choix opéré. Cette solution est celle adoptée ici. Le **véritable problème (Observable)** survient lorsque toutes les solutions mènent à l'échec. (Corpus Scientext, TAL)

(15) Nous pensons en particulier à un **véritable outil (Objet construit)** d'aide à la rédaction en Langue Contrôlée, dont les caractéristiques exactes restent à spécifier. (Corpus Scientext, TAL)

Dans les exemples (12) et (13), l'adjectif *véritable* est axiologique, il se combine avec les noms *théorie* et *critère* pour faire allusion à la légitimité de cette théorie ou de ce critère qu'on utilise. En revanche, *véritable problème* et *véritable outil* dans les exemples (14) et (15) ne signifient pas un problème ou un outil légitime ou pertinent, ils veulent plutôt mettre l'accent sur le vrai problème qui se cache et sur un vrai outil qui peut s'employer dans la réalité. L'adjectif *véritable* ici n'émet pas de jugement sur la valeur du

problème et de l'outil. En conséquence, il est bien non axiologique.

À part l'adjectif *véritable*, nous avons d'autres adjectifs qui sont pareils à *véritable*. À titre d'exemple : *original, épineux, faible*.

6.2.1.2. Analyses sémantiques sur des collocations non axiologiques

Les analyses quantitatives sur les collocations non axiologiques aboutissent à mettre en évidence les quatre sous-catégories qui partagent les collocations recensées : Degré, Importance, Complexité et Nouveauté. Nous allons regarder attentivement les quatre dimensions avec leurs préférences en matière d'association avec des noms scientifiques ainsi que les différences disciplinaires.

Nous regardons d'abord la distribution des collocations non axiologiques.

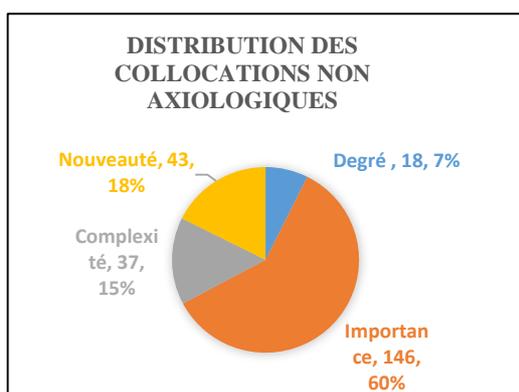


Figure 34 : Distribution des collocations non axiologiques

La dimension Importance prend une proportion plus forte, Degré demeure le plus faible, Nouveauté et Complexité se rapprochent et se trouvent au milieu.

Degré : Il se focalise sur la description de l'intensité, la quantité des noms scientifiques. Dans notre corpus, tout comme dans le corpus de Tutin (2010) sur la linguistique et l'économie, cette dimension n'est pas très représentée. Nous recensons 18 occurrences parmi 244 occurrences. Les adjectifs sont *fin* (9 occurrences), *complet* (2), *véritable* (4), *vrai* (1) et *faible* (1). Ils se combinent largement avec des noms d'objets construits comme *description, analyse, étude*, et quelques noms d'observables comme *phénomène*, et de caractérisation comme *caractéristique*.

(16)L'étude de celles-ci nécessite une **analyse (Objet construit)** linguistique plus **fine**, qui n'a pas été effectuée dans cette recherche. (Corpus Scientext, TAL)

(17)Par contre, d'un point de vue didactique, une **description (Objet construit)** plus **fine** des problèmes lexicaux des étudiants s'avère nécessaire, puisqu'elle constitue une première

étape vers des interventions visant à les corriger. (Corpus Scientext, Linguistique)

(18) Le **phénomène (Observable)** demeure relativement **faible** jusqu'à la crise de 1929 mais avant le début des années 1960 le taux de chômage ne dépasse jamais 2 % Il ne commence à augmenter qu'en 1967, de manière certes lente (de 1,6 % en 1966, à 3 % en 1974) mais apparemment irréversible; on observe cependant une baisse de sa durée. (Corpus Scientext, Sciences de l'éducation)

Cette dimension contient moins de collocations, et semble moins cruciale que les trois autres dimensions pour la stratégie rhétorique dans l'introduction et la conclusion des écrits scientifiques.

Complexité : cette dimension décrit la difficulté à comprendre, à analyser les produits ou les observations de l'activité scientifique. Dans cette dimension, nous trouvons des adjectifs courants comme *difficile, simple, complexe, délicat*, etc. Dans notre corpus, les adjectifs repérés sont : *complexe* (31 occurrences), *difficile* (4), *facile* (1) et *épineux* (1). Ils se combinent principalement avec les noms d'objets construits (29) comme *système, modèle, structure* etc. et d'observables (6) comme *question, phénomène*, mais aussi avec certains noms de processus comme *choix*, de caractérisation comme *fonction*. Les sciences appliquées et expérimentales ont tendance à évaluer principalement les noms d'objets (17/18), dont le plus fréquent renvoie à *système* (13/18). Les sciences humaines s'ouvrent à plusieurs dimensions (Processus, Objets construits, Observables, Caractérisation).

(19) J'y ai conduit des travaux de modélisation et d'identification de **systèmes (Objet construit) complexes**, dans le domaine de la dynamique véhicule et du contrôle adaptatif de freinage. (Corpus Scientext, Mécanique)

(20) Cette valeur moyenne de 1,96 se trouve déplacée vers le bas (~ 1,935) lorsque l'on considère les cas des monosubstitutions { Cys ? Ser } dans les ferrédoxines de type plante ou { Thiolate ? Hydroxyl } dans les **complexes modèles (Objet construit)**. (Corpus Scientext, Électronique)

(21) En effet, le découpage en phrases et la détermination de l'information empaquetable dans une phrase sont identifiés en Génération Automatique comme des **problèmes (Observable) complexes**, situés plutôt en amont dans le processus de Génération Automatique et souvent esquivés par les chercheurs de cette communauté (cf.). (Corpus Scientext, TAL)

(22) Notre objectif est de tester la Hiérarchie de Sélection des Auxiliaires de Sorace (2000) selon laquelle le **choix (Processus)** des auxiliaires est plus **facile** pour certains verbes que pour d'autres. (Corpus Scientext, Psychologie)

Du point de vue rhétorique, la dimension Complexité vise à souligner les difficultés rencontrées lors de l'activité scientifique, ce qui concourt dans un autre sens à mettre en valeur le poids de la contribution que l'auteur a réalisée dans la communauté scientifique. Même si elle n'est pas autant représentée que les dimensions Nouveauté ou Importance, elle joue toujours un rôle important dans l'introduction et la conclusion des écrits scientifiques, car les recherches ont prioritairement pour objectif d'observer les phénomènes, de chercher des solutions pour soulever les difficultés ou de proposer des explications aux phénomènes complexes. C'est aussi ce qui fait l'intérêt des recherches. Cela pourrait aussi aider l'auteur à expliquer au lecteur la difficulté de trouver une solution simple et impeccable.

Nouveauté : cette dimension se trouve la deuxième place parmi les quatre dimensions repérées, juste après la dimension Importance. À l'inverse du constat de Tutin (2010), que la dimension Nouveauté est accompagnée souvent par la dimension Temporalité, nos corpus ne fournissent aucune occurrence de cette dernière dimension.

Du point de vue évaluatif, cette dimension comporte deux polarités. Soit elle consiste à juger le côté original ou inédit des études effectuées, et dans ce sens les adjectifs employés seront nouveau, original, inédit qui met l'accent sur la nouveauté de l'objet évalué. Soit elle reflète les grandes traditions ou critères déjà établis, opposés à la nouveauté. Les deux polarités peuvent être utilisées ensemble.

Dans nos corpus, les adjectifs repérés sont : classique (20 occurrences), nouveau (19), original (2), innovant (2). Cette dimension se combine avec des noms d'objets construits (27 occurrences parmi 43) et d'observable (16/43). Les noms d'objets construits regroupent des noms comme *définition, modèle, méthode, conception...* Les noms d'observables mobilisés ici sont *facteur, élément, phénomène...* Nouveauté semble de préférence évaluer les noms des appareils scientifiques comme *méthode, phénomène, cas, élément.*

Classique peut pourtant s'accompagner non seulement des noms d'appareil scientifiques, mais aussi des noms relevant de la conceptualisation scientifique comme

théorie, définition, description.

(23) Concernant l'aspect cardinal, Piaget s'appuie sur sa **définition (Objet construit)** mathématique **classique** en termes de correspondance : deux collections sont de cardinal identique si, et seulement si, leurs éléments peuvent être mis en correspondance terme à terme. (Corpus Scientext, Biologie)

(24) Nous avons ensuite présenté, dans le chapitre II, différentes méthodes utilisées dans la littérature pour résoudre des écoulements bidimensionnels de la **théorie (Objet construit) classique** des jets de fluide parfait. Notre propre méthode de résolution est ensuite développée. (Corpus Scientext, Mécanique)

(25) On peut ajouter des dimensions estimées manquantes à des théories existantes, quitte à justifier que ces **nouveaux outils (Objet construit)** sont nécessaires, et qu'ils permettent de capter des **phénomènes (Observable) nouveaux** par rapport à ce que permettaient les anciens outils. (Corpus Scientext, Science de l'éducation)

L'adjectif *classique* dans les exemples (23) et (24) permet à l'auteur de se situer dans la lignée de la définition et de la théorie déjà établies et acceptées par sa communauté scientifique. Cette polarité manifeste une grande importance dans l'écrit scientifique, car selon Berthelot, un texte scientifique doit essayer de s'inscrire dans la communauté scientifique adéquate, et de manifester « un apport de connaissances reconnu par une communauté savante » (Berthelot, 2003:33).

L'exemple (25) explicite la fonction rhétorique remplie par l'adjectif *nouveau* : il envisage d'ajouter une dimension inédite ou originale à ce qui existe déjà. Bien que nous ne trouvions pas d'exemples de paires indiquant Nouveauté et Temporalité dans nos listes de collocations extraites automatiquement par ScienQuest, nous trouvons des cooccurrences en opposition sémantique dans l'exemple (25) :

Théories existantes, ancien outil vs nouveaux outils, phénomènes nouveaux

D'une discipline à l'autre, nous ne trouvons pas de grandes différences : *original* apparaît seulement dans la linguistique et le TAL ; *innovant* se trouve seulement dans l'électronique ; au contraire, *classique* et *nouveau* sont omniprésents dans toutes les disciplines. Les deux classes de noms en occurrences transcendent également toutes les huit disciplines.

Importance : cette dimension est prédominante. Elle apprécie si le rôle, l'intérêt, l'influence, le pouvoir ou les conséquences sont considérables. Les adjectifs repérés dans cette dimension sont : *principal, importance, fondamental, essentiel, majeur, central, crucial, déterminant, significatif, sérieux*. La figure 35 met un commun tous les adjectifs repérés.

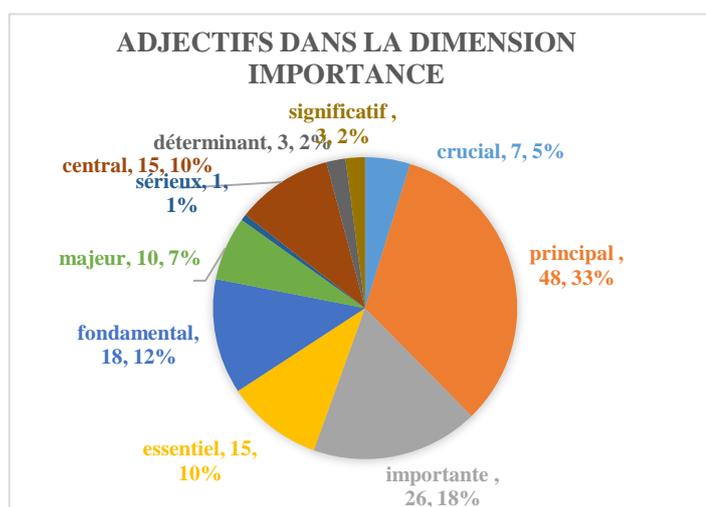


Figure 35 : Distribution des adjectifs dans la dimension Importance

Les adjectifs dans cette dimension se combinent dans la même proportion avec les quatre classes de noms scientifiques : la plupart se concentrent dans les noms d'objets construits et d'observables et de caractérisation, peu dans les noms de processus. Il faut signaler que les noms de caractérisation sont plus représentés dans cette dimension que les autres dimensions. Nous trouvons des collocations comme *rôle majeur, rôle important, fonction essentielle, caractéristique principale...*

(26)Le jeune enfant a donc une capacité analogique permettant de produire des approximations sémantiques tant nominales que verbales qui jouent un **rôle (Caractérisation) fondamental** dans la constitution progressive du lexique chez le jeune enfant et qui sont des indices de sa flexibilité cognitive. (Corpus Scientext, Linguistique)

(27)Les catégories Nom, Verbe et Adjectif portant une **fonction (Caractérisation)** sémantique **importante** nécessitent la saisie d'un peu plus de 2 caractères en moyenne avant d' être trouvées par l'utilisateur dans la liste des propositions. (Corpus Scientext, TAL)

Principal est le plus utilisé dans cette dimension, il peut travailler avec les quatre classes des noms scientifiques. Sa combinaison favorite est le nom d'objets construits

objectif, on compte 12 occurrences de la collocation *objectif principal* qui se produit dans cinq disciplines. Ce qui peut être compris comme une preuve incontestable de la caractéristique transdisciplinaire de ces collocations.

(28) **L'objectif (Objet construit) principal** de notre travail est d'étudier le fonctionnement des usages sociaux du langage en situation de communication. (Corpus Scientext, Linguistique)

(29) **Notre objectif (Objet construit) principal** était une meilleure compréhension de la nature et de l'évolution de ce type d'événements dans le cas d'une couche cisailée turbulente décollée, situation rendue plus complexe par la présence d'une recirculation. (Corpus Scientext, Mécanique)

Important est le deuxième adjectif le plus fréquent dans la dimension Importance, il s'associe en général avec des noms d'objets construits, d'observables et de caractérisation. Il partage avec l'adjectif *principal* beaucoup de noms comme *rôle*, *caractéristique*, *contribution*, *méthode* etc.

Fondamental se trouve en troisième place. Il se combine avec des noms de processus comme *choix*, des noms d'objets construits comme *concept*, *idée*, des noms d'observables comme *question*, *problème* et des noms de caractérisation comme *rôle*, *caractéristique*. Il a l'air plus sujet à se produire dans les sciences humaines tandis que l'adjectif *essentiel*, quatrième représentatif, présente davantage dans les sciences non humaines (9 occurrences dans les sciences appliquées et expérimentées, 6 occurrences dans les sciences humaines qui ont une taille de corpus beaucoup plus large).

Central, *déterminé*, et *significatif* ont été repérés seulement dans les sciences humaines avec quatre classes de noms.

Sérieux est un mot polysémique, il peut avoir une évaluation bien axiologique, mais dans notre cas, nous préférons le ranger dans la dimension Importance que dans la classe axiologique.

(30) À son stade le plus avancé, elle provoque la chute des dents. La forte prévalence de ces maladies, leurs conséquences cliniques et les coûts importants liés à leurs traitements ou à ceux de leurs séquelles en font un **problème (Observable) sérieux** de santé publique. (Corpus Scientext, Médecine)

Sérieux dans l'exemple (30) peut être interprété comme « grave, important, qui ne peut pas être traité à la légère ». Cet adjectif permet de juger l'importance, la gravité d'un problème ici de santé publique.

Dans nos listes de collocations, nous avons une autre collocation avec cet adjectif, mais nous l'avons classé dans les collocations axiologiques, car cette fois-ci, l'adjectif *sérieux* permet de juger la bonne qualité d'un évènement (des études).

(31) En effet, pour nous, le traitement automatique des langues ne peut être abordé sans **étude** linguistique **sérieuse** ; toutes les réalisations de l'équipe répondent à cette condition. (Corpus Scientext, TAL)

Dans l'exemple (31), *sérieux* est utilisé pour décrire une étude approfondie, bien faite. En conséquence, il est plus convenable de traiter la collocation *étude sérieuse* comme l'une des collocations axiologiques.

Au niveau rhétorique, cette dimension paraît cruciale pour l'écrit scientifique. L'auteur utilise abondamment des adjectifs d'importance pour évaluer les noms scientifiques afin de bien guider le lecteur. Il semble que l'auteur soit censé sélectionner des informations essentielles qui facilitent la poursuite de la lecture (Tutin, 2010). Bien sûr, par ce processus de sélection consciencieuse, l'auteur montre sa capacité à gérer des informations obtenues lors des recherches.

6.2.1.3. Mise en commun des analyses sur la combinaison entre les constituants des collocations

Nous inspirant du travail de Tutin sur les collocations évaluatives, nous proposons à la fin des analyses sémantiques des constituants des collocations un tableau récapitulatif :

Dimension des adjectifs	Classes des noms scientifiques	Exemples	Fonction rhétorique essentielle
Axiologiques	Largement utilisés avec Objet construit, Observation, peu de Caractérisation et de Processus	Traitement efficace Analyse pertinente Communication riche Bonne qualité	Juger la valeur des objets à évaluer, souvent positive, pour apprécier le travail de l'auteur afin de le promouvoir.
Degré	Se combiner	Analyse fine	Décrire l'intensité et la quantité

	principalement avec Objet construit, peu de noms d'observable et de caractérisation	Véritable outil Vrai phénomène Caractéristique fine	des objets à évaluer. Cette dimension s'avère moins cruciale que les autres.
Complexité	Se combinent principalement avec Objet construit et Observable, mais aussi avec certains noms de processus, de caractérisation.	Système complexe Structure complexe Phénomène difficile Question difficile Fonction complexe Choix facile	Mettre l'accent sur la difficulté des recherches pour concourir à soulever l'intérêt des recherches et expliquer pourquoi une solution simple est difficile à proposer.
Nouveauté	Les noms d'objet construit Les noms d'observable	Définition classique Structure originale Nouvelle méthode Méthode innovante	Montrer les connaissances dans le domaine. Souligner la nouveauté, l'originalité des travaux de l'auteur.
Importance	Les noms d'objet construit Les noms d'observable Les noms de caractérisation Les noms de processus	Structure principal Facteur essentiel Rôle important Choix déterminant	Particulièrement crucial pour la stratégie rhétorique dans l'introduction et la conclusion. L'auteur sélectionne des éléments fondamentaux pour mieux guider la lecture, et aussi pour monter sa compétence de mieux gérer des informations.

Tableau 17 : Association principales entre des constituants des collocations évaluatives et leurs
fonctions rhétoriques

6.2.2. Analyses sur les patrons syntaxiques et syntagmatiques des collocations évaluatives

L'une des propriétés des collocations porte sur le fait qu'elles peuvent subir des changements syntaxiques et que même sous la même structure syntaxique, des variations syntagmatiques peuvent se réaliser. Dans cette partie, nous allons interroger les patrons syntaxiques les plus présents dans les corpus et puis regarder au sein de chacun les variations syntagmatiques afin de souligner des phénomènes courants avec leur fonction

rhétorique.

En principal, nous rencontrons deux structures syntaxiques :

1. Le collocatif et la base constituent un syntagme nominal, dont le premier joue le rôle d'épithète pour qualifier la base. La structure syntaxique correspond à la forme canonique:

Nom scientifique + Adjectif évaluatif (épithète)①

Cet ordre pourrait être inversé en fonction des adjectifs. Tout comme ce que les exemples (32) et (33) montrent.

(32)Au plan fondamental, elle représente un **bon exemple** de la validité de l'approche de "désorphanisation" des récepteurs orphelins pour l'identification de nouveaux peptides bio actifs et de nouvelles cibles thérapeutiques. (Corpus Scientext, Biologie)

(33)En outre, le bénéfice de la consigne d'imagerie sur les performances de mémorisation des mots présentés est conforme aux **résultats classiques** qui rendent compte de l'effet renforçateur du double codage sur l'organisation des éléments à apprendre (Paivio , 1991). (Corpus Scientext, Sciences de l'éducation)

2. La base et le collocatif sont liés par un verbe support. Le collocatif fait office de l'attribut de la base pour mettre l'accent sur le collocatif. La forme correspondante est :

Nom scientifique + Verbe support + Adjectif évaluatif (attribut).....②
--

(34)Globalement, les **résultats sont satisfaisant**, même si des améliorations sont encore possibles, puisque avec ou sans mot outils, l'extraction LEM avec PC et ABIJ se situe à environ 10 % de précision et 13 % de rappel des résultats optimaux (définis par rapport aux contraintes de notre implémentation). (Corpus Scientext, TAL)

(35)Après avoir rappelé que l'adaptation en puissance entre des cellules intégrées sur silicium fonctionnant autour de 950 MHz n'était pas indispensable, nous avons démontré que les **notions** habituellement utilisées par les concepteurs **restaient valables** pour des cellules non adaptées. (Corpus Scientext, Électronique)

Dans cette structure, le verbe support le plus fréquent est le verbe *être*, mais nous trouvons aussi d'autres verbes support comme *rester*, *demeurer* (l'exemple 18).

Dans nos corpus, la plupart des collocations respectent la structure canonique et peu de collocations suivent le patron avec l'insertion d'un verbe support. La figure 36 permet

un accès immédiat de leur proportion dans les huit disciplines.

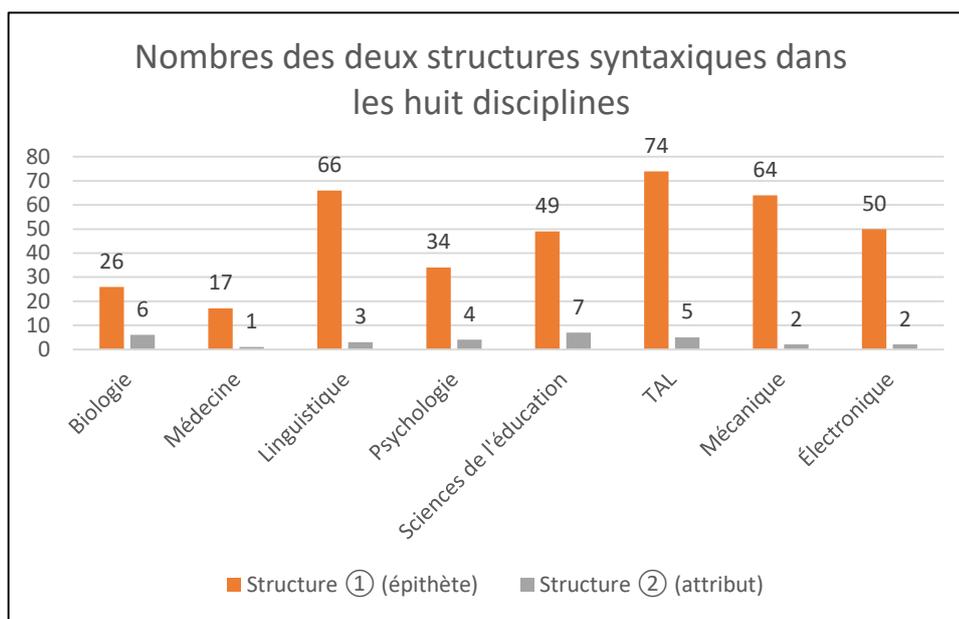


Figure 36 : Distribution des variations syntaxiques des collocations évaluatives

(N_{scientifique} + Adj_{évaluatif})

La figure 36 montre que pour toutes les disciplines, la structure syntaxique préférée est la structure canonique où l'adjectif fait fonction d'épithète. Néanmoins, une autre étude en anglais montre d'autres résultats qui nous intéressent beaucoup. Pérez-Llantada Auría (2008) a mené une étude sur les adjectifs évaluatifs en anglais dans l'introduction des deux disciplines (Linguistique appliquée et Science informatique). Ses recherches ont abouti au constat de la supériorité de la présence des adjectifs attribués par rapport aux adjectifs épithètes dans la partie introductive de ces deux disciplines. Après avoir comparé les deux études, il faut remarquer que Pérez-Llantada Auría prend en compte la structure *it + copula verb + evaluative adjectif + to* (Il + copule + adjectif évaluatif + pour) qui n'a pas été prise en compte dans notre étude, car la nôtre est consacrée aux collocations évaluatives (l'association entre des noms scientifiques et des adjectifs évaluatifs), et pas seulement aux adjectifs évaluatifs.

Par rapport à la structure canonique, la structure ② (Nom scientifique + Verbe support + Adjectif évaluatif (attribut)) exprime plus d'emphase lorsque l'adjectif évaluatif est placé après le verbe, parce que les informations importantes sont sujettes à se placer

après le verbe.

A l'exception des deux structures syntaxiques, les collocations évaluatives en question accepteraient des insertions comme les adjectifs, les adverbes pour renforcer ou affaiblir la dimension évaluative sans changer le sens original. Parmi les insertions, nous en trouvons trois types.

1. Ajout d'une autre épithète terminologique

(36) Les catégories Nom, Verbe et Adjectif portant une **fonction sémantique importante** nécessitent la saisie d'un peu plus de 2 caractères en moyenne avant d'être trouvées par l'utilisateur dans la liste des propositions. (Corpus Scientext, TAL)

(37) Ce risque important est en partie attribuable à une prévalence accrue de **facteurs de risque "classiques"**. (Corpus Scientext, Médecine)

L'ajout d'une épithète spécifique ou terminologique peut aider à préciser le nom scientifique. Dans l'exemple (36), *sémantique* restreint le nom *fonction* au domaine linguistique en ajoutant une dimension terminologique : c'est la fonction de la sémantique qui est en question, pas d'autres aspects linguistique comme la syntaxe, la rhétorique, etc. L'épithète pourrait être un adjectif ou un nom introduit par une préposition. *Facteurs de risque* est bien le cas.

2. Ajout d'un adverbe

(38) Cette **question** est **particulièrement importante** étant donné la fréquence avec laquelle les techniques d'imagerie guidée sont utilisées dans les situations d'entretiens thérapeutiques ou judiciaires. (Corpus Scientext, Psychologie)

(39) Les **données** provenant d'études épidémiologiques et cliniques sont **actuellement insuffisantes** pour préciser le rôle et l'interaction des différents facteurs. (Corpus Scientext, Médecine)

L'adverbe a pour fonction de modifier ou circonscrire l'évaluation exprimée par l'adjectif évaluatif. Dans l'exemple (38), l'adverbe *particulièrement* amplifie l'importance de la question discutée dans le texte. Nos corpus montrent une tendance forte pour des adverbes d'intensité ou d'amplification comme *très*, *particulièrement*, *relativement*, *réellement*, *bien sûr*, *bien*, *assez*, mais aussi des adverbes circonstanciels comme *actuellement*, *potentiellement*.

3. Le comparatif et le superlatif de l'adjectif évaluatif

Une autre stratégie pour modifier le sens d'évaluation des adjectifs évaluatifs est de mettre les adjectif évaluatifs en comparatif ou superlatif. Cette stratégie vise à graduer l'évaluation par degré ou à permettre une comparaison avec d'autres éléments.

(40) Nous avons également démontré la présence de moyens d'évitement de la complexité syntaxique, caractérisés par une prédominance des **structures (Objet construit) les moins complexes** (ne nécessitant pas de flexion verbale). (Corpus Scientext, Linguistique)

(41) Notre objectif est de tester la Hiérarchie de Sélection des Auxiliaires de Sorace (2000) selon laquelle le **choix (Processus)** des auxiliaires est plus facile pour certains verbes que pour d'autres. (Corpus Scientext, Psychologie)

Parfois, nous pouvons trouver deux ajouts dans une phrase :

(42) Elles ont globalement des **projets (Objet construit) professionnels moins ambitieux** car pour elles, vie professionnelle et vie familiale entreront souvent en compétition. (Corpus Scientext, Sciences de l'éducation)

Les études sur les structures syntaxiques nous permettent de comprendre les caractéristiques syntaxiques de nos collocations évaluatives du genre $N_{\text{scientifique}} + \text{Adj}_{\text{évaluatif}}$. Les insertions possibles décrivent de façon sommaire leur utilisation réelle dans le contexte avec d'autres éléments comme l'adverbe, ce qui va contribuer à la conceptualisation des propositions didactiques dans le dernier chapitre.

6.3. Transdisciplinarité des collocations évaluatives

Les analyses quantitative et qualitative révèlent certaines différences et similarités disciplinaires. Dans cette partie, nous rechercherons des preuves de la transdisciplinarité des collocations évaluatives. Cette confirmation va jeter des fondements pour la légitimité de l'enseignement des collocations évaluatives au profit des apprenants non natifs provenant de différentes disciplines.

Nous listons ici les 13 collocations évaluatives les plus fréquentes (dont le nombre d'occurrences ≥ 5) dans nos listes :

N°	Collocations	Nombres d'occurrences dans l'ensemble des corpus (introduction, conclusion)	Disciplines
----	--------------	---	-------------

		des huit disciplines)	
1	structure cohérente	32	Mécanique
2	système complexe	15	Mécanique, Électronique, Biologique
3	objectif principal	12	Linguistique, Psychologie, Science de l'éducation, TAL, Mécanique
4	rôle important	11	Biologie, Linguistique, Psychologie, Sciences de l'éducation, TAL
5	meilleur résultat	9	Linguistique, TAL, Électronique
6	résultat principal	7	Linguistique, Psychologie, Sciences de l'éducation, TAL
7	nouveau modèle	6	Linguistique, Mécanique, Électronique
8	résultat encourageant	6	Linguistique, Psychologie, TAL, Mécanique
9	résultat satisfaisant	6	Linguistique, Psychologie, TAL, Mécanique
10	facteur classique	5	Médecine
11	facteur principal	5	Linguistique, Science de l'éducation, Électronique
12	structure complexe	5	Linguistique, TAL, Électronique
13	question principal	5	Biologie, Linguistique, Science de l'éducation

Tableau 18 : Les 13 collocations évaluatives les plus fréquentes et leurs disciplines

Parmi les 13 collocations évaluatives, seulement deux collocations n'appartiennent qu'à une seule disciplines (*structure cohérente* pour la Mécanique, *facteur classique* pour la Médecine), les autres sont toutes partagées par au moins deux disciplines. Comme la taille de nos corpus est limitée (nous rappelons ici que pour un corpus de spécialité, le seuil de la taille de corpus s'élève à 500 000 mots), nous supposons ici que les collocations seront présentes dans plus de disciplines à la suite de l'augmentation de la taille. Malgré cette taille relativement petite, nos résultats fournissent des preuves bien concluantes sur la transdisciplinarité des collocations évaluatives du genre $N_{\text{scientifique}} + \text{Adj}_{\text{évaluatif}}$.

Bilan

Ce chapitre était consacré entièrement aux analyses des collocations évaluatives que nous avons repérées dans la deuxième partie du mémoire. Les analyses quantitatives montrent que tant l'introduction que la conclusion ont tendance à utiliser plus de

collocations non axiologiques que de collocations axiologiques, ces dernières se présentant davantage dans la partie conclusive pour évaluer et promouvoir le travail fait par l'auteur. Les collocations non axiologiques dans nos corpus se classent en quatre dimensions : Degré, Complexité, Importance et Nouveauté, elles se combinent avec des noms scientifiques également catégorisés en quatre classes : des noms d'objets construits, d'observations, de processus et de caractérisation, dont les deux premiers sont plus représentatifs que les deux classes restantes. La combinaison entre les dimensions des adjectifs et des classes de noms varie en fonction des disciplines. En revanche, les fonctions rhétoriques remplies par les collocations axiologiques et les quatre dimensions des collocations non axiologiques ne sont pas différentes d'une discipline à l'autre. En observant de plus près les structures syntaxiques des collocations évaluatives, nous trouvons que le collocatif (l'adjectif évaluatif) prend plus souvent la fonction d'épithète que d'attribut dans le texte, et dans ce domaine, il n'y a pas de différence manifeste entre les disciplines. À la fin, l'étude sur les collocations les plus fréquentes démontre que les collocations en question sont bien transdisciplinaires, ce qui rend possible la didactique des éléments collocationnels auprès des apprenants non natifs s'inscrivant dans différents domaines scientifiques.

Limite des résultats

Tous nos résultats sont obtenus à travers l'observation de notre corpus : les parties introductive et conclusive dans les huit disciplines. En raison de la taille inégale des parties dans chaque discipline, et la taille en générale plus petite que la norme (500 000 mots pour un corpus spécialisé (Williams, 1999), les résultats d'analyses ne peuvent pas représenter de façon incontestable la situation de l'emploi des collocations évaluatives du patron $N_{\text{scientifique}} + \text{Adj}_{\text{évaluatif}}$, mais ils peuvent quand même nous fournir des hypothèses plus précises pour une étude plus approfondie sur un corpus plus large et équilibré.

Partie IV - Propositions didactique

Chapitre 7. Propositions didactiques à l'aide du corpus Scientext

Notre dernière partie est dédiée aux propositions didactiques pour aider les étudiants non natifs à mieux évaluer dans l'introduction et la conclusion de l'écrit scientifique à travers des collocations transdisciplinaires évaluatives du genre $N_{\text{scientifique}} + \text{adj}_{\text{évaluatif}}$. Nos analyses et discussions effectuées dans la troisième partie nous permettent de dégager un accord sur les caractéristiques de l'évaluation dans les deux parties textuelle de l'écrit scientifique :

1. L'introduction et la conclusion mobilisent moins de collocations axiologiques que de collocations non axiologiques ;

2. La conclusion a tendance à utiliser plus de collocations axiologiques que l'introduction ;

3. Parmi les collocations axiologiques, la polarité positive est prédominante, c'est à dire l'auteur de l'écrit scientifique a tendance à apprécier ses travaux ;

4. Parmi les collocations non axiologiques, les dimensions Importance et Nouveauté sont les plus présentes pour indiquer l'importance et le caractère inédit des recherches effectuées ;

5. Les dimensions Degré et Complexité sont moins présentes.

Les cinq remarques que nous avons faites sur l'utilisation des collocations évaluatives dans l'introduction et la conclusion vont amener à faire des propositions didactiques sur l'enseignement de l'évaluation dans les deux parties textuelles de l'écrit scientifique. De plus, les classements préétablis des collocations évaluatives constituent un support de l'enseignement pour guider et vérifier les acquis des apprenants lors de leur apprentissage des collocations ainsi que leurs fonctions rhétoriques connotées.

7. 1. Corpus Scientext comme outil d'enseignement

Nos propositions didactiques se situent dans l'approche de corpus pour l'enseignement et l'apprentissage de la langue française étrangère. Les linguistes utilisent

les corpus pour mener des recherches linguistiques, les didacticiens tels que Boulton (2007, 2010, 2014), Chambres (2005, 2010), ont aussi remarqué l'intérêt de leur application dans le champ didactique des langues.

7.1.1. Corpus comme outil d'enseignement : avantages

Le corpus, regroupant un nombre énorme d'occurrences des éléments linguistiques des natifs, accorde aux apprenants

un accès aux intuitions de milliers de locuteurs natifs (Frankenberg-Garcia 2005 : 192) et une richesse de contextes pertinents [...] promouvant ainsi une sensibilisation à la langue dans des contextes authentiques. (Boulton, 2010: 100)

En ce qui concerne l'apprentissage du lexique, le corpus peut manifester aussi son avantage en permettant à l'apprenant un apprentissage efficace du lexique à l'aide de l'analyse des concordances. Le lexique à rechercher est toujours aligné dans les concordances, ce qui facilite leur sélection et observation, et favorise par la suite l'apprentissage inductif qui vise à laisser découvrir à l'apprenant les caractéristiques syntaxiques, sémantiques et rhétoriques des éléments linguistiques en question (Cavalla, 2013: 168).

7.1.2. Scientext comme outil d'enseignement pour les collocations N_{scientifique} + Adj_{évaluatif}

Malgré le corpus Scientext est initialement conçu pour un but linguistique consistant à découvrir la phraséologie dans l'écrit scientifique, il peut aussi s'employer comme un outil d'enseignement et d'apprentissage pour aider les apprenants à mieux rédiger leur écrit scientifique. Cavalla a indiqué également le rôle didactique que le corpus Scientext peut jouer dans ce sens : « pour faire face aux difficultés des apprenants de Master et doctorat à rédiger leurs écrits universitaires scientifiques, Scientext constitue un outil qui peut les aider à découvrir certaines des attentes, en matière d'écrits, des universités françaises » (Cavalla, 2013: 168).

Nos collocations évaluatives ne sont pas seulement un genre de collocations qui sont fréquentes dans l'écrit scientifique, elles manifestent aussi la stratégie rhétorique de

l'auteur pour évaluer et s'évaluer. Scientext est un outil parfait qui regroupe des collocations du genre $N_{\text{scientifique}} + \text{Adj}_{\text{évaluatif}}$ et propose des contextes bien scientifiques variés transcendant des disciplines. Comment l'exploiter pour finalité didactique reste donc une question centrale.

7.2. Exploitation du corpus Scientext : principes et processus

7.2.1. Principes d'exploitation du Corpus Scientext

L'exploitation des corpus pour l'enseignement des langues est toujours l'objet de discussions. Principalement, nous avons deux approches : approche directe et indirecte (Chambers 2010 : 15-16).

1. L'approche directe place l'apprenant directement en face des corpus numériques. D'une part, elle favorise l'autonomie de l'apprenant en l'encourageant à formuler ses propres hypothèses et à créer ses propres concordances, d'autre part, elle aide à l'autocorrection des écrits scientifiques de l'apprenant, ce qui lui permet de constituer sa confiance en soi. Pourtant, l'accès direct aux corpus à grande taille pose aussi des problèmes concernant les équipements, la formation des outils numériques, etc., ce qui peut être résolu par l'approche indirecte.

2. L'approche indirecte renvoie à l'enseignant pour la création d'un petit corpus constitué des concordances sélectionnées à partir des corpus numériques. Ses avantages consistent en quatre aspects :

- a. Le nombre des concordances est limité mais les contextes pourraient être variés, car ils sont sélectionnés spécifiquement par l'enseignant ;
- b. L'enseignant peut concevoir des séquences didactiques pour guider l'apprenant ;
- c. Il ne nécessite pas d'outils numériques, il est donc facile à prendre en main par l'apprenant ;
- d. L'approche indirecte peut constituer comme une période de transition vers l'approche directe.

La mise en lumière des deux approches d'exploitation des corpus nous permet de

comprendre qu'elles sont dans une certaine mesure complémentaire. Nous proposons par conséquent de commencer par l'enseignement des collocations à partir des concordances sélectionnées dans le corpus Scientext afin de laisser l'apprenant se familiariser avec l'apprentissage direct sur des corpus, et par la suite, nous le guiderons vers l'autonomie d'apprentissage en lui apprenant à employer des outils numériques du corpus Scientext pour qu'il puisse consulter directement les corpus numériques à tout moment où il en a besoin.

7.2.2. Processus d'enseignement dans l'approche indirecte

Notre première étape d'enseignement consiste à enseigner les collocations évaluatives en créant des séquences didactiques à partir des concordances sélectionnées dans le corpus Scientext. Avant d'entrer dans l'élaboration des séquences didactiques, il faut d'abord considérer la constitution des petits corpus.

7.2.2.1. Constitution des petits corpus à enseigner

Bien sélectionner des concordances ou des petits corpus à enseigner constitue une étape centrale. Principalement, nous avons deux formes de composition pour de petits corpus :

- a. Soit nous fabriquons notre petit corpus avec des concordances : nous regardons une seule fonction rhétorique, par exemple la dimension Importance dans l'introduction et la conclusion. Nous choisissons les collocations évaluatives qui sont apparentées à cette dimension pour construire notre liste de concordances. De ce fait, chaque liste va traiter une dimension sémantique qui peut être mise en commun à la fin de l'enseignement.
- b. Soit nous produisons un corpus regroupant quelques introductions ou conclusions intégrales (en cas de la présence simultanée des deux parties textuelles, leur nombre seront identiques, par exemple 2 introductions + 2 conclusions). Ce genre de corpus permet à l'apprenant d'avoir une compréhension globale des fonctions rhétoriques mobilisées dans les deux parties textuelles, ce qui donne la possibilité de découvrir certaines similarités ou différences entre l'emploi des fonctions rhétoriques, ainsi que

celles entre les deux parties textuelles.

Dans la constitution des deux genres de petits corpus, il faut faire attention à la question des lexiques terminologiques dans les concordances ou les textes choisis. Pour la première composition, les concordances peuvent être tirées de différentes disciplines, mais l'enseignant doit faire attention au nombre des lexiques terminologiques inhérents à la disciplines. Dans le but de ne pas créer des difficultés de compréhension, il faut choisir des concordances qui ont moins de termes disciplinaires ou les concordances dont les termes dans la phrase ne posent pas de problème pour comprendre les collocations. Quant à la deuxième composition qui regroupe l'introduction ou la conclusion intégrale, elle conviendrait davantage à un public homogène qui partage le même domaine scientifique. Il est donc conseillé de choisir les textes provenant de la discipline où s'inscrivent les apprenants: les étudiants en biologie vont comprendre mieux les concordances dans des textes biologiques par exemple.

Le support des petits corpus sélectionnés pourraient être le papier ou la fiche électronique sur l'ordinateur.

7.2.2.2. Processus d'enseignement des collocations évaluatives

Pour mieux guider l'apprentissage des éléments linguistiques, l'enseignant est censé créer des activités pédagogiques autour des petits corpus constitués. Les activités se déroulent souvent en quatre phases (inspiré par Cavalla 2013: 8) :

1. Repérage des collocations évaluatives dans le petit corpus ;
2. Observation des propriétés sémantique et syntaxique des collocations évaluatives, ainsi que leurs fonctions rhétoriques et contextes d'utilisation ;
3. Mémorisation et consolidation des collocations ;
4. Réemploi des collocations évaluatives par l'activité d'écrit.

Maintenant, nous allons détailler les quatre phases en prenant un exemple basé sur la première composition de petits corpus que nous venons de décrire en haut.

Exemple :

- Activité 01

La liste des concordances traitant la dimension Importance : (nous choisissons les collocations qui comportent les six adjectifs les plus fréquents dans le corpus : *principal, important, fondamental, central, essentiel, majeur.*)

1. Les catégories Nom, Verbe et Adjectif portant une fonction sémantique importante nécessitent la saisie d'un peu plus de 2 caractères en moyenne avant d'être trouvées par l'utilisateur dans la liste des propositions. (Corpus Scientext, Linguistique)
2. Mais comme nous l'avons montré, la question du paramétrage des modèles mis en œuvre est centrale , le succès d'une technique dépendant largement de ce type de réglage. (Corpus Scientext, TAL)
3. Cette question est particulièrement importante étant donné la fréquence avec laquelle les techniques d'imagerie guidée sont utilisées dans les situations d'entretiens thérapeutiques ou judiciaires. (Corpus Scientext, Psychologie)
4. L' objectif principal est d'étudier si certains facteurs agissant sur la sélection et l'emploi des auxiliaires influencent la compréhension de phrase . (Corpus Scientext, Psychologie)
5. Une autre caractéristique fondamentale des théories élevées au rang de modèles, est leur caractère prédictif. C'est un point que nous examinerons dans le cas de la TSD. (Corpus Scientext, Sciences de l'éducation)
6. Chez les végétaux, les mécanismes épigénétiques ont pour fonction essentielle d'inactiver les éléments génétiques envahissants comme les virus. (Corpus Scientext, Biologie)
7. Nous avons vu que la manipulation lexicale joue un rôle très important dans le développement du scamto. (Corpus Scientext, Linguistique)
8. Néanmoins, le problème majeur de ces systèmes de suppléance réside, comme on peut s'en douter, dans la lenteur de la composition des messages. (Corpus Scientext, TAL)
9. C'est donc par clics successifs qu'il va sélectionner les éléments de son message sur le clavier virtuel, élément central de l'interaction. (Corpus Scientext, TAL)
10. Les principes fondamentaux de la théorie des primitives sémantiques (Goddard, 1994). (Corpus Scientext, Psychologie)
11. D'autre part, et ce grâce à cette intégration, ce travail permet d'apporter une contribution importante aux recherches existantes. (Corpus Scientext, Psychologie)
12. Dans un second temps, nous avons entrepris l' étude majeure de ce travail de thèse. (Corpus Scientext, Électronique)
13. Il semble qu'on soit maintenant assez délivré du comportementalisme pour oser en retenir une idée centrale : même si l'acquisition et l'usage du langage sont des activités complexes et spécifiques, il

est bien probable qu' ils soient réalisés à partir de capacités très simples et très générales. (Corpus Scientext, Linguistique)
14. L'efficacité d'une méthode traitement des congestions suivant ces critères essentiels : Elle doit être techniquement faisable et la plus flexible d'utilisation. (Corpus Scientext, Électronique)
15. Avant tout, nous avons donc voulu revenir sur les principes physiques de base conduisant au bruit de phase dans un oscillateur, en faisant appel aux paramètres essentiels mis en jeu dans ce processus. (Corpus Scientext, Électronique)
16. Dans le second Chapitre, nous allons décrire les principales méthodes de traitement des congestions ayant été appliquées jusqu'à maintenant, en les explicitant sur le plan théorique. (Corpus Scientext, Électronique)

Observez les phrases listées dans le tableau qui sont extraites du concordancier du corpus Scientext et répondez aux questions ci-dessous par groupe :

1. Repérez les collocations du genre Nom + Adjectif dans les phrases, quels sont les adjectifs mobilisés dans les collocations ?

2. Observez les collocations repérées. Quelles sont leurs structures syntaxiques ? Remplissez le tableau ci-dessous pour les classer selon leurs structures syntaxiques.

	Structures syntaxiques	Collocations
1		
2		
3		

3. Observez les contextes gauche et droit des collocations, résumez ce que vous avez remarqué.

4. Quelle valeur véhiculent les adjectifs dans les collocations, quelle fonction est remplie par des collocations dans les phrases?

La première question consiste à faire découvrir aux apprenants par groupes les collocations évaluatives dans les phrases. Nous avons choisi plusieurs noms scientifiques différents avec six adjectifs pour que l'apprenant se focalise sur la valeur de l'adjectif et la fonction rhétorique des collocations évaluatives. La diversité des noms scientifiques comme *travail, rôle, question* peut laisser comprendre que les collocations en question sont très courantes dans l'écrit scientifique et que les apprenants peuvent les utiliser

fréquemment dans leur propre écrit scientifique.

Les deuxième et troisième questions demandent aux apprenants de regarder au plus près les collocations évaluatives pour récapituler leurs syntaxiques possibles dans les contextes, ce qui contribue au repérage des structures syntaxiques des collocations en question. Par exemple dans cet exercice, les collocations répondent aux deux structures principales : Nom + adjectif comme *rôle important*, Nom + Verbe + adjectif comme *question est centrale*. Parfois, au sein de chaque structure, certains éléments comme un autre adjectif, ou adverbe peuvent s'y intégrer, comme *fonction sémantique importante*, *question est particulièrement importante*. De plus, pour certaines collocations, les contextes des deux côtés sont assez intéressants :

Collocation	Contexte gauche	Contexte droit
<i>Rôle important</i>	Verbe <i>jouer</i>	Préposition <i>dans</i>
<i>Contribution importante</i>	Verbe <i>apporter</i>	Préposition <i>à</i>

Les contextes gauche et droit de la première collocation dans le tableau ci-dessus constituent avec cette collocation une autre collocation verbale : *jouer un rôle important dans*. La deuxième collocation permet de rendre compte que le nom *contribution* est souvent en cooccurrence avec la préposition *à*. Avec ces deux exercices, l'apprenant comprend mieux la morphologie des collocations évaluatives ainsi que les contextes d'utilisation.

La dernière question cherche à faire maîtriser par l'apprenant la valeur sémantique des adjectifs ainsi que la fonction rhétorique des collocations dans les phrases.

La première activité consiste à traiter les différentes dimensions sémantiques et rhétoriques où s'inscrivent les collocations évaluatives, pour d'autres dimensions, nous pouvons sensibiliser de même façon comme la dimension Importance. En ce qui concerne les collocations axiologiques (nous proposons de les enseigner sous le nom *collocations de jugement* pour éviter le terme *axiologique* qui poserait un problème à l'apprenant), nous pouvons ajouter une question supplémentaire: discriminer des collocations en fonction de leur valeur positive ou négative.

Ce genre d'activités correspond aux deux premières phrases de l'enseignement/apprentissage : repérage de collocations et sensibilisation sur leurs propriétés syntaxiques, sémantiques et rhétoriques. Par la suite, nous proposons la deuxième activité ayant pour objectif la mémorisation et la consolidation des collocations évaluatives.

Activité 02 :

La liste des collocations de plusieurs valeurs sémantiques :

1. Il semble qu'on soit maintenant assez délivré du comportementalisme pour oser en retenir une idée centrale : même si l'acquisition et l'usage du langage sont des activités complexes et spécifiques, il est bien probable qu'ils soient réalisés à partir de capacités très simples et très générales. (Corpus Scientext, Linguistique)
2. Cela signifie que la représentation du phénomène à travers un outil n'est représentatif que de certaines caractéristiques du phénomène (dépendante de l'outil) et non pas du phénomène complet . (Corpus Scientext, Mécanique)
3. Les résultats sont satisfaisants , le taux d'économie de saisies est d'environ 55 % pour une liste de 9 propositions et d'environ 50 % pour une liste de 5 propositions. (Corpus Scientext, Sciences de l'éducation)
4. Son rôle est essentiel car elle contient tout le programme nécessaire au déroulement de la vie cellulaire. (Corpus Scientext, Biologie)
5. Nous avons délaissé le critère classique de synonymie - deux façons de dire la même chose - pour appliquer systématiquement le critère d'identité du fonctionnement structural. (Corpus Scientext, Linguistique)
6. Fait surprenant : après le détour de décennies de recherche en TAL, qui ont montré à quel point il était important d'intégrer une description fine des phénomènes linguistiques mis en jeu par la traduction. (Corpus Scientext, TAL)
7. Or, cette notion est difficile à définir à l'oral et peu accessible à des enfants qui n'ont aucune connaissance de l'écrit. (Corpus Scientext, Linguistique)
8. Il nous semble que le facteur principal pour transformer les attitudes des usagers d'une année sur l'autre, d'un groupe à l'autre, est le facteur temps : temps d'appropriation et d'utilisation de la plate-forme. (Corpus Scientext, Sciences de l'éducation)
9. Ainsi, les données provenant d'études d'intervention sont actuellement insuffisantes pour établir des recommandations précises quant à la correction des différents facteurs de risque. (Corpus Scientext, Médecine)
10. Dans la deuxième partie, les questionnements sur les niveaux de structuration des textes et le jeu des marques formelles se poursuivent avec la mise en relation de nouveaux modèles et la

focalisation sur un objet textuel la définition dans des corpus de textes scientifiques ou techniques. (Corpus Scientext, Linguistique)
11. Offrir une solution pertinente , crédible, générique, robuste et facile de réglage était déjà un fait délicat pour un moteur essence dans une configuration standard. (Corpus Scientext, Mécanique)
12. Les enseignants et les étudiants évoquent des difficultés liées à l'apprentissage des concepts fondamentaux et des modes de raisonnement en informatique. (Corpus Scientext, Sciences de l'éducation)
13. Si le bi-texte permet de mettre en regard des structures complexes , la détection de régularités peut concerner tout phénomène : diathèse, temps et modes verbaux, parties du discours, fonctions grammaticales, cas, etc. (Corpus Scientext, TAL)
14. Bien plus, elles posent des problèmes nouveaux et intéressants à la linguistique théorique, par exemple en leur soumettant comme sujet d'étude des structures linguistiques originales , propres aux textes techniques.(Corpus Scientext, TAL)
15. La production ne paraît toutefois pas nécessaire à la réalisation de traitements efficaces si ceux-ci sont opérationnalisés par d'autres tâches. (Corpus Scientext, Psychologie)

Consigne: Dans les exercices précédents, vous avez déjà appris les collocations exprimant plusieurs sémantiques. Maintenant, c'est à vous de les classer en fonction de leurs valeurs sémantiques. Repérez les collocations du genre Nom + Adjectif dans les phrases et classez-les. Pour les collocations de jugement, indiquez si elle est positive ou négative.

Dans la phase de mémorisation, nous demandons à l'apprenant de classer des collocations en fonction des valeurs sémantiques et rhétoriques afin de l'inciter à apprendre à créer son propre tableau de classement des collocations évaluatives rencontrées et à l'emmagasiner dans ses propres données au fur et à mesure, parce que le classement du lexique « permet à l'apprenant de se concentrer sur l'objet, voire de trouver sa place d'acteur « linguiste » » (Boche et Frier, 2015: 99)

Nous proposons des collocations dont les bases (noms scientifiques) sont différentes par rapport aux collocations dans l'activité 01, mais avec les collocatifs (adjectifs) déjà rencontrés par l'apprenant dans les activités précédentes. Cette considération possède deux avantages :

a. Elle permet à l'apprenant de connaître plus de combinaisons entre les noms

scientifiques et les adjectifs évaluatifs.

- b. Le classement sémantique et rhétoriques ne dépend que des valeurs sémantiques des adjectifs. Ce processus renforce donc la mémorisation et la consolidation des valeurs sémantiques des adjectifs évaluatifs chez apprenant.

L'activité 03 se situe dans le cadre de la dernière phase d'enseignement : le réemploi des collocations évaluatives dans l'écrit scientifique de l'apprenant. Nous pouvons demander à l'apprenant d'écrire une introduction sur un thème auquel il s'intéresse dans son domaine en réemployant des collocations évaluatives qu'il a déjà appris dans les activités 01 et 02.

L'exemple vise un public d'apprenant dont les membres ne proviennent pas de la même discipline, la sélection des concordances ne se limitera donc pas à une discipline précise. Si le public est plus homogène, nous tendrons davantage à utiliser l'introduction ou la conclusion intégrale (la deuxième composition de corpus) dans les textes de la discipline des apprenants afin de mieux observer comment dans leur domaine commun l'auteur français natif exprime son évaluation envers les noms scientifiques dans les deux parties textuelles. Le déroulement des exercices s'articulera également autour des quatre phases, à savoir le repérage, la sensibilisation, la mémorisation et le réemploi. Un travail de comparaison entre l'introduction et la conclusion pourrait être mis en place pour faire différencier les deux parties textuelles en matière d'évaluation. Comme nous n'avons pas d'accès aux introductions et conclusions intégrales des textes dans le corpus Scientext, nous ne montrons pas ici un exemple concret.

Bien qu'utiliser l'intégralité de l'introduction ou de la conclusion favorise la découverte et la compréhension des collocations évaluatives et leurs propriétés sémantiques, syntaxiques et rhétoriques dans les deux parties textuelles, cette composition de corpus se heurte à une difficulté cruciale : l'introduction ou la conclusion intégrales sont souvent d'une dimension beaucoup plus grande par rapport aux concordances tout en contenant peu de collocations évaluatives. En conséquence, pour qu'un apprenant maîtrise l'ensemble des valeurs sémantiques et des fonctions rhétoriques, il serait obligé de lire un

nombre assez grand des introductions ou de conclusions pour obtenir des données suffisantes. Ce travail laborieux pourrait décourager l'apprenant.

Bilan

Cette partie a discuté l'exploitation du corpus Scientext par l'enseignant dans l'enseignement des collocations évaluatives en question. Nous proposons en premier temps de privilégier l'approche indirecte qui ne donne pas l'accès direct au corpus Scientext pour ne pas rebuter l'apprenant par l'énormité des concordances issues de la requête. L'approche indirecte peut aussi avoir des variantes lors de l'exploitation du corpus. Dans ce sens, nous avons analysé deux façons de constitutions des petits corpus pour l'apprenant : les listes des concordances sélectionnées dans n'importe quelle discipline pour le public hétérogène, les introductions ou conclusions intégrales pour le public homogène. Les quatre phases pour élaborer une séquence didactique fonctionnent en effet dans les deux cas de constitution des petits corpus, juste que la deuxième composition de corpus pose le problème de la quantité de travail à cause de la taille relativement plus grande que des concordances.

Tous nos réflexions dans cette partie se limitent encore à l'approche indirecte, mais comme nous avons signalé au début de cette partie que l'approche indirecte peut servir de la phase de transition pour aboutir à l'approche directe qui place l'apprenant en face des outils informatiques au sein du corpus Scientext afin d'établir son apprentissage autonome chez apprenant, nous invitons à regarder le dernier point : vers l'autonomie de l'apprenant.

7.3. Vers l'autonomie de l'apprenant non natif

Est-ce que l'apprenant non natif peut enfin apprendre à manipuler des outils numériques conçus dans le corpus? Et s'orienter vers l'autonomie d'apprentissage? Des recherches sur l'emploi des corpus par l'apprenant ont déjà apporté des réponses. Yoon et Hirvela (2004 : 278) constatent dans leur étude que les apprenants d'anglais apprécient la consultation de corpus lors de l'apprentissage de l'écrit, en particulier sur les usages communs, les collocations. La consultation de corpus leur permet de construire la confiance dans leur écrit. Boulton (2010) a mené également une étude sur la consultation

de corpus pour l'autocorrection de l'écrit sans l'intervention de l'enseignant auprès des étudiants d'anglais d'un niveau modeste sans formation préalable. Cette étude demandait à 43 étudiants de corriger leurs traductions à travers la consultation de corpus. Ses résultats montrent que les étudiants effectuent plus de corrections positives que négatives. Les résultats de l'enquête à la fin de l'expérimentation ont prouvé en même temps que la plupart des étudiants appréciaient la consultation de corpus pour l'autocorrection.

Étant donné les avantages de la consultation directe de corpus et sa fiabilité prouvée par les études que nous venons d'évoquer, nous conseillons que la consultation directe suive la consultation indirecte qui joue déjà le rôle de tremplin pour guider l'apprenant vers l'autonomie d'apprentissage.

Nous proposons une progression pour la découverte et la maîtrise du corpus Scientext en prenant celle établie par Cavalla (2014: 164) :

1. Présenter le corpus Scientext et son outil informatique ScienQuest
2. Guider strictement la recherche sur le corpus Scientext
 - soit consulter par la requête sémantique en choisissant *Adjectifs d'évaluation*, soit consulter par la requête libre en extrayant les concordances d'un nom précis + adjectif.
 - relever les collocations évaluatives N_{scientifique} + Adj_{évaluatif}
 - classer les collocations repérées en fonction de leurs valeurs sémantiques (importance, degré...)
3. Recherches libres sur le corpus Scientext
4. Demander aux apprenants de consulter le corpus pour la correction de leur écrit scientifique.

La première étape tente d'introduire brièvement le corpus Scientext et l'outil ScienQuest de manière à ce que l'apprenant obtienne au moins les connaissances principales du corpus Scientext.

La deuxième étape s'avère sans doute la plus cruciale. Pour que l'apprenant puisse trouver les collocations évaluatives en question, nous le guidons à la requête sémantique du corpus, ce qui va lui donner directement des collocations évaluatives $N_{\text{scientifique}} + \text{Adj}_{\text{évaluatif}}$ à classer ultérieurement. Pour un nom scientifique plus concret, nous pouvons guider l'apprenant vers la requête libre pour extraire toutes les combinaisons entre ce nom et des adjectifs. Cette étape demande en effet à l'apprenant de s'exercer plusieurs fois pour garantir sa bonne maîtrise de la consultation de corpus.

Les troisième et quatrième étapes demandent déjà l'autonomie de l'apprenant. Ce dernier peut essayer de trouver les phénomènes linguistiques dont qu'il a besoin. Dans cette phase, l'enseignant peut se mettre en retrait et laisser l'apprenant se débrouiller seul devant le corpus Scientext.

Bilan

Notre dernier chapitre a essayé de traiter la question de la didactique des collocations évaluatives qui ont fait l'objet de recherches dans le chapitre précédent. Bien que le Corpus Scientext ne soit pas conçu pour une finalité didactique, nous proposons de l'exploiter comme outil d'enseignement et d'apprentissage. La consultation indirecte de corpus doit être guidée par l'enseignant, et ce dernier, est censé créer des séquences didactiques à partir de la constitution de petits corpus tirés minutieusement des concordances énormes du corpus Scientext. Une fois que l'apprenant se serait familiarisé avec le travail sur de petits corpus, nous suggérons d'intégrer la consultation directe de corpus dans la classe pour que l'apprenant développe l'autonomie en apprentissage de la langue française, surtout sur les collocations $N_{\text{scientifique}} + \text{Adj}_{\text{évaluatif}}$.

Conclusion et perspective

Notre étude sur les combinaisons des noms scientifique transdisciplinaires avec des adjectifs évaluatifs qui expriment le positionnement évaluatif de l'auteur dans l'introduction et la conclusion des textes scientifiques dans huit disciplines (Biologie, Médecine, Linguistique, Psychologie, Sciences de l'éducation, TAL -Traitement Automatique des Langues, Mécanique, Électronique) a apporté certains résultats très intéressants.

Les sciences expérimentales et appliquées mobilisent plus de collocations évaluatives dans l'introduction et la conclusion de l'écrit scientifique que les sciences humaines.

Il semble que les auteurs natifs ont tendance à utiliser plus de collocations non axiologiques que des collocations axiologiques dans l'introduction et la conclusion dans la plupart des disciplines (7 sur 8 disciplines). Nous avons relevé aussi que le cas de la seule discipline dans laquelle la proportion des collocations axiologiques est supérieure à celle des collocations non axiologiques a pour explication l'utilisation abusive d'un seul auteur. De ce fait, nous pouvons affirmer que l'écrit scientifique n'utilise pas fréquemment des jugements positifs ou négatifs, tout comme ce que Tutin et Cavalla ont constaté (2008).

En ce qui concerne l'utilisation des collocations axiologiques dans le corpus scientifique en français, nous avons découvert que la conclusion est apte à utiliser plus de collocations axiologiques que l'introduction. Ceci fait que dans la conclusion, l'auteur cherche à valoriser l'étude faite et les résultats obtenus.

Quant aux collocations non axiologiques, notre corpus nous a permis de repérer seulement quatre dimensions sémantiques et rhétoriques : Importance, Nouveauté, Degré et Complexité, faut d'un corpus assez large. Parmi les quatre dimensions, Importance et Nouveauté l'emportent sur les autres dimensions. Ce qui dévoile la stratégie rhétorique de l'auteur natif dans ce genre d'écrit :

La dimension Importance s'avère particulièrement cruciale pour la stratégie

rhétorique dans l'introduction et la conclusion, parce qu'elle permet à l'auteur de montrer ses compétences à mieux gérer des informations en sélectionnant des éléments fondamentaux pour son lecteur. La dimension Nouveauté souligne la nouveauté, l'originalité des travaux de l'auteur afin de les promouvoir, lorsque la dimension Complexité met l'accent sur la difficulté des recherches, ce qui concourt à augmenter leur intérêt et à expliquer pourquoi une solution simple est difficile à proposer. La dimension Degré semble plus restreinte dans l'écrit scientifique. Tous ces constats que nous avons fait sur les quatre dimensions semblent parfaitement correspondre aux résultats de Tutin et Cavalla (2008). Néanmoins, pour la dimension Temporalité, qui s'est faite remarquer souvent en cooccurrence avec la dimension Nouveauté dans l'étude des deux auteurs, est totalement absente dans notre corpus. En conséquence, nous ne pouvons pas nous autoriser de nous prononcer sur la stratégie de la dimension Temporalité, ni sur sa relation avec la dimension Nouveauté.

En plus des valeurs sémantiques des collocations, nous avons observé aussi leurs propriétés syntaxiques. Les résultats sont évidents : la structure Nom scientifique + Adjectif évaluatif où l'adjectif évaluatif fait fonction d'épithète est largement plus présente que la structure Nom scientifique + verbe support + Adjectif évaluatif dont l'adjectif travaille en tant qu'attribut. De plus, pour les deux structures, les insertions sont possibles, souvent de la nature adverbiale ou adjectivale.

À l'intérieur des collocations évaluatives, la combinaison entre des noms scientifiques et des adjectifs évaluatifs ne se fait pas de façon aléatoire, au contraire, il existe des préférences combinatoires. Notre étude montre que l'adjectif d'évaluation se combine souvent avec quatre catégories des noms scientifiques : les noms d'objets construits, d'observables, de processus et de caractérisation.

Notre étude montre aussi que ce genre de collocations sont susceptibles de transcender les disciplines et peuvent donc servir à l'enseignement d'étudiants non natif quelle que soit leur discipline.

Il faut signaler enfin que les résultats que nous avons obtenus ne sont pas

incontestables, compte tenu de la taille inégale des textes scientifiques dans les huit disciplines, et aussi de la taille restreinte de notre corpus. Mais nous pouvons dire que les résultats observés peuvent servir de base à de nouvelles hypothèses pour une étude sur des corpus plus larges, plus représentatifs et de taille égale.

Les résultats obtenus peuvent ouvrir des pistes pour l'enseignement des collocations en question. Nous trouvons que le corpus Scientext n'est pas seulement un outil pour des linguistes, mais aussi un outil d'enseignement et d'apprentissage avec beaucoup de potentialités. La consultation du corpus Scientext se répartit en deux approches : indirecte et directe. Dans notre travail, nous avons proposé de faire consulter d'abord les concordances extraites du corpus par l'enseignant afin de laisser l'apprenant se familiariser avec les concordances et puis le diriger vers l'autonomie d'apprentissage à l'aide de la consultation directe de corpus Scientext.

Au cours de cette étude sur les collocations évaluatives, j'ai vraiment beaucoup appris et compris que l'écrit scientifique n'est pas un écrit purement objectif, car l'auteur doit toujours savoir bien se positionner et évaluer pour mettre en avant son travail et le promouvoir. Ce qui m'impressionne le plus, c'est la dimension interpersonnelle dans l'écrit scientifique. Des ressources linguistiques comme les collocations évaluatives dans notre étude manifestent non seulement la spécificité en matière des valeurs sémantiques, des fonctions rhétoriques, mais reflètent particulièrement une dimension interpersonnelle où l'auteur essaie de faire engager le lecteur dans son écrit. Aussi, si je peux poursuivre cette thèse, je voudrais étudier comment les phénomènes phraséologiques aident à construire la dimension interpersonnelle dans l'écrit scientifique en français.

Références bibliographiques :

- Adam, J.-M. . (1999). *Linguistique textuelle*. Paris: Nathan.
- Allouche, V. , & Maurer, B. (2010). *L'écrit en FLE: travail du style et maîtrise de la langue*. Montpellier: Presses universitaires de la Méditerranée.
- Aouadi, L. (2015). *L'expression de la subjectivité dans le discours scientifique: Cas des mémoires de magister français* (masters). Université Mohamed Khider - Biskra. Consulté à l'adresse <http://thesis.univ-biskra.dz/2218/>
- Arbach, N., & Ali, S. (2013). Aspects théoriques et méthodologiques de la représentativité des corpus. *Corela. Cognition, représentation, langage*, (HS-13). <https://doi.org/10.4000/corela.3029>
- Archer, V. (2006). Acquisition semi-automatique de collocations à partir de corpus monolingues et multilingues comparables. In *RECITAL 2006* (p. 651-660). Leuven, Belgium. Consulté à l'adresse <https://hal.archives-ouvertes.fr/hal-00983425>
- Bally, C. (1951). *Traité de Stylistique Française*, Paris : Librairie Klincksieck.
- Baneyx, A., Malaisé, V., Charlet, J., Zweigenbaum, P., & Bachimont, B. (2005). Synergie entre analyse distributionnelle et patrons lexico-syntaxiques pour la construction d'ontologies différentielles. *rue, 1*, U729.
- Banks, D. (2001). Analyses des discours spécialisés. *Revue française de linguistique appliquée*, VI(2), 7-16.
- Banks, D. (2011). La Linguistique Systémique Fonctionnelle : une approche sémantique et sociale. *Amadis*, (9), 265-278.
- Barton, E. L. (1993). Evidentials, Argumentation, and Epistemological Stance. *College English*, 55(7), 745-769. <https://doi.org/10.2307/378428>
- Benveniste, E. (1966). *Problèmes de linguistique générale. [I]*. Paris: Gallimard.
- Berthelot, J.-M. (2003). *Figures du texte scientifique*. Paris: Presses universitaires de France.
- Bronckart, J. P. (1996). Genres de textes, types de discours et opérations psycholinguistiques. *Voies livres*, (78), 1-20.
- Binon, J., & Verlinde, S. (1999). La contribution de la lexicographie pédagogique à l'apprentissage et à l'enseignement d'une langue étrangère ou seconde. *Études de Linguistique Appliquée*, 0, 453.

- Binon, J., & Verlinde, S. (2004). Les collocations: clef de voûte de l'enseignement et de l'apprentissage du vocabulaire d'une langue étrangère ou seconde. *Romanesque*, 29(2), 16-24.
- Boch, F. et Fanny Rinck. (2010). « Enonciation et rhétorique dans l'écrit scientifique [Special Issue] ». *Lidil: Revue de Linguistique et de Didactique des Langues* 41: 5–169. Print.
- Boch, F. (2013). Former les doctorants à l'écriture de la thèse en exploitant les études descriptives de l'écrit scientifique. *Linguagem em (Dis)curso*, 13(3), 543-568.
- Boch, F. (2017). *Écrire dans l'enseignement supérieur: Des apports de la recherche aux outils pédagogiques*. UGA Éditions.
- Bondi, M. (2006). *Academic Discourse Across Disciplines*. Peter Lang.
- Boulton, A. (2007). Esprit de corpus: Promouvoir l'exploitation de corpus en apprentissage des langues. *Texte et Corpus.*, 3, 37-46.
- Boulton, A., & Tyne, H. (2014). *Des Documents Authentiques aux Corpus: Démarches pour l'Apprentissage des Langues* (p. 309). Didier.
- Boulton, A. (2009, July). Corpora for all? Learning styles and data-driven learning. In *5th Corpus Linguistics Conference*.
- Boulton, A. (2010). Consultation de corpus et styles d'apprentissage. *Recherche et pratiques pédagogiques en langues de spécialité. Cahiers de l'Aplut*, 29(1), 98-115.
- Cavalla, C. (2007). Réflexion pour l'aide à l'écrit universitaire auprès des étudiants étrangers entrant en Master et Doctorat. In J. G. et J.-M. Mangiante (Éd.), *L'accueil des étudiants étrangers dans les universités francophones* (p. 37-48). Artois Presses Université. Consulté à l'adresse <https://hal.archives-ouvertes.fr/hal-00380482>
- Cavalla, C. (2008a). Les collocations dans les écrits universitaires : un lexique spécifique pour les apprenants étrangers. In O. B. et I. Schaffner (Éd.), *Apprendre une langue de spécialité : enjeux culturels et linguistiques* (p. 93-104). Editions Ecole Polytechnique. Consulté à l'adresse <https://hal.archives-ouvertes.fr/hal-00397684>
- Cavalla, C. (2008c). Propositions didactiques pour l'enseignement d'éléments phraséologiques en FLE. In D. V. R. et M.-E. Damar (Éd.), *Enseigner les structures langagières en FLE* (p. <http://tmaes.no-org/ME/>). Gramm-FLE. Consulté à l'adresse <https://hal.archives-ouvertes.fr/hal-00397807>
- Cavalla, C., & Tutin, A. (2008). Etude des collocations évaluatives dans les écrits scientifiques. In *Europhras2008*. Helsinki, Finland. Consulté à l'adresse <https://hal.archives-ouvertes.fr/hal-01099187>

- Cavalla, C. (2009). La phraséologie en classe de FLE. Les Langues Modernes. <http://www.aplv-languesmodernes.org/spip.php?article2292>.
- Cavalla, C. (2010). Les écrits universitaires des étudiants étrangers : quelles normes présenter ? In I. S. O. Bertrand (Éd.), *Quel français enseigner ? La question de la norme dans l'enseignement / apprentissage* (p. 203-214). Polytechnique. Consulté à l'adresse <https://hal.archives-ouvertes.fr/hal-01098816>
- Cavalla, C., & Grossmann, F. (2005). Caractéristiques sémantiques de quelques « Noms scientifiques » dans l'article de recherche en français. In T. Kinn (Éd.), *Akademisk prosa* (p. 47-59). Skrifter fra KIAP Romansk institutt, Universitetet i Bergen. Consulté à l'adresse <https://hal.archives-ouvertes.fr/hal-00394208>
- Cavalla, C., & Loiseau, M. (2013). Scientext: a corpus to teach. In A. Tutin & F. Grossmann (Éd.), *L'écrit scientifique : du lexique au discours. Autour de Scientext* (p. 163-182). Presses Universitaires de Rennes. Consulté à l'adresse <https://hal.archives-ouvertes.fr/hal-01154754>
- Cavalla, C. (2014). Collocations transdisciplinaires: réflexion pour l'enseignement. 151-169.
- Chafe, Wallace L., and Johanna Nichols (editors). 1986. *Evidentiality: The linguistic coding of epistemology*. Advances in Discourse Processes 20. Norwood, NJ: Ablex. 9683101704.
- Chambers, A. (2005). Integrating corpus consultation in language studies.
- Chambers, A. (2010). L'apprentissage de l'écriture en langue seconde à l'aide d'un corpus spécialisé. *Revue française de linguistique appliquée*, 15(2), 9-20.
- Charaudeau, P., & Maingueneau, D. (2002). *Dictionnaire d'analyse du discours*. Paris: Éddu Seuil.
- Chavez, I. (2008). *La démarcation dans les écrits scientifiques - Les collocations transdisciplinaires comme aide à l'écrit universitaire auprès des étudiants étrangers*. Mémoire de Master Français Langue Etrangère Recherche, Université Stendhal-Grenoble3, Grenoble.
- Cislaru, G.), Claudel, C.), & Vlad, M. (2009). *L'écrit universitaire en pratique: applications et exercices autocorrectifs, rédaction d'un mémoire, présentation PowerPoint*. Bruxelles: De Boeck.
- Costa, R. (2006). Texte, terme et contexte. *Septième Journées Scientifiques AUF-LTT «Mots, termes et contextes», Bruxelles*, 8.
- Delaporte. A., *L'importance d'unités complexes méconnues: les collocations et*

- collocations*, mémoire de master 1, AMIOT, Dany. (directeur), Sciences humaines et sociales, Université de Lille 3, Lille, 2015-2016.
- Dostie, G. (2013). Les associations de marqueurs discursifs de la cooccurrence libre à la collocation. *Linguistik Online*, 62(5), 15.
- Drouin, P. (2007). Identification automatique du lexique scientifique transdisciplinaire. *Revue française de linguistique appliquée*, XII(2), 45-64.
- ÉVALUER : Etymologie de ÉVALUER. (s. d.). Consulté 18 avril 2017, à l'adresse <http://www.cnrtl.fr/etymologie/%C3%A9valuer>
- Falkert, A. (2013). À propos de l'existence d'une compétence phraséologique chez l'apprenant de FLE. Une étude expérimentale. *ResearchGate*, 153-172.
- Fløttum, Kjersti, Trine Dahl, et Torodd Kinn. (2006). *Academic Voices: Across Languages and Disciplines*. John Benjamins Publishing, 2006. Print.
- Frankenberg-Garcia, A., Flowerdew, L., & Aston, G. (2011). *New Trends in Corpora and Language Learning*. A&C Black.
- Garcia-Debanc, C., Masseron, C., Ronveaux, C., Association internationale pour la recherche en didactique du français, & Centre d'études et de documentation pour l'enseignement du français . Namur, B. (2013). *Enseigner le lexique*. Namur: Presses universitaires de Namur.
- Gardin, P. (2009). Application de la théorie de l'Appraisal à l'analyse d'opinion , MajecSTIC.
- Geens, D. (1984). Semantic Analysis Automated for Large Computer Corpora and Their Exploitation in Recent Developments in the Use of Computer Corpora in English Language Research. *Costerus*, 45, 211-229.
- Gledhill, C., & Frath, P. (s. d.). Collocation, phrasème, dénomination : vers une théorie de la créativité phraséologique. *La linguistique*, 43(1), 63-88.
- Gledhill, C. J. (2000). *Collocations in Science Writing*. Gunter Narr Verlag.
- Gonzalez, Rey, I. (2015). *La phraséologie du français* (2e édition revue et corrigée). Toulouse: Presses Universitaires du Mirail.
- Grossmann, F., & Tutin, A. (2003). *Les collocations: analyse et traitement*. Amsterdam: De Werelt.
- Grossmann, F., Tutin, A., & Silva, P. P. G. D. (2009). Filiation et transfert d'objets scientifiques dans les écrits de recherche. *Pratiques. Linguistique, littérature, didactique*, (143-144), 187-202. <https://doi.org/10.4000/pratiques.1447>

- Grossmann, F., & Tutin, A. (2010). Evidential Markers in French Scientific Writing : the Case of the French Verb voir. In G. Diewald & E. Smirnova (Éd.), *Linguistic Realization of Evidentiality in European Languages (Empirical Approaches to Language Typology)* (p. 279-307). Berlin, New York: De Gruyter Mouton.
- Guespin, L. (1971). *Le discours politique*. Paris: Didier Larousse.
- Halliday, M.A.K. (1985). *An introduction to Functional Grammar*. London:Arnold.
- Hatier, S., Tutin, A., Jacques, M.-P., Jacquey, E., & Kister, L. (2014). *Catégorisation sémantique des noms simples du lexique scientifique transdisciplinaire*. Présentation présenté à Congrès ACFAS, colloque 330 : Étude de lexiques à vocation particulière : approches théoriques, méthodologiques, pédagogiques et multidisciplinaires, 14 mai, Montréal. Consulté à l'adresse <http://www.acfas.ca/evenements/congres/programme/82/300/330/c>
- Hausmann F. J. (1989). "Le dictionnaire de collocations". In Hausmann F.J., Reichmann O., Wiegand H.E., Zgusta L. (eds), *Wörterbücher : ein internationales Handbuch zur Lexicographie*. Dictionaries. Dictionnaires. Berlin/New-York, De Gruyter, 1010-1019.
- Hausman, F. J., & Blumenthal, P. (2006). Présentation : collocations, corpus, dictionnaires. *Langue française*, 150(2), 3-13. <https://doi.org/10.3406/lfr.2006.6850>
- Hunston, S. (1994). Evaluation and Organization in a Sample of Written Academic Discourse. In M. Coulthard (Ed.), *Advances in Written Text Analysis* (pp. 191-218). London: Routledge. - References - Scientific Research Publish. (s. d.). Consulté 30 mars 2017, à l'adresse [http://www.scirp.org/\(S\(oyulxb452alnt1aej1nfow45\)\)/reference/ReferencesPapers.aspx?ReferenceID=1729664](http://www.scirp.org/(S(oyulxb452alnt1aej1nfow45))/reference/ReferencesPapers.aspx?ReferenceID=1729664)
- Hunston et Sinclair (2000). A local grammar of evaluation in *Evaluation in Text : Authorial Stance and the Construction of Discourse: Authorial Stance and the Construction of Discourse*. Oxford University Press, UK.
- Hunston, S., & Thompson, G. (2000). *Evaluation in Text : Authorial Stance and the Construction of Discourse: Authorial Stance and the Construction of Discourse*. Oxford University Press, UK.
- Hyland, K. (2002). Authority and invisibility: Authorial identity in academic writing. *Journal of pragmatics*, 34(8), 1091–1112.
- Hyland, K. (2003). Options of Identity in Academic Writing. *ELT Journal*, 56(4), 351–58.
- Hyland, Ken, and Marina Bondi. 2006, ed. *Academic discourse across disciplines*. Bern: Peter Lang.

- Hyland, K. (2008). As can be seen: Lexical bundles and disciplinary variation. *English for specific purposes*, 27(1), 4–21.
- Hyland, K., & Guinda, C. (2012). Stance and voice in written academic genres. *Journal of English for Academic Purposes*, 16. <https://doi.org/10.1057/9781137030825>
- Jaubert, A. (1990). *La lecture pragmatique*. Paris: Hachette supérieur.
- Jaubert, A. (2002). Corpus et champs disciplinaires. Le rôle du point de vue. *Corpus*, (1). Consulté à l'adresse <http://corpus.revues.org/13>
- Kerbrat-Orrechioni, Catherine. 1980. *L'énonciation: De la subjectivité dans le langage*. Paris: Armand Colin.
- La notion de Discours - Analyse-du-discours.com. (s. d.). Consulté 10 mars 2017, à l'adresse <http://www.analyse-du-discours.com/la-notion-de-discours>
- Landure, C., & Boulton, A. (2010). Corpus et autocorrection pour l'apprentissage des langues. *ASp. la revue du GERAS*, (57), 11-30.
- Le Robert. (2016). *Le Robert illustré & son dictionnaire en ligne: [2017]* (Nouvelle édition.). Paris: Le Robert.
- Leistyna, P., & Meyer, C. F. (2003). *Corpus Analysis: Language Structure and Language Use*. Rodopi.
- Lewis M. (2001). *Teaching collocation*, Londres, Language Teaching Publications.
- Ligas, P. (2008). Pour une typologie des définitions d'adjectifs et d'adverbes dans le «Dictionnaire alphabétique et analogique du français des activités physiques et sportives» (à paraître, 2009). In *Congrès Mondial de Linguistique Française* (p. 112). EDP Sciences. <https://doi.org/10.1051/cmlf08093>
- Lilly, R. S. (1968). The qualification of evaluative adjectives by frequency adverbs. *Journal of Verbal Learning and Verbal Behavior*, 7(2), 333-336. [https://doi.org/10.1016/S0022-5371\(68\)80011-8](https://doi.org/10.1016/S0022-5371(68)80011-8)
- Loffler-Laurian, A. M. (1983). Typologie des discours scientifiques: deux approches. *Études de linguistique appliquée*, 51, 8.
- Maingueneau, D. (1987). *Nouvelles tendances en analyse du discours*. Paris: Hachette.
- Maingueneau, D. (1976). *Initiation aux méthodes de l'analyse du discours: problèmes et perspectives*. Paris: Hachette.
- Maingueneau, D. (1991). *L'analyse du discours: introduction aux lectures de l'archive*. Paris: Hachette supérieur.

- Maingueneau, D. (1992). Le tour ethnolinguistique de l'analyse du discours. *Langages*, 26(105), 114-125. <https://doi.org/10.3406/lgge.1992.1628>
- Malrieu & Rastier : Genres et variations morphosyntaxiques. (s. d.). Consulté 9 mars 2017, à l'adresse http://www.revue-texto.net/Inedits/Malrieu_Rastier/Malrieu-Rastier_Genres.html
- Mangiante, J.-M., & Parpette, C.). (2011). Le français sur objectif universitaire. Presses universitaires de Grenoble.
- Martin, A. V. (1976). Teaching Academic Vocabulary to Foreign Graduate Students. *TESOL Quarterly*, 91-98.
- Martin, J. (1997). Du bon usage des corpus dans la recherche sur le discours spécifique. *ASp. la revue du GERAS*, (15-18), 75-83. <https://doi.org/10.4000/asp.2984>
- Martin, J., & White, P. (2005). *The Language of Evaluation: Appraisal in English*. Springer.
- Master, P. « J. Swales, Genre Analysis: English in Academic and Research Settings ». *Journal of Pragmatics* 17.3 (1992): 286. Print.
- Matthiessen, C. M., & Halliday, M. A. K. (2009). Systemic functional grammar: a first step into the theory.
- Mel'čuk, I., Iordanskaja, L., Arbatchewsky-Jumarie, N., & Lessard, A. (1983). Trois principes de description sémantique d'une unité lexicale dans un dictionnaire explicatif et combinatoire. *Canadian Journal of Linguistics/Revue canadienne de linguistique*, 28(02), 105-122.
- Melčuk, I. (1993). La phraséologie et son rôle dans l'enseignement-apprentissage d'une langue étrangère. *Études de Linguistique Appliquée*, 92. Consulté à l'adresse <http://search.proquest.com.ezproxy.univ-paris3.fr/docview/1307652747/citation/DF2D361F75D04CB5PQ/1>
- Mel'čuk, I. (2003). Collocations: définition, rôle et utilité. *Travaux et recherches en linguistique appliquée. Série E, Lexicologie et lexicographie*, (1), 23-31.
- Mel'čuk, I. (2013). Tout ce que nous voulions savoir sur les phrasèmes, mais. *Cahiers de lexicologie*, 102, 129-149.
- Mroue, M. (2014, octobre 23). *Ecrit de recherche universitaire : éléments pour une sensibilisation au positionnement scientifique à travers la phraséologie transdisciplinaire* (phdthesis). Université Grenoble Alpes. Consulté à l'adresse <https://tel.archives-ouvertes.fr/tel-01321057/document>

- Ochs, E. (1989). *The Pragmatics of Affect*. Mouton de Gruyter.
- Pecman, M. (2004). *Phraséologie contrastive anglais-français: analyse et traitement en vue de l'aide à la rédaction scientifique*. Thèse de doctorat, Université de Nice-Sophia Antipolis.
- Pecman, M. (2007). Approche onomasiologique de la langue scientifique générale, Abstract. *Revue française de linguistique appliquée*, XII(2), 79-96.
- Pelpel, P. (2003). *Accueillir, accompagner, former des enseignants. Guide de réflexion et d'action*. Lyon : Chronique sociale.
- Perez - Llantada Auria, M. C. (2008). Stance and academic promotionalism: a cross-disciplinary comparison in the soft sciences.(Report). *Atlantis, revista de la Asociacion Espanola de Estudios Anglo-Norteamericanos*, 30(1), 129.
- Rastier, F. (1998). Le problème épistémologique du contexte et le statut de l'interprétation dans les sciences du langage. *Langages*, 32(129), 97-111. <https://doi.org/10.3406/lgge.1998.2149>
- Reuter, Y. (2004). Analyser les problèmes de l'écriture de recherche en formation. *Pratiques*, 121(122), 9-27.
- Rinck, F., Boch, F., & Grossmann, F. (2006). Quelques lieux de variation du positionnement énonciatif dans l'article de recherche. *Filologia e Linguística Portuguesa*, 0(8), 451-464.
- Rinck, F. (2011). Former à (et par) l'écrit de recherche. Quels enjeux, quelles exigences ? *Le français aujourd'hui*, (174), 79-89.
- Saussure, F. de. (1979). *Cours de linguistique générale*. Paris: Payot.
- 宋成方. (2010) 评价理论的社会主体间性模式及其特征 in *评价理论研究*, 外语教学与研究出版社. 北京 (traduction personnelle) Song Chengfang (2010) Le modèle social intersubjectif de la théorie de l'évaluation et ses caractéristiques in *Studies of appraisal theory*, ed: Forgeign Language Teaching and Research Presse, Beijing.
- Shaw, P. (2003). Evaluation and promotion across languages. *Journal of English for Academic Purposes*, 2(4), 343-357. [https://doi.org/10.1016/S1475-1585\(03\)00050-X](https://doi.org/10.1016/S1475-1585(03)00050-X)
- Sinclair J. 1996. *Preliminary recommendations on Corpus Typology*, EAG--TCWG--CTYP/P. Version May 1996. Pisa : EAGLES. <http://www.ilc.pi.cnr.it/EAGLES96/corpuSTyp.htm>
- Silva, R., Costa, R., & Ferreira, F. (2004). Entre langue générale et langue de spécialité une question de collocations. *Ela. Études de linguistique appliquée*, no 135(3), 347-359.

- Stotesbury, H. (2003). Evaluation in research article abstracts in the narrative and hard sciences. *Journal of English for Academic Purposes*, 2(4), 327-341. [https://doi.org/10.1016/S1475-1585\(03\)00049-3](https://doi.org/10.1016/S1475-1585(03)00049-3)
- Sumpf, J., & Dubois, J. (1969). Problèmes de l'analyse du discours. *Langages*, 4(13), 3-7. <https://doi.org/10.3406/lgge.1969.2506>
- Swales, J. (1990). *Genre Analysis: English in Academic and Research Settings*. Cambridge University Press.
- Teubert, W. (2009). La linguistique de corpus : une alternative [version abrégée]. *Semen. Revue de sémio-linguistique des textes et discours*, (27), 185-211.
- Teubert, W. (2009). Corpus Linguistics: An Alternative. *Semen. Revue de Sémio-Linguistique Des Textes et Discours*, (27). Consulté à l'adresse <http://semen.revues.org/8912?lang=en>
- Textes réunis par David Banks. (2014 apr. J.-C.). *LES MARQUEURS LINGUISTIQUES DE LA PRÉSENCE DE L'AUTEUR*. Editions Harmattan.
- Tran, T. T. H. (2010). Aides logicielles à la rédaction universitaire. L'exemple des collocations, 167.
- Tutin, A. (2010). Showing phraseology in context: an onomasiological access to lexico-grammatical patterns in corpora of French scientific writings. In S. Granger & M. Paquot (Éd.), *eLexicography in the 21st century: new applications, new challenges* (p. 303-312). Louvain: Presses de l'Université de Louvain.
- Tutin, A., & Grossmann, F. (2002). Collocations régulières et irrégulières : esquisse de typologie du phénomène collocatif, Abstract. *Revue française de linguistique appliquée*, VII(1), 7-25.
- Tutin, A. (2007). Autour du lexique et de la phraséologie des écrits scientifiques. *Revue française de linguistique appliquée, Lexique et écrits scientifiques*(XII(2)), 5-14.
- Tutin, A. (2007). Modélisation linguistique et annotation des collocations : une application au lexique transdisciplinaire des écrits scientifiques. In S. Koeva, D. Maurel, & M. Silberstein (Éd.), *Formaliser les langues avec l'ordinateur : de INTEX à Nooj* (Vol. 8, p. 189-215). Besançon: Presses Universitaires de Franche Comté.
- Tutin, Agnès. (2007) « Traitement sémantique par analyse distributionnelle des noms transdisciplinaires des écrits scientifiques ». *Actes de la 14ème conférence annuelle sur le Traitement Automatique des Langues Naturelles*. Toulouse

- Tutin, A. (2007a). Collocations du lexique transdisciplinaire des écrits scientifiques: annotation et extraction des propriétés linguistiques dans la perspective d'une application didactique. *Cahiers de l'Institut de linguistique de Louvain*, 31(2-4), 247–262.
- Tutin, A. (Éd.). (2007b). Lexique des écrits scientifiques. *Revue française de linguistique appliquée*, XII(2). Consulté à l'adresse <http://www.cairn.info/revue-francaise-de-linguistique-appliquee-2007-2.htm>
- Tutin, A. (2008). For an extended definition of lexical collocations (p. 0). Présenté à Euralex. Consulté à l'adresse <https://halshs.archives-ouvertes.fr/halshs-00371418/document>
- Tutin, Agnès, Francis Grossmann, Achille Falaise, and Olivier Kraif. 2009. “Autour Du Projet Scientext : Étude Des Marques Linguistiques Du Positionnement de L'auteur Dans Les Écrits Scientifiques.” In Lorient, France. http://www.licorn-ubs.com/jlc6/ACTES/Souque_JLC09.pdf.
- Tutin, A. (2010). Évaluative adjectives in academic writing in the humanities and social sciences. In R. Lores-Duenas, P. Mur-Duenas, & E. Lafuente-Milan (Éd.), *Constructing Interpersonality: Multiple Perspectives on Written Academic Genres*. Cambridge: Cambridge Scholars Publishing.
- Tutin, A. (2013). Les collocations lexicales : une relation essentiellement binaire définie par la relation prédicat-argument, Lexical collocations ; a binary and predicate-argument based relation. *Langages*, (189), 47-63.
- Tutin, A., & Grossmann, F. (Éd.). (2014). *L'Écrit scientifique : du lexique au discours. Autour de Scientext*. Rennes, France: Presse universitaire de Rennes.
- Tutin, A. (2014b). La phraséologie transdisciplinaire des écrits scientifiques : des collocations aux routines sémantico-rhétoriques. In A. Tutin & F. Grossmann (Éd.), *L'écrit scientifique : du lexique au discours. Autour de Scientext* (p. 27-44). Rennes: Presses Universitaires de Rennes.
- Tutin, A., Tran, T. T. H., Kraif, O., & Hatier, S. (2015). French collocations of cross-disciplinary scientific lexicon. Présenté à PARSEME 5th general meeting (COST action), Iași, Romania.
- 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.
- Williams, Geoffrey C. (1999). Les réseaux collocationnels dans la construction et l'exploitation d'un corpus dans le cadre d'une communauté de discours scientifique.

2011, thèse de Doctorat, Université de Nantes. Consulté à l'adresse:
<http://geoffrey.williams.pagesperso-orange.fr/thesis/frameintro.htm>

- Williams G, (2003), "Les collocations et l'école contextualiste", in Grossmann F, Tutin A, Collocations : analyse et traitement, Travaux et recherches en linguistique appliquée, série E no1, éd de Werelt, 33-44.
- Williams, Geoffrey. 2006. La linguistique de corpus: une affaire prépositionnelle. In F. Rastier and M. Ballabriga (eds.). Corpus en lettres et sciences sociales: des documents numériques à l'interprétation. Actes du colloque international d'Albi, juillet 2006, 151–158. Paris: Texto
- Williams, G., & Millon, C. (2012). Palmer, Firth and Internet: Drawing together collocational threads. In *Corpus Linguistics 2011*. Consulté à l'adresse http://halshs.archives-ouvertes.fr/docs/00/72/53/01/PDF/Williams_and_Millon_CL2011_v10_GCW.pdf
- Yan, R., & Tutin, A. (2013). *Un modèle lexicographique basé sur le corpus Scientext comparé avec d'autres ressources lexicographiques : l'exemple des constructions verbales*. Présenté à 7ème Journées Internationales de linguistique de corpus, 4-6 septembre, Lorient, France.
- Yoon, H., & Hirvela, A. (2004). ESL student attitudes toward corpus use in L2 writing. *Journal of second language writing*, 13(4), 257-283.
- Yu-Jieh LIN. (2005). *L'enseignement des collocations sur corpus comparable multilingue en FLE avec public sinophone comme exemple*. Mémoire de Master Français Langue Etrangère Recherche, Université Sorbonne Nouvelle.

Annexe 01 : Tableaux des collocations évaluatives N_{scientifique} + Adj_{évaluatif} en question

La biologie :

collocations	nombre d'occurrences
rôle essentiel	2
question principale	1
définition classique	1
question essentielle	1
question fondamentale	1
outil performant	1
facteur essentiel	1
problème classique	1
idée originale	1
traitement efficace	1

Introduction (10)

collocations	nombre d'occurrences
rôle important	3
nouveau outil	2
système complexe	1
analyse fine	1
traitement efficace	1
structure principale	1
outil performant	1
résultats intéressants	1
rôle central	1
élément majeur	1
élément important	1
rôle crucial	1
choix crucial	1
résultat difficile	1
modèle cohérent	1
méthode importante	1
outil essentiel	1
bon exemple	1
fonction essentielle	1

communication riche	1
---------------------	---

Conclusion (20)

La médecine :

collocations	nombre d'occurrences
facteur classique	4
facteur principal	1
meilleur résultat	1
meilleur traitement	1
nouveau cas	1
rôle principal	1
problème sérieux	1

Introduction (7)

collocations	nombre d'occurrences
donnée insuffisante	2
outil intéressant	1
résultat encourageant	1
système complexe	1
rôle principal	1
facteur classique	1
rôle essentiel	1

Conclusion (7)

La linguistique :

Introduction (24)

Collocations	nombre d'occurrences
rôle important	2
rôle central	2
résultat principal	2
description fine	2
traitement satisfaisant	1
notion valide	1
question essentielle	1
résultat encourageant	1
nouveau élément	1
approche prometteuse	1
définition claire	1
élément majeur	1
question principale	1
excellent cas	1

objectif principal	1
principe majeur	1
problème majeur	1
rôle majeur	1
traitement complexe	1
modèle nouveau	1
notion complexe	1
principe fondamental	1
analyse pertinente	1

Conclusion (36)

Collocations	Nombre d'occurrences
rôle important	2
principe fondamental	2
meilleur résultat	2
rôle crucial	2
élément pertinent	2
structure complexe	2
facteur pertinent	1
critère fondamental	1
description complète	1
objectif ambitieux	1
critère essentiel	1
problème épineux	1

communication efficace	1
résultat solide	1
outil essentiel	1
description fine	1
élément déterminant	1
structure originale	1
résultat satisfaisant	1
étude fine	1
idée centrale	1
analyse fine	1
étude rigoureuse	1
phénomène classique	1
qualité principale	1
notion difficile	1
résultat encourageant	1
test classique	1
résultat principal	1
critère classique	1
outil prometteur	1
critère pertinent	1
facteur intéressant	1
rôle fondamental	1
facteur principal	1

meilleure description	1
-----------------------	---

véritable théorie	1
question importante	1

Introduction (13)

La psychologie :

collocations	nombre d'occurrences
objectif principal	3
théorie principale	1
modèle principal	1
travail important	1
outil classique	1
résultat satisfaisant	1
modèle classique	1
principe fondamental	1
choix facile	1
système central	1
fonction complexe	1

collocations	nombre d'occurrences
meilleur outil	2
système complexe	2
objectif principal	2
projet fondamental	1
bonne méthode	1
traitement pertinent	1
cas complexe	1
traitement efficace	1
contribution importante	1
caractéristique pertinente	1

objectif ambitieux	1
rôle important	1
bon exemple	1
idée cohérente	1
résultat classiques	1
résultat principaux	1
véritable modèle	1
modèle cohérent	1
Solution intéressante	1
résultat encourageant	1

Conclusion (16)

Les Sciences de l'éducation :

collocations	nombre d'occurrences
modèle erroné	2
objectif principal	2
question principale	2
résultat principal	2
élément central	2
caractéristique fondamentale	1
concept central	1
conception claire	1
définition concise	1

modèle classique	1
meilleur modèle	1
bonne méthode	1
outil efficace	1
nouveau outil	1
outil pertinent	1
nouveau phénomène	1
vrai problème	1
meilleure qualité	1
question pertinente	1
rôle fondamental	1
théorie classique	1
théorie principale	1
élément essentiel	1
élément pertinent	1
étude significative	1

Introduction (25)

collocations	nombre d'occurrences
conception erronée	4
rôle central	2
concept fondamental	1
conception classique	1
conception cohérente	1

facteur déterminant	1
facteur important	1
facteur principal	1
fonction centrale	1
idée fondamentale	1
outil intéressant	1
phénomène faible	1
principe essentiel	1
projet ambitieux	1
question principale	1
résultat cohérent	1
mauvais résultat	1
rôle important	1
structure acceptable	1
élément central	1
élément majeur	1
élément significatif	1

Conclusion (22)

Le TAL :

collocations	nombre d'occurrences
caractéristique principale	2
fonction importante	2
problème majeur	2
bon résultat	2
rôle important	2
caractéristique importante	1

choix fondamental	1
description cohérente	1
description fine	1
donnée cruciale	1
bonne définition	1
problème complexe	1
problème fondamental	1
problème intéressant	1
nouveau problème	1
projet important	1
mauvaise qualité	1
meilleure qualité	1
question centrale	1
résultat principal	1
résultat significatif	1
bonne solution	1
structure complexe	1
élément central	1
étude sérieuse	1

Introduction (25)

collocations	nombre d'occurrences
meilleur résultat	5
résultat satisfaisant	3

méthode classique	2
objectif principal	2
proposition complexe	2
analyse correcte	1
analyse fine	1
choix déterminant	1
contribution principale	1
critère central	1
critère crucial	1
critère insuffisant	1
critère pertinent	1
véritable critère	1
définition correcte	1
définition cruciale	1
définition épineuse	1
nouveau exemple	1
véritable outil	1
paramètre important	1
phénomène difficile	1
nouveau phénomène	1
phénomène pertinent	1
problème complexe	1
problème fondamental	1

véritable problème	1
proposition principale	1
qualité acceptable	1
qualité intéressante	1
meilleure qualité	1
question centrale	1
question difficile	1
bon résultat	1
résultat encourageant	1
meilleure solution	1
structure complexe	1
structure originale	1
nouveau test	1
élément essentiel	1
élément principal	1

Conclusion (40)

La mécanique :

collocations	nombre d'occurrences
structure cohérente	13
système complexe	4

phénomène cohérent	2
nouveau modèle	2
objectif principal	1
résultat encourageant	1
théorie classique	1
théorie principale	1

Introduction (8)

collocations	nombre d'occurrences
structure cohérente	21

système complexe	3
phénomène cohérent	2
résultat cohérent	2
caractéristique cohérente	1
caractéristique fine	1
caractéristique intéressante	1
nouveau concept	1
objectif principal	1
phénomène classique	1
phénomène complet	1
résultat important	1
résultat intéressant	1
résultat majeur	1
résultat satisfaisant	1
solution pertinente	1
travail important	1

Conclusion (17)

L'électronique :

collocations	nombre d'occurrences
--------------	----------------------

méthode principale	3
système complexe	2
caractéristique principale	1
critère essentiel	1
facteur principal	1
modèle satisfaisant	1
paramètre essentiel	1
résultat intéressant	1
meilleure solution	1
système classique	1
étude majeure	1

Introduction (11)

collocations	nombre d'occurrences
modèle complexe	3
nouveau modèle	3
choix judicieux	2
méthode innovante	2
paramètre important	2
bonne qualité	2
structure adéquate	2
système complexe	2
nouveau cas	1
contribution effective	1

contribution importante	1
contribution principale	1
facteur principal	1
modèle adéquat	1
modèle fondamental	1
notion principale	1
notion valable	1
outil efficace	1
mauvaise qualité	1
question efficace	1

question fondamentale	1
résultat cohérent	1
meilleur résultat	1
rôle fondamental	1
solution prometteuse	1
solution satisfaisante	1
structure complexe	1
véritable travail	1

Conclusion (28)

Annexe 02 : Tableaux de classement des collocations évaluatives N_{scientifique} + Adj_{évaluatif}

Les collocations dans l'introduction sont en bleu, les collocations dans la conclusion sont en violet.

Adjectif : Épithète : dans un syntagme nominal pour qualifier le nom (non indiqués)

Adjectif : Attribut : introduit par un verbe pour qualifier le nom : ~~~~~~

La Biologie :

		Processus de l'activité scientifique	Objets construits par l'activité scientifique	Observables de l'activité scientifique	Supports de la rédaction scientifique	Caractérisation	Relation logique
Adjectifs modaux							

Adjectifs axiologique Polarité +/-		<u>Communicati on riche +</u>	Traitement efficace + Outil performant + Idée originale + Outil performant + <u>Modèle cohérent +</u> Traitement efficace +	Résultat intéressant + Bon exemple +			
Adjectifs non axiologiques	Degré(quantité) et intensité		Analyse fine				
	Comparaison						
	Importance	<u>choix crucial</u>	Structure principale <u>Méthode importante</u> Outil essentiel	Question fondamentale Question essentielle Question principale Facteur essentiel Élément majeur Élément important		Rôle essentiel <u>Rôle essentiel</u> Rôle important (2) Rôle crucial Fonction essentielle	
	Complexité		Système complexe	<u>Résultat difficile</u>			
	Nouveauté Polarité +/-		Définition classique Nouveau outil (2)	Problème classique			

	Temporalité						
	Autre						

La Médecine :

		Processus de l'activité scientifique	Objets construits par l'activité scientifique	Observables de l'activité scientifique	Supports de la rédaction scientifique	Caractérisation	Relation logique
Adjectifs modaux							
Adjectifs axiologique Polarité +/-			Meilleur traitement + Outil intéressant +	Meilleur résultat + <u>Donnée insuffisante</u> (2)- Résultat encourageant+			
Adjectifs non axiologiques	Degré(quantité) et intensité						
	Comparaison						
	Importance			Problème sérieux Facteur principal		Rôle principal Rôle principal Rôle essentiel	
	Complexité		Système complexe				

	Nouveauté		Nouveau cas	Facteur classique (4)			
	Polarité +/-			Facteur classique			
	Temporalité						
	Autre						

La Linguistique :

	Processus de l'activité scientifique	Objets construits par l'activité scientifique	Observables de l'activité scientifique	Supports de la rédaction scientifique	Caractérisation	Relation logique
Adjectifs modaux						
Adjectifs axiologique Polarité +/-	Communication efficace +	Notion valide + Approche prometteuse + Définition claire + Traitement satisfaisant + Analyse pertinente + Objectif ambitieux + Étude rigoureuse + Outil prometteur + Critère pertinent +	Excellent cas + Résultat encourageant + Meilleur résultat + (2) Élément pertinent + (2) Facteur pertinent + Résultat satisfaisant + Résultat encourageant ±			

			Meilleure description +	Facteur intéressant + Problème épineux - Résultat solide			
Adjectifs non axiologiques	Degré (quantité) et intensité		Description fine Description fine Description fine Description complète Étude fine Analyse fine				
	Comparaison						
	Importance		Question principale Objectif principal Principe majeur Principe fondamental Principe fondamental (2) Critère fondamental Critère essentiel Outil essentiel Idée centrale	Problème majeur Résultat principal (2) Élément majeur Question essentielle Élément déterminant Résultat principal Facteur principal		Rôle important (2) Rôle central (2) Rôle majeur Rôle important (2) Rôle crucial (2) Qualité principale Rôle fondamental	

	Complexité		Notion complexe Traitement complexe Structure complexe (2) Notion difficile				
	Nouveauté Polarité +/-		Nouveau modèle Structure originale Critère classique Test classique	Nouveau élément Phénomène classique			
	Temporalité						
	Autre						

La Psychologie :

		Processus de l'activité scientifique	Objets construits par l'activité scientifique	Observables de l'activité scientifique	Supports de la rédaction scientifique	Caractérisation	Relation logique
Adjectifs modaux							

Adjectifs axiologique Polarité +/-			Véritable théorie + Meilleur outil (2) + Bonne méthode + Objectif ambitieux + <u>Idée cohérente +</u> Modèle cohérent + Solution intéressante + Traitement pertinent + Traitement efficace +	Résultat satisfaisant+ Bon exemple + <u>Résultat encourageant</u> ±		Caractéristique pertinente +	
Adjectifs non axiologiques	Degré(quantité) et intensité		Véritable modèle				
	Comparaison						
	Importance	Contribution importante	Objectif principal (3) <u>Question importante</u> Théorie principale Modèle principal Principe fondamental Système central Travail important	Résultat principal		Rôle important	

			Objectif principal (2) Projet fondamental				
	Complexité	<u>Choix facile</u>	Système complexe (2)	Cas complexe		Fonction complexe	
	Nouveauté		Outil classique	Résultat classique			
	Polarité +/-		Modèle classique				
	Temporalité						
	Autre						

Les Sciences de l'éducation:

		Processus de l'activité scientifique	Objets construits par l'activité scientifique	Observables de l'activité scientifique	Supports de la rédaction scientifique	Caractérisation	Relation logique
Adjectifs modaux							
Adjectifs axiologique	Polarité +/-		<u>Modèle erroné - (2)</u> Définition concise + Meilleur modèle + Bonne méthode + Outil efficace + Outil pertinent +	Question pertinente + Élément pertinent + <u>Résultat cohérent +</u> Mauvais résultat -		Meilleure qualité +	

			Conception claire + Conception erronée - (4) <u>Conception cohérente +</u> Outil intéressant + Projet ambitieux - Structure acceptable +				
Adjectifs non axiologiques	Degré(quantité) et intensité			Vrai problème <u>Phénomène faible</u>			
	Comparaison						
	Importance		Étude significative + Objectif principal (2) Question principale (2) Concept central Théorie principale Concept fondamental Concept principal Idée fondamentale Principe essentiel	Résultat principal (2) Élément central (2) Élément essentiel Facteur déterminant Facteur important Facteur principal Élément central Élément majeur Élément significatif		Caractéristique fondamental Rôle fondamental Rôle central (2) Rôle important Fonction centrale	

	Complexité						
	Nouveauté Polarité +/-		Modèle classique Nouveau outil Théorie classique Conception classique	Nouveau phénomène			
	Temporalité						
	Autre						

Le TAL :

		Processus de l'activité scientifique	Objets construits par l'activité scientifique	Observables de l'activité scientifique	Supports de la rédaction scientifique	Caractérisation	Relation logique
Adjectifs modaux							
Adjectifs axiologique	Polarité +/-		Description cohérente + Bonne définition + Bonne solution + Étude sérieuse + Critère insuffisant -	Bon résultat + (2) Problème intéressant + Meilleur résultat + (5) Résultat satisfaisant + (3)		Mauvaise qualité - Meilleure qualité + Qualité acceptable + Qualité intéressante +	

			<p>Critère pertinent +</p> <p>Véritable critère+</p> <p>Définition correcte +</p> <p>Meilleure solution +</p> <p>Analyse correcte +</p>	<p>Phénomène (non) pertinent -</p> <p>Bon résultat +</p> <p>Résultat encourageant +</p>		<p>Meilleure qualité +</p>	
Adjectifs non axiologiques	Degré (quantité) et intensité		<p>Description fine</p> <p>Analyse fine</p> <p>Véritable problème</p> <p>Véritable outil</p>				
	Comparaison						
	Importance	<p>Choix fondamental</p> <p>Choix déterminant</p> <p>Contribution principale</p>	<p>Projet important</p> <p>Question centrale</p> <p>Objectif principal (2)</p> <p>Critère central</p> <p>Critère crucial</p> <p><u>Définition cruciale</u></p> <p>Proposition principale</p>	<p>Résultat significatif</p> <p>Problème majeur (2)</p> <p>Problème fondamental</p> <p>Donnée cruciale</p> <p>Résultat principal</p> <p>Élément central</p> <p>Paramètre important</p> <p>Élément essentiel</p>		<p>Caractéristique principale (2)</p> <p><u>Fonction importante</u></p> <p>Fonction importante</p> <p>Rôle important (2)</p> <p>Caractéristique important</p>	

				Élément principal <u>Question centrale</u> Problème fondamental			
	Complexité		Structure complexe Proposition complexe (2) <u>Définition épineuse</u> Structure complexe	Problème complexe Phénomène difficile Question difficile Problème complexe			
	Nouveauté Polarité +/-		Méthode classique (2) Nouveau test Structure originale	Nouveau problème Nouveau exemple Nouveau phénomène			
	Temporalité						
	Autre						

La Mécanique:

		Processus de l'activité scientifique	Objets construits par l'activité scientifique	Observables de l'activité scientifique	Supports de la rédaction scientifique	Caractérisation	Relation logique
Adjectifs modaux							

Adjectifs axiologique Polarité +/-			Structure cohérente + (12) Structure cohérente + (21) Solution pertinente +	Phénomène cohérent + (2) <u>Résultat encourageant</u> + Phénomène cohérent + (2) Résultat cohérent + (2) Résultat intéressant + Résultat satisfaisant +		Caractéristique cohérente + Caractéristique intéressante +	
Adjectifs non axiologiques	Degré(quantité) et intensité			Phénomène complet		Caractéristique fine	
	Comparaison						
	Importance		Objectif principal Théorie principale Objectif principal Travail important	Résultat important Résultat majeur			
	Complexité		<u>Système complexe (4, 1 avec être)</u> Système complexe (3)				
	Nouveauté		Nouveau modèle (2)	Phénomène classique			

	Polarité +/-		Théorie classique Nouveau concept				
	Temporalité						
	Autre						

L'Électronique:

		Processus de l'activité scientifique	Objets construits par l'activité scientifique	Observables de l'activité scientifique	Supports de la rédaction scientifique	Caractérisation	Relation logique
Adjectifs modaux							
Adjectifs axiologique	Polarité +/-	<p>Choix judicieux + (2)</p> <p>Contribution effective +</p>	<p>Modèle satisfaisant +</p> <p>Meilleure solution +</p> <p>Structure adéquat + (2)</p> <p>Modèle adéquat +</p> <p><u>Notion valable</u> +</p> <p>Outil efficace +</p> <p><u>Modèle efficace</u> +</p> <p>Solution prometteuse +</p>	<p>Résultat intéressant +</p> <p>Résultat cohérent +</p> <p>Meilleur résultat +</p>		<p>Bonne qualité + (2)</p> <p>Mauvaise qualité -</p>	

			Solution satisfaisante +				
Adjectifs non axiologiques	Degré(quantité) et intensité		Véritable travail				
	Comparaison						
	Importance	Contribution importante Contribution principale	Méthode principale (3) Critère essentiel Étude majeure Modèle fondamental Notion principale	Facteur principal Paramètre essentiel Paramètre important (2) Facteur principal Question fondamentale		Caractéristique principale Rôle fondamental	
	Complexité		Système complexe (2) Modèle complexe (3) Système complexe (2) Structure complexe				
	Nouveauté Polarité +/-		Système classique Nouveau modèle (3) Méthode innovante (2)	Nouveau cas			
	Temporalité						

	Autre						
--	-------	--	--	--	--	--	--

MOTS-CLÉS

Phraséologie, collocation, Corpus Scientext, écrit scientifique en français, sémantique, didactique

RÉSUMÉ

Les apprenants non-natifs de la langue française éprouvent souvent des difficultés en matière de rédaction des écrits scientifiques, tant sur la stratégie rhétorique que sur le lexique à utiliser. Notre contribution a donc interrogé les aspects linguistiques et rhétoriques des collocations évaluatives qui marquent le positionnement de l'auteur dans le corpus Scientext, afin de proposer des pistes didactiques de ces éléments phraséologiques qui facilitent la rédaction des écrits scientifiques en français pour des apprenants non natifs.

KEY-WORDS

Phraseology, collocation, Corpus Scientext, academic writing in French, semantics, didactics

ABSTRACT

Non-native learners of the French language often find it difficult to write academic writings on both the rhetorical strategy and the lexicon to be used. Our contribution therefore questioned the linguistic and rhetorical aspects of the evaluative collocations that mark the author's evaluation in the Scientext corpus in order to propose didactic tracks of these phraseological elements that facilitate the writing of scientific writings in French for non-native learners.